DGK Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C | Dissertationen | Heft Nr. 802

Lena Albert

Simultane Klassifikation der Bodenbedeckung und Landnutzung unter Verwendung von Conditional Random Fields

München 2017

Verlag der Bayerischen Akademie der Wissenschaften

ISSN 0065-5325 | ISBN 978-3-7696-5214-7

Diese Arbeit ist gleichzeitig veröffentlicht in:

Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik
der Leibniz Universität Hannover

ISSN 0174-1454, Nr. 335, Hannover 2017

Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C | Dissertationen | Heft Nr. 802

Simultane Klassifikation der Bodenbedeckung und Landnutzung unter Verwendung von Conditional Random Fields

Von der Fakultät für Bauingenieurwesen und Geodäsie

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor-Ingenieurin (Dr.-Ing.)

genehmigte Dissertation

Vorgelegt von

M. Sc. Lena Albert

Geboren am 19.06.1986 in Rinteln

München 2017

Verlag der Bayerischen Akademie der Wissenschaften

ISSN 0065-5325 ISBN 978-3-7696-5214-7

Diese Arbeit ist gleichzeitig veröffentlicht in:
Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover
ISSN 0174-1454, Nr. 335, Hannover 2017

Adresse der DGK:

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften (DGK)

Alfons-Goppel-Straße 11 ● D – 80 539 München

Telefon +49 – 331 – 288 1685 ● Telefax +49 – 331 – 288 1759

E-Mail post@dgk.badw.de ● http://www.dgk.badw.de

Prüfungskommission:

Vorsitzender: Prof. Dr.-Ing. Winrich Voß

Referent: Prof. Dr. techn. Franz Rottensteiner

Korreferenten: Prof. Dr.-Ing. Monika Sester
Prof. Dr.-Ing. Stefan Hinz (KIT Karlsruhe)

Tag der mündlichen Prüfung: 10.07.2017

ISSN 0065-5325 ISBN 978-3-7696-5214-7

Zusammenfassung

Im Rahmen dieser Arbeit wird ein neuer Ansatz zur simultanen, kontextbasierten Klassifikation der Bodenbedeckung und Landnutzung anhand von aktuellen Sensordaten vorgestellt. Zur Klassifikation werden Conditional Random Fields (CRF) [Kumar & Hebert, 2006] verwendet, die einen flexiblen, leistungsfähigen Rahmen für die kontextbasierte Klassifikation auf Basis von graphischen Modellen bilden. CRF sind ein überwachtes Verfahren, d.h. die zugrunde liegenden Klassifikatoren werden anhand von repräsentativen Trainingsdaten gelernt. Im Rahmen der Klassifikation wird den GIS-Objekten eines gegebenen räumlichen Landnutzungsdatenbestandes eine Nutzungsart und kleinräumigen Bildsegmenten (Superpixeln) eine Bodenbedeckungsart zugewiesen. Der Ansatz ist konzipiert für monotemporale, multispektrale, hochauflösende Luftbilddaten sowie Höheninformation. Im Gegensatz zu vielen bestehenden Arbeiten ist dieser Ansatz nicht auf einen geographischen Anwendungsbereich, bestimmte Sensordaten oder eine spezifische Klassenstruktur beschränkt, sondern lässt sich flexibel an neue Gegebenheiten anpassen. Die Klassifikation der Landnutzung bildet die Grundlage für einen semi-automatischen Verifikations- und Aktualisierungsprozess von Landnutzungsdatenbeständen.

Den ersten wissenschaftlichen Beitrag dieser Arbeit bildet die Integration von Bodenbedeckung und Landnutzung in ein gemeinsames graphisches Modell, bestehend aus zwei Ebenen. Auf diese Weise entsteht eine einheitliche, konsistente Modellierung, die es erlaubt, Unsicherheiten im Rahmen der Klassifikation zu berücksichtigen. Die gemeinsame Modellierung bietet außerdem den Vorteil, Abhängigkeiten zwischen beliebigen Bildprimitiven, semantischen Klassenstrukturen und den Daten explizit zu modellieren. Dies ermöglicht die Berücksichtigung von lokalem als auch regionalem Kontext im Klassifikationsprozess. Den lokalen Kontext bilden die räumlichen Interaktionen zwischen benachbarten Bildprimitiven, die als paarweises Interaktionspotential modelliert sind. Im Gegensatz zu bereits bestehenden kontextbasierten Verfahren zur Klassifikation der Bodenbedeckung und Landnutzung, in denen die Modellierung der Nachbarschaftsbeziehungen auf a priori Wissen basiert, werden die Abhängigkeiten in dem integrierten Ansatz aus realen Daten gelernt. Bei dem regionalen Kontext handelt es sich um die bidirektionalen statistischen Abhängigkeiten zwischen der Landnutzung und der Bodenbedeckung. Diese Kontextinformation beschreibt die komplexen Abhängigkeiten zwischen mehreren Bildprimitiven einer Region, wodurch sie sich explizit nur als Potential höherer Ordnung modellieren lässt. Die Berücksichtigung dieser Kontextinformation in beide Richtungen stellt eine Neuerung dar. Um eine effiziente Inferenz in dem CRF höherer Ordnung zu ermöglichen, wird Kontextinformation zwischen den Bildprimitiven in dem graphischen Modell mit Hilfe eines iterativen Inferenzalgorithmus ausgetauscht, der den zweiten wissenschaftlichen Beitrag dieser Arbeit darstellt. Die statistischen Abhängigkeiten zwischen der Bodenbedeckung und der Landnutzung werden mit Hilfe von Kontextmerkmalen in den Klassifikationsprozess integriert, welche die komplexen Abhängigkeiten zwischen der Bodenbedeckung und der Landnutzung beschreiben. Diese Kontextmerkmale bilden die Grundlage für einen diskriminativen Klassifikator, der das Potential höherer Ordnung im Rahmen des iterativen Inferenzalgorithmus approximiert. Die Kontextmerkmale bilden den dritten Beitrag dieser Arbeit.

Die Experimente werden anhand von zwei repräsentativen Testgebieten durchgeführt. Mit der Berücksichtigung von Kontext reduziert sich die Anzahl der Fehlklassifikationen in beiden Klassifikationsergebnissen im Vergleich zu einer unabhängigen, nicht-kontextbasierten Klassifikation. Für die Bodenbedeckungsklassifikation wird ein homogeneres Ergebnis und eine verbesserte geometrische Abgrenzung erzielt im Vergleich zu einer unabhängigen Klassifikation. Bei der Landnutzungsklassifikationen werden speziell für GIS-Objekte mit untypischen Eigenschaften oder Ähnlichkeiten zu anderen Landnutzungsklassen Verbesserungen erzielt. Die Größe der Bodenbedeckungssegmente hat einen Einfluss auf den Detailgrad, den Glättungseffekt und die Rechenzeit.

Schlagworte: Kontext, Klassifikation, Bodenbedeckung, Landnutzung, Conditional Random Fields

Abstract

In this thesis, a new approach is proposed for the simultaneous classification of land cover and land use based on current sensor data considering contextual information. For this purpose, Conditional Random Fields (CRF) [Kumar & Hebert, 2006] are applied, which provide a flexible, powerful statistical framework for contextual classification based on graphical models. The presented method is supervised, i.e. it requires representative training data to learn the classifier. Land cover classification is carried out at the level of small image segments (super-pixels), whereas land use is determined at the level of objects from a geospatial database, represented by polygons. This approach is designed for monotemporal, high-resolution, multispectral aerial image data and height information. In contrast to many existing techniques, the approach is not limited to a certain application area, class structure or specific type of sensor data. Due to its flexibility, the approach can be adapted easily to different circumstances. Classification is the first step of a semi-automatic scheme for the verification and update of geospatial land use databases.

In order to realize a simultaneous classification of land cover and land use, both classification tasks are combined in a graphical model consisting of two layers. The design of the graphical model forms the first scientific contribution of this thesis. A main benefit of this approach is that it represents a unified model, where uncertainties of class predictions are considered. By jointly modelling land cover and land use in a two-layer CRF, dependencies between arbitrary image sites, semantic class structures and data can be modelled explictly. This is done for spatial relationships between adjacent image sites, i.e. local context information, as well as for complex dependencies between several image sites in a larger neighbourhood. Spatial dependencies between neighbouring image sites are modelled by pairwise interaction potentials. In contrast to existing approaches, where prior knowledge is used to model this relationship, the spatial dependencies in the two-layer CRF are learned from real-world occurrences in representative training data. The complex statistical dependencies between land cover and land use require the formulation of a high-order potential. By using a high-order potential, both tasks interact in the classification procedure, i.e. contextual information is considered in both directions. In order to enable efficient inference in the two-layer higher order CRF, an iterative inference procedure is presented in which the two classification tasks mutually influence each other. The iterative inference procedure forms the second scientific contribution of this thesis. Contextual relations between land cover and land use are integrated in the classification process by using contextual features describing the complex dependencies of all nodes in a high-order clique. These contextual features represent the third contribution of this thesis. These features are incorporated in a discriminative classifier which approximates the high-order potentials during the inference procedure.

Experiments are carried out on two test sites to evaluate the performance of the proposed method. The experiments show that by considering context the classification results are improved compared to the results of a non-contextual classifier. For land cover classification, the result is much more homogeneous and the delineation of land cover segments is improved. For land use classification, an improvement is mainly achieved for land use objects showing non-typical characteristics or similarities to other land use classes. Furthermore, the size of the super-pixels has an influence on the level of detail of the classification result, but also on the degree of smoothing induced by the segmentation method as well as on the computation time. The smoothing effect is especially beneficial for land cover classes covering large, homogeneous areas, whereas land cover classes mainly representing small structures benefit from a high level of detail.

Keywords: Contextual classification, land cover, land use, Conditional Random Fields

Nomenklatur

Abkürzungen

2D zweidimensional

3D dreidimensional

ALKIS Amtliches Liegenschaftskatasterinformationssystem

CART Classification and Regression Trees

CRF Conditional Random Fields

DGM Digitales Geländemodell

DOM Digitales Oberflächenmodell

DOP Digitales Orthophoto

EM Expectation Maximisation

GIS Geoinformationssystem

GLCM Grey Level Co-Occurrence Matrix

HOG Histogram of Oriented Gradients

HSV Hue-Saturation-Value

LBP Loopy Belief Propagation

LGLN Landesamt für Geoinformation und Landesvermessung Niedersachsen

LVermGeo SH Landesamt für Vermessung und Geoinformation Schleswig-Holstein

MAP Maximum-A-Posteriori

MRF Markov Random Fields

nDOM normalisiertes Digitales Oberflächenmodell

NDVI Normalised Difference Vegetation Index

NIR Nahes Infrarot

OOB Out-of-Bag

RAM Random-Access Memory

RF Random Forests

RGB Rot-Grün-Blau

SAR Synthetic Aperture Radar

SLIC Simple Linear Iterative Clustering

SVM Support Vector Machines

TN Tatsächliche Nutzung

Liste von Symbolen

Allgemein

b, h, i, j, k, l, o, t	Indizes
d	euklidische Distanz
δ_K	Kronecker-Delta-Funktion
μ	Mittelwert
π	Pi-Zahl
σ	Standardabweichung

Klassifikation

$\mathcal{L}$	Menge diskreter Klassenlabels
$\mathcal{S}$	Menge von Bildprimitiven
$\mathcal{T}$	Menge von Trainingsbeispielen
$\mathbb{R}^n$	n-dimensionaler euklidischer Raum
$\boldsymbol{C}$	Konfusionsmatrix
D	Dimension des Merkmalsvektors
G	Gesamtgenauigkeit
K	Korrektheit
M	Anzahl der Klassen
N_S	Anzahl der Bildprimitive in $\mathcal{S}$
N_T	Anzahl der Bildprimitive in $\mathcal{T}$
P	Wahrscheinlichkeitsverteilung
Q	Qualität
V	Vollständigkeit
$\mathbf{f}$	Merkmalsvektor
$\mathbf{x}$	Beobachtungen
$\mathbf{y}$	Vektor von Zufallsvariablen
l	diskretes Klassenlabel
p_0	tatsächliche Übereinstimmung zw. Klassifikationsergebnis und Referenz
p_c	erwartete Übereinstimmung zw. Klassifikationsergebnis und Referenz
y	Zufallsvariable
κ	Kappa-Index

Merkmale

$\mathcal{R}$	Menge von Pixeln innerhalb einer Bildregion

A	Flächeninhalt
A_{mbr}	Flächeninhalt des minimal umschließenden Rechtecks
N_F	Anzahl der Merkmale
N_G	Anzahl der Grauwerte eines Bildes
N_R	Anzahl von Pixeln in $\mathcal{R}$
$P_{\Delta,\alpha}$	GLCM für räumliche Konfiguration (Δ, α)
U	Umfang
b_{mbr}	Breite des minimal umschließenden Rechtecks
c	Spalte (Bildkoordinate)
d_{max}	maximaler Abstand der Polygonpunkte vom Schwerpunkt
d_{min}	minimaler Abstand der Polygonpunkte vom Schwerpunkt
f	Merkmal
g	Grauwert eines Bildes
g_B	Spektralwert Blau
g_G	Spektralwert Grün
g_{NIR}	Spektralwert Nahes Infrarot
g_R	Spektralwert Rot
l_{mbr}	Länge des minimal umschließenden Rechtecks
p, q	Ordnung der Bildmomente
r	Zeile (Bildkoordinate)
Δ	Pixelabstand
α	Richtung
ε	Quantil
μ_{pq}	zentrale Bildmomente der Ordnung p, q
$\bar{\mu}_{pq}$	normalisierte zentrale Bildmomente der Ordnung p, q

Random Forests

$\mathcal{T}_b$	Bootstrap-Datensatz
B	Anzahl der Entscheidungsbäume
L, R	Bezeichner linker bzw. rechter Zweig des Entscheidungsbaumes
$N_{T,min}$	Mindestanzahl der Samples für nicht-terminierende Knoten
$N_{T,max}$	maximale Anzahl an Trainingsbeispielen pro Klasse
N_{T_b}	Anzahl der Bildprimitive in $\mathcal{T}_b$
N_f	Anzahl der zu testenden Merkmale
N_l	Summe der Stimmen pro Klassenlabel
N_n	Anzahl der Knoten im Pfad
T_{max}	maximale Tiefe der Entscheidungsbäume
$\mathbf{w}$	Parameter der Hyperebene
θ	Schwellwert

Conditional Random Fields

$\mathcal{H}$	Menge der Cliquen höherer Ordnung
$\mathcal{N}$	Menge der benachbarten Knoten
Z	Partitionsfunkion
$\mathbf{m}$	Vektor der Kontextmerkmale
c	Bezeichner Bodenbedeckungsebene
e	Kante
e_R	räumliche Kanten
e_S	semantische Kanten
n	Knoten
n_{It}	Anzahl der Iterationen des iterativen Inferenzalgorithmus
n_{LBP}	Anzahl der Iterationen in jedem LBP-Teilschritt
u	Bezeichner Landnutzungsebene
$\boldsymbol{\mu}$	Vektor der Interaktionsmerkmale
Ω	Menge der Parameter ω
β	Grad der Datenabhängigkeit im kontrastsensitiven Potts-Modell
ϕ	Assoziationspotential
ψ	Interaktionspotential
ω	Parameter zur relativen Gewichtung der Potentiale

Inhaltsverzeichnis

1. Einleitung

1.1. Motivation

Die tatsächliche Landnutzung wird typischerweise als räumlicher Datenbestand in Geoinformationssystemen (GIS) detailliert und flächendeckend vorgehalten. Die Landnutzung beschreibt die sozioökonomische Funktion eines Teilbereichs der Erdoberfläche, der verschiedene Arten von Bodenbedeckungen enthalten kann. Demgegenüber beschreibt die Bodenbedeckung das physische Material der Erdoberfläche, wobei es sich sowohl um natürlichen Bewuchs, wie z.B. Gras, oder künstliche Bodenbeläge, wie z.B. Asphalt, handeln kann. Die Landnutzungsinformation stellt einen wichtigen Bestandteil der Geodateninfrastruktur dar und ist von großer Bedeutung für verschiedenste Nutzergruppen, wie z.B. zur Realisierung einer nachhaltigen Stadtentwicklung. Die Anforderungen an diesen Datenbestand haben sich in den letzten Jahren gewandelt hin zu einem hohen Detailgrad in Form von kleinräumigen geometrischen Einheiten sowie einer feinen Unterscheidung von Landnutzungsklassen [Arnold et al., 2017]. Allerdings führen schnelle Nutzungsänderungen, z.B. im Zuge von städtebaulichen Maßnahmen, in weiten Bereichen zu Einbußen in der Aktualität dieser Datenbestände. Die derzeitigen Aktualisierungsprozesse stoßen hierbei häufig an ihre Grenzen, da sie für derart detaillierte Datenbestände mit einem hohen Aufwand verbunden sind [Champion, 2007].

Um den Aktualisierungsprozess zu vereinfachen und damit die Aktualität der Datenbestände zu gewährleisten, streben die Vermessungsbehörden langfristig die Entwicklung und Durchführung eines automatischen Aktualisierungsprozesses an. Für die automatische Verifikation und Aktualisierung eines räumlichen, großmaßstäbigen Landnutzungsdatenbestandes bieten sich insbesondere aktuelle, hochauflösende Luft- bzw. Satellitenbilder und daraus automatisch ableitbare Produkte an, wie z.B. Orthophotos, digitale Gelände- (DGM) und Oberflächenmodelle (DOM). Die Verwendung dieser Daten bietet den Vorteil, dass sie eine großräumige und, dank der hohen geometrischen Auflösung, detaillierte Bearbeitung ermöglichen. Zudem handelt es sich hierbei um Daten, die typischerweise von den Vermessungsbehörden in regelmäßigen Abständen im Rahmen von Standardprozessen erfasst bzw. erzeugt werden und damit ohne Mehraufwand zur Verfügung stehen.

Ein automatisierter Aktualisierungsprozess lässt sich am effizientesten auf der Grundlage einer automatischen Klassifikation der Landnutzung anhand von Sensordaten realisieren (z.B. [Helmholz et al., 2014]). Der Erfolg des automatisierten Verfahrens wird dabei maßgeblich durch die Güte der Klassifikation beeinflusst. Daher werden an die Klassifikation verschiedene Anforderungen gestellt. Zum einen sollte die Klassifikation eine feine Klassenstruktur unterscheiden, die sich möglichst nah an der Klassenstruktur des zu verifizierenden Datenbestandes orientiert. Zum anderen sollte das Verfahren universell für verschiedene Regionen einsetzbar sein, d.h. sich nicht auf ein Anwendungsgebiet

bzw. eine Charakteristik (urban, ländlich) spezialisieren. Nicht zuletzt sollte die Klassifikation eine hohe Genauigkeit liefern. Nur auf der Grundlage eines korrekten, detaillierten und flächendeckenden Klassifikationsergebnisses lässt sich eine Landnutzungsdatenbank effizient und weitgehend automatisiert verifizieren bzw. aktualisieren. Zur Klassifikation werden in letzter Zeit vermehrt überwachte, statistische Verfahren eingesetzt, die im Gegensatz zu modellbasierten Ansätzen deutlich flexibler sind hinsichtlich der unterschiedlichen Ausprägungen der Klassen in den Eingangsdaten. Darüber hinaus lässt sich die Methodik in einfacher Weise auf andere Datensätze übertragen. Bei einem überwachten Ansatz wird zunächst ein Klassifikator anhand von Trainingsgebieten gelernt. Für die Bildprimitive in den Trainingsgebieten sind die Merkmale sowie die Klassenlabels als gewünschte Ausgabe bekannt, sodass der Zusammenhang zwischen den Merkmalen und den Klassen gelernt werden kann. Anschließend erfolgt die eigentliche Klassifikation, die jedem unbekannten Bildprimitiv ein Klassenlabel auf Basis seiner Merkmale zuweist.

Für diesen Zweck werden standardmäßig Klassifikationsverfahren eingesetzt, die jedes Bildprimitiv unabhängig von allen anderen nur anhand der Sensordaten klassifizieren. Diese Vorgehensweise ist für die Klassifikation der Landnutzung jedoch nur bedingt geeignet, da die Landnutzung im Gegensatz zur Bodenbedeckung typischerweise nicht direkt aus den Sensordaten abgeleitet werden kann. Neben den spektralen Eigenschaften liefert vielmehr die Zusammensetzung und Anordnung verschiedener Bodenbedeckungsarten einen Hinweis auf die vorherrschende Landnutzung. Gewisse Landnutzungsarten zeichnen sich durch eine charakteristische Zusammensetzung an Bodenbedeckungsarten aus. Beispielsweise sind in der Landnutzung *Wohnbaufläche* typischerweise die Bodenbedeckungsarten *Gebäude*, *Versiegelung* und *Gras* oder *Baum* enthalten.

Diverse Arbeiten haben bereits bestätigt, dass sich diese Art von Kontextinformation positiv auf die Klassifikation der Landnutzung auswirkt. So hat es sich bewährt, Kontextinformation über die Abhängigkeit von der Bodenbedeckung implizit in Form von Merkmalen in den Klassifikationsprozess zu integrieren. Diese Vorgehensweise stellt jedoch hohe Anforderungen an die Auswahl geeigneter Merkmale, die das Klassifikationsergebnis maßgeblich beeinflussen. Eine Erweiterung hierzu bilden kontextbasierte Klassifikationsverfahren, die Abhängigkeiten zwischen den Klassen bzw. Daten in benachbarten Bildprimitiven explizit in Zufallsfeldern modellieren, z.B. in *Conditional Random Fields* (CRF) [Kumar & Hebert, 2006].

Im Zusammenhang mit der Klassifikation der Landnutzung existieren weitere Arten von Kontextinformation, die das Ergebnis positiv beeinflussen können, wie z.B. die räumliche Abhängigkeit benachbarter Landnutzungsobjekte. Grundsätzlich ist zu beobachten, dass bestimmte Konfigurationen von Nutzungsarten häufiger auftreten. Beispielsweise wird ein Wohngrundstück üblicherweise von einer Straße erschlossen, sodass Landnutzungsobjekte der Klassen *Wohnbaufläche* und *Straßenverkehr* typischerweise räumlich benachbart sind. Hingegen treten andere Kombinationen nicht oder nur selten auf, was u.a. auch darin begründet sein kann, dass sie durch Vorgaben der Stadt- oder Flächennutzungsplanung untersagt sind. Diese Zusatzinformation kann dazu genutzt werden, um wahrscheinliche Konfigurationen von Landnutzungsklassen den unwahrscheinlicheren vorzuziehen.

Die Klassifikation der Bodenbedeckung profitiert ebenfalls von der Berücksichtigung von Kontext. Analog zur Landnutzung sind hierbei räumliche Abhängigkeiten benachbarter Bildprimitive zu

beobachten, deren Ausprägung jedoch in Abhängigkeit der Größe der Bildprimitive variiert. Je kleiner die Ausdehnung der Bildprimitive auf der Erdoberfläche, desto höher ist die Wahrscheinlichkeit, dass benachbarte Bildprimitive zur gleichen Klasse gehören. Für den Fall, dass die Bildprimitive mit den Begrenzungen der Bodenbedeckungen konsistent sind, d.h. sich von angrenzenden Bildprimitiven hinsichtlich ihrer Bodenbedeckungsart unterscheiden, lässt sich die räumliche Abhängigkeit als Wahrscheinlichkeit für das gemeinsame Auftreten von Bodenbedeckungsarten interpretieren. Darüber hinaus hängt die Bodenbedeckungsart von der vorherrschenden Landnutzung ab. Die Wahrscheinlichkeit für das Auftreten einer Bodenbedeckungsart variiert in Abhängigkeit der Nutzungsart, so enthalten Wälder vornehmlich die Bodenbedeckungsart *Baum*, wohingegen andere Bodenbedeckungsarten, wie z.B. *Gebäude*, normalerweise nicht vorkommen. Die Bestimmung der Bodenbedeckung kann somit von dem Wissen über die vorliegende Nutzungsart profitieren. Dieser zusätzliche Hinweis ist besonders hilfreich bei hochauflösenden Eingangsdaten. Eine hohe geometrische Auflösung führt – speziell in urbanen Gebieten – zu einem heterogenen Erscheinungsbild vieler Bodenbedeckungsarten, was die Klassifikation der Bodenbedeckung maßgeblich erschwert.

Zusammenfassend lässt sich festhalten, dass die Bodenbedeckung und die Landnutzung ein komplexes Geflecht an kontextuellen Abhängigkeiten bilden, was eine isolierte Betrachtung beider Aufgaben wenig zweckmäßig erscheinen lässt. Mit der Berücksichtigung von Kontext wird die Entscheidungsfindung mit zusätzlichen Hinweisen unterstützt, die zudem ein konsistentes Klassifikationsergebnis forcieren. Trotz der mittlerweile etablierten Nutzung von Kontext in beiden Klassifikationsaufgaben, werden die Bodenbedeckung und die Landnutzung bislang separat klassifiziert. Um die Abhängigkeiten jedoch vollständig in der Klassifikation abzubilden, bedarf es einer aufgabenübergreifenden, konsistenten Modellierung beider Klassifikationsaufgaben.

1.2. Ziel und wissenschaftlicher Beitrag der Arbeit

Das Ziel dieser Arbeit besteht in der Entwicklung eines kontextbasierten, probabilistischen Verfahrens zur Klassifikation der Landnutzung anhand von aktuellen Sensordaten. Da die Landnutzung in wesentlichem Maße von der vorliegenden Bodenbedeckung abhängt, soll zusätzlich die Bodenbedeckung bestimmt und im Klassifikationsprozess berücksichtigt werden. Aus der allgemeinen Zielsetzung leiten sich somit zwei Teilziele ab. Zum einen gilt es, die aktuelle Nutzungsart für die GIS-Objekte eines gegebenen räumlichen Landnutzungsdatenbestandes zu prädizieren und zum anderen kleinen Bildsegmenten eine Bodenbedeckungsklasse zuzuweisen. Der Ansatz ist konzipiert für monotemporale, multispektrale, hochauflösende Luftbilder und daraus abgeleitete Produkte, wie z.B. Orthophotos, DOM und DGM.

Das Klassifikationsergebnis bildet die Grundlage für eine sich anschließende Verifikation und Aktualisierung des Landnutzungsdatenbestandes. Diese Schritte liegen jedoch nicht im Fokus dieser Arbeit und werden daher nur rudimentär realisiert. Hierbei ist grundsätzlich ein manueller Nachbearbeitungsaufwand erforderlich, da die Ergebnisse von einem Bearbeiter vor der Übernahme in den Datenbestand zu prüfen und ggf. zu verbessern sind.

Die Grenzen der GIS-Objekte werden im Rahmen der Klassifikation nicht geprüft, geschweige denn

korrigiert. Folglich bleibt die topologische Konsistenz und die Konsistenz mit den Eigentumsgrenzen gewahrt. Zur Änderung von Nutzungsartengrenzen bedarf es einer manuellen Nachbearbeitung. Eine Automatisierung dieses Schrittes ist zwar für die praktische Anwendbarkeit des Verfahrens von Bedeutung, liegt aber außerhalb des Fokus dieser Arbeit.

Im Gegensatz zu einigen bereits bestehenden Ansätzen zur Klassifikation der Landnutzung ist der entwickelte Ansatz nicht auf einen bestimmten Anwendungsbereich spezialisiert, sondern lässt sich flexibel an neue Gegebenheiten anpassen. Dies betrifft zum einen die Fähigkeit zur Unterscheidung von Nutzungsarten in urbanen sowie ländlich geprägten Gebieten und zum anderen die Übertragbarkeit des Ansatzes auf neue Testgebiete mit unterschiedlichen Charakteristiken. Der entwickelte Ansatz unterscheidet eine feine Klassenstruktur und bildet damit eine gute Ausgangsbasis für einen effizienten Verifikations- und Aktualisierungsprozess. Mit der Berücksichtigung von Kontext wird die Anzahl der Fehlklassifikationen in den Ergebnissen der Bodenbedeckung und Landnutzung im Vergleich zu einer unabhängigen, nicht-kontextbasierten Klassifikation reduziert. Nicht zuletzt liefert das Verfahren als Nebenprodukt Informationen über die aktuelle Bodenbedeckung, die für verschiedenste Anwendungen von Interesse sind, wie z.B. für die Flächenstatistik oder das Umweltmonitoring [Arnold et al., 2017]. Angesichts der großen Bedeutung der Bodenbedeckungsinformation streben die Vermessungsbehörden in Deutschland aktuell an, die Bodenbedeckung explizit und getrennt von der Landnutzung auszuweisen bzw. in ihren Datenbeständen vorzuhalten[1].

Den ersten wissenschaftlichen Beitrag dieser Arbeit bildet die Integration von Bodenbedeckung und Landnutzung in einen gemeinsamen Klassifikationsansatz unter Berücksichtigung von Kontext. Es handelt sich hierbei um ein überwachtes Verfahren, d.h. der Klassifikator wird anhand von repräsentativen Trainingsdaten gelernt. Als Klassifikationsmethode werden CRF [Kumar & Hebert, 2006] verwendet, die einen flexiblen, leistungsfähigen Rahmen für die kontextbasierte Klassifikation auf Basis von graphischen Modellen bilden. Die Integration beider Klassifikationsaufgaben erfolgt durch deren Zusammenführung in einem gemeinsamen graphischen Modell, welches aus zwei Ebenen besteht. Auf diese Weise entsteht eine einheitliche, konsistente Modellierung, die es erlaubt, Unsicherheiten im Rahmen der Klassifikation zu berücksichtigen. Die gemeinsame Modellierung bietet zudem den Vorteil, Abhängigkeiten zwischen beliebigen Bildprimitiven, semantischen Klassenstrukturen und den Daten explizit zu modellieren. Folglich ist es möglich, sowohl lokalen als auch regionalen Kontext im Klassifikationsprozess zu berücksichtigen. Den lokalen Kontext bilden die räumlichen Interaktionen zu benachbarten Bildprimitiven. Im Gegensatz zu bereits bestehenden kontextbasierten Verfahren zur Klassifikation der Bodenbedeckung und der Landnutzung, die a priori Wissen über die Nachbarschaftsbeziehungen in das Modell einbinden, werden die Abhängigkeiten in dem integrierten Ansatz aus realen Daten gelernt. Bei dem regionalen Kontext handelt es sich um die bidirektionalen statistischen Abhängigkeiten zwischen der Landnutzung und der Bodenbedeckung. Diese Kontextinformation beschreibt die komplexen Abhängigkeiten zwischen mehreren Bildprimitiven einer Region, wodurch sie sich explizit nur als Potential höherer Ordnung modellieren lässt. Die Berücksichtigung dieser Kontextinformation in beide Richtungen stellt eine Neuerung dar. Der Austausch der Kontextinformation zwischen den Bildprimitiven in dem graphischen Modell erfolgt mit Hilfe eines iterativen Inferenz-

[1] Unveröffentlichter Beschluss des Plenums der Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland 128/3 gem. Nr. 5.2 der GO-AdV 2016: GeoInfoDok, Fortschreibung der AAA-Fachschemata.

algorithmus, der den zweiten Beitrag dieser Arbeit darstellt. Im Rahmen des iterativen Inferenzalgorithmus wird Kontextinformation mittels diskriminativer Kontextmerkmale übertragen, welche die Abhängigkeiten zwischen der Bodenbedeckung und der Landnutzung in geeigneter Weise beschreiben. Die Kontextmerkmale bilden den dritten Beitrag dieser Arbeit.

Dem vorgestellten Ansatz liegen zwei Annahmen zugrunde, die die Relevanz der modellierten Kontextinformation betreffen. Es wird angenommen, dass sowohl der räumliche Kontext als auch der bidirektionale Austausch von Kontextinformation wichtig sind für die Klassifikation der Bodenbedeckung und der Landnutzung.

In diversen Arbeiten hat sich gezeigt, dass die erste Annahme bzgl. der Relevanz von räumlichen Kontext gerechtfertigt ist (z.B. bei der Klassifikation der Bodenbedeckung in [Schindler, 2012]). Darüber hinaus wurde diese Art von Kontextinformation bereits von existierenden Verfahren zur Klassifikation der Landnutzung herangezogen (z.B. [Montanges et al., 2015; Novack & Stilla, 2015]), wenn auch in Form einer vereinfachten Modellierung. Um die räumlichen Abhängigkeiten der Landnutzung realistischer beschreiben zu können, bietet sich ein Modell an, das anhand von realen Daten gelernt wird. Den Zusammenhang zwischen der Bodenbedeckung und der Landnutzung haben bereits andere Verfahren genutzt (z.B. [Hermosilla et al., 2012; Helmholz et al., 2014]), allerdings beschränkt sich die Modellierung in diesen Ansätzen auf eine Richtung, d.h. von der Bodenbedeckung zur Landnutzung. Obwohl nicht direkt von einer Nutzungsart auf eine Bodenbedeckung geschlossen werden kann, so lassen sich doch unwahrscheinliche Bodenbedeckungsarten innerhalb einer Landnutzung ausschließen, was insgesamt zu einer Verbesserung der Ergebnisse beiträgt. Aus diesem Grund ist eine gemeinsame Modellierung beider Klassifikationsaufgaben anzustreben, die den bidirektionalen Austausch von Kontextinformation ermöglicht. Nach bestem Wissen der Autorin sind sowohl der bidirektionale Austausch von Kontextinformation zwischen der Klassifikation der Bodenbedeckung und der Landnutzung als auch ein Modell der räumlichen Abhängigkeiten, das anhand von realen Daten gelernt wird, im Rahmen der Landnutzungsklassifikation bisher noch nicht realisiert.

Die Annahmen werden im Rahmen von Experimenten anhand von realen kartographischen Datensätzen in zwei repräsentativen Testgebieten überprüft. Die Experimente sollen dazu dienen, einerseits die Bedeutung von Kontext für die Klassifikation zu belegen und andererseits die Vorteile der Interaktion zwischen beiden Klassifikationsaufgaben herauszustellen.

1.3. Struktur der Arbeit

Die Arbeit ist wie folgt gegliedert. Zu Beginn wird in Kapitel 2 auf den Stand der Forschung getrennt für den anwendungsorientierten und den methodischen Aspekt dieser Arbeit eingegangen. Anschließend beschreibt das Kapitel 3 die für das neu entwickelte Verfahren relevanten Grundlagen aus dem Bereich der Bildanalyse und dem maschinellen Lernen. Die Methodik des neuen Klassifikationsansatzes wird in Kapitel 4 ausführlich erläutert. Dieses Kapitel geht insbesondere auf die Struktur des graphischen Modells sowie den implementierten Inferenzalgorithmus ein. Das Kapitel 5 dokumentiert die durchgeführten Experimente und führt eine umfassende Evaluation der Ergebnisse an. Die Arbeit schließt mit einem Fazit und einem Ausblick in Kapitel 6.

2. Stand der Forschung

Im folgenden Kapitel wird der Stand der Forschung separat für den anwendungsorientierten und den methodischen Aspekt dieser Arbeit vorgestellt. Bezüglich der Anwendung konzentriert sich die Darstellung in Kapitel 2.1 auf Arbeiten zur Klassifikation der Landnutzung anhand von Sensordaten. Dieses Kapitel gibt nur einen allgemeinen Überblick über die Vielzahl an Realisierungen und erhebt damit keinen Anspruch auf Vollständigkeit. Der Fokus liegt vielmehr auf den verschiedenen Varianten zur Integration von Kontext in die Landnutzungsklassifikation. Demgegenüber bilden statistische Modelle von Kontext, insbesondere CRF, den Schwerpunkt der zweiten, eher methodisch orientierten Übersicht in Kapitel 2.2. In Kapitel 2.3 werden die Grenzen der existierenden Arbeiten ausführlich diskutiert, aus denen sich Anforderungen und Rahmenbedingungen für die neue Methodik ableiten.

2.1. Klassifikation der Landnutzung

In der Literatur findet sich eine Vielzahl von Ansätzen zur Klassifikation der Landnutzung anhand von Sensordaten. Grundsätzlich unterscheiden sich diese Ansätze hinsichtlich der allgemeinen Vorgehensweise, der verwendeten Eingangsdaten, der Klassenstruktur, der räumlichen Klassifikationseinheiten sowie der verwendeten Merkmale und Klassifikationsmethoden.

Die meisten Verfahren verwenden Satelliten- oder Luftbilddaten unterschiedlicher räumlicher Auflösung als Datengrundlage für die Klassifikation der Landnutzung, (z.B. [Walde et al., 2014; Hermosilla et al., 2012]). Daneben werden aber auch vereinzelt andere Sensordaten zur Klassifikation herangezogen, wie z.B. flugzeuggestützte Laserscannerdaten (z.B. [Yoshida & Omae, 2005]) oder Radardaten (z.B. [Novack & Stilla, 2015]). Darüber hinaus integrieren diverse Ansätze zusätzlich thematische Information in den Klassifikationsprozess, wobei es sich zumeist um Angaben aus topographischen oder Katasterdatenbanken handelt (z.B. [Banzhaf & Höfer, 2008]).

Entgegen der einheitlichen Zielsetzung zur Klassifikation der Landnutzung, die allen Verfahren gemein ist, variieren sie hinsichtlich der zu unterscheidenden Klassenstruktur. Diese Unterschiede prägen sich meist in Form einer Spezialisierung auf Teilbereiche der Landnutzung oder Variationen im Hinblick auf den Detailgrad aus. Sie resultieren in erster Linie aus der Zielsetzung des jeweiligen Klassifikationsansatzes. Darüber hinaus wird die unterscheidbare Klassenstruktur durch die gewählten Eingangsdaten sowie die Charakteristik der verwendeten Testgebiete weiter eingeschränkt. So lassen sich beispielsweise in niedrig aufgelösten Satellitendaten nur wenige Details erkennen, die zur Unterscheidung einer fein aufgelösten Klassenstruktur im Allgemeinen unentbehrlich sind. Bei überwachten Klassifikationsverfahren können grundsätzlich nur jene Klassen unterschieden werden, die auch in den jeweiligen Trainingsgebieten auftreten. Die meisten Verfahren haben sich auf die Landnutzung in

urbanen Gebieten spezialisiert, die als urbane Strukturarten (engl.: *urban structure types*) bezeichnet werden (z.B. [Hermosilla et al., 2012]). Daneben existieren weitere Verfahren, die ihren Fokus auf andere Bereiche legen, wie z.B. die Unterscheidung von landwirtschaftlichen Nutzungsarten (z.B. [Helmholz et al., 2014]). Eine solche Spezialisierung schränkt die universelle Anwendbarkeit der Verfahren ein. Der Detailgrad der Klassenstruktur variiert von einer sehr groben Einteilung der Landschaft in zwei Landnutzungsklassen [Taubenböck et al., 2013] bis hin zu sehr detaillierten Hierarchien von mehr als 15 Landnutzungsklassen [Banzhaf & Höfer, 2008]. Um einen derartig hohen Detailgrad zu erzielen, integrieren Banzhaf & Höfer [2008] zusätzlich thematische Daten in ihren Klassifikationsprozess. Hierbei handelt es sich insbesondere um Informationen zur Funktion von Bauwerken (z.B. Krankenhaus, Schule), die sich in der Regel nicht aus den Sensordaten ableiten lassen, aber zur Unterscheidung einer Vielzahl an Landnutzungsarten erforderlich sind. Allerdings stehen solche Informationen typischerweise nicht in der erforderlichen Qualität bzw. Aktualität für die Klassifikation der Landnutzung zur Verfügung. Dies ist insbesondere dann der Fall, wenn die Klassifikation der Landnutzung die Grundlage für die Verifikation und Aktualisierung von Landnutzungsdatenbeständen bildet. Hierbei ist in der Regel davon auszugehen, dass der zu prüfende Datensatz nicht überall die erforderliche Aktualität aufweist, um den Klassifikationsprozess effektiv zu unterstützen.

Darüber hinaus unterscheiden sich die Verfahren hinsichtlich der räumlichen Einheiten auf der Erdoberfläche, denen im Zuge der Klassifikation eine Nutzungsart zugewiesen wird. Grundsätzlich sollte es sich hierbei um sinnvolle strukturelle Einheiten gleicher Nutzungsart handeln, wobei die konkrete räumliche Ausdehnung variieren kann. Die Definition der Interessensgebiete erfolgt für jeden Klassifikationsansatz in Abhängigkeit der jeweiligen Zielsetzung sowie der Eigenschaften der verwendeten Eingangsdaten. Die Abgrenzung kann einerseits manuell anhand von Luftbilddaten [Herold et al., 2003] oder automatisch mit Hilfe von Bildanalysemethoden erfolgen. Bei einer automatischen Herangehensweise lassen sich sowohl unregelmäßig geformte Regionen mittels Segmentierungsverfahren [Barr & Barnsley, 1997] als auch gleichmäßig geformte Zellen mittels Rasterbildung oder einem über das Bild gleitenden Fenster ableiten [Montanges et al., 2015]. Darüber hinaus besteht die Möglichkeit, zusätzliche Informationen zur Abgrenzung der Gebiete heranzuziehen, wovon in der Literatur viele Ansätze Gebrauch machen. Hierbei handelt es sich in erster Linie um kartographische Einheiten, z.B. Häuserblöcke in Stadtgebieten, abgeleitet aus digitalen Landschaftsmodellen (z.B. [Walde et al., 2014]), oder Verwaltungseinheiten, z.B. Flur- bzw. Grundstücke (z.B. [Hermosilla et al., 2012]). Im Gegensatz zu automatisch abgeleiteten Regionen oder Rasterzellen bieten diese Einheiten den Vorteil, dass sie die wesentlichen Objekte in der Realität prägnanter abgrenzen. Des Weiteren bleibt die Begrenzung in der Regel über die Zeit sehr stabil, selbst wenn sich zwischenzeitlich die zugehörige Landnutzung geändert hat (z.B. bei der Umwandlung von Grün- in Ackerland innerhalb des Bestandes).

Im Allgemeinen wird zur Bestimmung der Landnutzung eine zweistufige Vorgehensweise verfolgt (z.B. [Hermosilla et al., 2012; Helmholz et al., 2014]). In einem ersten Schritt wird zuvor definierten Bildprimitiven (z.B. Pixel, Segmente) eine Bodenbedeckungsklasse zugewiesen. Diese Information kann entweder mittels Klassifikation aus den Sensordaten abgeleitet oder aus bestehenden räumlichen Datenbeständen übernommen werden. In einem zweiten Schritt erfolgt die eigentliche Klassifikation der Landnutzung, die u.a. die Information über die aktuelle Bodenbedeckung in die Entscheidung einbezieht. Zu diesem Zweck wird jedes Landnutzungsobjekt hinsichtlich der Zusammensetzung und

Anordnung von Bodenbedeckungsarten innerhalb dessen Begrenzung analysiert und anschließend unter anderem auf Basis dieser Information klassifiziert. Bei dieser Vorgehensweise werden die charakteristischen Eigenschaften implizit in Form von Kontextmerkmalen in den Klassifikationsprozess integriert. Diese Vorgehensweise ermöglicht es, Kontextinformation über die Abhängigkeit zwischen Bodenbedeckung und Landnutzung im Rahmen der Landnutzungsklassifikation zu berücksichtigen. Die Verwendung von Merkmalen birgt jedoch das Problem, dass ein gewisses Maß an Vorwissen über die Eigenschaften und die Abhängigkeiten der Landnutzungsklassen vorliegen muss, um geeignete Merkmale für die Klassifikation zu erhalten.

Die zur Klassifikation der Landnutzung verwendete Methodik reicht von einfachen regelbasierten Ansätzen (z.B. [Banzhaf & Höfer, 2008]) bis hin zu überwachten Klassifikationsmethoden, wie beispielsweise *Random Forests* (RF) (z.B. [Walde et al., 2014; Albert et al., 2014a]) oder *Support Vector Machines* (SVM) (z.B. [Montanges et al., 2015]). Überwachte Klassifikatoren bieten den Vorteil, dass sie anhand von repräsentativen Trainingsgebieten gelernt und somit in einfacher Weise auf andere Gebiete übertragen werden können, wohingegen regelbasierte Ansätze zunächst aufwändig an die veränderten Eigenschaften des neuen Testgebietes anzupassen sind.

In den meisten zweistufigen Ansätzen zur Klassifikation der Landnutzung wird Kontextinformation implizit in Form von sogenannten Kontextmerkmalen in den Klassifikationsprozess integriert (z.B. [Hermosilla et al., 2012]). Die Kontextmerkmale dienen der Beschreibung der statistischen Abhängigkeit der Landnutzung von der Bodenbedeckung. Sie lassen sich grundsätzlich in zwei Gruppen einteilen: *räumliche Metriken* und *graphenbasierte Merkmale* [Hermosilla et al., 2012]. Räumliche Metriken charakterisieren die Zusammensetzung und räumliche Anordnung der Bodenbedeckungssegmente innerhalb eines Landnutzungsobjekts. Sie beschreiben die Größe und Form der Bodenbedeckungssegmente, z.B. in Form des Flächenanteils von Gebäudepixeln an der Gesamtfläche [Hermosilla et al., 2012], sowie ihre räumliche Anordnung in Bezug auf das Landnutzungsobjekt, z.B. in Form der Position von Gebäudesegmenten in Relation zu den Nutzungsgrenzen oder anderen Gebäudesegmenten [Novack & Stilla, 2015]. Die graphenbasierten Merkmale leiten sich aus der Graphentheorie ab und wurden erstmals von Barnsley & Barr [1996] zur Klassifikation der Landnutzung eingesetzt. Graphenbasierte Merkmale beschreiben die räumliche Nachbarschaft (Kanten) zwischen Bodenbedeckungselementen (Knoten) innerhalb eines Landnutzungsobjekts. Folglich charakterisieren sie die Frequenz und die gegenseitige Anordnung der Bodenbedeckungselemente innerhalb eines Landnutzungsobjekts. Zu diesem Zweck berechnen Barnsley & Barr [1996] separat für jedes Landnutzungsobjekt eine Co-Occurrende Matrix, die sogenannte *Adjacency-Event-Matrix*, welche die Nachbarschaftsbeziehungen der Bodenbedeckungselemente auf Pixelniveau beschreibt. Diese Matrix bildet die Grundlage für die Ableitung von Merkmalen, wie z.B. die normalisierte Anzahl von Kanten zwischen bestimmten Bodenbedeckungsklassen. Walde et al. [2014] verwenden eine adaptierte Version der Matrix, die die Nachbarschaftsbeziehungen nicht auf der Ebene von Pixeln sondern auf der Ebene von Segmenten analysiert. Die Verwendung von Kontextmerkmalen im Rahmen der Klassifikation beschränkt sich nicht nur auf die Klassifikation der Landnutzung. Kontextmerkmale finden auch in anderen Bereichen Anwendung, wie z.B. zur Klassifikation von terrestrischen Bildern [Munoz et al., 2010] oder 3D-Punktwolken [Xiong et al., 2011].

Neben der Abhängigkeit von der Bodenbedeckung existieren im Zusammenhang mit der Klassifikation der Landnutzung weitere Arten von Kontext. Hierbei handelt es sich in erster Linie um die

räumliche Abhängigkeit benachbarter Landnutzungsobjekte. Diese Art von Kontextinformation wird von einigen Ansätzen zusätzlich im Klassifikationsprozess berücksichtigt, wobei die Integration auf unterschiedliche Weise erfolgt. Hermosilla et al. (2012) verwenden Kontextmerkmale zur Beschreibung der räumlichen Abhängigkeiten von Landnutzungsobjekten. Konkret modellieren sie Landnutzungsobjekte und deren Nachbarschaftsbeziehungen in einem Graphen, aus dem anschließend diverse Merkmale, wie z.B. die Anzahl von benachbarten Objekten gleicher Klasse, abgeleitet werden. Bei der Verwendung von Kontextmerkmalen wird Kontext implizit in den Klassifikationsprozess integriert. Alternativ lassen sich Beziehungen und Abhängigkeiten zwischen den verschiedenen Bildprimitiven auch explizit in einem graphischen Modell, wie z.B. einem CRF, modellieren. Novack & Stilla [2015] sowie Montanges et al. [2015] nutzen graphische Modelle zur expliziten Modellierung der räumlichen Abhängigkeiten benachbarter Landnutzungsobjekte. Novack & Stilla [2015] klassifizieren die Landnutzung in urbanen Gebieten anhand von satellitengestützten synthetischen Apertur Radar (SAR) Daten. Die Autoren bestimmen in einem ersten Schritt initiale Klassenlabels mit Hilfe von drei verschiedenen Klassifikatoren (RF, Logistische Regression, Nächste-Nachbarn-Klassifikation). Die zugehörigen Wahrscheinlichkeiten werden mittels arithmetischer Mittelwertbildung zusammengefasst. Die Anwendung von CRF erfolgt in einem Nachbearbeitungsschritt, der die Glättung der Ergebnisse zum Ziel hat. Das verwendete CRF modelliert die räumlichen Abhängigkeiten zwischen benachbarten Landnutzungsobjekten. Montanges et al. [2015] führen eine rasterförmige Klassifikation der Landnutzung durch, d.h. sie bestimmen für jede Zelle eines 200 m-Rasters die Landnutzung anhand von Satellitendaten, DOM und thematischer Geoinformation. Zur Klassifikation kommen zwei verschiedene Arten von probabilistischen graphischen Modellen zum Einsatz: *Markov Random Fields* (MRF) und CRF. In beiden graphischen Modellen werden die Ergebnisse eines SVM Klassifikators mit einem Modell der Interaktionen räumlich benachbarter Zellen kombiniert. Die SVM Klassifikation basiert auf Kontextmerkmalen, die sich aus den Bodenbedeckungsergebnissen ableiten, sowie strukturellen Informationen. Dem Modell der räumlichen Abhängigkeit liegt wie bei Novack & Stilla [2015] die Annahme zu Grunde, dass benachbarte Objekte typischerweise zu einer Klasse gehören. Diese Annahme ist insbesondere bei großen Klassifikationseinheiten, die zumeist Flächen gleicher Landnutzung von anderen abgrenzen, oder funktionalen Einheiten, wie z.B. GIS-Objekten, nicht gerechtfertigt.

2.2. Kontextbasierte Klassifikation

Eine gängige Herangehensweise zur Klassifikation von Bildprimitiven besteht darin, jedes Bildprimitiv unabhängig von allen anderen zu klassifizieren. Die Klassifikation erfolgt allein anhand der beobachteten Merkmale des jeweiligen Bildprimitivs. Kontextinformation, wie z.B. die Klassen benachbarter Bildprimitive, wird bei der Entscheidung hingegen nicht berücksichtigt. Der Klassifikation liegt somit die Annahme zugrunde, dass die Bildprimitive in keinerlei Zusammenhang zueinander stehen. Hierbei handelt es sich allerdings um eine grobe Vereinfachung der Realität. Objekte haben typischerweise eine gewisse Ausdehnung in den Sensordaten, die meist mehrere Bildprimitive überdeckt. Innerhalb eines Objekts nehmen die Bildprimitive das gleiche Klassenlabel an. Unterschiedliche Klassenlabels treten nur an den Objekträndern auf, die in Abhängigkeit der Größe der Primitive und der Objekte in den Daten einen flächenmäßig geringen Anteil ausmachen. Folglich sind die Klassenlabels benachbarter

Primitive korreliert. Diese Abhängigkeit wird bei einer unabhängigen Klassifikation vernachlässigt. Die isolierte Betrachtung der einzelnen Bildprimitive führt zu einem verrauschten Ergebnis, das im Allgemeinen nicht der tatsächlichen Konfiguration von Klassenlabels entspricht. In manchen Fällen weisen verschiedene Objektklassen sehr ähnliche Eigenschaften auf, sodass allein anhand der beobachteten Merkmale keine eindeutige Zuweisung zu einer Klasse erfolgen kann. Mit der Berücksichtigung von Kontext, wie z.B. der Information über Klassenlabels benachbarter Bildprimitive, lassen sich Mehrdeutigkeiten auflösen.

Es existieren grundsätzlich zwei verschiedene Herangehensweisen zur Berücksichtigung von Kontext in statistischen Klassifikationsverfahren. Zum einen lässt sich Kontext implizit in Form von Merkmalen in den Klassifikationsprozess integrieren. Hierbei wird zwar jedes Bildprimitiv unabhängig von allen anderen klassifiziert, jedoch mit dem Unterschied, dass zusätzlich Merkmale zur Beschreibung des Kontexts verwendet werden. Bei diesen Merkmalen kann es sich um die Merkmale benachbarter Primitive [Wolf & Bileschi, 2006] oder eigens aus den Klassenlabels oder Merkmalen benachbarter Primitive abgeleitete Größen handeln, die den Kontext adäquat beschreiben. In Abhängigkeit der gewählten Merkmale führt dies zu teils hochdimensionalen Merkmalsvektoren, die eine Klassifikation erschweren. Alternativ lassen sich Kontextbeziehungen, wie z.B. gegenseitige Abhängigkeiten zwischen benachbarten Primitiven, explizit in einem graphischen Modell modellieren. Hierfür kommen insbesondere *Markoff-Zufallsfelder* (MRF) [Geman & Geman, 1984] sowie *bedingte Zufallsfelder* (CRF) [Kumar & Hebert, 2006] in Betracht [Förstner, 2013].

Der Fokus dieser Arbeit liegt auf CRF, die von Lafferty et al. [2001] zur Klassifikation sequentieller Daten in der Texterkennung eingeführt wurden. Kumar & Hebert [2006] wenden bedingte Zufallsfelder erstmals zur Klassifikation von Bilddaten an. CRF sind ungerichtete graphische Modelle zur kontextbasierten Klassifikation, bestehend aus Knoten und ungerichteten Kanten, mit deren Hilfe lokaler Kontext in der Umgebung eines Bildprimitivs statistisch modelliert werden kann. Knoten repräsentieren die zu klassifizierenden Bildprimitive und Kanten modellieren die statistischen Abhängigkeiten zwischen den Bildprimitiven. Im CRF sind standardmäßig ausschließlich paarweise Beziehungen zwischen Bildprimitiven modelliert, d.h. Paare von Knoten sind jeweils mit einer Kante verbunden. Die Knoten entsprechen Zufallsvariablen, die im Fall einer Klassifikation die unbekannten, diskreten Klassenlabels repräsentieren. Außerdem ist jedem Knoten ein Merkmalsvektor zugeordnet, der aus den Eingangsdaten abgeleitet wird und somit die Eigenschaften des jeweiligen Bildprimitivs beschreibt. Im Zuge der Klassifikation wird für das gesamte graphische Modell simultan die wahrscheinlichste Konfiguration an Klassenlabels bestimmt. Die Berechnung stützt sich auf einen Daten- und einen Interaktionsterm. Der Datenterm modelliert den Zusammenhang zwischen den beobachteten Merkmalen und den Klassenlabels der Bildprimitive. Der Interaktionsterm beschreibt das Kontextmodell, das den Zusammenhang zwischen den Klassenlabels benachbarter Knoten und den Daten modelliert. CRF stellen einen flexiblen Rahmen zur Verfügung, der sich individuell an die jeweilige Aufgabenstellung anpassen lässt. So kann für beide Potentiale jeder beliebige probabilistische Klassifikator gewählt werden. Die Flexibilität von CRF spiegelt sich insbesondere in der Vielfalt der verwendeten Klassifikatoren für den Datenterm wider, u.a. verallgemeinerte lineare Modelle (z.B. [Kumar & Hebert, 2006]) und RF (z.B [Schindler, 2012]). Für das Interaktionspotential werden in der Regel einfachere Modelle zur Modellierung von Glattheitsbedingungen verwendet, wie z.B. das Ising- bzw. Potts-Modell [Besag,

1986; Geman & Geman, 1984] oder das kontrastsensitive Potts-Modell [Boykov & Jolly, 2001; Kumar & Hebert, 2006]. Für paarweise CRF existieren effiziente Inferenzmethoden, wie z.B. *Loopy Belief Propagation* (LBP) [Frey & MacKay, 1998] oder *Graph cuts* [Boykov et al., 2001] (siehe [Szeliski et al., 2008] für einen Vergleich der Inferenzverfahren).

CRF haben sich als flexibles, leistungsfähiges Verfahren zur kontextbasierten Klassifikation in vielfältigen Anwendungen im Bereich des computergestützten Sehens (engl.: *Computer Vision*) sowie im Bereich der Photogrammetrie und Fernerkundung bewährt. In der Computer Vision kommen CRF mit paarweisen Relationen insbesondere bei der Objekterkennnung und -segmentierung in Bildern zum Einsatz, wie z.B. in den Arbeiten von Shotton et al. [2009].

Im Bereich der Photogrammetrie und Fernerkundung finden paarweise CRF beispielsweise in folgenden Arbeiten Anwendung: Zhong & Wang [2007] verwenden mehrere Varianten von CRF zur Detektion von bebauten Gebieten in optischen Satellitenbildern. Lu et al. [2009] nutzen CRF zur Filterung von Geländemodellen aus 3D-Punktwolken, die mittels flugzeuggestütztem Laserscanning erfasst wurden. Niemeyer et al. [2014] klassifizieren flugzeuggestützte 3D-Laserpunktwolken punktweise hinsichtlich ihrer Objektzugehörigkeit. Schindler [2012] vergleicht verschiedene Methoden zur Berücksichtigung von Kontextinformation bei der Klassifikation der Bodenbedeckung anhand von Luftbilddaten und belegt, dass sich Glattheitsbedingungen am effektivsten mit Hilfe von CRF in den Klassifikationsprozess integrieren lassen. Aber auch die zuvor genannten Arbeiten attestieren den CRF eine verbesserte Klassifikationsgenauigkeit gegenüber unabhängigen Ansätzen.

In allen hier dargestellten Ansätzen modelliert das CRF die Klassifikation von Eingangsdaten eines bestimmten Maßstabs und Zeitpunkts. Die Kanten verbinden direkt benachbarte Bildprimitive und beschreiben daher nur lokale Interaktionen. Allerdings beschränkt sich die Abhängigkeit zwischen den Bildprimitiven in vielen Fällen nicht auf die direkte Nachbarschaft. Vielmehr bestehen weitreichendere Abhängigkeiten z.B. räumlich gesehen in größerer Entfernung oder zeitlich gesehen zwischen unterschiedlichen Zeitpunkten. Eine Möglichkeit zur Modellierung von Interaktionen zwischen verschiedenen Zeitpunkten oder Maßstäben besteht darin, Klassifikationsaufgaben einer Szene in unterschiedlichen Maßstäben bzw. zu unterschiedlichen Zeitpunkten in einem graphischen Modell zusammenzuführen, wobei jede Klassifikationsaufgabe eine Ebene des graphischen Modells bildet. Dies bietet den Vorteil, dass Interaktionen über zeitliche als auch räumliche Distanzen explizit modelliert werden können und somit im Rahmen der Klassifikation Berücksichtigung finden.

CRF mit mehreren Ebenen wurden in der Photogrammetrie und Fernerkundung für verschiedenste Fragestellungen angewendet, wie z.B. zur Realisierung eines hierarchischen Ansatzes zur Klassifikation von Gebäudefassaden [Yang & Förstner, 2011]. In dem Ansatz von Yang & Förstner [2011] werden zunächst terrestrische Bilder von Fassaden in mehreren Skalenniveaus mittels einer multiskaligen Wasserscheidentransformation segmentiert. Anschließend modellieren die Autoren räumliche und multiskalige Relationen zwischen den Segmenten in einem Mehr-Ebenen-CRF, wobei die Anzahl der Skalenniveaus der Anzahl der Ebenen im CRF entspricht. Außerdem wurden Mehr-Ebenen-CRF bereits zur Klassifikation von Luftbildern und DOM unter Berücksichtigung von Verdeckungen [Kosov et al., 2013] und zur multitemporalen und multiskalaren Klassifikation der Bodenbedeckung anhand von Satellitenbildern [Hoberg et al., 2015] eingesetzt. Kosov et al. [2013] modellieren die Klassenlabels

des verdeckten und verdeckenden Objekts für jedes Bildprimitiv in zwei separaten Ebenen eines CRF. Die Knoten in beiden Ebenen sind jeweils durch eine Kante im graphischen Modell verbunden, die deren Abhängigkeit explizit modelliert. Hoberg et al. [2015] modellieren multitemporale und multiskalige Abhängigkeiten von Fernerkundungsdaten zu verschiedenen Zeitpunkten in einem Mehr-Ebenen-CRF, wobei jede Ebene im CRF den Zustand einer Epoche beschreibt. Die Auflösung der Satellitenbilder kann dabei zwischen den Epochen variieren. Das graphische Modell enthält zwei verschiedene Arten von Kanten: Kanten innerhalb einer Ebene modellieren die räumliche Abhängigkeit der Bildprimitive einer Epoche und Kanten zwischen den Ebenen die zeitliche Abhängigkeit der Bildprimitive zu unterschiedlichen Zeitpunkten.

In den meisten Mehr-Ebenen-CRF-Modellen wird in jeder Ebene die gleiche Klassenstruktur unterschieden. Lediglich in multiskaligen Ansätzen kann die Klassenstruktur zwischen den einzelnen Ebenen in geringem Maße variieren. Dies liegt darin begründet, dass sich das Erscheinungsbild der Objekte im Skalenraum verändert. Lassen sich beispielsweise in der Originalauflösung noch die einzelnen Bestandteile eines Objekts, z.B. Autoreifen, Karosserieteile, etc., differenzieren, so ist in einer gröberen Auflösung meist nur noch das zusammengesetzte Objekt, z.B. das Auto, zu erkennen.

Die genannten Ansätze verwenden ausschließlich paarweise Potentiale um das gemeinsame Auftreten von Klassen oder Glattheitsbedingungen zu modellieren. Somit beschränken sie sich hauptsächlich auf die Modellierung paarweiser Interaktionen. Bei vielen Fragestellungen, wie z.B. auch bei der gemeinsamen Klassifikation der Bodenbedeckung und der Landnutzung, bietet es sich hingegen an, mehrere Bildprimitive in einem größeren Radius miteinander zu verknüpfen, d.h. ein komplexes Geflecht an Interaktionen über größere Distanzen (engl.: *long-range interactions*) [Luo & Sohn, 2014; Gould et al., 2008], explizit in die Modelle zu integrieren. Hierbei stoßen paarweise Modelle allerdings an ihre Grenzen. Beispielsweise lassen sich komplexe Abhängigkeiten zwischen mehr als zwei Zufallsvariablen, wie z.B. die Konfiguration mehrerer Bodenbedeckungssegmente innerhalb eines Landnutzungsobjekts, nur bedingt mit Hilfe von paarweisen Potentialen modellieren.

Verbindungen von mehr als zwei Zufallsvariablen bilden Cliquen höherer Ordnung, die über ein Potential höherer Ordnung verknüpft sind. Diese Art von Funktion verbindet Gruppen von Knoten anstatt Paaren in einem graphischen Modell und ermöglicht so die explizite Modellierung der komplexen Abhängigkeiten zwischen mehr als zwei Knoten. Potentiale höherer Ordnung werden beispielsweise von Kohli et al. [2009] zur Bildklassifikation verwendet. Hierbei besteht das Ziel darin, unterschiedliche Segmentierungsergebnisse eines Bildes in der Art zu kombinieren, um eine optimale Abgrenzung und Klassifizierung von Objekten zu erreichen. Zu diesem Zweck führen die Autoren eine neue Klasse von Potentialen höherer Ordnung ein, das sogenannte P^N *Potts-Modell*. Dieses Potential verknüpft alle Pixel, die einem Segment zugehörig sind, zu einer Clique höherer Ordnung und belegt die Pixel mit einem Strafterm, sofern nicht der überwiegende Anteil der Pixel zur gleichen Klasse gehört. Folglich favorisiert dieser Term die Konsistenz von Klassenlabels innerhalb eines Segments, d.h. dass alle Pixel innerhalb des Segments das gleiche Klassenlabel annehmen. Es handelt sich hierbei um eine Verallgemeinerung des Potts-Modells für Cliquen höherer Ordnung.

In der Photogrammetrie und Fernerkundung wurden Potentiale höherer Ordnung basierend auf dem P^N Potts-Modell beispielsweise zur Extraktion von Straßennetzen [Wegner et al., 2013] bzw.

Gebäuden [Montoya-Zegarra et al., 2015] aus Luftbildern eingesetzt. Beide Ansätze verwenden CRF höherer Ordnung, um a priori Wissen über die zu extrahierenden Objekte (Straßen [Wegner et al., 2013], Gebäude [Montoya-Zegarra et al., 2015]) in den Klassifikationsprozess zu integrieren. Zu diesem Zweck werden jene Bildprimitive in einer Clique höherer Ordnung zusammengefasst, die mit hoher Wahrscheinlichkeit zu einem Objekt gehören. Die Kombination von Bildprimitiven in einer Clique und die Auswahl relevanter Cliquen erfolgt im Rahmen eines Samplingansatzes, der neben den spektralen Eigenschaften der Bildprimitive zusätzlich die typische Form der Objekte in die Auswahl mit einbezieht. Die Cliquen werden unter Verwendung des P^N Potts-Modells als Objekthypothesen in das CRF eingeführt. Ladický et al. [2009] wenden das P^N Potts-Modell in einem hierarchischen CRF-Modell an.

Die Inferenz stellt den limitierenden Faktor bei der Anwendung von CRF höherer Ordnung dar. Kohli et al. [2009] zeigen, dass die spezielle Struktur des P^N Potts-Modells eine effiziente Inferenz mit Hilfe von Graph Cuts [Boykov et al., 2001] ermöglicht. Speziell für allgemeine Formulierungen der Potentiale höherer Ordnung gestaltet sich die Inferenz jedoch schwierig. So lassen sich die standardmäßigen Inferenzmethoden (z.B. Loopy Belief Propagation (LBP) [Frey & MacKay, 1998]) nur effektiv auf Potentiale höherer Ordnung anwenden, die auf einer begrenzten Anzahl an Variablen basieren. Potetz & Lee [2008] beschreiben zwar eine Erweiterung von LBP für Potentiale höherer Ordnung, die jedoch ebenfalls Einschränkungen im Hinblick auf die dabei genutzten Modelle mit sich bringt.

