

Leibniz Universität Hannover  
Institut für Photogrammetrie und Geoinformation

## Bachelorarbeit

# Erkennung von Veränderungen durch überwachte Klassifikation von Luft- und Satellitenbildern auf Basis von Superpixeln

Studiengang Geodäsie und Geoinformatik

**Betreuer:** 1. apl. Prof. Dr. techn. Franz Rottensteiner  
2. M. Sc. Alina Maas

**eingereicht von:** Mirjana Voelsen

**eingereicht am:** 30. September 2016

# Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst habe.

Ich versichere, dass ich keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe, und dass die eingereichte Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist.

Mirjana Voelsen

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>3</b>
<b>2</b>	<b>Methodik</b>	<b>4</b>
2.1	SLIC-Simple linear interative clustering . . . . .	4
2.1.1	Der SLIC-Algorithmus . . . . .	4
2.2	Merkmale . . . . .	6
2.2.1	Spektrale Merkmale . . . . .	6
2.2.2	Texturmerkmale . . . . .	7
2.2.3	Strukturelle Merkmale . . . . .	7
2.2.4	3D-Merkmale . . . . .	8
2.2.5	Zusammenfassung der Merkmale . . . . .	9
2.3	Klassifikationsmethoden . . . . .	9
2.3.1	Random Forest . . . . .	10
2.3.2	Logistische Regression . . . . .	12
2.4	Evaluierungsmethoden . . . . .	15
2.4.1	Konfusionsmatrix . . . . .	15
<b>3</b>	<b>Vorgehen</b>	<b>17</b>
3.1	Vorstellung der Datensätze . . . . .	17
3.1.1	Datensatz 1: Vaihingen . . . . .	17
3.1.2	Datensatz 2: Saudi-Arabien . . . . .	18
3.2	Einzelpixel- und Superpixelreferenz . . . . .	19
3.3	Vorgehen zur Parameterauswahl für den SLIC-Algorithmus . . . . .	19
3.4	Vorgehen zur Parameterauswahl für Random Forest . . . . .	19
3.5	Vorgehen zur Parameterauswahl für die logistische Regression . . . . .	20
3.6	Vergleich von segment- und pixelbasierter Klassifikation . . . . .	21
<b>4</b>	<b>Ergebnisse</b>	<b>21</b>
4.1	SLIC-Algorithmus . . . . .	21
4.1.1	Vaihingen . . . . .	21
4.1.2	Saudi-Arabien . . . . .	27
4.2	Random Forest . . . . .	32
4.2.1	Vaihingen . . . . .	32
4.2.2	Saudi-Arabien . . . . .	35
4.3	Logistische Regression . . . . .	37
4.3.1	Vaihingen . . . . .	37
4.3.2	Saudi-Arabien . . . . .	39
4.4	Vergleich Random Forest - Logistische Regression . . . . .	42

4.5	Einzelpixeluntersuchung . . . . .	43
4.5.1	Random Forest . . . . .	43
4.5.2	Logistische Regression . . . . .	48
<b>5</b>	<b>Fazit und Ausblick</b>	<b>51</b>
<b>6</b>	<b>Anhang</b>	<b>55</b>

## Kurzfassung

Luft- und Satellitenbilder werden immer mehr dazu genutzt, Landbedeckungen oder Änderung in der Landbedeckung ohne eine Ortsbegehung zu detektieren. Hierfür müssen die Bilder klassifiziert werden, was bedeutet das die Flächen auf den Bildern einer Objektklasse zugewiesen werden. Oft wird hierzu die überwachte Klassifikation genutzt. Dabei stehen für jede Objektklasse Trainingsdaten bereit, anhand derer der Klassifikator die Parameter jeder Klasse bestimmen kann. Anschließend können unbekannte Primitive der richtigen Klasse zugeordnet werden.

Die Klassifikation erfolgt entweder pixelbasiert oder segmentbasiert. Die pixelbasierte Klassifikation trainiert anhand vieler Trainingpixel und klassifiziert anschließend jedes einzelne Pixel. Diese Methode benötigt viel Speicher und Zeit. Die segmentbasierte Klassifikation teilt die einzelnen Pixel vorher in zusammenpassende Segmente auf. Wichtig ist dabei, dass Grenzen im Bild auch von den Segmentgrenzen eingehalten werden. Die Prozessierungszeit wird dadurch sehr stark reduziert, jedoch gehen die Informationen der einzelnen Pixel verloren.

In dieser Arbeit wird eine Methode untersucht um diese Segmente zu erzeugen. Genutzt wird der SLIC (Simple linear iterative clustering)-Algorithmus, der Superpixel erzeugt. Superpixel sind zu Segmenten zusammengefasste Pixel die keine Ecken oder Farbsprünge aufweisen. Ziel ist es, herauszufinden wie gut die Genauigkeiten der Klassifikationsergebnisse sind, die mit Superpixeln erzeugt wurden.

Für die Untersuchung sind zwei Datensätze gegeben. Für beide werden jeweils die Parameter des SLIC-Algorithmus' untersucht, sowie die Parameter der zwei verwendeten Klassifikationsmethoden. Nach der Klassifikation erfolgt jeweils ein Vergleich mit einer per Hand erstellten Referenz. Auf dieser Referenz sind die Objektklassen für jedes Pixel zu sehen. Die Referenz kann einmal für jedes Pixel erstellt werden, aber auch für jedes Superpixel. Die Ergebnisse werden zum Schluss mit den Ergebnissen der Einzelpixelklassifikation derselben Daten verglichen.

Das Ergebnis zeigt, dass die Ergebnisse der Superpixel etwas schlechter sind, als die Ergebnisse der Einzelpixelklassifikation. Allerdings sind die Differenzen in der Gesamtgenauigkeit mit 3 - 4% nicht groß. Die Rechenzeit wird von etwa 24 Stunden für die Einzelpixelklassifikation auf 15 bis 20 Minuten für die Superpixelklassifikation reduziert.

Somit können Superpixel gut genutzt werden, um Landbedeckungen oder deren Änderungen berührungslos zu erkennen. Wichtig dafür sind allerdings ein ausreichend großer Umfang an Trainingsdaten, ein gutes Höhenmodell und eine gute Auswahl an Parametern für den SLIC-Algorithmus und die Klassifikatoren. Nur wenn diese Voraussetzungen erfüllt sind, können gute Ergebnisse erwartet werden.

## Abstract

Aerial and satellite images are more and more used to detect land cover or changes in land cover without site visit. Therefore the images have to be classified. This means that the areas in the images have to be assigned to an object class. For this, the supervised classification is used in many cases. In the supervised classification training samples for every object class are given. With these training samples the classifier can calculate the parameters for each object class. Afterwards primitives with unknown object classes can be assigned to the right class. The classification can be done segment based or pixel based. The pixelbased classification trains with every single pixel from the training samples and classifies every single unknown pixel. This takes very long time and a lot of memory. The segment based classification divides the pixel in groups before the classification algorithm starts. For this, it is very important that borders in the image are borders of the segments too. The computation time is reduced through this very significant but the information of the single pixel get lost.

In this study a method to create these segments is examined. For this the SLIC (Simple linear iterative clustering)-algorithm is used. This algorithm creates superpixel. Superpixel are segments of pixel without edges or color jumps. The aim is to work out whether the quality from the superpixel based classification is as good as the pixel based segmentation or not.

For the examination two sets of data are given. For both, the parameters for the SLIC-Algorithm and the two used classification methods are determined. After the classification the results are compared to the hand made reference. In this reference, the object class of every pixel or superpixel is known. The reference is given pixel based and segment based. In the end the results are compared to the results of the pixel based classification.

The outcome shows, that the results with the superpixel classification are worse than the pixel based results. The differences for the overall accuracy lie between three and four percent and are not very large. The computation time is reduced from 24 hours for the pixel based classification to 15 to 20 minutes for the segment based classification.

Consequently superpixel can be used for the detection of land cover or changes in it without site visit. Important is, that there are enough training samples, a good calculated elevation model and a good selection of parameters for the superpixel and the classification algorithms. Only when these conditions are complied, results with high accuracies can be expected.

# 1 Einleitung

Die automatisierte Klassifikation von Luft- und Satellitenbildern wird heutzutage mehr denn je dazu genutzt, die Landbedeckung großräumig und berührungsfrei zu ermitteln. Diese Daten bilden die Grundlage für Planungen, Erkennung von Veränderungen oder können Informationen über den Klimawandel liefern [Bach et al., 2013].

Die Klassifikation dieser Daten findet häufig durch überwachte Klassifikation statt. Dabei stehen Trainingsdaten für jede Objektklasse bereit, anhand derer die Klassifikationsparameter bestimmt werden [Rottensteiner, 2015].

Das Training und die anschließende Klassifikation neuer Daten kann pixel- oder segmentbasiert durchgeführt werden. Bei der pixelbasierten Klassifikation wird jedes Pixel der vorhandenen Daten nacheinander einer bestimmten Objektklasse zugewiesen. Für große Datenmengen wird der benötigte Speicher und die Prozessierungszeit entsprechend groß, wenn alle vorhandenen Pixel genutzt werden. Um diese Durchlaufzeit zu verkürzen, werden in der segmentbasierten Klassifikation die einzelnen Pixel zunächst zu größeren Segmenten zusammengefasst, bevor diese anschließend als Ganzes einer Objektklasse zugewiesen werden. Die Herausforderung ist, die Segmente so zusammenzufassen, dass Grenzen zwischen Objekten im Bild auch den Grenzen der Segmente entsprechen. Nur wenn das der Fall ist, können in der Klassifikation die Segmente der richtigen Objektklasse zugeordnet werden. Die Informationen der einzelnen Pixel werden dabei zusammengefasst.

Ein Verfahren, diese Segmente zu erzeugen, ist die Benutzung von Superpixeln. Ein Superpixel besteht aus einer bestimmten Anzahl von Ausgangspixeln und enthält keine Objektkanten oder Farbsprünge[Achanta et al., 2010].

In dieser Arbeit wird der SLIC (Simple linear iterative clustering)-Algorithmus zur Erstellung der Superpixel genutzt. Dieser Algorithmus wird mithilfe von zwei Klassifikationsmethoden für zwei unterschiedliche Datensätze bewertet. Ein Vergleich mit der Einzelpixelklassifikation derselben Daten soll zeigen, ob die Genauigkeiten der segmentweisen Klassifikation mit der pixelweisen Klassifikation mithalten können, und ob sie für die Erkennung von Veränderungen in Luftbildern genutzt werden können.

Zu Beginn der Arbeit werden die verwendeten Methoden beschrieben. Dieser Abschnitt beginnt mit der Beschreibung des Algorithmus zur Erstellung der Superpixel, geht anschließend auf die verwendeten Merkmale der Pixel und Superpixel ein, und wendet sich dann den Klassifikations- und Evaluierungsmethoden zu.

Im nächsten Kapitel werden die Untersuchungen, die im Rahmen dieser Arbeit durchgeführt wurden, beschrieben. Hier werden zu Beginn die genutzten Datensätze beschrieben und anschließend der Untersuchungsablauf. Die Ergebnisse werden im nächsten Kapitel dokumentiert und analysiert.

Es folgen ein Fazit und ein Ausblick für mögliche weitere Untersuchungen und Vorgehen.

## 2 Methodik

### 2.1 SLIC-Simple linear interactive clustering

#### 2.1.1 Der SLIC-Algorithmus

Der SLIC-Algorithmus generiert Superpixel mit einem einfachen iterativen Algorithmus.

Dem Algorithmus muss ein Eingabeparameter  $k$  übergeben werden, der die gewünschte Anzahl an Superpixeln darstellt. In dieser Arbeit wird jeweils die gewünschte Superpixelgröße angegeben, die sich aus  $k$  und der Pixelanzahl  $N$  des Bildes ergibt, da diese leichter zu vergleichen ist, wenn unterschiedlich große Bilder verwendet werden.

Im ersten Schritt werden  $k$  Cluster-Zentren regelmäßig über das Bild verteilt. Dies geschieht im Abstand  $S = \frac{N}{k}$ . Jedes Cluster erhält einen Suchbereich mit der Größe  $2S \cdot 2S$ , welcher in Abb.1 dargestellt ist. Durch diese Begrenzung des Suchbereichs ist die Anzahl der benötigten Distanzberechnungen stark eingeschränkt und die Komplexität des Algorithmus ist linear [Achanta et al., 2010].

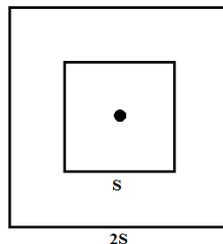


Abbildung 1: Skizze des Suchbereichs um ein Cluster-Zentrum.  $S$  ist der Abstand der Clusterzentren.