Roig et al. [2011] überwinden das Problem der ineffizienten Inferenz in CRF höherer Ordnung mit Hilfe eines iterativen Inferenzalgorithmus. Der Ansatz hat zum Ziel, Objekte in mehreren Ansichten einer Szene simultan zu klassifizieren. Die Autoren bestimmen eine genäherte Lösung für das zugrunde liegende CRF höherer Ordnung, indem sie in jeder Iteration des iterativen Inferenzalgorithmus die Energiefunktion approximieren. Dies wird ermöglicht, indem sie die Zufallsvariablen der einzelnen Ansichten zur Bestimmung der Potentiale höherer Ordnung in jeder Iteration als konstant annehmen, womit sich die Terme höherer Ordnung zu Datentermen vereinfachen. Durch diesen Schritt ist es möglich, die Energiefunktion mit standardmäßigen Inferenzmethoden zu minimieren. Die Autoren bestimmen in jeder Iteration auf Grundlage der approximierten Energiefunktion eine Zwischenlösung, die in der nächsten Iteration zur erneuten Berechnung der Zufallsvariablen und damit zur Aktualisierung der Terme höherer Ordnung herangezogen wird. Dieses Verfahren wird solange fortgesetzt bis der Optimierungsprozess konvergiert. Die verwendeten Potentiale höherer Ordnung modellieren relativ einfache Abhängigkeiten; sie dienen der Berücksichtigung von Verdeckungen zwischen Objekten innerhalb einer Szene sowie der Wahrung der Konsistenz der Klassifikationsergebnisse zwischen verschiedenen Ansichten.

2.3. Diskussion

Die meisten der in Kapitel 2.1 vorgestellten Ansätze zur Klassifikation der Landnutzung unterscheiden nur eine geringe Anzahl von Nutzungsarten oder sind auf bestimmte Anwendungsgebiete spezialisiert. Um die Verifikation eines detaillierten räumlichen Landnutzungsdatenbestandes effektiv durchführen zu können, ist es jedoch erforderlich, dass das Klassifikationsergebnis bereits eine hohe semantische

Auflösung liefert. Zur Unterscheidung einer feinen Klassenstruktur wird z.B. von Banzhaf & Höfer [2008] zusätzlich thematische Information herangezogen. In einem Verifikationsprozess steht diese Art von Daten jedoch meist nicht in der erforderlichen Qualität bzw. Aktualität zur Verfügung. Aus der allgemeinen Zielsetzung zur Realisierung eines automatisierten Verifikations- und Aktualisierungsprozesses leiten sich verschiedene Anforderungen an die Klassenstruktur ab. Zum einen sollte der Ansatz eine möglichst feine Klassenstruktur unterscheiden, wobei ausschließlich Sensordaten verwendet werden sollen. Damit ist das Verfahren weitgehend unabhängig von der Verfügbarkeit zusätzlicher Datenquellen. Lediglich zur Bereitstellung von Trainingsdaten (in Form von Geometrie und Semantik in Trainingsgebieten) sowie zur Definition der Klassifikationseinheiten (in Form der Geometrie im gesamten Anwendungsgebiet) sind Zusatzinformationen in Form von räumlichen Bodenbedeckungs- und Landnutzungsdatenbeständen erforderlich. Zum anderen sollte die Klassenstruktur möglichst alle gängigen Nutzungsarten enthalten, um die Anwendbarkeit des Verfahrens in urbanen sowie ländlichen Gebieten zu gewährleisten und somit die Landschaft vollständig beschreiben zu können [Arnold et al., 2017].

Wie in Kapitel 2.1 beschrieben, wird die aktuelle Nutzungsart typischerweise in einem zweistufigen Ansatz prädiziert (z.B. [Hermosilla et al., 2012; Walde et al., 2014]). Bei der zweistufigen Vorgehensweise werden beide Klassifikationsaufgaben konsekutiv prozessiert. Ein Austausch von Kontextinformation findet nur in eine Richtung statt, d.h. von der Bodenbedeckung zur Landnutzung. Bei der Klassifikation der Bodenbedeckung findet die Zusatzinformation über die vorliegende Nutzungsart hingegen keine Berücksichtigung. Vor dem Hintergrund, dass die Bodenbedeckung und Landnutzung eine gegenseitige Abhängigkeit aufweisen, wird die semantische Abhängigkeit bei einer zweistufigen Vorgehensweise nicht vollständig berücksichtigt. Daneben besteht ein wesentlicher Nachteil eines zweistufigen Verfahrens darin, dass Fehlklassifikationen der Bodenbedeckung, die im ersten Schritt auftreten, im zweiten Schritt nicht wieder behoben werden können. Dies ist insbesondere vor dem Hintergrund kritisch, dass die Bodenbedeckung einen unmittelbaren Einfluss auf das Ergebnis der Landnutzungsklassifikation hat. Folglich können falsch klassifizierte Bodenbedeckungen leicht zu Fehlern in der Landnutzungsklassifikation führen. Darüber hinaus werden bei den meisten zweistufigen Ansätzen lediglich die Klassenlabels weiterverarbeitet, wohingegen die mitunter aussagekräftigen Unsicherheiten der Klassenlabels keine Berücksichtigung finden.

Mit der Integration der Bodenbedeckung und der Landnutzung in einen gemeinsamen Klassifikationsansatz wird eine einheitliche und konsistente Modellierung der Abhängigkeiten realisiert, die in dieser Form nach bestem Wissen der Autorin noch nicht umgesetzt ist. Zur Modellierung von statistischen Abhängigkeiten haben sich CRF bewährt. In Mehr-Ebenen-CRF lassen sich Abhängigkeiten zwischen verschiedenen Klassifikationsaufgaben modellieren, indem die Bildprimitive in den Ebenen miteinander verbunden werden. Die gleichzeitige Bestimmung der wahrscheinlichsten Konfiguration an Klassenlabels in beiden Ebenen bietet verschiedene Vorteile. Zum einen wird hierbei vermieden, dass mangels ausreichender Erkenntnisse verfrüht falsche Entscheidungen getroffen werden. Fehlklassifikationen können im Laufe des Verfahrens wieder behoben werden, insbesondere dann, wenn die Entscheidung mit einer hohen Unsicherheit behaftet ist, d.h. eine geringe Konfidenz aufweist. Zum anderen ermöglicht die gemeinsame Modellierung in einem CRF, dass sich die Klassifikationsaufgaben während der Inferenz gegenseitig beeinflussen und damit in ihrer Entscheidungsfindung unterstützen

können. Im Gegensatz zu den in Kapitel 2.2 genannten Mehr-Ebenen-CRF, die in jeder Ebene eine identische oder zumindest ähnliche Klassenstruktur unterscheiden, wird bei der Zusammenführung der Bodenbedeckungs- und Landnutzungsklassifikation in einem Mehr-Ebenen-CRF in jeder Ebene eine andere Klassenstruktur unterschieden.

In diversen Arbeiten wurde bereits gezeigt, dass die räumliche Abhängigkeit bei der Klassifikation der Landnutzung von Bedeutung ist. In der Literatur existieren verschiedene Ansätze zur Berücksichtigung dieser Abhängigkeiten im Klassifikationsprozess. Hermosilla et al. [2012] integrieren die räumlichen Nachbarschaftsbeziehungen in Form von Merkmalen in die Klassifikation. Novack & Stilla [2015] und Montanges et al. [2015] modellieren die räumlichen Abhängigkeiten als paarweise Interaktionspotentiale in einem CRF, wobei es sich nur um sehr einfache Modelle auf Grundlage der Annahme handelt, dass benachbarte Bildprimitive der gleichen Nutzungsart angehören. Diese Annahme ist insbesondere bei großen Klassifikationseinheiten, die zumeist Flächen gleicher Landnutzung von anderen abgrenzen, nicht gerechtfertigt. Je größer die Einheiten, desto größer ist die Wahrscheinlichkeit, dass sich die Klassenlabels benachbarter Objekte voneinander unterscheiden. In dem in dieser Arbeit entwickelten CRF-Modell werden separat in jeder Ebene die räumlichen Abhängigkeiten zwischen benachbarten Bildprimitiven, und zwar Superpixeln im Fall der Bodenbedeckungsklassifikation und Segmenten im Fall der Landnutzungsklassifikation, modelliert. Jedoch werden anders als bei Novack & Stilla [2015] und Montanges et al. [2015] nicht generell gleiche Klassenlabels, sondern wahrscheinliche Konfigurationen von Klassenlabels an benachbarten Bildprimitiven favorisiert. Dieses Modell hat sich in vorangegangenen Arbeiten [Albert et al., 2014a, 2015] als für diese Art von Abhängigkeit besser geeignet herausgestellt. Die Wahrscheinlichkeit für das Auftreten bestimmter Konfigurationen von Nutzungsarten bei gegebenen Beobachtungen wird aus realen Daten gelernt. Die Verwendung von Klassifikatoren zur Bestimmung des Interaktionspotentials hat sich u.a. in den Arbeiten von [Nowozin et al., 2011; Niemeyer et al., 2014] als vorteilhaft für die Klassifizierung von komplexen Interaktionen zwischen den Klassen erwiesen.

Die Modellierung der statistischen Abhängigkeit zwischen der Bodenbedeckung und der Landnutzung erfolgte bislang nur implizit in Form von Kontextmerkmalen, z.B. [Hermosilla et al., 2012]. Mit der Vereinigung beider Aufgaben in einem CRF-Modell ist es möglich, neben den räumlichen Abhängigkeiten benachbarter Bildprimitive zusätzlich die semantischen Abhängigkeiten zwischen den Ebenen explizit zu modellieren. Ein solches Zwei-Ebenen-CRF stellt eine konsistente Modellierung dar, in der die Abhängigkeit zwischen Bodenbedeckung und Landnutzung anhand von repräsentativen Trainingsdaten gelernt wird.

Für ein solches Zwei-Ebenen-Modell ist entscheidend, dass die gegenseitige Abhängigkeit zwischen der Bodenbedeckung und der Landnutzung in geeigneter Weise modelliert ist. In Albert et al. [2014b] beschränkt sich die Modellierung auf ein relativ einfaches Modell, das nur paarweise Abhängigkeiten berücksichtigt. Konkret wird hierbei die Wahrscheinlichkeit für das gemeinsame Auftreten von Klassen unter Berücksichtigung der Daten bestimmt. Bei dieser Herangehensweise ist jedes Superpixel mit dem räumlich überlagernden Landnutzungsobjekt mit einer Kante verbunden, für die das paarweise Interaktionspotential als gemeinsame a posteriori Wahrscheinlichkeit modelliert ist. Die Beschränkung auf paarweise Relationen hat sich allerdings als ungeeignet herausgestellt, da sich die jeweiligen Klassen der Bodenbedeckung und der Landnutzung nicht eindeutig einander zuordnen lassen,

d.h. aus einer Bodenbedeckung lässt sich nicht eindeutig auf eine Landnutzung schließen und umgekehrt. Zur Bestimmung der Landnutzung ist oftmals nicht eine einzelne Bodenbedeckung, sondern vielmehr das Arrangement verschiedener Bodenbedeckungen von Bedeutung. Folglich besteht zwischen den Landnutzungsobjekten und den darin enthaltenen kleinräumigen Bodenbedeckungssegmenten ein komplexes Abhängigkeitsverhältnis, das sich nicht in paarweisen Modellen abbilden lässt.

Um die komplexen Abhängigkeiten in adäquater Weise zu modellieren, bietet es sich an, die zueinander in Beziehung stehenden Primitive in einer Clique höherer Ordnung zu verbinden und deren Abhängigkeit mit einem Interaktionspotential höherer Ordnung zu modellieren. Nach bestem Wissen der Autorin existiert bis dato kein Verfahren zur Klassifikation der Landnutzung, das die komplexen Abhängigkeiten zwischen Bodenbedeckung und Landnutzung explizit in einem Graphen modelliert. Die Modellierung von Potentialen höherer Ordnung unterliegt diversen Einschränkungen, z.B. hinsichtlich der Inferenz. Daher wurden in der Literatur spezielle Funktionen eingeführt, die eine effiziente Inferenz ermöglichen, wie z.B. das robuste P^N Potts-Modell [Kohli et al., 2009]. Dieses Modell verknüpft alle Pixel innerhalb eines Segments und favorisiert Segmentierungen, in denen die Mehrheit der Pixel der gleichen Klasse zugehörig sind [Kohli et al., 2009]. Die Abhängigkeit zwischen den Bodenbedeckungs- und Landnutzungsprimitiven ist allerdings von einer anderen Natur. So ist es nicht das Ziel, allen Bodenbedeckungssegmenten innerhalb eines Landnutzungsobjekts das gleiche Klassenlabel zuzuweisen. Vielmehr zeichnet sich ein Landnutzungsobjekt durch eine spezifische Konfiguration unterschiedlicher Bodenbedeckungsarten aus. Beispielsweise enthält ein Objekt der Nutzungsart *Wohnbaufläche* typischerweise die Bodenbedeckungen *Gebäude*, *Vegetation* und *Versiegelung*. Diese Abhängigkeit lässt sich nicht mit Hilfe eines P^N Potts-Modell abbilden, sondern erfordert ein allgemeineres Modell. So ist es das Ziel, das gemeinsame Auftreten der Bodenbedeckungs- und Landnutzungsklassen anhand von repräsentativen Trainingsdaten zu lernen. Darüber hinaus ist in der von Kohli et al. [2009] vorgestellten Formulierung nicht vorgesehen, eine Klasse für das übergeordnete Segment zu bestimmen, d.h. das Segment dient allein zur Verknüpfung der zusammengehörigen Bildprimitive. In dem Zwei-Ebenen-CRF-Modell entspricht das überlagernde Segment dem Landnutzungsobjekt, dessen Klassenlabel gemäß der Zielsetzung des Verfahrens ebenfalls zu bestimmen ist. Bei der generischen Formulierung von Potentialen höherer Ordnung resultieren die größten Einschränkungen aus der Anzahl der Zufallsvariablen innerhalb einer Clique. Eine Clique höherer Ordnung kann grundsätzlich eine hohe Zahl an Knoten enthalten. Mit der Erhöhung der Anzahl beteiligter Knoten an einer Clique steigt die Anzahl aller möglichen Klassenkonfigurationen exponentiell an. Dies führt dazu, dass es bereits ab einer geringen Anzahl von Zufallsvariablen rechnerisch schwierig bzw. unlösbar ist, mit Hilfe eines statistischen Klassifikators alle möglichen Klassenkonfigurationen zu lernen. Für bestimmte Potentiale, und zwar jene die die Submodularitätsbedingung erfüllen, existieren zwar effiziente Inferenzalgorithmen [Kohli et al., 2009], jedoch lässt sich bei einer allgemeinen Formulierung in der Regel nicht garantieren, dass diese Funktion die Submodularitätsbedingung erfüllt. Folglich existieren keine effizienten Inferenzmethoden für CRF höherer Ordnung, die eine allgemeine Formulierung der Potentiale höherer Ordnung verwenden bei einer gleichzeitig hohen Anzahl an Zufallsvariablen pro Clique.

In dieser Arbeit wird ein iterativer Inferenzalgorithmus vorgestellt, der die optimale Konfiguration von Klassenlabels nicht simultan für das gesamte graphische Modell bestimmt, wie es üblicherweise in CRF erfolgt, sondern im Rahmen einer iterativen Prozedur. Die Grundidee zur iterativen Prozes-

sierung orientiert sich an dem *Expectation Maximisation* (EM) Ansatz [Dempster et al., 1977], der im gegenseitigen Abhängigkeitsverhältnis stehende Parametergruppen abwechselnd unter Festhaltung der jeweils anderen Parameter bestimmt und die Parameterwerte somit kontinuierlich verfeinert. In Analogie zu EM werden im Rahmen der iterativen Inferenzprozedur anstelle von Parametern die Klassenlabels in beiden Ebenen abwechselnd bestimmt, wobei der Grundgedanke darin besteht, dass die Klassenlabels der jeweils anderen Ebene als Kontextinformation die Klassifikation in beiden Ebenen verbessert, was wiederum zu einer verfeinerten und aussagekräftigeren Kontextinformation führt. Ein iterativer Inferenzalgorithmus bietet ein hohes Maß an Flexibilität zur Anpassung an die jeweilige Klassifikationsaufgabe. Somit lässt sich mit diesem Verfahren auch eine Variante des etablierten zweistufigen Verfahrens abbilden, indem nur ein Iterationsschritt durchgeführt wird. Im Gegensatz zum zweistufigen Verfahren erfolgt die zweistufige Prozessierung simultan in beide Richtungen, d.h. die Vorteile für die Bodenbedeckung aus der gemeinsamen Modellierung der Klassifikationsaufgaben bleiben bestehen.

Der hier entwickelte iterative Inferenzalgorithmus ist inspiriert von Roig et al. [2011]. In ihrem Verfahren wird Kontextinformation zwischen den verschiedenen Ansichten einer Szene ausgetauscht, indem die zu Datentermen vereinfachten Potentiale höherer Ordnung in jeder Iteration aktualisiert werden. Mit der Vereinfachung der Potentiale höherer Ordnung vereinfachen sich auch die Trainings- und die Inferenzprozesse, weil die entsprechenden Algorithmen in einfacher Weise von den paarweisen CRF übernommen und im Rahmen der iterativen Prozedur ausgeführt werden können.

Die Aktualisierung der Potentiale höherer Ordnung erfolgt in [Roig et al., 2011] ausschließlich auf Basis der Klassenlabels der jeweils aktuellen Zwischenlösung. Zusatzinformationen, wie z.B. die zugehörigen Konfidenzwerte, werden hingegen nicht berücksichtigt. Diese Art von Informationen ist besonders dann relevant, wenn sich der Klassifikator unsicher ist und eine andere Klasse unter Umständen nahezu gleich wahrscheinlich ist. Darüber hinaus modellieren die Potentiale höherer Ordnung in dem Ansatz von Roig et al. [2011] relativ einfache Abhängigkeiten. Insbesondere nehmen die Potentiale nur binäre Zustände an, ganz gleich, wie viele Knoten beteiligt sind. Das Potential zur Modellierung der Relation zwischen der Bodenbedeckung und der Landnutzung bildet weitaus komplexere Abhängigkeiten ab und nimmt in Abhängigkeit der Anzahl beteiligter Knoten und deren Zustände unterschiedliche Werte an. Da die Anzahl der Zufallsvariablen nicht beschränkt ist, ist es nicht möglich, einen Klassifikator für alle Kombinationen anzulernen.

Aus diesem Grund wird die Energiefunktion weiter vereinfacht, indem die zu Datentermen vereinfachten Potentiale höherer Ordnung nicht direkt auf Grundlage der einzelnen Knoten, sondern auf Basis von Kontextmerkmalen bestimmt werden, welche den Zustand aller beteiligten Knoten charakterisieren. Wie bereits in Kapitel 2.1 dargestellt, handelt es sich hierbei um eine weit verbreitete Strategie zur Berücksichtigung von Kontext im Rahmen der Klassifikation der Landnutzung. Im Gegensatz zu den in der Literatur verwendeten Kontextmerkmalen, die sich ausschließlich aus den Klassenlabels ableiten, werden in dieser Arbeit Merkmale ergänzt, welche die Unsicherheit der Klassenlabels berücksichtigen. Diese Art von Merkmalen ist inspiriert von Munoz et al. [2010] und Xiong et al. [2011], die Kontextmerkmale in ihren Ansatz integrieren, um Abhängigkeiten zwischen verschiedenen Hierarchieebenen zu beschreiben.

3. Grundlagen

Die Darstellung der Grundlagen beginnt mit einer kurzen allgemeinen Einführung in die Aufgabe der Klassifikation von Bildern in Kapitel 3.1. Die nachfolgenden Kapitel beschreiben die Grundlagen des in dieser Arbeit entwickelten Klassifikationsverfahrens mit besonderem Fokus auf die Merkmalsextraktion in Kapitel 3.2, die Klassifikation mit RF in Kapitel 3.3 und die statistische Modellierung unter Verwendung von CRF in Kapitel 3.4.

3.1. Klassifikation von Bildern

Im Rahmen der Klassifikation von Bildern gilt es, eine Menge von N_S Bildprimitiven (Pixel, Segmente) $\mathcal{S} = \{1, \ldots, N_S\}$ zu klassifizieren. Der Vektor $\mathbf{x}_i$ bezeichnet die in einem Bildprimitiv $i \in \mathcal{S}$ beobachteten Merkmale (z.B. die in einem Pixel aufgezeichnete Strahlungsintensität). Die Gesamtheit aller Beobachtungen eines Bildes wird beschrieben durch $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_{N_S})^T$. Jedem Primitiv i ist ein D-dimensionaler Merkmalsvektor $\mathbf{f}_i = (f_{i_1}, \ldots, f_{i_D})^T \in \mathbb{R}^D$ zugeordnet, der neben den beobachteten Merkmalen weitere Merkmale enthalten kann, die Funktionen der ursprünglichen Beobachtungen einzelner Pixel bzw. aller Pixel innerhalb eines Segments oder einer lokalen Umgebung eines Pixels darstellen. Die Abhängigkeit des Merkmalsvektors $\mathbf{f}_i$ eines Bildprimitivs i von Teilen oder der Gesamtheit der Beobachtungen wird ausgedrückt durch $\mathbf{f}_i = \mathbf{f}_i(\mathbf{x})$ [Rottensteiner, 2017]. Das Kapitel 3.2 gibt einen Überblick über die in dieser Arbeit verwendeten Merkmale.

Das allgemeine Ziel einer Klassifikation ist es, jedem Bildprimitiv i anhand des zugehörigen Merkmalsvektors $\mathbf{f}_i(\mathbf{x})$ ein Klassenlabel als Zufallsvariable $y_i \in \mathcal{L} = \{l_1, \ldots, l_M\}$ zuzuweisen. Das Klassenlabel entspricht einer diskreten, kategorischen, ungeordneten Variable und repräsentiert die Zugehörigkeit zu einer thematischen Klasse. $\mathcal{L}$ repräsentiert die Menge aller Klassenlabels und M entspricht der Anzahl der Elemente in $\mathcal{L}$ und somit der Anzahl der Klassen. Die Klassenlabels aller N_S Bildprimitive werden in einem Labelvektor $\mathbf{y} = (y_1, \ldots, y_i, \ldots, y_{N_S})^T$ zusammengefasst. Es gilt, eine Funktion zu definieren, welche die wahrscheinlichste Konfiguration der Klassenlabels $\mathbf{y}$ aus den Beobachtungen $\mathbf{x}$ prädiziert. Bei einer lokalen Klassifikation wird ein Klassenlabel y_i individuell für jedes Bildprimitiv i, d.h. unabhängig von allen anderen Primitiven, als Funktion des zugehörigen Merkmalsvektors $\mathbf{f}_i(\mathbf{x})$ bestimmt [Rottensteiner, 2017]. Zu den lokalen Klassifikationsverfahren zählt beispielsweise der in Kapitel 3.3 beschriebene RF Klassifikator, der ein Element der Methodik dieser Arbeit bildet. Graphische Modelle, wie z.B. CRF, bestimmen die optimale Konfiguration aller Klassenlabels $\mathbf{y}$ eines Bildes bei gegebenen Daten $\mathbf{x}$ unter Berücksichtigung von statistischen Abhängigkeiten (Kontext). Die Entscheidung basiert somit nicht allein auf dem individuellen Merkmalsvektor, sondern zusätzlich auf den statistischen Abhängigkeiten der Bildprimitive untereinander. Das Kapitel 3.4 erläutert die theoreti-

schen Grundlagen für die statistische Modellierung von Kontext mittels CRF, die das Grundgerüst der entwickelten Methodik bilden. Bei einer überwachten Klassifikation, wie sie auch in dieser Arbeit Anwendung findet, wird der Klassifikator anhand von Trainingsdaten gelernt. Zu diesem Zweck steht ein Trainingsdatensatz $\mathcal{T} = \{(\mathbf{x}, y)^i\}_{i=1,\ldots,N_T}$ zur Verfügung, der sich aus einer Menge von N_T unabhängigen Trainingsbeispielen zusammensetzt, für die jeweils der Merkmalsvektor $\mathbf{f}_i(\mathbf{x})$ als Eingabedaten und das Klassenlabel y_i als gewünschte Ausgabe bekannt ist [Hänsch & Hellwich, 2017].

3.2. Merkmale

Die Klassifikation von Bildern basiert auf Merkmalen, die für die jeweiligen Bildprimitive (Pixel, Segmente) in einem Vorverarbeitungsschritt aus den Eingangsdaten abgeleitet werden. Die extrahierten Merkmale beschreiben die Eigenschaften der Bildprimitive. Das Ziel der Merkmalsextraktion ist es, aus den Eingangsdaten diskriminative Merkmale abzuleiten. Die extrahierten Merkmale werden in einem Merkmalsvektor $\mathbf{f}_i(\mathbf{x})$ pro Bildprimitiv i zusammengefasst, der die Grundlage für die Klassifikationsentscheidung bildet.

Das Klassifikationsergebnis hängt maßgeblich von der Wahl geeigneter Merkmale ab. Eine gute Trennung der Klassen auf der Grundlage von diskriminativen Merkmalen ist nur möglich, wenn sich die Ballungen (Cluster) der Klassen im Merkmalsraum nicht überlappen. In vielen realen Anwendungen treten jedoch Überlagerungen auf, was zu Fehlklassifikationen führen kann. Die intuitive Annahme, dass eine größere Anzahl von Merkmalen zu einer besseren Unterscheidung der Klassen beiträgt, ist vielfach nicht gerechtfertigt. Eine Verbesserung wird nur erzielt, wenn die neu hinzugefügten Merkmale zusätzliche Information enthalten, d.h. nicht mit anderen Merkmalen korreliert sind, und sich für die einzelnen Klassen unterschiedlich ausprägen. Außerdem besagt das *Hughes-Phänomen* [Hughes, 1968], dass eine Erhöhung der Dimensionalität des Merkmalsraums bei gleich bleibenden Trainingsdaten ab einem bestimmten Punkt zu einer Verschlechterung der Klassifikationsgenauigkeit führt. Demzufolge ist es sinnvoll, die Dimension des Merkmalsvektors durch eine Auswahl wichtiger Merkmale zu begrenzen. Zu diesem Zweck gilt es, jene Merkmale aus der Gesamtheit der Merkmale zu selektieren, die zur Lösung des Problems am besten geeignet sind. Die Auswahl geeigneter Merkmale kann beispielsweise manuell erfolgen. Dies erfordert allerdings ein gewisses Maß an Modellwissen über die zu unterscheidenden Klassen. Bei dieser Vorgehensweise ist es leicht möglich, dass relevante Merkmale übersehen oder Korrelationen zwischen den Merkmalen unerkannt bleiben. Daher bietet es sich an, zunächst einen umfangreichen Satz an Merkmalen abzuleiten und diesen während des Trainings automatisch auf die relevanten Merkmale zu reduzieren [Rottensteiner, 2017]. Die Auswahl kann hierbei anhand eines Relevanzmaßes erfolgen, z.B. dem in Kapitel 3.3 vorgestellten RF Wichtigkeitsmaß [Breiman, 2001], das die einzelnen Merkmale hinsichtlich ihres Einflusses auf die Klassifikation bewertet.

Die folgende Darstellung konzentriert sich auf jene Merkmale, die in der Arbeit verwendet werden, und erhebt damit keinen Anspruch auf Vollständigkeit. Die Merkmale gliedern sich in bildbasierte (Kapitel 3.2.1), dreidimensionale (Kapitel 3.2.2) und geometrische Merkmale (Kapitel 3.2.3). Darüber hinaus präsentiert das Kapitel 3.2.4 eine Gruppe von Merkmalen zur Beschreibung von Kontext im Rahmen der Landnutzungsklassifikation.

3.2.1. Bildbasierte Merkmale

Die bildbasierten Merkmale gliedern sich in spektrale, texturelle und strukturelle Merkmale. Sie werden direkt aus den Spektralwerten eines Bildes abgeleitet.

Spektrale Merkmale Die spektralen Merkmale beschreiben die Reflexions- und Absorptionseigenschaften der zu klassifizierenden Objekte für bestimmte Wellenlängen direkt in Form der originären Spektralwerte (Grauwerte, Farbinformation) oder indirekt als Funktionen dieser Spektralwerte. Die Merkmale können individuell für jedes Pixel oder als statistische Parameter (z.B. Mittelwert, Standardabweichung) aller Pixel innerhalb eines Segments oder einer lokalen Umgebung um ein Pixel berechnet werden [Rottensteiner, 2017]. Bei der Transformation vom RGB- in den HSV-Farbraum und der Ableitung des Normalisierten Differenzierten Vegetationsindex (engl.: *Normalised Difference Vegetation Index*; NDVI) handelt es sich um Funktionen der Spektralwerte. Mit der Transformation in den HSV-Farbraum wird aus den RGB-Farbkomponenten g_R, g_G, g_B eines Bildes die Helligkeit f_{Val}, die Sättigung f_{Sat} und der Farbton f_{Hue} wie folgt abgeleitet [Burger & Burge, 2005]:

$$\text{Helligkeit (V)} \qquad f_{Val} = max(g_R, g_G, g_B) \tag{3.1}$$

$$\text{Sättigung (S)} \qquad f_{Sat} = \begin{cases} \dfrac{f_{Val} - min(g_R, g_G, g_B)}{f_{Val}}, & \text{für } f_{Val} > 0, \\ 0, & \text{sonst.} \end{cases} \tag{3.2}$$

$$\text{Farbton (H)} \qquad f_{Hue} = \begin{cases} \dfrac{60(g_G - g_B)}{f_{Val} - min(g_R, g_G, g_B)}, & \text{für } f_{Val} = g_R, \\ \dfrac{120 + 60(g_B - g_R)}{f_{Val} - min(g_R, g_G, g_B)}, & \text{für } f_{Val} = g_G, \\ \dfrac{240 + 60(g_R - g_G)}{f_{Val} - min(g_R, g_G, g_B)}, & \text{für } f_{Val} = g_B. \end{cases} \tag{3.3}$$

Für den Fall $g_R, g_G, g_B \in [0, 1]$, liegen die Werte für f_{Val} und f_{Sat} im Intervall $[0, 1]$. f_{Hue} kodiert den Farbton als Winkel auf einem Farbkreis in dem Wertebereich $[0°, 360°]$. Der NDVI berechnet sich aus den Spektralwerten im nahen Infrarot- (g_{NIR}) und Rotkanal (g_R) eines Bildes nach

$$f_{NDVI} = \frac{g_{NIR} - g_R}{g_{NIR} + g_R}. \tag{3.4}$$

Der NDVI nimmt Werte an im Intervall $[-1, 1]$, wobei Vegetationsflächen hohe Werte induzieren. Damit ist der Index besonders gut zur Trennung von Vegetations- und Nicht-Vegetationsflächen geeignet [Lillesand & Kiefer, 2000].

Texturelle Merkmale Die Texturmerkmale beschreiben regelmäßige, charakteristische Muster innerhalb eines Segments oder in einer lokalen Umgebung um ein Pixel. Als ein Standardverfahren hat sich die statistische Analyse der Textur auf Grundlage einer Co-Occurrence Matrix, bezeichnet als *Grey Level Co-Occurrence Matrix* (GLCM) [Haralick et al., 1973], etabliert. Die GLCM erfasst die

Häufigkeit des gemeinsamen Auftretens von zwei Grauwerten g_1 und g_2 an Pixeln in einer bestimmten räumlichen Konfiguration innerhalb einer Bildregion. Die räumliche Konfiguration der Grauwertpaare wird durch den Pixelabstand Δ und die Richtung α festgelegt. Für jede Kombination (Δ, α) ergibt sich eine GLCM der Dimension $N_G \times N_G$, wobei N_G der Anzahl der möglichen Grauwerte entspricht. Die normalisierte GLCM wird wie folgt aufgestellt [Tönnies, 2005]:

$$\text{GLCM} \qquad P_{\Delta,\alpha}(g_1, g_2) = \frac{1}{N_R} \sum_{(r,c)\in\mathcal{R}} \delta_K(g(r,c), g_1) \cdot \delta_K(g(r+\Delta r, c+\Delta c), g_2), \tag{3.5}$$

$$\text{mit} \qquad (\Delta r, \Delta c) = \Delta \cdot (\sin\alpha, \cos\alpha). \tag{3.6}$$

In der Gleichung 3.5 entspricht $\mathcal{R}$ der Menge aller N_R Pixel innerhalb der betrachteten Bildregion, $g(r,c)$ gibt den Grauwert eines Bildes an der Position (r,c) an und $\delta_K(\cdot,\cdot)$ repräsentiert die Kronecker-Delta-Funktion, die wie folgt definiert ist:

$$\delta_K(g_i, g_j) = \begin{cases} 1, & \text{falls } g_i = g_j, \\ 0, & \text{sonst.} \end{cases} \tag{3.7}$$

Die GLCM bildet die Grundlage für die Ableitung von aussagekräftigen Merkmalen zur Beschreibung der Textur, den sogenannten *Haralick Merkmalen* [Haralick et al., 1973]. Hiervon werden häufig die Merkmale Energie, Kontrast, Homogenität und Korrelation zur Klassifikation verwendet, die wie folgt berechnet werden können [Tönnies, 2005]:

$$\text{Energie} \qquad f_{Energ} = \sum_{g_1} \sum_{g_2} P^2_{\Delta,\alpha}(g_1, g_2) \tag{3.8}$$

$$\text{Kontrast} \qquad f_{Kontr} = \sum_{g_1} \sum_{g_2} (g_1 - g_2)^2 \cdot P_{\Delta,\alpha}(g_1, g_2) \tag{3.9}$$

$$\text{Homogenität} \qquad f_{Hom} = \sum_{g_1} \sum_{g_2} \frac{1}{1 + |g_1 - g_2|} \cdot P_{\Delta,\alpha}(g_1, g_2) \tag{3.10}$$

$$\text{Korrelation} \qquad f_{Korr} = \sum_{g_1} \sum_{g_2} \frac{(g_1 - \mu_1)(g_2 - \mu_2)}{\sigma_1 \cdot \sigma_2} \cdot P_{\Delta,\alpha}(g_1, g_2) \tag{3.11}$$

In den Gleichungen 3.8 bis 3.11 bezeichnen μ_1, μ_2 die Mittelwerte und σ_1, σ_2 die Standardabweichungen der Grauwerte g_1, g_2. Die Merkmale f_{Energ} und f_{Hom} nehmen Werte im Intervall $[0,1]$ an, f_{Kontr} liegt im Intervall $[0, N_G^2]$ und f_{Korr} im Intervall $[-1,1]$.

Strukturelle Merkmale Die strukturellen Merkmale dienen der Beschreibung der lokalen Bildstruktur. Spezielle Strukturen innerhalb eines Segments prägen sich z.B. in Form von Grauwertkanten aus. Die Analyse der Grauwertgradienten kann auf der Grundlage eines gewichteten Histogramms von Gradientenrichtungen erfolgen (engl.: *Histogram of Oriented Gradients*; HOG) [Dalal & Triggs, 2005]. Zu diesem Zweck werden zunächst die Orientierungen und Amplituden der Gradienten aus einem Intensitätsbild abgeleitet. Für die Gradientenorientierungen wird anschließend ein Histogramm pro Segment aufgestellt, wobei jeder Eintrag einer Orientierung mit der zugehörigen Amplitude gewichtet wird. Demnach ist in jeder Klasse im Histogramm die Summe der Amplituden all jener Gradienten aufge-

tragen, deren Orientierungswerte innerhalb des zur Klasse gehörigen Intervalls liegen [Hoberg et al., 2015]. Dieses Histogramm kann entweder direkt als Merkmal für die Klassifikation verwendet werden [Dalal & Triggs, 2005] oder zur Ableitung von charakteristischen Kenngrößen dienen, wie z.B. Mittelwert, Varianz, minimale und maximale Werte oder aus diesen Werten abgeleitete Verhältniszahlen [Helmholz et al., 2014; Hoberg et al., 2015].

3.2.2. Dreidimensionale Merkmale

Die dreidimensionalen Merkmale charakterisieren ein Bildprimitiv in Bezug auf dessen Höhe, die u.a. in DGM und DOM erfasst ist. Ein wichtiges Merkmal für die Klassifikation von Fernerkundungsdaten bildet die Höhe über Grund [Chehata et al., 2011], die sich aus der Differenz beider Höhenmodelle ergibt und als normalisiertes Digitales Oberflächenmodell (nDOM) bezeichnet wird [Weidner & Förstner, 1995]. Für den Fall, dass kein DGM vorhanden ist, lässt sich die Geländehöhe auch mittels geeigneter Filtermethoden [Sithole & Vosselman, 2004] aus dem DOM ableiten oder näherungsweise als die niedrigste Objekthöhe innerhalb einer lokalen Umgebung bestimmen [Mallet et al., 2008]. Dieses Merkmal ist besonders gut geeignet, um von der Geländeoberfläche hervortretende Objekte (z.B. Bäume, Gebäude) von jenen auf der Geländeoberfläche (z.B. Gras, Straße) zu unterscheiden. Analog zu den spektralen Merkmalen können diese Merkmale für jedes Pixel oder in Form von statistischen Kenngrößen für ein Segment oder eine lokale Umgebung um ein Pixel berechnet werden.

3.2.3. Geometrische Merkmale

Die geometrischen Merkmale charakterisieren die Fläche oder Form von Segmenten. Sie berechnen sich aus der polygonalen Repräsentation der Segmente. Für eine Auswahl von geometrischen Merkmalen ist nachfolgend die Berechnungsformel angegeben. Zudem werden häufig der Flächeninhalt A und der Umfang U eines Segments als Merkmal verwendet.

Kompaktheit	$f_{Komp} = \frac{4 \cdot \pi \cdot A}{U^2}$	(3.12)
Formindex	$f_{Form} = \frac{U}{4 \cdot \sqrt{A}}$	(3.13)
Fraktale Dimension	$f_{Frakt} = 2 \cdot \frac{log(U)}{log(A)}$	(3.14)
Seitenverhältnis	$f_{Seitv} = \frac{b_{mbr}}{l_{mbr}}$	(3.15)
Länglichkeit	$f_{Lang} = \frac{A}{b_{mbr}^2}$	(3.16)
Elongation	$f_{Elong} = \left\lvert log\left(\frac{l_{mbr}}{b_{mbr}}\right)\right\rvert$	(3.17)
Füllungsgrad	$f_{Konv} = \frac{A}{A_{mbr}}$	(3.18)
Polarer Abstand	$f_{Polar} = \frac{d_{min}}{d_{max}}$	(3.19)

In den Gleichungen 3.12 bis 3.19 bezeichnet A_{mbr} den Flächeninhalt des minimal umschließenden Rechtecks mit den Seitenlängen l_{mbr} und b_{mbr}. Die Parameter d_{max} bzw. d_{min} geben den maximalen bzw. minimalen Abstand aller Polygonpunkte eines Segments von dessen Schwerpunkt an. Die Kompaktheit, der Formindex [Bogaert et al., 2000] und die fraktale Dimension [Krummel et al., 1987; McGarigal & Marks, 1995] setzen den Flächeninhalt auf unterschiedliche Weise ins Verhältnis zum Umfang und quantifizieren damit insbesondere die Abweichung eines Segments von einer kompakten Form (d.h. Kreis oder Quadrat) bzw. die Komplexität der Form. Aus dem minimal umschließenden Rechteck des Segments leiten sich u.a. die Merkmale Seitenverhältnis und Elongation ab, welche die Länglichkeit des Segments charakterisieren. Das Verhältnis der Fläche des Segments zur Fläche des minimal umschließenden Rechtecks beschreibt den Füllungsgrad eines Segments, der mit der Konvexität des Segments zusammenhängt. Der polare Abstand [Bässmann & Besslich, 1991] analysiert die Anordnung der einzelnen Eckpunkte des Segments zum Schwerpunkt.

3.2.4. Kontextmerkmale

Die nachfolgende Beschreibung von Kontextmerkmalen konzentriert sich auf Merkmale zur Beschreibung der statistischen Abhängigkeit der Landnutzung von der Bodenbedeckung. Wie in Kapitel 2.1 beschrieben, werden Merkmale dieser Art in der Literatur in zwei Gruppen unterteilt: *räumliche Metriken* und *graphenbasierte Merkmale* [Hermosilla et al., 2012].

Räumliche Metriken Räumliche Metriken charakterisieren die Zusammensetzung und räumliche Anordnung der Bodenbedeckungssegmente innerhalb eines Landnutzungsobjekts. Zur Beschreibung der Zusammensetzung eignen sich insbesondere Merkmale, die den relativen Flächenanteil einer Bodenbedeckungsart l an der Gesamtfläche eines Landnutzungsobjekts berechnen gemäß

$$f_{relArea,l} = \frac{1}{\sum_{l \in \mathcal{L}^c} \sum_{(r,c) \in \mathcal{R}} \delta_K(y_{(r,c)}, l)} \sum_{(r,c) \in \mathcal{R}} \delta_K(y_{(r,c)}, l), \tag{3.20}$$

mit der Kronecker-Delta-Funktion $\delta_K(\cdot,\cdot)$, der Menge $\mathcal{L}^c$ aller Bodenbedeckungsklassen und der Menge $\mathcal{R}$ aller Pixel mit den Bildkoordinaten (r, c) in dem jeweiligen Landnutzungsobjekt. Diese Art von Merkmalen wird in vielen Ansätzen speziell zur Beschreibung des relativen Flächenanteils von Gebäuden ($l = \mathit{Gebäude}$) an der Gesamtfläche genutzt (z.B. [Hermosilla et al., 2012; Van de Voorde et al., 2009]), kann aber auch in der gleichen Weise für alle anderen Bodenbedeckungsarten $l \in \mathcal{L}^c$ berechnet werden. Zur Beschreibung der räumlichen Anordnung bieten sich beispielsweise Maße an, welche die Bodenbedeckungssegmente im Hinblick auf ihre Position zu den Nutzungsgrenzen beschreiben, z.B. durch den Mittelwert der quadratischen minimalen Distanzen der einzelnen Konturpunkte eines Bodenbedeckungssegments von den Nutzungsgrenzen.

Graphenbasierte Merkmale Die graphenbasierten Merkmale beschreiben die räumliche Nachbarschaft (Kanten) zwischen Bodenbedeckungselementen (Knoten) innerhalb eines Landnutzungsobjekts auf Basis einer Co-Occurrence Matrix, der sogenannten *Adjacency-Event-Matrix* [Barnsley & Barr, 1996]. Analog zur GLCM für Grauwerte beschreibt diese Matrix das gemeinsame Auftreten von Bo-

denbedeckungsarten an räumlich benachbarten Pixeln oder Segmenten. Im Gegensatz zur GLCM wird allerdings nur eine Matrix pro Segment aufgestellt, welche die Kombination von Bodenbedeckungsarten in einer festen Distanz von eins, d.h. nur für direkt benachbarte Bildprimitive, und gemeinsam für alle Richtungen erfasst, d.h. im Fall einer pixelbasierten Betrachtung gemeinsam für alle unabhängigen Richtungen der räumlichen 4er- oder 8er-Nachbarschaft. Die Matrix bildet die Grundlage für die Ableitung von aussagekräftigen Merkmalen. Im Gegensatz zur GLCM können aber auch direkt die einzelnen normalisierten Matrixeinträge als Merkmal verwendet werden, wie z.B. die normalisierte Anzahl von Kanten zwischen gewissen Bodenbedeckungsklassen. Dies ist vor allem deshalb möglich, weil die Dimension dieser Matrix mit $M^c \times M^c$ (M^c entspricht der Anzahl der Bodenbedeckungsklassen) i.d.R. kleiner ist als die der GLCM mit $N_G \times N_G$.

3.3. Random Forests

Random Forests (RF; „Zufallswälder“) wurden von Breiman [2001] eingeführt und haben sich seitdem als eine der erfolgreichsten Methoden des maschinellen Lernens bewährt [Hänsch & Hellwich, 2017]. Sie zeichnet sich besonders durch ihre Vielseitigkeit aus, so kann sie beispielsweise zur überwachten Klassifikation, Regression oder zum Clustering angewendet werden [Criminisi & Shotton, 2013]. Darüber hinaus ist das Verfahren nicht auf einen bestimmten Datentyp beschränkt, sondern kann sowohl diskrete als auch kontinuierliche, gelabelte oder ungelabelte Daten verarbeiten [Criminisi & Shotton, 2013]. Die nachfolgende Darstellung der theoretischen Grundlagen bezieht sich auf die Anwendung von RF zur überwachten Klassifikation, wobei die grundlegenden Konzepte in ähnlicher Weise auch für die Regression und das Clustering gelten. RF gehört zu der Gruppe der diskriminativen, nichtprobabilistischen Klassifikationsverfahren.

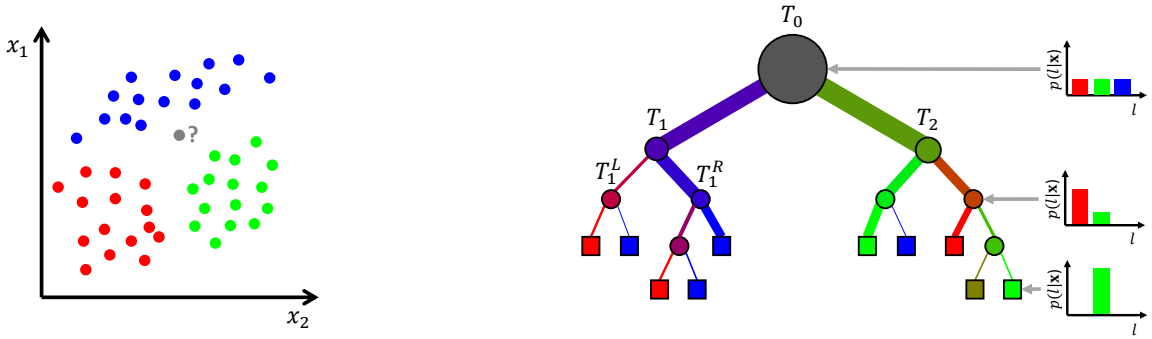
RF bestehen aus einem Ensemble (= Wald) an randomisierten, d.h. nach dem Zufallsprinzip gelernten Entscheidungsbäumen. Damit vereinen sie zwei grundlegende Konzepte des maschinellen Lernens: Entscheidungsbäume und Ensemble-Methoden. Die Entwicklung und Anwendung der Entscheidungsbäume im Bereich des maschinellen Lernens wurde maßgeblich durch die Arbeit von Breiman et al. [1984] beeinflusst, in der die sogenannten *Klassifikations- und Regressionsbäume* (engl.: *Classification and Regression Trees*; CART) eingeführt wurden. Die Kombination von mehreren schwachen Klassifikatoren zu einem starken Klassifikator, d.h. die Verwendung eines Ensembles, hat sich in den 1990er Jahren als ein wirkungsvolles Konzept herausgestellt, mit dem sich höhere Genauigkeiten und eine verbesserte Generalisierung erzielen lassen [Hänsch & Hellwich, 2017].

Den Grundbaustein von RF bilden Entscheidungsbäume, die komplexe Klassifikationsentscheidungen auf eine Reihe von binären Tests herunterbrechen. Diese Vorgehensweise findet sich auch im Alltag wieder, wobei ein binärer Test als Entscheidungsfrage interpretiert werden kann. So lassen sich viele Probleme durch die Beantwortung einer Reihe von Entscheidungsfragen lösen, wobei die nächste Frage jeweils von der zuvor gegebenen Antwort abhängt. Mit jeder Frage wird der Raum möglicher Lösungen weiter eingeschränkt. Eine solche Abfolge an Fragen lässt sich mit einem gerichteten Entscheidungsbaum abbilden [Criminisi & Shotton, 2013]. Ein Baum ist eine besondere Form eines Graphen, in dem die Knoten und Kanten hierarchisch angeordnet sind. Den Anfang eines gerichteten

Baumes markiert der Wurzelknoten, der mit mindestens zwei nachfolgenden Knoten (Kindknoten) verbunden ist. Handelt es sich bei den nachfolgenden Knoten um nicht-terminierende Knoten, weisen diese ebenfalls Verbindungen zu nachfolgenden Knoten auf. Mit jeder Verzweigung an einem Knoten wird ein neuer Teilbaum aufgespannt. Die letzte Ebene des Baumes bilden die Blattknoten, die über keine ausgehenden Kanten verfügen [Criminisi & Shotton, 2013]. Jeder nicht-terminierende Knoten repräsentiert eine Frage, deren Beantwortung darüber entscheidet an welchem der nachfolgenden Knoten die Befragung fortgesetzt wird. Dabei ist es wichtig, dass die Antwort eindeutig einer Richtung zugewiesen werden kann. Das Blatt repräsentiert die wahrscheinlichste Antwort auf die Ausgangsfragestellung, z.B. im Fall einer Klassifikation das wahrscheinlichste Klassenlabel.

Der Erfolg der Entscheidungsfindung in einem solchen Baum ist im Wesentlichen von zwei Faktoren abhängig, zum einen von den gestellten Fragen bzw. den vollzogenen Tests in jedem Knoten und zum anderen von den daraus gezogenen Schlussfolgerungen in jedem Blatt [Criminisi & Shotton, 2013]. Für einfache Klassifikationsprobleme können diese Angaben manuell festgelegt werden, da es sich in diesen Fällen um sehr einfache und leicht zu interpretierende Konstrukte handelt [Criminisi & Shotton, 2013]. Diese Vorgehensweise bietet sich insbesondere an, um Expertenwissen als a priori Information in die Klassifikation zu integrieren [Duda et al., 2001]. Allerdings hängt das Ergebnis sehr stark von den gewählten Parametern ab, was die manuelle Konstruktion von Entscheidungsbäumen für komplexe Probleme erschwert bzw. unmöglich macht [Criminisi & Shotton, 2013]. Daher bietet es sich an, die Entscheidungsbäume hinsichtlich ihrer Struktur und Parameter automatisch aus Trainingsdaten zu lernen.

Die Klassifikations- und Regressionsbäume (CART) [Breiman et al., 1984] sind randomisierte Entscheidungsbäume, die aus Trainingsdaten unter Einbeziehung des Zufalls gelernt werden. Das Grundgerüst bilden hierbei binäre Bäume, d.h. jeder nicht-terminierende Knoten hat exakt zwei Kindknoten. In der Abbildung 3.1 ist die Struktur eines solchen Baumes schematisch am Beispiel einer Klassifikationsaufgabe dargestellt. Zur Vereinfachung der Notation bezeichnet $\mathbf{x}$ in diesem Kapitel den Merkmalsvektor $\mathbf{f}_i(\mathbf{x})$ eines Bildprimitivs i. Die Klassifikation hat zum Ziel, den unbekannten, in der Abbildung 3.1a grau dargestellten Punkten im Merkmalsraum ein Klassenlabel zuzuweisen. Zu diesem Zweck wird der zugehörige Merkmalsvektor $\mathbf{x}$ den Knoten $n = \{n_0, ..., n_j, ..., n_{N_n}\}$ des Baumes in hierarchischer Reihenfolge, beginnend bei dem Wurzelknoten n_0, präsentiert, wobei N_n der Anzahl der Knoten im Pfad entspricht. Jeder Knoten n_j wendet einen binären Test auf den Merkmalsvektor an, z.B. ob ein Element des Merkmalsvektors einen Schwellwert über- oder unterschreitet. In Abhängigkeit der binären Testentscheidung wird der Merkmalsvektor an den linken oder rechten Kindknoten weitergereicht, in dem er erneut einem Test unterzogen wird. Jede Verzweigung teilt den Merkmalsraum in Teilgebiete auf. Die hierarchische Testsequenz endet, sobald der Merkmalsvektor ein Blatt erreicht. Das Klassenlabel des jeweiligen Blattknotens ist aus dem Training bekannt und wird dem Merkmalsvektor als Ausgabe zugewiesen [Hänsch & Hellwich, 2017]. Jeder Knoten n_j wendet eine unterschiedliche Testfunktion an, die im Rahmen des Trainings bestimmt wird und während der Klassifikation konstant bleibt [Criminisi & Shotton, 2013]. Die Art der Tests hängt u.a. von der Art der Merkmale (diskret, kontinuierlich) ab und ist vorab zu definieren. Die Parameter der Tests können hingegen gelernt werden [Hänsch & Hellwich, 2017]. Die Testfunktion kann einerseits nur ein Merkmal x mit einem Schwellwert θ vergleichen oder eine Teilmenge oder die Gesamtheit der Merkmale in die



(a) Merkmalsraum (b) Entscheidungsbaum

Abbildung 3.1.: Trainingsdaten und darauf angelernter Entscheidungsbaum. (a) Farbige Punkte kennzeichnen die Position der Trainingsbeispiele im 2D-Merkmalsraum, wobei die Farbe die Zugehörigkeit zu einer Klasse kodiert. Graue Punkte beschreiben zu klassifizierende Objekte, für die ein Klassenlabel zu bestimmen ist. (b) Binärer Entscheidungsbaum zur Klassifikation. Die Dicke der Kanten ist proportional zu der Anzahl und die Farbe der Kanten kennzeichnet die Zusammensetzung der Klassen aller Trainingsbeispiele, die diese Kanten durchlaufen. Während des Trainings wird eine Menge an Trainingsbeispielen $\mathcal{T}_0$ genutzt, um die Parameter des Baumes zu lernen. Die Entropie der Verteilung der Klassen an den Knoten nimmt mit zunehmender Entfernung vom Wurzelknoten ab. (Abbildung in abgeänderter Form übernommen aus [Criminisi & Shotton, 2013]).

Entscheidung einbeziehen. In diesem Fall kann der Merkmalsraum beispielsweise mit einer linearen Entscheidungsgrenze unterteilt werden, d.h. die Trennfläche bildet die Hyperebene $\mathbf{w}^T \cdot \mathbf{x} + w_0 = 0$. Diese Vorgehensweise bietet den Vorteil, dass gleichzeitig mehrere Merkmale in der Testentscheidung berücksichtigt werden können. Infolgedessen entstehen Entscheidungsgrenzen im Merkmalsraum, die nicht zwangsläufig parallel zu den Koordinatenachsen verlaufen und sich damit besser an die tatsächliche Form der Cluster im Merkmalsraum anpassen.

Der Entscheidungsbaum wird anhand von Trainingsdaten gelernt. Im Zuge dessen gilt es, die Parameter der Testfunktionen und die Struktur des Baumes zu bestimmen sowie jedem Blatt eine Klasse zuzuweisen. Während des Trainings werden die zur Verfügung stehenden Trainingsdaten rekursiv in Untermengen abnehmender Größe unterteilt. Den Ausgangspunkt bildet der Wurzelknoten, der Zugriff auf die Trainingsdaten $\mathcal{T}_0$ hat, die zum Lernen der Parameter der Testfunktion zur Verfügung stehen [Hänsch & Hellwich, 2017]. An jedem Knoten n_j wird aus mehreren zufällig vorgeschlagenen Tests jener ausgewählt, der die Trainingsdaten am besten unterteilt. Die Beurteilung der Güte der Aufteilung erfolgt anhand eines zuvor definierten Qualitätskriteriums, z.B. des Gini-Index [Breiman, 2001] oder des Informationszuwachses [Duda et al., 2001]. Die Merkmale und die Parameter der Tests werden zufällig gewählt, wobei die Anzahl N_f der zu testenden Merkmale vorzugeben ist (standardmäßig wird $N_f = \sqrt{D}$ gewählt [Breiman, 2001]). Die aktuelle Datenmenge $\mathcal{T}_j \subseteq \mathcal{T}_0$ am Knoten n_j wird anhand der gewählten Testfunktion in zwei disjunkte Teilmengen $\mathcal{T}_j^L$ und $\mathcal{T}_j^R$ aufgespalten, mit $\mathcal{T}_j^L \cup \mathcal{T}_j^R = \mathcal{T}_j$ und $\mathcal{T}_j^L \cap \mathcal{T}_j^R = \emptyset$ [Criminisi & Shotton, 2013]. Die hochgestellten Indizes L bzw. R kennzeichnen hierbei die Zuordnung der Teilmengen zu dem linken bzw. rechten Zweig. In jedem Zweig wird ein Kindknoten eingefügt, dem die jeweilige Teilmenge zugewiesen wird. Der Trainings-

prozess setzt sich rekursiv für alle neu erzeugten Knoten im Graph fort. Die Konstruktion weiterer Äste an einem Knoten endet, wenn ein Abbruchkriterium erfüllt ist. Der terminierende Knoten wird zum Blatt deklariert [Duda et al., 2001]. Sobald jeder Pfad in dem Baum mit einem Blatt endet, ist der Trainingsprozess beendet. Die Baumstruktur wird in erster Linie in Abhängigkeit des gewählten Abbruchkriteriums bestimmt, das entscheidet, wann die Rekursion, d.h. die weitere Unterteilung der Daten, endet. Für eine perfekte Unterteilung der Daten wird der Prozess solange fortgesetzt bis die Knoten „rein“ sind, d.h. nur noch Trainingsbeispiele einer Klasse enthalten [Duda et al., 2001]. Das garantiert zwar eine eindeutige Zuordnung eines Blattes zu einer Klasse, führt aber unter Umständen zu Überanpassung und sehr tiefen Bäumen. Daher ist es sinnvoll, zusätzliche Abbruchkriterien zu definieren. Die Rekursion kann beispielsweise enden, sobald eine maximale Tiefe T_{max} erreicht ist, die Anzahl an Trainingsbeispielen in einem Knoten eine vorgegebene Mindestanzahl $N_{T,min}$ unterschreitet oder ein gewähltes Qualitätskriterium erfüllt ist.

Im Anschluss an den Trainingsprozess werden die restlichen Trainingsdaten, die nicht zum Lernen der Parameter der Testfunktionen verwendet wurden, genutzt, um für jeden Blattknoten die Zugehörigkeit zu einer Klasse zu ermitteln. Die Trainingsbeispiele durchlaufen den Baum bis sie jeweils ein Blatt erreichen. Im Anschluss daran wird für jedes Blatt anhand der Trainingsbeispiele, die es erreichen, ein normiertes Histogramm $P(l)$ der Klassenlabels $l \in \mathcal{L}$ aufgestellt. Die statistische Information über die Verteilung der Klassenlabels in einem Blattknoten wird im Rahmen der Klassifikation genutzt, um den Merkmalsvektoren, die in diesem Blatt landen, ein Klassenlabel zuzuweisen. Das Histogramm lässt sich hierbei als a posteriori Wahrscheinlichkeit $P(y_i = l|\mathbf{x})$ interpretieren, wobei mit der Abhängigkeit von den Beobachtungen $\mathbf{x}$ ausgedrückt wird, dass die Klassenverteilung von dem jeweiligen Blattknoten abhängt, den der unbekannte, zu klassifizierende Merkmalsvektor $\mathbf{x}$ erreicht [Criminisi & Shotton, 2013]. Das Ergebnis der Klassifikation bildet in den meisten Fällen das Klassenlabel, für das die a posteriori Wahrscheinlichkeit maximal ist. Für die Weiterverarbeitung in einem probabilistischen Kontext bietet es sich jedoch an, nicht ein einzelnes Klassenlabel, sondern die Gesamtverteilung auszugeben, um die Unsicherheiten der Entscheidung berücksichtigen zu können.

Der CART-Algorithmus zeichnet sich insbesondere durch seine Flexibilität aus und erzielt selbst für fehlerhafte und unvollständige Daten ein gutes Ergebnis [Hänsch & Hellwich, 2017]. Durch die Aneinanderreihung von typischerweise sehr einfachen Abfragen ist das Verfahren sehr schnell beim Training und beim Klassifizieren [Duda et al., 2001]. Die Instabilität der Entscheidungsbäume resultiert aus der zufälligen Wahl der Entscheidungsgrenzen und äußert sich insbesondere dadurch, dass kleine Änderungen der Trainingsdaten große Änderungen der Entscheidungsgrenzen bewirken können [Duda et al., 2001]. Zudem neigt das Verfahren zur Überanpassung an die Trainingsdaten. Dieser Nachteil, der sich insbesondere in Form von tiefen Bäumen und einer hohen Anzahl an Blättern ausprägt, kann beispielsweise durch das Zurückschneiden („Pruning“) der Bäume behoben werden [Duda et al., 2001]. Eine weitere Möglichkeit, die Generalisierung bei der Klassifikation mit Entscheidungsbäumen zu verbessern, bildet das Ensemble-Lernen. Die Grundidee besteht hierbei darin, verschiedene Instanzen eines Basisklassifikators zu einem starken Klassifikator zu kombinieren [Hänsch & Hellwich, 2017].

Ein Ensemble von B randomisierten Entscheidungsbäumen bildet den *Zufallswald* [Breiman, 2001]. Jeder Baum $b = \{1, ..., B\}$ wird unabhängig von den anderen anhand von Trainingsdaten gelernt. Das Training basiert dabei zu einem gewissen Grad auf Zufallsentscheidungen, indem zum

einen die Trainingsbeispiele zum Anlernen der Bäume zufällig gewählt werden [Breiman, 2001] und zum anderen die Bestimmung optimaler Testfunktionen auf einer zufälligen Auswahl an Merkmalen basiert. Die zufällige Auswahl von Trainingsbeispielen führt dazu, dass jeder Baum b auf Grundlage einer individuellen Untermenge $\mathcal{T}_b$ der Trainingsdaten angelernt wird. Der Trainingsdatensatz $\mathcal{T}_b$ pro Baum wird nach dem *Bootstrap*-Prinzip erzeugt. Hierfür werden für jeden Baum b $N_{T_b} < N_T$ Trainingsbeispiele unabhängig und mit Zurücklegen aus der Gesamtheit der Trainingsdaten $\mathcal{T}_0$ gezogen und zu der Menge $\mathcal{T}_b$ hinzugefügt ($\mathcal{T}_b \subset \mathcal{T}_0$) [Hänsch & Hellwich, 2017]. Die auf diese Weise gezogenen Bootstrap-Datensätze gelten als unabhängig. *Bagging* (*Bootstrap AGGregatING*) bezeichnet die Nutzung von Bootstrap-Datensätzen zur Klassifikation [Breiman, 1996]. Jeder dieser Datensätze dient zum Anlernen eines Klassifikators, d.h. im Fall von RF dem Trainieren eines Entscheidungsbaumes. Da sich die Struktur eines Entscheidungsbaumes in erster Linie aus den jeweils zur Verfügung stehenden Trainingsdaten sowie ihrer Unterteilung in den Knoten bestimmt, führt die zufällige Variation der Trainingsdaten zu verschiedenartigen Instanzen und damit zur Dekorrelation der Entscheidungsbäume [Hänsch & Hellwich, 2017]. Durch die Verwendung von vielen unabhängigen Untermengen der Trainingsdaten wird Überanpassung vermieden [Breiman, 2001], was zu verbesserten Generalisierungseigenschaften von RF beiträgt. Im Gegenzug erhöht sich allerdings der Trainingsfehler, weil insgesamt weniger Trainingsbeispiele zum Lernen der einzelnen Bäume zur Verfügung stehen [Hänsch & Hellwich, 2017].

Die Vorgehensweise bei der Klassifikation unter Verwendung aller Bäume eines RF ist in der Abbildung 3.2 schematisch dargestellt. Im Rahmen der Klassifikation gilt es, die Klasse eines unbe-

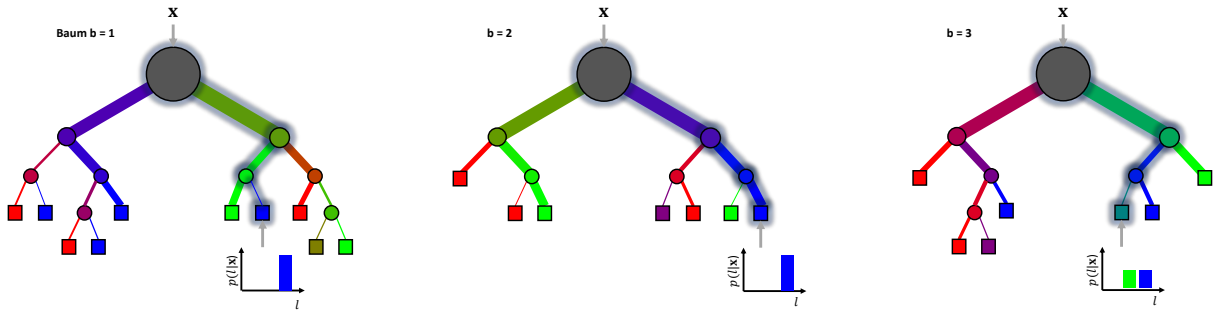


Abbildung 3.2.: Klassifikation unter Verwendung aller Bäume eines RF. Der Merkmalsvektor $\mathbf{x}$ eines zu klassifizierenden Primitivs durchläuft jeden Baum. In jedem Knoten wird ein Test auf den Merkmalsvektor angewendet, der über die Weiterleitung in den rechten oder linken Zweig entscheidet. Der Prozess endet, sobald der Merkmalsvektor ein Blatt erreicht. Die in dem jeweiligen Blatt gespeicherte a posteriori Wahrscheinlichkeit $P_b(l|\mathbf{x})$ aller Bäume wird zur finalen Klassifikationsentscheidung kombiniert. (Abbildung in abgeänderter Form übernommen aus [Criminisi & Shotton, 2013]).

kannten Samples zu bestimmen. Zu diesem Zweck durchläuft der zugehörige Merkmalsvektor $\mathbf{x}$ alle B Bäume des Waldes bis es in jedem Baum einen Blattknoten erreicht [Hänsch & Hellwich, 2017]. Die in den B Blättern gespeicherten a posteriori Wahrscheinlichkeiten $P_b(y_i = l|\mathbf{x})$ werden zur finalen Klassifikationsentscheidung kombiniert. Dieser Vorgehensweise liegt die Annahme zugrunde, dass die kombinierte Entscheidung im Durchschnitt genauer und robuster ist als die Einzelentscheidungen

der Bäume [Hänsch & Hellwich, 2017]. Die Ergebnisse der einzelnen Bäume können dabei auf unterschiedliche Weise zu einer Gesamtentscheidung kombiniert werden. Eine Möglichkeit besteht darin, dass jeder Baum b zunächst individuell eine Entscheidung für das wahrscheinlichste Klassenlabel l^b auf Grundlage der im Blatt gespeicherten Klassenverteilung trifft. Diese Einzelentscheidungen fließen als Votum für die wahrscheinlichste Klasse in die Gesamtentscheidung ein. Die Stimmen N_l werden pro Klasse $l \in \mathcal{L}$ gezählt [Hänsch & Hellwich, 2017]

$$N_l = \sum_{b=1}^{B} \delta_K(l^b, l), \tag{3.21}$$

wobei $\delta_K(\cdot,\cdot)$ wieder die Kronecker-Delta-Funktion beschreibt. Die finale Entscheidung fällt nach dem Mehrheitsprinzip auf die Klasse mit den meisten Stimmen [Breiman, 2001]. Darüber hinaus lässt sich aus den Einzelstimmen in einfacher Weise ein Konfidenzmaß der getroffenen Klassifikationsentscheidung ableiten. So ergibt sich die Gesamtwahrscheinlichkeit pro Klasse l aus der Summe aller Stimmen N_l für diese Klasse dividiert durch die Anzahl aller Bäume [Rottensteiner, 2017]:

$$P(y_i = l|\mathbf{x}) = \frac{N_l}{B}. \tag{3.22}$$

Die Abstimmung nach dem Mehrheitsprinzip hat allerdings den Nachteil, dass die Unsicherheit der Schätzung eines Klassenlabels innerhalb eines Baumes nicht in die endgültige Klassifikationsentscheidung einfließt [Hänsch & Hellwich, 2017]. Zur Berücksichtigung dieser Unsicherheit bietet es sich daher an, die in den Blättern gespeicherten relativen Klassenhäufigkeiten zu mitteln:

$$P(y_i = l|\mathbf{x}) = \frac{1}{B} \sum_{b=1}^{B} P_b(y_i = l|\mathbf{x}), \tag{3.23}$$

was direkt zu einer a posteriori Wahrscheinlichkeit führt, aus dem das endgültige Klassenlabel nach dem Maximum-A-Posteriori (MAP) Kriterium abgeleitet wird.