#### Die Distanzberechnung

Für jedes Pixel werden nun die Distanzen zu den Cluster-Zentren berechnet, in deren Suchbereichen das Pixel liegt. Das Pixel wird anschließend dem Kluster mit dem geringsten Abstand zugeordnet.

Für die Distanzberechnung wird für jedes Pixel ein 5-dimensionaler Vektor  $\begin{bmatrix} dl & da & db & dx & dy \end{bmatrix}^T$  aufgestellt. Der Vektor  $\begin{bmatrix} dl & da & db \end{bmatrix}^T$  steht dabei für die Pixelfarbe im CIELAB-Farbraum. Dieser dreidimensionale Farbraum unterteilt Farben anhand ihres Farbtons sowie ihrer Helligkeit (Schwarz-Weiß-Skala). Auf diese Weise können Farbunterschiede als Distanz beschrieben werden. Der Vektor  $\begin{bmatrix} dx & dy \end{bmatrix}$  stellt den räumlichen Abstand auf dem Bild dar. Der Abstand zwischen den Farbwerten ist auf die Größe des CIELAB-Farbraums limitiert und der räumliche Abstand im Bild auf die Bildgröße. Deshalb muss der räumliche Abstand normalisiert werden,

damit beide Distanzen den gleichen Einfluss auf das Ergebnis haben.

$$d_{lab} = \sqrt{(dl_k - dl_i)^2 + (da_k - da_i)^2 + (db_k - db_i)^2}$$

$$d_{xy} = \sqrt{(dx_k - dx_i)^2 + (dy_k - dy_i)^2}$$

$$D_s = d_{lab} + \frac{m}{S} d_{xy}$$

Hierbei sind  $d_{lab}$  und  $d_{xy}$  die Distanzen im Farbraum sowie auf dem Bild. Die Summe der beiden ist  $D_s$ , wobei der räumliche Abstand mithilfe des Klusterabstands  $S$  normalisiert wird. Außerdem kann mit der Variable  $m$  der Einfluss des räumlichen Abstands variiert werden. Je größer der Wert für  $m$  dabei gewählt wird, desto kompakter werden die Superpixel und desto weniger passen sie sich an Farbsprünge im Bild an. In Abb. 2 werden für die Kompaktheit die Werte 10, 40 und 80 eingesetzt. Man erkennt deutlich, dass bei einem kleinen Wert für die Kompaktheit die Grenzen im Bild noch sehr gut erkennbar sind und bei einem hohen Wert kaum noch.

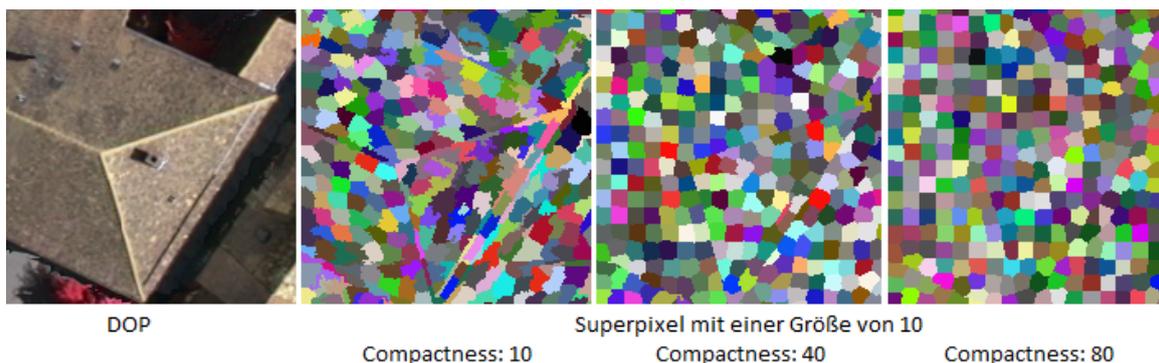


Abbildung 2: Superpixel, dargestellt mit Farben mit unterschiedlichen Werten für die Kompaktheit und einer Superpixelgröße von 20x20 (Bearbeitet mit ArcMap)

Im nächsten Schritt werden die Cluster-Zentren aller Superpixel neu berechnet. Hierzu wird der Mittelwert für jeden Eintrag des Distanzvektors aus allen Pixeln des Clusters gebildet. Nun werden wieder die Distanzen zu allen Cluster-Zentren, in deren Suchbereichen das jeweilige Pixel liegt, berechnet und der Algorithmus wird erneut durchlaufen. Dies geschieht iterativ bis sich die Cluster-Zentren kaum noch bewegen, was nach maximal 10 Iterationen der Fall ist [Achanta et al., 2010].

Insgesamt benötigt der SLIC-Algorithmus folgende Eingabeparameter:

- Kanäle 1-3 : Hier können drei Merkmale eingetragen werden, die später auf den CIELAB-Farbraum skaliert werden und so einen Einfluss auf die Grenzen der Superpixel haben. Ist ein Merkmal z.B. die Höhe so werden später an Häuserkanten etc. Superpixel-Grenzen

sein. Hier gilt es herauszufinden welche Merkmale sich am besten für die Erzeugung der Superpixel eignen.

- Minimale und maximale Größe der Superpixel: Minimale und Maximale Größe, die die Superpixel haben sollen. Hier wird für beide Werte immer derselbe Wert eingegeben damit alle Superpixel in etwa die gleiche Größe haben. Pixelgröße 10 bedeutet dabei, dass die Pixel (wären sie genau rechteckig) 10x10=100 Pixel groß sind. Die tatsächlichen Pixelgrößen können natürlich etwas von diesem eingestellten Wert abweichen, wenn z.B. Grenzen im Bild eingehalten werden müssen.
- Kompaktheit m der Superpixel: Dieser Parameter gibt an wie kompakt die einzelnen Pixel sein sollen. Je größer der Wert für m, desto mehr Einfluss hat der räumliche Abstand und desto weniger passen sich die Superpixel an Farbsprünge im Bild an. Ein Wert von 10 ist dabei sehr wenig kompakt, ein Wert von 80 (höchster Wert) ist sehr kompakt.

## 2.2 Merkmale

Das Ziel der Klassifikation ist es, jedes Primitiv (Pixel oder Segment) einer bestimmten Klasse zuzuordnen. Klassen sind in diesem Fall verschiedene Arten der Landbedeckung wie Gebäude, Straße oder Vegetation. Die Eigenschaften des Bildprimitivs  $i$  werden durch einen  $k$ -dimensionalen Merkmalsvektor  $x_i = \begin{bmatrix} f_{i_1} & f_{i_2} & \dots & f_{i_k} \end{bmatrix}^T$  beschrieben. Dabei kann es sich um direkt beobachtete Merkmale handeln oder um Funktionen aus einigen oder allen von ihnen [Rottensteiner, 2015]. Die Merkmale können unterschieden werden nach spektralen Merkmalen, Texturmerkmalen und 3D-Merkmalen.

### 2.2.1 Spektrale Merkmale

#### NDVI (Normalized Difference Vegetation Index)

Dieser normalisierte Vegetationsindex hat einen Wertebereich von -1 bis +1. Vegetation liefert für diesen Index Werte nahe 1. Durch die Normierung haben Geländeneigung und Beleuchtung keinen Einfluss auf den Index. Er berechnet sich mithilfe der Differenz zwischen dem nahen infraroten Wellenlängenbereich (ca. 760nm) und dem Wellenlängenbereich des sichtbaren roten Lichts (ca 660nm) [Albertz and Wiggenhagen, 2009].

$$NDVI = \frac{NIR - Rot}{NIR + Rot}$$

Der NDVI kann für jedes einzelne Pixel berechnet werden. Für Superpixel wird der Index ebenfalls für jedes Pixel berechnet und es können der Mittelwert, die Standardabweichung, der kleinste sowie der größte Wert des NDVI für das Superpixel berechnet werden.

## Der IHS-Farbraum

Der IHS (Intensität - Farbton - Sättigung) Farbraum unterscheidet zwischen der Helligkeit und der Farbe. Die Helligkeit wird durch die Intensität beschrieben, wobei schwarz eine Intensität von 0%, und weiß eine Intensität von 100% besitzt. Die Farbe wird durch den Farbton (grün, blau, rot etc.) beschrieben und durch die Sättigung dieses Farbtons. Hat ein Farbton keine (0%) Sättigung so ist er grau (oder im Extremfall schwarz oder weiß). Bei 100% Sättigung ist die Farbe sehr kräftig und erscheint für das menschliche Auge strahlend [Albertz and Wiggenhagen, 2009]. Intensität und Sättigung können für jedes Pixel angegeben werden. Für die Superpixel werden aus den Werten jedes einzelnen Pixels der Mittelwert, die Standardabweichung, der kleinste und der größte Wert für die Intensität und Sättigung für jedes Superpixel berechnet.

### 2.2.2 Texturmerkmale

#### Co-Occurance-Matrix

Die Co-Occurance-Matrix beschreibt die räumlichen Abhängigkeiten von Grauwerten in einer bestimmten Richtung und Distanz. In der Matrix werden in jedem Element die Häufigkeiten eingetragen, mit der zwei Grauwerte in einer gewissen räumlichen Anordnung auftreten. Der räumliche Abstand für diese Berechnung beträgt eins und die Co-Occurance-Matrix wird für vier unabhängige Richtungen aufgestellt (so ergeben sich auch vier Matrizen) und anschließend gemittelt. Mithilfe dieser Matrix lassen sich verschiedene Merkmale berechnen die auf unterschiedliche Weise die Texturen untersuchen.

Die Co-Occurance Matrix wird für Einzelpixel für einen Bildausschnitt berechnet, in dessen Mitte das jeweils aktuelle Pixel liegt. Die Größe dieser Matrix beträgt 5x5 Pixel. Bei der segmentierten Klassifikation wird die Co-Occurance Matrix für jedes Superpixel einmal berechnet. Eins dieser Textur-Merkmale ist die Energieverteilung  $Ev$ . Dafür wird die Summe über alle quadrierten Einträge der Co-Occurance Matrix berechnet. Vorher werden die Anzahlen in der Matrix mit der Gesamtanzahl normiert. Der Wertebereich der Energieverteilung liegt zwischen 0 und 1. Ein Wert von 1 würde herauskommen, wenn nur ein einziger Grauwert im Bild vorkommt und so nur ein Eintrag der Co-Occurance-Matrix besetzt ist [Haralick, 1979].

$$Ev = \sum_{ij} P_{ij}^2$$

$P_{ij}$  steht dabei für den jeweiligen Eintrag der Co-Occurance-Matrix mit  $i$  Spalten und  $j$  Zeilen.

### 2.2.3 Strukturelle Merkmale

#### Gradientenhistogramm

Im Gradientenhistogramm werden die Größen (Normen) von Gradientenrichtungen aufsummiert. Der Grauwertgradient  $\nabla g(x, y)$  beschreibt, wie stark sich der Grauwert im Bild in  $x$ -

und y-Richtung verändert.

$$\nabla g(x, y) = \begin{pmatrix} g_x \\ g_y \end{pmatrix} = \begin{pmatrix} \frac{\delta g}{\delta x} \\ \frac{\delta g}{\delta y} \end{pmatrix}$$

$$Norm : \|\nabla g(x, y)\| = \sqrt{g_x^2 + g_y^2}$$

$$Richtungswinkel : \theta = \text{atan}\left(\frac{g_y}{g_x}\right)$$

Für den Grauwertgradienten  $\nabla g(x, y)$  werden in x und y Richtung die Grauwertänderungen nach x ( $g_x$ ) und nach y ( $g_y$ ) abgeleitet. Die Norm und die Richtung des Grauwertgradienten ergeben sich anschließend aus den Einträgen ( $g_x$  und  $g_y$ ) des Gradientenvektors.

Für das Gradientenhistogramm werden die Richtungen in  $6^\circ$  Intervalle eingeteilt. Fällt die Gradientenrichtung eines Bildprimitivs in eins dieser Intervalle wird die Norm des Primitivs auf die Gesamtsumme des Intervalls addiert. Für Einzelpixel geschieht dies für einen Bereich von  $2 \times 2$  Pixel um das mittige Pixel herum. Für Superpixel wird ein Histogramm für das gesamte Segment aufgestellt. Aus dem Histogramm kann nun der Maximalwert, der Minimalwert oder auch der Mittelwert der Amplituden berechnet werden [Rodehorst, 1997].