Durch die Kombination mehrerer Bäume wird eine bessere Generalisierung und höhere Stabilität erzielt. Das Verfahren ist relativ robust gegenüber Ausreißern sowie ungenauen Trainingsdaten und liefert genaue Ergebnisse [Breiman, 2001]. Das Trainieren und Klassifizieren der einzelnen Bäume lässt sich gut parallelisieren [Breiman, 2001], wodurch das Verfahren äußerst recheneffizient ist. Der Ansatz ist für Mehrklassenprobleme geeignet und in der Lage, hochdimensionale Merkmalsvektoren zu verarbeiten. Die Ausgabe lässt sich probabilistisch interpretieren, wodurch die Weiterverarbeitung der Ergebnisse in einem probabilistischen Kontext ermöglicht wird [Hänsch & Hellwich, 2017].

Vor der Anwendung des Klassifikators gilt es, eine bestimmte Anzahl an Parametern zu definieren (vgl. Breiman [2001]), u.a. die Gesamtanzahl B der Entscheidungsbäume und die Anzahl N_f der zu testenden Merkmale. Weitere Parameter dienen der Festlegung des Abbruchkriteriums, wie z.B. die maximale Tiefe T_{max} der Bäume sowie die Mindestanzahl $N_{T,min}$ der Samples für nicht-terminierende Knoten. Die Parameter sollten möglichst individuell auf die jeweilige Klassifikationsaufgabe abgestimmt werden.

Der RF Klassifikator bietet zudem als eine Art Nebenprodukt die Funktionalität zur Berech-

nung eines Wichtigkeitsmaßes pro Merkmal [Breiman, 2001], das dessen Einfluss auf die Klassifikation bewertet. Dieses Relevanzmaß ist gut zur automatischen Selektion von Merkmalen für die Klassifikation geeignet. Die Berechnung erfolgt im Rahmen des Trainings unter Verwendung der sogenannten *Out-of-Bag* (OOB)-Samples, die typischerweise etwa ein Drittel des gesamten Trainingsdatensatzes ausmachen [Breiman, 2001], jedoch nicht zum Trainieren verwendet werden. Um die Wichtigkeit eines Merkmals zu berechnen, werden die zugehörigen Merkmalswerte in dem OOB-Datensatz jedes Baumes zufällig vertauscht. Die übrigen Merkmale bleiben unverändert. Die Idee ist hierbei, dass das zufällige Vertauschen der Merkmalswerte mit dem Fehlen des jeweiligen Merkmals gleichzusetzen ist [Chehata et al., 2011]. Die Wichtigkeit eines Merkmals berechnet sich aus der Abweichung der Klassifikationsgenauigkeit der OOB-Samples vor und nach dem Vertauschen der Merkmalswerte, gemittelt über alle Bäume [Breiman, 2001].

3.4. Conditional Random Fields

Conditional Random Fields (CRF) sind ungerichtete, graphische Modelle zur kontextbasierten Klassifikation, die aus Knoten n und Kanten e bestehen. Die Knoten repräsentieren die Klassenlabels der zu klassifizierenden Bildprimitive, z.B. Pixel oder Segmente. In einem paarweisen CRF verbinden Kanten adjazente Knoten und modellieren die statistischen Abhängigkeiten zwischen den Klassenlabels an benachbarten Bildprimitiven und den Daten. Im Rahmen der Klassifikation gilt es, allen Knoten eines graphischen Modells simultan die wahrscheinlichsten Klassenlabels $\mathbf{y}$ aus einer Menge möglicher Klassenlabels $\mathcal{L}$ unter Berücksichtigung der Beobachtungen $\mathbf{x}$ zuzuweisen. CRF gehören zu den diskriminativen Klassifikationsverfahren, d.h. sie modellieren direkt die a posteriori Wahrscheinlichkeit $P(\mathbf{y}|\mathbf{x})$ des Labelvektors $\mathbf{y}$ bei gegebenen Daten $\mathbf{x}$ [Kumar & Hebert, 2006]:

$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \left(\prod_{i \in \mathcal{S}} \phi(y_i, \mathbf{x}) \cdot \prod_{i \in \mathcal{S}} \prod_{j \in \mathcal{N}_i} \psi(y_i, y_j, \mathbf{x})^{\omega} \right). \tag{3.24}$$

In der Gleichung 3.24 wird $\phi(y_i, \mathbf{x})$ als *Assoziationspotential* oder *unäres Potential* und $\psi(y_i, y_j, \mathbf{x})$ als *Interaktionspotential* oder *binäres Potential* bezeichnet. Das Assoziationspotential $\phi(y_i, \mathbf{x})$ modelliert die Relation zwischen dem Klassenlabel y_i an dem Bildprimitiv i und den Beobachtungen $\mathbf{x}$. Jedem Knoten ist ein Merkmalsvektor $\mathbf{f}_i(\mathbf{x})$ zugeordnet, der prinzipiell von allen verfügbaren Beobachtungen $\mathbf{x}$ abhängen kann [Kumar & Hebert, 2006]. Das Assoziationspotential lässt sich beispielsweise proportional zur Wahrscheinlichkeit von y_i bei gegebenen Beobachtungen $\mathbf{x}$ modellieren, d.h. $\phi(y_i, \mathbf{x}) \propto P(y_i|\mathbf{f}_i(\mathbf{x}))$. Das Interaktionspotential $\psi(y_i, y_j, \mathbf{x})$ modelliert die Abhängigkeiten zwischen den Klassenlabels y_i und y_j adjazenter Knoten unter Berücksichtigung der Beobachtungen $\mathbf{x}$. Die Beobachtungen werden hierbei in Form eines Merkmalsvektors $\boldsymbol{\mu}_{ij}(\mathbf{x})$ pro Kante berücksichtigt, der sich beispielsweise aus den aneinandergehängten Merkmalsvektoren $\mathbf{f}_i(\mathbf{x})$ und $\mathbf{f}_j(\mathbf{x})$ adjazenter Knoten n_i und n_j, deren elementweisen Differenzen oder der zwischen ihnen liegenden euklidischen Distanz zusammensetzen kann. Die Partitionsfunktion $Z(\mathbf{x})$ stellt eine Normalisierungskonstante dar, mit der die Potentiale in Wahrscheinlichkeiten transformiert werden. $\mathcal{N}_i$ definiert die Nachbarschaft des Bildprimitivs i. Der Parameter ω gewichtet das Interaktionspotential relativ zum Assoziationspotential

und bestimmt somit den Einfluss des Interaktionspotentials im Rahmen der Klassifikation.

CRF geben nur einen allgemeinen Rahmen zur kontextbasierten Klassifikation vor, in den verschiedene funktionale Modelle für die Potentiale eingebunden werden können [Kumar & Hebert, 2006]. Für beide Potentiale können demnach beliebige diskriminative Klassifikatoren verwendet werden, sofern sie eine Wahrscheinlichkeit oder ein der Wahrscheinlichkeit ähnliches Maß als Ausgabe liefern. Für das Assoziationspotential verwenden Kumar & Hebert [2006] beispielsweise verallgemeinerte lineare Modelle. Darüber hinaus haben sich auch andere Klassifikatoren bewährt, wie z.B. der RF Klassifikator [Schindler, 2012].

In den meisten Fällen werden für das Interaktionspotential einfachere Modelle verwendet, die identische Klassenlabels an benachbarten Knoten bevorzugen, indem sie Klassenübergänge mit höheren Kosten bestrafen. Diese Modelle bewirken eine Glättung der Ergebnisse (siehe [Schindler, 2012] für einen Vergleich der Modelle). Das sogenannte Ising- (im binären Fall) bzw. Potts-Modell (im Mehrklassenfall) [Besag, 1986; Geman & Geman, 1984] stellt hierbei das einfachste Modell dar, bei dem der Glättungseffekt ausschließlich von den Klassenlabels an benachbarten Knoten abhängt [Kumar & Hebert, 2006]. Bei allgemeineren Modellen werden zusätzlich die Beobachtungen berücksichtigt. Damit wird eine datenabhängige Glättung erzielt. In diesem Fall kann das Interaktionspotential als die Wahrscheinlichkeit, dass beide Knoten das gleiche Klassenlabel annehmen gegeben die Interaktionsmerkmale $\boldsymbol{\mu}_{ij}(\mathbf{x})$, d.h. $\psi(y_i, y_j, \mathbf{x}) \propto P(y_i = y_j | \boldsymbol{\mu}_{ij}(\mathbf{x}))$, interpretiert werden. Zu dieser Gruppe zählt beispielsweise das kontrastsensitive Potts-Modell [Boykov & Jolly, 2001], wobei es sich um eine Erweiterung des Potts-Modells handelt. Bei diesem Modell wird gemäß der Gleichung 3.25 der Grad der Glättung durch das Interaktionsmerkmal beeinflusst, wobei es sich um die euklidische Distanz d_{ij} zwischen den Beobachtungen $\mathbf{x}_i$ und $\mathbf{x}_j$ an zwei adjazenten Knoten n_i und n_j handelt. Das kontrastsensitive Potts-Modell ist wie folgt definiert [Boykov & Jolly, 2001]:

$$\psi(y_i, y_j, \mathbf{x}) = \begin{cases} \exp\left(\beta + (1-\beta) \cdot e^{-\frac{d_{ij}^2}{2\sigma^2}}\right), & \text{falls } y_i = y_j, \\ 1, & \text{sonst.} \end{cases} \tag{3.25}$$

In der Gleichung 3.25 bestimmt der Parameter $\beta \in [0, 1]$ das relative Gewicht zwischen dem datenunabhängigen und dem datenabhängigen Glättungsterm und regelt somit den Einfluss der Daten auf die Glättung. Nimmt β den Wert Eins an, entspricht das Modell dem Potts-Modell [Schindler, 2012] und die Glättung der Klassifikationsergebnisse erfolgt unabhängig vom Bildinhalt. Nimmt β hingegen den Wert Null an, wird der Grad der Glättung ausschließlich durch den datenabhängigen Term bestimmt. Der Parameter σ^2 beschreibt den Mittelwert der quadrierten Distanzen d_{ij}^2, dessen Wert während des Trainings gelernt wird. Das kontrastsensitive Potts-Modell hat sich in diversen Anwendungen als sehr effektiv herausgestellt, da es zufriedenstellende Ergebnisse [Schindler, 2012] in angemessener Rechenzeit erzielt. Bei bestimmten Anwendungen ist hingegen ein reiner Glättungseffekt nicht gewünscht. In diesen Fällen bietet es sich an, ein komplexeres Modell für das Interaktionspotential zu wählen. Dieses Modell sollte wahrscheinlichere Klassenrelationen unter Berücksichtigung der Daten favorisieren, anstatt generell ungleiche Klassenrelationen zu bestrafen. Der Zusammenhang zwischen den Beobachtungen und den Klassenlabels an benachbarten Bildprimitiven lässt sich anhand

von repräsentativen Trainingsdaten lernen. Das Interaktionspotential wird hierbei modelliert als die gemeinsame Wahrscheinlichkeit beider Klassenlabels y_i und y_j gegeben die Interaktionsmerkmale $\boldsymbol{\mu}_{ij}$, d.h. $\psi(y_i, y_j, \mathbf{x}) \propto P(y_i, y_j | \boldsymbol{\mu}_{ij}(\mathbf{x}))$. Diese Gruppe formuliert die Bestimmung des Interaktionspotentials als eine normale Klassifikationsaufgabe, die Klassenrelationen anstatt einzelner Klassen lernt. Demzufolge bildet jedes Paar von Klassenlabels eine eigene Klasse, was in Abhängigkeit der Anzahl an Klassen zu einer hohen Zahl an Zuständen führen kann. Analog zu dem Assoziationspotential kann hierbei ebenfalls jeder beliebige diskriminative Klassifikator mit einer probabilistischen Ausgabe verwendet werden [Kumar & Hebert, 2006], wie z.B. Entscheidungsbäume [Nowozin et al., 2011] oder verallgemeinerte lineare Modelle [Niemeyer et al., 2012; Zhong & Wang, 2011].

Bei CRF handelt es sich um ein überwachtes Klassifikationsverfahren, d.h. die Parameter des Modells werden anhand von repräsentativen Trainingsdaten gelernt. Hierbei handelt es sich zum einen um die Parameter der einzelnen Potentiale und zum anderen um das Gewicht ω. Das Training lässt sich im Allgemeinen nur approximativ realisieren [Vishwanathan et al., 2006], da hierfür keine effizienten Methoden zur Verfügung stehen, mit denen sich eine exakte Lösung bestimmen lässt. Um das Training zu vereinfachen, können die Klassifikatoren der einzelnen Potentiale separat gelernt werden [Shotton et al., 2009]. Der Gewichtsparameter ω kann im Anschluss an das Trainieren der Potentiale z.B. durch Kreuzvalidierung bestimmt werden [Shotton et al., 2009].

Im Rahmen der Inferenz wird die wahrscheinlichste Konfiguration von Klassenlabels $\mathbf{y}$ simultan für alle Knoten des CRF bestimmt. Hierzu gilt es, die a posteriori Wahrscheinlichkeit $P(\mathbf{y}|\mathbf{x})$ der Klassenlabels bei gegebenen Beobachtungen zu maximieren, d.h. $\mathbf{y} = \arg\max P(\mathbf{y}|\mathbf{x})$. Im Allgemeinen ist eine exakte Inferenz für Mehrklassenprobleme rechnerisch nicht lösbar [Kumar & Hebert, 2006]. Daher werden zur Inferenz in CRF – von wenigen Ausnahmen abgesehen – ausschließlich approximative Methoden verwendet [Szeliski et al., 2008]. Die Ausnahme hiervon bildet das Verfahren Graph cuts [Boykov et al., 2001], das für den Zweiklassenfall eine exakte Lösung garantiert. Hierbei wird das graphische Modell um zwei Knoten erweitert, bezeichnet als *Quelle* und *Senke*, die jeweils eine Klasse repräsentieren und mit allen Bildprimitiven des Graphen verbunden sind. Für die Inferenz werden die Potentiale in Kosten umgewandelt, d.h. invertiert, wobei die unären Potentiale der Bildprimitive jeweils den abgehenden Kanten zu der Quelle und Senke zugeordnet sind. Im Rahmen der Inferenz werden Kanten entfernt, sodass zwei disjunkte Teilgraphen entstehen, wobei jeder Teilgraph einer Klasse entspricht. Es werden jene Kanten entfernt, für die die Summe der Kosten minimal ist. Dieses Verfahren lässt sich allerdings nur für paarweise Interaktionspotentiale anwenden, welche die Submodularitätsbedingung erfüllen:

$$\psi(-1,+1) + \psi(+1,-1) \leq \psi(-1,-1) + \psi(+1,+1), \tag{3.26}$$

wie es für das Ising-, das Potts- sowie das kontrastsensitive Potts-Modell der Fall ist [Kolmogorov & Zabin, 2004; Park & Gould, 2012]. Für eine allgemeine Formulierung eines Klassifikators kann die Einhaltung der Submodularitätsbedingung jedoch nicht garantiert werden [Rottensteiner, 2017]. Graph cuts lassen sich mittels bestimmter Algorithmen (α-β-*swaps*, α-*expansion*) auch für den Mehrklassenfall anwenden [Boykov et al., 2001]. Mit dem Übergang auf den Mehrklassenfall ist eine exakte Lösung jedoch nicht mehr garantiert. Graph cuts gelten als leistungsfähiges Inferenzverfahren, des-

sen Anwendbarkeit allerdings durch die Einhaltung der Submodularitätsbedingung eingeschränkt ist [Rottensteiner, 2017]. Ein weiteres Standardverfahren zur Inferenz in paarweisen CRF bildet Loopy Belief Propagation (LBP) [Frey & MacKay, 1998], das gegenüber Graph cuts den Vorteil bietet, dass es nicht auf submodulare Potentiale beschränkt ist. Dieses Verfahren approximiert eine Lösung für die optimale Labelkonfiguration mit Hilfe eines iterativen Optimierungsansatzes, der eine Erweiterung von *Message-Passing*-Algorithmen auf Graphen mit Zyklen darstellt. Im Rahmen von LBP kommunizieren die Knoten über Nachrichten. Im ersten Schritt senden die Knoten eine Nachricht an alle ausgehenden Kanten. Im zweiten Schritt werden die Nachrichten von den eintreffenden Kanten in den Knoten zusammengeführt und zur Bestimmung eines Konfidenzwertes, dem sogenannten *Belief*, herangezogen [Rottensteiner, 2017]. Dieses Vorgehen wird iterativ fortgesetzt. In jeder Iteration wird der Konfidenzwert anhand neuer Nachrichten aktualisiert. In der *Sum-Product*-Variante beschreibt der Konfidenzwert die Randverteilung der Klassenlabels an dem jeweiligen Bildprimitiv [Frey & MacKay, 1998], wohingegen der Konfidenzwert in der *Max-Product*-Variante die a posteriori Wahrscheinlichkeit approximiert [Kolmogorov, 2006]. LBP garantiert keine optimale Lösung, wenngleich das Verfahren in den meisten Fällen eine gute Näherungslösung liefert [Szeliski et al., 2008].

4. Methodik

4.1. Überblick

In der Abbildung 4.1 ist der im Rahmen dieser Arbeit entwickelte Ansatz zur Verifikation und Aktualisierung eines räumlichen Landnutzungsdatenbestandes schematisch dargestellt. Der Ansatz besteht im Wesentlichen aus drei Komponenten: den Eingangsdaten, der Methodik zur gemeinsamen Klassifizierung der Bodenbedeckung und der Landnutzung über den Austausch von Kontextinformation sowie der anschließenden Verifikation und Aktualisierung des Landnutzungsdatenbestandes. Der Klassifikations-

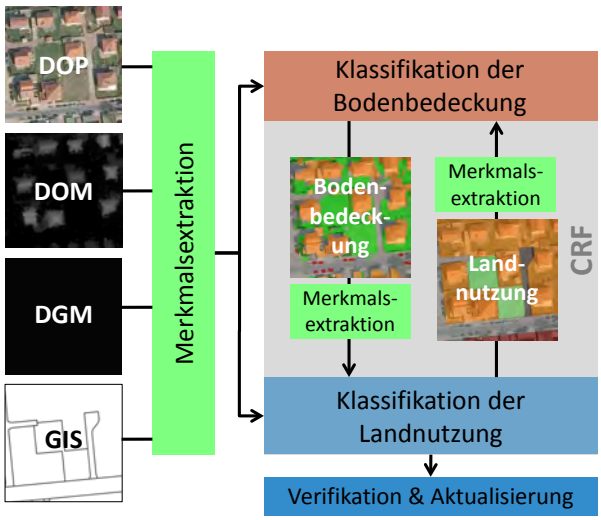


Abbildung 4.1.: Schematische Darstellung des Ansatzes zur simultanen Klassifizierung der Bodenbedeckung und Landnutzung. DOP: Digitales Orthophoto.

ansatz wurde für Eingangsdaten entwickelt, die sich aus hochauflösenden, multispektralen Bilddaten sowie Höheninformation ableiten (z.B. DOP, DOM und DGM in Abb. 4.1). Für die geometrische Abgrenzung der Landnutzungsobjekte werden die GIS-Objekte eines räumlichen Landnutzungsdatenbestandes benötigt. Darüber hinaus handelt es sich um ein überwachtes Verfahren, das Trainingsdaten für die Bodenbedeckung und Landnutzung erfordert. Die Trainingsdaten für die Landnutzung umfassen GIS-Objekte mit bekannten Nutzungsarten. Hierbei ist sicherzustellen, dass die Klassenlabels korrekt sind, was eine manuelle Überprüfung durch einen versierten Bearbeiter erfordert. Die Trainingsdaten für die Bodenbedeckung bestehen aus vollständig ausgelabelten Bildausschnitten, die typischerweise manuell zu erfassen sind. Die Gesamtheit der Eingangsdaten muss in einem einheitlichen Koordinatensystem referenziert sein. Die geometrische Auflösung der Eingangsdaten im Rasterformat kann prinzipiell variieren, jedoch sollte mindestens ein Eingangsdatensatz Informationen in der finalen geometri-

schen Auflösung der Ergebnisse für die Klassifikation bereitstellen. Die Mindestauflösung ist abhängig von der Aufgabenstellung bzw. dem gewünschten Detailgrad der Ergebnisse, jedoch sollte eine geometrische Auflösung von 1 bis 2 m nicht unterschritten werden, um noch Einzelobjekte, wie z.B. Gebäude, in den Daten auflösen zu können. Die zentrale Komponente der Methodik bildet ein CRF-Modell, das theoretisch durch geringfügige Änderungen flexibel an diverse Sensordaten angepasst werden kann. Zu diesem Zweck bedarf es lediglich der Extraktion anderer Merkmale. Gleichwohl empfiehlt sich die Verwendung von Luftbildern mit einem Nahen Infrarotkanal sowie von Höheninformation, da sie insbesondere nützliche Informationen zur Unterscheidung verschiedener Bodenbedeckungsarten beitragen [Chehata et al., 2011; Rottensteiner et al., 2007]. Das CRF-Modell verknüpft die Klassifikation der Bodenbedeckung und der Landnutzung in einem gemeinsamen graphischen Modell. Das CRF modelliert einerseits die Relationen der Bildprimitive zu den Eingangsdaten und andererseits die kontextuellen Abhängigkeiten zwischen den Bildprimitiven in Form von Potentialen. Im Rahmen der Inferenz interagieren die Klassifikationsaufgaben indem sie untereinander Information in Form von Klassifikationsergebnissen austauschen (vgl. Abb. 4.1). Infolge der Approximation des Inferenzprozesses werden die Ergebnisse jedoch nicht direkt übertragen, sondern dienen der Ableitung von aussagekräftigen Kontextmerkmalen, die in einem iterativen Prozess der jeweils anderen Klassifikationsaufgabe kommuniziert werden. Die finale Komponente der Methodik bildet die Verifikation und Aktualisierung eines Landnutzungsdatenbestandes. Dies erfolgt auf der Grundlage der klassifizierten Nutzungsarten. Durch den Abgleich mit dem Landnutzungsdatenbestand werden Änderungshinweise abgeleitet. Für die Aktualisierung des Datenbestandes prädiziert der Klassifikationsansatz eine aktuelle Nutzungsart. Die Verifikation und Aktualisierung des Datenbestandes liegen nicht im Fokus dieser Arbeit und sind daher nur rudimentär realisiert. Die Effektivität der Verifikation und Aktualisierung wird im Rahmen dieser Arbeit deshalb auch nicht näher untersucht.

Im folgenden Abschnitt wird die Methodik des Verfahrens detailliert beschrieben, wobei der Fokus auf der zentralen Komponente des Verfahrens, d.h. der simultanen Klassifizierung der Bodenbedeckung und Landnutzung, liegt. Die Darstellung der Methodik konzentriert sich auf die drei wesentlichen wissenschaftlichen Beiträge dieser Arbeit. Den ersten Beitrag bildet die Struktur des graphischen Modells inklusive der Modellierung der statistischen Abhängigkeiten in Form von Potentialen, die in den Kapiteln 4.2.1 und 4.2.2 ausführlich beschrieben werden. Das Kapitel 4.2.3 geht anschließend auf den zweiten Beitrag dieser Arbeit ein, und zwar die Struktur der entwickelten Inferenzprozedur. In engem Zusammenhang hierzu steht das Training, das in Kapitel 4.2.4 thematisiert wird. Den dritten Beitrag bilden die im Rahmen dieser Arbeit entwickelten Kontextmerkmale, die das Kapitel 4.3 vorstellt. Die Darstellung der Methodik wird komplettiert durch die Beschreibung der verfolgten Verifikations- und Aktualisierungsstrategie in Kapitel 4.4. Das Kapitel 4.5 befasst sich abschließend mit einigen Aspekten, die mit der konkreten Implementierung der Methodik in Zusammenhang stehen.

4.2. Zwei-Ebenen-CRF-Modell

4.2.1. Graphisches Modell

Die Klassifikation der Bodenbedeckung und der Landnutzung werden in einem graphischen Modell zusammengeführt, das zwei Ebenen umfasst: eine Bodenbedeckungs- und eine Landnutzungsebene. Auf diese Weise wird eine simultane Prozessierung ermöglicht, im Verlauf derer sich beide Klassifikationsaufgabe

Ebenen des graphischen Modells entsprechen zwei Hierarchieebenen, die übereinander angeordnet sind, sodass sich die zugehörigen Bildprimitive räumlich überlagern. Jede Ebene besteht aus Knoten n und *räumlichen Kanten* e_R. Die beiden Ebenen sind durch *semantische Kanten* e_S miteinander verbunden. Die Struktur des graphischen Modells ist in Abbildung 4.2 schematisch dargestellt.

Methodik: Zwei-Ebenen-Modell

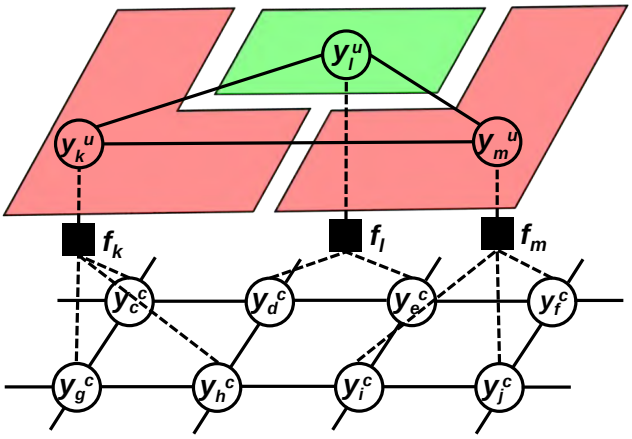
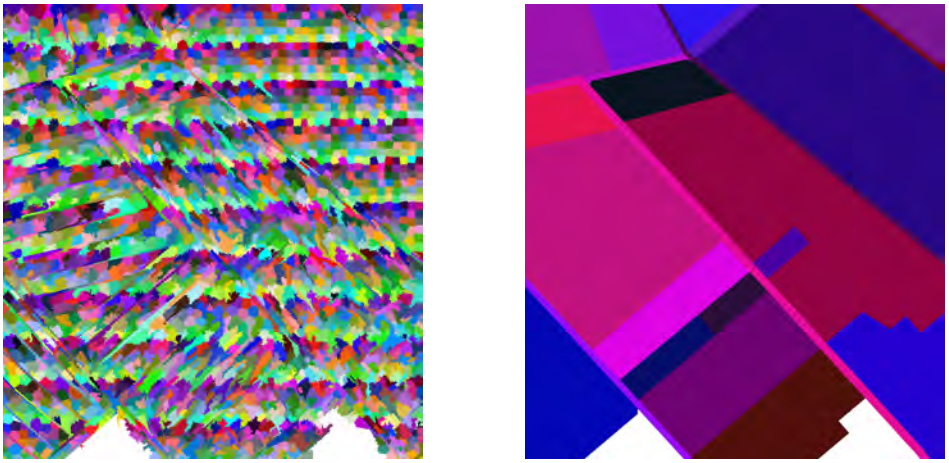


Abbildung 4.2: Graphisches Modell bestehend aus zwei Ebenen: Bodenbedeckungs- (Index c) and Landnutzungsebene (Index u). Knoten sind als Kreise und räumliche Kanten als durchgezogene Linien dargestellt. Semantische Kanten verbinden alle räumlich überlagernden Bildprimitive in beiden Ebenen, was zu Cliquen höherer Ordnung führt, repräsentiert durch die Faktorknoten f_k, f_l und f_m (schwarze Quadrate).

Die Ebenen des graphischen Modells unterscheiden sich hinsichtlich der durch die Knoten repräsentierten Bildprimitive und der verwendeten Merkmale. Darüber hinaus sind die Klassenstrukturen in beiden Ebenen verschieden, d.h. $y_i^c \in \mathcal{L}^c = [l_1^c, ..., l_{M^c}^c]$ in der Bodenbedeckungs- und $y_k^u \in \mathcal{L}^u = [l_1^u, ..., l_{M^u}^u]$ in der Landnutzungsebene, wobei M^c und M^u die Anzahl der Klassen in der jeweiligen Ebene angeben. Die Knoten der Bodenbedeckungsebene entsprechen Superpixeln, die vorweg mit Hilfe eines Segmentierungsverfahrens aus den Eingangsdaten abgeleitet werden [Achanta et al., 2012]. Demgegenüber repräsentieren die Knoten in der Landnutzungsebene die GIS-Objekte eines räumlichen Landnutzungsdatenbestandes. Die Geometrie der Bildprimitive wird als fest angenommen und im Rahmen der Klassifikation nicht verändert. Bezüglich der Landnutzungsobjekte liegt dem Verfahren somit die Annahme zugrunde, dass die Begrenzungen der GIS-Objekte korrekt sind. Diese Annahme ist dadurch begründet, dass in Gebieten wie z.B. Mitteleuropa, wo auch das Untersuchungsgebiet dieses Forschungsprojektes lokalisiert ist, die Nutzungsartengrenzen häufig einen Bezug zu den Eigentumsgrenzen haben. Die Eigentumsgrenzen werden von den Vermessungsbehörden aufgrund ge-

setzlicher Bestimmungen aktuell gehalten, wohingegen für die Anzeige von Landnutzungsänderungen keine gesetzliche Verpflichtung besteht. Folglich ist die semantische Information in einem Landnutzungsdatenbestand im Gegensatz zu den geometrischen Begrenzungen häufig veraltet. Die in dieser Arbeit entwickelte Methode dient der automatischen Ableitung der semantischen Informationen aus aktuellen Sensordaten. Sofern zur Aktualisierung eine höhere semantische Auflösung erforderlich ist als sie die Klassifikation liefern kann, bildet die Klassifikation den ersten Schritt einer semi-automatischen Prozesskette, die die Aktualisierung einer gegebenen Datenbank zum Ziel hat. Selbst für den Fall, dass die Annahme der korrekten Begrenzungen verletzt ist, geht eine Grenzänderung häufig mit einer Änderung der Landnutzung einher, sodass das Verfahren implizit einen Hinweis auf Grenzänderungen gibt. Die Abbildung 4.3 zeigt die Form beider Bildprimitive am Beispiel einer ländlich geprägten Szene.



(a) Superpixel (b) Landnutzungsobjekte

Abbildung 4.3.: Eingefärbte Regionen repräsentieren die Bildprimitive der Bodenbedeckungs- (a) und Landnutzungsebene (b). Farben sind zufällig gewählt.

Bei Superpixeln handelt es sich um kleine Gruppen von räumlich aneinander grenzenden Pixeln mit ähnlichen spektralen Eigenschaften. Die Segmentierung von Superpixeln erfolgt in dieser Arbeit mit Hilfe der von Achanta et al. [2012] entwickelten Methode *Simple Linear Iterative Clustering* (SLIC). Hierbei handelt es sich um eine modifizierte Version des *K-Means Clusterings*. Die Größe und Kompaktheit der segmentierten Superpixel lässt sich mit zwei Parametern einstellen. Insbesondere der Kompaktheitsparameter hat einen Einfluss darauf, inwieweit sich die Superpixel in heterogenen Gebieten an Grauwertsprünge anpassen. In homogenen Gebieten nehmen SLIC Superpixel eine kompakte Form (quadratisch) an. Die Verwendung von Superpixeln anstelle von Pixeln bietet zum einen den Vorteil, dass sich die Rechenzeit deutlich reduziert bei vergleichsweise geringen Einbußen in der Genauigkeit. Darüber hinaus bewirken die Superpixel einen zusätzlichen Glättungseffekt, d.h. isolierte Fehlklassifikationen an vereinzelten Pixeln werden vermieden, da diese Pixel der sie umgebenden, dominanten Klasse zugewiesen werden, weil allen Pixeln innerhalb eines Superpixels die Klasse des Superpixels zugeordnet wird. Des Weiteren ist für die Klassifikation der Landnutzung ohnehin nicht die exakte pixelbasierte Verteilung der Bodenbedeckungsarten, sondern vielmehr die Information über deren Vorkommen und ungefähre Anordnung innerhalb eines Landnutzungsobjekts von Interesse. Für

diesen Zweck sind kleinräumige Superpixel in gleichem Maße geeignet wie Pixel, deren Klassifizierung jedoch deutlich mehr Rechenzeit erfordert.

Die räumlichen Kanten e_R modellieren die räumlichen Abhängigkeiten zwischen benachbarten Knoten innerhalb einer Ebene. Zwei Knoten gelten innerhalb einer Ebene als benachbart, wenn die Randpolygone der zugehörigen Bildprimitive über eine gemeinsame Kante verfügen. Die statistischen Abhängigkeiten zwischen der Bodenbedeckung und der Landnutzung werden mit Hilfe der semantischen Kanten e_S modelliert. Diese Kanten verbinden alle räumlich überlagernden Bildprimitive in beiden Ebenen miteinander, d.h. jeder Knoten der Landnutzungsebene ist mit allen Knoten der Bodenbedeckungsebene verbunden, die sich mit dem Landnutzungsobjekt räumlich überlappen. Dies führt zur Bildung von Cliquen höherer Ordnung, repräsentiert durch die Faktorknoten in Abbildung 4.2.

Das allgemeine Ziel besteht darin, die Klassenlabels der Bodenbedeckung y_i^c und Landnutzung y_k^u zu bestimmen. Die Klassenlabels sind als Zufallsvariablen der Knoten n_i bzw. n_k der jeweils zugehörigen Ebene modelliert, wobei $i \in \mathcal{S}^c$ und $k \in \mathcal{S}^u$ die Indizes der Bildprimitive und $\mathcal{S}^c$ und $\mathcal{S}^u$ die Menge aller Bildprimitive in der Bodenbedeckungs- bzw. der Landnutzungsebene beschreiben. Die Klassenlabels aller Bildprimitive werden in einem Labelvektor $\mathbf{y}^o = [y_1^o, ..., y_i^o, ..., y_n^o]^T$ pro Ebene $o \in \{c, u\}$ zusammengefasst. Ziel ist es, die wahrscheinlichste Konfiguration an Klassenlabels $\mathbf{y}^T = [\mathbf{y}^{c^T}, \mathbf{y}^{u^T}]$ mit y^c aus der Menge der Bodenbedeckungsklassen $\mathcal{L}^c$ und y^u aus der Menge der Landnutzungsklassen $\mathcal{L}^u$ simultan für alle Knoten des CRF unter Einbeziehung der Beobachtungen $\mathbf{x}$ zu bestimmen. Das CRF zur Klassifikation der Bodenbedeckung und der Landnutzung modelliert die a posteriori Wahrscheinlichkeit $P(\mathbf{y}|\mathbf{x})$ gemäß:

$$
\begin{aligned}
P(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \Bigg(& \prod_{i \in \mathcal{S}^c} \phi^c(y_i^c, \mathbf{x})^{\omega^1} \cdot \prod_{i \in \mathcal{S}^c} \prod_{j \in \mathcal{N}_i^c} \psi^c(y_i^c, y_j^c, \mathbf{x})^{\omega^2} \\
& \cdot \prod_{k \in \mathcal{S}^u} \phi^u(y_k^u, \mathbf{x})^{\omega^3} \cdot \prod_{k \in \mathcal{S}^u} \prod_{l \in \mathcal{N}_k^u} \psi^u(y_k^u, y_l^u, \mathbf{x})^{\omega^4} \\
& \cdot \prod_{h \in \mathcal{H}} \psi^{cu}(\mathbf{y}_h^c, \mathbf{y}_h^u)^{\omega^5} \Bigg).
\end{aligned}
\tag{4.1}
$$

In der Gleichung 4.1 bezeichnen $\phi^c(y_i^c, \mathbf{x})$ und $\phi^u(y_k^u, \mathbf{x})$ die *Assoziationspotentiale*, die jeweils die Relation zwischen den Klassenlabels y_i^c bzw. y_k^u und den Daten $\mathbf{x}$ modellieren. Die *paarweisen räumlichen Interaktionspotentiale* $\psi^c(y_i^c, y_j^c, \mathbf{x})$ und $\psi^u(y_k^u, y_l^u, \mathbf{x})$ modellieren die räumlichen Abhängigkeiten zwischen benachbarten Bildprimitiven innerhalb der Ebenen unter Berücksichtigung der Daten $\mathbf{x}$. $\mathcal{N}_i^c$ und $\mathcal{N}_k^u$ definieren die Nachbarschaft der Knoten n_i^c bzw. n_k^u. $\psi^{cu}(\mathbf{y}_h^c, \mathbf{y}_h^u)$ beschreibt das *semantische Potential höherer Ordnung*, mit Hilfe dessen die Relation zwischen den Klassenlabels y_j aller Knoten n_j modelliert wird, die zu der Clique $h \in \mathcal{H}$ gehören (d.h. dem Labelvektor $\mathbf{y}_h^c$ in der Bodenbedeckungs- und dem Labelvektor $\mathbf{y}_h^u$ in der Landnutzungsebene), wobei h den Index der Clique und $\mathcal{H}$ die Menge aller Cliquen bezeichnet. In einer Clique höherer Ordnung sind alle räumlich überlagernden Bildprimitive aus beiden Ebenen miteinander verbunden. Die Partitionsfunktion $Z(\mathbf{x})$ bewirkt die Normalisierung der Potentiale zu Wahrscheinlichkeiten. Die Parameter $\Omega = (\omega^1, \omega^2, \omega^3, \omega^4, \omega^5)$ bestimmen den Einfluss der einzelnen Potentialterme relativ zu dem ersten Potentialterm (d.h. $\omega^1 \equiv 1$). Ein höherer Wert für einen Parameter bewirkt einen geringeren Einfluss des Potentials wegen $\psi, \phi \in [0, 1]$.

4.2.2. Potentiale

4.2.2.1. Assoziationspotentiale

Das Assoziationspotential ist in dieser Arbeit definiert als die Wahrscheinlichkeit für die Zugehörigkeit des Knotens n_i zur Klasse y_i gegeben der Beobachtungen $\mathbf{x}$. Die Beobachtungen werden jedem Knoten der Bodenbedeckungs- (n_i^c) bzw. Landnutzungsebene (n_k^u) in Form von Merkmalsvektoren $\mathbf{f}_i^c(\mathbf{x})$ und $\mathbf{f}_k^u(\mathbf{x})$ zugeordnet. Diese Merkmalsvektoren können sich prinzipiell aus der Gesamtheit der Beobachtungen $\mathbf{x}$ ableiten. Bei der Klassifikation der Bodenbedeckung und der Landnutzung setzt sich der Merkmalsvektor aus bildbasierten, dreidimensionalen und geometrischen Merkmalen zusammen, wobei der Vektor in jeder Ebene aus einem unterschiedlichen Satz an Merkmalen bestehen kann. Die Potentiale nehmen für jeden Knoten einen Wert an, der gleich der Wahrscheinlichkeit von y_i^c und y_k^u gegeben der Merkmalsvektoren $\mathbf{f}_i^c(\mathbf{x})$ bzw. $\mathbf{f}_k^u(\mathbf{x})$ ist, d.h. $\phi^c(y_i^c, \mathbf{x}) = P(y_i^c|\mathbf{f}_i^c(\mathbf{x}))$ für die Bodenbedeckungs- und $\phi^u(y_k^u, \mathbf{x}) = P(y_k^u|\mathbf{f}_k^u(\mathbf{x}))$ für die Landnutzungsebene. Zur Bestimmung der Assoziationspotentiale wird in beiden Ebenen der RF Klassifikator [Breiman, 2001] verwendet. Der Wert des Assoziationspotentials pro Klasse berechnet sich dabei aus dem Verhältnis der Summe der Stimmen für die jeweilige Klasse zu der Gesamtanzahl der Stimmen, wobei jeder Entscheidungsbaum des Gesamtensembles eine Stimme abgibt (vgl. Gl. 3.22 in Kapitel 3.3). RF hat sich in zahlreichen Anwendungen, u.a. auch in der Photogrammetrie und Fernerkundung (z.B. [Schindler, 2012]), als effizienter Klassifikator bewährt. Darüber hinaus erfüllt RF die Voraussetzungen, die an die funktionalen Modelle der Potentiale in einem CRF gestellt werden [Kumar & Hebert, 2006]. Zum einen handelt es sich um ein diskriminatives Klassifikationsverfahren, d.h. RF modelliert direkt die a posteriori Wahrscheinlichkeit. Zum anderen lässt sich aus den Ergebnissen ein der Wahrscheinlichkeit ähnliches Maß ableiten, obwohl es sich bei dem RF Klassifikator originär um ein nicht-probabilistisches Verfahren handelt. Einige Parameter des RF Klassifikators müssen vorab definiert werden. Hierbei handelt es sich u.a. um die Mindestanzahl $N_{T,min}$ der Samples für nicht-terminierende Knoten, die maximale Tiefe T_{max} eines Baumes sowie die Anzahl B der Bäume im Ensemble (vgl. [Breiman, 2001]). Zusätzlich wird die maximale Anzahl $N_{T,max}$ an Trainingsbeispielen pro Klasse definiert, die zum Trainieren des Klassifikators verwendet werden. Mit der Einbeziehung einer gleichen Anzahl an Trainingsbeispielen pro Klasse wird sichergestellt, dass alle Klassen im Trainingsprozess gleichwertig repräsentiert sind [Chen et al., 2004]. Die Auswahl der $N_{T,max}$ Trainingsbeispiele pro Klasse erfolgt zufällig aus der Menge $\mathcal{T}$ der verfügbaren Trainingsdaten.

4.2.2.2. Paarweise räumliche Interaktionspotentiale

Das paarweise Interaktionspotential modelliert die räumliche Abhängigkeit von Klassenlabels an benachbarten Knoten n_i und n_j innerhalb einer Ebene unter Berücksichtigung der Beobachtungen $\mathbf{x}$. Jeder Kante, d.h. Verbindung benachbarter Knoten, ist ein Vektor von Kantenmerkmalen $\boldsymbol{\mu}_{ij}(\mathbf{x})$ zugeordnet. Die Merkmalsvektoren $\boldsymbol{\mu}_{ij}^c(\mathbf{x})$ und $\boldsymbol{\mu}_{kl}^u(\mathbf{x})$ in beiden Ebenen leiten sich aus den Merkmalsvektoren der jeweils verbundenen Knoten ab, wobei diese entweder aneinandergereiht oder elementweise subtrahiert werden können. Alternativ kann die euklidische Distanz zwischen den Knotenmerkmalen als Interaktionsmerkmal verwendet werden, was sich jedoch in vorangegangenen Tests als für diese

Aufgabe ungeeignet herausgestellt hat und daher an dieser Stelle nicht verwendet wird.

Die Bestimmung der paarweisen Interaktionspotentiale erfolgt in beiden Ebenen mit Hilfe eines statistischen Klassifikators. Im Gegensatz zu anderen Modellen, wie z.B. dem kontrastsensitiven Potts-Modell, forciert ein statistischer Klassifikator keine generelle Glättung der Ergebnisse, sondern favorisiert wahrscheinliche Konfigurationen von Klassenlabels unter Berücksichtigung der Daten. Der Klassifikator lernt die Wahrscheinlichkeit für das Auftreten einer bestimmten Relation anhand von repräsentativen Trainingsdaten. Dies stellt eine realistischere Beschreibung der räumlichen Abhängigkeit benachbarter Superpixel und Landnutzungsobjekte dar als die schlichte Annahme gleicher Klassenlabels. Folglich ist das Interaktionspotential modelliert als die gemeinsame a posteriori Wahrscheinlichkeit der Klassenlabels y_i^c und y_j^c gegeben $\boldsymbol{\mu}_{ij}^c(\mathbf{x})$, d.h. $\psi^c(y_i^c, y_j^c, \mathbf{x}) = P(y_i^c, y_j^c | \boldsymbol{\mu}_{ij}^c(\mathbf{x}))$ für die Bodenbedeckungsebene, sowie der Klassenlabels y_k^u und y_l^u gegeben $\boldsymbol{\mu}_{kl}^u(\mathbf{x})$, d.h. $\psi^u(y_k^u, y_l^u, \mathbf{x}) = P(y_k^u, y_l^u | \boldsymbol{\mu}_{kl}^u(\mathbf{x}))$ für die Landnutzungsebene. Die Bestimmung des Potentials entspricht damit einer herkömmlichen Klassifikationsaufgabe. Folglich ist es möglich, jeden beliebigen lokalen diskriminativen Klassifikator mit probabilistischem Output anzuwenden [Kumar & Hebert, 2006]. Der Unterschied besteht lediglich darin, dass jede Kombination von Klassenlabels (y_i, y_j) an benachbarten Knoten n_i und n_j als separate Klasse unterschieden wird, d.h. bei M zu unterscheidenden Klassen gilt es M^2 Klassenkombinationen zu unterscheiden. Um den Klassifikator adäquat trainieren zu können, müssen die Trainingsdaten für jede Klassenkombination eine ausreichende Anzahl an Trainingsbeispielen vorhalten, wodurch sich der erforderliche Umfang der Trainingsdaten deutlich erhöht. Analog zu dem Assoziationspotential erfolgt die Bestimmung der paarweisen Interaktionspotentiale in beiden Ebenen mit Hilfe des RF Klassifikators.

4.2.2.3. Semantisches Potential höherer Ordnung

Das semantische Potential höherer Ordnung modelliert die statistischen Abhängigkeiten der Klassenlabels der Bodenbedeckung y_i^c und der Landnutzung y_k^u aller Knoten n_i^c und n_k^u in beiden Ebenen, die zu einer Clique höherer Ordnung h verbunden sind. Das semantische Potential höherer Ordnung sollte einen Wert annehmen, der proportional zur gemeinsamen a posteriori Wahrscheinlichkeit der Menge an Klassenlabels $\mathbf{y}_h^c$ in der Bodenbedeckungs- und $\mathbf{y}_h^u$ in der Landnutzungsebene aller zu einer Clique h verbundenen Knoten n_h ist, d.h. $\psi^{cu}(\mathbf{y}_h^c, \mathbf{y}_h^u) \propto P(\mathbf{y}_h^c, \mathbf{y}_h^u)$. Die Beobachtungen $\mathbf{x}$ fließen nicht in die Bestimmung des Potentials ein. Somit ist das Potential ausschließlich von den Klassenlabels abhängig, was in diesem Zusammenhang die Konfidenzwerte aller möglichen Klassenlabels einschließt.

Wie bereits in Kapitel 2.3 diskutiert, existieren keine effizienten Inferenzmethoden für CRF höherer Ordnung, die eine generische Formulierung der Potentiale höherer Ordnung verwenden bei einer gleichzeitig hohen Anzahl an Zufallsvariablen pro Clique. Ziel ist es, das Modell anhand von repräsentativen Trainingsdaten zu lernen und zugleich eine effiziente Inferenz zu ermöglichen. Zu diesem Zweck wurde ein iterativer Inferenzalgorithmus entwickelt, im Rahmen dessen das semantische Potential höherer Ordnung approximiert wird. Der iterative Inferenzalgorithmus wird in Kapitel 4.2.3 vorgestellt.

4.2.3. Iterativer Inferenzalgorithmus

Im Rahmen der Inferenz wird die wahrscheinlichste Konfiguration von Klassenlabels $\mathbf{y}$ simultan für alle Knoten des CRF bestimmt. Hierzu wird die a posteriori Wahrscheinlichkeit $P(\mathbf{y}|\mathbf{x})$ der Klassenlabels bei gegebenen Daten maximiert. Die Bestimmung einer exakten Lösung ist für Mehrklassenprobleme rechnerisch nicht möglich [Kumar & Hebert, 2006]. Daher werden zur Inferenz in CRF ausschließlich approximative Methoden verwendet, wie z.B. Loopy Belief Propagation (LBP) [Frey & MacKay, 1998]. Die Inferenz in dem Zwei-Ebenen-CRF-Modell erfolgt ebenfalls auf Grundlage von LBP, allerdings in einer modifizierten Form. Wie zuvor erwähnt, führt die Verwendung von Potentialen höherer Ordnung zu Einschränkungen in Bezug auf die Inferenz. Um die Inferenz trotzdem effizient sicherstellen zu können, wird die optimale Konfiguration des graphischen Modells in einem iterativen Ansatz bestimmt, der konzeptionell an ein Verfahren von Roig et al. [2011] angelehnt ist. Das Grundkonzept ist hierbei, dass die Inferenz zunächst unabhängig für jede Ebene durchgeführt wird. Anschließend werden die Ergebnisse der anderen Ebene kommuniziert, wo sie dann in die nächste Iteration der Inferenz einfließen. Die iterative Vorgehensweise erlaubt es, dass sich die Klassifikationsaufgaben während der Inferenz durch den gegenseitigen Austausch von Kontextinformation beeinflussen, während die Trennung eine effiziente Inferenz ermöglicht. Diese Vorgehensweise orientiert sich an dem Expectation Maximisation Ansatz [Dempster et al., 1977], wobei es sich um ein allgemeines Konzept zur Bestimmung von mehreren Parametergruppen handelt, die in einem Abhängigkeitsverhältnis zueinander stehen und sich daher nicht simultan bestimmen lassen. Hierbei wird abwechselnd eine partielle Lösung für die Parameter der einen Gruppe festgehalten, auf Basis dessen die Parameter der zweiten Gruppe berechnet werden. Diese Vorgehensweise wird iterativ wiederholt bis sich die Parameter nur noch unwesentlich ändern.

In jeder Iteration t wird die wahrscheinlichste Konfiguration an Klassenlabels separat für jede Ebene bestimmt. Das Ergebnis bilden die Zwischenlösungen $\widetilde{\mathbf{y}}_t^c$ für die Bodenbedeckungs- und $\widetilde{\mathbf{y}}_t^u$ für die Landnutzungsebene. Dieses Vorgehen wird durch die Approximation der Energiefunktion ermöglicht, indem die Potentiale höherer Ordnung $\psi^{cu}(\mathbf{y}_h^c, \mathbf{y}_h^u)$ durch Faktorisierung vereinfacht werden zu einem unären Term pro Ebene gemäß

$$\psi^{cu}(\mathbf{y}_h^c, \mathbf{y}_h^u) \propto \prod_{i \in \mathcal{S}^c} \phi^{u \to c}(y_i^c, \mathbf{y}_h^u)^{\omega^{5,u \to c}} \cdot \prod_{k \in \mathcal{S}^u} \phi^{c \to u}(y_k^u, \mathbf{y}_h^c)^{\omega^{5,c \to u}}. \tag{4.2}$$

Der unäre Term pro Ebene ist definiert als die Wahrscheinlichkeit von y_i^c bzw. y_k^u an den Knoten n_i^c bzw. n_k^u gegeben der Klassenlabels $\mathbf{y}_h^u$ bzw. $\mathbf{y}_h^u$ der räumlich überlagernden Bildprimitive n_h in der jeweils anderen Ebene, die gemeinsam mit den Knoten n_i^c bzw. n_k^u eine Clique höherer Ordnung h bilden, d.h. $\phi^{u \to c}(y_i^c, \mathbf{y}_h^u) = P(y_i^c|\mathbf{y}_h^u)$ für die Bodenbedeckungs- und $\phi^{c \to u}(y_k^u, \mathbf{y}_h^c) = P(y_k^u|\mathbf{y}_h^c)$ für die Landnutzungsebene. Infolgedessen lassen sich zur Minimierung der Energiefunktion approximative Inferenzmethoden anwenden, die für paarweise CRF konzipiert sind, wie z.B. LBP. Die Approximation der Potentiale höherer Ordnung durch je zwei unäre Terme wird durch zwei vereinfachende Annahmen realisiert: Erstens hängt das Potential höherer Ordnung für die Knoten einer Ebene nur von den Klassenlabels der anderen Ebene ab, und zweitens werden die Klassenlabels der anderen Ebene in jeder Iteration als konstant angenommen, d.h. die zugehörigen Zufallsvariablen werden zu Konstanten. Infolgedessen ist es möglich, die Energiefunktion in eine Funktion pro Ebene aufzuspalten und die

Teilfunktionen separat zu lösen. Die konstanten Werte leiten sich aus den Zwischenlösungen $\widetilde{\mathbf{y}}^c_{t-1}$ und $\widetilde{\mathbf{y}}^u_{t-1}$ der vorausgegangenen Iteration ab. Im Gegensatz zu Roig et al. [2011], die ihre Potentiale ausschließlich auf Basis der aktuellen Klassenlabels bestimmen, werden bei dem hier vorgestellten Ansatz zusätzlich die Konfidenzwerte aller Klassenlabels berücksichtigt. Der Inferenzalgorithmus wird solange fortgesetzt bis sich das Klassifikationsergebnis nicht mehr signifikant ändert.

Die Bestimmung der vereinfachten Potentialterme erfolgt in beiden Ebenen mit Hilfe eines statistischen Klassifikators. Allerdings werden die Klassenlabels inklusive der zugehörigen Konfidenzwerte nicht direkt dem Klassifikator übergeben, sondern dienen vielmehr der Ableitung von diskriminativen Kontextmerkmalen. Diese Merkmale charakterisieren die komplexen Abhängigkeiten innerhalb einer Clique höherer Ordnung. Sie werden pro Bildprimitiv aus dem Zwischenergebnis der anderen Ebene abgeleitet, während die ursprünglichen Eingangsdaten keine Berücksichtigung finden. Für beide Klassifikationsaufgaben wird ein unterschiedlicher Satz an Kontextmerkmalen extrahiert, die den Knoten in jeder Ebene in Form von zusätzlichen Merkmalsvektoren $\mathbf{m}^c_i(\mathbf{y}^u_h)$ und $\mathbf{m}^u_k(\mathbf{y}^c_h)$ zugeordnet sind. Bei den Merkmalsvektoren handelt es sich um Funktionen der Klassenlabels (inklusive Konfidenzwerte) ohne die Berücksichtigung der Daten. Diese Merkmale fließen in einen diskriminativen Klassifikator ein, der die Potentiale höherer Ordnung während der Inferenz approximiert. Folglich basieren die Potentiale auf der Wahrscheinlichkeit von y^c_i und y^u_k gegeben der Merkmalsvektoren $\mathbf{m}^c_i(\mathbf{y}^u_h)$ und $\mathbf{m}^u_k(\mathbf{y}^c_h)$, d.h. $\phi^{u\rightarrow c}(y^c_i, \mathbf{y}^u_h) = P(y^c_i|\mathbf{m}^c_i(\mathbf{y}^u_h))^{\omega^{5,u\rightarrow c}}$ für die Bodenbedeckungs- und $\phi^{c\rightarrow u}(y^u_k, \mathbf{y}^c_h) = P(y^u_k|\mathbf{m}^u_k(\mathbf{y}^c_h))^{\omega^{5,c\rightarrow u}}$ für die Landnutzungsebene. Zur Bestimmung der Potentiale wird in beiden Ebenen ein RF Klassifikator [Breiman, 2001] verwendet.

Der detaillierte Ablauf des iterativen Inferenzalgorithmus basierend auf LBP ist nachfolgend beschrieben. Zu Beginn wird in jeder Ebene separat eine bestimmte Anzahl n_{LBP} von Iterationen des LBP-Verfahrens durchgeführt. Vorab werden die Kantenpotentiale an jedem Knoten mit der Gesamtanzahl eingehender Kanten normalisiert, um zu vermeiden, dass die Kantenpotentiale an Knoten mit einer hohen Anzahl an eingehenden Kanten einen stärkeren Einfluss haben als an Knoten mit einer geringeren Anzahl [Hoberg et al., 2015]. Das Potential höherer Ordnung wird im Rahmen von LBP als zusätzlicher unärer Term berücksichtigt, der während der Inferenz mit LBP konstant bleibt. Als Ergebnis liefert LBP für jeden Knoten die temporären Konfidenzwerte (sogenannte *Beliefs*) für jedes Klassenlabel, die zusammengefasst für alle Knoten einer Ebene die aktuellen Zwischenergebnisse für die Bodenbedeckung und die Landnutzung ergeben. Anschließend wird das LBP-Verfahren unterbrochen, um die zu unären Termen vereinfachten Potentiale höherer Ordnung in jeder Ebene auf Basis der Zwischenergebnisse zu aktualisieren. Zu diesem Zweck erfolgt zunächst die Aktualisierung der Kontextmerkmale auf Grundlage der aktuellen Zwischenergebnisse. Anschließend werden auf Grundlage der aktualisierten Merkmalsvektoren $\mathbf{m}^c_i$ und $\mathbf{m}^u_k$ neue Werte für das Potential höherer Ordnung abgeleitet. Aufgrund der Aufspaltung des Potentials in einen Term pro Ebene erfolgt die Neubestimmung auf Basis neuer Merkmalsvektoren separat für jede Ebene des Zwei-Ebenen-CRF-Modells. Die Ergebnisse fließen erneut als unäre Terme in die separat durchgeführten LBP-Prozesse ein, die im Anschluss an den Aktualisierungsschritt erneut mit n_{LBP} Iterationen fortgesetzt werden. In Folge des Aktualisierungsschrittes haben sich die Potentiale höherer Ordnung geändert, was die weitere Entwicklung der Nachrichten in dem LBP-Verfahren maßgeblich beeinflusst. Dieses Vorgehen wird so lange wiederholt bis eine vorgegebene Anzahl n_{It} an Iterationen im Inferenzalgorithmus erreicht

ist. Abschließend erfolgt die Berechnung der finalen Konfidenzwerte für jeden Knoten. Jedem Knoten wird das Klassenlabel mit dem maximalen Konfidenzwert zugewiesen. Die Anzahl n_{It} der Iterationen des iterativen Inferenzalgorithmus (äußere Iterationen) sowie die Anzahl n_{LBP} der Iterationen in jedem LBP-Teilschritt (innere Iterationen) sind manuell vorzugeben. Für die Inferenz mit LBP wird im Rahmen dieser Arbeit die *Sum-Product*-Variante verwendet [Frey & MacKay, 1998].

In der ersten Iteration stehen noch keine Klassifikationsergebnisse zur Ableitung der Kontextmerkmale zur Verfügung. Daher wird vorab eine initiale Klassifikation der Bodenbedeckung und der Landnutzung auf Basis von Bild- und Höhendaten durchgeführt, deren Ergebnisse zur Ableitung von initialen Merkmalswerten dienen. Diese Klassifikation erfolgt nur anhand der Assoziationspotentiale unabhängig für jede Ebene.

4.2.4. Training

CRF gehören zur Gruppe der überwachten Klassifikationsverfahren, d.h. die Parameter der Potentiale werden gelernt. Das Training der Assoziations-, der paarweisen Interaktionspotentiale und des semantischen Potentials höherer Ordnung erfolgt separat anhand von repräsentativen Trainingsdaten. Dies beinhaltet das Trainieren der zugehörigen RF Klassifikatoren. Daneben gilt es, die Parameter ω^2, ω^4, $\omega^{5,c\rightarrow u}$ und $\omega^{5,u\rightarrow c}$, die Anzahl n_{It} der Iterationen der Inferenzprozedur, die Anzahl n_{LBP} der Iterationen eines LBP-Teilschrittes und diverse Parameter der RF Klassifikatoren zu definieren. Diese Parameter werden empirisch festgelegt. Der Einfluss der Parameter wird im Rahmen einer Sensitivitätsanalyse analysiert (siehe Kapitel 5.2.4). Bei den RF Parametern handelt es sich um die maximale Anzahl $N_{T,max}$ der Trainingsbeispiele pro Klasse, die Mindestanzahl $N_{T,min}$ der Samples für nicht-terminierende Knoten, die maximale Tiefe T_{max} eines Baumes und die Anzahl B der Entscheidungsbäume im Ensemble. Da sich die Klassifikationsaufgaben strukturell unterscheiden, müssen diese Parameter individuell gewählt werden. Für das Training der paarweisen räumlichen Interaktionspotentiale sind vollständig ausgelabelte Trainingsdaten erforderlich, um die Relationen zwischen adjazenten Bildprimitiven adäquat lernen zu können.

Wie zuvor beschrieben, basiert die Klassifikation auf Kontextmerkmalen. Um die RF Klassifikatoren der zu unären Potentialen vereinfachten Potentiale höherer Ordnung korrekt anlernen zu können, müssen diese Merkmale bereits für das Training zur Verfügung stehen. Allerdings stehen die Eingangsdaten für die Extraktion der Kontextmerkmale, d.h. die Klassifikationsergebnisse, während des Trainingsschrittes noch nicht zur Verfügung. Die Extraktion könnte alternativ auf der Grundlage von Trainingsdaten erfolgen, jedoch müssten hierzu für alle Pixel innerhalb der Landnutzungsobjekte, die zum Training verwendet werden, die Klassenlabels bekannt sein. Vor dem Hintergrund der zumeist weitläufigen Ausdehnung von Landnutzungsobjekten würde dies einen hohen Aufwand bei der Erstellung von Trainingsdaten bedeuten. Aus diesem Grund wird vorweg eine initiale Klassifikation durchgeführt, deren Ergebnisse dann als Eingangsdaten für die initiale Berechnung der Kontextmerkmale dienen. Die initiale Klassifikation basiert ausschließlich auf den Assoziationspotentialen ohne die Berücksichtigung von Kontext.

Durch den Segmentierungsprozess kann es unter Umständen zu Misch-Superpixeln kommen, die mehrere Bodenbedeckungsarten in einem Segment vereinen. Folglich sind diese Superpixel nicht re-

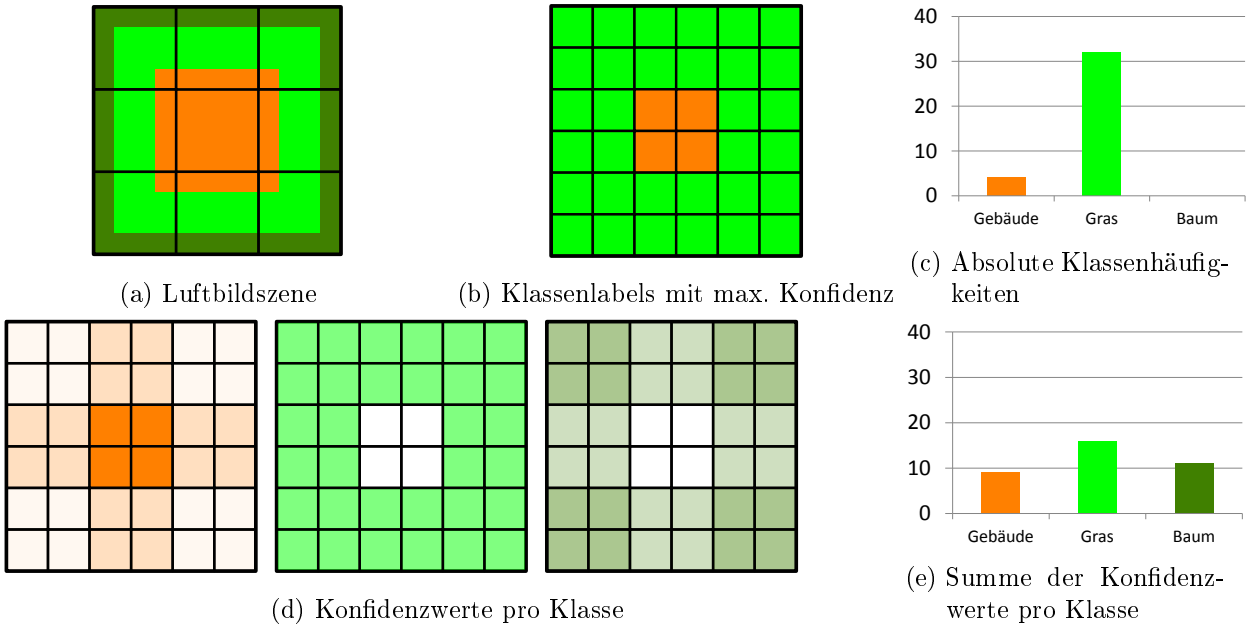
präsentativ für eine Klasse und erfüllen damit nicht die Qualitätsansprüche, die an Trainingsdaten im Allgemeinen gestellt werden. Um diese Unsicherheit im Rahmen des Trainings zu berücksichtigen, werden für das Training der Assoziationspotentiale $\phi^c(y_i^c, \mathbf{x})$ und der paarweisen räumlichen Interaktionspotentiale $\psi^c(y_i^c, y_j^c, \mathbf{x})$ in der Bodenbedeckungsebene nur jene Superpixel verwendet, bei denen mindestens 75 % der Pixel bzgl. ihrer Referenzlabel zu einer Klasse gehören.