#### 2.2.4 3D-Merkmale

##### NDOM (Normalisiertes digitales Oberflächenmodell)

Ein NDOM ist ein normiertes digitales Oberflächenmodell. Es berechnet sich aus der Differenz zwischen DOM (digitales Oberflächenmodell) und dem DGM (digitales Geländemodell). Im digitalen Oberflächenmodell sind dabei, im Gegensatz zum Geländemodell, alle Objekte auf der Erdoberfläche mit enthalten. Durch die Differenzbildung haben (wie in Abb.3 zu sehen) Geländehöhen und -neigungen keinen Einfluss auf das Modell. Auf diese Weise können flache und hohe Objekte direkt auf Grundlage des NDOM unterschieden werden [Haala and Brenner, 1999].

Da für den gegebenen Datensatz kein Geländemodell vorliegt, wird dieses zunächst aus dem Oberflächenmodell extrahiert. Dabei ist es vor allem wichtig Häuser und Bäume herauszufiltern. Dies geschieht durch den morphologischen Filter „Opening“. Dabei werden Umrisse glatter und Objekte die kleiner sind als das genutzte Strukturelement entfernt. Die Größe der Strukturmatrix beträgt hier  $30 \times 30$  Meter, womit alle Bäume und Gebäude entfernt werden sollten. Erst nach dieser Filterung kann das NDOM berechnet werden [Weidner, 1997].

Das NDOM liefert so für jedes Pixel eine Höhe. Für Superpixel werden die Höhen aller Pixel im jeweiligen Segment mit einbezogen. So können wieder der Mittelwert, die Standardabweichung sowie der minimale und maximale Wert angegeben werden.

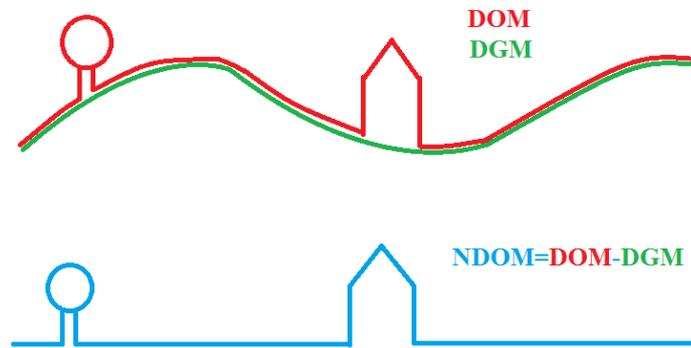


Abbildung 3: Graphische Darstellung der Berechnung des NDOM

### 2.2.5 Zusammenfassung der Merkmale

Da für die Untersuchungen mit Abkürzungen für die genutzten Merkmale gearbeitet wird, sind in Tabelle 1 einmal alle Merkmale mit einer kurzen Erklärung und ihrer jeweiligen Abkürzung aufgelistet.

Abkürzung	Erklärung
Farbkanal 1-3	Der jeweilige Wert für den Farbkanal des Primitivs.
NDVI	Normalisierter Vegetationsindex. Dieser Index hat für Primitive mit viel Vegetation einen Wert nahe eins.
Intensität	Beschreibt die Helligkeit der Farbe eines Primitivs
Sättigung	Beschreibt wie kräftig der Farbton eines Primitivs ist.
Energieverteilung	Energieverteilung, wird berechnet mithilfe der Co-Occurance-Matrix und ist ein Texturmerkmal
Gradient	Beschreibt das Gradientenhistogramm. Es können Mittelwert, Maximum (max) und Minimum (min) angegeben werden
NDOM	Höhe von Objekten auf der Oberfläche gegenüber dem Gelände

Tabelle 1: Zusammenfassung der Merkmale

## 2.3 Klassifikationsmethoden

Die segmentbasierte Klassifikation soll anhand von zwei Klassifikatoren mit der Einzelpixelklassifikation verglichen werden. Hierzu werden die Klassifikationsmethoden Logistische Regression und Random Forest genutzt. Beide arbeiten mit der überwachten Klassifizierung. Das bedeutet, dass Trainingsdaten für alle Objektklassen zur Verfügung stehen, anhand derer die Algorithmen die Merkmalsvektoren der einzelnen Klassen bestimmt werden können. Trainingsdaten

sind Primitive, deren Klassen bekannt sind und durch deren Merkmalsvektoren die Klassifikatoren angelernt werden können. Dies geschieht im Training. In der anschließenden Klassifikation werden unbekannte Pixel oder Segmente ihrer wahrscheinlichsten Klasse zugeordnet [Rottensteiner, 2015].

### 2.3.1 Random Forest

#### Entscheidungsbäume

Die Klassifikation mit Random Forest basiert auf einer Zuordnung der Pixel in verschiedene Klassen durch eine Kombination von Entscheidungsbäumen. Ein Entscheidungsbaum ist ein binärer Baum, der an jedem Knoten durch eine festgelegte Merkmalsgrenze die Primitive in zwei Unterklassen aufteilt. Dabei kann entweder auf ein oder mehrere Merkmale eingegangen werden. Am nächsten Knoten geschieht das gleiche Vorgehen für das nächste Merkmal.

In Abb.4 ist hierfür ein Beispiel dargestellt. Die erste Abfrage soll die Pixel aufteilen in die Klassen Vegetation und keine Vegetation. Für diese Abfrage müsste für den NDVI-Wert eine passende Grenze gefunden werden, ab dem es sich bei einem Pixel mit großer Wahrscheinlichkeit um Vegetation handelt. Anschließend sollen die beiden Unterklassen in jeweils zwei weitere Klassen eingeteilt werden. Auf der einen Seite soll die Vegetation unterteilt werden in Wald oder keinen Wald. Hier kann z.B. eine Grenze für die Höhe eingeführt werden und ein Pixel wird der Klasse Wald zugeordnet, wenn seine Höhe größer ist als dieser Wert [Albertz and Wiggenhagen, 2009].

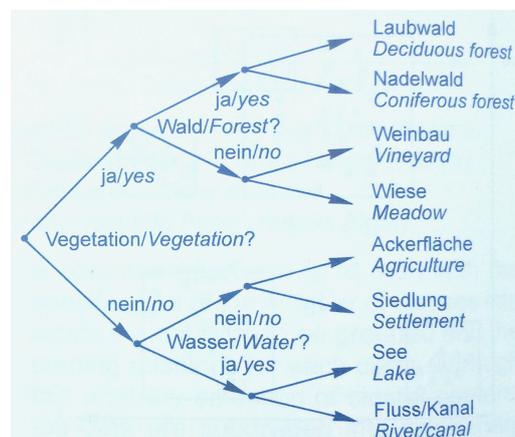


Abbildung 4: Beispiel für einen Entscheidungsbaum mit drei Entscheidungsstufen [Albertz and Wiggenhagen, 2009]

Auf diese Weise kann mit beliebig vielen Grenzen der Merkmalsraum immer weiter unterteilt werden. Der Nutzer kann so durch die Wahl der Schwellwerte sehr leicht eingreifen und die Klassifikation beeinflussen. Dies ist jedoch gleichzeitig ein großer Nachteil, denn so ist es sehr schwer eine optimale Grenze zu finden. [Albertz and Wiggenhagen, 2009].

## CART - Classification and Regression Trees

Das CART-Verfahren wird genutzt, um einen Entscheidungsbaum zu trainieren und Grenzen zwischen den Klassen optimal fest zu legen. Außerdem soll festgelegt werden, wann ein Blatt erreicht ist und welches Blatt welcher Klasse angehört.

Die Grenzen zwischen den Klassen sind lineare Entscheidungsgrenzen, wodurch mehrere Merkmale pro Knoten berücksichtigt werden können.

In jedem Knoten werden nun Tests durchgeführt. Hierfür werden ca.  $\frac{1}{3}$  der Trainingsdaten genutzt. In dem genutzten Algorithmus werden an jedem Knoten alle ausgewählten Merkmale  $n$  mal zufällig gewichtet. Wird ein Merkmal mit Null gewichtet wird es nicht mit genutzt. Es werden nun  $x$  verschiedene Entscheidungsgrenzen für jede Gewichtung zufällig generiert und mithilfe der ausgewählten Trainingsdaten wird geprüft, welche Entscheidungsgrenze für welche Gewichtung die Daten am besten trennt und für den Knoten beibehalten wird. Dies kann durch den Informationszuwachs  $\Delta E$  berechnet werden [Heipke, 2016].

$$\Delta E = \frac{N_1}{N_1 + N_2} \cdot \sum_k P_1(C^k) \cdot \log_2[P_1(C^k)] + \frac{N_2}{N_1 + N_2} \cdot \sum_k P_2(C^k) \cdot \log_2[P_2(C^k)]$$

Mit:

- Entscheidungsgrenze als Ebene der Form:  $w^T \cdot x + w_0 = 0$
- Zwei Teilmengen  $M_1, M_2$  die durch die Entscheidungsgrenze entstehen:
  - $M_1 = w^T \cdot x + w_0 \leq 0$
  - $M_2 = w^T \cdot x + w_0 > 0$
- Histogramm über die Teilmengen  $M_1, M_2$  mit den Histogrammeinträgen  $P_1(C^k), P_2(C^k)$
- $N_1, N_2$  sind die Anzahl der Trainingsbeispiele die durch die getestete Entscheidungsgrenze in  $M_1$  oder  $M_2$  fallen

Je größer der Wert für  $\Delta E$  ist, desto besser trennt die Entscheidungsgrenze mit der dazugehörigen Gewichtung der Merkmale die beiden Unterklassen  $M_1, M_2$ . Somit wird die Grenze und die Merkmalsgewichtung ausgewählt, die den größten Wert für den Informationszuwachs hat.

Am nächsten Knoten wiederholt sich nun das Vorgehen für alle Trainingsdaten, die in diesen Knoten eingeteilt wurden. Es gibt mehrere Abbruchkriterien, die festlegen, wann keine erneute Teilung stattfindet, sondern ein Blatt erreicht ist:

- Wenn eine maximale Tiefe des Entscheidungsbaums erreicht ist wird das Verfahren beendet. Sonst droht der Algorithmus zur Überanpassung und Klassen werden immer weiter in Unterklassen aufgeteilt.

- Wenn der Informationszuwachs  $\Delta E$  sehr klein ist wird keine weitere Aufteilung benötigt und der Knoten wird zu einem Blatt.
- Ist die Anzahl der Trainingsdaten in einem Blatt zu klein, um eine weitere gute Entscheidungsgrenze zu finden um diese aufzuteilen, entsteht ebenfalls ein Blatt
- Sind die Anzahl und Arten den Klassen bekannt, so kann das Verfahren beendet werden, wenn in einem Blatt nur noch Pixel einer Klasse zu finden sind.

Um anschließend jedes Blatt einer Klasse zuzuordnen, werden nun die restlichen Trainingsdaten anhand des Entscheidungsbaums in ein Blatt eingeordnet. In jedem Blatt liegen jetzt eine bestimmte Anzahl an Daten mit bekannter Klasse. Die Blätter werden nun derjenigen Klasse zugeordnet, die am Häufigsten in sie eingeordnet wurde.

Folgt nun ein zu klassifizierendes Primitiv, wird dies ebenfalls anhand der Entscheidungsgrenzen in ein Blatt eingeordnet und kann der Klasse zugeordnet werden, die diesem Blatt entspricht.

### **Bootstrapping**

Für Random Forest müssen nun noch Bootstrap-Datensätze angelegt werden. Dafür werden aus allen Trainingsdaten eine vorher festgelegte Anzahl  $n$  zufällig ausgewählt. Die Auswahl erfolgt mit Zurücklegen, wodurch ein Pixel auch mehrmals in einem Bootstrap-Datensatz vorkommen kann. Auf diese Weise werden  $B$  Bootstrap Datensätze, die unabhängig voneinander sind, angelegt. Mit jedem dieser Datensätze wird nun ein Klassifikator trainiert. In der anschließenden Klassifikation wird nun für jedes Pixel für jeden Klassifikator das Ergebnis bestimmt und letztendlich die Klasse ausgewählt, die am Häufigsten ausgewählt wurde [Breiman, 2001].

### **Random Forest**

Random Forest besteht aus einer Zusammensetzung von Bootstrapping und CART. Es werden  $T$  Entscheidungsbäume nach dem CART-Verfahren angelegt und jeder wird auf einem anderen Bootstrap-Datensatz trainiert. In der Klassifikation wird jedes Primitiv anhand aller Bäume einer Klasse zugeordnet. Die Klasse die dabei am Häufigsten gewählt wurde, wird als Klasse des Primitivs ausgewählt. Auf diese Weise erhält man eine hohe Stabilität und eine bessere Generalisierung als wenn z.B. nur auf einem Baum trainiert wird [Breiman, 2001].