Für das Trainieren der zu unären Termen vereinfachten Potentiale höherer Ordnung $\phi^{u\to c}(y_i^c, \mathbf{y}_h^u)$ stehen in der Bodenbedeckungsebene typischerweise nicht genügend Trainingsdaten zur Verfügung, um die Relation adäquat lernen zu können. Dies liegt vor allem darin begründet, dass die Gebiete, in denen Referenzdaten für die Bodenbedeckung vorliegen, räumlich begrenzt sind und infolgedessen nur von wenigen Landnutzungsobjekten überlagert werden. Folglich ist die Abhängigkeit der Bodenbedeckung von der Landnutzung nur eingeschränkt in den Trainingsdaten repräsentiert. Eine großräumige Erfassung von Referenzdaten für die Bodenbedeckung ist vor allem durch den hohen manuellen Bearbeitungsaufwand nicht praktikabel. Aus diesem Grund werden ergänzend zu den verfügbaren Referenzdaten die Klassenlabels aus einer initialen Klassifikation als Referenzdaten zum Trainieren der zu unären Termen vereinfachten Potentiale höherer Ordnung $\phi^{u\to c}(y_i^c, \mathbf{y}_h^u)$ in der Bodenbedeckungsebene verwendet, sofern sie mit einer Wahrscheinlichkeit von mindestens 75 % bestimmt wurden. Mit diesem Schritt wird erreicht, dass die Abhängigkeit der Bodenbedeckung von der Landnutzung nicht nur in den räumlich begrenzten Bereichen in den Trainingsgebieten gelernt wird, in denen Referenzdaten für die Bodenbedeckung vorliegen, sondern darüber hinaus auch an anderen Stellen in den Trainingsgebieten, sofern eine hohe Konfidenz der initialen Klassifikationsergebnisse gegeben ist. Dies führt zu einer deutlichen Erhöhung der Menge an Trainingsdaten, die allerdings mit einer höheren Unsicherheit dieser Trainingsdaten einhergeht.

4.3. Kontextmerkmale

Die Kontextbeziehungen zwischen den Ebenen, d.h. die statistischen Abhängigkeiten zwischen Bodenbedeckung und Landnutzung, werden mit Hilfe von Kontextmerkmalen modelliert. Für die Klassifikation werden verschiedene Arten von Kontextmerkmalen verwendet, die sich ausnahmslos aus den Klassifikationsergebnissen räumlich überlagernder Bildprimitive in beiden Ebenen ableiten. Hingegen fließen die ursprünglichen Sensordaten nicht in die Berechnung ein. Die Kontextmerkmale werden separat für jedes Bildprimitiv der Clique höherer Ordnung berechnet und beschreiben die komplexe Konfiguration an Klassenlabels und Konfidenzwerten innerhalb der zugehörigen Clique. Die Kontextmerkmale werden in einem separaten Merkmalsvektor $\mathbf{m}_i^o$ pro Knoten n_i^o in beiden Ebenen $o \in \{c, u\}$ zusammengefasst. Diese Merkmalsvektoren bilden jeweils die Grundlage für die Bestimmung der zu unären Termen vereinfachten Potentiale höherer Ordnung. Wie bereits in Kapitel 2.1 und 3.2.4 beschrieben, lassen sich die Kontextmerkmale in zwei Gruppen unterteilen: räumliche Metriken und graphenbasierte Merkmale. Neben den in Kapitel 3.2.4 vorgestellen Kontextmerkmalen werden weitere Merkmale aus der Gruppe der räumlichen Metriken extrahiert, deren Berechnung nachfolgend dargestellt ist. Für die Klassifikation der Bodenbedeckung und der Landnutzung wird ein unterschiedlicher Satz an Merkmalen verwendet.

Konfidenz-gewichtete relative Flächenanteile Die erste Gruppe von Merkmalen stellt eine Erweiterung der in Kapitel 3.2.4 vorgestellten Merkmale dar, welche den absoluten bzw. relativen Flächenanteil einer Klasse an einem Segment beschreiben. Die Abbildung 4.4 veranschaulicht anhand eines einfachen konstruierten Beispiels das grundlegende Prinzip zur Berechnung der neu entwickelten Merkmale und zeigt Unterschiede zu den in Kapitel 3.2.4 vorgestellten Merkmalen auf. Die Abbildung 4.4a



(a) Luftbildszene (b) Klassenlabels mit max. Konfidenz (c) Absolute Klassenhäufigkeiten

(d) Konfidenzwerte pro Klasse (e) Summe der Konfidenzwerte pro Klasse

Abbildung 4.4.: Schematische Darstellung einer Luftbildszene innerhalb eines quadratischen Landnutzungsobjekts (a), Ergebnisse der Bodenbedeckungsklassifikation, d.h. Klassenlabels mit maximalen Konfidenzwert (b) und Konfidenzwerte für die Klassen *Gebäude* (d, links), *Gras* (d, Mitte) und *Baum* (d, rechts), wobei die Sättigung der Farben proportional zum Konfidenzwert zunimmt. Die Farbe markiert die Zugehörigkeit zu den Klassen Gebäude (orange), Gras (hellgrün) und Baum (dunkelgrün). Die schwarzen, durchgehenden Linien kennzeichnen in (a) die Ausdehnung der Superpixel und in (b) und (d) das Pixelraster. In dem Histogramm in (c) sind die absoluten, pixelweisen Klassenhäufigkeiten und in (e) die Summe der pixelbasierten Konfidenzwerte pro Klasse aufgetragen.

zeigt die tatsächliche Bodenbedeckung innerhalb eines quadratischen Landnutzungsobjekts, das es zu klassifizieren gilt. Die kleineren Quadrate in Abbildung 4.4a repräsentieren die Bildprimitive der Bodenbedeckung, d.h. die Superpixel, denen im Rahmen der Klassifikation eine Bodenbedeckungsart zuzuweisen ist. Das Klassifikationsergebnis der Bodenbedeckung ist in der Abbildung 4.4b gezeigt, wobei das Ergebnis pro Superpixel bestimmt und anschließend auf Pixelniveau abgebildet wurde. Aus diesem Ergebnis leitet sich der in Abbildung 4.4c pro Klasse aufgetragene absolute Flächenanteil von 4 Pixeln für die Klasse *Gebäude* und 32 Pixeln für die Klasse *Gras* ab. Diese Merkmale werden in dem Sinne erweitert, dass in die Berechnung zusätzlich die Konfidenz des Klassifikationsergebnisses einfließt, wie es auch in den Arbeiten von Munoz et al. [2010] und Xiong et al. [2011] erfolgt. Bei den erweiterten Merkmalen handelt es sich um die relativen Flächenanteile pro Klasse an der Gesamtfläche eines Segments, wobei jedes Pixel nicht mit dem gleichen Einfluss, sondern gewichtet mit der vom Klassifikator mitgelieferten Konfidenz in die Berechnung einfließt. Das Berechnungskonzept ist für

Landnutzungsobjekte schematisch in Abbildung 4.4 visualisiert. Analog zur Berechnung der absoluten Flächenanteile wird das Klassifikationsergebnis der Bodenbedeckung auf das Pixelniveau abgebildet, allerdings mit dem Unterschied, dass jedem Pixel nicht ein einzelnes Klassenlabel, sondern die Konfidenzwerte aller Klassen zugewiesen wird. Zur vereinfachten Darstellung sind die Konfidenzwerte pro Klasse in der Abbildung 4.4d jeweils in einem eigenen Ausschnitt pixelbasiert aufgetragen, wobei die Sättigung der Farben das jeweilige Konfidenzniveau widerspiegelt (maximale Sättigung entspricht einer Konfidenz von 100 %). In die Berechnung des Flächenanteils fließen diese Konfidenzwerte als Gewichte ein. Folglich ist es möglich, Unsicherheiten des Klassifikationsergebnisses bei der Merkmalsextraktion zu berücksichtigen. In dem in Abbildung 4.4d skizzierten Beispiel treten Unsicherheiten in den beiden äußeren Ringen des Landnutzungssegments auf, die sich dadurch äußern, dass für den äußeren Ring die Klasse *Baum* und in dem zweiten Ring die Klasse *Gebäude* nahezu gleich wahrscheinlich zur Klasse *Gras* ist. Diese Unsicherheiten können beispielsweise daraus resultieren, dass ein Superpixel mehrere Bodenbedeckungen einschließt und damit nicht eindeutig einer Klasse zugewiesen werden kann. In dem skizzierten Beispiel wird allerdings für die Pixel in beiden Ringen eine höhere Konfidenz für die Klasse *Gras* erzielt, sodass den Pixeln dieses Klassenlabel als Ergebnis zugewiesen wird (vgl. Abb. 4.4b). Folglich ist die Klasse *Baum* in den in Abbildung 4.4c dargestellten absoluten Flächenanteilen nicht enthalten, aber sehr wohl in den gewichteten Flächenanteilen in Abbildung 4.4e. Darüber hinaus wirkt sich die Berücksichtigung der Konfidenz auch auf die Verhältnisse zwischen den Klassen aus. So ist das Landnutzungsobjekt, nach den absoluten Flächenanteilen zu urteilen, durch die Bodenbedeckung *Gras* dominiert, wohingegen sich die gewichteten Flächenanteile von *Gebäude* und *Gras* einander annähern und somit besser die in Abbildung 4.4a dargestellte tatsächliche Verteilung der Bodenbedeckungsarten widerspiegeln. Um eine bessere Vergleichbarkeit zwischen den Landnutzungsobjekten zu gewährleisten, wird das Histogramm der gewichteten Flächenanteile normalisiert, d.h. jeder Eintrag wird durch die Summe der pixelbasierten Konfidenzwerte aller Klassen in dem betrachteten Landnutzungsobjekt dividert. Die Kontextmerkmale $f_{belArea,l}$ pro Klasse $l \in \mathcal{L}$ sind für jedes zu klassifizierende Segment wie folgt definiert:

$$f_{belArea,l} = \frac{1}{\sum_{l \in \mathcal{L}} \sum_{(r,c) \in \mathcal{R}} P(y_{(r,c)} = l)} \sum_{(r,c) \in \mathcal{R}} P(y_{(r,c)} = l), \tag{4.3}$$

wobei $\mathcal{R}$ der Menge aller Pixel an den Positionen (r, c) in dem jeweiligen Segment entspricht. $P(y_{(r,c)} = l)$ bezeichnet die für die Zufallsvariable $y_{(r,c)}$ ermittelten Konfidenzwerte für die Klassen $l \in \mathcal{L}$. Mit dem Index (r, c) der Zufallsvariable wird ausgedrückt, dass die originär pro Bildprimitiv bestimmten Verteilungen auf das Pixelniveau abgebildet werden. Diese Art von Merkmalen wird sowohl für die GIS-Objekte in der Landnutzungs- als auch für die Superpixel in der Bodenbedeckungsebene berechnet. Bei der Berechnung der Merkmalsvektoren $\mathbf{m}_i^c$ auf Basis der Superpixel bezieht sich $\mathcal{R}$ auf die Menge der Pixel, die das Superpixel bilden, und $\mathcal{L} = \mathcal{L}^u$ bezeichnet die Landnutzungsklassen. Bei der Berechnung der Merkmalsvektoren $\mathbf{m}_k^u$ auf Ebene der Landnutzungsobjekte verweist $\mathcal{R}$ auf die Pixel innerhalb des GIS-Objekts und $\mathcal{L} = \mathcal{L}^c$ auf die Bodenbedeckungsklassen.

Zentrale Momente Die zweite Gruppe von Merkmalen bilden die zentralen Momente erster und zweiter Ordnung der räumlichen Verteilung der Bodenbedeckungsarten innerhalb eines Landnutzungsobjekts, die folglich nur für die Bildprimitive der Landnutzungsebene, d.h. $\mathbf{m}_k^u$, abgeleitet werden.

Diese Merkmale beschreiben die pixelbasierte Verteilung der unterschiedlichen Bodenbedeckungsarten innerhalb eines Landnutzungsobjekts. Für deren Berechnung wird das Klassifikationsergebnis der Bodenbedeckung erneut auf das Pixelniveau abgebildet. Um eine Vergleichbarkeit der Merkmale zu gewährleisten, werden die pixelbasierten Bodenbedeckungen vom Objekt- in ein lokales Bildkoordinatensystem transformiert, dessen Ursprung im Schwerpunkt des Landnutzungspolygons gelagert ist und das entsprechend dessen Hauptrichtung, d.h. in Richtung des zum größten Eigenwert gehörigen Eigenvektors, ausgerichtet ist. Die zentralen Momente μ_{pq} der Ordnung p, q berechnen sich für die Region $\mathcal{R}$ in einem Binärbild [Burger & Burge, 2005] gemäß

$$\mu_{pq} = \sum_{(r,c)\in\mathcal{R}} (r - \bar{r})^p \cdot (c - \bar{c})^q, \tag{4.4}$$

mit den Bildkoordinaten $(\bar{r}, \bar{c})$ des Schwerpunkts der Region und den Bildkoordinaten (r, c) der in der Region enthaltenen Pixel. Die Berechnung der zentralen Momente erfolgt separat für jede Bodenbedeckungsklasse, d.h. die Segmente der verschiedenen Bodenbedeckungsarten bilden jeweils ein eigenes Binärbild, für das die zentralen Momente zu bestimmen sind. Die Berechnung unterscheidet sich von der allgemeinen Formel in Gleichung 4.4 in dem Punkt, dass sich die zentralen Momente nicht auf den Schwerpunkt der jeweiligen Region im Binärbild beziehen, sondern auf den Schwerpunkt des Landnutzungspolygons, der den Ursprung des lokalen Koordinatensystems bildet. Damit charakterisieren die zentralen Momente die Verteilung der Bodenbedeckungsarten in Bezug auf das Landnutzungspolygon. Die Werte der zentralen Momente sind von der absoluten Größe der Region $\mathcal{R}$ abhängig, daher werden sie üblicherweise mit der Gesamtanzahl der Pixel innerhalb der Region normiert [Burger & Burge, 2005]. Um auch hier eine Vergleichbarkeit auf Ebene der Landnutzungspolygone zu gewährleisten, werden die zentralen Momente mit der Gesamtanzahl N_R der Pixel pro Landnutzungspolygon normalisiert. Folglich ergeben sich die normalisierten zentralen Momente gemäß

$$\bar{\mu}_{pq} = \mu_{pq} \cdot \left(\frac{1}{N_R}\right)^{(p+q+2)/2}. \tag{4.5}$$

4.4. Verifikation und Aktualisierung

Zum Zwecke der Verifikation wird für jedes GIS-Objekt die prädizierte Landnutzungsklasse mit dem möglicherweise veralteten Eintrag in der Datenbank verglichen. Somit lassen sich Widersprüche zwischen der Datenbank und den aktuellen Sensordaten aufdecken. Die Effektivität des Verifikationsprozesses wird dabei maßgeblich durch die semantische Auflösung der Klassenstruktur beeinflusst. So können nur jene Nutzungsarten eindeutig verifiziert werden, die eine eigene Klasse in der Klassifikation bilden. In vielen Fällen sind jedoch mehrere Nutzungsarten des Objektartenkatalogs der Datenbank zu einer Klasse zusammengefasst, für die eine Trennung auf Basis von Sensordaten nicht möglich ist. In diesen Fällen kann im Rahmen der Verifikation nur festgestellt werden, ob der Eintrag in der Datenbank mit der klassifizierten, übergeordneten Klasse konsistent ist, d.h. ein Element dieser Klasse darstellt. Bei einer Übereinstimmung kann allerdings nicht ausgeschlossen werden, dass sich die Nutzung innerhalb der übergeordneten Klasse geändert hat. Folglich ist es das Ziel, eine möglichst feine Klassenstruktur zu unterscheiden, um einen effektiven Verifikationsprozess zu ermöglichen.

Die Ergebnisse der Analyse werden dem Nutzer in einem GIS als Überlagerung des Orthophotos visualisiert, und zwar in Form der entsprechend dem Analyseergebnis eingefärbten Landnutzungspolygone. Die Einfärbung erfolgt nach dem intuitiv verständlichen *Ampel-Prinzip*, d.h. in den Farben *grün*, *gelb* und *rot*. Mit der Farbe *grün* sind jene Polygone markiert, deren Klassifikationsergebnisse mit den Nutzungsarten in der Datenbank konsistent sind. Die Farben *rot* und *gelb* weisen hingegen auf Widersprüche hin, wobei der Änderungshinweis bei *rot* aufgrund eines zugrundeliegenden höheren Konfidenzwertes der Klassifikationsentscheidung zuverlässiger ist als bei *gelb*. Der Schwellwert für die Differenzierung in die Klassen *rot* und *gelb* kann individuell an die jeweiligen Anforderungen angepasst werden. Es findet keine automatische Fortführung der Datenbestände statt. Der Algorithmus liefert lediglich Hinweise für den Bearbeiter an welcher Stelle die Landnutzung zu überprüfen ist. Darüber hinaus wird im Rahmen der Klassifikation die wahrscheinlichste neue Nutzungsart prädiziert, die als Vorschlag für die neue Klasse in den Bearbeitungsprozess einfließt. Dieser Vorschlag ist von einem Bearbeiter manuell zu prüfen und sofern zutreffend zu präzisieren bzw. andernfalls zu korrigieren. Im Rahmen des manuellen Fortführungsprozesses gilt es außerdem, die zugehörigen Nutzungsgrenzen manuell auf ihre Gültigkeit hin zu überprüfen und ggf. anzupassen.

Trotz des manuellen Nachbearbeitungsaufwandes führt diese Vorgehensweise insgesamt zu einer Arbeitsersparnis. Zum einem müssen die in grün dargestellten, verifizierten GIS-Objekte nicht überprüft werden, und zum anderen können im Falle von Änderungen automatisch generierte Vorschläge den Fortführungsprozess unterstützen.

Die Methodik zur Verifikation und Aktualisierung des Landnutzungsdatenbestandes steht nicht im Fokus dieser Arbeit und wird daher im Rahmen der Experimente in Kapitel 5 nicht evaluiert.

4.5. Implementierungsaspekte

Die Implementierung des hier vorgestellten Verfahrens erfolgte im Wesentlichen auf Grundlage der frei verfügbaren Bibliothek OpenCV [OpenCV, 2017] sowie einer am Institut für Photogrammetrie und Geoinformation der Leibniz Universität Hannover entwickelten Bibliothek für Graphische Modelle [Kosov et al., 2013]. Für den RF Klassifikator wird die OpenCV-Implementierung verwendet. Bei der Testfunktion handelt es sich um eine einfache Schwellwertfunktion, d.h. jeder zufällig erstellte Test verwendet exakt ein Merkmal. Für die Anzahl der zu testenden Merkmale N_f pro Knoten wird der Standardwert $N_f = \sqrt{D}$ verwendet, d.h. die Wurzel aus der Anzahl D aller verfügbaren Merkmale. Die Beurteilung der Güte der Aufteilung erfolgt auf Basis des Gini-Index [Breiman, 2001]. Die Knoten eines Entscheidungsbaumes werden zu einem Blattknoten erklärt, d.h. es werden keine weiteren Knoten mehr hinzugefügt, wenn die vorgegebene maximale Tiefe T_{max} des Baumes erreicht ist, die Samples innerhalb eines Knotens „rein“ sind oder die Anzahl der Samples pro Knoten eine vorgegebene Mindestanzahl $N_{T,min}$ unterschreitet. Es werden solange Entscheidungsbäume hinzugefügt bis die vorgegebene maximale Anzahl an Bäumen erreicht ist.

Bei der segmentbasierten Extraktion der bildbasierten Merkmale wird vorab eine Erosion mit einer Fenstergröße von 3×3 Pixeln auf die binäre Repräsentation der Segmente angewendet, um Mischpixel an den Rändern zu entfernen.

Die Prozessierung erfolgt sowohl im Training als auch im Rahmen der Klassifikation kachelweise. Die maximal mögliche Kachelgröße richtet sich nach der Größe der Superpixel sowie den zur Verfügung stehenden Rechnerkapazitäten. Eine überlappungsfreie Prozessierung der Kacheln würde dazu führen, dass speziell für GIS-Objekte an den Kachelrändern nicht alle Nachbarn im Rahmen der Klassifikation berücksichtigt werden können. Aus diesem Grund werden zusätzlich überlappende Bereiche in die Klassifikation einbezogen. Alle GIS-Objekte, die ganz oder teilweise in einem Umring von 256 Pixeln um die Kachel liegen, werden in das Training und die Klassifikation einbezogen. Angeschnittene GIS-Objekte werden nach maximal 512 Pixeln Abstand vom erweiterten Kachelrand abgeschnitten. Diese Parameter sind im Rahmen dieser Arbeit fest eingestellt, lassen sich jedoch prinzipiell an unterschiedliche Gegebenheiten (z.B. Kachelgrößen, Eingangsdaten, Rechnerkapazität) anpassen. Die Überlappung wird jedoch nicht über die Grenzen der Trainings- und Testgebiete hinweg angebracht, d.h. in den Randbereichen verfügen die GIS-Objekte nur über eingeschränkte Kontextinformation. Um zu vermeiden, dass GIS-Objekte, die sowohl in Trainings- als auch in Testgebieten liegen, zugleich zum Trainieren als auch zur Evaluierung verwendet werden, werden die betreffenden GIS-Objekte aus dem Training ausgeschlossen.

5. Experimente

5.1. Aufbau der Experimente

5.1.1. Datensätze

Im Rahmen der Experimente soll der entwickelte Ansatz anhand von Testdaten evaluiert werden. Die Testdaten dienen zum einen dem Anlernen des Klassifikators anhand von repräsentativen Trainingsgebieten und zum anderen dem Evaluieren der Ergebnisse anhand von Referenzdaten.

Hierfür stehen zwei Testgebiete zur Verfügung, ein vorwiegend urban geprägtes Testgebiet in Niedersachsen (Hameln) und ein ländlich geprägtes Testgebiet in Schleswig-Holstein (Schleswig) (siehe Abb. A.1 und A.2 im Anhang). Das Testgebiet *Hameln* hat eine Ausdehnung von 2 km x 6 km. Das Testgebiet *Schleswig* deckt mit einer Ausdehnung von 6 km x 6 km einen größeren Bereich ab. Beide Testgebiete weisen eine unterschiedliche Charakteristik auf. Das Testgebiet Hameln umfasst im Wesentlichen urbane Bereiche, in denen unterschiedlich strukturierte Typen von Bebauung vorkommen, wie z.B. eine dichte, geschlossene Bebauung im Innenstadtbereich, eine offene Bebauung mit Einfamilienhäusern in den vorgelagerten Wohngebieten sowie eine Bebauung mit großflächigen Hallen in Industriegebieten. Ländliche Strukturen, wie z.B. Wald, Acker- und Grünlandflächen, sind ebenfalls im Testgebiet enthalten, machen aber insgesamt einen kleineren Anteil aus. Die Charakteristik dieser Szene wird im besonderen Maße von dem Fluss *Weser* geprägt, der die Szene durchkreuzt. Demgegenüber weist das Testgebiet Schleswig überwiegend ländliche Strukturen auf, wie z.B. Acker- und Grünlandflächen sowie Waldgebiete unterschiedlicher Ausdehnung. Daneben befinden sich in diesem Testgebiet vereinzelt kleine Dörfer sowie eine Kleinstadt, die größtenteils durch eine offene Bebauung geprägt sind. Darüber hinaus zeichnet sich diese Szene durch eine hohe Zahl an Gewässerflächen aus, u.a. Seen, Teiche, Bäche und Teile einer Meeresfläche.

Für beide Testgebiete liegen hochauflösende Orthophotos (Bodenpixelgröße je 20 cm, 4 Kanäle), digitale Oberflächen- und Geländemodelle (unterschiedliche Rasterweiten) vor. Die originalen Luftbildaufnahmen, die Orthophotos und die DGM wurden von dem Landesamt für Geoinformation und Landesvermessung Niedersachsen (LGLN) und dem Landesamt für Vermessung und Geoinformation Schleswig-Holstein (LVermGeo SH) für die Durchführung des Projektes zur Verfügung gestellt. Die DOM wurden im Zuge eines Vorverarbeitungsschrittes mittels Bildzuordnung aus den originalen Luftbildaufnahmen abgeleitet. Bei den verwendeten DGM handelt es sich um einen Datenbestand, der bei den Landesvermessungsbehörden originär vorgehalten wird. Historisch bedingt ist dieser Datenbestand mit heterogenen Erfassungsmethoden entstanden (photogrammetrische Auswertung, Laserscanning, topographische Aufnahme etc.), wodurch die Qualität der Daten räumlich stark variiert. Bei den Orthophotos und den DGM handelt es sich um Standardprodukte der Landesvermessungsbehörden.

Die Spezifikationen zu den verwendeten Datensätzen können der Tabelle 5.1 entnommen werden. Die beiden Testdatensätze unterscheiden sich u.a. hinsichtlich ihres Erfassungszeitpunktes, der einen maßgeblichen Einfluss auf die Repräsentation der Vegetation in den Bildern hat. Die Luftbilder in Hameln wurden im Frühjahr erfasst. Zu diesem Zeitpunkt haben die Laubbäume typischerweise noch keine Blätter. Demgegenüber stammen die Daten für das Testgebiet Schleswig aus einer Luftbildbefliegung im Sommer, sodass beispielsweise Laubbäume eine dichte Belaubung aufweisen (vgl. Abb. A.1). Infolgedessen prägen sich Laub- und Nadelbäume in den Luftbildern in ähnlicher Weise aus, was deren Unterscheidung erschwert.

Daten	Parameter	Hameln	Schleswig
DOP	Erfassungszeitpunkt	Frühjahr 2010	Sommer 2013
	Bodenpixelgröße	20 cm	
	Anzahl Kanäle	4 Kanäle (3 Farbkanäle (RGB), nahes Infrarot)	
DOM	Rasterweite	50 cm	28 cm
DGM	Rasterweite	5 m	1 m

Tabelle 5.1.: Spezifikationen zu den Eingangsdaten der Testgebiete Hameln und Schleswig.

Die Klassifikation der Landnutzung erfolgt für Bildprimitive, die den GIS-Objekten eines räumlichen Landnutzungsdatenbestandes entsprechen, der in beiden Fällen das gesamte Testgebiet abdeckt. Hierbei handelt es sich um den Objektartenbereich „Tatsächliche Nutzung“ (TN) des Amtlichen Liegenschaftskatasterinformationssystems (ALKIS®) [Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland, 2008], der deutschlandweit, flächendeckend und überschneidungsfrei in einer hohen geometrischen und semantischen Auflösung vorgehalten wird. Die Objekte der „Tatsächlichen Nutzung“ beschreiben Flächen gleicher Landnutzung. Sie können sich aus mehreren (oder auch Teilen von) Flurstücken gleicher Nutzungsart zusammensetzen[1]. Die Polygone des Vektordatenbestandes geben die geometrischen Begrenzungen der Bildprimitive vor, die für diese Arbeit als fix angenommen werden.

Die Klassifikation der Bodenbedeckung erfolgt auf Basis von Superpixeln. Die Superpixel werden mit Hilfe des Segmentierungsverfahrens SLIC [Achanta et al., 2012] aus den Bilddaten extrahiert. Die verwendete SLIC-Implementierung basiert auf dreikanaligen Eingangsbildern. In Kapitel 5.2.2 werden verschiedene Arten von Eingangsdaten untersucht, zum einen die primäre Bildinformation (Grauwerte) und zum anderen die aus den Sensordaten abgeleiteten sekundären Bildinformationen nDOM, Intensität und NDVI. Der Verwendung speziell dieser drei Sekundärdaten liegt die Annahme zugrunde, dass sich die auf diese Weise extrahierten Superpixel besser an die Grenzen der Bodenbedeckungssegmente anpassen. Erfahrungsgemäß lassen sich die verschiedenen Bodenbedeckungsarten bereits gut anhand dieser Sekundärinformationen differenzieren. Folglich treten speziell an den Übergängen Höhensprünge oder große Differenzen im NDVI auf. Im Rahmen einer Segmentierung werden Pixel mit ähnlichen Eigenschaften zu Segmenten zusammengefasst und von Pixeln mit abweichenden Eigenschaften abgegrenzt, sodass diese Übergänge typischerweise die Grenzen der Segmente definieren. Die Größe und Kompaktheit der Superpixel werden dem Verfahren als Parameter übergeben und bleiben für eine Seg-

[1] Unveröffentlichte, interne Arbeitsanweisung der Vermessungs- und Katasterverwaltung Niedersachsen vom 16.02.2016: Erhebung der Topografie für den Nachweis im Liegenschaftskataster (Topografie-Handbuch).

mentierung konstant. Für die Experimente in Kapitel 5.2.2 werden Superpixel verschiedener Größen extrahiert, um den Einfluss der Größe der Superpixel auf das Klassifikationsergebnis zu untersuchen. Die Größe ist definiert als die Anzahl der Pixel innerhalb eines Superpixels. Im Rahmen dieser Arbeit werden Superpixel der Größe 2.500, 400 und 100 Pixel extrahiert, was bei der Auflösung der verwendeten Eingangsdaten einer Fläche von 100 m^2, 16 m^2 und 4 m^2 entspricht. Diese Größen wurden exemplarisch zur Repräsentation von drei verschiedenen Detailstufen gewählt. Die Kompaktheit variiert im Rahmen der Experimente in Kapitel 5.2.2 im Wertebereich [1,80].

Für das Training und die Evaluierung stehen Referenzdaten zur Verfügung. Die Referenzdaten für die Bodenbedeckung bestehen aus pixelweise manuell gelabelten Bildkacheln der Größe 200 m x 200 m. Für das Testgebiet Hameln bzw. Schleswig stehen 40 bzw. 37 solcher Kacheln zur Verfügung. Die Kacheln sind jeweils gleichmäßig über die Testgebiete verteilt und repräsentieren die wesentlichen charakteristischen Strukturen der Testgebiete. Das Referenzlabel eines Superpixels entspricht der am häufigsten unter den konstituierenden Pixeln vorkommenden Klasse. Die Zuweisung eines Referenzlabels nach der „Winner-takes-all"-Strategie führt zu Ungenauigkeiten in den Referenzdaten. Im Rahmen des Verfahrens werden diese Unsicherheiten berücksichtigt, indem zum Trainieren des Assoziations- und des paarweisen räumlichen Interaktionspotentials in der Bodenbedeckungsebene ausschließlich Superpixel herangezogen werden, bei denen mindestens 75 % aller Pixel zu einer Klasse gehören (vgl. Kapitel 4.2.4). Die Referenzdaten für die Klassifikation der Landnutzung bildet der ALKIS®-TN-Datenbestand, der für beide Testgebiete manuell sowohl im Hinblick auf die Semantik als auch die Geometrie korrigiert wurde.

5.1.2. Klassenstruktur

Die Klassifikation der Bodenbedeckung unterscheidet zwischen den 8 Klassen *Gebäude (Geb.)*, *Versiegelung (Vers.)*, *Boden (Bod.)*, *Gras*, *Baum*, *Wasser (Was.)*, *Fahrzeug (Fhzg.)* und *Sonstiges (Sonst.)*, die typischerweise in urbanen und ländlichen Umgebungen auftreten und in der geometrischen Auflösung der Bilddaten unterschieden werden können. Die Anzahl der Landnutzungsklassen variiert im Rahmen der Experimente zwischen minimal 4 und maximal 37 Klassen. Die Definition der Landnutzungsklassen erfolgt auf Grundlage des ALKIS®-Objektartenkatalogs [Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland, 2008]. Der ALKIS®-TN-Datenbestand unterscheidet 26 Objektarten, die sich in die vier übergeordneten Objektartengruppen *Siedlung*, *Verkehr*, *Vegetation* und *Gewässer* gliedern. Neben der Untergliederung in Objektarten erfolgt eine weitere Differenzierung auf Basis von Attributarten (z.B. Funktion, Vegetationsmerkmal, Zustand), welche die spezifische Nutzung innerhalb einer Objektart detailliert beschreiben. Insgesamt werden in der feinsten Auflösung ca. 245 verschiedene Nutzungsarten unterschieden, was einer sehr detaillierten Beschreibung der Landnutzung entspricht. Die ALKIS®-Kennung setzt sich der hierarchischen Struktur folgend aus der Kennung der Objektart sowie dem Bezeichner (z.B. *FKT* für Funktion) und dem Wert einer Attributart zusammen, wobei ein TN-Objekt mehrere Attributarten innerhalb einer Objektart belegen kann[2]. In den zur Verfügung stehenden Testdatensätzen ist pro TN-Objekt maximal eines der

[2]Unveröffentlichte, interne Arbeitsanweisung der Vermessungs- und Katasterverwaltung Niedersachsen vom 16.02.2016: Erhebung der Topografie für den Nachweis im Liegenschaftskataster (Topografie-Handbuch).

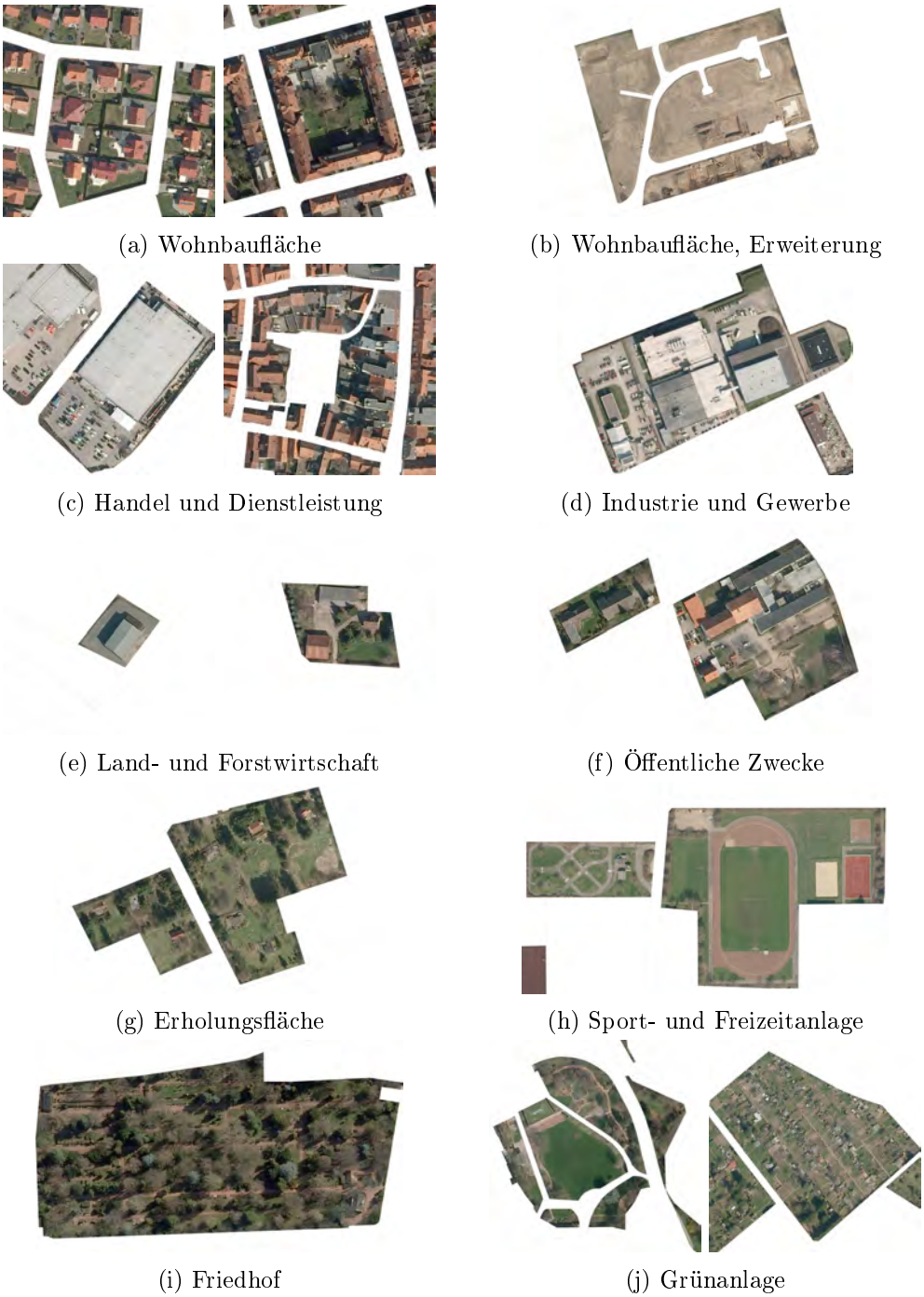
zuvor genannten Attributarten belegt, sodass eine Klasse in der feinsten Auflösung definiert ist als Kombination aus einer Objektart mit, sofern vorhanden, maximal einem Attributwert.

In der Objektartengruppe *Siedlung* sind alle bebauten und unbebauten Flächen enthalten, die für Wohn-, Wirtschafts- und Erholungszwecke genutzt werden. In der Abbildung 5.1 sind ausgewählte Beispiele für Nutzungsarten dieser Kategorie dargestellt, die einen guten Eindruck von dem hohen Detailgrad des TN-Datenbestandes vermitteln. Die bebauten und unbebauten Flächen, die dem Verkehr dienen, sind in der Objektartengruppe *Verkehr* zusammengefasst. Zur Kategorie *Vegetation* zählen alle Flächen außerhalb von Siedlungsgebieten, die entweder land- und forstwirtschaftlich genutzt oder durch einen natürlichen Bewuchs oder andere natürliche Landschaftselemente (z.B. Felsen) geprägt sind. Die Objektartengruppe *Gewässer* umfasst alle mit Wasser bedeckten Flächen der Erdoberfläche, die sich geometrisch i.d.R. durch ihre Uferlinie von benachbarten Nutzungen abgrenzen[3].

Der Detailgrad der Klassenstruktur, der zumindest theoretisch noch in den Testgebieten unterschieden werden kann, ist in erster Linie durch das Vorkommen der Klassen in den zur Verfügung stehenden Trainingsgebieten limitiert. In der Tabelle 5.2 ist die Klassenstruktur hierarchisch von der gröbsten Auflösung auf Ebene der Objektartengruppen bis hin zur feinsten Gliederungsstufe auf Ebene der Attributarten dargestellt, die den Experimenten in Kapitel 5.2.1 zugrunde liegt. Es werden nur Klassen bis zur dritten Hierarchieebene unterschieden, d.h. Objektartengruppe, Objektart und die erste Obergruppe der Attribute, die sich unter Umständen aus zahlreichen, sehr differenzierten Untergruppen zusammensetzt. Zudem sind in der Klassenstruktur nur jene Klassen enthalten, von denen mindestens fünf Objekte in den Testgebieten vorliegen. Diesem Grenzwert liegt die Annahme zugrunde, dass bei weniger als fünf Objekten der Klassifikator nicht korrekt gelernt und evaluiert werden kann. Bei weniger als fünf Objekten wird die zugehörige Klasse einer anderen, semantisch ähnlichen Klasse zugeordnet bzw. mit dieser zu einer gemeinsamen Klasse zusammengefasst. Beispiele hierfür sind die Zusammenfassung von *Sport- und Freizeitanlagen*, die in dem ALKIS®-Objektartenkatalog eigentlich separate Funktionen beschreiben, sowie die Verknüpfung der Objektarten *Heide* und *Moor* zu einer gemeinsamen Klasse (vgl. Tab. 5.2). Die Klassenstruktur für beide Testgebiete weicht voneinander ab, da einzelne Nutzungsarten in jeweils einem der Testgebiete nicht oder nur selten auftreten oder nicht bis zu der gleichen Hierarchiestufe differenziert werden. Beispielsweise wird die Objektart *Weg* im Testgebiet Schleswig – anders als in Hameln – mit Hilfe der Attribute nicht weiter differenziert in die Nutzungsarten *Fahrweg* und *Fuß- und Radweg* (vgl. Tab. 5.2).

In der Tabelle 5.2 sind ergänzend die Anzahl der GIS-Objekte pro Klasse getrennt für die beiden Testgebiete angegeben. Die größere Ausdehnung des Testgebietes Schleswig spiegelt sich in einer höheren Gesamtanzahl von 3739 Objekten gegenüber 2812 Objekten im Testgebiet Hameln wider. Die Verteilung der Objekte gibt einen Hinweis auf die Charakteristik der Testgebiete. Beispielsweise prägt sich die überwiegend ländliche Charakteristik im Testgebiet Schleswig in einer hohen Anzahl an Vegetationsflächen aus, die insgesamt ca. 24 % aller Objekte ausmachen. Im Gegensatz dazu zählen im Testgebiet Hameln nur ca. 8 % aller Objekte zu dieser Kategorie, was die überwiegend urbane Struktur des Testgebietes belegt. Zu den größeren Unterschieden zählen zudem die höhere Anzahl an Gewässerflächen in Schleswig. Darüber hinaus sind bestimmte Klassen, wie z.B. *Heide*, *Moor*, *Sumpf* und

[3] Unveröffentlichte, interne Arbeitsanweisung der Vermessungs- und Katasterverwaltung Niedersachsen vom 16.02.2016: Erhebung der Topografie für den Nachweis im Liegenschaftskataster (Topografie-Handbuch).



(a) Wohnbaufläche (b) Wohnbaufläche, Erweiterung

(c) Handel und Dienstleistung (d) Industrie und Gewerbe

(e) Land- und Forstwirtschaft (f) Öffentliche Zwecke

(g) Erholungsfläche (h) Sport- und Freizeitanlage

(i) Friedhof (j) Grünanlage

Abbildung 5.1.: Ausgewählte Beispiele für TN-Objekte der Kategorie *Siedlung*. Darstellung je Spalte in einheitlichem Maßstab (mit Ausnahme von (d)). Quelle der Orthophotos: Auszug aus den Geobasisdaten der Niedersächsischen Vermessungs- und Katasterverwaltung, ©2010 LGLN.

Speicherbecken, in Schleswig jedoch nicht in Hameln vertreten, wohingegen andere Nutzungsarten, wie z.B. *Fluss* und *Bahnverkehr*, das Testgebiet Hameln in wesentlichem Maße prägen, jedoch in Schleswig nicht bzw. nur selten auftreten. Daneben finden sich weitere Unterschiede, beispielsweise im Hinblick auf die Nutzungsarten *Mischnutzung mit Wohnen* oder *Fußgängerzone*, die nicht in der Charakteristik der Testgebiete, sondern vielmehr in unterschiedlichen Vorgehensweisen und Regelungen zur Erfassung

Gruppe	Objektart	Attribut	Kurzform	Anzahl Obj. HM	SL
Siedlung	Wohnbaufläche	Wohnbaufläche	Wbfl.	477	681
		Erweiterung	Erw.	25	60
	Industrie- und Gewerbefläche	Industrie und Gewerbe	Ind.	86	32
		Handel und Dienstleistung	Dnstl.	173	126
		Versorgungsanlage	Vers.	53	53
		Entsorgungsanlage	Ents.		6
	Fläche gemischter Nutzung	Mischnutzung mit Wohnen	Mischn.	-	113
		Land- und Forstwirtschaft	Landw.	11	43
	Fläche besond. funkt. Prägung	Öffentliche Zwecke	Öffntl.	107	127
		Historische Anlage	Hist.		12
	Sport-, Freizeit- & Erholungsfl.	Gebäude- und Freifläche	Sprtgeb.	6	20
		Sport- und Freizeitanlage	Sprtanl.	6	26
		Erholungsfläche	Erhol.	23	308
		Grünanlage	Grnanl.	219	
	Friedhof		Frhf.	6	8
Verkehr	Straßenverkehr	Straße	Str.	426	578
		Verkehrsbegleitfläche	Strbgl.	148	69
		Fußgängerzone	Fußgz.	19	-
	Weg	Fahrweg	Fahrw.	268	332
		Fuß- und Radweg	Fußw.	328	
	Platz	Parkplatz	Parkpl.	31	58
		Platz (div. Funktionen)	Pl.	9	13
	Bahnverkehr	Bahntrasse	Bhntr.	23	zu Str.
		Gebäude- und Freifläche	Bhngeb.	10	-
		Verkehrsbegleitfläche	Bhnbgl.	59	-
	Schiffsverkehr	Landfläche	Schiffv.	5	5
Vegetation	Landwirtschaft	Grünland	Grünl.	47	334
		Ackerland	Ackerl.	61	154
		Gartenland	Gartenl.		6
		Brachland	Brachl.	10	7
	Wald	Laubholz	Laubw.	27	99
		Nadelholz	Nadelw.	15	31
		Laub- und Nadelholz	Mischw.		98
	Gehölz		Gehlz.	43	78
	Heide, Moor		Moor	-	11
	Sumpf		Sumpf	-	38
	Unland / Vegetationslose Fl.	Unland	Unland	-	34
		Gewässerbegleitfläche	Gewbgl.	37	5
Gewässer	Fließgewässer	Fluss	Fluss	10	-
		Graben	Graben	8	69
		Bach	Bach	31	18
	Stehendes Gewässer	See, Meer	See	5	12
		Teich	Teich		67
		Speicherbecken	Spbeck.	-	8
			Gesamtanzahl	2812	3739

Tabelle 5.2.: Hierarchische Klassenstruktur gemäß den Spezifikationen des ALKIS®-Objektartenkatalogs für den Objektartenbereich „Tatsächliche Nutzung“ und die Anzahl der GIS-Objekte pro Klasse in den Testgebieten Hameln (HM) und Schleswig (SL). Die Zusammenfassung von Nutzungsarten ist gekennzeichnet durch verschmolzene Zellen oder einen textlichen Hinweis auf die Zuordnung.

der TN in den Bundesländern begründet liegen.

Im Rahmen der experimentellen Evaluation in Kapitel 5.2.1 wird die maximal mögliche semantische Auflösung bestimmt, d.h. die feinste Klassenstruktur, die sich in den Testgebieten noch unterscheiden lässt. Diese Klassenstruktur wird in den nachfolgenden Experimenten in den Kapiteln 5.2.2 bis 5.2.5 festgehalten.

5.1.3. Merkmale

In einem Vorverarbeitungsschritt werden aus den zur Verfügung stehenden Eingangsdaten verschiedene Arten von Merkmalen abgeleitet. Diese lassen sich in zwei Gruppen unterteilen: bildbasierte, dreidimensionale und geometrische Merkmale, die während der Inferenz konstant bleiben, und Kontextmerkmale, die in jeder Iteration des Inferenzprozesses aktualisiert werden. Die Kontextmerkmale leiten sich aus den Zwischenlösungen ab, die in jedem Schritt der iterativen Inferenzprozedur bestimmt werden. In den Zwischenlösungen sind je Bildprimitiv das wahrscheinlichste Klassenlabel, d.h. das Klassenlabel mit dem maximalen Konfidenzwert, und die Konfidenzwerte für alle Klassen enthalten. Die zur Bodenbedeckungs- und Landnutzungsklassifikation verwendeten Merkmale beziehen sich auf unterschiedliche Bildprimitive, d.h. GIS-Objekte in der Landnutzungs- und Superpixel in der Bodenbedeckungsebene. In der nachfolgenden Beschreibung bezieht sich der Begriff *Segment* auf beide Arten von Bildprimitiven.

Für die Bildprimitive in beiden Ebenen wird ein identischer Satz an bildbasierten, dreidimensionalen und geometrischen Merkmalen extrahiert, der in der Tabelle 5.3 aufgelistet ist. Die theoretischen Grundlagen zu den einzelnen Merkmalen sind in den Kapiteln 3.2.1, 3.2.2 und 3.2.3 angegeben. Aus der Gruppe der spektralen Merkmale werden die Spektralwerte der vier Farbkanäle, der NDVI, der Farbton, die Sättigung und die Intensität extrahiert. Die Berechnung erfolgt zunächst separat für jedes Pixel innerhalb eines Segments. Um Merkmalswerte pro Segment zu erhalten, werden jeweils der Mittelwert, die Standardabweichung und der minimale und maximale Wert der pixelbasierten Werte innerhalb eines Segments berechnet. Das pro Segment aufgestellte Histogramm der Gradientenorientierungen (HOG) teilt die pixelweise berechneten Gradientenrichtungen im Intervall $[0°, 180°]$ in 30 Klassen ein, was einer Klassenbreite von 6° entspricht. Das Histogramm bildet die Grundlage für die Ableitung von 13 verschiedenen Merkmalen, u.a. das Minimum und Maximum sowie das höchste Nebenminimum und -maximum (hier bezeichnet als erstes und zweites Minimum bzw. Maximum), Mittelwert, Standardabweichung sowie verschiedene Verhältniszahlen der minimalen und maximalen Werte. Weitere Merkmale bilden die Anzahl der Klassen im Histogramm (nSignifikant), deren Einträge die Summe aus Mittelwert und Standardabweichung überschreiten und damit eine besonders starke Ausprägung im Histogramm darstellen, sowie die Winkeldistanz zwischen dem ersten und zweiten Maximum (nOriDiff). Bei den Texturmerkmalen handelt es sich um die Haralick Merkmale Energie, Kontrast, Homogenität und Korrelation [Haralick et al., 1973]. Deren Berechnung basiert auf den vier GLCM $P_{1,0°}$, $P_{1,45°}$, $P_{1,90°}$, $P_{1,135°}$, die das gemeinsame Auftreten der Intensitätswerte an Pixeln in einem Abstand $\Delta = 1$ in vier unabhängigen Richtungen der Achternachbarschaft ($\alpha = 0°, 45°, 90°, 135°$) beschreiben. Die Merkmalswerte werden pro GLCM berechnet und anschließend gemittelt, um richtungsunabhängige Merkmale zu erhalten. Die dreidimensionalen Merkmale umfassen die statistischen

Gruppe	Grundlage / Kanal	Merkmal
Spektral	Spektralwert 'Rot'	Mw., Std., Min., Max.
	Spektralwert 'Grün'	Mw., Std., Min., Max.
	Spektralwert 'Blau'	Mw., Std., Min., Max.
	Spektralwert 'Nahes Infrarot'	Mw., Std., Min., Max.
	NDVI	Mw., Std., Min., Max.
	Farbton	Mw., Std., Min., Max.
	Sättigung	Mw., Std., Min., Max.
	Intensität	Mw., Std., Min., Max.
Textur	GLCM	Haralick Energie
		Haralick Kontrast
		Haralick Homogenität
		Haralick Korrelation
Struktur	HOG	Mw., Std., Min.(1.,2.), Max.(1.,2.), V.(1.Min/1.Max), V.(2.Min/1.Max), V.(1.Min/2.Max), V.(2.Min/2.Max), V.(2.Max/1.Max), nSignifikant, nOriDiff
3D	nDOM	Mw., Std., Min., Max.
Geometrie	Segment	Fläche
		Umfang
		Kompaktheit
		Formindex
		Fraktale Dimension
		Seitenverhältnis
		Länglichkeit
		Elongation
		Füllungsgrad
		Polarer Abstand

Tabelle 5.3.: Gesamtheit an spektralen, texturellen, strukturellen, dreidimensionalen (3D) und geometrischen Merkmalen, die pro Segment extrahiert werden. Die spektralen, strukturellen und 3D-Merkmale sind die statistischen Parameter Mittelwert (Mw.), Standardabweichung (Std.), Minimum (Min.) und Maximum (Max.) abgeleitet aus den pixelbasierten Merkmalswerten innerhalb eines Segments. Der Satz an strukturellen Merkmalen umfasst zusätzlich die ersten (1.) und zweiten (2.) minimalen und maximalen Werte sowie aus diesen Werten abgeleitete Verhältniszahlen (V.), die Anzahl der Klassen (nSignifikant), deren Histogrammeinträge die Summe aus Mw. und Std. überschreiten, sowie die Winkeldistanz (nOriDiff) zwischen dem 1. und 2. Maximum.

Parameter Mittelwert, Standardabweichung, Minimum und Maximum der Höhe über Grund (nDOM) aller Pixel innerhalb des Segments. Die geometrischen Merkmale berechnen sich aus der polygonalen Repräsentation der Segmente. Die Polygone der Landnutzungsobjekte gibt der TN-Datenbestand vor. Die Berechnung der geometrischen Merkmale für die Superpixel basiert auf deren Kontur. Zu den extrahierten geometrischen Merkmalen zählen Fläche, Umfang, Kompaktheit, Formindex, fraktale Dimension, Füllungsgrad, polarer Abstand sowie Seitenverhältnis, Länglichkeit und Elongation des minimal umschließenden Rechtecks.

Für die Klassifikation der Bodenbedeckung und der Landnutzung wird ein unterschiedlicher Satz an Kontextmerkmalen extrahiert. Die theoretischen Grundlagen zu den extrahierten Merkmalen sind in Kapitel 3.2.4 für die aus der Literatur übernommenen und in Kapitel 4.3 für die im Rahmen der vorliegenden Arbeit entwickelten Merkmale angegeben. Für die Klassifikation der Bodenbedeckung wird aus der Gruppe der räumlichen Metriken der relative und der mit der Konfidenz gewichtete relative Flächenanteil der Landnutzungsart an der Gesamtfläche eines Superpixels berechnet. Darüber hinaus bildet der Mittelwert der quadratischen minimalen Distanzen der Konturpunkte des Superpixels von den Nutzungsgrenzen ein weiteres Merkmal. Der jeweils berechnete minimale Abstand ist normalisiert mit der maximalen Ausdehnung des Landnutzungsobjekts. Für die Klassifikation der Landnutzung werden neben dem relativen und dem Konfidenz-gewichteten relativen Flächenanteil der Bodenbedeckungsarten an der Gesamtfläche des Landnutzungsobjekts weitere Arten von Merkmalen extrahiert. Aus der Gruppe der räumlichen Metriken zählen hierzu die zentralen Momente erster und zweiter Ordnung zur Charakterisierung der Verteilung der Bodenbedeckungsarten innerhalb eines Landnutzungsobjekts. Aus der Gruppe der graphenbasierten Merkmale werden die normalisierten Matrixeinträge der auf Pixelniveau aufgestellten Co-Occurrence Matrix der Bodenbedeckungsarten innerhalb des Landnutzungsobjekts als Kontextmerkmal für die Klassifikation verwendet.

Die Gesamtheit der extrahierten Merkmale bilden einen Pool an Merkmalen, von denen alle oder nur eine Teilmenge zur Klassifikation verwendet werden. Insgesamt umfasst dieser Pool an Merkmalen für die Klassifikation der Bodenbedeckung und der Landnutzung 63 bildbasierte, dreidimensionale und geometrische Merkmale. Die Anzahl der Kontextmerkmale pro Ebene variiert im Rahmen der Experimente in Abhängigkeit der Anzahl der Klassen, die in der jeweils anderen Ebene unterschieden werden. Insgesamt werden auf Ebene der Landnutzungsklassifikation 76 Merkmale extrahiert. Auf der Ebene der Bodenbedeckungsklassifikation variiert die Anzahl der Kontextmerkmale zwischen minimal 9 und maximal 75 Merkmalen, je nachdem wie viele Landnutzungsklasssen aktuell unterschieden werden. Die Merkmalswerte werden auf das Intervall $[0, 1]$ skaliert und anschließend in einem Merkmalsvektor pro Potentialterm zusammengeführt. Eine Möglichkeit zur Normalisierung der Merkmalswerte auf das Intervall $[0, 1]$ besteht darin, die minimalen und maximalen Merkmalswerte auf die unteren und oberen Intervallgrenzen abzubilden und dazwischen linear zu transformieren. Dies hat allerdings zur Folge, dass gegebenenfalls Ausreißer dazu führen, dass nicht der gesamte Wertebereich ausgenutzt wird. Diesem Effekt wird vorgebeugt, indem ein gewisser Prozentsatz am unteren und oberen Rand des Histogramms der Merkmalswerte jeweils auf die unteren und oberen Grenzwerte abgebildet wird. Im Rahmen dieser Arbeit definiert das Quantil ε die obere Grenze, d.h. für ein Quantil von 99 % nehmen 99 % der Daten einen kleineren Wert an und 1 % der Daten werden auf die obere Grenze abgebildet. Aufgrund der symmetrischen Definition des Intervalls folgert für die untere Grenze ein Prozentsatz von $1 - \varepsilon$, d.h. für $\varepsilon = 99\,\%$ nehmen 1 % der Daten einen kleineren Wert an und werden auf die untere Grenze abgebildet. Ein Quantil von 100 % entspricht dabei dem Fall, dass der gesamte Wertebereich auf das Intervall $[0, 1]$ skaliert wird. Der Vektor der Interaktionsmerkmale $\boldsymbol{\mu}_{ij}^{o}(\mathbf{x})$ entspricht in jeder Ebene $o \in \{c, u\}$ entweder den aneinandergehängten Merkmalsvektoren der Knoten n_i^o und n_j^o, die durch die Kante e_{ij} verbunden sind, oder den elementweisen Differenzen dieser Merkmalsvektoren. Im Rahmen der Experimente in Kapitel 5.2.3 wird der optimale Wert für das Quantil zur Normalisierung der Merkmalswerte sowie die optimale Art der Zusammensetzung der Interaktionsmerkmale

gesondert für jede Klassifikationsaufgabe bestimmt. Die Merkmale sind teilweise hochgradig korreliert, sodass die Gesamtheit der Merkmale auf die wesentlichen Merkmale zu reduzieren ist. Die Selektion von Merkmalen erfolgt in Kapitel 5.2.3 anhand eines Relevanzmaßes, das von dem RF Klassifikator bereitgestellt wird (vgl. Kapitel 3.3).

5.1.4. Durchführung der Experimente

Die Zielsetzung der Experimente besteht darin, die Leistungsfähigkeit des entwickelten Ansatzes zur simultanen kontextbasierten Klassifikation der Bodenbedeckung und der Landnutzung zu evaluieren. Im Rahmen der Untersuchungen gilt es, die in Kapitel 1.2 aufgestellten Annahmen mit Experimenten auf ihre Gültigkeit zu überprüfen. Darüber hinaus werden verschiedene Einflussgrößen untersucht, die sich in mehrere Gruppen unterteilen lassen:

- Klassenstruktur
- Parameter der Superpixelsegmentierung
- Merkmale
- Modell- und Inferenzparameter

Die Untersuchung der verschiedenen Einflussgrößen und die Bestimmung optimaler Parameterwerte erfolgt empirisch auf Basis einer Sensitivitätsanalyse, im Rahmen derer die Trainingsdaten abwechselnd zum Trainieren und Evaluieren verwendet werden. Zu diesem Zweck werden die Testdaten in Bearbeitungsblöcke der Ausdehnung 1 km x 1 km unterteilt. In Hameln führt dies zu 12 und in Schleswig zu insgesamt 36 Blöcken. Aus diesen Blöcken werden Gruppen gebildet, wobei jede Gruppe im Rahmen der empirischen Sensitivitätsanalyse einmal zur Evaluation beiträgt, währenddessen die jeweils verbleibenden Gruppen zum Training verwendet werden. In Hameln setzt sich jede Gruppe aus nur einem Block zusammen, wodurch sich 12 Gruppen und damit 12 Testdurchläufe ergeben. In Schleswig werden jeweils 18 Blöcke zu einer Gruppe zusammengefasst, sodass insgesamt 2 Gruppen entstehen. In dem Testgebiet Hameln wird jeder Block individuell prozessiert. Auf diese Weise wird sichergestellt, dass in jedem Testdurchlauf möglichst viele, d.h. die verbleibenden 11 Blöcke, für das Training zur Verfügung stehen. Die Notwendigkeit resultiert daraus, dass in diesem Testgebiet insgesamt nur eine geringe Anzahl an Trainingsbeispielen für die Landnutzung zur Verfügung steht. Die im Vergleich zum Testgebiet Schleswig geringere Anzahl an Trainingsbeispielen resultiert aus der flächenmäßig kleineren Ausdehnung des Testgebietes Hameln. Obwohl in beiden Testgebieten flächendeckend Trainingsdaten vorliegen, führt die flächenmäßig große Ausdehnung der TN-Objekte zu einer vergleichsweise geringen Anzahl an Trainingsbeispielen pro Testgebiet.

Um die Vorteile der Integration von Kontext zu evaluieren, werden die Ergebnisse mit denen von zwei weiteren Verfahren verglichen. Das erste Vergleichsverfahren bildet die unabhängige Klassifikation mit RF, die auf den Assoziationspotentialen des in Kapitel 4.2 beschriebenen graphischen Modells basiert ohne die Berücksichtigung von Kontext (d.h. $\omega^2 = \omega^4 = \omega^{5,c\to u} = \omega^{5,u\to c} = 0$). Bei dem zweiten Vergleichsverfahren handelt es sich um eine Klassifikation unter Verwendung von CRF, die nur den räumlichen Kontext modelliert (d.h. $\omega^{5,c\to u} = \omega^{5,u\to c} = 0$). Zur Evaluation des iterativen Inferenzalgorithmus werden die Ergebnisse des Verfahrens bei einmaliger (d.h. $n_{It} = 1$) und fünfmaliger

(d.h. $n_{It} = 5$) Iteration der Inferenzprozedur verglichen. Mit der einmaligen Integration von Kontext entspricht das Verfahren einer Variante des zweistufigen Ansatzes zur Klassifikation der Landnutzung, wie es in Albert et al. [2014a] vorgestellt wurde. Bei dem zweistufigen Ansatz sind beide Klassifikationsaufgaben ebenfalls als CRF modelliert. Im Gegensatz zu den Arbeiten in Albert et al. [2014a] wird jedoch Kontextinformation in beide Richtungen übertragen (d.h. $\omega^{5,c\rightarrow u} = \omega^{5,u\rightarrow c} = 1$) und zudem erfolgt die Klassifikation der Bodenbedeckung auf der Basis von Superpixeln anstelle von Pixeln, um die Vergleichbarkeit der Ergebnisse zu gewährleisten. Darüber hinaus werden die Kontextmerkmale in einem separaten Potential modelliert und bilden nicht wie in [Albert et al., 2014a] gemeinsam mit den bildbasierten, dreidimensionalen und geometrischen Merkmalen die Grundlage für die Bestimmung des Assoziationspotentials.

Sofern nicht explizit andere Parameter angegeben sind, werden im Rahmen der experimentellen Evaluation die in der Tabelle 5.4 aufgelisteten Standardparameter verwendet. Die Anzahl $N_{T,max}$ der

Parameter	Bodenbedeckung			Landnutzung		
Random Forests	AP	pIP	hoP	AP	pIP	hoP
$N_{T,max}$	10.000	10.000	10.000	1.000	1.000	1.000
T_{max}	25	25	25	25	25	25
$N_{T,min}$	5	5	5	5	5	5
B	150	150	150	150	150	150
Inferenz						
ω^i	1,0	1,0	1,0	1,0	1,0	1,0
n_{It}	5					
n_{LBP}	5					

Tabelle 5.4.: Standardparameter für die RF Klassifikatoren der Assoziationspotentiale (AP), der paarweisen räumlichen Interaktionspotentiale (pIP) und der aufgespaltenen Potentialterme des semantischen Potentials höherer Ordnung (hoP) in der Bodenbedeckungs- und Landnutzungsebene sowie die Parameter für die Inferenzprozedur. Die Parameter der RF Klassifikatoren sind die maximale Anzahl $N_{T,max}$ der Trainingssamples pro Klasse, die maximale Tiefe T_{max} der Bäume, die Mindestanzahl $N_{T,min}$ der Samples für nichtterminierende Knoten und die Anzahl B der Bäume. Die Parameter der Inferenzprozedur sind die Parameter ω^i pro Potentialterm, die Anzahl n_{LBP} der Iterationen in jedem LBP-Teilschritt und die Anzahl n_{It} der Iterationen im Rahmen der Inferenzprozedur.

Trainingsbeispiele pro Klasse variiert für die Potentialterme in der Bodenbedeckungs- und Landnutzungsebene, weil dieser Parameter an die insgesamt zur Verfügung stehende Anzahl an Trainingsdaten anzupassen ist, die für die Landnutzung geringer ist als für die Bodenbedeckung. Dieser Unterschied liegt vor allem in der großen räumlichen Ausdehnung der Landnutzungsobjekte begründet, aus der trotz flächendeckender Verfügbarkeit eine insgesamt kleinere Anzahl an Trainingsbeispielen resultiert. Die Ausprägung der Superpixel entspricht in den Experimenten, sofern nichts anderes angegeben ist, standardmäßig einer Größe von 400 Pixeln und einer Kompaktheit von 20 und basiert auf den in Kapitel 5.1 genannten sekundären Bildinformationen.