### **2.3.2 Logistische Regression**

Bei der Logistischen Regression wird nach einer optimalen Trennfläche im Merkmalsraum gesucht um die Klassen bestmöglich voneinander zu trennen. Auf dieser Grundlagen können später, abhängig vom Merkmalsvektor, Wahrscheinlichkeiten errechnet werden, mit denen ein Pixel der richtige Klasse zugeordnet wird. Basis ist eine logistische Sigmoidfunktion, welche auf dem Theorem von Bayes aufbaut [Bishop, 2006].

Zunächst werden die a-posteriori Wahrscheinlichkeiten  $p(C^k|x)$  für jede Klasse  $C^k$  mithilfe der Softmax-Funktion aufgestellt.

$$p(C^k|x_n) = \frac{e^{a_k(x_n)}}{\sum_j e^{a_j(x_n)}}$$

$$a_k(x_n) = \ln[p(x_n|C^k)] + \ln[p(C^k)]$$

Mit:

- $x_n$  - Merkmalsvektor des Punktes n
- k - Klasse deren a-posteriori Wahrscheinlichkeit bestimmt werden soll
- j - Die Summe läuft einmal für jede Klasse j durch
- $p(x_n|C^k)$  - Wahrscheinlichkeit den Merkmalsvektor x zu Beobachten, wenn x zur Klasse  $C^k$  gehört
- $p(C^k)$  - Wahrscheinlichkeit das die Klasse k auftritt

### Feature Space Mapping

Lassen sich die einzelnen Klassen nicht mehr durch eine lineare Funktion voneinander trennen, muss auf eine komplexe Funktion für  $a(x)$  zurückgegriffen werden. Anstatt aber ein komplexes Modell für  $a_k$  einzuführen, wird beim Feature Space Mapping der Merkmalsraum in einen Merkmalsraum mit höherer Dimension transformiert, in dem  $a(\Phi(x))$  linear bleibt:

$$\Phi(x)_k = [\Phi_1(x), \Phi_2(x), \dots, \Phi_N(x)]$$

$\Phi(x)_k$  ist dabei eine Funktion des Merkmalsvektors x und N die Dimension des transformierten Merkmalsvektors der Klasse k, der meist größer ist als die Dimension von x. Auf diese Weise ist  $a(\Phi(x))$  weiterhin linear.

Ein Beispiel für eine Transformation ist die quadratische Erweiterung, die hier für den Merkmalsvektor  $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$  dargestellt wird:

$$\Phi(x) = [1, x_1, x_2, x_1 \cdot x_2, x_1^2, x_2^2]^T$$

Mit dieser Erweiterung ergeben sich die folgenden Formeln.

$$a_k(x_n) = a_k(\Phi(x_n)) = w_k^T \cdot \Phi(x_n)$$

$$p(C^k|x_n) = \frac{e^{w_k^T \cdot \Phi(x_n)}}{\sum_j e^{w_j^T \cdot \Phi(x_n)}} = y_{nk}$$

Der Parametervektor  $w_k$  muss nun für jede Klasse  $k$  im Training bestimmt werden.

### Training

Ziel des Trainings ist es, den Parametervektor  $w_k$  für jede Klasse zu bestimmen, sodass bei der späteren Klassifikation die Pixel eindeutig einer Klasse zugeordnet werden können.

Als Trainingsdaten sind  $x_n$  Merkmalsvektoren mit ihrer bekannten Klassenzugehörigkeit  $t_{nk}$  gegeben. Die Klassenzugehörigkeit kann dabei die Werte null und eins annehmen. Gehört  $x_n$  zu der Klasse  $k$ , für die gerade der Parametervektor bestimmt wird, so ist der Wert eins, gehört  $x_n$  nicht dazu, so ist der Wert null.

Der Parametervektor  $w_k$  soll nun so bestimmt werden, dass die Wahrscheinlichkeit, dass der Merkmalsvektor  $x_n$  entweder zu ( $t_{nk} = 1$ ) oder nicht zu ( $t_{nk} = 0$ ) der Klasse  $k$  gehört, maximal wird. Dies geschieht im Maximum Likelihood-Training.

$$p(t_{nk}|w_k, x_1, x_2, \dots, x_N) \rightarrow maximal$$

$$p(t_{nk}|w_k, x_1, x_2, \dots, x_N) = \prod_{k=1}^K \prod_{n=1}^N y_{nk}^{t_{nk}}$$

$K$  ist dabei die Anzahl der Klassen und  $N$  die Anzahl der vorhandenen Merkmalsvektoren  $n$  zum Trainieren.  $t_{nk}$  ist die Klassenzugehörigkeit, wie oben beschrieben.

Anstelle der Berechnung des Maximums dieser Funktion, wird das Minimum des negativen Logarithmus gesucht.

$$E(w_k) = -\ln(p(t|w_k, x_1, x_2, \dots, x_N)) \rightarrow minimal$$

$$E(w_k) = -\sum_{n=1}^N \sum_{k=1}^K (t_{nk} \cdot \ln(y_{nk}))$$

Die Ermittlung des Minimums erfolgt an dieser Stelle iterativ, da  $E(w)$  eine nichtlineare Funktion von  $w$  ist. Bestimmt wird das Minimum mithilfe des Gradienten  $\nabla E(w) = 0$  mithilfe des Newton-Raphson-Verfahrens, ausgehend von der Näherung  $w^{\tau-1}$ .

$$w_k^\tau = w_k^{\tau-1} - H^{-1} \cdot \nabla E \cdot w_k^{\tau-1}$$

Mit:

$$\nabla E(w_k) = \sum_{n=1}^N y_{nk} - t_{nk} \cdot \Phi(x_n)$$

$$H = \nabla \nabla E(w_k) = \sum_{n=1}^N y_{nk} \cdot (1 - y_{nk}) \cdot \Phi(x_n) \cdot \Phi(x_n)$$

$w_k^T$  ist dabei immer der in der Iteration neu berechnete Parametervektor der Klasse  $k$  und  $w_k^{T-1}$  der Parameter der letzten Iteration. Zu Beginn wird dieser mit Nullen besetzt.

Das beschriebene Maximum Likelihood-Training neigt zur Überanpassung, wenn beispielsweise die Klassenzentren so weit auseinander liegen, dass eine eindeutige Grenze zwischen ihnen gezogen werden kann. Liegen die Klassen sehr weit auseinander (Abstand geht theoretisch gegen unendlich), so geht die Sigmoidfunktion in eine Stufenform über. Um das zu vermeiden, sollen die numerischen Werte von  $w$  möglichst klein bleiben. Dafür wird der negative Logarithmus  $E(w)$  mithilfe eines Gauss-Modells mit Erwartungswert  $E(w) = 0$  ergänzt:

$$E(w_k) = - \sum_{n=1}^N \sum_{k=1}^K (t_{nk} \cdot \ln(y_{nk})) + \frac{w^T \cdot w}{2 \cdot \sigma^2}$$

Der Hyperparameter  $\sigma$  beschreibt die Standardabweichung des Gauss-Modells wird hierfür vorgegeben. Wird diese Funktion nun minimiert, so werden die Werte von  $w$  klein (Der Gradient als auch die Hesse-Matrix werden ebenfalls entsprechend erweitert) [Bishop, 2006].

## 2.4 Evaluierungsmethoden

### 2.4.1 Konfusionsmatrix

Um die Ergebnisse beurteilen zu können, werden die klassifizierten Bilder mit verschiedenen Referenzen verglichen. Für die Superpixelklassifikation erfolgt dieser Vergleich einmal anhand der gegebenen Einzelpixelreferenz und einmal anhand einer erstellten Superpixelreferenz. Im ersten Fall wird das Ergebnis Pixel für Pixel mit der Einzelpixelreferenz verglichen, während im zweiten Fall der Vergleich Superpixel für Superpixel stattfindet. Für die Einzelpixelklassifikation findet nur der Vergleich mit der Einzelpixelreferenz statt.

Anhand der Konfusionsmatrix können verschiedene Genauigkeitsmaße berechnet werden, um die Ergebnisse vergleichbar zu machen. Die Konfusionsmatrix besteht aus  $n \times n$  Zeilen und Spalten, wobei  $n$  die Anzahl der Klassen darstellt. Die Zeilen der Matrix entsprechen den Klassen der Referenz und die Spalten den Klassen im klassifizierten Bild. Für jeden Eintrag in der Matrix wird nun gezählt, wie oft ein Pixel der jeweiligen Referenz in der Klassifikation der jeweiligen Klasse zugeordnet wurde. Auf der Diagonalen stehen somit die Einträge die in der Referenz und in der Klassifikation derselben Klasse zugeordnet wurden.

Es werden nun einige Genauigkeitsmaße berechnet, um das Ergebnis besser bewerten und vergleichen zu können. Die Gesamtgenauigkeit (Overall Accuracy) ergibt sich aus der Summe der Diagonaleinträge der Matrix. Diese werden vorher mit der Gesamtanzahl der Pixel oder Segmente normiert. Auf diese Weise wird eine durchschnittliche Genauigkeit berechnet, die allerdings nicht darauf eingeht welche Klassen besonders gute oder besonders schlechte Ergebnisse liefern

[Stoffels, 2009].

$$OA = \frac{1}{n} \sum_{n=1}^N K(n, n) \cdot 100$$

N ist dabei die Anzahl der Klassen und  $K(n, n)$  der jeweilige Diagonaleintrag in der Konfusionsmatrix. Der Parameter n ist die Gesamtanzahl der Primitive aller Klassen.

Um auch die Ergebnisse der einzelnen Klassen unabhängig von einander beurteilen zu können werden die Pixel in vier verschiedene Gruppen eingeordnet. Dies geschieht nacheinander für jede Klasse einmal [Rutzinger et al., 2009]:

- True positive (TP): Anzahl der Pixel die, in der Referenz zu der aktuellen Klasse gehören und in der Klassifikation auch dieser Klasse zugeordnet wurden. In der Konfusionsmatrix ist dies der jeweilige Diagonaleintrag.
- True negative (TN): Anzahl der Pixel, die nicht zu der aktuellen Klasse gehören, dieser aber auch nicht zugeordnet wurden.
- False positive (FP): Anzahl der Pixel die, in der Referenz nicht zu der aktuellen Klasse gehören, ihr aber zugeordnet wurden.
- False negative (FN): Anzahl der Pixel die, in der Referenz zu der aktuellen Klasse gehören, ihr aber nicht zugeordnet wurden

Aus diesen Gruppen lassen sich nun folgende drei Genauigkeitsmaße für jede Klasse berechnen [Rutzinger et al., 2009]:

- Correctness : Beschreibt wie viele Pixel einer Klasse auf dem Bild mit der Referenz übereinstimmen.

$$Correctness = \frac{TP}{TP + FP}$$

- Completeness: Dieses Maß gibt an, wie viele Pixel der Referenzklasse im Ergebniss korrekt detektiert werden.

$$Completeness = \frac{TP}{TP + FN}$$

- Quality: Gibt die Gesamtqualität der Klasse an. Dieses Maß ist nur hoch, wenn sowohl die Correctness als auch die Completeness gute Werte liefern. Aus diesem Grund wird die

Qualität der einzelnen Klassen später auch zum Vergleich herangezogen, da sie Correctness und Completeness zusammenfasst

$$Quality = \frac{TP}{TP + FP + FN}$$

## 3 Vorgehen

Für die Untersuchungen sind zwei Datensätze gegeben, die im Folgenden vorgestellt werden. Für beide Datensätze werden zunächst die Parameter des SLIC-Algorithmus festgestellt und anschließend die logistische Regression sowie Random Forest untersucht. Anschließend erfolgt ein Vergleich mit der Klassifikation von Einzelpixeln.

### 3.1 Vorstellung der Datensätze

#### 3.1.1 Datensatz 1: Vaihingen

Gegeben sind 16 TOPs mit einer Bodenpixelgröße (GSD) von 8 Zentimetern. Aufgenommen wurden die Bilder im Sommer 2008 über der kleinen Stadt Vaihingen durch die Firma „RWE Power“. Genutzt wurde die digitale Großformat-Luftbildkamera Integraph/ZI DMC mit einer Kamerakonstante von 120mm. Die radiometrische Auflösung beträgt 11 Bit. Diese beschreibt, wie viele verschiedene Grauwerte ein Pixel annehmen kann [?]. 11 Bit bedeutet, dass  $2^{11} = 2048$  unterschiedliche Werte angenommen werden können. Die Flughöhe über Grund beträgt 900 Meter und die Längs- und Querüberdeckung 60% ??.

Die Bilder wurden von der deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) für die Erprobung verschiedener digitaler Luftbildkameras genutzt und werden jetzt zusammen mit einem digitalen Oberflächenmodell (DOM) und einer Referenz zur Verfügung gestellt.