Die Bewertung der Klassifikationsergebnisse erfolgt auf Grundlage der Konfusionsmatrix und daraus abgeleiteter Qualitätsparameter [Foody, 2002]. Zur Aufstellung der Konfusionsmatrix werden die Klassifikationsergebnisse separat für jede Ebene mit den Referenzdaten verglichen. Die quantitative

Qualitätsanalyse der Klassifikationsergebnisse in Kapitel 5.2 gründet sich dabei auf zwei verschiedene Konfusionsmatrizen, wovon eine auf einem Vergleich der Klassifikationsergebnisse mit den Referenzdaten auf Grundlage der Bildprimitive (d.h. Superpixel bzw. GIS-Objekte) und die andere auf Grundlage der Pixel basiert. Für den Vergleich auf Ebene der Superpixel werden anders als beim Training auch solche Superpixel für den Vergleich herangezogen, die zu weniger als 75 % konsistent sind. Die Konfusionsmatrix bildet die Grundlage für die Ableitung verschiedener Genauigkeitsmaße, die der quantitativen Qualitätsanalyse der Ergebnisse dienen. Hierbei handelt es sich um die Maße Gesamtgenauigkeit, Cohen's Kappa-Index [Foody, 2002], sowie um Korrektheit, Vollständigkeit und Qualität pro Klasse [Heipke et al., 1997; Rutzinger et al., 2009]. Gegeben eine Konfusionsmatrix $\mathbf{C}$, in der die Klassen aus der Klassifikation in den Zeilen und die Referenzlabels in den Spalten aufgetragen sind, ergeben sich die Genauigkeitsmaße zu:

Gesamtgenauigkeit [%]
$$G = 100 \cdot \frac{\sum_i C_{ii}}{\sum_i \sum_j C_{ij}} \tag{5.1}$$

Kappa-Index [%]
$$\kappa = 100 \cdot \frac{p_0 - p_c}{1 - p_c}, \tag{5.2}$$

$$\text{mit } p_0 = \frac{\sum_i C_{ii}}{\sum_i \sum_j C_{ij}} = G, \tag{5.3}$$

$$\text{und } p_c = \frac{\sum_i \left(\sum_j C_{ji} \cdot \sum_j C_{ij} \right)}{\left(\sum_i \sum_j C_{ij} \right)^2}. \tag{5.4}$$

Korrektheit [%]
$$K_i = 100 \cdot \frac{C_{ii}}{\sum_j C_{ji}} \tag{5.5}$$

Vollständigkeit [%]
$$V_i = 100 \cdot \frac{C_{ii}}{\sum_j C_{ij}} \tag{5.6}$$

Qualität [%]
$$Q_i = 100 \cdot \frac{K_i \cdot V_i}{K_i + V_i - K_i \cdot V_i} \tag{5.7}$$

In den Gleichungen 5.1 bis 5.7 bezeichnet C_{ij} den Eintrag in der Konfusionsmatrix in der Zeile i und der Spalte j. Die Gesamtgenauigkeit beschreibt den prozentualen Anteil aller korrekten Zuordnungen an der Gesamtanzahl der klassifizierten Bildprimitive. Dieses Maß hat den Nachteil, dass es nur einen Durchschnittswert über alle Klassen berechnet und sich daher ggf. schlechte Ergebnisse für Klassen mit wenigen Samples nicht auf das Gesamtmaß auswirken. Dieses Problem adressiert der Cohen's Kappa-Index, der sich gegenüber kleinen Klassen sensitiver verhält. In dessen Berechnung fließen sowohl die tatsächliche Übereinstimmung p_0 zwischen dem Klassifikationsergebnis und der Referenz ein (entsprechend der Gesamtgenauigkeit) als auch die erwartete Übereinstimmung p_c, die sich aus dem Produkt der Zeilen- und Spaltensummen der Konfusionsmatrix über alle Klassen berechnet. Im Allgemeinen sind die Gesamtmaße allein nicht ausreichend, um die Qualität des Klassifikationsergebnisses differenziert zu analysieren. So liefern sie beispielsweise keine Information zur Verteilung der Fehler über die Klassen. Für eine tiefergehende Analyse bieten sich vielmehr Qualitätsmaße an, die pro Klasse aus der Konfusionsmatrix abgeleitet werden. Hierzu zählen die Korrektheit und Vollständigkeit, die in der Fernerkundung auch als Genauigkeit aus Sicht des Nutzers (*user's accuracy*) bzw. des Herstellers (*producer's accuracy*) bezeichnet werden, sowie die Qualität. Die Korrektheit quantifiziert die Anzahl

der Bildprimitive einer Klasse im Klassifikationsergebnis, die auch tatsächlich dieser Klasse entsprechen. Demgegenüber quantifiziert die Vollständigkeit die Anzahl der Bildprimitive einer Klasse in den Referenzdaten, die korrekt erkannt wurden. Ein gutes Klassifikationsergebnis weist sowohl eine hohe Vollständigkeit als auch Korrektheit auf. Die Qualität stellt ein Maß zur Abwägung von Vollständigkeit und Korrektheit dar, das definitionsgemäß einen geringeren Wert annimmt als die Eingangswerte. Die Qualität bietet sich insbesondere für das Ranking von Ergebnissen in Bezug auf eine Klasse an. Für eine tiefergehende Analyse der Klassifikationsergebnisse sollten jedoch zusätzlich die Genauigkeitsmaße Vollständigkeit und Korrektheit betrachtet werden.

Die Angaben zur Rechenzeit beziehen sich jeweils auf einen Intel(R) Core(TM) i5-2400 Prozessor mit 16 GB RAM und 4 Prozessoren. Bei der Angabe von Rechenzeiten ist grundsätzlich zu beachten, dass die Zeiten in Abhängigkeit der verwendeten Hard- und Softwareumgebung variieren können. Darüber hinaus wurde das entwickelte Programm nicht konsequent hinsichtlich der Laufzeit optimiert, z.B. sind in der aktuellen Version lediglich das Training und die Berechnung der einzelnen Potentiale sowie Teile des Inferenzalgorithmus parallelisiert. Das Programm bietet Potential für weitere Optimierungen, mit denen sich die Laufzeit voraussichtlich deutlich reduzieren ließe. Folglich sind die angegebenen Rechenzeiten nur für einen relativen Vergleich der Ergebnisse geeignet. Die Laufzeitunterschiede resultieren in erster Linie aus der unterschiedlichen Komplexität der graphischen Modelle. Die Angaben zur Rechenzeit beziehen sich jeweils auf eine Prozessierungseinheit, d.h. einen Block der Größe 1 km x 1 km.

5.2. Evaluation

5.2.1. Semantische Auflösung der Klassenstruktur

5.2.1.1. Zielsetzung

Die nachfolgende Untersuchung nimmt sich der semantischen Auflösung der Klassenstruktur an. Die Klassenstruktur formt gemeinsam mit der Ausprägung der Bildprimitive sowie den Merkmalen den Rahmen für jede zu lösende Klassifikationsaufgabe. Die Untersuchung zur semantischen Auflösung der Klassenstruktur erfolgt an dieser Stelle ausschließlich für die Landnutzungsklassifikation. Die Klassenstruktur der Bodenbedeckungen wird von der Untersuchung ausgenommen, da die wesentlichen, landschaftsprägenden Bodenbedeckungsarten im urbanen und ländlichen Umfeld enthalten sind und eine feinere Unterscheidung für die gestellte Aufgabe keinen Mehrwert bietet.

Um einen detaillierten, räumlichen Landnutzungsdatenbestand effizient im Hinblick auf dessen Richtigkeit verifizieren zu können, gilt es im Rahmen der Landnutzungsklassifikation eine möglichst feine Klassenstruktur zu unterscheiden, die sich möglichst nah an der des Landnutzungsdatenbestandes orientiert. Einer sehr fein untergliederten Klassenstruktur steht jedoch entgegen, dass die Klassifikationsgenauigkeit bei zunehmender Verfeinerung der Klassenstruktur sinkt und vermehrt Fehlzuweisungen zwischen den Klassen auftreten. Bei der Entwicklung des Ansatzes wurde implizit die Annahme getroffen, dass Kontext die Unterscheidung einer detaillierten Klassenstruktur unterstützt. Daher ist zu erwarten, dass der entwickelte Ansatz in der Lage ist, eine hohe semantische Auflösung zu unter-

scheiden. Zur Überprüfung dieser Annahme gilt es festzustellen, in welcher Detailstufe Nutzungsarten mit Hilfe des entwickelten Ansatzes noch unterschieden werden können, und zu bewerten, ob dieser Detailgrad einer hohen semantischen Auflösung entspricht. Diesen Fragestellungen nimmt sich die nachfolgende Untersuchung an.

5.2.1.2. Strategie

Zur Bestimmung des maximal zu erzielenden Detailgrads wird die Klassenstruktur je Objektartengruppe zunächst schrittweise hierarchisch verfeinert, bis die feinste Auflösung erreicht ist. Der Klassifikation jedes Schrittes liegt demnach eine feinere Klassenstruktur zugrunde als im vorherigen Schritt, d.h. mit jedem Schritt nimmt die Anzahl der Klassen, die es im Rahmen der Klassifikation zu bestimmen gilt, zu. Die Verfeinerung folgt der hierarchischen Struktur des zugrundeliegenden Objektartenkatalogs. Die erzielten Klassifikationsergebnisse werden einer detaillierten, empirischen Untersuchung unterzogen, die insbesondere Konfusionen mit anderen Klassen analysiert. Im Anschluss an den Verfeinerungsprozess werden auf Grundlage dieser Erkenntnisse Klassen, zwischen denen vermehrt Verwechslungen auftreten, schrittweise zu Gruppen zusammengefasst. Infolgedessen reduziert sich in jedem Schritt die Anzahl der Klassen. Dieser Prozess wird solange wiederholt bis die kontextbasierte Klassifikation für jede einzelne Klasse bzw. Gruppe von Klassen ein akzeptables Genauigkeitsniveau erreicht. Die Beurteilung der Klassifikationsergebnisse erfolgt auf der Grundlage von klassenspezifischen Genauigkeitsmaßen, die aus einer objektbasierten Evaluation abgeleitet werden. Die finale Klassenstruktur wird im Folgenden als maximal mögliche semantische Auflösung bezeichnet.

Die Klassenstruktur im ersten Schritt (vgl. Experiment 1 in Tab. 5.5 und 5.6) bildet mit vier Klassen, und zwar den in Kapitel 5.1.2 definierten vier Hauptkategorien bzw. Objektartengruppen, die gröbste Auflösung. In den folgenden Schritten (vgl. Experimente 2 bis 9 in Tab. 5.5 und 5.6) wird jeweils eine der Hauptkategorien hierarchisch verfeinert, während die anderen Kategorien in der gröbsten Auflösung verbleiben. Beispielweise wird die Objektartengruppe *Siedlung* zunächst unterteilt in die Objektarten *Wohnbaufläche*, *Industrie- und Gewerbefläche*, *Fläche gemischter Nutzung*, *Fläche besonderer funktionaler Prägung*, *Sport-, Freizeit- und Erholungsfläche* sowie *Friedhof* (vgl. Experiment 2). Im nächsten Schritt wird jede Objektart weiter untergliedert, sodass z.B. die Klasse *Wohnbaufläche* differenziert wird in *Wohnbaufläche* (bebaut) und *Erweiterung* (unbebaut) entsprechend des Attributs „Zustand“. Die Klassifikation in Schritt 10 unterscheidet die feinste semantische Auflösung aller Objektartengruppen, die, wie der Tabelle 5.2 entnommen werden kann, für die Testgebiete Hameln und Schleswig variiert. Die Unterteilung der Klassenstrukturen inklusive der jeweils erzielten Genauigkeiten sind in den Tabellen 5.5 und 5.6 separat für die Testgebiete Hameln und Schleswig aufgetragen.

Auf Grundlage der Konfusionsmatrizen werden Fehlzuweisungen dahingehend analysiert, zwischen welchen Klassen häufig Verwechslungen auftreten. Klassen mit wechselseitigen Fehlzuweisungen formen Ballungen, die zwei oder mehr Klassen umfassen können. Die zu einer Ballung gehörigen Klassen weisen zumeist starke Ähnlichkeiten in ihren Merkmalswerten auf, die in einem ähnlichen Erscheinungsbild der zugehörigen TN-Objekte in der Realität begründet liegen. Aus diesem Grund gilt es, neben der Art der Fehlzuweisung auch die Ursachen für Fehlzuweisungen zu beleuchten, d.h. zu analysieren, ob die verschiedenen Klassen überhaupt genügend charakteristische Merkmale aufweisen,

um eine Trennung im Merkmalsraum zu ermöglichen. Auf der Grundlage dieser Erkenntnisse wird die Klassenstruktur schrittweise zusammengefasst bis ein akzeptables Genauigkeitsniveau erreicht ist (vgl. Experimente 11 bis 13 in Tab. 5.5 bzw. 11 bis 14 in Tab. 5.6). In jedem Schritt werden Klassen mit einer Qualität unterhalb eines bestimmten Grenzwertes einer in Bezug auf die Semantik und die Eigenschaften ähnlichen Klasse zugewiesen. Zwei Klassen gelten hinsichtlich ihrer Eigenschaften als ähnlich, wenn sie sich in ihren Merkmalswerten nur geringfügig voneinander unterscheiden. Die Ähnlichkeit in den Eigenschaften spiegelt sich direkt in der Häufigkeit der Fehlzuweisungen wider. Bei der Auswahl geeigneter Verknüpfungen wird neben der Häufigkeit auch die semantische Ähnlichkeit berücksichtigt. So werden jene Verknüpfungen favorisiert, die in der Hierarchie eng verwandt sind, z.B. einer gemeinsamen Objektart oder Objektartengruppe zugehörig sind. Die Zusammenfassung erfolgt in mehreren Stufen, wobei in jeder Stufe der Schwellwert für die Qualität um 10 % erhöht wird. Im Rahmen dieser Arbeit gilt ein Genauigkeitsniveau als akzeptabel, wenn die Qualität einen Wert von 30 % überschreitet. Hierbei ist zu beachten, dass die Qualität generell geringere Werte annimmt als die Maße Korrektheit und Vollständigkeit, so wird z.B. eine Qualität von etwa 30 % bei einer Korrektheit und Vollständigkeit von jeweils ca. 50 % erzielt. Dieser Grenzwert hat sich als geeignet herausgestellt, da selbst für Klassen mit einer Qualität nahe 30 % korrekte Zuweisungen gegenüber Fehlklassifikationen überwiegen. Bei einer Erhöhung des Grenzwertes würde sich die Anzahl der Klassen unter Umständen drastisch reduzieren, was der zu Beginn gestellten Anforderung bzgl. einer feinen Klassenstruktur entgegensteht. Die Klassenstruktur, für die im Rahmen der Klassifikation ausnahmslos Qualitätswerte von größer als 30 % erzielt werden, ist definiert als die maximal mögliche semantische Auflösung. Zu beachten ist, dass die resultierenden Klassenstrukturen in beiden Testgebieten ggf. voneinander abweichen. Dies liegt vor allem in der unterschiedlichen Charakteristik der Testgebiete begründet. Für die weiteren Untersuchungen wird daher eine gemeinsame Klassenstruktur aus den Ergebnissen der Testgebiete abgeleitet. Für den Fall, dass eine Nutzungsart in den resultierenden Klassenstrukturen unterschiedlich fein aufgelöst ist, orientiert sich die gemeinsame Klassenstruktur an der jeweils gröberen Auflösung. Die gemeinsame Klassenstruktur entspricht der maximal möglichen semantischen Auflösung, die in beiden Testgebieten noch unterschieden werden kann.

Die Klassifikation erfolgt in allen Testdurchläufen auf Basis des Gesamtmodells (d.h. inklusive der Potentiale höherer Ordnung) unter Verwendung aller verfügbaren und berechenbaren Merkmale und der Anwendung der in der Tabelle 5.4 aufgetragenen Standardparameter. Die einzelnen Merkmalssätze sind jeweils auf das 100 %-Quantil skaliert.

5.2.1.3. Beschreibung der Ergebnisse

Siedlung Für die Kategorie *Siedlung* wird in der gröbsten Auflösung eine Qualität in der Höhe von durchschnittlich ca. 84 % für das Testgebiet Hameln und ca. 81 % für das Testgebiet Schleswig erzielt. Diese Kategorie wird nur in geringem Maße von der Verfeinerung der Klassenstruktur in den anderen Kategorien beeinflusst, was sich in beiden Testgebieten in einer maximalen Abweichung der Qualität in Höhe von 1,2 % zeigt. Folglich wirkt sich die Verfeinerung der anderen Objektartengruppen nur marginal auf das Klassifikationsergebnis für diese Kategorie aus.

Bei der Unterscheidung auf Ebene der Objektarten, d.h. der ersten Hierarchiestufe (vgl. Experi-

Klasse	Qualität [%] 1	2	3	4	5	6	7	8	9	10	11	12	13
Wbfl.	84,5	68,7	67,8	84,3	84,6	85,1	84,0	84,2	84,5	53,4	67,4	69,8	71,2
Erw.			16,0							23,3	25,9	15,4	Grnanl.
Ind.		51,0	25,5							18,4	23,5	21,8	62,8
Dnstl.			30,5							11,8	32,9	47,3	
Vers.			33,3							30,0	37,0	41,0	35,4
Landw.		-	-							-	Dnstl.	Dnstl.	Dnstl.
Öffntl.		7,4	12,3							10,6	10,4		
Sprtgeb.		51,8	-							1,2			
Sprtanl.			-							-	46,8	52,9	56,4
Erhol.			-							-			
Grnanl.			47,1							23,0			
Frhf.		-	-							-			
Str.	87,2	87,2	84,8	63,4	73,9	86,6	85,5	86,8	86,7	54,2	73,3	75,1	74,6
Fußgz.					-					-			
Strbgl.				63,3	33,9					6,1	41,0	45,2	46,4
Fahrw.					32,2					22,4	33,7	35,2	36,5
Fußw.					43,7					21,0	44,9	45,0	47,0
Parkpl.				12,1	11,4					18,2	16,7	Str.	Str.
Pl.					-					-			
Bhntr.				8,6	13,8					12,1	27,6	17,2	
Bhngeb.					-					-	Strbgl.	Strbgl.	Strbgl.
Bhnbgl.					18,3					15,2			
Schiffv.				-	-					-	Str.	Str.	Str.
Grünl.	62,0	61,9	59,7	62,8	63,2	64,0	38,3	61,2	59,7	19,8	38,0	41,7	41,8
Brachl.							-			11,8	-		
Ackerl.							58,4			42,3	56,6	61,7	60,6
Laubw.						50,8	42,2			12,0	51,6	41,4	61,6
Mischw.							15,0			7,4			
Gehlz.						33,3	26,5			11,6	30,2	29,0	
Gewbgl.						9,3	2,3			13,7	18,2	Strbgl.	Strbgl.
Fluss	33,8	31,9	31,1	36,5	31,6	30,6	25,3	26,5	47,1	38,9	44,4	40,0	35,2
Graben									-	-	23,5	10,4	
Bach									25,6	16,4			
See								-	-	16,7	-	Fluss	
G [%]	90,3	81,3	76,9	81,1	76,0	89,5	87,6	89,9	89,8	43,4	63,5	68,4	73,1
κ [%]	83,6	73,7	68,1	73,7	68,5	82,2	79,1	83,0	82,9	37,8	59,2	64,2	69,2

Tabelle 5.5.: Qualität [%] pro Klasse, Gesamtgenauigkeit G [%] und Kappa-Index κ [%] abgeleitet aus objektbasierter Evaluation der Klassifikationsergebnisse bei hierarchischer Verfeinerung der Klassenstruktur in den Experimenten 1 - 10 und der schrittweisen Zusammenfassung zu Gruppen in den Experimenten 11 - 13 für das Testgebiet Hameln. Abkürzungen siehe Tab. 5.2.

ment 2 in Tab. 5.5 und 5.6), werden für die Objektarten *Wohnbaufläche*, *Industrie- und Gewerbefläche* und *Sport-, Freizeit- und Erholungsfläche* im Testgebiet Hameln Qualitätswerte von größer als 50 % erzielt. In dem Testgebiet Schleswig zeigt sich jedoch ein anderes Bild. Für diese drei Objektarten werden zwar ebenfalls die höchsten Qualitätswerte erzielt, jedoch sind die Werte für die Objektarten *Industrie- und Gewerbefläche* sowie *Sport-, Freizeit- und Erholungsfläche* um bis zu 23,5 % niedriger im Vergleich zum Testgebiet Hameln. Bei einer weiteren Untergliederung der Objektarten (vgl. Experiment 3 in Tab. 5.5 und 5.6) lassen sich die meisten Klassen nicht mehr adäquat voneinander

Klasse	Qualität [%] 1	2	3	4	5	6	7	8	9	10	11	12	13	14
Wbfl.	81,5	60,7	64,2	81,9	80,7	81,2	80,8	81,0	81,1	26,3	63,1	65,7	68,7	69,6
Erw.			27,3							22,5	27,3	24,3	Grnanl.	Grnanl.
Ind.		27,5	7,8							6,0	48,9	51,6	53,1	55,8
Dnstl.			18,4							8,5				
Vers.			16,9							18,9	24,3	18,1		
Ents.			-							-				
Mischn.		21,1	25,0							9,8	Ind.	Ind.		
Landw.			20,0							4,0				
Öffntl.		13,5	18,1							8,6				
Hist.			-							-	Grnanl.	Grnanl.	Grnanl.	Grnanl.
Sprtgeb.		34,6	-							1,0	Ind.	Ind.	Ind.	Ind.
Sprtanl.			-							14,7	6,3	38,9	47,1	46,9
Grnanl.			36,6							11,9	36,7			
Frhf.		-	-							6,7				
Str.	77,8	78,4	73,9	71,7	78,5	77,7	75,8	77,5	78,0	50,6	68,6	69,4	71,5	70,1
Strbgl.					2,5					5,6				
Weg				54,0	54,0					31,6	49,5	53,5	56,0	55,7
Parkpl.				14,7	13,3					11,8	23,3	30,6	14,5	Str.
Pl.					-					-				
Schiffv.				-	-					-	Str.	Str.	Str.	
Grünl.	79,4	78,0	76,5	78,8	77,9	73,9	58,9	78,2	77,6	37,0	55,8	57,6	62,9	62,7
Brachl.							-			-	Unl.	Unl.		
Ackerl.							54,7			36,8	48,5	55,7	43,6	45,9
Gartenl.							-			-				
Laubw.						67,0	23,0			19,3	24,9	23,5	68,5	67,4
Nadelw.							10,5			24,5	10,9	46,9		
Mischw.							32,9			27,2	32,4			
Gehlz.						16,5	31,5			12,6	27,7	23,2		
Moor						-	-			-	21,8	23,5	Grünl.	Grünl.
Sumpf						14,0	20,8			8,9				
Unland						2,1	2,3			4,2				
Gewbgl.							-			-				
Graben	40,5	41,4	36,6	39,7	33,8	35,8	37,4	34,6	37,8	30,4	43,4	44,6	47,6	46,2
Bach									-	-				
See								38,0	22,7	27,3	26,1	22,7	37,4	35,4
Teich									17,7	15,2	22,7	23,6		
Spbeck.									-	-				
G [%]	87,7	76,1	73,8	84,7	84,0	84,2	79,5	87,2	86,8	36,4	65,8	68,4	74,8	75,6
κ [%]	81,8	69,4	67,0	78,5	77,7	77,8	71,6	81,1	80,5	32,2	61,5	64,3	71,0	71,8

Tabelle 5.6.: Qualität [%] pro Klasse, Gesamtgenauigkeit G [%] und Kappa-Index κ [%] abgeleitet aus objektbasierter Evaluation der Klassifikationsergebnisse bei hierarchischer Verfeinerung der Klassenstruktur in den Experimenten 1 - 10 und der schrittweisen Zusammenfassung zu Gruppen in den Experimenten 11 - 14 für das Testgebiet Schleswig. Abkürzungen siehe Tab. 5.2.

trennen. In dieser Detailstufe werden nur noch für die Nutzungsarten *Wohnbaufläche* und *Grünanlage* Qualitätswerte von größer als 35 % erzielt. Wenngleich dieser Sachverhalt in beiden Testgebieten zu beobachten ist, spiegeln die Qualitätswerte in Schleswig erneut ein geringeres Genauigkeitsniveau wider. Der Tabelle 5.2 ist zu entnehmen, dass für diese beiden Nutzungsarten in beiden Testgebieten die meisten Trainingsbeispiele in der Kategorie *Siedlung* zur Verfügung stehen.

Die Ergebnisse belegen, dass eine feingliedrige Unterscheidung bebauter Nutzungsarten mit Ausnahme der Klasse *Wohnbaufläche* schwierig ist. Für die bebauten Klassen *Industrie und Gewerbe*, *Handel und Dienstleistung*, *Öffentliche Zwecke* und *Ver- und Entsorgung* werden in beiden Testgebieten deutlich geringere Qualitätswerte erzielt (von max. 33,3 % in Hameln und max. 18,4 % in Schleswig). Unter den genannten Nutzungsarten erzielt die Klasse *Ver- und Entsorgung* im Testgebiet Hameln mit 33,3 % die höchste Qualität und weist folglich sehr spezielle Eigenschaften auf, die sich von den übrigen bebauten Nutzungsarten unterscheiden. Dies bestätigt sich auch bei einer Analyse der zugehörigen TN-Objekte. Die meisten TN-Objekte dieser Klasse beherbergen Umspannstationen, die sich sowohl hinsichtlich der Fläche der Bauwerke (wenige m^2) als auch der GIS-Objekte deutlich von den anderen bebauten Nutzungsarten unterscheiden. Demgegenüber weisen die übrigen bebauten Nutzungsarten zum Teil große Ähnlichkeiten untereinander auf, wie es auch in der Abbildung 5.1 zu sehen ist. Beispielsweise zeigen die TN-Objekte der Klasse *Handel und Dienstleistung* im linken Teil der Abbildung 5.1c einen Supermarkt, der aus großflächigen Flachdachhallen besteht, und sich somit kaum von den Produktionshallen eines Industrie- und Gewerbebetriebes, dargestellt in Abbildung 5.1d, unterscheidet. Weitere Ähnlichkeiten sind in Abbildung 5.1 auch noch zwischen anderen bebauten Nutzungsarten zu beobachten, wie z.B. zwischen den Klassen *Wohnbaufläche* und *Handel und Dienstleistung* bei geschlossener Bebauung (vgl. Abb. 5.1a, rechts und 5.1c, rechts). Infolgedessen kommt es bei einer zunehmenden Verfeinerung der Klassenstruktur zu vermehrten Fehlzuweisungen zwischen den bebauten Nutzungsarten, was sich in einer Verschlechterung der Genauigkeit manifestiert. Lediglich die Klasse *Öffentliche Zwecke* profitiert von der feingliedrigen Unterscheidung, so steigt die Qualität im Zuge des Verfeinerungsprozesses in beiden Testgebieten an. Trotz der Verbesserung bleibt die Qualität jedoch auf einem geringen Niveau (unterhalb 20 %), womit eine individuelle Klassifizierung nach den oben genannten Kriterien ausgeschlossen ist.

Die Unterscheidung der unbebauten Nutzungsarten gestaltet sich deutlich schwieriger. Neben der Klasse *Grünanlage* werden nur noch geringe Anteile der Nutzungsart *Erweiterung (Wohnbaufläche)* korrekt klassifiziert. Auch hier liegt die mangelnde Trennbarkeit der Klassen hauptsächlich in einer ähnlichen Charakteristik begründet. In der Abbildung 5.1 sind TN-Objekte verschiedener unbebauter Klassen dargestellt, u.a. *Erholungsfläche* (Abb. 5.1g), *Sport- und Freizeitanlage* (Abb. 5.1h), *Friedhof* (Abb. 5.1i) und *Grünanlage* (Abb. 5.1j), die allesamt durch einen überwiegenden Anteil unversiegelter Flächen geprägt sind.

Daneben existieren in beiden Testgebieten weitere Nutzungsarten, die sich gar nicht unterscheiden lassen. Hierzu zählen die Klassen *Land- und Forstwirtschaft* und *Erholungsfläche* in Hameln, *Entsorgungsanlage* und *Historische Anlage* in Schleswig sowie *Gebäude- und Freifläche (Sport-, Freizeit- und Erholungsfläche)*, *Sport- und Freizeitanlage* und *Friedhof* in beiden Testgebieten. Wie der Tabelle 5.2 zu entnehmen ist, handelt es sich hierbei um die Klassen aus der Kategorie *Siedlung* mit der geringsten Anzahl an Trainingsbeispielen in dem jeweiligen Testgebiet (maximal 26 Objekte pro Klasse). Allerdings lässt sich aus den Ergebnissen kein linearer Zusammenhang zwischen der Anzahl der Trainingsbeispiele und der Genauigkeit ableiten. So übertrifft beispielsweise die Genauigkeit für die Klasse *Ver- und Entsorgung* im Testgebiet Hameln die Genauigkeit für einige Klassen mit der dreifachen Menge an Trainingsbeispielen (z.B. *Handel und Dienstleistung*). Nichtsdestotrotz belegen die Ergebnisse, dass eine geringe Anzahl von Trainingsbeispielen die Unterscheidung einer Klasse er-

schwert. Welche Anzahl zur korrekten Unterscheidung einer Klasse notwendig ist, ist in Abhängigkeit der Eigenschaften der Klasse individuell unterschiedlich. Beispielsweise werden die TN-Objekte der Klasse *Erweiterung (Wohnbaufläche)* im Testgebiet Hameln zumindest teilweise korrekt klassifiziert, wohingegen für die Klasse *Erholungsfläche* trotz einer ähnlichen Anzahl an Trainingsbeispielen kein Objekt korrekt zugeordnet werden kann.

Auf Basis dieser Erkenntnisse werden die Klassen sukzessive in Abhängigkeit der erzielten Qualität zusammengefasst bis hin zur finalen Unterscheidung von vier Gruppen im Testgebiet Hameln bzw. drei Gruppen im Testgebiet Schleswig (vgl. Experiment 13 in Tab. 5.5 und 5.6). In der endgültigen Auflösung werden in beiden Testgebieten die Klassen *Wohnbaufläche*, *Sonstige Bebauung* und *Urbane Grünfläche* unterschieden, wobei die Klasse *Sonstige Bebauung* im Testgebiet Hameln noch differenziert wird in *Sonstige Bebauung* und *Ver- und Entsorgungsanlagen*.

Verkehr Für die Kategorie *Verkehr* wird in der gröbsten Auflösung eine hohe Qualität von durchschnittlich ca. 86 % in Hameln und ca. 77 % in Schleswig erzielt. Im Gegensatz zu der Kategorie *Siedlung*, die nur in geringem Maße von der Verfeinerung der Klassenstruktur in den anderen Kategorien beeinflusst wird, sind hier größere Variationen zu beobachten. Beispielweise führt die feingliedrige Unterscheidung der Kategorie *Siedlung* zu Einbußen in der Qualität gegenüber der gröberen Auflösung, d.h. die Qualität reduziert sich in Hameln um 2,4 % und in Schleswig um 3,9 % gegenüber der Unterscheidung aller Kategorien in der gröbsten Auflösung. Die Verfeinerung der Kategorien *Vegetation* und *Gewässer* hat hingegen eine geringere Auswirkung.

Bei der Unterscheidung der Nutzungsarten der Kategorie *Verkehr* auf Ebene der Objektarten werden für die Objektarten *Straßenverkehr* und *Weg* relativ hohe Genauigkeiten erzielt; die Qualität für beide Klassen überschreitet in Hameln 60 % und in Schleswig 50 %. Für die Objektarten *Platz* und *Bahnverkehr* (nur in Hameln) werden deutlich geringere Genauigkeiten erzielt, d.h. Qualitätswerte von kleiner als 15 %. Die Klassifizierung der Objektart *Schiffsverkehr* scheitert in beiden Testgebieten.

Für das Testgebiet Hameln werden selbst bei einer weiteren Verfeinerung der Klassenstruktur für die Unterkategorien *Straße*, *Straßenbegleitfläche*, *Fahrweg* sowie *Fuß- und Radweg* Qualitätswerte von größer als 32 % erzielt. Für die Unterkategorien *Parkplatz*, *Bahntrasse* und *Bahnbegleitfläche* wird dieses Genauigkeitsniveau hingegen nicht erreicht; die Qualität beträgt hierbei weniger als 20 %. Darüber hinaus ist eine Unterscheidung der Unterkategorien *Fußgängerzone*, *Platz (diverse Funktionen)* und *Gebäude- und Freifläche (Bahnverkehr)* im Testgebiet Hameln nicht möglich. Zusammen mit der Klasse *Schiffsverkehr*, die ebenfalls nicht korrekt klassifiziert wird, handelt es sich hierbei um die vier Klassen mit den wenigsten Trainingsbeispielen in dieser Kategorie im Testgebiet Hameln (maximale Anzahl von 19 Objekten pro Klasse). Demnach bestätigt sich die Beobachtung, dass ein Zusammenhang zwischen der Anzahl der Trainingsbeispiele und der Genauigkeit besteht, wenngleich abermals kein linearer Zusammenhang festzustellen ist.

Im Testgebiet Schleswig zeigt sich ein weniger differenziertes Bild, da der feinste Detailgrad schon alleine durch das Nichtvorhandensein bestimmter Klassen im Testgebiet (z.B. *Bahntrasse*) oder durch die weniger detaillierte Differenzierung bestimmter Nutzungsarten (z.B. der Objektart *Weg* in *Fahrweg* und *Fuß- und Radweg*) eingeschränkt ist. Aus diesem Grund bezieht sich die nachfolgende weiterge-

hende Analyse der Experimente für die Kategorie *Verkehr* auf das Testgebiet Hameln.

Bei Verfeinerung der Kategorie *Verkehr* reduziert sich die Qualität für die meisten Klassen. Ausgenommen hiervon sind die Nutzungsart *Straße*, deren Qualität gegenüber der gröberen Auflösung um ca. 10 % zunimmt, sowie die Unterkategorien *Bahntrasse* und *Bahnbegleitfläche* der Objektart *Bahnverkehr*, deren Qualitätswerte um jeweils mehr als 5 % ansteigen (vgl. Experimente 4 und 5 in Tab. 5.5). Da die Unterkategorien der Objektart *Bahnverkehr* jedoch trotz signifikantem Anstieg nicht die Marke von 30 % überschreiten, wird von einer separaten Ausweisung dieser Klassen abgesehen.

Insgesamt lassen sich in dieser Gruppe drei Ballungen von Nutzungsarten identifizieren, die von wechselseitigen Fehlzuweisungen betroffen sind. Dies sind zum einen die versiegelten Verkehrsflächen, wozu beispielsweise die Klassen *Straße*, *Fußgängerzone*, *Platz*, *Bahntrasse* und *Schiffsverkehr (Landfläche)* zählen. Die Fehlzuweisungen treten zwar vor allem innerhalb der Gruppe auf, jedoch kommt es beispielsweise bzgl. der Klasse *Platz* auch zu Verwechslungen mit allen Klassen der Gruppe *Urbane Grünfläche* aus der Kategorie *Siedlung* und mit landwirtschaftlichen Flächen sowie der Klasse *Gehölz* aus der Kategorie *Vegetation*. Die zweite Gruppe bilden die Kategorien *Fahrweg* und *Fuß- und Radweg*, die hauptsächlich von gegenseitigen Fehlzuweisungen betroffen sind, neben wenigen Verwechslungen mit den Klassen *Straße* und *Straßenbegleitfläche*. Die dritte Gruppe bilden die überwiegend unversiegelten Verkehrsbegleitflächen inklusive baulicher Anlagen, die sowohl an Straßen- als auch Bahnverkehrsflächen zu finden sind. Da sich die Verkehrsbegleitflächen an Straßen und Bahntrassen sehr ähneln (längliche Form, unversiegelt) treten vermehrt gegenseitige Fehlzuweisungen auf. Daneben kommt es zu Verwechslungen mit den *Gewässerbegleitflächen* aus der Kategorie *Vegetation*, die sich sehr ähnlich in den Daten ausprägen. Bei Zusammenfassung der Klassen zu den zwei Gruppen versiegelte Verkehrsflächen bzw. unversiegelte Verkehrsbegleitflächen wird im Testgebiet Hameln jeweils ein Qualitätsniveau von 74,6 % bzw. 46,4 % erreicht (vgl. Experiment 13 in Tab. 5.5). Die Zusammenfassung wirkt sich zudem positiv auf die Klassen *Fahrweg* und *Fuß- und Radweg* aus, deren Qualität sich dadurch um 2 - 3 % erhöht. In dem finalen Testdurchlauf wird für die Klasse *Fuß- und Radweg* eine Qualität von 47 % erreicht und für die Klasse *Fahrweg* eine etwas geringere Qualität von 36,5 %. Diese Werte überschreiten die 30 %-Marke, sodass in der endgültigen Auflösung im Testgebiet Hameln von der Zusammenfassung zu einer Klasse abgesehen wird (vgl. Experiment 13 in Tab. 5.5).

Vegetation In Bezug auf die Kategorie *Vegetation* weist das Testgebiet Schleswig im Gegensatz zu dem Testgebiet Hameln eine deutlich differenziertere Klassenstruktur auf. Dies liegt vor allem in der überwiegend ländlichen Struktur des Testgebietes begründet. Die Qualität ist zwar in beiden Testgebieten relativ hoch, jedoch übersteigt die durchschnittliche Qualität in der gröbsten Auflösung im Testgebiet Schleswig mit einem Wert von 78 % deutlich jene in Hameln (62 %). Die Qualität für die Gesamtkategorie wird in beiden Testgebieten von der Verfeinerung der anderen Kategorien beeinflusst. Die maximale Abweichung beträgt 3,5 % in Hameln und 2,9 % in Schleswig.

Die Unterscheidung auf Ebene der Objektarten liefert in beiden Testgebieten gute Ergebnisse für die Objektarten *Landwirtschaft* und *Wald*, wobei das Qualitätsniveau im Testgebiet Schleswig bzgl. der Objektart *Landwirtschaft* um ca. 10 % und bzgl. der Objektart *Wald* sogar um 16 % höher ist als im Testgebiet Hameln. Bei der weiteren Verfeinerung zeigt sich, dass bzgl. der landwirtschaftlichen

Flächen nur die Klassen *Grünland* und *Ackerland* gemäß der oben genannten Kriterien zuverlässig getrennt werden können. Die übrigen Klassen *Brachland* und *Gartenland* (nur in Schleswig) lassen sich hingegen nicht unterscheiden. Die feine Untergliederung der Objektart *Wald* führt in beiden Testgebieten zu starken Genauigkeitseinbußen. Lediglich eine Unterkategorie dominiert unter den erzielten Genauigkeiten mit einer Qualität von größer als 30 % (*Mischwald* in Schleswig und *Laubwald* in Hameln). Ein Unterschied zwischen den Testgebieten prägt sich im Hinblick auf die Objektart *Gehölz* aus, die in Hameln bereits in der ersten Hierarchiestufe eine Qualität von größer als 30 % erreicht und dann bei einer weiteren Verfeinerung an Genauigkeit verliert. Im Testgebiet Schleswig profitiert diese Klasse hingegen von der zunehmenden Verfeinerung und erreicht erst in der höchsten Auflösung eine Qualität von ca. 30 %. Die übrigen Klassen *Moor*, *Sumpf*, *Unland* und *Gewässerbegleitfläche* werden nicht oder nur mit schlechter Genauigkeit klassifiziert, wobei diese Ergebnisse auch wieder mit einer geringen Anzahl an verfügbaren Trainingsbeispielen einhergehen (maximale Anzahl von 38 Objekten). Für Klassen mit insgesamt weniger als 11 Objekten wird keines der Objekte korrekt klassifiziert. Insgesamt werden aber trotz einer relativ geringen Anzahl an Trainingsbeispielen pro Klasse (maximal 61 Objekte im Testgebiet Hameln) relativ gute Qualitätswerte erzielt, z.B. 58,4 % für *Ackerland* mit 61 Objekten in Hameln, welche die Genauigkeit für die Klasse *Bahnbegleitfläche* mit einer ähnlichen Anzahl an Objekten deutlich übertrifft. Aus diesem Vergleich lässt sich folgern, dass nicht alleine die Anzahl sondern vielmehr die Repräsentativität der Trainingsdaten sowie die allgemeine Charakteristik der Klasse in Bezug auf die Merkmale entscheidend ist für die Trennung der Klasse, die im Fall der Klasse *Ackerland* offensichtlich besser gegeben ist.

In der Konfusionsmatrix prägen sich insbesondere drei Ballungen aus, zwischen denen vermehrt Fehlzuweisungen auftreten. Die erste Ballung bildet die Klasse *Grünland* mit den assoziierten Klassen *Brachland* sowie *Moor*, *Sumpf* und *Unland* (nur in Schleswig). Die assoziierten Klassen weisen in den meisten Fällen ein sehr ähnliches Erscheinungsbild zu Grünlandobjekten auf, d.h. großflächige Einheiten mit Grasbewuchs. Die Durchsättigung mit Wasser, die in Moor- und Sumpfgebieten gegeben ist, prägt sich nicht in den zur Verfügung stehenden Testdaten aus, was die Abgrenzung von Grünlandobjekten erschwert. Als Folge der fehlenden Bewirtschaftung von Objekten der Klassen *Brachland* und *Unland* weisen diese Flächen zumeist einen niedrigen Bewuchs auf, der kaum von dem Grasbewuchs von Grünlandobjekten zu unterscheiden ist. Die zweite Gruppe bildet die Klasse *Ackerland*, der in Schleswig noch die Klasse *Gartenland* zugeordnet ist. Bei der dritten Gruppe handelt es sich um bewaldete Flächen, zu denen die Klassen *Laubwald*, *Nadelwald* (nur in Schleswig), *Mischwald* und *Gehölz* zählen. Eine feine Untergliederung scheitert in beiden Testgebieten an unterschiedlichen Gründen. Die verschiedenen Waldtypen lassen sich insbesondere dann unterscheiden, wenn sie sich in den Sensordaten unterschiedlich ausprägen. Dies ist insbesondere im Winter oder Frühjahr gegeben, wenn die Laubbäume keine Blätter haben. Die Luftbildbefliegung für den Testdatensatz Schleswig erfolgte allerdings im Sommer, wodurch sich die unterschiedlichen Waldtypen in den Eingangsdaten sehr ähnlich ausprägen und damit eine Trennung erschweren. In Hameln erfolgte die Befliegung zwar zu einem geeigneteren Zeitpunkt, jedoch liegen hier – speziell für die Klasse *Mischwald* – nur sehr wenige Trainingsbeispiele vor, die offensichtlich nicht ausreichen, um den Klassifikator in geeigneter Weise anzulernen. Eine Besonderheit bildet die Klasse *Gewässerbegleitfläche*, die sich in beiden Testgebieten infolge unterschiedlicher Grundsätze bei der Erfassung der Landnutzung unterschiedlich ausprägt. In

Hameln sind die Uferstreifen detailliert als *Gewässerbegleitfläche* erfasst, wohingegen die Uferstreifen in Schleswig meist nicht gesondert ausgewiesen sind. Infolgedessen existieren in dem gesamten Testgebiet Schleswig nur 5 Objekte dieser Klasse. Im Gegensatz zu den länglich geformten Uferstreifen in Hameln handelt es sich hierbei vielmehr um die großflächigen Uferzonen eines stehenden Gewässers, die von ihrer Charakteristik eher der von Sumpfgebieten ähneln und daher der Gruppe *Grünland* zugewiesen werden. Demgegenüber treten in Hameln aufgrund ihrer charakteristischen länglichen Form überwiegend Fehlzuweisungen zu der Gruppe der unversiegelten Verkehrsbegleitflächen auf, sodass diese Klassen in den Experimenten 12 und 13 in Tabelle 5.5 zusammengefasst werden.

Nach der schrittweisen Zusammenfassung werden für die Gruppen *Grünland* und *Ackerland* Qualitätswerte von größer als 40 % bzw. 60 % erzielt (der jeweils höhere Wert für die Klasse mit der größeren Anzahl an Trainingsbeispielen pro Testgebiet). Die Qualität für die Klasse *Wald* überschreitet in beiden Testgebieten einen Wert von 60 %.

Gewässer Die Genauigkeit, die für die Kategorie *Gewässer* in der gröbsten Auflösung erzielt wird, ist mit durchschnittlich ca. 32 % in Hameln und ca. 38 % in Schleswig insgesamt deutlich schlechter als die der übrigen Kategorien. Darüber hinaus hat die Verfeinerung der anderen Kategorien den stärksten Einfluss, was sich in großen Variationen der Qualität von bis zu 11,2 % in Hameln und 7,6 % in Schleswig ausdrückt.

Die Verfeinerung dieser Kategorie führt in beiden Testgebieten zu anderen Ergebnissen, was vor allem in der unterschiedlichen Struktur der Testgebiete begründet liegt. Mit Ausnahme der Klasse *Fluss*, die nur in Hameln vorkommt, sind die übrigen Klassen häufiger im Testgebiet Schleswig vertreten. In Hameln scheitert die Unterscheidung bereits auf Ebene der Objektarten, was vermutlich in der ungenügenden Anzahl an Trainingsdaten für die Objektart *Stehendes Gewässer* begründet liegt (insgesamt 5 Objekte). Im Zuge der weiteren Verfeinerung zeigt sich, dass neben der Klasse *Fluss*, für die in der feinsten Auflösung eine Qualität von 47,1 % erzielt wird, keine weiteren Klassen unterschieden werden können. Ein anderes Bild ergibt sich im Testgebiet Schleswig. Hier liefert die Klassifikation auf Ebene der Objektarten Qualitätswerte von größer als 30 %. Bei der weiteren Verfeinerung kristallisiert sich heraus, dass neben der Klasse *Graben* keine weiteren Klassen adäquat klassifiziert werden können. Selbst für die Klasse *Teich*, die über eine vergleichbar hohe Anzahl an Trainingsbeispielen verfügt, bleibt die Qualität unter 20 %. Wie auch schon zuvor festgestellt, hat aber nicht die Anzahl allein einen maßgeblichen Einfluss auf das Klassifikationsergebnis, sondern vielmehr die Repräsentativität der Trainingsbeispiele sowie die allgemeine Charakteristik der jeweiligen Klasse. So liegen für die Klasse *Fluss* in Hameln mit einer geringen Anzahl von 10 zwar wenige TN-Objekte vor, die allerdings für diese Klasse repräsentativ genug sind, um die mit 47 % höchste Qualität in dieser Kategorie zu erzielen und damit die Ergebnisse vieler anderer Klassen mit weit mehr Trainingsdaten zu übertreffen. Führt man sich diesbezüglich das typische Erscheinungsbild der Klasse *Fluss* vor Augen, das in der Regel aus reinen Wasserflächen ohne Uferböschungen besteht (Uferlinie als Begrenzung), so wird deutlich, dass sich die einzelnen Objekte definitionsgemäß sehr ähnlich sind. Demgegenüber zeigen andere Klassen, wie z.B. die Klasse *Bach*, deutlich größere Variationen, da in diesen TN-Objekten – gemäß den Erfassungsgrundsätzen der Landesvermessungsbehörden – neben den deutlich schmaleren Wasserflächen auch Streifen von Uferböschungen enthalten sein können. Zudem werden schmale Gewässerläufe

häufig von Vegetation verdeckt, die über die Wasserfläche ragt. Da Vegetationsformen naturgemäß stärkere Variationen aufweisen als homogene Wasseroberflächen, wird trotz einer größeren Anzahl an Trainingsbeispielen eine weit geringere Genauigkeit erzielt.

Bei der Analyse der Konfusionsmatrix fällt auf, dass sich die Fehlklassifikationen, neben denen innerhalb der Kategorie, auf diverse Klassen anderer Kategorien verteilen, zu denen u.a. die Klassen *Straße*, *Weg*, *Verkehrsbegleitflächen* (*Bahn*, *Straße*), *Ackerland*, *Laubwald*, *Gehölz* und *Gewässerbegleitfläche* gehören. Folglich lässt sich kein Schwerpunkt ausmachen, was einen Hinweis auf die Variabilität der Objekte dieser Kategorie gibt. Im Testgebiet Schleswig werden bei Unterscheidung auf Ebene der Objektarten Qualitätswerte von größer als 30 % erzielt, sodass in der finalen Klassenstruktur die Gruppen *Fließgewässer* und *Stehendes Gewässer* unterschieden werden können (vgl. Experiment 14 in Tab. 5.6). Anders verhält es sich für das Testgebiet Hameln. Hier sind in der finalen Klassenstruktur alle Nutzungsarten dieser Kategorie zu einer Gruppe vereinigt (vgl. Experiment 13 in Tab. 5.5).

5.2.1.4. Diskussion

Die Experimente haben gezeigt, dass die maximal erreichbare semantische Auflösung im Wesentlichen durch drei Einflussfaktoren limitiert ist. Zum einen sind sich einige Klassen in ihrem Erscheinungsbild sehr ähnlich, was eine Trennung dieser Klassen erschwert bzw. unmöglich macht. Die Ähnlichkeit der Klassen kann dabei in Abhängigkeit der regional unterschiedlichen Charakteristik des Untersuchungsgebietes variieren. Insgesamt werden vor allem für jene Klassen gute Genauigkeiten erzielt, die sich mit charakteristischen Eigenschaften von den übrigen Klassen abgrenzen. Darüber hinaus haben die Eigenschaften der Sensordaten, insbesondere der Erfassungszeitpunkt, einen wesentlichen Einfluss auf die Auflösung der Klassenstruktur, speziell auf die Trennbarkeit der einzelnen Nutzungsarten der Kategorie *Vegetation*. Den dritten limitierenden Einflussfaktor bildet die Anzahl und Repräsentativität der Trainingsdaten. Dieser Faktor hat bei Landnutzungsobjekten eine besondere Relevanz, da die zugehörigen GIS-Objekte typischerweise flächenmäßig große Einheiten beschreiben, was selbst bei der Klassifizierung großer Gebiete (30.000^2 Pixel im Testgebiet Schleswig) zu einer insgesamt geringen Anzahl an Trainingsbeispielen (3739 Objekte) führt. Darüber hinaus enthält der Objektartenkatalog teilweise in Bezug auf das Erscheinungsbild und die Funktion sehr spezialisierte Nutzungsarten, die nur sehr selten in der Realität auftreten und damit auch nur vereinzelt in den Testdaten enthalten sind.

Als Ergebnis der Untersuchungen kann festgehalten werden, dass bei einer semantischen Auflösung von 12 Klassen in Hameln(vgl. Experiment 13 in Tab. 5.5) und 10 Klassen in Schleswig (vgl. Experiment 14 in Tab. 5.6) bei Anwendung eines kontextbasierten Klassifikators akzeptable Ergebnisse erzielt werden können. Diese Klassenstrukturen spiegeln jeweils die maximal erreichbare semantische Auflösung je Testgebiet wider. Sie sind in der Tabelle 5.7 in der dritten (für Hameln) und vierten Spalte (für Schleswig) zusammengefasst. Unter Anwendung dieser Klassenstrukturen wird in beiden Testgebieten ein ähnliches Genauigkeitsniveau bzgl. der Gesamtgenauigkeit erreicht, d.h. 73,1 % in Hameln und 75,6 % in Schleswig.

Aus den Klassenstrukturen beider Testgebiete wird eine gemeinsame Klassenstruktur nach dem Prinzip des kleinsten gemeinsamen Nenners ermittelt, die für die nachfolgenden Experimente fest-

gehalten wird. Die Klassenstrukturen beider Testgebiete inklusive der zugeordneten Nutzungsarten sowie die daraus abgeleitete gemeinsame Klassenstruktur sind in der Tabelle 5.7 dargestellt. Durch die Zusammenführung verliert man in beiden Testgebieten in Teilbereichen an semantischer Auflösung, z.B. bzgl. der Nutzungsartengruppen *Sonstige Bebauung*, *Verkehrsweg* und *Weg* in Hameln sowie *Gewässer* in Schleswig. Die Nutzungsart *Gewässerbegleitfläche* wird in den folgenden Untersuchungen – entsprechend der Einteilung für das Testgebiet Hameln – der Klasse *Verkehrsweg* zugewiesen, da diese Nutzungsart mit nur 5 Objekten in Schleswig unterrepräsentiert ist. Durch die Verwendung einer einheitlichen Klassenstruktur sind die Klassifikationsergebnisse in beiden Testgebieten vergleichbar, sodass für beide Testgebiete gültige Schlussfolgerungen gezogen werden können. Außerdem gilt es, eine möglichst allgemein anwendbare Klassenstruktur zu definieren, die bei der Anwendung des Verfahrens in der Praxis auch in vielen anderen Gebieten zufriedenstellende Ergebnisse liefert. Im Verifikationsschritt werden bei gering aufgelösten Klassenstrukturen Veränderungen zwar nicht im Detail erkannt, aber bei einer hohen Genauigkeit können zumindest in der gröberen semantischen Auflösung gesicherte Aussagen getroffen werden. Hierbei ist zu berücksichtigen, dass jede unsichere Klassifikationsentscheidung unter Umständen einen Änderungshinweis induziert. Folglich steigt mit einer höheren Unsicherheit auch der manuelle Nachbearbeitungsaufwand, der vor dem Hintergrund der Effizienzsteigerung möglichst auf ein Minimum zu begrenzen ist.

Klasse	Abk.	Hameln	Schleswig	Zugeordnete Nutzungsarten
Wohnbaufläche	Wbfl.	Wohnbaufläche	Wohnbaufläche	Wbfl.
Sonst. Bebauung	Sonst.	Sonst. Bebauung	Sonst. Bebauung	Ind., Dnstl., Mischn., Landw., Öffntl., Sprtgeb.
		Ver- und Entsorg.		Vers., Ents.
Urb. Grünfläche	Grünfl.	Urb. Grünfläche	Urb. Grünfläche	Erw., Hist., Sprtanl., Erhol., Grnanl., Frhf.
Verkehrsweg	Verkw.	Verkehrsweg	Verkehrsweg	Str., Fußgz., Parkpl., Pl., Bhntr., Schiffv.
		Begleitflächen		Strbgl., Bhngeb., Bhnbgl., (HM: Gewbgl.)
Weg	Weg	Fahrweg	Weg	Fahrw.
		Fuß- und Radweg		Fußw.
Ackerland	Ackerl.	Ackerland	Ackerland	Ackerl., Gartenl.
Grünland	Grünl.	Grünland	Grünland	Grünl., Brachl., Moor, Sumpf, Unl., (SL: Gewbgl.)
Wald	Wald	Wald	Wald	Laubw., Nadelw., Mischw., Gehlz.
Gewässer	Gew.	Gewässer	Fließgewässer	Fluss, Graben, Bach
			Stehend. Gew.	See, Teich, Spbeck.

Tabelle 5.7.: Maximal mögliche semantische Auflösung für die Testgebiete Hameln (dritte Spalte) und Schleswig (vierte Spalte) sowie die daraus abgeleitete gemeinsame Klassenstruktur (erste Spalte) inklusive der jeweils zugeordneten Nutzungsarten (fünfte Spalte). Abkürzungen (Abk.) der Nutzungsarten siehe Tab. 5.2. Für die finale Zuordnung der Klasse *Gewässerbegleitfläche* ist das Ergebnis für Hameln maßgeblich.

Die Klassenstrukturen für beide Testgebiete weisen eine im Vergleich zu der feinsten Gliederungsstufe in Tabelle 5.2 deutlich reduzierte Auflösung auf, d.h. 12 gegenüber ursprünglich 34 Klassen in Hameln (vgl. die Spalten für Hameln in den Tab. 5.2 und 5.7) und 10 gegenüber ursprünglich 37 Klassen in Schleswig (vgl. die Spalten für Schleswig in den Tab. 5.2 und 5.7). Die daraus abgeleitete gemeinsame

Klassenstruktur besteht aus 9 Klassen. In welcher Tiefe die Nutzungsarten hierbei noch unterschieden werden können, variiert in Abhängigkeit der Objektart. Für einzelne Objektarten ist eine Trennung auf Ebene der Attribute möglich, wie z.B. für die Objektarten *Wohnbaufläche* und *Landwirtschaft*, wohingegen andere Nutzungsarten nur auf Ebene der Objektarten, z.B. *Wald* und *Weg*, oder sogar nur auf Ebene der Objektartengruppe, z.B. *Gewässer*, unterschieden werden können. Darüber hinaus werden weitere Klassen als Konglomerat aus verschiedenen semantisch ähnlichen Objektarten einer Objektartengruppe gebildet, wie z.B. im Falle von *Sonstiger Bebauung* oder *Verkehrsweg*. Insgesamt kann allerdings festgehalten werden, dass die wesentlichen die Landschaft strukturierenden Nutzungsarten enthalten sind. Folglich kann auf Basis dieser Klassenstruktur eine sinnvolle und aussagekräftige Gliederung der Landschaft vorgenommen werden, die zudem einen effektiven Verifikationsprozess ermöglicht.

5.2.2. Parameter der Superpixelsegmentierung

5.2.2.1. Zielsetzung

Die folgenden Experimente haben zum Ziel, den Einfluss der Parameter der SLIC-Segmentierung [Achanta et al., 2012] auf die Klassifikationsergebnisse zu untersuchen. Diese Parameter wirken sich in erster Linie auf die Klassifikation der Bodenbedeckung aus, da sie die Form und Größe der zugrundeliegenden Bildprimitive, d.h. der Superpixel, bestimmen. Um den entwickelten Ansatz effektiv in der Praxis einsetzen zu können, werden an die Bodenbedeckungsklassifikation vielfältige Anforderungen gestellt. Hierzu zählt neben einer feingliedrigen räumlichen Unterteilung der Landschaft, einer detaillierten Unterscheidung von Bodenbedeckungsarten sowie einer hohen Korrektheit und Vollständigkeit der Klassifikationsergebnisse insbesondere auch eine zweckmäßige Rechenzeit. Die Rechenzeit ist von besonderer Relevanz, da sie einen limitierenden Faktor für den praktischen Einsatz darstellt. Aus diesem Grund soll die Analyse neben der Qualität der Ergebnisse zusätzlich die Rechenzeit als Kriterium einbeziehen, sofern hier ein Einfluss festzustellen ist. Ziel ist es, jene Parameter für die weitere Prozessierung auszuwählen, die einen guten Kompromiss zwischen Rechenzeit und Genauigkeit darstellen. Darüber hinaus gilt es zu untersuchen, inwiefern sich die Wahl der Parameter der Superpixel auf die Klassifikation der Landnutzung auswirkt.

5.2.2.2. Strategie

Die Auswahl geeigneter Parameter für die SLIC Superpixel erfolgt im Rahmen einer empirischen Untersuchung anhand des Testdatensatzes Hameln. In einem Vorverarbeitungsschritt werden auf Basis der Eingangsdaten Superpixel in unterschiedlicher Größe und Form generiert, wobei deren Ausprägung in Abhängigkeit der jeweils verwendeten Größe, Kompaktheit und Eingangsdaten variiert. Im Rahmen der Untersuchungen wird jeweils ein Parameter verändert, während die übrigen konstant bleiben (Standardwerte).

Die Untersuchungen zur Größe basieren auf Superpixeln, die 100, 400 bzw. 2.500 Pixel umfassen, und damit drei verschiedene Auflösungsstufen repräsentieren. Kleinere Einheiten stehen im Widerspruch zu der allgemeinen Motivation zur Verwendung von Superpixeln anstelle von Pixeln und

werden daher an dieser Stelle nicht weiter untersucht. Zu den Vorteilen von Superpixeln zählen neben dem Rechenzeitgewinn auch ein Glättungseffekt, der durch die Segmentierung erzielt wird. Kleinere Einheiten als 100 Pixel (ca. $2 \times 2\,m^2$) haben in der aktuellen Implementierung einen drastischen Anstieg der Rechenzeit zur Folge, der die Durchführung umfangreicher Sätze an Experimenten erschwert. Demgegenüber sind größere Einheiten als 2.500 Pixel (ca. $10 \times 10\,m^2$) ebenfalls nicht zweckmäßig, da sie nur noch eine grobe Gliederung der Landschaft ermöglichen.

Die Kompaktheit variiert im Rahmen der Experimente in dem Intervall $[1, 80]$. Als Eingangsdaten für die SLIC-Segmentierung dienen zum einen die sekundären Bildinformationen NDVI, nDOM und Intensität und zum anderen die originären Spektralkanäle R,G,B (primäre Bildinformation). Die primäre bzw. sekundäre Bildinformation wird vorab pro Pixel aus den Sensordaten abgeleitet und auf das 98 %-Quantil normalisiert.

Die Klassifikation erfolgt auf Basis des Gesamtmodells (d.h. inklusive der Potentiale höherer Ordnung) unter Verwendung der Standardparameter (siehe Tab. 5.4). In jedem Satz an Experimenten wird jeweils ein Parameter variiert, während die übrigen konstant bleiben. Die erzielten Klassifikationsergebnisse werden auf der Grundlage von Genauigkeitsmaßen verglichen, wobei neben der Gesamtgenauigkeit auch die Genauigkeit pro Klasse betrachtet wird. Die Genauigkeitsmaße für die Bodenbedeckung leiten sich aus dem pixelbasierten Vergleich der Klassifikationsergebnisse mit der Referenz ab.

Durch die Klassifikation von Superpixeln anstelle von Pixeln wird allen Pixeln innerhalb eines Superpixels das gleiche Klassenlabel zugewiesen, obwohl sie unter Umständen zu unterschiedlichen Klassen gehören. Dies trifft auch für das Referenzlabel zu, das mittels Mehrheitsentscheid aller Pixel innerhalb eines Superpixels bestimmt wird. Infolgedessen werden Klassen, die typischerweise nur kleine Gruppen von Pixeln bedecken, wie z.B. *Fahrzeug*, zu anderen Klassen hinzugefügt, und sind damit nicht adäquat in den Trainingsdaten repräsentiert. Demgegenüber bietet die Segmentierung jedoch den Vorteil, dass Superpixel Pixel mit homogenen Eigenschaften zusammenfassen. Dies ist selbst dann der Fall, wenn darunter vereinzelt Pixel auftreten, die untypische Eigenschaften aufweisen. Hiermit wird der Bodenbedeckungsklassifikation von vornherein ein Glättungseffekt auferlegt, der sich mit den natürlichen Gegebenheiten der Bodenbedeckung in der Realität deckt. Bei der Evaluation der Ergebnisse der Bodenbedeckungsklassifikation auf Basis von Pixeln, d.h. bezogen auf den Vergleich mit der pixelbasierten Referenz, ist eine maximale Genauigkeit von 100 % rein konzeptionell mit Superpixeln anstelle von Pixeln nicht zu erreichen. Aus diesem Grund wird zusätzlich eine Referenzgenauigkeit berechnet, die das maximale Genauigkeitsniveau beschreibt, das sich mit der jeweiligen Segmentierung im optimalen Fall erzielen lässt. Die Referenzgenauigkeit ergibt sich aus dem pixelweisen Vergleich der Referenz pro Superpixel zu den pixelbasierten Referenzdaten.

Auf der Grundlage der Genauigkeitsmaße sollen die Auswirkungen der unterschiedlichen Größen identifiziert und anhand geeigneter Beispiele visualisiert werden. Zudem hat die Größe der Superpixel einen erheblichen Einfluss auf die Rechenzeit. Eine feine Segmentierung führt zu einer hohen Anzahl der zu klassifizierenden Bildprimitive, was einen Anstieg der Rechenzeit zur Folge hat. Daher gilt es, zusätzlich den Einfluss der Größe der Superpixel auf die Rechenzeit zu analysieren. Das Ziel besteht darin, die Kriterien Qualität und Rechenzeit gegeneinander abzuwägen, wobei der Qualität ein höheres Gewicht beizumessen ist, und darauf basierend geeignete Parameter für die weitere Prozessierung

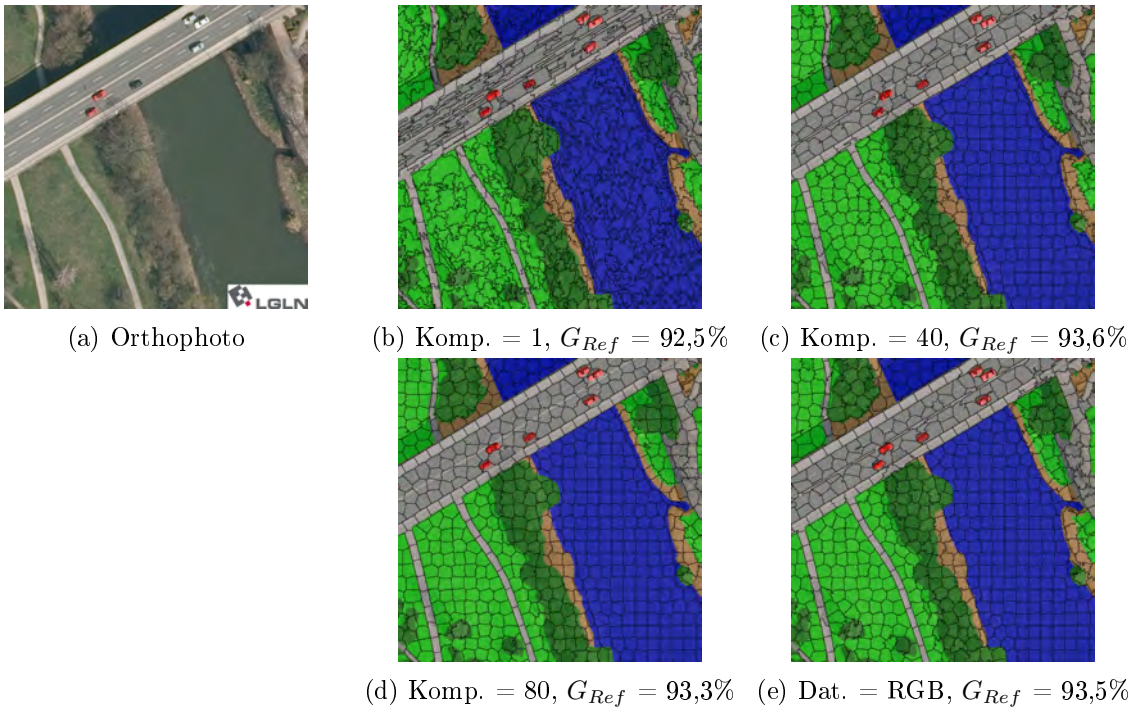
festzulegen. Die Größe ist so zu definieren, dass die Superpixel einerseits so groß wie möglich sind, um einen Glättungseffekt und einen Rechenzeitgewinn zu erzielen, und andererseits so klein wie nötig, um die wesentlichen Strukturen in der Szene zu erfassen. Bezüglich der Kompaktheit ist es das Ziel, einen guten Kompromiss zwischen möglichst kompakten Einheiten und einer flexiblen Anpassung an spektrale Übergänge zu erreichen. Die Eingangsdaten sind so zu wählen, dass auf deren Grundlage die Bodenbedeckungsarten möglichst eindeutig voneinander abgegrenzt werden können.

Darüber hinaus wird der Einfluss der Größe der Superpixel auf die Landnutzungsklassifikation evaluiert. Zur Beurteilung des Einflusses werden abermals die Genauigkeitsmaße herangezogen, die sich für die Landnutzungsklassifikation aus einem Vergleich der Klassifikationsergebnisse mit der Referenz auf Ebene der Landnutzungsobjekte ergeben.