Das DOM wurde durch dense-image-matching mit der Software Trimble INPHO 5.3 erstellt. Die Referenz gibt für jedes Pixel an, zu welcher Objektklasse es gehört. Hier wird unterschieden zwischen den Klassen Straße, Gebäude, niedrige Vegetation (Gras), hohe Vegetation (Bäume), Autos und Wasser. Da die Klasse Auto sehr klein ist wird sie mit der Klasse Straße zusammengefasst. Die Klasse Wasser ist nur in wenigen Bildern vorhanden und wird deshalb nur trainiert und nicht mit ausgewertet.

Für die Klassifikation werden die Bilder aufgeteilt in zwei Gruppen, wobei darauf geachtet wird, dass ähnliche Bilder in unterschiedliche Gruppen kommen. Es wird nun immer eine Gruppe zum

Training genutzt und die zweite für die Klassifikation um möglichst aussagekräftige Ergebnisse zu erhalten.

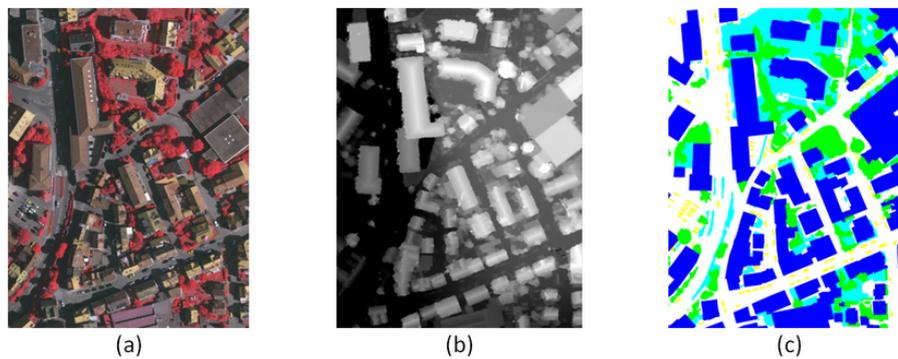


Abbildung 5: Beispiel für ein TOP (a), ein DOM (b) und die dazugehörige Referenz (c) für den ersten Datensatz aus Vaihingen [Sohn et al., 2013]

### 3.1.2 Datensatz 2: Saudi-Arabien

Der zweite Datensatz besteht aus einem Bildausschnitt, das für die Klassifikation in vier gleich große Bilder aufgeteilt wurde. Aufgenommen wurden es über der Stadt Riyadh aus dem Satelliten GeoEye, der in einer Höhe von etwa 660 Kilometer die Erde umkreist. Die Bodenpixelgröße beträgt 50 Zentimeter. Diese Daten sind die Gleichen, mit denen auch schon im Paper von [Alobeid et al., 2011] im Auftrag der ISPRS gearbeitet wurde.

Mithilfe der Software „ERDAS“ wurde sowohl das digitale Oberflächenmodell als auch die Referenz erstellt. Für das DOM wurde mit der Methode Semi Global Matching (SGM) gearbeitet. Außerdem wurden hier in der Nachbearbeitung mithilfe der morphologische Filter opening und closing Ausreißer eliminiert. Die per Hand erstellte Referenz besitzt hier die Klassen Asphalt, Baum (Vegetation), Gebäude, Straße und Boden.

Die vier Bilder werden ebenfalls in zwei Gruppen geteilt, wobei mit der einen trainiert und mit der zweiten klassifiziert wird.

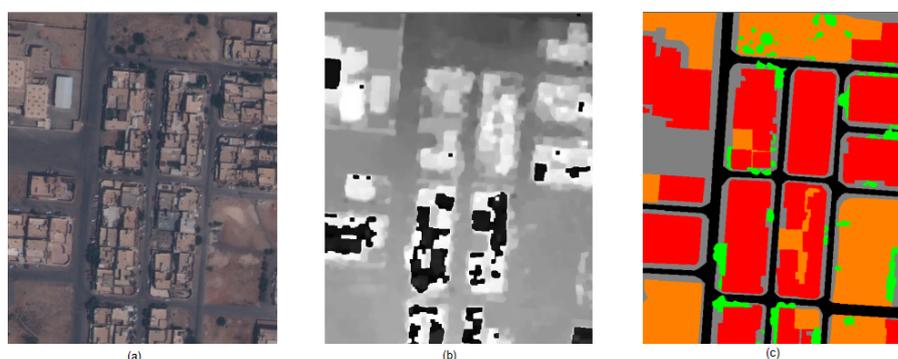


Abbildung 6: Beispiel für ein TOP (a), ein DOM (b) und die dazugehörige Referenz (c) vom zweiten Datensatz aus Saudi-Arabien

In Abb.6 erkennt man schwarze Lücken auf dem DOM (b). Diese befinden sich fast alle auf Gebäuden und entstehen durch Matching-Fehler bei der Erstellung des DOMs. Auf die Auswirkungen dieser Lücken im Höhenmodell wird später eingegangen.

### 3.2 Einzelpixel- und Superpixelreferenz

Die Berechnung der Konfusionsmatrizen erfolgt bei allen segmentbasierten Untersuchungen einmal mit der Einzelpixelreferenz und einmal mit einer Superpixelreferenz. Die Einzelpixelreferenz ist die gegebene Referenz, in der jedes Pixel einer Klasse zugewiesen ist. Diese Referenz wird Pixel für Pixel mit dem klassifizierten Bild verglichen. Die Superpixelreferenz wird für jedes klassifizierte Bild erstellt und beinhaltet genau dieselben Superpixel. Diese Superpixel werden mithilfe der Einzelpixelreferenz der Klasse zugeordnet, die am Häufigsten darin vorkommt.

Da die Superpixelreferenz dem Bild entspricht, das im Idealfall bei der Klassifikation der Superpixel herauskommen kann, wird zunächst für den Vergleich der Superpixelreferenz mit der Einzelpixelreferenz eine Konfusionsmatrix erstellt. Die entstehenden Genauigkeiten entsprechen den maximal erreichbaren Genauigkeiten des Einzelpixelvergleichs.

Anschließend erfolgt die Berechnung der Konfusionsmatrizen für das klassifizierte Bild einmal mit der Einzelpixel- und einmal mit der Superpixelreferenz.

### 3.3 Vorgehen zur Parameterauswahl für den SLIC-Algorithmus

Der SLIC-Algorithmus wird auf die Eingangsparameter Kanäle, Größe der Superpixel und Kompaktheit der Superpixel hin untersucht. Die drei Kanäle werden mit verschiedenen Kombinationen von Merkmalen besetzt und so getestet. Die Superpixelgröße und die Kompaktheit werden immer abhängig voneinander untersucht. Das bedeutet, dass eine Pixelgröße für eine Reihe von Kompaktheiten beibehalten wird. So entsteht ein Raster und es ist möglich eine Tendenz zu erkennen, mit welchen Kombinationen die Ergebnisse am besten werden.

Dieses Vorgehen wird anhand beider Datensätze durchgeführt. Mit den jeweils ausgewählten Parametern wird anschließend die Untersuchung der Klassifikationsmethoden durchgeführt.

### 3.4 Vorgehen zur Parameterauswahl für Random Forest

Sowohl Random Forest als auch die logistische Regression werden für jeden Datensatz mit der Hälfte der Bilder trainiert und mit der anderen klassifiziert.

Für den Datensatz aus Vahingen ergeben sich bei einer Superpixelgröße von 40x40 Einzelpixeln im Mittel 3000 Superpixel für ein Bild. Für alle acht Bilder die zum Training verwendet werden, ergibt sich so eine Anzahl von etwa 24.000 Superpixeln die insgesamt zum Training genutzt werden können. Für das Saudi-Arabien-Bild ergeben sich mit einer Superpixelgröße von 15x15 Einzelpixeln etwa 3300 Superpixel für beide Bildausschnitte die zum Training verwendet werden.

Trainiert wird mit 100 Bäumen und an jedem Knoten werden die Merkmale drei mal unterschiedlich gewichtet. Pro Gewichtung werden anschließend 4 Grenzen zufällig erzeugt und für jede wird der Informationszuwachs  $\Delta E$  berechnet. Als Abbruchkriterien ist eine maximale Tiefe von 9 angegeben und der minimale Informationszuwachs mit 0,01.

Es gilt nun herauszufinden, welche Merkmale sich für die Klassifikation mit Random Forest am besten eignen. Mit der Methodier „calcVarimportance“ kann eine Reihenfolge erstellt werden, die den Einfluss der einzelnen Merkmale auf das Klassifizierungsergebnis darstellt. Die Methode trainiert zunächst mit allen ausgewählten Merkmalen den Klassifikator und speichert das Ergebnis. Anschließend wird nacheinander jedes Merkmal einmal ausgewählt und für jedes Trainingsprimitiv mit zufälligen Werten gefüllt. Das Training erfolgt nun wie zu Anfang. Merkmale die einen großen Einfluss auf das Ergebnis haben, werden dadurch auffallen, dass die Gesamtgenauigkeit des Klassifikationsergebnis signifikant schlechter wird, wenn das Merkmal mit zufälligen Werten gefüllt ist. Je größer diese Abweichung zum Anfangsergebnis für ein Merkmal ist, desto höher steht das Merkmal in dem Ranking, das von der Methode ausgegeben wird.

Mithilfe dieses Ranking werden nun nacheinander Klassifikationen mit immer einem Merkmal mehr durchgeführt, wobei immer das nächste Merkmal des Ranking hinzugefügt wird. Durch einen Vergleich der Ergebnisse, die wieder jeweils mit der Einzelpixel- und der Superpixelreferenz verglichen werden, kann hinterher entschieden werden welche und wie viele Merkmale sinnvoll für eine gute Klassifikation sind, und ab wann die Klassifikation nicht mehr deutlich besser wird.

### 3.5 Vorgehen zur Parameterauswahl für die logistische Regression

Das Training der logistischen Regression erfolgt ebenfalls mit jeweils der Hälfte der gegebenen Bilder pro Datensatz. Dabei werden alle vorhandenen Trainingsprimitiv genutzt. Für das Features Space Mapping wird die quadratische Erweiterung gewählt und im Gauss-Modell wird  $\sigma = 10$  genutzt. Dieser Wert hat sich in vorherigen Untersuchungen ergeben.

Als Anhaltspunkt für die Merkmale die einen großen Einfluss auf das Training haben werden die Merkmale aus dem Ranking von Random Forest genutzt. Da für die logistische Regression nicht automatisch von der gleichen Reihenfolge ausgegangen werden kann, werden immer mehrere Merkmale getestet, um anschließend auszuwählen welches das beste Ergebnis liefert. So kommt es zu leicht anderen Reihenfolgen als bei Random Forest.

Anschließend können die Genauigkeitsmaße abhängig von der Anzahl der Merkmale verglichen werden. Dies geschieht ebenfalls anhand der Einzel- und Superpixelreferenz. So kann ausgewählt werden wie viele Merkmale sich für das Training und die anschließende Klassifikation mit der logistischen Regression eignen.

Sind die Untersuchungen für beide Klassifikatoren abgeschlossen erfolgt ein Vergleich von Random Forest und der logistischen Regression.

### 3.6 Vergleich von segment- und pixelbasierter Klassifikation

Um die Ergebnisse der segmentbasierten Klassifikation bewerten zu können erfolgt ein Vergleich mit einer pixelbasierten Klassifikation. Dies geschieht unabhängig voneinander für beide Klassifikationsmethoden. Aufgrund der sehr viel höheren Anzahl an Trainingsprimitiven (siehe Tabelle 2), benötigen beide Klassifikatoren hier deutlich mehr Zeit. Der Vergleich der Ergebnisse erfolgt mit den Superpixelgrößen die für jeden Datensatz die besten Ergebnisse geliefert haben und mithilfe der Merkmale die für den jeweiligen Klassifikator ausgewählt wurden. Dabei kann auf die Ergebnisse der Untersuchungen der segmentbasierten Klassifikation zurückgegriffen werden.

Das Ziel des Vergleichs ist eine abschließende Bewertung der segmentbasierten Klassifikation mit dem SLIC-Algorithmus.