5.2.2.3. Beschreibung der Ergebnisse

Kompaktheit der Superpixel, Eingangsdaten Im Rahmen der ersten Untersuchung wird der Einfluss des Kompaktheitsparameters sowie der Eingangsdaten auf die Ergebnisse der Superpixelsegmentierung untersucht. Die Untersuchung erfolgt anhand der erreichbaren Referenzgenauigkeit je Segmentierung. Hier zeigt sich, dass sowohl die Kompaktheit als auch die verwendeten Eingangsdaten nur einen geringen Einfluss auf die Gesamtgenauigkeit (Referenz) haben. Die Variation dieser Größen führt zu einer maximalen Abweichung in der Gesamtgenauigkeit (Referenz) von 1,1 %, die zwischen dem besten Ergebnis mit einer Gesamtgenauigkeit (Referenz) von 93,6 % bei einer Kompaktheit von 40 und dem schlechtesten Ergebnis mit einer Gesamtgenauigkeit (Referenz) von 92,5 % bei einer Kompaktheit von 1 auftritt, jeweils bei Verwendung der sekundären Bildinformationen als Eingangsdaten. Die Analyse der Referenzgenauigkeit pro Klasse zeigt ein ähnliches Bild. Die Qualität weicht für die Klassen *Gebäude*, *Versiegelung*, *Boden*, *Gras*, *Baum* und *Wasser* um maximal 2,4 % voneinander ab. Lediglich für die Klassen *Fahrzeug* und *Sonstiges* ist die maximale Abweichung zwischen den Qualitätswerten (Referenz) mit 9 % bzw. 6 % deutlich größer. Diese beiden Klassen profitieren von einer hohen Kompaktheit. Alleine die Erhöhung der Kompaktheit unter Verwendung der sekundären Bildinformationen als Eingangsdaten führt zu einer Steigerung von 5 % für die Klasse *Fahrzeug* (Maximum bei Kompaktheit 60) und 6 % für die Klasse *Sonstiges* (Maximum bei Kompaktheit 80). Die Verwendung der primären Bildinformation als Eingangsdaten bewirkt eine weitere Steigerung von 4 % für die Klasse *Fahrzeug*, wohingegen für die Klasse *Sonstiges* keine weitere Verbesserung erzielt wird.

Die Auswirkungen des Kompaktheitsparameters sowie der verwendeten Eingangsdaten auf die Superpixelsegmentierung sind in der Abbildung 5.2 anhand einer Beispielszene aus dem Testgebiet Hameln visualisiert. Die Erhöhung der Kompaktheit in dem Wertbereich $[1, 80]$ wirkt sich deutlich auf die Form der Superpixel aus. Bei einer Kompaktheit von 1 (vgl. Abb. 5.2b) weisen die Superpixel eine sehr ausgefranste Form auf, die sich mit Erhöhung dieses Parameters auf den maximalen Wert von 80 in weiten Teilen einer kompakten Form annähert (vgl. Abb. 5.2d). Die Strukturen in heterogenen Gebieten, wie z.B. dem Bereich der Uferböschungen, werden bei einer geringen Kompaktheit filigran in separaten Superpixeln erfasst (vgl. Abb. 5.2b). Die Superpixel passen sich hinsichtlich ihrer Umrandungen bestmöglich an die Grauwertübergänge an. Der Zwang zu einer möglichst hohen Kompaktheit, der durch einen hohen Parameterwert in die Segmentierung eingebracht wird, führt selbst in diesen Ge-



(a) Orthophoto (b) Komp. = 1, G_{Ref} = 92,5% (c) Komp. = 40, G_{Ref} = 93,6%

(d) Komp. = 80, G_{Ref} = 93,3% (e) Dat. = RGB, G_{Ref} = 93,5%

Abbildung 5.2.: Orthophoto einer Szene im Testgebiet Hameln (a), überlagert mit den Ergebnissen der Superpixelsegmentierung für eine Größe von 400 Pixeln und die Kompaktheitsparameter (Komp.) 1 (b), 40 (c) und 80 (d) unter Verwendung der sekundären Bildinformationen als Eingangsdaten sowie für eine Kompaktheit von 20 bei Verwendung der primären Bildinformation (Dat. = RGB) als Eingangsdaten (e). Die Grenzen der segmentierten Superpixel sind jeweils den pixelbasierten Referenzdaten überlagert. Die Farben markieren die Zugehörigkeit zur Bodenbedeckungsklasse: *Geb.* (orange), *Vers.* (grau), *Bod.* (braun), *Gras* (hellgrün), *Baum* (dunkelgrün), *Was.* (blau), *Fhzg.* (rot) und *Sonst.* (rosa). G_{Ref}: Gesamtgenauigkeit abgeleitet aus pixelweisen Vergleich der Referenzlabels pro Superpixel zu den pixelbasierten Referenzdaten.

bieten zu nahezu quadratischen Formen (vgl. Abb. 5.2d). Die geringen Grauwertunterschiede zwischen den einzelnen Strukturen werden hierbei vernachlässigt unter der Prämisse einer möglichst kompakten Form. Ausgeprägte Grauwertkanten, wie sie z.B. an dem Übergang zwischen der Brücke und dem Gewässer in der Abbildung 5.2 auftreten, werden hingegen in allen Segmentierungen gleichermaßen gut berücksichtigt (vgl. Abb. 5.2b bis 5.2e). In homogenen Gebieten, wie z.B. der Wasseroberfläche, nähern sich die Superpixel aller Segmentierungen, mit Ausnahme der Kompaktheit 1, einer quadratischen Form an. Insgesamt lässt sich festhalten, dass die Segmentierungen in den Abbildungen 5.2c bis 5.2e eine vergleichbar gute Qualität aufweisen. Dennoch sind Unterschiede zu erkennen. Beispielsweise spiegeln sich die Qualitätsunterschiede für die Klasse *Fahrzeug* in den Abbildungen wider. Bei einer Kompaktheit von 1 werden die Autos zwar teilweise in separaten Superpixeln erfasst, jedoch führt die hohe Sensitivität der Segmentierung gegenüber Grauwertunterschieden dazu, dass mitunter die Fahrzeuge mitsamt ihrer Schatten in einem Superpixel zusammengefasst sind oder Fahrzeuge in zwei Superpixeln aufgehen (vgl. Abb. 5.2b). Bei Erhöhung der Kompaktheit orientieren sich die Superpixel zunehmend weniger an den Grauwertkanten der Fahrzeuge. Infolgedessen werden Fahrzeugpixel bei einer Kompaktheit von 80 vermehrt mit angrenzenden Versiegelungsflächen in einem Superpixel

zusammengefasst (vgl. Abb. 5.2d). Das beste Ergebnis bezüglich der Klasse *Fahrzeug* erzielt die Segmentierung auf der Grundlage der primären Bildinformationen (vgl. Abb. 5.2e). Dies lässt sich vor allem damit erklären, dass Fahrzeuge zum einen nicht im nDOM enthalten sind und sich zum anderen auch nicht in Bezug auf den NDVI von ihrer Umgebung (typischerweise *Versiegelung*) abgrenzen. Den meisten der in Abbildung 5.2e dargestellten Fahrzeugen sind individuelle Superpixel zugeordnet, die sich hinsichtlich ihrer Begrenzungen an den Umringen der Fahrzeuge orientieren. Die Verbesserung wird insbesondere bei den zwei linken Fahrzeugen offensichtlich, die bei Verwendung der sekundären Bildinformationen als Eingangsdaten in mehrere Teile aufgespalten den umliegenden Superpixeln zugeordnet werden (vgl. Abb. 5.2c und 5.2d) und bei Verwendung der primären Bildinformation ein eigenes Superpixel bilden (vgl. Abb. 5.2e). Neben den aufgezeigten geringfügigen Unterschieden lassen sich keine immanenten Vor- und Nachteile bzgl. der Eingangsdaten aus den Segmentierungsergebnissen ableiten, die eindeutig für einen bestimmten Satz an Eingangsdaten sprechen. Der rein visuelle Vergleich der Segmentierungen im Testgebiet Hameln vermittelt jedoch einen etwas homogeneren Gesamteindruck für eine Kompaktheit von 40, der sich – wenn auch nicht signifikant – in der besten Referenzgenauigkeit bestätigt. Die Berücksichtigung der Höheninformation sowie des NDVI bietet keinen Mehrwert. In der Abbildung 5.2 ist die Abgrenzung von Erhebungen der Erdoberfläche, wie z.B. dem Brückenbauwerk, als auch von Vegetationsobjekten, wie z.B. den Baumreihen, für beide Eingangsdaten gleich gut realisiert.

Größe der Superpixel Der zweite Teil der Untersuchungen widmet sich dem Einfluss der Größe der Superpixel auf die Klassifikationsergebnisse. In der Tabelle 5.8 sind die Gesamtgenauigkeitsmaße für die Bodenbedeckung und die Landnutzung aufgetragen, die jeweils für Superpixel der Größen 100, 400 und 2.500 Pixel mittels des iterativen Inferenzalgorithmus erzielt wurden. Darüber hinaus sind je Superpixelgröße die erreichbare Referenzgenauigkeit sowie die Rechenzeit für das Training und die Inferenz angegeben.

		$CRF_{iter,100}$		$CRF_{iter,400}$		$CRF_{iter,2.500}$	
		Klass.	Ref.	Klass.	Ref.	Klass.	Ref.
Bodenbedeckung	G [%]	81,9	96,0	82,2	93,3	80,2	87,2
	κ [%]	76,7	94,8	77,0	91,4	74,3	83,5
Landnutzung	G [%]	78,5	–	78,4	–	78,6	–
	κ [%]	73,8	–	73,6	–	73,8	–
Rechenzeit	Training [s]	2411,95		277,84		135,01	
	Inferenz [s]	5373,54		559,98		133,51	

Tabelle 5.8.: Gesamtgenauigkeit (G) [%] und Kappa-Index (κ) [%] abgeleitet aus pixelbasierter Evaluation für die Bodenbedeckungs- und objektbasierter Evaluation für die Landnutzungsklassifikation (Klass.) für das Testgebiet Hameln unter Anwendung der iterativen Inferenzprozedur auf Basis von Superpixeln der Größe 100 ($CRF_{iter,100}$), 400 ($CRF_{iter,400}$) und 2.500 Pixeln ($CRF_{iter,2.500}$) und Kompaktheit 20. Ref.: Genauigkeitsmaße abgeleitet aus dem pixelweisen Vergleich der Referenzlabels pro Superpixel zu den pixelbasierten Referenzdaten. Zusätzlich ist die Rechenzeit [s] pro Block separat für Training und Inferenz angegeben.

In der Tabelle 5.8 fällt zunächst auf, dass sich die Genauigkeitsmaße der Landnutzungsklassifi-

kation für unterschiedliche Größen der Superpixel kaum unterscheiden. Demnach hat die Größe der Superpixel nur einen geringen Einfluss auf die Klassifikation der Landnutzung. In dem hier untersuchten Rahmen lässt sich somit feststellen, dass weder ein hoher geometrischer Detailgrad noch eine starke Glättung der Bodenbedeckungsinformation einen Einfluss auf die Klassifikation der Landnutzung hat.

Die Rechenzeit reduziert sich um den Faktor 9 bis 10 bei dem Übergang von der Größe 100 auf 400 Pixel und weiter um den Faktor 2 bis 4 bei Vergrößerung der Superpixel von 400 auf 2.500 Pixel. Der größere Faktor bezieht sich jeweils auf die Rechenzeit der Inferenz, d.h. hier wird ein im Vergleich zum Training größerer Rechenzeitgewinn erzielt. Bei den Angaben in der Tabelle 5.8 ist zu beachten, dass die Prozessierung eines Blocks kachelweise erfolgt, wobei die Größen der Kacheln in Abhängigkeit der Größe der Superpixel variieren. Die Erhöhung der Anzahl der Bildprimitive im Zuge einer feineren Segmentierung erfordert die Anpassung der Kachelgrößen, d.h. keine Unterteilung des Blocks für die Größe 2.500 Pixel, Einteilung in 4 Kacheln für die Größe 400 Pixel und in 25 Kacheln für die Größe 100 Pixel. Insgesamt führt der Übergang von 100 zu 400 Pixeln zu einer immensen Beschleunigung des Verfahrens von Stunden hin zu wenigen Minuten pro Bearbeitungseinheit (Block). Mit einer Reduzierung der Rechenzeit gewinnt das Verfahren zunehmend an Attraktivität für den praktischen Einsatz.

Der Anstieg der Referenzgenauigkeit mit kleiner Superpixelgröße spiegelt sich nicht in der Genauigkeit der Klassifikationsergebnisse wider. Die höchste Genauigkeit wird für eine Größe von 400 Pixeln erzielt (Gesamtgenauigkeit 82,2 %, Kappa-Index 77,0 %), die sich jedoch nicht signifikant von der Genauigkeit für die Größe 100 Pixel unterscheidet (Gesamtgenauigkeit 81,9 %, Kappa-Index 76,7 %). Folglich erzielt die Bodenbedeckungsklassifikation für Superpixel der Größen 100 und 400 Pixel das gleiche Genauigkeitsniveau, obwohl – wie durch die Referenzgenauigkeit belegt – theoretisch eine höhere Genauigkeit für Superpixel der Größe 100 Pixel möglich wäre. Von diesem Niveau fällt die Genauigkeit, die für Superpixel der Größe 2.500 Pixel erzielt wird, mit ca. 2 % bzgl. der Gesamtgenauigkeit und ca. 3 % bzgl. des Kappa-Index etwas deutlicher ab. Trotz eines insgesamt ähnlichen Genauigkeitsniveaus, weichen die Ergebnisse in Bezug auf einzelne Klassen deutlich voneinander ab. Diese Unterschiede werden im Folgenden näher analysiert.

In dem Diagramm in Abbildung 5.3 sind die Gesamtgenauigkeit und der Kappa-Index aus der Tabelle 5.8 sowie zusätzlich die Qualität pro Klasse für unterschiedliche Größen der Superpixel aufgetragen. Darüber hinaus ist die Abweichung zur Referenzgenauigkeit angegeben. Die Qualität der Klassen *Gebäude* und *Versiegelung* ist für die Größen 100 und 400 Pixel jeweils auf einem ähnlichen Niveau. Die Qualitätswerte für die Größe 2.500 Pixel weichen hiervon deutlich ab und zeigen Einbußen von ca. 6 % gegenüber den Ergebnissen für Superpixel der Größe 400 Pixel. Für die Klassen *Gras* und *Baum* zeigt sich ein ähnliches Bild, jedoch mit dem Unterschied, dass die Qualitätswerte der Größe 2.500 Pixel nur um ca. 1 % von den Ergebnissen der Größen 100 und 400 Pixel abweichen. Anders verhält es sich bei der Klasse *Boden*. Hier erzielen die Größen 400 und 2.500 Pixel ein ähnliches Genuigkeitsniveau und die Qualität der Größe 100 Pixel ist um ca. 3 % niedriger. Für die Klassen *Wasser*, *Fahrzeug* und *Sonstiges* hat die Vergrößerung der Superpixel eine kontinuierliche Verschlechterung der Qualität zur Folge. Für die Klasse *Wasser* verschlechtern sich die Ergebnisse lediglich um max. 4 %. Die weit größeren Einbußen treten bei den Klassen *Fahrzeug* und *Sonstiges* auf. Hier lässt sich bereits aus den Referenzgenauigkeiten darauf schließen, dass deutliche Genauigkeitseinbußen zu erwar-

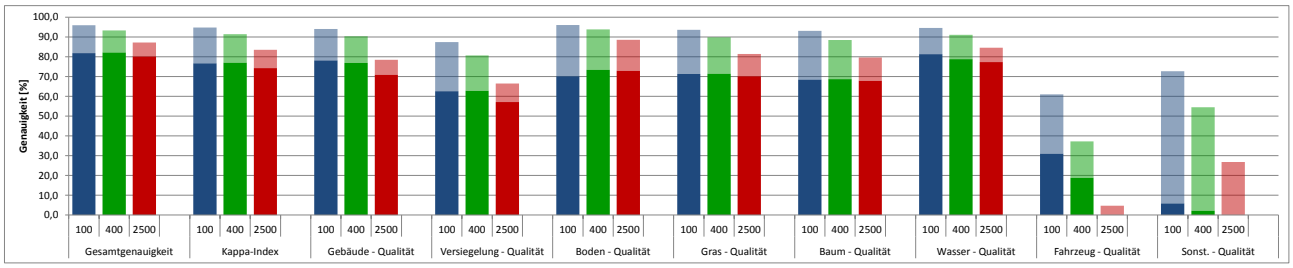
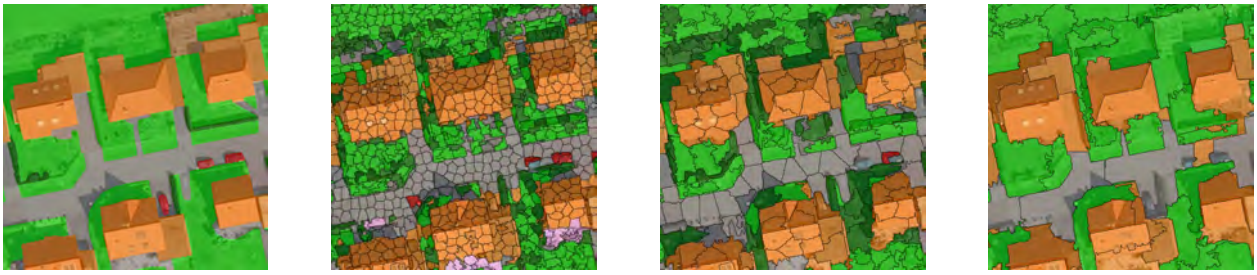


Abbildung 5.3.: Gesamtgenauigkeit [%], Kappa-Index [%] und Qualität [%] pro Klasse abgeleitet aus pixelbasierter Evaluation der Klassifikationsergebnisse für die Bodenbedeckungsklassen *Gebäude*, *Versiegelung*, *Boden*, *Gras*, *Baum*, *Wasser*, *Fahrzeug* und *Sonstiges* für das Testgebiet Hameln unter Anwendung der iterativen Inferenzprozedur auf Basis von Superpixeln der Größe 100 (blau), 400 (grün) und 2.500 Pixel (rot). Die transparenten Bereiche kennzeichnen die Differenz zur maximal zu erreichenden Referenzgenauigkeit.

ten sind, so betragen die Einbußen zwischen den Größen 100 und 2.500 Pixel ca. 56 % für die Klasse *Fahrzeug* und ca. 46 % für die Klasse *Sonstiges*. Die Klassifikation schöpft für die Größen 100 und 400 Pixel nur einen Bruchteil der erreichbaren Qualität für diese Klassen aus, d.h. jeweils die Hälfte für die Klasse *Fahrzeug* und zwischen 4 % (Größe 400) und 8 % (Größe 100) für die Klasse *Sonstiges*. Bei Verwendung einer Superpixelgröße von 2.500 Pixeln lassen sich diese Klassen gar nicht mehr von den anderen unterscheiden.

In der Abbildung 5.4 sind die Ergebnisse der iterativen Inferenzprozedur auf Basis der drei untersuchten Superpixelgrößen am Beispiel einer urbanen Szene dargestellt. Die Abbildungen haben zum Ziel, Unterschiede in den Ergebnissen aufzuzeigen und mögliche Ursachen für die zuvor genannten Genauigkeitsabweichungen zu ergründen. Die qualitative Beurteilung der Ergebnisse erfolgt anhand der in Abbildung 5.4a dargestellten pixelbasierten Referenzdaten.



(a) Referenz (b) Größe = 100 (c) Größe = 400 (d) Größe = 2.500

Abbildung 5.4.: Orthophoto einer urbanen Szene im Testgebiet Hameln, überlagert mit den pixelbasierten Referenzdaten (a) und den superpixelbasierten Ergebnissen der iterativen Inferenzprozedur auf der Basis von Superpixeln der Größe 100 (b), 400 (c) und 2.500 Pixel (d). Die Farben markieren die Zugehörigkeit zur Bodenbedeckungsklasse: *Geb.* (orange), *Vers.* (grau), *Bod.* (braun), *Gras* (hellgrün), *Baum* (dunkelgrün), *Was.* (blau), *Fhzg.* (rot) und *Sonst.* (rosa).

Auf den ersten Blick ist ersichtlich, dass der Detailgrad (bzgl. kleinräumiger Strukturen) mit zunehmender Größe der Superpixel abnimmt. Dieser Zusammenhang lässt sich besonders gut am Beispiel der Klasse *Fahrzeug* beobachten. Wird die Szene mit einer Größe von 100 Pixeln segmentiert,

sind einzelne Autos in individuellen Superpixeln erfasst (vgl. Abb. 5.4b). Im Zuge der Vergrößerung der Superpixel werden die Auto-Pixel den sie umgebenden Bodenbedeckungsarten hinzugefügt und sind infolgedessen nicht mehr durch ein individuelles Superpixel repräsentiert (vgl. Abb. 5.4c und 5.4d). Ein Klassifikator bestimmt in Abhängigkeit der Zusammensetzung eines Superpixels mit hoher Wahrscheinlichkeit das Klassenlabel der dominierenden Bodenbedeckungsart, d.h. mit dem größten Anteil an Pixeln innerhalb des Superpixels. Die Auswirkungen sind besonders gut in der Abb. 5.4d zu beobachten. In dem hier dargestellten Klassifikationsergebnis auf Basis von Superpixeln der Größe 2.500 Pixel sind Autos nicht mehr enthalten. Diese Problematik betrifft auch die Klasse *Sonstiges*, die in den verwendeten Testgebieten ebenfalls überwiegend kleinräumige Strukturen repräsentiert.

Neben einem höheren Detailgrad tragen kleinere Superpixel zu einer exakteren geometrischen Abgrenzung der Bodenbedeckungssegmente bei. Dies wird am Beispiel der Gebäude in Abbildung 5.4 deutlich. Wie in der Abbildung 5.4d zu sehen, sind die Begrenzungen der Superpixel der Größe 2.500 Pixel in vielen Fällen nicht konsistent zu den Umringen der Gebäude. Dieser Effekt resultiert u.a. aus den Ungenauigkeiten des zugrundeliegenden nDOM sowie ähnlichen spektralen Eigenschaften zwischen den Gebäuden und den angrenzenden Bodenbedeckungen, insbesondere den Versiegelungsflächen. Demgegenüber verbessert sich die geometrische Abgrenzung der Gebäude mit kleinerer Größe der Superpixel (vgl. Abb. 5.4b und 5.4c). Sind bei einer Größe von 400 Pixeln noch vereinzelt Abweichungen von den Gebäudeumrissen zu erkennen, so stimmen diese bei einer Größe von 100 Pixeln weitestgehend mit einer Kante der Superpixel überein. Der Vorteil der exakteren geometrischen Abgrenzung wirkt sich allerdings nur marginal auf die Klassifikationsgenauigkeit aus. Bei einem visuellen Vergleich der Klassifikationsergebnisse für eine Größe von 100 (vgl. Abb. 5.4b) und 400 Pixeln (vgl. Abb. 5.4c) fällt auf, dass die Flächen, die als Gebäude klassifiziert sind, in beiden Ergebnissen trotz unterschiedlicher Klassifikationseinheiten weitgehend übereinstimmen. Diese Beobachtung wird auch durch die Genauigkeitsmaße für die Klassifikation bestätigt, die für die Klasse *Gebäude* kaum Unterschiede aufweisen. Auf der anderen Seite sind bei einer feinen Segmentierung kleine Strukturen, die eine ähnliche Charakteristik wie Gebäude aufweisen, ebenfalls in individuellen Superpixeln erfasst. Infolgedessen kommt es vermehrt zu isolierten Fehlklassifikationen von Gebäuden, speziell in den Nutzungsarten *Wohnbaufläche* und *Sonstige Bebauung* (vgl. Abb. 5.4b).

Insgesamt vermittelt das Klassifikationsergebnis bei einer Größe von 100 Pixeln in Abbildung 5.4b im Vergleich zu Abbildung 5.4c und 5.4d einen verrauschteren Gesamteindruck. Mit Vergrößerung der Superpixel werden die Ergebnisse zunehmend geglättet, sodass nur noch vereinzelt isolierte Fehlklassifikationen auftreten (vgl. Abb. 5.4c und 5.4d). Dieser Glättungseffekt ist besonders vorteilhaft für Bodenbedeckungsarten, die typischerweise großräumige, homogene Flächen beschreiben, wie z.B. die Klasse *Boden*.

5.2.2.4. Diskussion

Die Analyse zeigt, dass sowohl die Kompaktheit als auch die Eingangsdaten nur einen geringen Einfluss auf die Referenzgenauigkeit haben. Die Auswirkungen auf die Klassifikationsergebnisse sind nicht signifikant. Die Annahme, dass sich eine Segmentierung auf Basis der sekundären Bildinformationen besser an die spektralen Übergänge anpasst, konnte nicht bestätigt werden. Unter den untersuchten

Parametereinstellungen liefert die Segmentierung mit einer Kompaktheit von 40 unter Verwendung der sekundären Bildinformationen als Eingangsdaten das – wenn auch nicht signifikant – beste Ergebnis bezogen auf die Referenzgenauigkeit und vermittelt zudem einen etwas homogeneren Gesamteindruck gegenüber der Segmentierung unter Verwendung der Standardwerte (Kompaktheit 20, sekundäre Bildinformationen als Eingangsdaten), sodass diese Konfiguration für die Experimente in Kapitel 5.2.5 verwendet wird.

Darüber hinaus haben die Experimente gezeigt, dass sich die Größe der Superpixel je nach Bodenbedeckungsart unterschiedlich auf das Klassifikationsergebnis auswirkt. Kleinere Superpixel bieten den Vorteil, dass sie eine exaktere geometrische Abgrenzung der Bodenbedeckungssegmente ermöglichen. Zudem werden kleinräumige Strukturen, wie z.B. Autos oder kleine Bäume, in individuellen Superpixeln erfasst, wodurch ein hoher Detailgrad der Klassifikationsergebnisse erzielt wird. Im Gegensatz dazu sind in größeren Superpixeln teilweise mehrere Bodenbedeckungsklassen enthalten, was zu Ungenauigkeiten führt. Auf der anderen Seite wird durch die Verwendung von Superpixeln anstelle von Pixeln ein Glättungseffekt erzielt, von dem besonders jene Bodenbedeckungsarten profitieren, die typischerweise großräumige, homogene Flächen bedecken, wie es z.B. für die Klasse *Boden* der Fall ist. Der Glättungseffekt prägt sich mit Vergrößerung der Superpixel stärker aus. Die Analyse hat ergeben, dass eine Größe von 400 Pixeln einen guten Kompromiss zwischen einem hohen Detailgrad, einer exakten geometrischen Abgrenzung, einem zweckmäßigen Glättungsgrad sowie einer sinnvollen Rechenzeit darstellt. Diese Größe wird für die nachfolgenden Experimente in den Kapiteln 5.2.3 bis 5.2.5 verwendet.

Ferner haben die Untersuchungen ergeben, dass die Größe der Superpixel im hier untersuchten Rahmen so gut wie keinen Einfluss auf die Landnutzungsklassifikation hat. Die Klassifikation der Landnutzung profitiert somit weder von einem hohen geometrischen Detailgrad noch einer starken Glättung der Bodenbedeckungsinformation.

5.2.3. Merkmale

5.2.3.1. Zielsetzung

Die Qualität der Klassifikationsergebnisse wird maßgeblich durch die verwendeten Merkmale beeinflusst. Die Auswahl geeigneter Merkmale erfordert Kenntnisse über die allgemeinen Eigenschaften der verschiedenen Bodenbedeckungs- bzw. Landnutzungsarten sowie über deren Relationen zueinander. Eine manuelle Auswahl von Merkmalen birgt die Gefahr, dass relevante Merkmale vergessen bzw. stark korrelierte Merkmale verwendet werden. Aus diesem Grund werden geeignete Merkmale aus einem großen Pool (siehe Kap. 5.2.3) auf Basis eines Wichtigkeitsmaßes selektiert, welches die Relevanz der extrahierten Merkmale für die Klassifikation objektiv beurteilt. Zu diesem Zweck wird der individuelle Einfluss der Merkmale auf das Klassifikationsergebnis analysiert. Die Beurteilung der Relevanz der Merkmale erfolgt auf Basis des RF Wichtigkeitsmaßes [Breiman, 2001] (siehe Kap. 3.3), das bereits in diversen Arbeiten zur Beurteilung der Relevanz von Merkmalen herangezogen wurde, wie z.B. in [Chehata et al., 2011; Niemeyer et al., 2014].

Ziel der Untersuchung ist es, einen minimalen Satz an Merkmalen für die weitere Prozessie-

rung auszuwählen, der die wesentlichen Merkmale für die Klassifikation enthält. Unter Verwendung dieses Merkmalssatzes würde jedes zusätzliche Merkmal aus dem Pool zu keiner signifikanten Verbesserung der Genauigkeit führen. Die Beschränkung auf die notwendige Mindestanzahl bietet die Vorteile, dass sich zum einen die Rechenzeit reduziert und zum anderen den ggf. mit einer hohen Anzahl an Merkmalen einhergehenden Genauigkeitseinbußen vorgebeugt wird (vgl. Ausführungen zu dem Hughes-Phänomen in Kapitel 3.2).

Neben der Art der Merkmale wirken sich auch zwei weitere Faktoren auf das Klassifikationsergebnis aus, und zwar wie die Knotenmerkmale normalisiert und wie sie in den Merkmalsvektoren der Kanten kombiniert werden. Daher gilt es, den Einfluss dieser Faktoren zu untersuchen und hierfür die bestmöglichen Einstellungen zu wählen.

5.2.3.2. Strategie

Die Untersuchung erfolgt separat für jedes Potential des Zwei-Ebenen-CRF-Modells, deren Berechnung sich im Falle des Assoziations- und des paarweisen Interaktionspotentials auf die bildbasierten, dreidimensionalen und geometrischen Merkmale und im Falle des Potentials höherer Ordnung auf die Kontextmerkmale stützt. Zu diesem Zweck werden separat für jedes Potential die zugrundeliegenden Merkmale hinsichtlich ihrer Relevanz analysiert. In der Literatur werden die Knoten- und Kantenmerkmale jedoch häufig als eine Einheit betrachtet, wobei sich die den Kanten zugehörigen Merkmalsvektoren aus den Merkmalen benachbarter Knoten zusammensetzen. Aus diesem Grund erfolgt an dieser Stelle keine individuelle Auswahl von Merkmalen für die Bestimmung der Knoten- und der räumlichen Kantenpotentiale, sehr wohl aber für die Bestimmung des semantischen Potentials höherer Ordnung. Die Untersuchung gliedert sich damit in zwei Teile, zum einen in die Analyse der bildbasierten, dreidimensionalen und geometrischen Merkmale, die während der Inferenz konstant bleiben, und zum anderen in die Analyse der Kontextmerkmale, die sich während der Inferenz ändern. Bei der Untersuchung einer Merkmalsgruppe haben die anderen Merkmale keinen Einfluss auf die Klassifikation, d.h. die zugehörigen Potentiale werden im Rahmen der Klassifikation nicht berücksichtigt.

Da sich die Relevanz der Merkmale unter Umständen auch in Abhängigkeit des zur Normalisierung verwendeten Quantils ε oder der Art der Zusammensetzung der Interaktionsmerkmale ändert, wird die Qualität der Klassifikationsergebnisse zunächst im Hinblick auf den Einfluss dieser Parameter analysiert. Zu diesem Zweck wird die im Rahmen der Klassifikation erzielte Genauigkeit zunächst für unterschiedlich skalierte Merkmalssätze analysiert. Der Unterschied in der Skalierung resultiert aus der Wahl der Quantile, welche die oberen und unteren Intervallgrenzen zur Skalierung der Merkmale definieren (siehe Kap. 5.1.3). Da die Knoten- und Kantenmerkmale gemeinsam betrachtet werden, sich aber der Einfluss der Normalisierung für die verschiedenen Zusammensetzungen der Kantenmerkmale unterschiedlich ausprägt, wird diese Untersuchung für zwei verschiedene Varianten der Kantenmerkmale durchgeführt: zum einen der Aneinanderreihung und zum anderen der elementweisen Differenzbildung der Knotenmerkmale. Die Untersuchung basiert auf der Gesamtheit der Merkmale. Die Auswahl geeigneter Merkmale erfolgt in einem zweiten Schritt, d.h. nach Festlegung guter Parameter für die Normalisierung der Merkmale sowie einer geeigneten Art der Zusammensetzung der Kantenmerkmale separat für jede Ebene.

Für die Auswahl geeigneter Merkmale wird das RF Wichtigkeitsmaß pro Merkmal eines jeden Potentials bestimmt, d.h. getrennt für jedes Potential des Zwei-Ebenen-CRF-Modells. Dies erfolgt unter Verwendung der zuvor bestimmten Parameter. Anschließend werden die Merkmale pro Potential entsprechend ihres Wichtigkeitsmaßes sortiert, wobei das wichtigste Merkmal an erster Position steht. Zur Bestimmung der Knoten- und räumlichen Kantenpotentiale wird zwar ein identischer Satz an Merkmalen verwendet, dennoch wird die Relevanz der Merkmale getrennt analysiert. Die Wichtigkeitsmaße für die Knoten- und Kantenmerkmale werden anschließend in einem Ranking zusammengeführt und bilden die Grundlage für die Selektion eines gemeinsamen Satzes an Merkmalen. In dem gemeinsamen Ranking sind die Knoten- und Kantenmerkmale entsprechend ihrer Wichtigkeit sortiert, wobei jedes Merkmal nur einmal enthalten ist. Die zwei Wichtigkeitsmaße pro Merkmal, die bei aneinandergefügten Merkmalsvektoren entstehen, werden vorab aufsummiert. Um die notwendige Anzahl an Merkmalen zu ermitteln, gilt es den Einfluss der gewählten Merkmale auf die Gesamtgenauigkeit zu analysieren. Zu diesem Zweck werden die Merkmale entsprechend ihrer Relevanz – beginnend mit dem Wichtigsten – sukzessive der Klassifikation hinzugefügt, d.h. die Klassifikation wird unter Verwendung der N_F wichtigsten Merkmale der sortierten Liste wiederholt durchgeführt, wobei N_F von 1 bis 30 variiert. Mit Erreichen des Sättigungspunktes, d.h. der Konvergenz der Gesamtgenauigkeit, führt jedes zusätzliche Merkmal zu keiner weiteren signifikanten Verbesserung der Gesamtgenauigkeit.

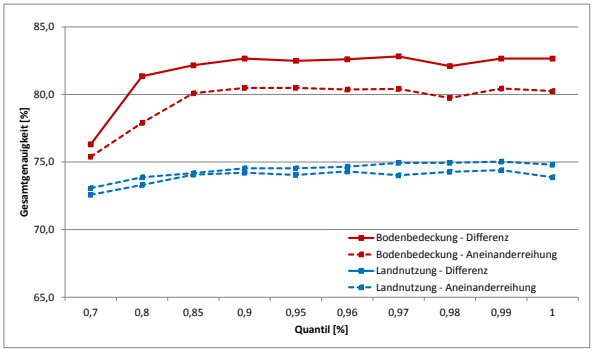
Die Klassifikation erfolgt in beiden Untersuchungen unter Anwendung der in Tabelle 5.4 aufgetragenen Standardparameter.

5.2.3.3. Beschreibung der Ergebnisse

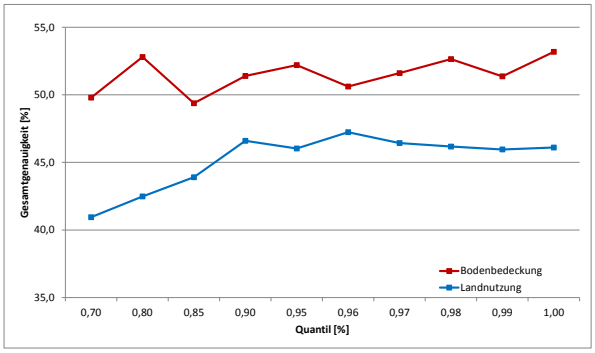
Die dargestellten Ergebnisse beziehen sich auf das Testgebiet Schleswig. Die Ergebnisse in Hameln zeigen ein weitgehend ähnliches Bild und werden daher an dieser Stelle zur Vermeidung von Redundanz nicht angeführt. Auf Unterschiede in den Ergebnissen der Testgebiete wird im Rahmen der Analyse hingewiesen.

Die erste Untersuchung dient der Analyse des Einflusses des Typs der Kantenmerkmale sowie des Parameters ε zur Normalisierung der Merkmale. Die Abbildung 5.5 zeigt die erzielten Gesamtgenauigkeiten bei Variation dieser Parameter für das Assoziations- und das räumliche Interaktionspotential sowie das semantische Potential höherer Ordnung, getrennt für die Bodenbedeckungs- und Landnutzungsebene. Bei der Interpretation der Diagramme ist zu beachten, dass die y-Achsen unterschiedlich skaliert sind.

In Bezug auf die Art der Kantenmerkmale zeigen sich in Abbildung 5.5a insgesamt größere Unterschiede bei der Klassifikation der Bodenbedeckung. Hier weichen die Gesamtgenauigkeiten um ca. 2 % voneinander ab. Die bessere Genauigkeit wird dabei für die Kantenmerkmale erzielt, die sich aus Differenzen der Knotenmerkmale ergeben. Im Gegensatz dazu sind die Unterschiede bei der Klassifikation der Landnutzung deutlich geringer (unter 1 %). Für die aneinandergereihten Knotenmerkmale werden zwar etwas bessere Gesamtgenauigkeiten erzielt, die sich aber nicht signifikant von denen der anderen Variante unterscheiden. Daraus folgert, dass für die Landnutzungsklassifikation die Wahl des Typs der Kantenmerkmale von untergeordneter Bedeutung ist. Der Unterschied bleibt bei beiden Klassifikationen über alle getesteten Quantile, mit Ausnahme des 70 %-Quantils, weitgehend konstant. In



(a) Klassifikationsgenauigkeit Assoziations- und räumliches Interaktionspotential



(b) Klassifikationsgenauigkeit semantisches Potential höherer Ordnung

Abbildung 5.5.: Gesamtgenauigkeit [%] der Klassifikation der Bodenbedeckung (rot) und Landnutzung (blau) auf Grundlage des Assoziations- und räumlichen Interaktionspotentials (a) sowie des semantischen Potentials höherer Ordnung (b) (jeweils ausschließlich) bei Variation der Kantenmerkmale (Differenz: durchgezogene Linie; Aneinanderreihung: gestrichelte Linie) und der Quantile ε zur Normalisierung im Wertebereich $[70\,\%, 100\,\%]$.

dem Testgebiet Hameln zeigt sich ein etwas anderes Bild. Die Unterschiede betragen hier insgesamt weniger als 1 %. Hiervon ausgenommen sind die Ergebnisse für ein Quantil von 70 %, die aber von der Genauigkeit in beiden Testgebieten deutlich abfallen. Diese Größenordnung wird in der weiteren Analyse nicht näher betrachtet. Für den Testdatensatz Hameln wechselt der Vorteil in Bezug auf die Genauigkeit zwischen den beiden Varianten der Kantenmerkmale bei Variation der Quantile. Insgesamt lässt sich auf Basis der Gesamtgenauigkeit keine eindeutige Präferenz für einen Typ erkennen. Die klassenspezifischen Genauigkeiten für die Landnutzungsklassen weisen jedoch auf Unterschiede hin. Die Differenz der Knotenmerkmale ist beispielsweise besonders geeignet zur Unterscheidung der Landnutzungsklassen *Wohnbaufläche*, *Sonstige Bebauung*, *Verkehrsweg* und *Weg*. Demgegenüber sind aneinandergereihte Knotenmerkmale insbesondere zur Bestimmung der Klassen *Ackerland*, *Grünland*, *Wald* und *Wasser* geeignet. Diese Art des Kantenmerkmals ist besonders vorteilhaft für die Klassen *Ackerland* und *Grünland*. In dem Testgebiet Schleswig profitieren nahezu alle Klassen gleichermaßen von je einer Variante der Kantenmerkmale. Es prägt sich in diesen Ergebnissen eine leichte Präferenz für eine Aneinanderreihung der Merkmalsvektoren im Falle der Landnutzungsklassifikation und der Differenzbildung im Falle der Bodenbedeckungsklassifikation aus.

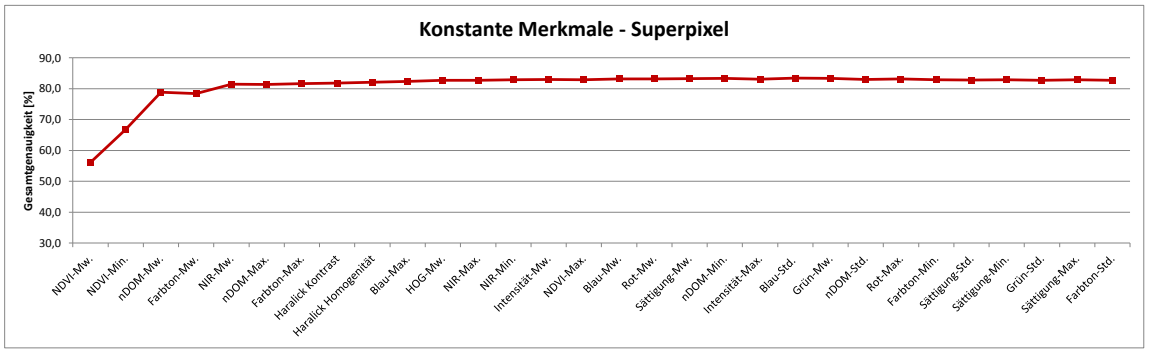
Die Änderung des Quantils ε hat insgesamt nur geringe Variationen der Gesamtgenauigkeiten von maximal 1 % zur Folge (ausgenommen der Ergebnisse für die Quantile 70 % und 80 %). Folglich hat das Quantil zur Normalisierung der Merkmale kaum Auswirkungen auf die Klassifikationsergebnisse, sofern ein Quantil von größer als 80 % gewählt wird. Die höchste Genauigkeit wird für die Bodenbedeckung bei einem 97 %-Quantil unter Verwendung der Differenzen als Merkmalsvektoren sowie für die Landnutzung bei einem 99 %-Quantil unter Verwendung aneinandergereihter Merkmalsvektoren erzielt. Die Unterschiede zu anderen Parametern sind jedoch nicht signifikant. Bei der Analyse der Qualität pro Klasse ergibt sich ein etwas differenzierters Bild. Bestimmte Klassen profitieren von einem hohen Wert des Quantils. Dies betrifft bzgl. der Bodenbedeckung insbesondere die Klasse *Wasser* (+6,7 %) und bzgl. der Landnutzung die Klassen *Weg* (+4,7 %) und *Gewässer* (+5,2 %), wobei nicht in allen Fällen die höchste Qualität bei einem Quantil von 100 % erzielt wird. In allen drei Fällen handelt es sich um

die Klassen mit einer vergleichsweise geringen Anzahl an Trainingsbeispielen im Testgebiet Schleswig. Zudem weisen die Klassen *Wasser* und *Gewässer* ein sehr charakteristisches Erscheinungsbild auf, wodurch sie sich von den übrigen Klassen deutlich abgrenzen. Zum Beispiel resultiert aus der nahezu vollständigen Überdeckung mit Wasser ein geringer NDVI, der am unteren Rand des Histogramms der Merkmalswerte angesiedelt ist. Bei insgesamt nur wenigen Objekten für diese Klassen kann ein zu kleiner Wert für ε dazu führen, dass die zugehörigen Merkmalswerte im Rahmen der Normalisierung abgeschnitten, d.h. auf die Intervallgrenzen abgebildet werden. Hierdurch gehen unter Umständen entscheidende Informationen zur Trennung dieser Klassen verloren. Nichtsdestotrotz zeigt sich auch für diese beiden Klassen, dass die Eliminierung von Ausreißern an den Rändern bis zu einem gewissen Grad vorteilhaft ist. So treten die höchsten Genauigkeiten bei einem Quantil von 99 % auf, d.h. 1 % der Merkmalswerte werden jeweils am oberen und unteren Rand auf die Intervallgrenzen abgebildet.

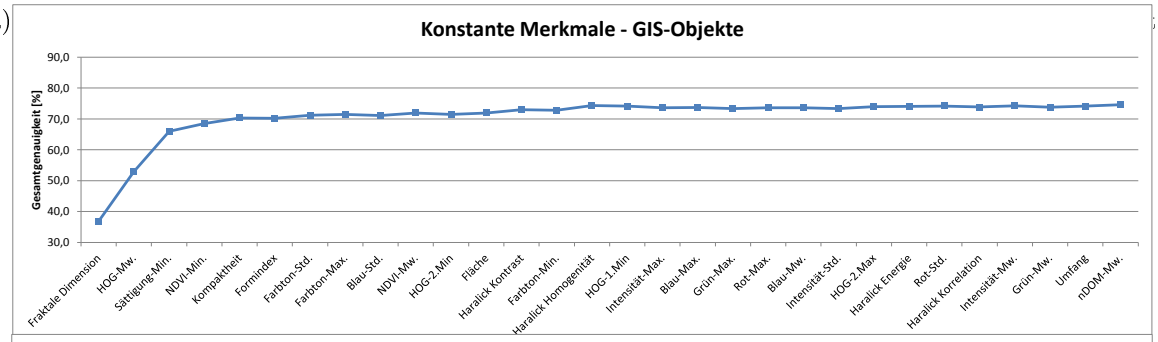
In Bezug auf die Gesamtgenauigkeit, die bei der Klassifikation der semantischen Potentiale höherer Ordnung erzielt wird, zeigen sich in Abbildung 5.5b deutlich größere Schwankungen. Bei Variation des Quantils ε treten zwischen den Ergebnissen große Sprünge auf. Eine eindeutige Tendenz ist bei der Klassifikation der Bodenbedeckung nicht zu erkennen. Die Ergebnisse für die Landnutzung lassen schon eher auf eine kontinuierliche Entwicklung schließen. Die höchste Genauigkeit erzielt ein Quantil von 100 % für die Bodenbedeckung und ein Quantil von 96 % für die Landnutzung. Die Unterschiede zu anderen Parametern sind jedoch nicht signifikant, beispielsweise wird für die Bodenbedeckung bei einem 80 %-Quantil ein ähnliches Genauigkeitsniveau erreicht. Aus den relativ großen Schwankungen lässt sich einerseits auf einen relativ großen Einfluss des gewählten Quantils auf die Klassifikation des semantischen Potentials höherer Ordnung schließen. Andererseits können die Schwankungen aber auch aus der hohen Unsicherheit des Klassifikators resultieren (Gesamtgenauigkeit von maximal 53 %). Für das folgende Experiment werden die oben genannten optimalen Parametereinstellungen bzgl. des Typs der Kantenmerkmale (Differenzen der Knotenmerkmale für Bodenbedeckung, Aneinanderreihung der Knotenmerkmale für Landnutzung) und des Quantils zur Normalisierung der Merkmale (Assoziations- und räumliches Interaktionspotential: 97 % für Bodenbedeckung, 99 % für Landnutzung; semantisches Potential höherer Ordnung: 100 % für Bodenbedeckung, 96 % für Landnutzung) verwendet.

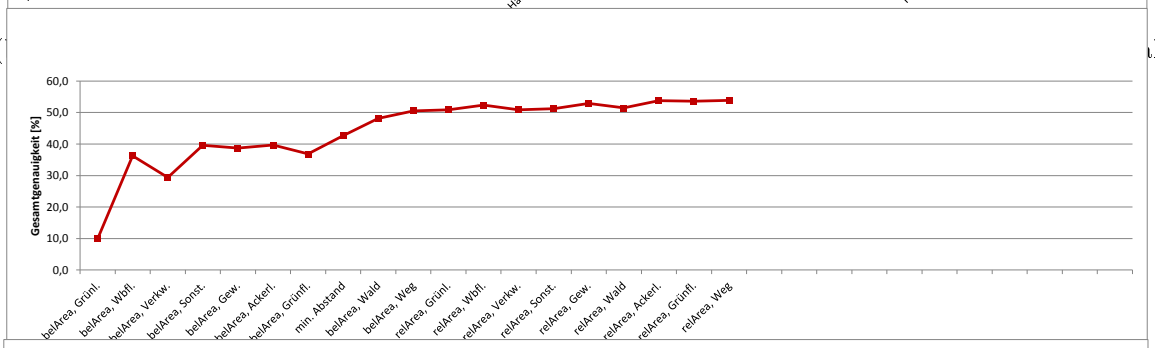
Die Abbildung 5.6 zeigt den Einfluss der Merkmale auf die Klassifikationsgenauigkeit der Assoziations- und der räumlichen Interaktionspotentiale sowie des semantischen Potentials höherer Ordnung jeweils getrennt für die Bodenbedeckungs- und Landnutzungsebene. Auf der x-Achse sind jeweils die 30 relevantesten Merkmale entsprechend ihrer Wichtigkeit angeordnet, wobei die Wichtigkeit von links nach rechts abnimmt. Eine Ausnahme hiervon bildet das semantische Potential höherer Ordnung für die Bodenbedeckungsklassifikation (Abb. 5.6c), für das insgesamt nur 19 Merkmale zur Verfügung stehen. In der Abbildung 5.6 ist zu beachten, dass die y-Achsen der oberen und unteren Paare von Grafiken unterschiedlich skaliert sind.

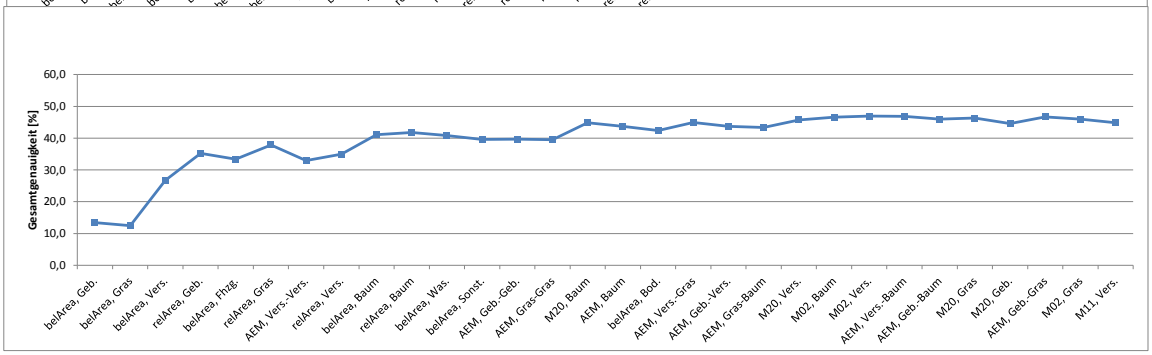
Für die Klassifikation der Bodenbedeckung und der Landnutzung ist ein unterschiedlicher Satz an Merkmalen von Relevanz. Unter den 30 wichtigsten Merkmalen für die Klassifikation der Bodenbedeckung auf Grundlage der Assoziations- und der räumlichen Interaktionspotentiale sind mit Ausnahme eines strukturellen Merkmals (Mittelwert des Gradientenhistogramms) ausschließlich bildbasierte und dreidimensionale Merkmale enthalten. Zu den bildbasierten Merkmalen zählen statistische Parameter (Mittelwert, Standardabweichung, Minimum, Maximum) des NDVI, der Intensität,



(a)







(d) Genauigkeit der Landnutzungsklassifikation - semantisches Potential höherer Ordnung

Abbildung 5.6.: Entwicklung der Gesamtgenauigkeit [%] bei Hinzunahme von Merkmalen nach Relevanz, beginnend mit dem Wichtigsten, zur Klassifikation der Assoziations- und räumlichen Interaktionspotentiale (a, b) sowie des semantischen Potentials höherer Ordnung (c, d) im Testgebiet Schleswig, getrennt für die Bodenbedeckungs- und Landnutzungsebene. Die Sortierung der Merkmale entspricht den RF Wichtigkeitsmaßen. Berechnung der Gesamtgenauigkeit für die Bodenbedeckung auf Ebene der Pixel und für die Landnutzung auf Ebene der GIS-Objekte. (belArea: Konfidenz-gewichtete relative Flächenanteile $f_{belArea,l}$. relArea: relative Flächenanteile $f_{relArea,l}$. AEM: Adjaceny-Event-Matrix. Mij: zentr. Bildmoment μ_{ij}).

des Farbtons, der Sättigung und aller Farbkanäle sowie die Haralick Texturmerkmale Kontrast und Homogenität. Die zwei wichtigsten Merkmale bilden der Mittelwert und das Minimum des NDVI. Demzufolge lassen sich Bodenbedeckungsarten sowie ihre Relationen am besten auf Basis von NDVI-Information unterscheiden. Daneben sind alle dreidimensionalen Merkmale, d.h. Mittelwert, Standardabweichung, Minimum und Maximum des nDOM, in dem Ranking enthalten, wobei der Mittelwert des nDOM den dritten Platz belegt. Dies bestätigt die Relevanz von Höheninformation zur Trennung von Bodenbedeckungsarten. Geometrische Merkmale sind hingegen weniger gut geeignet zur Klassifikation der Bodenbedeckung (Platz 35 für das wichtigste geometrische Merkmal). Dies liegt vor allem in der einheitlichen geometrischen Form der Superpixel begründet (z.B. annähernd gleiche Fläche aller Superpixel), die aus der angewendeten Segmentierungsmethode resultiert.

Demgegenüber sind geometrische Merkmale von wesentlich größerer Bedeutung für die Klassifikation der Landnutzung. Dies resultiert aus der charakteristischen geometrischen Form einiger Nutzungsarten, z.B. weisen *Verkehrswege* i.d.R. eine längliche Form auf, wohingegen *Wohnbauflächen* meist wesentlich kompakter strukturiert sind. Zu den wichtigsten geometrischen Merkmalen zählen die fraktale Dimension, die Kompaktheit, der Formindex, der Flächeninhalt und der Umfang, wobei die fraktale Dimension den ersten Platz im Ranking belegt. Neben den geometrischen Merkmalen sind unter den wichtigsten Merkmalen auch strukturelle Merkmale enthalten, wie z.B. der Mittelwert, das zweite Maximum sowie das erste und zweite Minimum des Histogramms der Gradientenorientierungen. Von diesen Merkmalen belegt der Mittelwert den zweiten Platz, was die Relevanz der strukturellen Information für die Landnutzungsklassifikation unterstreicht (siehe auch [Helmholz et al., 2014]). Daneben leisten insbesondere bildbasierte Merkmale – wie auch schon bei der Bodenbedeckungsklassifikation – einen entscheidenden Beitrag zur Klassifikation der Landnutzung. Das Maximum des Grünkanals, die Standardabweichung der Intensität und des Rotkanals sowie die Haralick Texturmerkmale Energie und Korrelation steuern für die Klassifikation der Landnutzung, nicht aber für die der Bodenbedeckung hilfreiche Informationen bei. Demgegenüber sind die dreidimensionalen Merkmale von untergeordneter Bedeutung für die Klassifikation der Landnutzung; so findet sich der Mittelwert des nDOM als einziges dreidimensionales Merkmal auf Platz 30 des Rankings wieder. Der Sättigungspunkt wird etwa bei einer Anzahl von 20 Merkmalen bei der Bodenbedeckungs- und 30 Merkmalen bei der Landnutzungsklassifikation erreicht, wobei der Punkt jeweils nicht eindeutig bestimmbar ist. Der Satz der 20 wichtigsten Merkmale stimmt für die Bodenbedeckungsklassifikation in beiden Testgebieten mit Ausnahme weniger Merkmale überein. Demgegenüber weichen die Merkmalssätze für die Landnutzungsklassifikation deutlich voneinander ab. Im Testgebiet Hameln sind beispielsweise die Haralick Texturmerkmale sowie die aus den R,G,B-Farbkanälen abgeleiteten Merkmale von geringerer Bedeutung, wohingegen den dreidimensionalen Merkmalen, diversen strukturellen Merkmalen sowie den aus dem nahen Infrarot abgeleiteten Merkmalen eine höhere Wichtigkeit beigemessen wird. Ungeachtet dieser Unterschiede sind die ersten drei Plätze bei der Bodenbedeckungsklassifikation und die ersten zwei Plätze bei der Landnutzungsklassifikation in beiden Testgebieten mit identischen Merkmalen besetzt.

Die Entwicklung der Gesamtgenauigkeit bei Hinzunahme der Kontextmerkmale zur Klassifikation der semantischen Potentiale höherer Ordnung in den Abbildungen 5.6c und 5.6d folgt keinem kontinuierlichen Anstieg, sondern zeigt wiederholt ein Schwanken der Genauigkeit (z.B. bei der Hinzunahme vom siebten bzw. achten Merkmal).

Zu den wichtigsten Kontextmerkmalen zählen in beiden Fällen, d.h. bei der Klassifikation der Bodenbedeckung und der Landnutzung, die Konfidenz-gewichteten relativen Flächenanteile, die sich aus den klassenspezifischen Konfidenzwerten der Zwischenlösungen ableiten. Bei der Klassifikation der Bodenbedeckung nehmen diese Merkmale die vorderen Plätze ein vor den relativen Flächenanteilen (vgl. Abb. 5.6c), die folglich für die Klassifikation von geringerer Bedeutung sind. Das Merkmal, welches den minimalen Abstand der Superpixel von den Nutzungsgrenzen beschreibt, belegt den achten Platz (vgl. Abb. 5.6c). Diesem Merkmal wird eine höhere Wichtigkeit zugeschrieben als den Konfidenz-gewichteten relativen Flächenanteilen der Klassen *Wald* und *Weg* (Platz 9 und 10). Die drei wichtigsten Merkmale quantisieren die räumliche Verteilung der Nutzungsarten *Grünland*, *Wohnbaufläche* und *Verkehrsweg* innerhalb eines Superpixels (vgl. Abb. 5.6c). Die Klassen *Grünland* und *Verkehrsweg* weisen zumeist eine homogene Bodenbedeckung auf (*Gras* im Falle von *Grünland* und *Versiegelung* im Falle von *Verkehrsweg*), sodass bei Vorliegen dieser Nutzungsarten eindeutig auf eine Bodenbedeckungsart geschlossen werden kann. Anders verhält es sich bei der Klasse *Wohnbaufläche*, die bzgl. der Bodenbedeckung typischerweise heterogener strukturiert ist. Jedoch ist in den verwendeten Untersuchungsgebieten ein regelmäßiges Muster mit den Anteilen *Gebäude*, Vegetation (*Gras*, *Baum*) und *Versiegelung* zu beobachten, sodass die Bodenbedeckung mit dem Wissen über die vorliegende Nutzungsart *Wohnbaufläche* auf drei von acht Klassen eingeschränkt werden kann.

Dieser Zusammenhang spiegelt sich auch in dem Ranking der Merkmale für die Landnutzungsklassifikation in Abbildung 5.6d wider; so belegen die zugehörigen Bodenbedeckungsarten *Gras*, *Versiegelung* und *Gebäude* in Form der Konfidenz-gewichteten relativen Flächenanteile die ersten drei Plätze im Ranking (vgl. Abb. 5.6d). Darüber hinaus sind unter den 30 wichtigsten Merkmalen diverse graphenbasierte Merkmale enthalten, die das gemeinsame Auftreten verschiedener Bodenbedeckungsklassen beschreiben (vgl. Abb. 5.6d). Unter diesen Merkmalen wird der Häufigkeit des gemeinsamen Auftretens der Bodenbedeckungsart *Versiegelung* an benachbarten Pixeln die höchste Relevanz zugeschrieben (Platz 7 im Ranking). Die zentralen Bildmomente zur Beschreibung der räumlichen Verteilung der Bodenbedeckungsarten innerhalb eines Landnutzungsobjekts sind ebenfalls mehrfach im Ranking vertreten (vgl. Abb. 5.6d). Das wichtigste Merkmal dieser Gruppe beschreibt die Verteilung der Bodenbedeckungsart *Baum* innerhalb eines Landnutzungsobjekts und belegt den 15. Platz im Ranking.

Bei der Klassifikation auf Grundlage des semantischen Potentials höherer Ordnung sind für die Bodenbedeckungsebene alle 19 Merkmale erforderlich, um den Sättigungspunkt zu erreichen. Für die Landnutzungsebene ist ein Sättigungseffekt bei ca. 25 Merkmalen zu beobachten, wenngleich die Genauigkeit zwischen 20 und 30 Merkmalen schwankt und somit kein eindeutiger Sättigungspunkt identifiziert werden kann. Im Testgebiet Hameln sind die 19 Kontextmerkmale für die Bodenbedeckungsklassifikation in einer ähnlichen Reihenfolge sortiert. Zudem stimmt der Merkmalssatz für die Kontextmerkmale der Landnutzungsklassifikation in beiden Testgebieten mit Ausnahme weniger Merkmale überein.

5.2.3.4. Diskussion

Die Analyse der Merkmale hat ergeben, dass die Klassifikation der Bodenbedeckung insbesondere von bildbasierten und dreidimensionalen Merkmalen profitiert. Von herausgehobener Bedeutung sind hierbei die NDVI- und die Höheninformation, die in besonderem Maße zur Unterscheidung der verschiedenen Bodenbedeckungsarten beitragen. Im Gegensatz dazu sind geometrische Merkmale aufgrund der ähnlichen geometrischen Eigenschaften der Superpixel von untergeordneter Bedeutung. Für die Klassifikation der Landnutzung sind die geometrischen Merkmale hingegen von besonderer Wichtigkeit, da einige Nutzungsarten durch eine charakteristische geometrische Form gekennzeichnet sind. Neben den geometrischen Merkmalen sind aber auch bildbasierte Merkmale von Relevanz für die Landnutzungsklassifikation. Zu den wichtigsten Merkmalen für die semantischen Potentiale höherer Ordnung zählen in beiden Ebenen die Konfidenz-gewichteten relativen Flächenanteile. Die Untersuchungen attestieren diesen Merkmalen eine höhere Effektivität im Vergleich zu den relativen Flächenanteilen, deren Berechnung sich alleine auf die Klassenlabels mit dem maximalen Konfidenzwert stützt. Die übrigen Kontextmerkmale sind zwar von nachgeordneter Bedeutung, tragen aber ebenfalls hilfreiche Informationen zur Klassifikation bei.

Die Selektion der Merkmale, die in den Experimenten in Kapitel 5.2.5 verwendet werden, erfolgt in Abhängigkeit des RF Wichtigkeitsparameters pro Merkmal. Die Mindestanzahl der Merkmale entspricht dabei dem Punkt bei sukzessiver Hinzunahme der Merkmale, an dem ein Sättigungseffekt eintritt. Für die Klassifikation der Assoziations- und der räumlichen Interaktionspotentiale werden die jeweils 20 wichtigsten Merkmale für die Klassifikation der Bodenbedeckung und die jeweils 30 wichtigsten Merkmale für die Klassifikation der Landnutzung verwendet. Für die Experimente in Kapitel 5.2.5 werden für beide Testgebiete individuelle Merkmalssätze verwendet, weil der Satz der wichtigsten Merkmale für die Landnutzungsebene deutlich variiert. Für die semantischen Potentiale höherer Ordnung werden für die Klassifikation der Bodenbedeckung die 19 verfügbaren Merkmale verwendet. Da sich bei Hinzunahme der Merkmale bei der Klassifikation der Landnutzung kein eindeutiger Sättigungspunkt identifizieren lässt, werden die jeweils 30 wichtigsten Merkmale verwendet. Auch hier unterscheiden sich die Merkmalssätze in beiden Testgebieten, wenngleich sie nur in wenigen Merkmalen voneinander abweichen. Die Wahl individueller Merkmalssätze pro Testgebiet dient im Rahmen dieser Arbeit der Erzielung bestmöglicher Ergebnisse für jedes Testgebiet. Allerdings gilt es diesbezüglich anzumerken, dass eine individuelle Auswahl der Merkmale in der Praxis eher ungeeignet ist, da hierfür eine Validierungsmenge der Trainingsdaten zur Bestimmung der Merkmale zurückgehalten werden muss. Um dieses Problem zu umgehen, bietet es sich daher an, in der Praxis alle verfügbaren Merkmale zu verwenden oder im Vorfeld einen möglichst allgemeingültigen Merkmalssatz anhand mehrerer Testgebiete zu bestimmen. Für die Klassifikation der Assoziations- und der räumlichen Interaktionspotentiale in den Experimenten in Kapitel 5.2.5 werden in der Bodenbedeckungsebene die Differenzen der Merkmalsvektoren und in der Landnutzungsebene aneinandergereihte Merkmalsvektoren verwendet. Da sich kein Quantil ε als besonders geeignet herausgestellt hat, d.h. zu keinem signifikant besseren Ergebnis führt als für den Standardwert ($\varepsilon = 100\,\%$), wird für die Experimente in Kapitel 5.2.5 der Standardwert für die Normalisierung aller Merkmale beibehalten.

5.2.4. Modell- und Inferenzparameter

5.2.4.1. Zielsetzung

Die nachfolgenden Experimente dienen der Analyse der einzelnen Modell- und Inferenzparameter. Die Parameterstudien haben zum Ziel, zu untersuchen, wie sensitiv der Algorithmus auf die einzelnen Parameter reagiert, und auf Basis dieser Erkenntnisse eine bestmögliche Konfiguration von Parametern für die Experimente in Kapitel 5.2.5 festzulegen.

5.2.4.2. Strategie

Die hier beschriebenen Experimente gliedern sich in zwei Teile. Der erste Teil beschäftigt sich mit den Parametern der einzelnen Potentiale, d.h. den Parametern der zugrundeliegenden Klassifikatoren. Der zweite Teil analysiert die Parameter Ω zur Gewichtung der Potentiale zueinander sowie die Inferenzparameter, d.h. die Anzahl der Iterationen.

Die Bestimmung bestmöglicher Werte für die Parameter erfolgt im Rahmen einer empirischen Untersuchung anhand des Testdatensatzes Schleswig. In dem ersten Teil der Untersuchung wird jeweils ein Satz an Experimenten für jeden Parameter der Potentiale durchgeführt. Die Untersuchung erfolgt gesondert für jedes Potential, d.h. die übrigen Potentiale fließen hierbei nicht in die Klassifikationsentscheidung ein (zugehörige Parameter ω^i werden zu Null gesetzt). Die Parameter variieren in einem vorgegebenen Wertebereich, während für alle anderen Parameter des jeweiligen Potentials die Standardwerte (siehe Tab. 5.4) verwendet werden. In dem zweiten Teil der Untersuchung werden die Potentiale nicht separat zum Klassifizieren verwendet, sondern hierbei basiert die Klassifikation auf dem Gesamtmodell unter Verwendung der Standardparameter (siehe Tab. 5.4), von denen lediglich der zu untersuchende Parameter abweicht. Für die Parameter w^2 und w^4 der räumlichen Interaktionspotentiale sowie die Parameter $\omega^{5,u\to c}$ und $\omega^{5,c\to u}$ der semantischen Potentiale höherer Ordnung in der Bodenbedeckungs- und Landnutzungsebene wird jeweils ein Satz an Experimenten gerechnet, in denen der Parameter im Wertebereich [0,1;2,0] variiert, während für alle übrigen Parameter die Standardwerte angenommen werden. Analog verhält es sich für die Anzahl der Iterationen n_{It} und n_{LBP}, die sich allerdings auf das Gesamtmodell beziehen und sich damit gleichzeitig auf die Klassifikation der Bodenbedeckung und der Landnutzung auswirken. Folglich werden zwei Sätze an Experimenten gerechnet, je ein Satz für n_{It} bzw. n_{LBP}, wobei die Parameter jeweils im Wertebereich [1,10] variieren. Ziel der Analyse ist es, die Sensitivität der Methode im Hinblick auf Änderungen der Parameter zu beurteilen, um festzustellen, inwieweit das Klassifikationsergebnis von der Wahl der Parameter abhängt. Die Bewertung erfolgt auf der Grundlage der Gesamtgenauigkeit, abgeleitet aus einer Evaluation der Klassifikationsergebnisse für die Bodenbedeckung auf Grundlage der Pixel und für die Landnutzung auf Grundlage der GIS-Objekte. Auf Basis dieser Erkenntnisse wird für die Experimente in Kapitel 5.2.5 eine bestmögliche Konfiguration von Modell- und Inferenzparametern definiert.

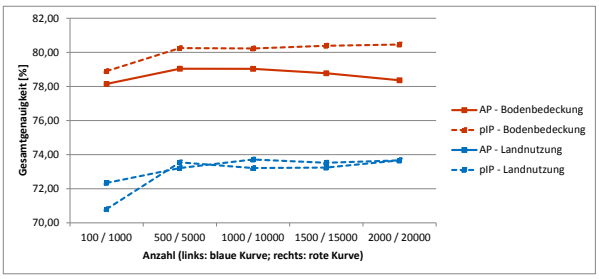
5.2.4.3. Beschreibung der Ergebnisse

Parameter der Random Forest Klassifikatoren Die Ergebnisse der Untersuchung des Einflusses der RF Parameter auf die Klassifikationsgenauigkeit sind in der Abbildung 5.7 für das Testgebiet Schleswig dargestellt. In den Diagrammen ist die Gesamtgenauigkeit als Funktion der RF Parameter aufgetragen, die bei exklusiver Anwendung des Assoziations-, des paarweisen räumlichen Interaktionspotentials oder des semantischen Potentials höherer Ordnung in der Bodenbedeckungs- bzw. Landnutzungsebene erzielt wird. Bei der Interpretation der Diagramme ist zu beachten, dass die y-Achsen zwar einen unterschiedlichen Wertebereich darstellen, jedoch gleich skaliert sind.

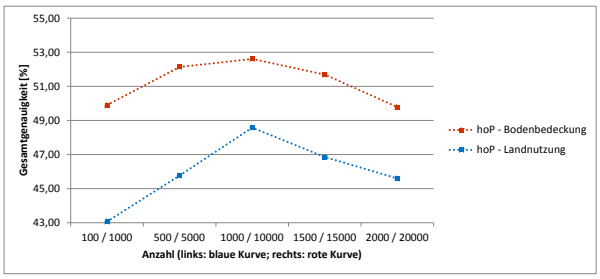
Die Abbildung 5.7a zeigt den Einfluss der maximalen Anzahl $N_{T,max}$ der Trainingsbeispiele pro Klasse für die Klassifikation des Assoziations- und räumlichen Interaktionspotentials pro Ebene. Die Genauigkeit für eine Anzahl von 100 Trainingsbeispielen pro Klasse im Falle der Landnutzungsklassifikation bzw. 1000 Trainingsbeispielen pro Klasse im Falle der Bodenbedeckungsklassifikation hebt sich bei beiden Klassifikationen negativ von den übrigen Ergebnissen ab. Für alle Parameter, die diesen Wert überschreiten, wird hingegen das gleiche Genauigkeitsniveau erreicht. Die Unterschiede in den Gesamtgenauigkeiten betragen hier weniger als 1 %, sowohl bezogen auf die Bodenbedeckungs- als auch die Landnutzungsebene. Der Einfluss dieses Parameters auf das semantische Potential höherer Ordnung ist in Abbildung 5.7b dargestellt. Hier zeigen sich deutlich größere Unterschiede in Höhe von ca. 3 % bei der Bodenbedeckungs- und ca. 5 % bei der Landnutzungsklassifikation, die ggf. aus einer insgesamt höheren Unsicherheit des Verfahrens resultieren (Gesamtgenauigkeit ist in beiden Fällen kleiner als 53 %). Nichtsdestotrotz zeichnet sich in den Diagrammen ab, dass dieses Potential sowohl bei einer zu geringen (kleiner als 100 bzw. 1000) als auch einer zu hohen Anzahl (größer als 2000 bzw. 20000) nicht adäquat gelernt werden kann. Die höchste Genauigkeit dieses Potentials wird in beiden Ebenen für die Standardwerte erzielt (1000 für die Landnutzung und 10.000 für die Bodenbedeckung).

Als zweites wird der RF Parameter T_{max} untersucht, der die maximale Tiefe der Bäume regelt. In der Abbildung 5.7c ist dessen Einfluss auf die Klassifikation des Assoziations- und räumlichen Interaktionspotentials pro Ebene dargestellt. Hier zeigt sich, dass eine geringe Tiefe von kleiner als 10 für beide Klassifikationsaufgaben zu Genauigkeitseinbußen bei der Klassifikation auf Basis der räumlichen Interaktionspotentiale führt. Die Genauigkeiten für diesen Wert weichen zwischen 3 % (Landnutzung) und 5 % (Bodenbedeckung) von den übrigen Genauigkeiten ab. Dieser Effekt lässt sich bei dem Assoziationspotential nicht in dem Ausmaß beobachten. Für die restlichen Größen betragen die Abweichungen weniger als 1 %, sodass die Parameter zwischen einer Anzahl von 20 und 40 gleichermaßen gut geeignet erscheinen. Demgegenüber treten bei dem semantischen Potential höherer Ordnung deutlichere Unterschiede in Bezug auf die Gesamtgenauigkeit auf (vgl. Abb. 5.7d). Die Schwankungen sind besonders ausgeprägt bei Verwendung von kleinen Werten (10, 20) für die Klassifikation der Bodenbedeckung. Hierbei ist keine klare Tendenz zu einem Wert zu erkennen. Ab einer Tiefe von 25 wird die Streuung jedoch geringer, woraus folgert, dass Werte von größer als 25 zu einer Stabilisierung der Klassifikation beitragen. Bei der Landnutzungsklassifikation wird für eine Tiefe von 25 die maximale Gesamtgenauigkeit erreicht. Die Streuung der Ergebnisse ist hierbei insgesamt etwas geringer.

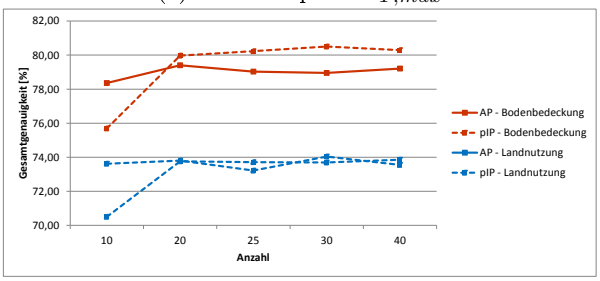
Die Abbildungen 5.7e und 5.7f untersuchen den Einfluss des RF Parameters $N_{T,min}$, der die minimale Anzahl an Samples unterschiedlicher Klassen vorgibt, die vorhanden sein muss, um den



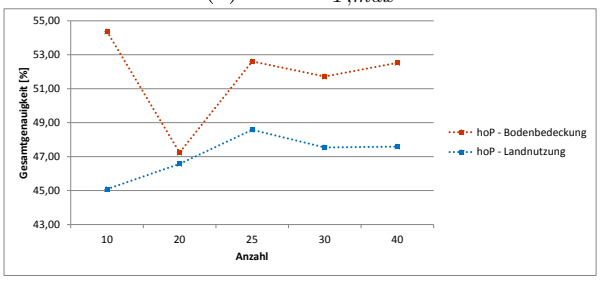
(a) AP und pIP: $N_{T,max}$



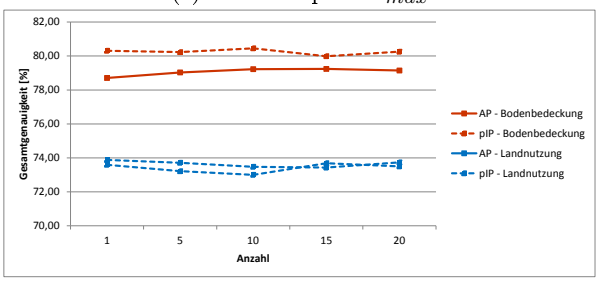
(b) hoP: $N_{T,max}$



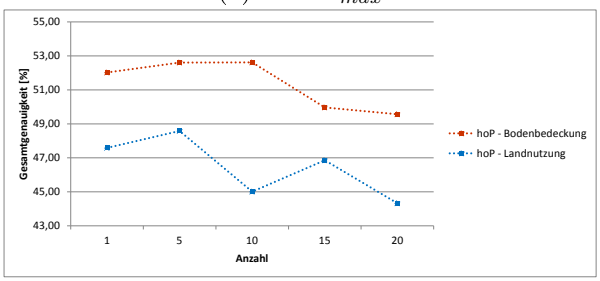
(c) AP und pIP: T_{max}



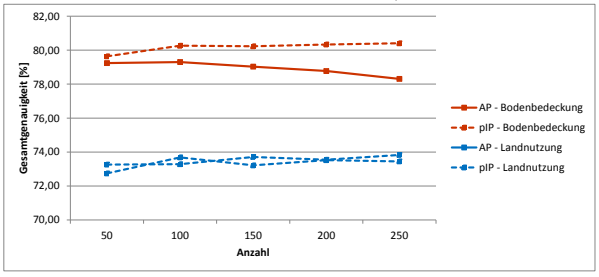
(d) hoP: T_{max}



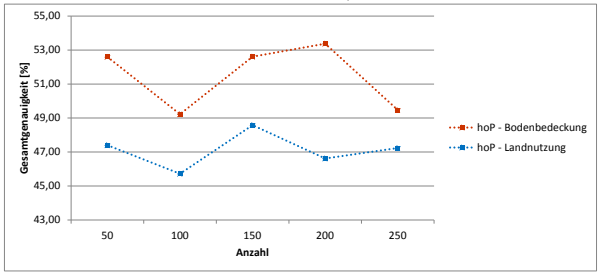
(e) AP und pIP: $N_{T,min}$



(f) hoP: $N_{T,min}$



(g) AP und pIP: B



(h) hoP: B

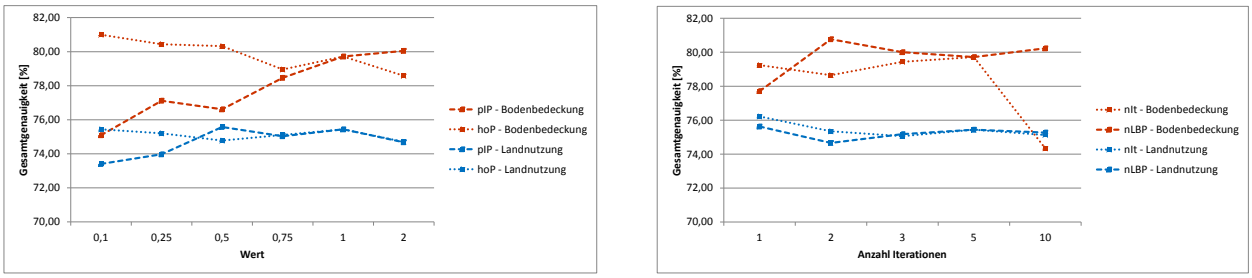
Abbildung 5.7.: Gesamtgenauigkeit [%] für die Klassifikation der Bodenbedeckung (rot) und Landnutzung (blau) auf Grundlage des Assoziations- (AP, durchgezogene Linie), des paarweisen räumlichen Interaktionspotentials (pIP, gestrichelte Linie) (beide: a,c,e,g) oder des semantischen Potentials höherer Ordnung (hoP, gepunktete Linie) (b,d,f,h) als Funktion der RF Parameter max. Anzahl $N_{T,max}$ von Trainingsbeispielen pro Klasse, max. Tiefe T_{max} der Bäume, min. Anzahl $N_{T,min}$ der Samples für nicht-terminierende Knoten, max. Anzahl B der Bäume. Die Berechnung der Gesamtgenauigkeit erfolgt für die Bodenbedeckung auf Basis der Pixel und für die Landnutzung auf Basis der GIS-Objekte. In (a) und (b) bezieht sich der linke Eintrag auf der x-Achse auf die Landnutzungsklassifikation und der rechte Eintrag auf die Bodenbedeckungsklassifikation.

Knoten weiter aufzuspalten. Aus der Abbildung 5.7e lässt sich folgern, dass dieser Parameter auf die Klassifikation der Assoziations- und räumlichen Interaktionspotentiale nur einen sehr geringen Einfluss

hat, da die Abweichungen zwischen den Genauigkeiten beider Klassifikationsergebnisse deutlich kleiner als 1 % sind. Bei der Klassifikation des semantischen Potentials höherer Ordnung kristallisiert sich hingegen ein leicht abfallender Trend mit zunehmender Anzahl heraus (vgl. Abb. 5.7f), der allerdings in Anbetracht der starken Schwankungen nicht mit Sicherheit verifiziert werden kann. Hierbei wird für beide Klassifikationsaufgaben unter Verwendung des Standardwertes (Mindestanzahl von 5 Samples) die höchste Genauigkeit erzielt.

In der Abbildung 5.7g ist der Einfluss der maximalen Anzahl B von Bäumen auf die Klassifikationsgenauigkeit des Assoziations- und räumlichen Interaktionspotentials dargestellt. Trotz des leicht abfallenden Trends der Genauigkeiten mit zunehmender Anzahl der Bäume im Rahmen der Klassifikation des räumlichen Interaktionspotentials in der Bodenbedeckungsebene weichen die Genauigkeiten insgesamt weniger als 1 % voneinander ab. Daraus lässt sich folgern, dass die untersuchten Parameterwerte gleichermaßen gut geeignet sind für die Klassifikationsaufgaben. In dem Diagramm in Abbildung 5.7h ist nicht zuletzt durch die hohe Streuung der Ergebnisse kein klarer Trend zu erkennen. Die Ergebnisse für den Standardwert (150 Bäume) weichen hierbei um weniger als 1 % von der höchsten Genauigkeit ab.

Parameter Ω und Parameter der Inferenzprozedur Die Ergebnisse der Untersuchung des Einflusses der Parameter $\omega^i \in \{\omega^2, \omega^4, \omega^{5,u\to c}, \omega^{5,c\to u}\}$ pro Potentialterm sowie der Parameter der Inferenzprozedur sind in der Abbildung 5.8 für das Testgebiet Schleswig dargestellt. In den Diagrammen ist die Gesamtgenauigkeit als Funktion der Parameter ω^i sowie der Anzahl der Iterationen n_{It} und n_{LBP} aufgetragen, die bei gemeinsamer Anwendung des Assoziations-, des paarweisen räumlichen Interaktionspotentials und des semantischen Potentials höherer Ordnung im Rahmen des iterativen Inferenzalgorithmus erzielt wird.



(a) Parameter ω^i pro Potentialterm

(b) Anzahl Iterationen n_{It}, n_{LBP}

Abbildung 5.8.: Gesamtgenauigkeit [%] für die Klassifikation der Bodenbedeckung (rot) und Landnutzung (blau) auf Grundlage des paarweisen räumlichen Interaktionspotentials (pIP) und des semantischen Potentials höherer Ordnung (hoP) als Funktion des Parameters ω^i pro Potentialterm (pIP (ω^2, ω^4): gestrichelte Linie; hoP ($\omega^{5,u\to c}, \omega^{5,c\to u}$): gepunktete Linie) (a) und als Funktion der Anzahl der Iterationen des iterativen Inferenzalgorithmus (nIt, gepunktete Linie) und der Iterationen eines LBP-Teilschritts (nLBP, gestrichelte Linie) (b).

In der Abbildung 5.8a ist der Einfluss des Parameters ω^i gesondert für die paarweisen räumlichen Interaktionspotentiale (d.h. ω^2, ω^4) und die semantischen Potentiale höherer Ordnung (d.h. $\omega^{5,u\to c}, \omega^{5,c\to u}$) in der Bodenbedeckungs- und Landnutzungsebene dargestellt, der den zugehörigen

Potentialterm relativ zum Assoziationspotential gewichtet. Alle Parameter außer ω^i nehmen die Standardwerte an. Für den Parameter ω^3 wird analog zu ω^1 in allen Untersuchungen der Wert 1 angesetzt, d.h. $\omega^3 \equiv 1$. Für die paarweisen räumlichen Interaktionspotentiale steigt die Genauigkeit bei beiden Klassifikationsaufgaben im Durchschnitt leicht an mit Erhöhung der Parameter ω^2 bzw. ω^4, wobei der Anstieg bei der Klassifikation der Bodenbedeckung stärker ausgeprägt ist als bei der Landnutzung (Differenz in Höhe von ca. 5 % zwischen dem minimalen Wert bei 0,1 und dem maximalen Wert bei 2,0). Von dem kontinuierlichen Anstieg weichen die Ergebnisse für den Parameter 0,5 in der Bodenbedeckungsebene und die Parameter 0,5 und 2,0 in der Landnutzunsgebene leicht ab. Da sich der Parameter ω^i invers proportional zum tatsächlichen Einfluss verhält, folgert daraus, dass sich die Verringerung des Einflusses des räumlichen Interaktionspotentials positiv auf die Gesamtgenauigkeit auswirkt. Folglich führt eine zu hohe Gewichtung dieses Potentials zu einer Verschlechterung der Klassifikationsergebnisse. Bei der Bodenbedeckungsklassifikation erzielt der Parameter $\omega^2 = 2,0$ das beste Ergebnis. Bei der Landnutzungsklassifikation erzielt der Parameter ω^4 für die Werte 0,5 und 1,0 ein ähnliches Genauigkeitsniveau (Unterschiede von ca. 0,1 %). Die Genauigkeitsunterschiede betragen hierbei insgesamt weniger als 2,2 %. Für die Parameter ω^i für das semantische Potential höherer Ordnung (d.h. $\omega^{5,u\to c}, \omega^{5,c\to u}$) ist im Bereich zwischen den Werten 0,1 und 0,5 ein leicht gegenläufiger Trend zu beobachten, d.h. Abnahme der Genauigkeit mit Erhöhung des Parameters. In der Bodenbedeckungsebene setzt sich dieser Trend auch für die Werte 0,75 bis 2,0 fort, wenngleich die Genauigkeit für den Wert 1,0 von dem kontinuierlichen Verlauf abweicht, d.h. die Genauigkeiten für die Werte 0,75 und 2,0 übersteigt, nicht jedoch die Genauigkeiten für Parameterwerte kleiner als 0,5. Der maximale Genauigkeitsunterschied beträgt 2,4 % zwischen dem maximalen Wert bei 0,1 und dem minimalen Wert bei 2,0. In der Landnutzungsebene setzt sich der abfallende Trend jedoch nicht fort. Hier steigt die Genauigkeit mit Erhöhung des Parameters $\omega^{5,c\to u}$ bis zum Wert 1,0 wieder an, sodass die Genauigkeiten für die Werte 0,1 und 1,0 auf dem gleichen Niveau sind. Für den Wert 2,0 wird die schlechteste Gesamtgenauigkeit erzielt, die allerdings von den maximalen Werten bei 0,1 und 1,0 um weniger als 1 % abweicht. Die Beobachtungen zu den Parametern ω^i werden im Testgebiet Hameln im Wesentlichen bestätigt. Hier weichen lediglich die Ergebnisse für die Parameter des semantischen Potentials höherer Ordnung in der Bodenbedeckungsebene von den Ergebnissen in Schleswig in der Form ab, dass hier für einen Parameter von $\omega^{5,u\to c} = 0,5$ die höchste Gesamtgenauigkeit erzielt wird.

In der Abbildung 5.8b ist die Gesamtgenauigkeit als Funktion der Anzahl der Iterationen dargestellt, jeweils getrennt für die Anzahl n_{It} der Iterationen der iterativen Inferenzprozedur sowie der Anzahl n_{LBP} der Iterationen in einem LBP-Teilschritt. Für die Anzahl n_{LBP} der Iterationen in einem LBP-Teilschritt weichen die Ergebnisse, mit Ausnahme des Ergebnisses der Bodenbedeckungsklassifikation für die Anzahl 1, nur minimal voneinander ab; so betragen die Abweichungen bei beiden Klassifikationsaufgaben maximal 1 %. Folglich hat dieser Parameter nur einen geringen Einfluss auf die Klassifikationsergebnisse. Bei 1 Iteration ergeben sich für die Klassifikation der Bodenbedeckung deutliche Genauigkeitseinbußen, wohingegen für die Klassifikation der Landnutzung mit 1 Iteration die beste Gesamtgenauigkeit erzielt wird. Umgekehrt verhält es sich bei 2 Iterationen, d.h. hier steht der höchsten Genauigkeit bei der Bodenbedeckungsklassifikation die niedrigste Genauigkeit bei der Landnutzungsklassifikation gegenüber. In Bezug auf die Anzahl n_{It} der Iterationen der Inferenzprozedur zeigt sich für beide Klassifikationsaufgaben zunächst ein leicht abfallender Trend von 1 hin zu 2

Iterationen. Bei der Landnutzungsklassifikation setzt sich dieser Trend jedoch nicht fort, sondern die Gesamtgenauigkeit bleibt ab einer Anzahl von 2 Iterationen auf einem ähnlichen Niveau (max. Unterschiede von 0,4 %). Demgegenüber steigt die Gesamtgenauigkeit bei der Bodenbedeckungsklassifikation zwischen einer Anzahl von 2 und 5 Iterationen wieder an, sodass bei einer Anzahl von 5 Iterationen die höchste Gesamtgenauigkeit erzielt wird. Das Klassifikationsergebnis für die Bodenbedeckung zeigt bei einer Anzahl von 10 Iterationen der Inferenzprozedur deutliche Genauigkeitseinbußen, die ggf. darin begründet liegen, dass die einzelnen Potentialterme mit den Standardwerten für ω^i (siehe Tab. 5.4) nicht optimal zueinander gewichtet sind, sodass die Gesamtenergie divergiert. Im Testgebiet Hameln sind die Genauigkeitsunterschiede beider Klassifikationsergebnisse mit maximal 0,8 % für n_{It} und n_{LBP} sehr klein, sodass sich daraus keine Aussagen ableiten lassen.

5.2.4.4. Diskussion

Insgesamt haben die RF Parameter in den untersuchten Wertebereichen, abgesehen von wenigen Ausnahmen, nur einen geringen Einfluss auf die Klassifikationsergebnisse. Folglich ist der Algorithmus nicht besonders sensitiv gegenüber der Wahl der RF Parameter. Die gewählten Standardparameter zeigen für alle Potentiale ein vergleichsweise gutes Ergebnis, sodass sie in den folgenden Experimenten beibehalten werden.

Im Gegensatz dazu zeigen die Experimente einen Einfluss der Parameter ω^i auf das Klassifikationsergebnis, der sich in maximalen Variationen der Gesamtgenauigkeit in Höhe von 5 % ausprägt. Aus diesem Grund werden die Parameterwerte gewählt, für die die bestmöglichen Ergebnisse erzielt wurden. Weichen die maximal erzielten Gesamtgenauigkeiten allerdings nur unwesentlich von den Genauigkeiten für die Standardwerte ab, werden die Standardwerte für die Parameter ω^i aus der Tabelle 5.4 in den nachfolgenden Untersuchungen beibehalten. Die gewählten Parameter sind in der Tabelle 5.9 separat für das Testgebiet Hameln und Schleswig aufgetragen. Die Werte unterscheiden sich nur hinsichtlich des Parameters $\omega^{5,u\rightarrow c}$, der dem semantischen Potential höherer Ordnung in der Bodenbedeckungsebene im Testgebiet Schleswig einen höheren Einfluss beimisst. Für die Anzahl der Iterationen haben die Untersuchungen ergeben, dass der gewählte Standardwert von 5 Iterationen für den LBP-Teilschritt einen guten Kompromiss für beide Klassifikationsaufgaben darstellt. Der iterative Inferenzalgorithmus erzielt sowohl für eine Anzahl von 1 bzw. 5 Iterationen gute Ergebnisse, sodass beide Werte in den nachfolgenden Untersuchungen in Kapitel 5.2.5 angewendet und näher analysiert werden. Mit einer Iteration entspricht das Verfahren einer Variante des zweistufigen Verfahrens (vgl. Ausführungen in Kap. 5.1.4).

Parameter ω^i	Bodenbedeckung			Landnutzung		
	AP	pIP	hoP	AP	pIP	hoP
Hameln	1,0	2,0	0,5	1,0	1,0	1,0
Schleswig	1,0	2,0	0,1	1,0	1,0	1,0

Tabelle 5.9.: Bestmögliche Konfiguration der Parameter ω^i pro Potentialterm für die Testgebiete Hameln und Schleswig. Für die Parameter der RF Klassifikatoren sowie die Anzahl der Iterationen n_{It} und n_{LBP} werden die Standardwerte beibehalten (siehe Tab. 5.4).

5.2.5. Kontextbasierte Klassifikation

5.2.5.1. Zielsetzung

Die nachfolgenden Experimente dienen der Analyse der Leistungsfähigkeit der in dieser Arbeit entwickelten Methodik zur simultanen Klassifikation der Bodenbedeckung und der Landnutzung. Die Untersuchungen haben zum Ziel, die zu Beginn formulierten Annahmen auf ihre Gültigkeit zu überprüfen, d.h. zu evaluieren, ob die Berücksichtigung von Kontextinformation zu einer verbesserten Klassifikation der Bodenbedeckung und Landnutzung beiträgt. Zu diesem Zweck gilt es zu analysieren, inwiefern sich die Berücksichtigung der einzelnen kontextuellen Abhängigkeiten auf die Klassifikationsgenauigkeit auswirkt. Da der Einfluss von Kontext in wesentlichem Maße auch von der Art und Weise abhängt, wie die Information in die Klassifikationsentscheidung einbezogen wird, gilt es zusätzlich die Effektivität des iterativen Inferenzalgorithmus im Vergleich zu einem einmaligen Austausch von Kontextinformation zu bewerten.

5.2.5.2. Strategie

Zur Analyse der Leistungsfähigkeit des Verfahrens wird die Klassifikation mit unterschiedlichen Ansätzen durchgeführt, und anschließend werden die Ergebnisse verglichen. Als Vergleichsverfahren für den iterativen Inferenzalgorithmus dienen die unabhängige RF Klassifikation und die CRF Klassifikation, die in Kapitel 5.1.4 näher beschrieben sind. Der Vergleich mit den Ergebnissen einer unabhängigen Klassifikation dient der Untersuchung des Einflusses der Gesamtheit an Kontextinformation auf die Klassifikationsergebnisse, d.h. der räumlichen Abhängigkeit benachbarter Bildprimitive sowie der semantischen Abhängigkeit zwischen der Bodenbedeckung und der Landnutzung. Im Gegensatz dazu wird mit dem Vergleich zur CRF Klassifikation speziell der Einfluss der semantischen Abhängigkeit analysiert. Zur Evaluation des iterativen Inferenzalgorithmus werden die Ergebnisse des Verfahrens bei einmaliger (d.h. $n_{It} = 1$) und fünfmaliger (d.h. $n_{It} = 5$) Iteration der Inferenzprozedur verglichen. Mit der einmaligen Integration von Kontext entspricht das Verfahren einer Variante des zweistufigen Ansatzes zur Klassifikation der Landnutzung, wie es in [Albert et al., 2014a] vorgestellt wurde (vgl. Ausführungen in Kap. 5.1.4).

Die Bewertung des in dieser Arbeit entwickelten Verfahrens erfolgt auf der Grundlage einer differenzierten Analyse der Klassifikationsergebnisse. Der Vergleich der Ergebnisse erfolgt auf Basis von Genauigkeitsmaßen für das Gesamtergebnis sowie die einzelnen Klassen. Weitere Erkenntnisse liefert ein rein visueller, qualitativer Vergleich der erzielten Klassifikationsergebnisse. Die mittels qualitativer und quantitativer Analyse identifizierten Vor- und Nachteile des entwickelten Verfahrens werden anhand von positiven und negativen Beispielen veranschaulicht und hinsichtlich möglicher Ursachen kritisch diskutiert. Die Beispiele zeigen dabei typische Fälle für Verbesserungen sowie Verschlechterungen gegenüber den Vergleichsverfahren. Ziel ist es, sowohl die Vorteile des Verfahrens hervorzuheben als auch die Probleme und Grenzen der Methode aufzuzeigen, wobei die Analyse detailliert auf die Unterschiede zwischen den Ergebnissen des iterativen Inferenzalgorithmus mit $n_{It} = 1$ und $n_{It} = 5$ eingeht.

Für die unterschiedlichen Klassifikationsverfahren werden die identischen Klassenstrukturen, Merk-

male sowie Superpixel-, Modell- und Inferenzparameter verwendet, sofern sie in den jeweiligen Klassifikationsansatz einfließen. Die Festlegung der Werte erfolgt auf Grundlage der Erkenntnisse aus den Voruntersuchungen in den Kapiteln 5.2.1 bis 5.2.4. Folglich handelt es sich hierbei um die empirisch bestimmte, bestmögliche Konfiguration der Parametereinstellungen für die verwendeten Testgebiete. Die Untersuchung der maximal möglichen semantischen Auflösung der Landnutzungsklassifikation in Kapitel 5.2.1 hat für beide Testgebiete einen Detailgrad von 9 Klassen ergeben. Auf der Ebene der Bodenbedeckung werden 8 Klassen unterschieden. Die Untersuchungen in Kapitel 5.2.3 haben ergeben, dass die optimalen Merkmalssätze der beiden Testgebiete aufgrund der unterschiedlichen Charakteristik der Sensordaten teilweise voneinander abweichen, sodass die nachfolgenden Untersuchungen auf individuellen Merkmalssätzen je Testgebiet basieren. Die Anzahl und die zur Normalisierung verwendeten Quantile stimmen jedoch in beiden Testgebieten überein. Von den bildbasierten, dreidimensionalen und geometrischen Merkmalen werden die jeweils 20 wichtigsten Merkmale für die Klassifikation der Superpixel und die 30 wichtigsten Merkmale für die Klassifikation der Landnutzungsobjekte verwendet. Der Satz von Kontextmerkmalen zur Klassifikation des semantischen Potentials höherer Ordnung besteht in der Bodenbedeckungsebene aus 19 Merkmalen (d.h. allen verfügbaren) und in der Landnutzungsebene aus 30 Merkmalen. Die Klassifikation der Bodenbedeckung basiert auf Superpixeln der Größe 400 Pixel mit einer Kompaktheit von 40, die auf Grundlage der sekundären Bildinformationen extrahiert werden. Diese Parameter haben sich in den Untersuchungen in Kapitel 5.2.2 als für diese Aufgabe geeignet herausgestellt. Für die Parameter der RF Klassifikatoren und die Anzahl der Iterationen werden die in der Tabelle 5.4 aufgelisteten Standardwerte verwendet, da sich in Kapitel 5.2.4 gezeigt hat, dass andere Werte zu keiner signifikanten Verbesserung der Ergebnisse führen. Da jedoch für die Parameter ω^i ein Einfluss festzustellen war, werden die in der Tabelle 5.9 aufgetragenen Parameterwerte für die nachfolgenden Experimente angewendet.

5.2.5.3. Beschreibung der Ergebnisse

Klassifikation der Bodenbedeckung In den Tabellen 5.10 und 5.11 sind die Ergebnisse einer quantitativen Evaluation der Klassifikationsergebnisse auf Pixelebene dargestellt für den iterativen Inferenzalgorithmus für $n_{It} = 1$ und $n_{It} = 5$ unter Einbeziehung der Gesamtheit an Kontextinformation sowie zu Vergleichszwecken für eine unabhängige Klassifikation mittels RF und für die Klassifikation unter Verwendung eines standardmäßigen CRF-Modells in den Testgebieten Hameln und Schleswig. Die Klassifikation basiert in allen Fällen auf Superpixeln der Größe 400 Pixel.

Der Vergleich der Gesamtgenauigkeitsmaße in den Tabellen 5.10 und 5.11 zeigt, dass die Integration von Kontext zu einem Genauigkeitsgewinn führt gegenüber einer unabhängigen Klassifikation. Bei Anwendung eines unabhängigen RF Klassifikators wird eine Gesamtgenauigkeit von 81,6 % in Hameln und 78,2 % in Schleswig erzielt. Aus der Berücksichtigung der räumlichen Abhängigkeiten im Rahmen eines standardmäßigen CRF-Modells resultiert eine Verbesserung der Gesamtgenauigkeit in Höhe von 1,1 % in Hameln und 5,8 % in Schleswig. Wird zusätzlich die Abhängigkeit zwischen der Bodenbedeckung und Landnutzung einmalig im Klassifikationsprozess berücksichtigt, verbessert sich die Gesamtgenauigkeit um weitere 0,6 % in Hameln und 0,3 % in Schleswig. Die iterative Prozessierung bewirkt nur in Schleswig einen leichten Anstieg der Gesamtgenauigkeit in Höhe von 0,5 %, wohingegen

	Testgebiet Hameln											
	RF_{400}			CRF_{400}			$CRF_{iter,1,400}$			$CRF_{iter,5,400}$		
	K [%]	V [%]	Q [%]	K [%]	V [%]	Q [%]	K [%]	V [%]	Q [%]	K [%]	V [%]	Q [%]
Geb.	87,9	89,6	79,7	88,7	87,5	78,7	87,2	90,4	79,9	86,9	89,8	79,1
Vers.	74,3	81,7	63,7	72,5	84,0	63,7	74,8	83,4	65,1	73,5	83,6	64,3
Bod.	76,9	70,9	58,5	92,7	78,3	73,8	95,1	77,7	74,7	95,3	78,0	75,1
Gras	86,4	80,4	71,4	87,2	80,5	72,0	87,8	81,0	72,7	88,0	80,6	72,6
Baum	78,1	82,9	67,3	77,9	86,2	69,3	78,5	86,7	70,1	78,5	86,5	69,9
Was.	91,0	85,3	78,6	96,0	84,8	81,9	96,0	84,7	81,7	95,9	83,5	80,6
Fhzg.	44,9	28,3	21,0	65,5	17,1	15,7	65,8	17,4	16,0	65,8	17,4	16,0
Sonst.	22,1	12,9	8,9	17,1	0,2	0,2	29,9	0,2	0,2	30,7	0,4	0,4
G		81,6 %			82,7 %			83,3 %			83,1 %	
κ		76,4 %			77,6 %			78,5 %			78,2 %	

Tabelle 5.10.: Gesamtgenauigkeit (G) [%], Kappa-Index (κ) [%], Vollständigkeit (V), Korrektheit (K) und Qualität (Q) [%] abgeleitet aus einer pixelbasierten Evaluation für die Bodenbedeckungsklassen *Geb.*, *Vers.*, *Bod.*, *Gras*, *Baum*, *Was.*, *Fhzg.* und *Sonst.* im Testgebiet Hameln bei Anwendung einer unabhängigen RF Klassifikation (RF_{400}), eines CRF-Modells (CRF_{400}) und der iterativen Inferenzprozedur mit 1 Iteration ($CRF_{iter,1,400}$) bzw. 5 Iterationen ($CRF_{iter,5,400}$) basierend auf Superpixeln der Größe 400 Pixel.

	Testgebiet Schleswig											
	RF_{400}			CRF_{400}			$CRF_{iter,1,400}$			$CRF_{iter,5,400}$		
	K [%]	V [%]	Q [%]	K [%]	V [%]	Q [%]	K [%]	V [%]	Q [%]	K [%]	V [%]	Q [%]
Geb.	82,7	86,8	73,4	84,8	87,6	75,7	84,5	87,9	75,7	85,1	88,2	76,3
Vers.	65,2	80,2	56,2	72,9	80,5	61,9	73,2	80,5	62,2	73,6	80,8	62,7
Bod.	85,2	47,4	43,8	95,0	81,0	77,7	96,2	80,7	78,2	96,7	83,8	81,5
Gras	75,3	79,7	63,2	82,3	84,1	71,2	82,9	84,1	71,7	83,7	84,7	72,7
Baum	83,0	88,0	74,6	85,6	91,1	79,0	85,4	91,9	79,4	85,8	91,9	79,8
Was.	96,7	67,3	65,8	98,3	73,2	72,3	98,5	73,3	72,5	98,4	73,2	72,4
Fhzg.	39,1	14,9	12,1	69,2	5,6	5,5	68,4	5,3	5,2	51,7	3,3	3,2
Sonst.	30,2	11,3	9,0	89,4	4,0	4,0	90,8	4,1	4,1	91,2	4,3	4,3
G		78,2 %			84,0 %			84,3 %			84,8 %	
κ		72,0 %			79,5 %			79,8 %			80,5 %	

Tabelle 5.11.: Gesamtgenauigkeit (G) [%], Kappa-Index (κ) [%], Vollständigkeit (V), Korrektheit (K) und Qualität (Q) [%] abgeleitet aus einer pixelbasierten Evaluation für die Bodenbedeckungsklassen *Geb.*, *Vers.*, *Bod.*, *Gras*, *Baum*, *Was.*, *Fhzg.* und *Sonst.* im Testgebiet Schleswig bei Anwendung einer unabhängigen RF Klassifikation (RF_{400}), eines CRF-Modells (CRF_{400}) und der iterativen Inferenzprozedur mit 1 Iteration ($CRF_{iter,1,400}$) bzw. 5 Iterationen ($CRF_{iter,5,400}$) basierend auf Superpixeln der Größe 400 Pixel.

die Gesamtgenauigkeit in Hameln um 0,2 % sinkt. Daraus folgert, dass der größere Beitrag aus der Berücksichtigung räumlicher Abhängigkeiten resultiert. Von dieser Art der Kontextinformation profitieren nahezu alle Klassen gleichermaßen mit Ausnahme der Klassen *Fahrzeug* und *Sonstiges* in beiden Testgebieten sowie den Klassen *Gebäude*, *Versiegelung* und *Gras* im Testgebiet Hameln. Anders verhält es sich hinsichtlich der Berücksichtigung der Abhängigkeiten zwischen der Bodenbedeckung und der Landnutzung. Von dieser Art der Kontextinformation profitieren nur einzelne Klassen. Ungeachtet kleinerer Unterschiede erreichen die Ergebnisse der kontextbasierten Klassifikationsverfahren in beiden Testgebieten jeweils ein ähnliches Genauigkeitsniveau im Bereich zwischen 83 % und 85 %.

Allerdings unterliegen die Ergebnisse für die einzelnen Klassen größeren Variationen, was sich in den klassenspezifischen Genauigkeitsmaßen, d.h. der Korrektheit, Vollständigkeit und Qualität pro Klasse, in den Tabellen 5.10 und 5.11 abbildet. Diese Genauigkeitsmaße belegen, dass alle Klassen im Testgebiet Schleswig in Bezug auf mindestens ein Genauigkeitsmaß von der Berücksichtigung der räumlichen Abhängigkeit im Rahmen der Klassifikation profitieren, was sich mit Ausnahme der Klassen *Fahrzeug* und *Sonstiges* auch in einem Anstieg der Qualität abbildet. Für die Klassen *Fahrzeug* und *Sonstiges* verbessert sich zwar die Korrektheit um mehr als 30 %, was allerdings mit Einbußen in der Vollständigkeit zwischen 7 % und 9 % einhergeht, sodass die Qualität insgesamt sinkt. Für die übrigen Klassen wird eine Steigerung der Qualität erzielt, d.h. konkret ein Anstieg in Höhe von 2,3 % für *Gebäude*, 5,7 % für *Versiegelung*, 8 % für *Gras*, 4,4 % für *Baum*, 6,5 % für *Wasser* und 33,9 % für die Klasse *Boden*, die somit den größten Genauigkeitsgewinn erzielt und folglich am meisten von der Berücksichtigung der räumlichen Abhängigkeiten profitiert. Im Testgebiet Hameln wird zwar ebenfalls die größte Qualitätsverbesserung für die Klasse *Boden* erzielt, die jedoch mit 15,3 % geringer ausfällt als in Schleswig. Daneben profitieren lediglich noch die Klassen *Baum* und *Wasser* von der Berücksichtigung der räumlichen Abhängigkeiten, was sich in einem Anstieg der Qualität in Höhe von 2 % für die Klasse *Baum* und 3,3 % für die Klasse *Wasser* manifestiert. Für die Klassen *Versiegelung* und *Gras* bleibt die Qualität auf einem ähnlichen Niveau wie bei der unabhängigen Klassifikation, d.h. die Berücksichtigung der räumlichen Abhängigkeiten ist für diese Klassen von untergeordneter Bedeutung. Ein Verlust in der Qualität ist – wie auch schon in Schleswig – für die Klassen *Fahrzeug* und *Sonstiges* zu verzeichnen. Zusätzlich büßt allerdings auch die Klasse *Gebäude* an Qualität ein (ca. 1 %), was insbesondere auf Verluste in der Vollständigkeit in Höhe von 2,1 % zurückzuführen ist.

Der einmalige Austausch von semantischer Kontextinformation hat im Testgebiet Schleswig keine Auswirkungen auf die Qualität im Vergleich zur CRF Klassifikation. Lediglich die Korrektheit steigt leicht an für die Klassen *Boden* (+1,2 %) und *Sonstiges* (+1,4 %). Im Rahmen der iterativen Prozessierung verbessert sich zusätzlich die Vollständigkeit für die Klasse *Boden* (+3,1 %) und die Korrektheit für die Klasse *Gras* (+0,8 %), sodass für diese Klassen eine Verbesserung der Qualität in Höhe von 3,3 % für die Klasse *Boden* und in Höhe von 1 % für die Klasse *Gras* im Vergleich zur einmaligen Iteration erzielt wird. Demgegenüber büßt die Klasse *Fahrzeug* im Rahmen der iterativen Prozessierung an Korrektheit und Vollständigkeit ein, sodass die Qualität um 2 % gegenüber der einmaligen Iteration sinkt. Im Testgebiet Hameln wirkt sich der einmalige Austausch der semantischen Kontextinformation positiv auf die Klassen *Gebäude* und *Versiegelung* aus, was sich in einem Anstieg der Qualität in Höhe von 1,2 % für *Gebäude* und 1,4 % für *Versiegelung* gegenüber der CRF Klassifikation abbildet. Speziell der Anstieg der Qualität für die Klasse *Gebäude* geht auf eine verbesserte Vollständigkeit (+2,9 %) zurück, die begleitet wird von leichten Einbußen in der Korrektheit (–1,5 %). Zudem verbessert sich die Korrektheit der Klassen *Boden* (+2,4 %) und *Sonstiges* (+12,8 %), was jedoch keinen Anstieg der Qualität zur Folge hat. Die iterative Prozessierung wirkt sich hingegen positiv für die Klasse *Boden* aus, so verbessert sich die Qualität im Vergleich zur CRF Klassifikation um 1,3 %.

In der Abbildung 5.9 sind beispielhaft vier Ausschnitte der Ergebnisse der Bodenbedeckungsklassifikation für das Testgebiet Hameln dargestellt, die für Superpixel der Größe 400 Pixel mit einer unabhängigen Klassifikation mittels RF, einer Klassifikation unter Anwendung eines CRF-Modells sowie des erweiterten Zwei-Ebenen-CRF-Modells mittels iterativer Prozessierung für $n_{It} = 5$ erzielt

wurden. Die Ergebnisse des iterativen Inferenzalgorithmus bei einmaliger Iteration sind aufgrund der geringen Unterschiede zur mehrfachen Iteration nicht dargestellt. Zur besseren visuellen Bewertung der Ergebnisse sind in der Abbildung 5.9 zusätzlich die pixelbasierten Referenzdaten pro Beispielausschnitt dargestellt.

Die Bilder in der ersten Zeile in Abbildung 5.9 zeigen zwei einander kreuzende Feldwege in einer landwirtschaftlichen Umgebung. Die an die Wege angrenzenden Ackerflächen sind jeweils in unterschiedlichen Stadien des Vegetationszyklus, d.h. die Ackerflächen im linken Bereich sind bereits von niedriger Vegetation bedeckt, wohingegen die Ackerflächen auf der rechten Seite noch keine Vegetation aufweisen. In allen drei Klassifikationsansätzen wird die niedrige Vegetation größtenteils korrekt der Klasse *Gras* zugewiesen. In den mit *Boden* bedeckten Bereichen sind allerdings Unterschiede zwischen den Klassifikationsergebnissen zu erkennen. Bei einer unabhängigen Klassifikation werden einige Superpixel der Klasse *Boden* fälschlicherweise als *Versiegelung* klassifiziert, was zu einem insgesamt sehr verrauschten Gesamteindruck der Klassifikationsergebnisse, speziell in diesen Bereichen, führt (vgl. Abb. 5.9b). Die Klasse *Boden* bedeckt typischerweise großräumige, homogene Flächen, woraus folgert, dass benachbarte Superpixel mit hoher Wahrscheinlichkeit ebenfalls zur Klasse *Boden* gehören. Bei Berücksichtigung der Kontextinformation über die räumliche Abhängigkeit im Rahmen der Klassifikation werden isolierte Fehlklassifikationen größtenteils eliminiert, sodass homogene Flächen gleicher Bodenbedeckung entstehen, wie in diesem Beispiel für die Klasse Boden (vgl. Abb. 5.9c). Allerdings ist zu beobachten, dass der Glättungseffekt von der Ähnlichkeit der Merkmale abhängt. So werden die Superpixel, welche exakt die mit Erde verdichteten Ackerspuren enthalten und somit hinsichtlich ihrer Merkmale von den mit Gras bedeckten Superpixeln abweichen, als *Versiegelung* klassifiziert und nicht im Sinne einer Glättung der angrenzenden Grasfläche zugewiesen (vgl. Abb. 5.9c). Davon abgesehen führt die Berücksichtigung der räumlichen Abhängigkeiten zu einer insgesamt homogeneren Lösung verglichen mit den Ergebnissen der unabhängigen Klassifikation (vgl. Abb. 5.9b und 5.9c). Die zusätzliche Berücksichtigung der vorherrschenden Landnutzung hat kaum Auswirkungen auf die Klassifikation dieser Szene (vgl. Abb. 5.9d).

Das zweite Beispiel ist in der zweiten Zeile in Abbildung 5.9 dargestellt und zeigt eine überwiegend bewaldete Szene, die von einem asphaltierten Waldweg durchkreuzt wird. Das Ergebnis des RF Klassifikators ist – wie auch schon im ersten Beispiel – relativ verrauscht, d.h. viele isolierte Superpixel werden fälschlicherweise den Klassen *Gebäude*, *Versiegelung*, *Boden*, *Gras* und *Wasser* zugewiesen. Fehlklassifikationen treten sowohl in der bewaldeten als auch in der versiegelten Fläche auf (vgl. Abb. 5.9f). Die korrekte Zuordnung der Superpixel zur Klasse *Baum* scheitert in den meisten Fällen daran, dass die Bäume kein Laub haben und daher nur sehr schwer von anderen Klassen zu unterscheiden sind, wie z.B. den Klassen *Boden* oder *Gras*, die typischerweise den Waldboden bedecken. Superpixel der Klasse *Baum* bedecken zumeist großräumige Flächen, sodass benachbarte Superpixel bei ähnlichen Eigenschaften mit hoher Wahrscheinlichkeit zur gleichen Klasse gehören. Durch die Berücksichtigung dieser Abhängigkeit im Rahmen der CRF Klassifikation werden die Ergebnisse korrigiert, indem viele isolierte Fehlklassifikationen in dem Waldgebiet der Klasse *Baum* zugewiesen werden (vgl. Abb. 5.9g). Eine Verbesserung ist auch für die Klasse *Versiegelung* zu beobachten. So werden insbesondere die Superpixel der Klasse *Wasser* sowie der überwiegende Teil der Gebäudeflächen innerhalb der Straßenfläche in Versiegelungsflächen umgewandelt (vgl. Abb. 5.9g). Durch die Berücksichtigung

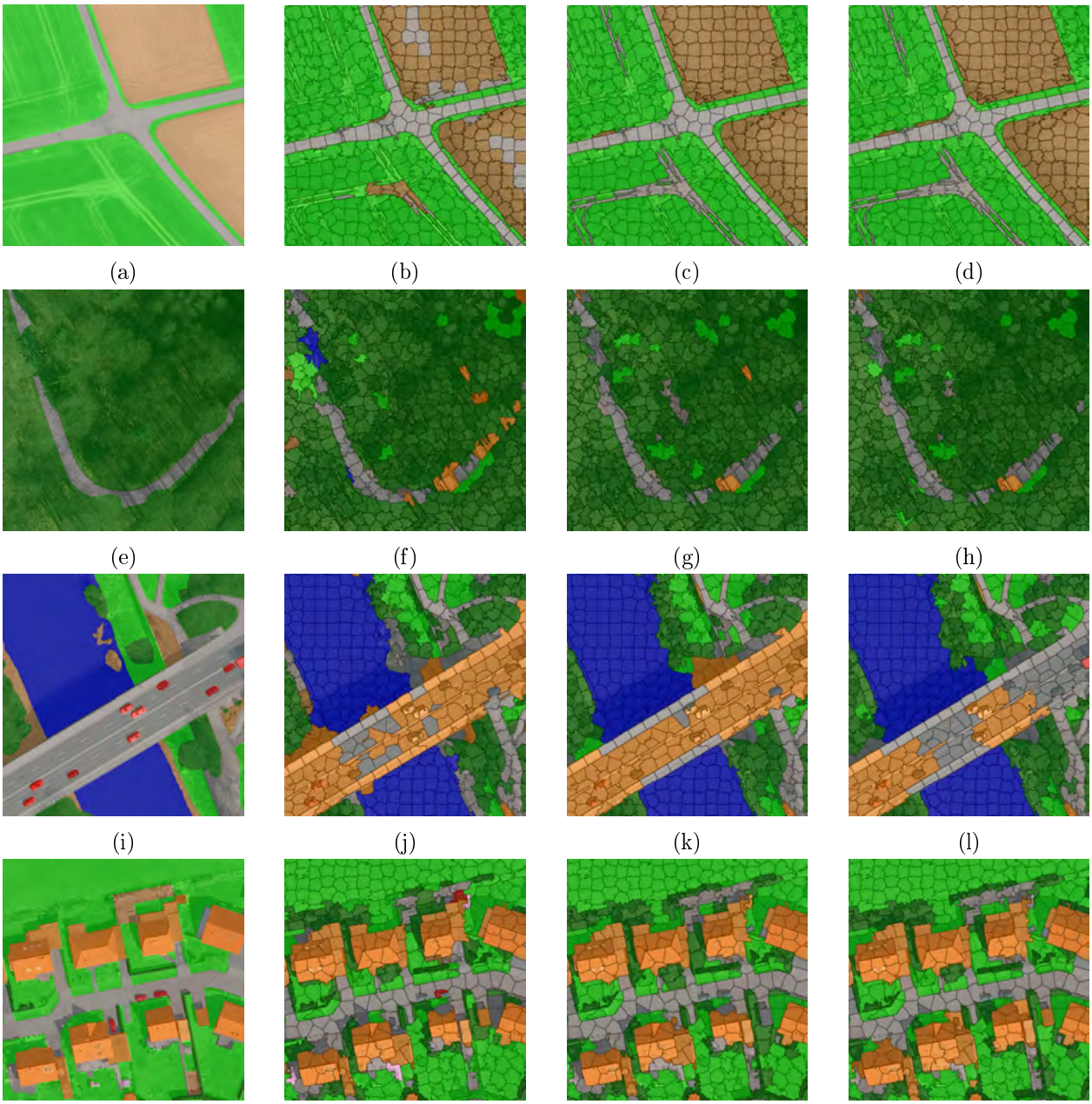


Abbildung 5.9.: Bilder von vier verschiedenen Szenen (jede Zeile repräsentiert eine Szene), welche die Referenzdaten pro Pixel (erste Spalte: a,e,i,m) und die Klassifikationsergebnisse pro Superpixel eines unabhängigen RF Klassifikators (zweite Spalte: b,f,j,n), einer CRF Klassifikation (dritte Spalte: c,g,k,o) und des Zwei-Ebenen-CRF-Modells unter Anwendung des iterativen Inferenzalgorithmus (vierte Spalte: d,h,l,p) als Überlagerung eines Orthophotos darstellen. Die Klassifikation basiert jeweils auf Superpixeln der Größe 400 Pixel, deren Begrenzungen linienhaft angedeutet sind. Die Farben markieren die Zugehörigkeit zur Bodenbedeckungsklasse: *Geb.* (orange), *Vers.* (grau), *Bod.* (braun), *Gras* (hellgrün), *Baum* (dunkelgrün), *Was.* (blau), *Fhzg.* (rot) und *Sonst.* (rosa).

der aktuell vorliegenden Nutzungsart im Rahmen der iterativen Inferenzprozedur werden unwahr-

scheinliche Bodenbedeckungsarten eliminiert und durch jene Klassen ersetzt, die typischerweise in den jeweiligen Landnutzungsobjekten vorkommen, z.B. die Bodenbedeckungsklasse *Baum* in Waldflächen (vgl. Abb. 5.9h). Folglich werden Fehlklassifikationen in den Waldflächen weiter reduziert, womit der iterative Inferenzalgorithmus eine realistischere Abbildung der realen Welt liefert. Nichtsdestotrotz bleiben vereinzelt Fehlklassifikationen bestehen, wie z.B. die Gebäudeflächen auf der Straße sowie versiegelte Flächen im Wald (vgl. Abb. 5.9h), deren Merkmale offensichtlich deutlich die jeweilige Klasse unterstützen, sodass zusätzliche Kontextinformation keinen Einfluss hat.

Die Bilder in der dritten Zeile in Abbildung 5.9 zeigen einen Flusslauf, der von einer Brücke überlagert ist. Ohne die Berücksichtigung von Kontext treten insbesondere am Übergang zwischen Wasser- und Landfläche bzw. an den Rändern der Brücke vermehrt Fehlklassifikationen auf. Superpixel werden in diesen Bereichen häufig den Klassen *Gebäude* oder *Versiegelung* zugewiesen (vgl. Abb. 5.9j). Bei näherer Betrachtung fällt auf, dass sich innerhalb dieser Superpixel häufig zwei Bodenbedeckungsarten partiell überlagern (z.B. am rechten Ufer oberhalb der Brücke); konkret handelt es sich hierbei um Sträucher, die zumindest teilweise die Wasseroberfläche verdecken. Infolgedessen sind die Eigenschaften beider Bodenbedeckungsarten, insbesondere im Hinblick auf die NDVI-Information, in den zugehörigen Merkmalswerten vermischt. Die Merkmale sind somit nicht eindeutig einer Klasse zuzuordnen. Dies hat unter Umständen zur Folge, dass der unabhängige RF Klassifikator, dessen Klassifikationsentscheidung ausschließlich auf den Merkmalen basiert, die Superpixel fälschlicherweise den Klassen *Gebäude* und *Versiegelung* zuordnet. Die Berücksichtigung der räumlichen Abhängigkeiten im Rahmen der Klassifikation unter Verwendung eines CRF-Modells hat zur Folge, dass diese Superpixel der angrenzenden Klasse zugewiesen werden mit denen die Merkmalswerte am ehesten übereinstimmen, d.h. in diesem Fall überwiegend der Klasse *Baum* (vgl. Abb. 5.9k). Das Ergebnis wirkt geglättet. Isolierte Fehlklassifikationen sind größtenteils eliminiert. Mit der Erweiterung der Kontextinformation um die Komponente der semantischen Abhängigkeit zwischen der Bodenbedeckung und der Landnutzung stehen im Rahmen der Klassifikation die Art und die geometrische Abgrenzung der Landnutzung zur Verfügung. Hier hat insbesondere das Wissen über den Verlauf der Gewässerkante den Effekt, dass die Superpixel in dem Übergangsbereich mehrheitlich korrekt der Wasser- oder Landfläche zugewiesen werden (vgl. Abb. 5.9l). Der gleiche Effekt ist am Beispiel der Straße zu beobachten. Das Wissen über die Existenz und die Ausdehnung der Straße bestärkt den Klassifikator darin, den Superpixeln die Klasse *Versiegelung* anstatt *Gebäude* zuzuweisen, trotz eines ähnlichen spektralen Erscheinungsbildes (vgl. rechter Teil der Straße in Abb. 5.9l). Dieser Effekt ist allerdings nicht durchgehend für den gesamten Straßenverlauf zu beobachten, was zwei verschiedene Ursachen hat: Zum einen wurde dem Brückenabschnitt im linken Teil der Abbildung 5.9l im Rahmen des iterativen Inferenzalgorithmus fälschlicherweise die Nutzungsart *Sonstige Bebauung* zugewiesen, sodass die Bodenbedeckung *Gebäude* in diesem Bereich als wahrscheinlicher eingestuft wird gegenüber der Bodenbedeckung *Versiegelung* (vgl. Abb. 5.9l). Zum anderen ist der Brückenabschnitt, der den Fluss überlagert, nicht in einem separaten Landnutzungsobjekt ausgewiesen, sondern bildet als Überlagerungsobjekt gemäß der Erfassungsrichtlinien[4] einen Bestandteil der darunter liegenden Nutzung, d.h. des Flusses. Infolgedessen werden die Superpixel zwar größtenteils korrekt der Klasse *Versiegelung* zugeordnet, jedoch treten

[4] Unveröffentlichte, interne Arbeitsanweisung der Vermessungs- und Katasterverwaltung Niedersachsen vom 16.02.2016: Erhebung der Topografie für den Nachweis im Liegenschaftskataster (Topografie-Handbuch).

insbesondere in den Randbereichen Ungenauigkeiten auf (vgl. Abb. 5.9l).

Das letzte Beispiel in der vierten Zeile der Abbildung 5.9 dient der Visualisierung der Grenzen des Verfahrens am Beispiel einer urbanen Szene. Die Evaluation der Ergebnisse anhand der Genauigkeitsmaße hat ergeben, dass die Korrektheit der Klasse *Gebäude* bei Anwendung des iterativen Inferenzalgorithmus sinkt, wohingegen die Vollständigkeit zunimmt. Dieser Effekt resultiert in erster Linie aus Unterschieden an den Gebäuderändern (vgl. Abb. 5.9p). Die Gebäude werden zwar an einigen Stellen durch das Hinzufügen von Superpixeln zur Gebäudefläche im Vergleich zur unabhängigen Klassifikation und CRF Klassifikation vervollständigt, jedoch werden vermehrt auch Superpixel hinzugefügt, die überwiegend zu einer anderen Klasse gehören (z.B. drittes Gebäude von links in der unteren Reihe in Abb. 5.9p). Fehlklassifikationen dieser Art treten insbesondere dann auf, wenn die Superpixel ein ähnliches spektrales Erscheinungsbild aufweisen, wie es für die Klasse *Versiegelung* typischerweise der Fall ist. Enthalten diese Superpixel kleine Anteile von Gebäudeflächen, führt die Fehlzuweisung zu einer Verbesserung der Vollständigkeit und gleichzeitig zu einer Verschlechterung der Korrektheit für die Klasse *Gebäude*.

Klassifikation der Landnutzung In den Tabellen 5.12 und 5.13 sind die Ergebnisse einer quantitativen Evaluation der Klassifikationsergebnisse auf Basis der GIS-Objekte angegeben für den iterativen Inferenzalgorithmus unter Einbeziehung der Gesamtheit an Kontextinformation für $n_{It} = 1$ und $n_{It} = 5$ sowie zu Vergleichszwecken für eine unabhängige Klassifikation mittels RF und für die Klassifikation unter Verwendung eines standardmäßigen CRF-Modells in den Testgebieten Hameln und Schleswig. Die Landnutzungsklassifikation bei Anwendung des iterativen Inferenzalgorithmus (mit einmaliger und fünfmaliger Iteration) basiert auf den Ergebnissen einer Bodenbedeckungsklassifikation für Superpixel der Größe 400 Pixel.

	Testgebiet Hameln												
	Anz.	RF			CRF			$CRF_{iter,1,400}$			$CRF_{iter,5,400}$		
	Obj.	K [%]	V [%]	Q [%]	K [%]	V [%]	Q [%]	K [%]	V [%]	Q [%]	K [%]	V [%]	Q [%]
Wbfl.	477	85,7	78,3	69,3	88,8	78,5	71,5	89,2	79,1	72,2	88,9	79,3	72,1
Sonst.	436	72,8	79,3	61,1	72,4	80,5	61,6	75,0	82,4	64,7	75,4	81,5	64,4
Grünfl.	279	65,3	79,8	56,1	65,1	77,9	54,9	66,8	79,8	57,1	66,7	80,8	57,5
Verkw.	767	78,5	75,0	62,2	84,9	71,8	63,7	83,8	73,6	64,4	85,6	72,6	64,7
Weg	596	75,8	80,8	64,2	73,0	89,9	67,5	74,6	89,9	68,8	74,1	91,1	69,1
Ackerl.	61	78,4	78,4	64,4	76,3	82,4	65,6	76,3	78,4	63,0	76,9	81,1	65,2
Grünl.	57	63,5	50,0	38,8	63,5	60,6	44,9	63,8	56,1	42,5	66,1	59,1	45,3
Wald	85	74,1	59,4	49,2	84,0	67,3	59,6	87,8	71,3	64,9	81,9	76,2	65,3
Gew.	54	83,3	43,5	40,0	95,2	29,0	28,6	95,7	31,9	31,4	100,0	27,5	27,5
G		76,2 %			77,5 %			78,6 %			78,7 %		
κ		70,9 %			72,5 %			73,8 %			74,0 %		

Tabelle 5.12.: Gesamtgenauigkeit (G) [%], Kappa-Index (κ) [%], Vollständigkeit (V), Korrektheit (K) und Qualität (Q) [%] abgeleitet aus einer objektbasierten Evaluation für die Landnutzungsklassen *Wbfl.*, *Sonst.*, *Grünfl.*, *Verkw.*, *Weg*, *Ackerl.*, *Grünl.*, *Wald*, *Gew.* im Testgebiet Hameln bei Anwendung eines unabhängigen RF Klassifikators (RF), eines CRF-Modells (CRF) und der iterativen Inferenzprozedur mit 1 Iteration ($CRF_{iter,1,400}$) bzw. 5 Iterationen ($CRF_{iter,5,400}$), jeweils basierend auf den Ergebnissen einer Bodenbedeckungsklassifikation für Superpixel der Größe 400 Pixel.

	Testgebiet Schleswig												
	Anz.	*RF*			*CRF*			$CRF_{iter,1,400}$			$CRF_{iter,5,400}$		
	Obj.	K	V	Q	K	V	Q	K	V	Q	K	V	Q
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
Wbfl.	681	76,7	77,4	62,7	79,9	80,5	66,9	82,6	82,2	70,1	82,7	80,6	69,0
Sonst.	520	62,4	66,6	47,5	64,6	69,6	50,4	68,9	72,3	54,5	68,2	73,0	54,4
Grünfl.	414	58,1	62,2	42,9	61,4	59,8	43,4	63,8	64,3	47,1	65,2	60,2	45,6
Verkw.	723	85,1	77,5	68,2	84,0	76,8	67,0	84,4	78,8	68,8	83,3	79,4	68,5
Weg	332	65,5	76,1	54,3	66,1	76,9	55,1	65,6	76,3	54,5	66,4	75,3	54,5
Ackerl.	160	90,6	73,6	68,4	90,8	65,0	61,0	91,1	57,4	54,3	94,6	53,3	51,7
Grünl.	429	78,8	83,1	67,9	73,3	87,1	66,1	72,8	88,7	66,6	69,5	89,5	64,3
Wald	306	80,7	84,1	70,0	76,7	86,9	68,8	79,3	87,2	71,1	76,2	88,1	69,0
Gew.	174	69,8	45,6	38,1	86,6	36,8	34,8	86,4	39,4	37,1	83,5	36,8	34,3
G		73,8 %			74,3 %			75,8 %			75,1 %		
κ		69,8 %			70,3 %			72,0 %			71,2 %		

Tabelle 5.13.: Gesamtgenauigkeit (G) [%], Kappa-Index (κ) [%], Vollständigkeit (V), Korrektheit (K) und Qualität (Q) [%] abgeleitet aus einer objektbasierten Evaluation für die Landnutzungsklassen *Wbfl.*, *Sonst.*, *Grünfl.*, *Verkw.*, *Weg*, *Ackerl.*, *Grünl.*, *Wald*, *Gew.* im Testgebiet Schleswig bei Anwendung eines unabhängigen RF Klassifikators (*RF*), eines CRF-Modells (*CRF*) und der iterativen Inferenzprozedur mit 1 Iteration ($CRF_{iter,1,400}$) bzw. 5 Iterationen ($CRF_{iter,5,400}$), jeweils basierend auf den Ergebnissen einer Bodenbedeckungsklassifikation für Superpixel der Größe 400 Pixel.