	Vaihingen (8 Bilder)	Saudi-Arabien (2 Bilder)
Anz. Pixel	40.000.000	700.000
Superpixelgröße	40	15
Anz. Superpixel	24.000	3.100

Tabelle 2: Anzahl der Pixel und Superpixel für die Trainingsbilder aus Vaihingen und Saudi-Arabien

## 4 Ergebnisse

### 4.1 SLIC-Algorithmus

#### 4.1.1 Vaihingen

##### Die drei Kanäle:

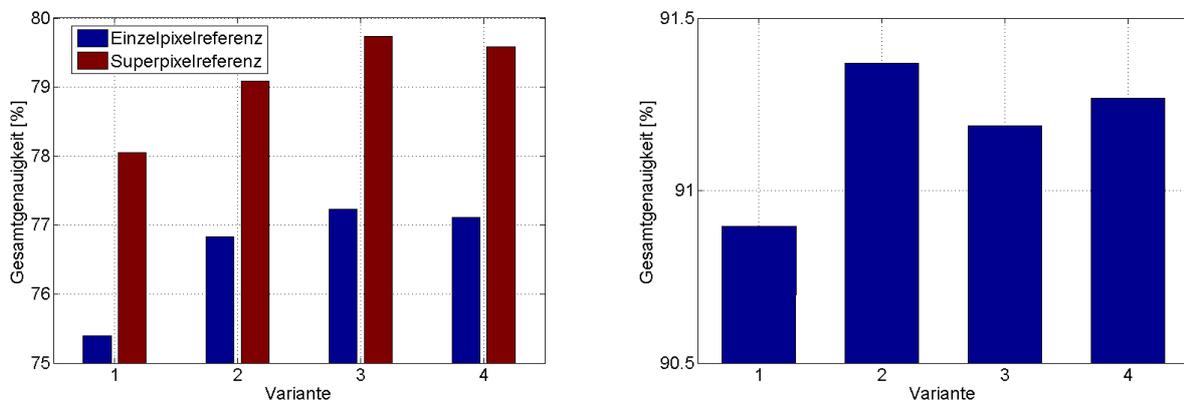
Für die Untersuchung der Kanäle für den SLIC-Algorithmus wird mit den in Kapitel 2.2 vorgestellten Merkmalen gearbeitet. Diese werden in unterschiedlichen Kombinationen zusammengefasst, und die Klassifikationsergebnisse werden im Anschluss miteinander verglichen. Alle anderen Parameter (wie Superpixelgröße, Kompaktheit und die Parameter des Klassifikators) werden dabei nicht verändert und sind in jedem Durchlauf gleich.

Im Folgenden wird eine Auswahl an Kombinationen graphisch dargestellt, die sowohl die besten als auch die schlechtesten Kombinationen enthält. Als Klassifikator wird hierzu die logistische Regression gewählt.

- Variante 1: Farbkanal 1, Farbkanal 2, Farbkanal 3

- Variante 2: Sättigung, Intensität, Farbkanal 1
- Variante 3: NDVI, NDOM, Farbkanal 1
- Variante 4: NDVI, NDOM, Intensität

In Abb.7-(a) sind die Gesamtgenauigkeiten der einzelnen Varianten zu sehen. Dabei wird für jede Variante der Vergleich mit der Einzelpixelreferenz sowie der Vergleich mit der Superpixelreferenz dargestellt. In Teil (b) der Abbildung sind die maximal zu erreichenden Gesamtgenauigkeiten für die Einzelpixelreferenz zu sehen.



(a) Gesamtgenauigkeiten der Varianten 1 bis 4, jeweils für die Einzelpixel- und die Superpixelreferenz (b) Maximal Erreichbare Gesamtgenauigkeiten beim Vergleich mit der Einzelpixelreferenz

Abbildung 7: Gesamtgenauigkeiten und die maximal erreichbare Gesamtgenauigkeit des Einzelpixelvergleichs

Die Differenzen zwischen dem schlechtesten und dem besten Ergebnis liegen für die Gesamtgenauigkeit bei knapp 2%. Dabei sind die Gesamtgenauigkeiten des Vergleichs mit der Superpixelreferenz etwa zwei bis drei Prozent besser als die des Vergleichs mit der Einzelpixelreferenz. Das liegt daran, dass die erreichbaren (Gesamt-) genauigkeiten des Vergleichs mit der Einzelpixelreferenz maximal 90,5 - 91,5% erreichen kann (siehe Abb.7-(b)). Beim Vergleich mit der Superpixelreferenz können hingegen theoretisch 100% erreicht werden.

Für die Qualitätsmaße der einzelnen Klassen sieht es genauso aus. Die Differenzen liegen hier bei maximal 4% und die Unterschiede zwischen der Einzelpixel- und der Superpixelreferenz bei maximal 6% (siehe Anhang).

Deutlich zu sehen ist in allen Fällen, dass Variante 1 jeweils das schlechteste Ergebnis liefert und Variante 3 sehr knapp das Beste aller getesteten Kombinationen. In Variante 1 wird nur mit den Farbkanälen des Bildes gearbeitet. Das schlechte Abschneiden dieser Variante zeigt, dass andere Merkmale wichtige Eigenschaften liefern, um Grenzen im Bild noch besser zu finden. Variante 3 besteht aus den Merkmalen NDVI, NDSM und dem ersten Farbkanal. Dies ist eine gute Kombination, da sowohl die Höhenunterschiede, als auch die Vegetation gut erfasst werden

können. Diese Variante wird für die weitere Untersuchung ausgewählt. Die Genauigkeitsmaße dieser Variante sind in Tabelle 3 zusammengefasst.

<b>Klasse</b>	<b>Straße</b>		<b>Gebäude</b>		<b>hohe Veg.</b>		<b>niedrige Veg.</b>	
<b>Vergleich</b>	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.
<i>Correctness</i> [%]	82,7	83,7	74,2	75,4	71,0	76,3	72,9	77,6
<i>Completeness</i> [%]	74,4	76,5	82,8	85,1	66,6	69,4	83,4	87,1
<i>Quality</i> [%]	64,5	66,7	64,7	67,0	52,0	56,6	62,5	68,5
<b>Gesamtgenauigkeiten</b> [%]	Einzelpixelreferenz				Superpixelreferenz			
	77,2				79,7			

Tabelle 3: Genauigkeitsmaße der ausgewählten Variante 3 mit den Merkmalen NDVI, NDOM und Farbkanal 1

Insgesamt haben die drei Kanäle allerdings keinen großen Einfluss auf die Güte der Superpixel. Abweichungen von bis zu 2% in der Gesamtgenauigkeit und bis zu 4% in den Qualitätsmaßen der einzelnen Klassen beeinflussen das Ergebnis nur wenig. Trotzdem wurde das bestmögliche Ergebnis ausgewählt.

### **Kompaktheit und Pixelgröße**

Die Parameter Kompaktheit und Pixelgröße werden gemeinsam untersucht. Dies geschieht in einem Raster, wobei die Kompaktheit Werte von 10 bis 80 annimmt und die Pixelgröße Werte zwischen 10 und 60. Die Ergebnisse für die Gesamtgenauigkeiten anhand des Vergleichs mit der Einzelpixelreferenz sind in Abb.8 graphisch dargestellt.

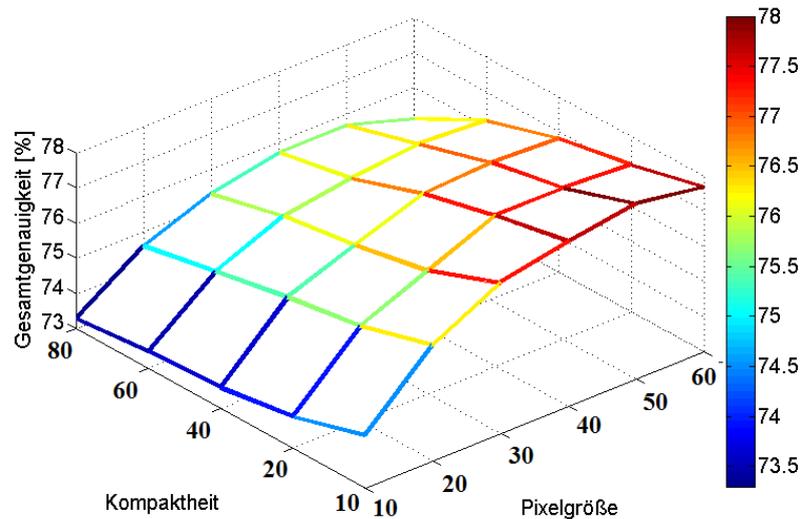
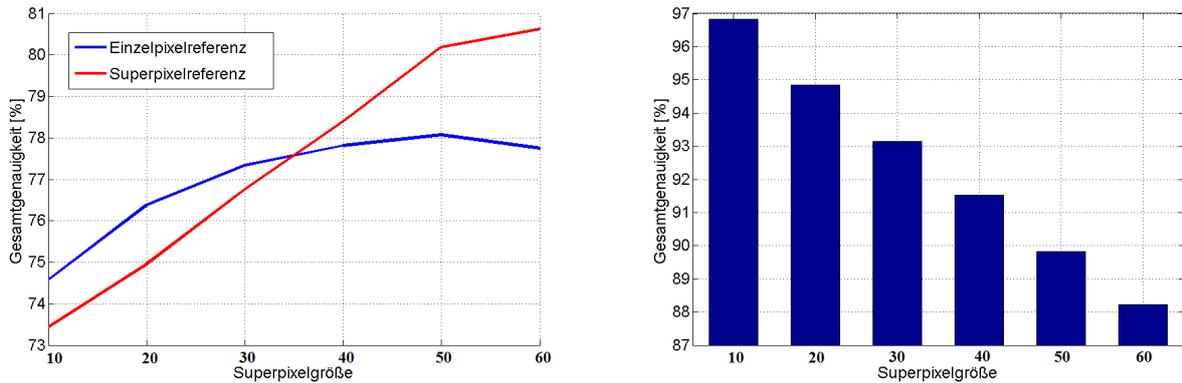


Abbildung 8: Gesamtgenauigkeiten abhängig von Pixelgröße und Kompaktheit für den SLIC-Algorithmus mit dem Datensatz aus Vaihingen, verglichen mit der Einzelpixelreferenz

Der segmentbasierte Vergleich erfolgt stichprobenartig. In Abb.9 sind für die Einzel- und auch die Superpixelreferenz - für eine Kompaktheit von 10 - die Gesamtgenauigkeiten, abhängig von der Superpixelgröße, dargestellt. Außerdem sind die maximal erreichbaren Genauigkeiten für den Vergleich mit der Einzelpixelreferenz abhängig von der Superpixelgröße aufgetragen.

Es fällt auf, dass die maximal erreichbare Gesamtgenauigkeit mit der Einzelpixelreferenz kleiner wird, je größer die Superpixel werden. Das liegt daran, dass in kleinen Superpixeln jedes einzelne Pixel einen größeren Einfluss auf das Klassifikationsergebnis hat. Sehr kleine Bereiche einer Klasse können so noch als diese dargestellt werden, während sie in einem größeren Superpixel durch die umliegenden Pixel einer anderen Klasse ebenfalls dieser zugeordnet werden.

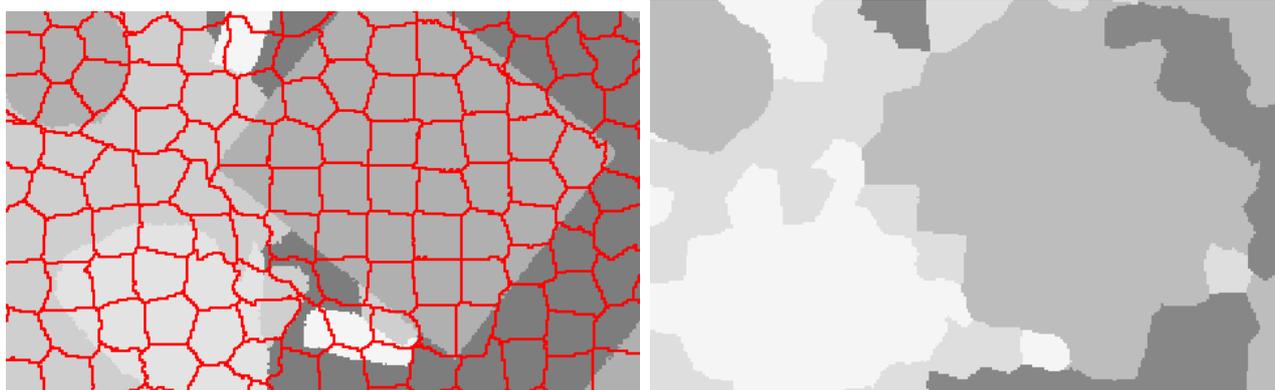
Dementsprechend sieht auch der Vergleich zwischen beiden Referenzen aus. Die Einzelpixelreferenz kann sich ab einer gewissen Superpixelgröße nicht mehr so stark verbessern, da ihre maximal erreichbare Genauigkeit immer kleiner wird. Die Superpixelreferenz kann in der Theorie immer die 100% Übereinstimmung erreichen. Man erkennt hier deutlich, dass die Klassifikationsergebnisse besser werden, je größer die Superpixel sind.