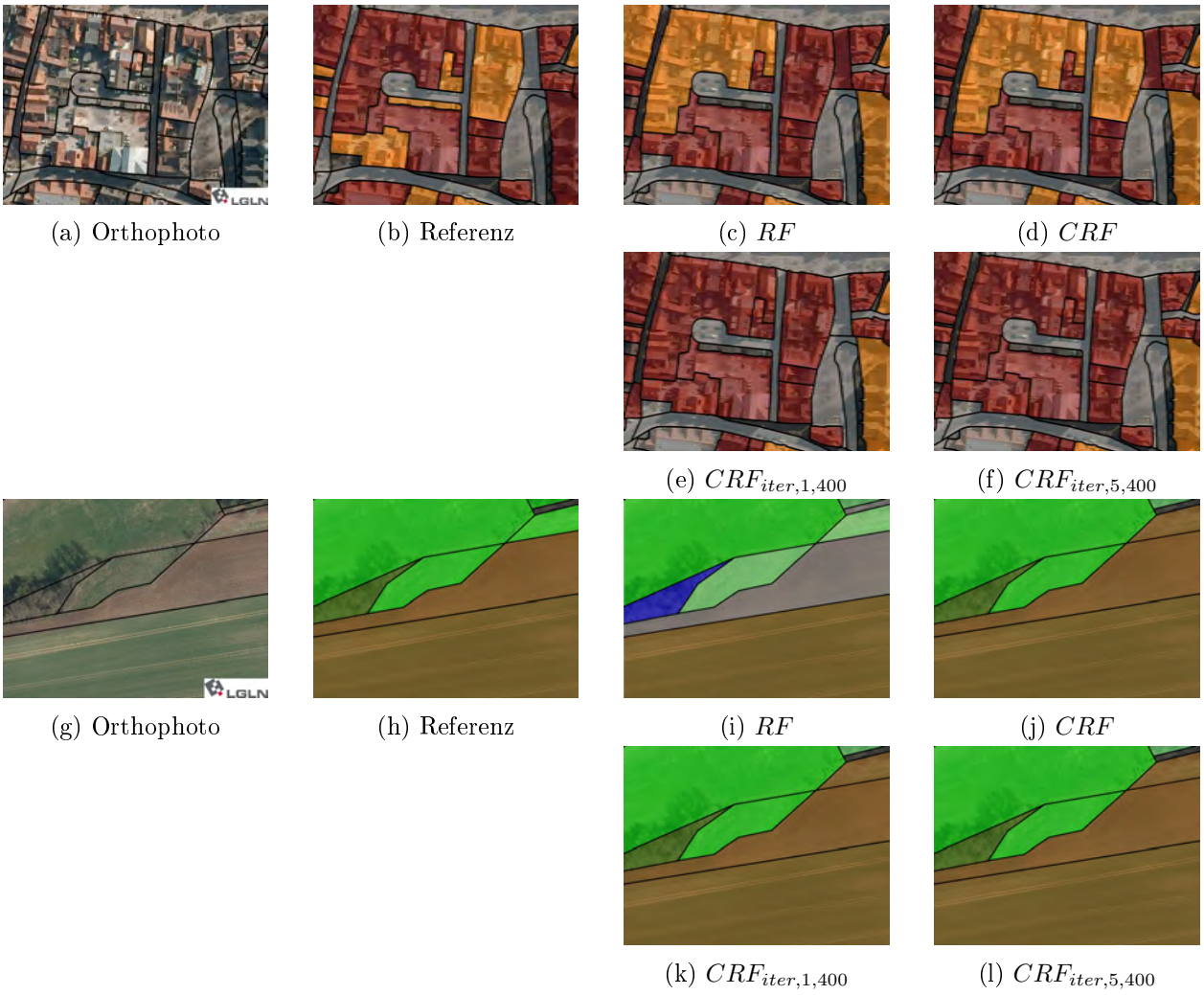
Die Gesamtgenauigkeit der Ergebnisse des iterativen Inferenzalgorithmus mit $n_{It} = 5$ beträgt 78,7 % in Hameln und 75,1 % in Schleswig und ist damit auf einem relativ hohen Niveau. Die kontextbasierten Verfahren erzielen jeweils eine Genauigkeitssteigerung gegenüber der unabhängigen RF Klassifikation von mindestens 0,5 %; speziell für den iterativen Ansatz beträgt der Gewinn 2,5 % im Testgebiet Hameln und 1,3 % im Testgebiet Schleswig. Die Ergebnisse des iterativen Inferenzalgorithmus bei einmaliger oder mehrfacher Iteration unterscheiden sich in beiden Testgebieten kaum in Bezug auf die Gesamtgenauigkeit, weisen jedoch bei näherer Betrachtung Unterschiede in den klassenspezifischen Genauigkeitsmaßen auf. Im Vergleich zur CRF Klassifikation verbessert sich die Gesamtgenauigkeit beider Varianten des iterativen Inferenzalgorithmus um mindestens 1,1 %.

Die Berücksichtigung der räumlichen Abhängigkeit wirkt sich in beiden Testgebieten unterschiedlich auf die Klassifikationsergebnisse aus. In Hameln profitieren, mit Ausnahme der Klassen *Sonstige Bebauung*, *Urbane Grünfläche* und *Gewässer*, alle Klassen von dieser Kontextinformation, wobei Qualitätsverbesserungen in Höhe von 2,2 % für *Wohnbaufläche*, 1,5 % für *Verkehrsweg*, 3,3 % für *Weg*, 1,2 % für *Ackerland*, 6,1 % für *Grünland* und 10,5 % für *Wald* erzielt werden. Die Klassen *Wald* und *Grünland* profitieren somit am meisten von der Berücksichtigung der räumlichen Abhängigkeiten im Rahmen der Klassifikation. Die Qualität für die Klasse *Sonstige Bebauung* bleibt auf einem ähnlichen Niveau zur unabhängigen Klassifikation. Für die Klasse *Urbane Grünfläche* verschlechtert sich die Qualität um 1,2 %, was insbesondere auf Einbußen in der Vollständigkeit (–1,9 %) zurückzuführen ist. Von der betragsmäßig größten Verschlechterung der Qualität in Höhe von 11,4 % ist die Klasse *Gewässer* betroffen, die bei einer Verbesserung der Korrektheit (+11,9 %) deutlich in der Vollständigkeit einbüßt (–14,5 %). Dieser Effekt ist in gleicher Weise im Testgebiet Schleswig zu beobachten. Hier reduziert sich die Qualität für diese Klasse jedoch nur um 3,3 % gegenüber der unabhängigen RF Klassifikation. Daneben sind in diesem Testgebiet weitere Klassen von Genauigkeitseinbußen betroffen, so reduziert

sich die Qualität für die Klassen *Verkehrsweg*, *Grünland* und *Wald* um bis zu 2 % und für die Klasse *Ackerland* sogar um 7,4 % gegenüber der RF Klassifikation. Lediglich die Klassen *Wohnbaufläche* und *Sonstige Bebauung* profitieren im Testgebiet Schleswig deutlich von der Berücksichtigung der räumlichen Abhängigkeiten, was sich in einem Anstieg der Qualität in Höhe von 4,2 % für *Wohnbaufläche* und 2,9 % für *Sonstige Bebauung* ausprägt.

Wird zusätzlich die semantische Kontextinformation einmalig in die Klassifikation integriert, verbessert sich die Qualität im Vergleich zur CRF Klassifikation für die Klassen *Sonstige Bebauung*, *Urbane Grünfläche*, *Wald* und *Gewässer* in beiden Testgebieten sowie für die Klasse *Weg* in Hameln und die Klassen *Wohnbaufläche* und *Verkehrsweg* in Schleswig. Die größten Qualitätsverbesserungen werden in Hameln für die Klasse *Wald* (+5,3 %) und in Schleswig für die Klasse *Sonstige Bebauung* (+4,1 %) erzielt. Die übrigen Verbesserungen betragen in beiden Testgebieten, mit Ausnahme der Klassen *Weg* (+1,3 %) und *Verkehrsweg* (+1,8 %), zwischen 2 % und 3 %. Ein negativer Effekt ist in beiden Testgebieten lediglich für die Klasse *Ackerland* sowie für die Klasse *Grünland* in Hameln festzustellen, deren Qualität sich bei einmaliger Berücksichtigung der semantischen Kontextinformation im Klassifikationsprozess um bis zu 6,7 % in Schleswig und bis zu 2,6 % in Hameln reduziert. Die Genauigkeitseinbußen werden allerdings im Testgebiet Hameln im Zuge der fünfmaligen Iteration wieder ausgeglichen. Im Gegensatz dazu führt die iterative Prozessierung im Testgebiet Schleswig zu einer weiteren Verschlechterung der Qualität für die Klassen *Ackerland* (–2,6 %) und *Grünland* (–2,3 %), obwohl sich die Korrektheit für die Klasse *Ackerland* mit 3,5 % deutlich verbessert gegenüber der einmaligen Iteration. Insgesamt wirkt sich die iterative Prozessierung im Testgebiet Schleswig eher negativ aus, was sich in Einbußen in der Qualität für die Klassen *Wohnbaufläche*, *Urbane Grünfläche*, *Wald* und *Gewässer* ausprägt, die jedoch maximal 2,8 % betragen. Im Testgebiet Hameln werden neben der Verbesserung der Qualität für die Klassen *Ackerland* und *Grünland* nur noch einzelne Genauigkeitsmaße (Korrektheit oder Vollständigkeit) verbessert, die jedoch keinen Anstieg der Qualität bewirken.

Die Bilder in der ersten und zweiten Zeile in Abbildung 5.10 zeigen ein Beispiel für die verbesserte Differenzierung der Klassen *Wohnbaufläche* und *Sonstige Bebauung*. Im Rahmen der unabhängigen Klassifikation werden drei Landnutzungsobjekte der Klasse *Sonstige Bebauung* fälschlicherweise der Klasse *Wohnbaufläche* zugewiesen. Die Berücksichtigung von Kontext bewirkt die Auflösung von zwei Fehlklassifikationen. Zwei der TN-Objekte weisen eine dichte Bebauung auf (links oben im bebauten Block). Vegetation ist kaum enthalten. Wie den Abbildungen 5.1a und 5.1c zu entnehmen ist, tritt diese Art von Bebauung im Innenstadtbereich sowohl bei den Nutzungsarten *Wohnbaufläche* als auch *Sonstige Bebauung* auf. Infolgedessen ist sich der RF Klassifikator unsicher über die Zuordnung zu einer der Klassen, was sich in nahezu gleichen Konfidenzwerten für diese Klassen ausdrückt (55 % für die Klasse *Wohnbaufläche* und 43 % für die Klasse *Sonstige Bebauung* für das zentrale TN-Objekt) und favorisiert fälschlicherweise die Klasse *Wohnbaufläche*. Durch die Berücksichtigung der räumlichen Abhängigkeiten zu benachbarten Nutzungsarten wird eines der Landnutzungsobjekte korrekt klassifiziert (vgl. TN-Objekt links oben in Abb. 5.10d). Das zweite Landnutzungsobjekt (zweites TN-Objekt von links) wird hingegen erst bei gemeinsamer Berücksichtigung der räumlichen und semantischen Kontextinformation korrekt klassifiziert (vgl. Abb. 5.10e). Das dritte TN-Objekt (in der rechten unteren Ecke) gehört ebenfalls zur Klasse *Sonstige Bebauung*. Aufgrund des hohen Vegetationsanteils (im



(a) Orthophoto (b) Referenz (c) RF (d) CRF

(e) $CRF_{iter,1,400}$ (f) $CRF_{iter,5,400}$

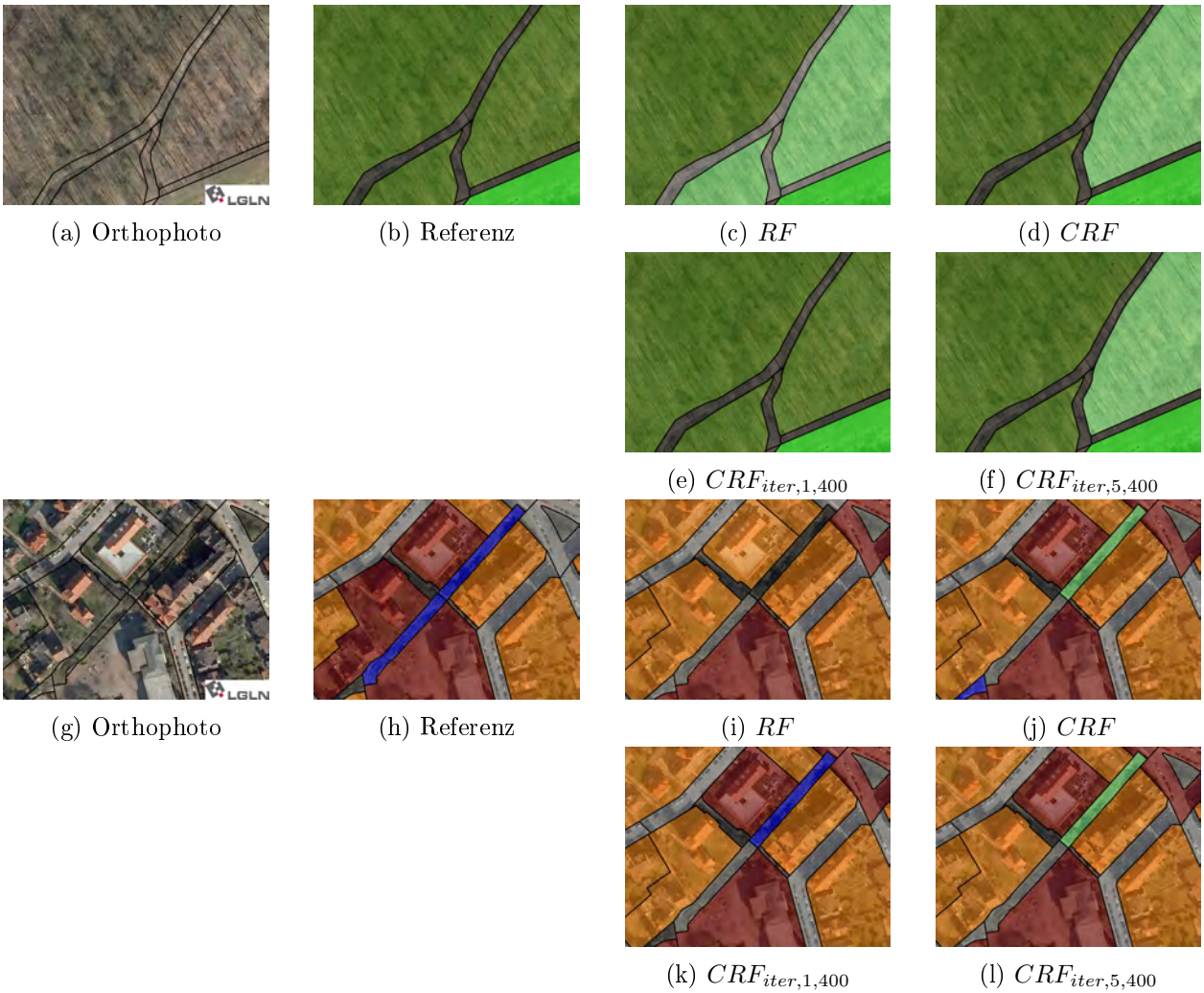
(g) Orthophoto (h) Referenz (i) RF (j) CRF

(k) $CRF_{iter,1,400}$ (l) $CRF_{iter,5,400}$

Abbildung 5.10.: Bilder von zwei verschiedenen Szenen (je zwei Zeilen repräsentieren eine Szene), welche das Orthophoto ebenso zeigt wie die Referenzdaten pro GIS-Objekt und die Klassifikationsergebnisse pro GIS-Objekt eines unabhängigen RF Klassifikators (RF), einer CRF Klassifikation (CRF) und des Zwei-Ebenen-CRF-Modells unter Anwendung des iterativen Inferenzalgorithmus mit 1 Iteration ($CRF_{iter,1,400}$) bzw. 5 Iterationen ($CRF_{iter,5,400}$), jeweils dem Orthophoto überlagert. Die Klassifikation basiert jeweils auf TN-Objekten, deren Begrenzungen linienhaft dargestellt sind. Die Farben markieren die Zugehörigkeit zur Landnutzungsklasse: *Wbfl.* (orange), *Sonst.* (rot), *Grünfl.* (hellgrün), *Verkw.* (grau), *Weg* (dunkelgrau), *Ackerl.* (braun), *Grünl.* (grün), *Wald* (dunkelgrün) und *Gew.* (blau).

dargestellten Ausschnitt nicht zu erkennen) wird dieses TN-Objekt von allen Verfahren (mit und ohne Berücksichtigung von Kontext) gleichermaßen falsch als *Wohnbaufläche* klassifiziert. In der Szene sind weitere typische Fehlklassifikation enthalten, so werden die Garagenanlagen nicht der korrekten Klasse *Wohnbaufläche* sondern der Klasse *Weg* zugewiesen. Dies liegt möglicherweise darin begründet, dass derartige TN-Objekte i.d.R. hinsichtlich ihrer länglichen Form große Ähnlichkeiten zu Verkehrsflächen haben. Darüber hinaus werden vereinzelt *Wohnbauflächen* in allen vier Klassifikationen fälschlicherweise der Klasse *Sonstige Bebauung* zugewiesen.

Die Bilder in der dritten und vierten Zeile in Abbildung 5.10 zeigen am Beispiel einer ländlich



(a) Orthophoto (b) Referenz (c) RF (d) CRF

(e) $CRF_{iter,1,400}$ (f) $CRF_{iter,5,400}$

(g) Orthophoto (h) Referenz (i) RF (j) CRF

(k) $CRF_{iter,1,400}$ (l) $CRF_{iter,5,400}$

Abbildung 5.11.: Bilder von zwei verschiedenen Szenen (je zwei Zeilen repräsentieren eine Szene), welche das Orthophoto ebenso zeigt wie die Referenzdaten pro GIS-Objekt und die Klassifikationsergebnisse pro GIS-Objekt eines unabhängigen RF Klassifikators (RF), einer CRF-Klassifikation (CRF) und des Zwei-Ebenen-CRF-Modells unter Anwendung des iterativen Inferenzalgorithmus mit 1 Iteration ($CRF_{iter,1,400}$) bzw. 5 Iterationen ($CRF_{iter,5,400}$), jeweils dem Orthophoto überlagert. Die Klassifikation basiert jeweils auf TN-Objekten, deren Begrenzungen linienhaft dargestellt sind. Die Farben markieren die Zugehörigkeit zur Landnutzungsklasse: *Wbfl.* (orange), *Sonst.* (rot), *Grünfl.* (hellgrün), *Verkw.* (grau), *Weg* (dunkelgrau), *Ackerl.* (braun), *Grünl.* (grün), *Wald* (dunkelgrün) und *Gew.* (blau).

geprägten Szene die Vorteile – speziell des räumlichen Kontexts – für die Klassifikation von landwirtschaftlichen Nutzungsarten. Ein unabhängiger RF Klassifikator weist den mittig in der Szene dargestellten TN-Objekten fälschlicherweise die Klassen *Gewässer* statt *Gehölz*, *Urbane Grünfläche* statt *Grünland* und *Verkehrsweg* statt *Ackerland* zu. Die Zuweisung der Klasse *Verkehrsweg* liegt vermutlich in der länglichen, schmalen Form der Ackerlandfläche begründet, die einem Straßenobjekt ähnelt. In Bezug auf das Grünlandobjekt ist generell in den Ergebnissen zu beobachten, dass sich der RF Klassifikator unsicher ist bezüglich der Unterscheidung zwischen *Grünland* und *Urbane Grünfläche*, was sich beispielsweise für dieses TN-Objekt in niedrigen Konfidenzwerten für beide Klassen ausprägt,

d.h. 36 % für die Klasse *Urbane Grünfläche* und 16 % für die Klasse *Grünland*. Die Berücksichtigung des räumlichen Kontexts im Klassifikationsprozess unterstützt in allen drei Fällen die korrekte Klassifikation. Für das schmale Grünlandobjekt am oberen Rand der Szene trägt die Berücksichtigung von Kontext jedoch zu keiner Verbesserung bei. Hierbei handelt es sich um einen typischen Fehler, der im Testgebiet Hameln häufiger zu beobachten ist, nämlich die fehlerhafte Klassifikation von brachliegenden Grünlandflächen zu den Klassen *Ackerland* (bei angedeuteten Bewirtschaftungsspuren) und *Urbane Grünfläche* (bei partiellen Besatz mit Sträuchern oder Bäumen).

Die Bilder in der ersten und zweiten Zeile in der Abbildung 5.11 zeigen eine Szene in einem Waldgebiet, anhand dessen die Verbesserung für die Klassen *Wald* und *Weg* beispielhaft visualisiert wird. Zwei der dargestellten Waldflächen werden im Rahmen der RF Klassifikation fälschlicherweise der Klasse *Urbane Grünfläche* zugewiesen. Beide Waldflächen werden bei einmaliger Berücksichtigung der an dieser Stelle vorliegenden Bodenbedeckung korrekt klassifiziert. Die kleinere Waldfläche profitiert bereits von der Berücksichtigung der räumlichen Nachbarschaft, im Zuge der sich die Konfidenz für die Klasse *Wald* auf den Wert 46 % erhöht. Durch die zusätzliche Berücksichtigung der Bodenbedeckung im Klassifikationsprozess erhöht sich die Konfidenz auf 99 %. Im Rahmen der iterativen Prozessierung wird jedoch allen Informationen, d.h. der Abhängigkeit von den Eingangsdaten in Form des Assoziationspotentials, dem räumlichen Kontext in Form des räumlichen Interaktionspotentials und der Abhängigkeit von der Bodenbedeckung in Form des semantischen Potentials höherer Ordnung, im Rahmen der Inferenz das gleiche Gewicht verliehen. Bei diesem Objekt wird das finale Ergebnis von der Unsicherheit des Assoziationspotentials beeinträchtigt, sodass die Konfidenz für die Klasse *Urbane Grünfläche* mit 61 % gegenüber der Konfidenz von 39 % für die Klasse *Wald* überwiegt. Darüber hinaus sind in dieser Szene mehrere Objekte der Klasse *Weg* dargestellt, die von einem RF Klassifikator fälschlicherweise der Klasse *Straße* zugeordnet werden. Hier zeigt sich, dass insbesondere die Nachbarschaft von Waldflächen die korrekte Zuordnung zur Klasse *Weg* unterstützt.

Die Bilder in der dritten und vierten Zeile in der Abbildung 5.11 visualisieren am Beispiel einer urbanen Szene die Auswirkungen von Kontext im Rahmen der Klassifikation von *Gewässer*. Die Szene wird durchkreuzt von einem Bach, der in zwei TN-Objekte aufgeteilt ist. Der RF Klassifikator weist keinem TN-Objekt die korrekte Klasse *Gewässer* zu und auch die Berücksichtigung der Nachbarschaft bewirkt keine Verbesserung. Wird hingegen die an der Stelle vorliegende Bodenbedeckung einmalig in die Klassifikation integriert, wird zumindest in einem TN-Objekt die Klasse *Gewässer* mit einem hohen Konfidenzwert von 85 % favorisiert. Im Rahmen des iterativen Inferenzalgorithmus wendet sich die Präferenz wieder zur Klasse *Urbane Grünfläche*, wie sie bereits in der CRF Klassifikation favorisiert wurde. Das andere TN-Objekt wird von allen Verfahren fälschlicherweise als Verkehrsweg klassifiziert. Hierbei handelt es sich um einen typischen Fehler in dem Testgebiet, der zum einen in der Ähnlichkeit zu den betreffenden Klassen (längliche Form, partielle Überlappung mit Bäumen) und zum anderen in einem Mangel an repräsentativen Trainingsbeispielen begründet liegt.

5.2.5.4. Diskussion

Klassifikation der Bodenbedeckung Die quantitative und qualitative Analyse der Klassifikationsergebnisse für die Bodenbedeckung hat die Annahme bestätigt, dass die Berücksichtigung von Kontext-

information zu einer signifikanten Verbesserung der Klassifikationsergebnisse beiträgt. Im Rahmen der Untersuchung hat sich gezeigt, dass sowohl die räumliche Abhängigkeit benachbarter Superpixel als auch die semantische Abhängigkeit von der vorherrschenden Landnutzung die Klassifikationsentscheidung effektiv unterstützt. Hierbei sind im Wesentlichen drei Arten von Verbesserungen festzustellen.

Die erste Verbesserung betrifft die Elimination von vereinzelten Fehlklassifikationen im Rahmen der Inferenz. Die korrekte Zuweisung einer Klasse wird dabei von zwei verschiedenen Arten von Kontextinformation hervorgerufen. Die räumliche Abhängigkeit zwischen den verschiedenen Bodenbedeckungsarten favorisiert jene Kombination von Klassen an benachbarten Superpixeln, die unter Berücksichtigung ihrer Merkmale mit hoher Wahrscheinlichkeit gemeinsam auftreten. Die Abhängigkeit von der vorliegenden Nutzungsart bevorzugt hingegen jene Bodenbedeckungsarten, die mit hoher Wahrscheinlichkeit in den jeweiligen Landnutzungsobjekten vorkommen, gegenüber den Klassen, die für diese Art von Landnutzung eher untypisch sind. Beide Arten von Kontextinformation bewirken ein homogeneres Erscheinungsbild der Klassifikationsergebnisse im Vergleich zu einer unabhängigen Klassifikation. Die Ergebnisse wirken geglättet. Von der Information über die vorliegende Nutzungsart profitieren insbesondere jene Bodenbedeckungsarten, die aufgrund eines ähnlichen spektralen Erscheinungsbildes nur sehr schwer voneinander zu unterscheiden sind, wie z.B. die Klassen *Versiegelung* und *Boden*, oder wenn die Sensordaten, z.B. aufgrund ungünstiger Aufnahmebedingungen, nur eine begrenzte Aussagekraft haben, wie z.B. für die Klasse *Baum* bei einer Befliegung zum Zeitpunkt fehlender Belaubung.

Die zweite Verbesserung ist an den Grenzen der Landnutzungsobjekte zu beobachten. Hier unterstützt die geometrische Abgrenzung der Landnutzungsobjekte die korrekte Zuordnung von Bodenbedeckungsarten im Grenzverlauf. Dies ist insbesondere dann vorteilhaft, wenn die Eigenschaften verschiedener Bodenbedeckungsarten in den Merkmalswerten eines Superpixels vermischt sind, z.B. infolge einer partiellen Überlagerung zweier Bodenbedeckungsarten.

Die dritte Verbesserung betrifft die geometrische Abgrenzung von Bodenbedeckungssegmenten, insbesondere von Gebäuden, innerhalb eines Landnutzungsobjekts. Mit der Information über die räumliche Abhängigkeit werden bereits viele Lücken in den Gebäudesegmenten geschlossen, wodurch sich die Abgrenzung der Gebäude verbessert. Die Berücksichtigung der Landnutzungsart im Rahmen der Inferenz bewirkt einerseits die Schließung noch vorhandener Lücken in den Gebäudesegmenten, und verursacht andererseits den Nebeneffekt, dass angrenzende Superpixel mit ähnlichen spektralen Eigenschaften fälschlicherweise zur Gebäudefläche hinzugefügt werden. Hierbei wird deutlich, dass der Klasse *Gebäude* speziell in bebauten Nutzungsarten im Rahmen der Inferenz ein höheres Gewicht verliehen wird als den übrigen Klassen. Infolgedessen wird für die Klasse *Gebäude* ein hoher Grad an Vollständigkeit erzielt, der aber zulasten der Korrektheit geht.

Demgegenüber wirkt sich die Berücksichtigung von Kontext im Rahmen der Klassifikation nachteilig auf bestimmte Klassen aus, wie z.B. die Klassen *Fahrzeug* und *Sonstiges*, die typischerweise nur kleinräumige Flächen bedecken. Diese Klassen sind nur selten in den Trainingsdaten vertreten, was dazu führt, dass der Klassifikator deren Vorkommen als unwahrscheinlich einstuft.

Klassifikation der Landnutzung Grundsätzlich lässt sich aus den durchgeführten Experimenten schließen, dass die Berücksichtigung von Kontext einen positiven Effekt auf die Genauigkeit der Landnutzungsklassifikation hat. Folglich trifft die Annahme, dass sich Kontext positiv auf die Klassifikation auswirkt, auch für die Klassifikation der Landnutzung zu, wenngleich nicht alle Klassen gleichermaßen und unter allen Umständen von Kontext profitieren. Die Berücksichtigung von Kontext wirkt sich innerhalb eines Testgebiets unterschiedlich auf einzelne Klassen aus, was insbesondere in den unterschiedlichen Ausprägungen der Nutzungsarten begründet liegt. Zudem prägen sich die Auswirkungen in den einzelnen Testgebieten unterschiedlich aus, was insbesondere auf die unterschiedliche Beschaffenheit der Testdaten zurückzuführen ist. Folglich lassen sich aus den Ergebnissen keine generellen Aussagen sowohl in Bezug auf die Gesamtaufgabe als auch für einzelne Klassen ableiten.

Die Klassifikation im Testgebiet Hameln profitiert in besonderem Maße von der zusätzlichen Kontextinformation, was insbesondere auf die geringe Anzahl an Trainingsbeispielen sowie den für die Klassifikation von Waldflächen ungeeigneten Erfassungszeitpunkt der Luftbilddaten zurückzuführen ist. Mit Ausnahme der Klasse *Gewässer*, die Qualität einbüßt, sowie der Klasse *Urbane Grünfläche*, für die Kontext zu keiner Veränderung beiträgt, profitieren alle Klassen von mindestens einer Art der Kontextinformation. Ein ähnliches Bild zeigt sich im Testgebiet Schleswig, wenngleich die Ergebnisse insgesamt etwas heterogener sind. Neben der Klasse *Gewässer*, sind die Klassen *Ackerland* und *Grünland* von Einbußen in der Qualität betroffen. Die Genauigkeitseinbußen für die Klassen *Ackerland* und *Grünland* sind ebenfalls auf einen zur Unterscheidung dieser Klassen ungeeigneten Erfassungszeitpunkt zurückzuführen. Alle anderen Klassen profitieren – wie auch schon in Hameln – von mindestens einer Art der Kontextinformation.

Eine Verbesserung wird insbesondere für jene Landnutzungsobjekte erzielt, denen aufgrund ihrer untypischen Eigenschaften oder ihrer Ähnlichkeit zu anderen Landnutzungsarten vom RF Klassifikator eine falsche Klasse zugewiesen wird. In diesen Fällen hilft Kontext den betreffenden TN-Objekten das korrekte Klassenlabel zuzuweisen.

Im Gegensatz zur Berücksichtigung von Kontext, deren positiver Effekt für nahezu alle Klassen festzustellen ist, variieren die Ergebnisse kaum in Abhängigkeit der gewählten Prozessierungsstrategie. Insgesamt lässt sich festhalten, dass einzelne Klassen von einer iterativen Prozessierung profitieren, wie z.B. die Klasse *Wald* im Testgebiet Hameln. Für Landnutzungsobjekte dieser Klasse liefert ein RF Klassifikator aufgrund der ungünstigen Eigenschaften der Sensordaten ungenaue Informationen über die Bodenbedeckung. Im Rahmen der iterativen Prozessierung wird kontinuierlich das Ergebnis der Bodenbedeckungsklassifikation innerhalb der Landnutzungsobjekte verfeinert. Dies hat zur Folge, dass in jeder Iteration aussagekräftigere Kontextinformationen zur Verfügung stehen, die wiederum die Landnutzungsklassifikation unterstützen.

Demgegenüber führt die iterative Inferenzprozedur zu Einbußen in der Qualität der Klassen mit den wenigsten Trainingsbeispielen innerhalb der Testgebiete, wie z.B. die Klasse *Gewässer*. Die zugehörigen Landnutzungsobjekte werden zunehmend Klassen mit ähnlichen Eigenschaften zugeordnet, die häufiger in den Trainingsdaten vorkommen. Folglich sind diese Klassen nicht adäquat in den Trainingsdaten repräsentiert, sodass kontextuelle Abhängigkeiten nicht angemessen gelernt werden können.

6. Fazit und Ausblick

6.1. Fazit

In dieser Arbeit wurde ein Verfahren zur simultanen Klassifikation der Bodenbedeckung und der Landnutzung anhand von aktuellen Sensordaten vorgestellt. Die Klassifikation basiert auf einem gemeinsamen graphischen Modell, d.h. der Vereinigung beider Klassifikationsaufgaben in einem Zwei-Ebenen-CRF-Modell. Das CRF modelliert verschiedene Arten von Kontextinformation: Räumliche Abhängigkeiten sind modelliert als paarweises Interaktionspotential jeweils separat in beiden Ebenen. Komplexe semantische Abhängigkeiten zwischen der Bodenbedeckung und der Landnutzung werden als Potential höherer Ordnung modelliert. Die Inferenz in dem CRF höherer Ordnung wird effizient durch einen iterativen Inferenzalgorithmus realisiert.

In Bezug auf den ersten Beitrag dieser Arbeit, d.h. die Integration der Bodenbedeckung und der Landnutzung in einem gemeinsamen graphischen Modell, haben die Experimente gezeigt, dass die Berücksichtigung von Kontext im Allgemeinen zu einer Verbesserung der Klassifikationsergebnisse beiträgt. Für die Bodenbedeckungsklassifikation wird ein homogeneres Ergebnis erzielt im Vergleich zu einer unabhängigen Klassifikation. Isolierte Fehlklassifikationen werden weitestgehend eliminiert. Zudem verbessert sich die Abgrenzung von Bodenbedeckungssegmenten. Bei der Landnutzungsklassifikationen werden speziell für Landnutzungsobjekte mit untypischen Eigenschaften oder Ähnlichkeiten zu anderen Landnutzungsklassen Verbesserungen durch die Berücksichtigung von Kontext erzielt. Grundsätzlich lässt sich in den Ergebnissen erkennen, dass nicht alle Klassen in gleichem Maße von der Berücksichtigung von Kontext profitieren. Im Rahmen der Bodenbedeckungsklassifikation resultiert der größere Genauigkeitsgewinn aus der räumlichen Kontextinformation; lediglich für die Klasse *Boden* bewirkt die Berücksichtigung der semantischen Kontextinformation einen weiteren Anstieg der Qualität. Ein anderes Bild zeigt sich für die Landnutzungsklassifikation. Hierbei profitieren zwar ebenfalls eine Vielzahl von Klassen von der räumlichen Kontextinformation, jedoch ist für eine größere Anzahl von Klassen auch ein positiver Effekt durch die Berücksichtigung der semantischen Kontextinformation zu beobachten, speziell für die Nutzungsarten *Wohnbaufläche*, *Sonstige Bebauung* und *Urbane Grünfläche* sowie für die Klasse *Wald*.

In Bezug auf den zweiten Beitrag des Verfahrens, und zwar den iterativen Inferenzalgorithmus, hat sich in den Experimenten gezeigt, dass eine mehrmalige Iteration gegenüber einem einmaligen Austausch von Kontextinformation keinen Mehrwert bietet. In beiden Fällen wird für die meisten Klassen ein ähnliches Genauigkeitsniveau erreicht. Daneben existieren vereinzelt Klassen, auf die sich die iterative Prozessierung positiv oder negativ auswirkt. Insgesamt lässt sich aus den Experimenten folgern, dass die allgemeine Vorgehensweise des iterativen Inferenzalgorithmus grundsätzlich für den

Austausch von Kontextinformation in dem CRF höherer Ordnung geeignet ist, jedoch eine einmalige Iteration ausreichend ist. Da es sich bei dem iterativen Inferenzalgorithmus mit einmaliger Iteration um eine Variante des vielfach in der Literatur genutzten zweistufigen Ansatzes handelt, haben die Experimente die Effektivität dieser etablierten Vorgehensweise bestätigt.

In Bezug auf den dritten Beitrag, d.h. die zur Klassifikation der semantischen Potentiale höherer Ordnung verwendeten Kontextmerkmale, haben die Experimente die Relevanz der Merkmale unterstrichen, bei denen die Konfidenzwerte im Rahmen der Merkmalsextraktion als Gewichte berücksichtigt werden. Diese Merkmale zählen sowohl bei der Klassifikation der Bodenbedeckung als auch der Landnutzung zu den wichtigsten Kontextmerkmalen.

Des Weiteren haben die Experimente gezeigt, dass die Größe der Superpixel einen Einfluss auf den Detailgrad, den Glättungseffekt und die Rechenzeit hat. Der Glättungseffekt, der aus der Segmentierungsmethode resultiert, wirkt sich besonders vorteilhaft aus für Bodenbedeckungsarten, die großräumige, homogene Flächen bedecken, wohingegen Bodenbedeckungsarten, die typischerweise kleinräumige Strukturen repräsentieren, von einem hohen Detailgrad profitieren.

Durch die Evaluation anhand von zwei verschiedenen Testgebieten konnte gezeigt werden, dass die Methode flexibel an unterschiedliche Gegebenheiten angepasst werden kann.

6.2. Ausblick

Zur Verbesserung der Ergebnisse des Zwei-Ebenen-CRF-Modells zur simultanen Klassifikation der Bodenbedeckung und Landnutzung sind diverse Weiterentwicklungen vorstellbar, die im Folgenden näher ausgeführt werden.

Eingangsdaten Der vorgestellte Ansatz ist speziell für die Verwendung von hochauflösenden Orthophotos sowie Höheninformation konzipiert. Allerdings bietet das Verfahren bereits ein hohes Maß an Flexibilität und lässt sich ohne großen Aufwand auch auf andere Sensordaten übertragen oder um weitere Sensordaten ergänzen. Hierfür ist es lediglich erforderlich, die Merkmale und Bildprimitive an die jeweiligen Sensordaten anzupassen. Durch diesen Schritt ist es möglich, von den Vorteilen anderer Sensordaten zu profitieren, wie z.B. der höheren zeitlichen Auflösung von Satellitenbilddaten, um beispielsweise Änderungen in der Landnutzung in kürzeren Abständen erfassen zu können.

Darüber hinaus hat sich im Rahmen der Experimente gezeigt, dass Fehlklassifikationen häufig auf Ungenauigkeiten der Eingangsdaten zurückzuführen sind. Daher liegt die Vermutung nahe, dass bereits eine höhere Qualität der aktuell verwendeten Eingangsdaten zu einer Verbesserung der Klassifikationsergebnisse beiträgt, wie z.B. durch die Verwendung eines *True-Orthophotos* oder aus flugzeuggestützten Laserscanning abgeleiteten Oberflächenmodellen.

Merkmale Das Klassifikationsergebnis wird maßgeblich durch die verwendeten Merkmale beeinflusst. Die Differenzierung von Klassen ist insbesondere dann eingeschränkt, wenn sich die Klassen nicht signifikant hinsichtlich ihrer Merkmalswerte unterscheiden. Die Hinzunahme weiterer Merkmale kann

hierbei Abhilfe schaffen und zu einer verbesserten Trennbarkeit von Klassen beitragen.

Statt der Erweiterung des Merkmalssatzes gilt es alternativ zu prüfen, ob die vorhandenen Merkmale verfeinert werden können, um deren individuelle Aussagekraft zu erhöhen. Von eingeschränkter Aussagekraft sind beispielsweise die Kontextmerkmale, die sich aus dem Ergebnis der Landnutzung ableiten. Diese Merkmale geben keinen Aufschluss darüber, an welcher Position innerhalb des TN-Objekts die Bodenbedeckungsarten lokalisiert sind. Es wird ausschließlich eine allgemeine Aussage getroffen zur Wahrscheinlichkeit des Vorkommens bestimmter Bodenbedeckungsarten innerhalb eines Landnutzungsobjekts. Diese Merkmale helfen zwar bestimmte Klassen auszuschließen, können aber wenig zur Verbesserung der räumlichen Verteilung innerhalb der TN-Objekte beitragen. Gould et al. [2008] präsentieren ein Modell, das die relative geometrische Anordnung von Segmenten berücksichtigt. Die Klassifikation erfolgt in zwei Stufen. In einem ersten Schritt wird eine unabhängige Klassifikation von Segmenten durchgeführt. Das Ergebnis bildet die Grundlage zur Bestimmung von Wahrscheinlichkeitsverteilungen für die relative Anordnung von Segmenten unterschiedlicher Klassen. Die Klassifikation des zweiten Schrittes basiert auf einem CRF, in dem die gelernte Wahrscheinlichkeitsverteilung der relativen Anordnung von Segmenten als a priori Information (*relative location prior*) einfließt. Dieses Vorgehen lässt sich allerdings nicht in einfacher Weise auf die Klassifikation der Bodenbedeckung übertragen. Im Gegensatz zu der Klassifikation von terrestrischen Bildern, in denen es eine eindeutige Bezugsrichtung gibt (Lotrichtung), anhand derer die Wahrscheinlichkeitsverteilung modelliert werden kann, liegt eine solche eindeutige Bezugsrichtung bei der Klassifikation von Fernerkundungsdaten nicht vor.

Demgegenüber beschränkt sich die geminderte Aussagekraft für bestimmte Merkmale nur auf räumlich begrenzte Bereiche, was an den Stellen meist auf Qualitätsmängel der Eingangsdaten an den Stellen zurückzuführen ist, z.B. für die dreidimensionalen Merkmale infolge von Ungenauigkeiten an Höhensprüngen. Abseits von diesen Bereichen tragen diese Merkmale jedoch wertvolle Information zur Klassifizierung bei, sodass von einem generellen Ausschluss abzuraten ist. Um die Ungenauigkeiten jedoch in angemessener Weise im Klassifikationsprozess zu berücksichtigen, wäre eine individuelle und räumlich variable Gewichtung der Merkmale vorteilhaft, wobei sich das Gewicht beispielsweise aus der Qualität der zugrundeliegenden Eingangsdaten ableiten lässt.

Die Extraktion bzw. Selektion von Merkmalen birgt das generelle Problem, dass ein gewisses Maß an Vorwissen über die Eigenschaften der einzelnen Klassen und ihrer Abhängigkeiten erforderlich ist, um geeignete Merkmale zu erhalten. Diese Problematik betrifft auch die Modellierung der komplexen Abhängigkeiten mit semantischen Potentialen höherer Ordnung. Im Rahmen des iterativen Inferenzalgorithmus werden die Potentiale höherer Ordnung zu unären Termen vereinfacht, deren Bestimmung auf Kontextmerkmalen beruht. Als Alternative zur manuellen Auswahl von Merkmalen, die das Risiko der unzureichenden Repräsentation birgt, haben sich in den letzten Jahren Verfahren zum Lernen von Merkmalen etabliert, z.B. [Farabet et al., 2013]. Hierbei werden komplexe Zusammenhänge von Pixeln in Bildern erfasst, indem mehrere Sätze von Filtern anhand der Daten gelernt werden. Das Ergebnis aller Filter bildet einen Merkmalsvektor.

Trainingsdaten Dem Vorteil von überwachten Klassifikationsverfahren, dass sie an neue Szenen durch Training angepasst werden können, steht der Nachteil gegenüber, dass hierfür korrekte und repräsentative Trainingsdaten benötigt werden. Trainingsdaten sind prinzipiell separat für jedes zu prozessierende Bildfluggebiet in einer hohen Qualität zu erheben. Die Erfassung von Trainingsdaten ist generell mit hohem Aufwand verbunden und schränkt damit die praktische Anwendbarkeit des Verfahrens in wesentlichem Maße ein. Um den Aufwand zu reduzieren ohne jedoch bei der Qualität der Klassifikationsergebnisse einzubüßen, könnte das Verfahren in verschiedene Richtungen weiterentwickelt werden. Zum einen wäre es vorstellbar einen Ansatz zu entwickeln, mit dem der einmal angelernte Klassifikator auf andere Szenen übertragen werden kann, ohne im gleichen Umfang Trainingsdaten bereitstellen zu müssen (*Transfer-Lernen* [Pan & Yang, 2010]). Alternativ wären Ansätze vorstellbar, die den Klassifikator robust machen gegenüber ungenauen Trainingsdaten (*Label Noise* [Frénay & Verleysen, 2014]). Hierdurch reduzieren sich die Anforderungen an die Qualität der Trainingsdaten, sodass sich der Erhebungsaufwand deutlich reduziert.

Mit der Verwendung von Klassifikatoren, die gegenüber *Label Noise* robust sind, ist es beispielsweise möglich, für die Klassifikation der Landnutzung direkt das Klassenlabel aus dem möglicherweise veralteten räumlichen Datenbestand für das Training zu verwenden. Für die Klassifikation der Bodenbedeckung kommen beispielsweise Trainingsdaten aus früheren Prozessierungen in Betracht, die speziell für das Gebiet erfasst wurden, jedoch zwischenzeitlich partiell veraltet sind, z.B. [Maas et al., 2016]. In bestimmten Fällen können Trainingsbeispiele für die Bodenbedeckung auch automatisch anhand von Landnutzungsobjekten abgeleitet oder evaluiert werden. Dies ist insbesondere dann möglich, wenn die Nutzungsarten durch eine homogene Bodenbedeckung geprägt sind, z.B. Ableitung der Bodenbedeckungsart *Gras* für Landnutzungsobjekte der Klasse *Grünland*.

Ein weiterer Aspekt, der zu ungenauen Trainingsdaten führt, ist die Zuweisung eines Referenzlabels pro Superpixel nach dem „Winner-takes-all"-Prinzip. In dem präsentierten Ansatz werden unsichere Superpixel nicht zum Training verwendet, obwohl sie Informationen enthalten die damit verloren gehen. Hierbei handelt es sich häufig um Superpixel an Klassenübergängen, die für das Lernen der paarweisen räumlichen Interaktionspotentiale von besonderer Relevanz sind.

In den Experimenten hat sich angedeutet, dass einige Probleme daraus resultieren, dass für einige Klassen, speziell der Landnutzung, in den verwendeten Testgebieten nur eine geringe Anzahl an Trainingsdaten zur Verfügung steht. Somit sind nicht alle Klassen einheitlich und adäquat in den Trainingsdaten repräsentiert. Die Anzahl der Trainingsdaten ließe sich beispielsweise in einfacher Weise erhöhen, indem die GIS-Objekte des flächendeckend vorliegenden TN-Datenbestands hinzugezogen werden. Da für diese TN-Objekte jedoch nicht garantiert werden kann, dass die zugehörigen Klassenlabels korrekt sind, gelten diese Trainingsdaten als ungenau. Dieses Problem ließe sich ebenfalls mit einem Klassifikator lösen, der gegenüber Label Noise robust ist.

Hierarchischer Ansatz Im Rahmen der Experimente hat sich gezeigt, dass im Zuge der Verfeinerung der Klassenstruktur vermehrt Fehlzuweisungen zwischen den Gruppen der obersten Hierarchieebene auftreten. Hier könnte ein hierarchischer Ansatz Abhilfe schaffen, der in einem ersten Schritt eine Klassifikation der Landnutzung entsprechend der Klassenstruktur der obersten Hierarchieebene durchführt.

In dem zweiten Schritt wird jedem Landnutzungsobjekt eine Klasse aus den zugehörigen Untergruppen zugewiesen. Durch die Modellierung der hierarchischen Klassenstruktur in einem CRF ließe sich alternativ eine simultane Klassifikation mehrerer Hierarchieebenen realisieren.

Aktualisierung der Geometrie Das vorgestellte Verfahren bildet den ersten Schritt zur Aktualisierung eines gegebenen Landnutzungsdatenbestands. Bisher erfolgt die Verifikation für die GIS-Objekte des Datenbestands, d.h. die Begrenzungen der Landnutzungsobjekte werden im Rahmen der Klassifikation als korrekt angenommen und nicht verändert. Vor dem Hintergrund, dass sich Nutzungsänderungen auch auf den Verlauf der Nutzungsgrenzen auswirken können, erscheint eine Überprüfung und ggf. Änderung der geometrischen Begrenzungen jedoch sinnvoll. Zur Überprüfung der Begrenzungen können die Klassifikationsergebnisse der Bodenbedeckung herangezogen werden, die in vielen Fällen einen Hinweis auf einen veränderten Grenzverlauf geben. Für den Fall, dass sich Nutzungsgrenzen auch in einer Änderung der Bodenbedeckung ausprägen, lässt sich überprüfen, ob sie im Rahmen der Genauigkeitsanforderungen von ALKIS® mit den Ergebnissen der Bodenbedeckungsklassifikation konsistent sind. Bei signifikanten Abweichungen gibt die Verteilung der Bodenbedeckung einen Hinweis auf den neuen Grenzverlauf. Ziel ist es, den neuen Grenzverlauf automatisch zu prädizieren, z.B. durch die iterative Verformung und Validierung der geometrischen Begrenzungen oder durch das abwechselnde Unterteilen und Zusammenfassen der Landnutzungsobjekte (*Split-and-Merge*-Ansatz).

Literaturverzeichnis

Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S (2012) SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34 (11): 2274–2282.

Albert L, Rottensteiner F, Heipke C (2014a) Land use classification using Conditional Random Fields for the verification of geospatial databases. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. II-4: 1–7.

Albert L, Rottensteiner F, Heipke C (2014b) A two-layer Conditional Random Field model for simultaneous classification of land cover and land use. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XL-3: 17–24.

Albert L, Rottensteiner F, Heipke C (2015) An iterative inference procedure applying Conditional Random Fields for simultaneous classification of land cover and land use. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. II-3/W5: 369–376.

Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (2008) ALKIS-Objektartenkatalog 6.0. http://www.adv-online.de/AAA-Modell/Dokumente-der-GeoInfoDok/ (Letzter Zugriff: 16.07.2014).

Arnold S, Kurstedt R, Riecken J, Schlegel B (2017) Paradigmenwechsel in der Landschaftsmodellierung - von der Tatsächlichen Nutzung hin zu Landbedeckung und Landnutzung. *Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, 142. Jahrgang (1/2017): 30–37.

Banzhaf E, Höfer R (2008) Monitoring urban structure types as spatial indicators with CIR aerial photographs for a more effective urban environmental management. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1 (2): 129–138.

Barnsley MJ, Barr SL (1996) Inferring urban land use from satellite sensor images using kernel-based spatial reclassification. *Photogrammetric Engineering and Remote Sensing*, 62 (8): 949–958.

Barr SL, Barnsley MJ (1997) A region-based, graph-theoretic data model for the inference of second-order thematic information from remotely-sensed images. *International Journal of Geographical Information Science*, 11 (6): 555–576.

Bässmann H, Besslich PW (1991) *Bildverarbeitung Ad Oculos*. Springer-Verlag, Berlin, Heidelberg, D.

Besag J (1986) On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society, Series B (Methodological)*, 48 (3): 259–302.

Bogaert J, Rousseau R, Van Hecke P, Impens I (2000) Alternative area-perimeter ratios for measurement of 2D shape compactness of habitats. *Applied Mathematics and Computation*, 111 (1): 71–85.

Boykov Y, Jolly MP (2001) Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 1: 105–112.

Boykov Y, Veksler O, Zabih R (2001) Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23 (11): 1222–1239.

Breiman L (1996) Bagging predictors. *Machine learning*, 24 (2): 123–140.

Breiman L (2001) Random forests. *Machine Learning*, 45 (1): 5–32.

Breiman L, Friedman J, Stone CJ, Olshen RA (1984) *Classification and regression trees*. Chapman and Hall/CRC, London, UK.

Burger W, Burge MJ (2005) *Digitale Bildverarbeitung: Eine Einführung mit Java und ImageJ*. Springer-Verlag, Berlin, Heidelberg, D.

Champion N (2007) 2D building change detection from high resolution aerial images and correlation digital surface models. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XXXVI-3/W49A: 197–202.

Chehata N, Mallet C, Boukir S, Guo L (2011) Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66 (1): 56–66.

Chen C, Liaw A, Breiman L (2004) *Using Random Forest to learn imbalanced data*. University of California, Berkeley, Technical report.

Criminisi A, Shotton J (2013) *Decision forests for computer vision and medical image analysis*. Springer-Verlag, London, UK.

Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 1: 886–893.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)*, B 39 (1): 1–38.

Duda R, Hart P, Stork D (2001) *Pattern Classification*. Wiley, New York, USA, second edition.

Farabet C, Couprie C, Najman L, LeCun Y (2013) Learning hierarchical features for scene labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35 (8): 1915–1929.

Foody GM (2002) Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80 (1): 185–201.

Förstner W (2013) Graphical models in geodesy and photogrammetry. *Photogrammetrie - Fernerkundung - Geoinformation*, 2013 (4): 255–267.

Frénay B, Verleysen M (2014) Classification in the presence of label noise: a survey. *IEEE Transactions on Neural Networks and Learning Systems*, 25 (5): 845–869.

Frey BJ, MacKay DJ (1998) A revolution: Belief propagation in graphs with cycles. In: *Proceedings of the Neural Information Processing Systems Conference*: 479–485.

Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6 (6): 721–741.

Gould S, Rodgers J, Cohen D, Elidan G, Koller D (2008) Multi-class segmentation with relative location prior. *International Journal of Computer Vision*, 80 (3): 300–316.

Hänsch R, Hellwich O (2017) Random forests. In: Heipke C (ed) *Handbuch Geodäsie: Band Photogrammetrie und Fernerkundung* (pp. 603–644). Springer-Verlag, Berlin, D.

Haralick RM, Shanmugam K, Dinstein IH (1973) Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3 (6): 610–621.

Heipke C, Mayer H, Wiedemann C, Jamet O (1997) Evaluation of automatic road extraction. In: *International Archives of Photogrammetry and Remote Sensing*. XXXII-3/4W2: 151–160.

Helmholz P, Rottensteiner F, Heipke C (2014) Semiautomatic verification of cropland and grassland using very high resolution mono-temporal satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 97: 204–218.

Hermosilla T, Ruiz L, Recio J, Cambra-López M (2012) Assessing contextual descriptive features for plot-based classification of urban areas. *Landscape and Urban Planning*, 106 (1): 124–137.

Herold M, Liu X, Clarke KC (2003) Spatial metrics and image texture for mapping urban land use. *Photogrammetric Engineering and Remote Sensing*, 69 (9): 991–1001.

Hoberg T, Rottensteiner F, Feitosa RQ, Heipke C (2015) Conditional Random Fields for multitemporal and multiscale classification of optical satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 53 (2): 659–673.

Hughes GF (1968) On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14 (1): 55–63.

Kohli P, Ladický L, Torr PH (2009) Robust higher order potentials for enforcing label consistency. *International Journal of Computer Vision*, 82 (3): 302–324.

Kolmogorov V (2006) Convergent tree-reweighted message passing for energy minimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28 (10): 1568–1583.

Kolmogorov V, Zabin R (2004) What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26 (2): 147–159.

Kosov S, Rottensteiner F, Heipke C (2013) Sequential Gaussian Mixture Models for two-level Conditional Random Fields. In: *Proceedings of the German Conference on Pattern Recognition (GCPR)*. LNCS 8142: 153–163.

Krummel J, Gardner R, Sugihara G, ONeill V, Coleman P (1987) Landscape patterns in a disturbed environment. *OIKOS*, 48 (3): 321–324.

Kumar S, Hebert M (2006) Discriminative random fields. *International Journal of Computer Vision*, 68 (2): 179–201.

Ladický L, Russell C, Kohli P, Torr PH (2009) Associative hierarchical CRFs for object class image segmentation. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*: 739–746.

Lafferty J, McCallum A, Pereira FC (2001) Conditional Random Fields: Probabilistic models for segmenting and labeling sequence data. In: *Proceedings of the International Conference on Machine Learning*: 282–289.

Lillesand T, Kiefer R (2000) *Remote sensing and image analysis*. Wiley, New York, USA.

Lu WL, Murphy KP, Little JJ, Sheffer A, Fu H (2009) A hybrid Conditional Random Field for estimating the underlying ground surface from airborne lidar data. *IEEE Transactions on Geoscience and Remote Sensing*, 47 (8): 2913–2922.

Luo C, Sohn G (2014) Scene-layout compatible Conditional Random Field for classifying terrestrial laser point clouds. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*. II-3: 79–86.

Maas A, Rottensteiner F, Heipke C (2016) Using label noise robust logistic regression for automated updating of topographic geospatial databases. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. III-7: 133–140.

Mallet C, Bretar F, Soergel U (2008) Analysis of full-waveform lidar data for classification of urban areas. *Photogrammetrie - Fernerkundung - Geoinformation*, 2008 (5): 337–349.

McGarigal K, Marks BJ (1995) *FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure*. Pacific Northwest Research Station, USDA-Forest Service, Portland, USA, Technical report PNW-GTR-351.

Montanges AP, Moser G, Taubenböck H, Wurm M, Tuia D (2015) Classification of urban structural types with multisource data and structured models. In: *Proceedings of the IEEE Joint Urban Remote Sensing Event*: 1–4.

Montoya-Zegarra JA, Wegner JD, Ladický L, Schindler K (2015) Semantic segmentation of aerial images in urban areas with class-specific higher-order cliques. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. II-3/W4: 127–133.

Munoz D, Bagnell JA, Hebert M (2010) Stacked hierarchical labeling. In: *Proceedings of the European Conference on Computer Vision (ECCV)*: 57–70.

Niemeyer J, Rottensteiner F, Soergel U (2012) Conditional Random Fields for lidar point cloud classification in complex urban areas. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. I-3: 263–268.

Niemeyer J, Rottensteiner F, Soergel U (2014) Contextual classification of lidar data and building object detection in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87: 152–165.

Novack T, Stilla U (2015) Discrimination of urban settlement types based on space-borne SAR datasets and a Conditional Random Fields model. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. II-3/W4: 143–148.

Nowozin S, Rother C, Bagon S, Sharp T, Yao B, Kohli P (2011) Decision tree fields. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*: 1668–1675.

OpenCV (2017) Dokumentation Open Source Computer Vision Library Version 3.0.0. http://docs.opencv.org/3.0.0/ (Letzter Zugriff: 07.01.2017).

Pan S, Yang Q (2010) A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22 (10): 1345–1359.

Park K, Gould S (2012) On Learning Higher-Order Consistency Potentials for Multi-class Pixel Labeling. In: *Proceedings of the European Conference on Computer Vision (ECCV)*: 202–215.

Potetz B, Lee TS (2008) Efficient belief propagation for higher-order cliques using linear constraint nodes. *Computer Vision and Image Understanding*, 112 (1): 39–54.

Roig G, Boix X, Shitrit BH, Fua P (2011) Conditional Random Fields for multi-camera object detection. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*: 563–570.

Rottensteiner F (2017) Kontextbasierte Ansätze in der Bildanalyse. In: Heipke C (ed) *Handbuch Geodäsie: Band Photogrammetrie und Fernerkundung* (pp. 555–602). Springer-Verlag, Berlin, D.

Rottensteiner F, Trinder J, Clode S, Kubik K (2007) Building detection by fusion of airborne laserscanner data and multi-spectral images: Performance evaluation and sensitivity analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62 (2): 135–149.

Rutzinger M, Rottensteiner F, Pfeifer N (2009) A comparison of evaluation techniques for building extraction from airborne laser scanning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2 (1): 11–20.

Schindler K (2012) An overview and comparison of smooth labeling methods for land-cover classification. *IEEE Transactions on Geoscience and Remote Sensing*, 50 (11): 4534–4545.

Shotton J, Winn J, Rother C, Criminisi A (2009) TextonBoost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *International Journal of Computer Vision*, 81 (1): 2–23.

Sithole G, Vosselman G (2004) Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59 (1): 85–101.

Szeliski R, Zabih R, Scharstein D, Veksler O, Kolmogorov V, Agarwala A, Tappen M, Rother C (2008) A comparative study of energy minimization methods for Markov Random Fields with smoothness-based priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30 (6): 1068–1080.

Taubenböck H, Klotz M, Wurm M, Schmieder J, Wagner B, Wooster M, Esch T, Dech S (2013) Delineation of central business districts in mega city regions using remotely sensed data. *Remote Sensing of Environment*, 136: 386–401.

Tönnies KD (2005) *Grundlagen der Bildverarbeitung*, volume 11. Pearson Studium, München, D.

Van de Voorde T, Canters F, van der Kwast J, Engelen G, Binard M, Cornet Y (2009) Quantifying intra-urban morphology of the Greater Dublin area with spatial metrics derived from medium resolution remote sensing data. In: *Proceedings of the IEEE Joint Urban Remote Sensing Event*: 1–7.

Vishwanathan S, Schraudolph NN, Schmidt MW, Murphy KP (2006) Accelerated training of Conditional Random Fields with stochastic gradient methods. In: *Proceedings of the International Conference on Machine learning*: 969–976.

Walde I, Hese S, Berger C, Schmullius C (2014) From land cover-graphs to urban structure types. *International Journal of Geographical Information Science*, 28 (3): 584–609.

Wegner JD, Montoya-Zegarra JA, Schindler K (2013) A higher-order CRF model for road network extraction. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*: 1698–1705.

Weidner U, Förstner W (1995) Towards automatic building reconstruction from high resolution digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50 (4): 38–49.

Wolf L, Bileschi S (2006) A critical view of context. *International Journal of Computer Vision*, 69 (2): 251–261.

Xiong X, Munoz D, Bagnell JA, Hebert M (2011) 3-D scene analysis via sequenced predictions over points and regions. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*: 2609–2616.

Yang MY, Förstner W (2011) A hierarchical Conditional Random Field model for labeling and classifying images of man-made scenes. In: *Proceedings of the IEEE International Conference on Computer Vision/ISPRS Workshop on Computer Vision for Remote Sensing of the Environment*: 196–203.

Yoshida H, Omae M (2005) An approach for analysis of urban morphology: Methods to derive morphological properties of city blocks by using an urban landscape model and their interpretations. *Computers, Environment and Urban Systems*, 29 (2): 223–247.

Zhong P, Wang R (2007) A multiple Conditional Random Fields ensemble model for urban area detection in remote sensing optical images. *IEEE Transactions on Geoscience and Remote Sensing*, 45 (12): 3978–3988.

Zhong P, Wang R (2011) Modeling and classifying hyperspectral imagery by CRFs with sparse higher order potentials. *IEEE Transactions on Geoscience and Remote Sensing*, 49 (2): 688–705.

A. Anhang

Abbildung A.1.: Orthophoto des Testgebiets Schleswig. Quelle: ©GeoBasis-DE/LVermGeo SH (www.LVermGeoSH.schleswig-holstein.de).

Abbildung A.2.: Orthophoto des Testgebiets Hameln. Quelle: Auszug aus den Geobasisdaten der Niedersächsischen Vermessungs- und Katasterverwaltung, ©2010 LGLN (www.lgln.de).

Danksagung

Ich möchte mich bei allen Personen und Institutionen bedanken, die zum Erfolg dieser Dissertation beigetragen haben.

Mein besonderer Dank gilt meinem Doktorvater Prof. Dr. techn. Franz Rottensteiner für die intensive fachliche Betreuung und gewissenhafte Anleitung während der gesamten Zeit meiner Promotion.

Darüber hinaus möchte ich Prof. Dr.-Ing. Monika Sester und Prof. Dr.-Ing. Stefan Hinz dafür danken, dass sie das Korreferat für diese Dissertation übernommen haben, sowie Prof. Dr.-Ing. Winrich Voß für die Übernahme des Vorsitzes der Prüfungskommission.

Außerdem möchte ich mich bei Prof. Dr.-Ing. Christian Heipke bedanken, mit dem ich viele fachliche Diskussionen über meine Forschungsarbeit führen konnte, die in wesentlichem Maße zum Erfolg dieser Dissertation beigetragen haben.

Außerdem gilt mein Dank dem Landesamt für Geoinformation und Landesvermessung Niedersachsen (LGLN) und dem Landesamt für Vermessung und Geoinformation Schleswig-Holstein (LVermGeo SH) für die finanzielle Unterstützung des Forschungsprojektes, in dessen Rahmen diese Dissertation entstand, sowie für die Bereitstellung der verwendeten Testdatensätze.

Darüber hinaus danke ich allen Freunden und Kollegen am Institut für Photogrammetrie und GeoInformation (IPI) der Leibniz Universität Hannover für die schöne gemeinsame Zeit und angenehme Arbeitsatmosphäre. Insbesondere möchte ich Alena und Joachim danken für die vielen anregenden Diskussionen, die Unterstützung in allen Belangen und die vielen aufbauenden Worte während unserer gemeinsamen Zeit in einem Büro. Ich werde diese Zeit in besonderer Erinnerung behalten. Alena danke ich außerdem dafür, dass sie die Dissertation Korrektur gelesen hat, aber vor allem für ihre Freundschaft, mit der sie mir in allen Lebenslagen zur Seite steht.

Ganz besonders möchte ich mich bei meinen Eltern Christine und Karlheinz und meiner Schwester Wiebke für ihre uneingeschränkte Liebe und Unterstützung bedanken. Nicht zuletzt gilt mein größter Dank meinem Freund Torben für seine bedingungslose Liebe und sein Verständnis in schwierigen Zeiten.

Lebenslauf

Persönliche Daten

Name	**Albert, Lena**
Geburtsdatum, -ort	19.06.1986 in Rinteln, Deutschland

Berufserfahrung

Zeiten	**Seit Oktober 2016**
Name und Anschrift des Arbeitgebers	Landesamt für Geoinformation und Landesvermessung Niedersachsen, Podbielskistr. 331, 30659 Hannover
Position bzw. Tätigkeit	Wissenschaftliche Assistenz

Zeiten	**November 2012 – Dezember 2016**
Name und Anschrift des Arbeitgebers	Institut für Photogrammetrie und GeoInformation, Leibniz Universität Hannover, Nienburger Str. 1, 30167 Hannover
Position bzw. Tätigkeit	Wissenschaftliche Mitarbeiterin

Ausbildung

Zeiten	**November 2010 – Oktober 2012**
Art der Ausbildung	Laufbahnausbildung für den höheren technischen Verwaltungsdienst - Fachrichtung Vermessungs- und Liegenschaftswesen
Qualifikation	Große Staatsprüfung
Name der Organisation	Landesamt für Geoinformation und Landentwicklung Niedersachsen

Zeiten	**Oktober 2008 – September 2010**
Art der Ausbildung	Masterstudium Geodäsie und Geoinformatik
Qualifikation	Master of Science
Name der Organisation	Leibniz Universität Hannover

Zeiten	**Oktober 2005 – Oktober 2008**
Art der Ausbildung	Bachelorstudium Geodäsie und Geoinformatik
Qualifikation	Bachelor of Science
Name der Organisation	Leibniz Universität Hannover

Zeiten	**1992 – 2005**
Art der Ausbildung	Schulausbildung
Qualifikation	Allgemeine Hochschulreife
Name der Organisation	Grundschule Heßlingen, Orientierungsstufe Hessisch Oldendorf, Gymnasium Ernestinum Rinteln