(a) Gesamtgenauigkeiten für den Vergleich mit der Einzel- und Superpixelreferenz (b) Maximal erreichbare Genauigkeiten abhängig von der Pixelgröße

Abbildung 9: Gesamtgenauigkeiten und maximal erreichbare Genauigkeiten für die Einzelpixelreferenz, abhängig von der Größe der Superpixel

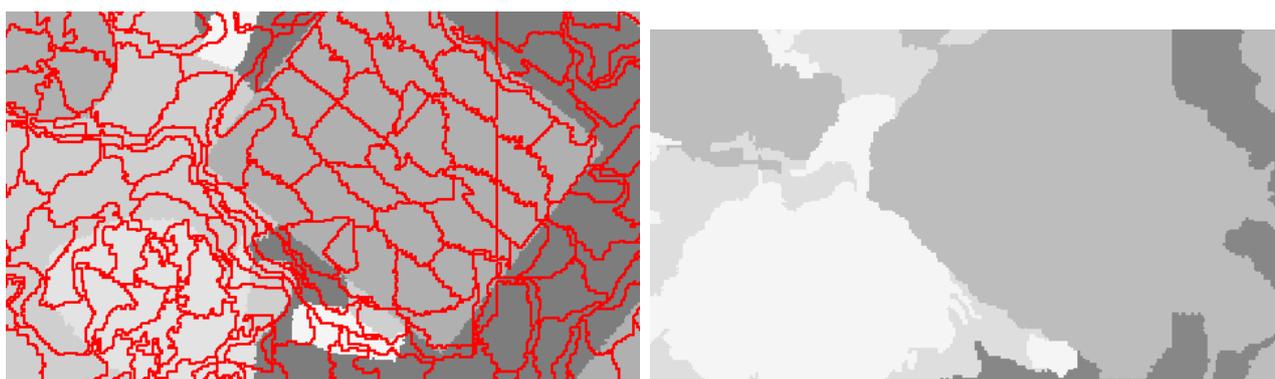
In Abb.8 ist zu sehen, dass die Ergebnisse für große Superpixelgrößen und kleine Werte für die Kompaktheit tendenziell besser werden. Um dies erklären zu können, werden in Abb.10 und 11 die Superpixel-Grenzen für eine Kompaktheit von 80 und für eine Kompaktheit von 10 gezeigt. Die Superpixelgröße beträgt in beiden Fällen 20. Es ist deutlich zu erkennen, dass durch einen hohen Wert für die Kompaktheit die Superpixel immer mehr die Form von Quadraten annehmen, und die Grenzen des Hauses und der anderen Objekte im Bild nicht mit den Grenzen der Superpixel übereinstimmen. Hier hat der räumliche Abstand der Superpixel einen viel höheren Einfluss auf die Erstellung der Superpixel als die Farben der einzelnen Pixel. Bei einem Wert von 10 für die Kompaktheit werden die Superpixel sehr unregelmäßig geformt. Hier hat die Farbe der einzelnen Pixel einen vergleichsweise größeren Einfluss als der räumliche Abstand. Die ideale Kompaktheit liegt also im Mittel zwischen diesen beiden Extremen. Dann werden die Superpixel nicht mehr unförmig und trotzdem hat die Farbe noch einen starken Einfluss auf ihre Erstellung.



(a) Superpixel vor dem Referenzbild

(b) Ergebnis der Klassifikation der einzelnen Superpixel

Abbildung 10: Ergebnis der Superpixel-Aufteilung und der anschließenden Klassifikation bei einer Kompaktheit von 80 und einer Pixelgröße von 20 für den ersten Datensatz



(a) Superpixel vor dem Referenzbild

(b) Ergebnis der Klassifikation der einzelnen Superpixel

Abbildung 11: Ergebnis der Superpixel-Aufteilung und der anschließenden Klassifikation bei einer Kompaktheit von 10 und einer Pixelgröße von 20 für den ersten Datensatz

Um noch besser beurteilen zu können, welche Werte sich für die Parameter am besten eignen, werden in Tabelle 4 die Flächen auf dem Boden für alle verwendeten Superpixelgrößen aufgezählt. Da es zwischen den Ergebnissen (Abb.8) für die Superpixelgrößen von 40, 50 und 60 kaum Differenzen gibt, wird als Superpixelgröße 40 ausgewählt. So sind die Pixel auf dem Boden mit  $10m^2$  nicht zu groß und es können auch Details noch erkannt werden.

Als Wert für die Kompaktheit wird 40 ausgewählt. Die etwas besseren Ergebnisse (etwa 1% besser in der Gesamtgenauigkeit) liefert zwar eine Kompaktheit von 10, allerdings werden die Superpixel dann sehr unregelmäßig geformt und auch die Klassengrenzen im klassifizierten Bild sind sehr unregelmäßig (siehe Abb.11). Mit einer Kompaktheit von 40 ist das nicht mehr der Fall.

<b>Superpixelgröße</b>	10	20	30	40	50	60
<b>Fläche [m<sup>2</sup>]</b>	0,64	2,56	5,76	10,24	16	23,04

Tabelle 4: Größe der Superpixel auf dem Boden bei einer Auflösung von 8 Zentimeter pro Einzelpixel

Die Genauigkeitsmaße der ausgewählten Kombination mit einer Superpixelgröße von 40 und einer Kompaktheit von 40 sind in Tabelle 5 zusammengefasst.

<b>Klasse</b>	<b>Straße</b>		<b>Gebäude</b>		<b>hohe Veg.</b>		<b>niedrige Veg.</b>	
<b>Vergleich</b>	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.
<i>Correctness</i> [%]	82,7	83,7	74,2	75,4	71,0	76,3	73,0	77,6
<i>Completeness</i> [%]	74,4	76,6	82,8	85,1	66,7	69,4	83,4	87,1
<i>Quality</i> [%]	64,5	66,7	64,7	67,0	52,0	56,6	62,5	68,5
<b>Gesamtgenauigkeiten</b> [%]	Einzelpixelreferenz				Superpixelreferenz			
	77,2				79,7			

Tabelle 5: Genauigkeitsmaße der ausgewählten Kombination mit einer Pixelgröße von 40 und einer Kompaktheit von 40

Nach der Untersuchung des SLIC-Algorithmus mit der logistischen Regression wird für ein etwas grober gewähltes Raster die gleiche Untersuchung mit Random Forest durchgeführt, um zu überprüfen, ob es hier signifikante Unterschiede gibt. Auch hier zeigt sich, dass die Ergebnisse besser werden, je größer die Superpixel gewählt werden und je geringer der Wert für die Kompaktheit ist. Aus denselben Gründen wie bei der Untersuchung mit der logistischen Regression (sehr unregelmäßige Superpixel für kleine Kompaktheiten) wird auch hier nicht der geringste Wert für die Kompaktheit genutzt, sondern ein Wert von 40. Ebenso wird der Wert von 40 für die Pixelgröße gewählt.

Insgesamt werden sowohl für die logistische Regression als auch für Random Forest folgende Parameter ausgewählt:

- Merkmale: NDVI, NDSM, 1. Farbkanal
- Superpixelgröße 40 bei einer Kompaktheit von 40

#### 4.1.2 Saudi-Arabien

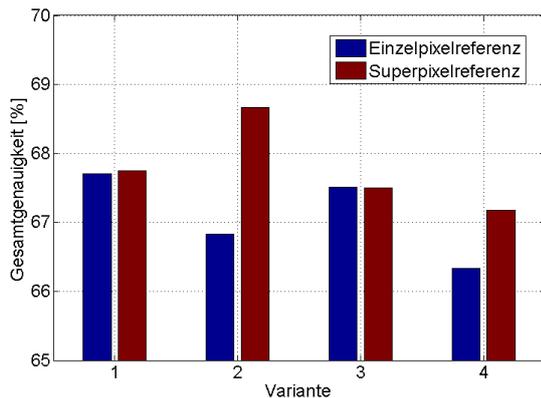
##### Merkmale:

Nun wird auch für die vier Bildausschnitte aus Saudi-Arabien ebenfalls untersucht, welche Merkmale sich zur Erstellung der Superpixel eignen. Hierzu werden wieder die Merkmale aus Kapitel 2.2 in unterschiedlichen Kombinationen untersucht. Der NDVI fällt in diesem Fall heraus, da kein Infrarotkanal vorliegt. Alle anderen Parameter werden dabei wieder nicht verändert.

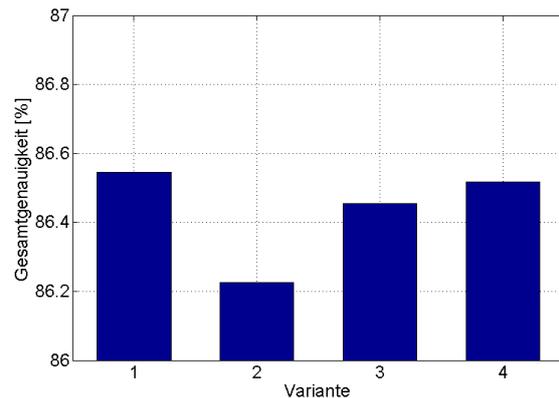
Im Folgenden wird eine Auswahl der Kombinationen graphisch dargestellt, die sowohl die besten als auch die schlechtesten Kombinationen enthält. Als Klassifikator wird hier ebenfalls die logistische Regression gewählt.

- Variante 1: Sättigung, NDOM, Farbkanal 1
- Variante 2: Intensität, NDOM, Farbkanal 1
- Variante 3: Intensität, Sättigung, NDOM
- Variante 4: Intensität, Sättigung, Farbkanal 1

In Abb. 12 sind die Gesamtgenauigkeiten der Varianten graphisch dargestellt. Für jede Variante erfolgt der Vergleich einmal mit der Einzelpixelreferenz und einmal mit der Superpixelreferenz. Die maximal erreichbaren Genauigkeiten für die Einzelpixelreferenz sind in Abb.12-(b) zu sehen. Die Qualitäten der einzelnen Klassen sind im Anhang zu finden.



(a) Gesamtgenauigkeit der Varianten



(b) Maximal erreichbare Genauigkeiten des Vergleichs mit der Einzelpixelreferenz

Abbildung 12: Gesamtgenauigkeiten und maximal erreichbare Genauigkeiten für die Einzelpixelreferenz, abhängig von der Größe der Superpixel für Saudi-Arabien

Die erste und zweite Variante liefern insgesamt das beste Ergebnis. Die Differenzen zwischen den Genauigkeiten sind allerdings für alle Varianten sehr gering. Bei Variante 2 fällt auf, dass die Differenz zwischen dem pixel- und superpixelweisen Vergleich mit 2% im Vergleich sehr hoch ist. Für Variante 1 liegen beide Ergebnisse gleichauf.

Insgesamt sticht keine Variante besonders heraus. Ausgewählt wird für die weitere Untersuchung die erste Variante, da hier der Einzel- sowie der Superpixelvergleich gute Ergebnisse liefern. Die verwendeten Kanäle lauten Sättigung, Höhe und Farbkanal 1. Die Genauigkeitsmaße dieser Kombination sind in Tabelle 6 zusammengefasst.

Die Klasse Asphalt wurde dabei jeweils am schlechtesten klassifiziert (*Correctness* und *Completeness* bei 42 - 46% während andere Klassen Genauigkeitsmaße von bis zu 80% haben). Darauf wird

im Kapitel „Ergebnisse der Klassifikationen“ noch näher eingegangen.

Klasse	Asphalt		hohe Veg.		Gebäude		Straßen		Boden	
	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.
<b>Vergleich</b>										
<i>Correctness</i> [%]	42,3	46,8	52,7	57,2	84,7	82,5	67,2	69,3	64,1	62,0
<i>Completeness</i> [%]	42,3	45,8	68,3	75,9	80,7	80,6	68,0	69,5	66,1	60,2
<i>Quality</i> [%]	26,7	29,9	41,8	47,5	70,4	68,8	50,9	53,0	48,2	43,9
<b>Gesamtgen. [%]</b>	Einzelpixelreferenz					Superpixelreferenz				
	67,7					67,8				

Tabelle 6: Genauigkeitsmaße der ausgewählten Variante 1 mit den Merkmalen Sättigung, NDOM und Farbkanal 1

### Kompaktheit und Pixelgröße

Die Parameter Kompaktheit und Superpixelgröße werden ebenfalls gemeinsam untersucht. Die Kompaktheit nimmt dabei Werte zwischen 10 und 80 an und die Pixelgröße Werte von 7 bis 30. Da hier die Bodenpixelgröße mit 0,5m viel größer ist als für Vaihingen (0,08m) haben hier auch die Superpixel auf dem Boden ganz andere Größen. Zum Vergleich sind diese einmal zusammen mit den Bodenpixelgrößen aus Vaihingen in Tabelle 7 dargestellt.

Pixelgröße Superpixel	7	10	15	20	30	40	50	60
<b>Fläche bei 0,08 [m<sup>2</sup>] Auflösung</b>	0,31	0,64	1,44	2,56	5,76	10,24	16	23,04
<b>Fläche bei 0,5 [m<sup>2</sup>] Auflösung</b>	12,25	25	56,25	100	225	400	625	900

Tabelle 7: Größe der Superpixel auf dem Boden

Somit ist eindeutig, dass im Fall der Saudi-Arabien-Bilder eine kleinere Superpixelgröße gewählt werden muss, um nicht zu viel Fläche mit einem Pixel abzudecken. Deshalb werden Superpixelgrößen ab 40x40 nicht untersucht.

Alle anderen Parameter werden wieder für die gesamte Untersuchung beibehalten. Als Klassifikator wird die logistische Regression gewählt.

In Abb.13 sind die Ergebnisse für die Gesamtgenauigkeiten anhand des Vergleichs mit der Einzelpixelreferenz graphisch dargestellt.

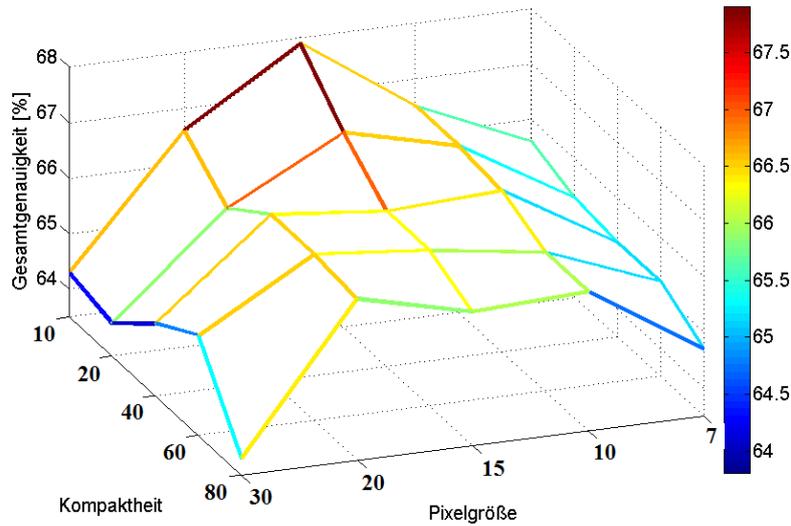
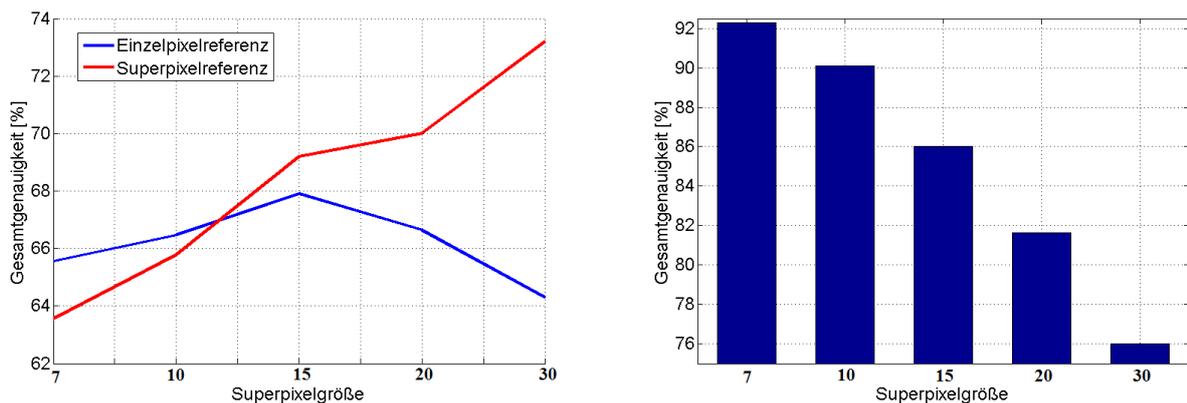


Abbildung 13: Gesamtgenauigkeiten abhängig von Pixelgröße und Kompaktheit für den SLIC-Algorithmus mit den zweiten Datensatz

Der Vergleich mit der Superpixelreferenz erfolgt wieder stichprobenartig. In Abb.14 sind die Gesamtgenauigkeiten für den Vergleich mit beiden Referenzen, abhängig von der Superpixelgröße für eine Kompaktheit von 10 dargestellt. Zu sehen sind außerdem die maximal erreichbaren Genauigkeiten, abhängig von der Superpixelgröße, für den Vergleich mit der Einzelpixelreferenz.



(a) Gesamtgenauigkeiten für den Vergleich mit der Einzel- und Superpixelreferenz (b) Maximal erreichbare Genauigkeiten abhängig von der Pixelgröße

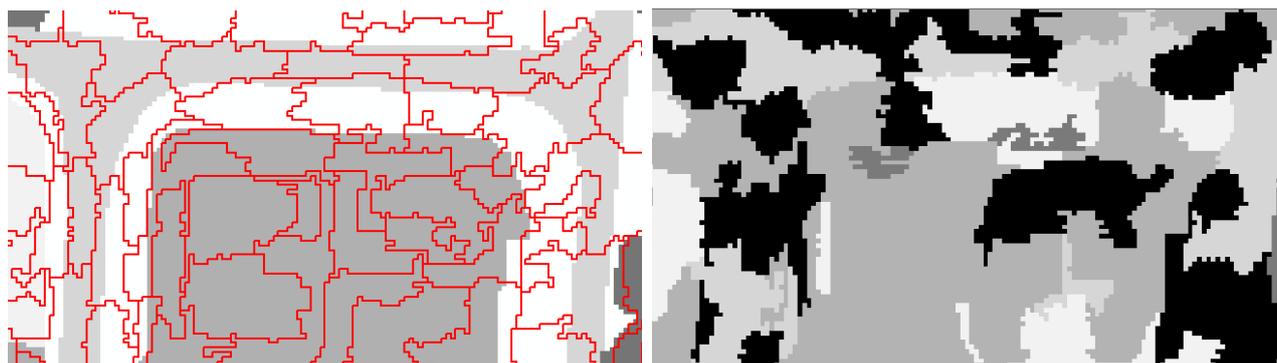
Abbildung 14: Gesamtgenauigkeiten und maximal erreichbare Genauigkeiten für die Einzelpixelreferenz

In Abb. 14 fällt wieder auf, dass die maximal zu erreichende Genauigkeit kleiner wird, je größer die Superpixel werden. Das ist wieder der Grund dafür, dass die Gesamtgenauigkeit ab einer bestimmten Superpixelgröße für den Vergleich mit der Einzelpixelreferenz nicht mehr steigt,

sondern sogar sinkt. Die Gesamtgenauigkeit beim Vergleich mit der Superpixelreferenz steigt wieder, je größer die Superpixel werden.

In Abb.13 ist ein „peak“ für eine Pixelgröße von 15 und einer Kompaktheit von 10 zu sehen. Das gilt jedoch nur für den Vergleich mit der Einzelpixelreferenz, da die maximal erreichbare Gesamtgenauigkeit hier immer weiter abnimmt, je größer die Superpixel sind. Insgesamt gilt wieder, dass das Ergebnis besser wird, je größer die Superpixel werden und je kleiner die Kompaktheit gewählt wird.

Für den Datensatz aus Vaihingen war bei einer Kompaktheit von 10 das Problem, dass die Superpixel teilweise sehr unförmig waren und sich nicht gut an Klassengrenzen angepasst haben. In Abb.15 sind für den Datensatz aus Saudi-Arabien die Superpixelgrenzen für eine Kompaktheit von 10 und eine Superpixelgröße von 15 zu sehen. Die Pixel sind hier viel weniger unförmig und passen sich gut an die Grenze des Gebäudes an (im Bild dunkelgrau). Zwischen den weißen (Asphalt) und hellgrauen (Straße) Flächen hingegen liegen die Superpixelgrenzen nicht auf den Grenzen im Referenzbild. Das liegt jedoch nicht an schlecht gewählten Parametern für den SLIC-Algorithmus, sondern daran, dass die Klassen Asphalt und Straße sehr ähnliche Eigenschaften besitzen und für den Klassifikator sehr schwer zu trennen sind. Diese Phänomen ist auch in Abb.15 -(b) zu erkennen. Viele Superpixel wurden hier falsch klassifiziert und die Form der Straße ist hier kaum noch zu erkennen. Darauf wird im Kapitel der Klassifikatoren genauer eingegangen.



(a) Superpixel vor dem Referenzbild

(b) Ergebnis der Klassifikation der einzelnen Superpixel

Abbildung 15: Ergebnis der Superpixelaufteilung und der anschließenden Klassifikation bei einer Kompaktheit von 10 und einer Superpixelgröße von 15 für den zweiten Datensatz

Aus diesen Gründen wird für die Kompaktheit ein Wert von 10 ausgewählt. Für die Superpixelgröße wird 15 ausgewählt, da so etwa  $56m^2$  Bodenfläche pro Superpixel abgedeckt sind. So können auch schmale Straßen oder andere Flächen noch einem Superpixel zugeordnet werden. In

Tabelle 8 sind die Genauigkeitsmaße für die ausgewählte Kombination zusammengefasst. Hier ist ebenfalls zu sehen, dass die Klasse Asphalt im Vergleich zu den anderen um 30-50% schlechtere Ergebnisse für *Correctness* und *Completeness* hat. Dies liegt an dem oben beschriebenen Phänomen, dass die Klassen Asphalt und Straße sehr schlecht voneinander trennbar sind.

Klasse	Asphalt		hohe Veg.		Gebäude		Straßen		Boden	
Vergleich	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.
<i>Correctness</i> [%]	33,4	37,1	50,8	58,1	86,9	85,6	71,1	73,5	67,4	62,6
<i>Completeness</i> [%]	41,6	44,7	66,7	73,6	80,4	79,7	64,6	69,0	66,4	60,2
<i>Quality</i> [%]	22,6	25,4	40,3	47,6	71,7	70,3	50,8	55,1	50,3	44,3
<b>Gesamtgen. [%]</b>	Einzelpixelreferenz					Superpixelreferenz				
	67,9					69,2				

Tabelle 8: Genauigkeitsmaße der ausgewählten Kombination mit einer Pixelgröße von 15x15 und einer Kompaktheit von 10.

Insgesamt werden für den zweiten Datensatz somit die folgenden Parameterwerte ausgewählt:

- Merkmale: Sättigung, NDSM, erster Farbkanal
- Pixelgröße 15 bei einer Kompaktheit von 10

Mit diesen Parametern kommen für den gegebenen Bildausschnitt die besten Ergebnisse heraus. Im Vergleich mit dem Datensatz aus Vaihingen liegen die erzielten Genauigkeiten aber rund 10-15% darunter. Das liegt vor allem am geringen Umfang der Trainingsdaten, die zur Verfügung stehen. Dadurch kann der Algorithmus die Eigenschaften der einzelnen Klassen nur begrenzt lernen.

## 4.2 Random Forest

### 4.2.1 Vaihingen

Zu Beginn wird für alle verfügbaren Merkmale das Ranking erstellt. In Tabelle 9 sind die ersten zehn Merkmale dieses Ranking zu sehen. Die Untersuchung erfolgt nun in der gleichen Reihenfolge. Das bedeutet, dass zunächst nur mit dem Merkmal trainiert wird, für das sich die höchste Differenz ergeben hat, als es mit zufälligen Werten besetzt wurde. Anschließend wird immer das nächste Merkmal hinzugenommen. Für die Erstellung der Superpixel werden die Parameter gewählt, die in Kapitel 4.1 ausgewählt wurden. Abhängig von der Anzahl der genutzten Merkmale kann nun der Vergleich mit der Referenz erfolgen. Dies geschieht wieder anhand der Einzel- und anhand der Superpixelreferenz. Die Ergebnisse für die Gesamtgenauigkeiten sind in Abb.16 graphisch dargestellt.

Nummer	Merkmal	Einfluss [%]
1	Maximum der Sättigung	7,5
2	Mittelwert des ersten Farbkanals	7,5
3	Mittelwert des NDOM	4,8
4	Mittelwert es zweiten Farbkanals	4,8
5	Minimum des NDVI	4,7
6	Energieverteilung	4,4
7	Standardabw. des zweiten Farbkanals	4,0
8	Minimum des NDOM	3,9
9	Mittelwert des NDVI	3,8
10	Mittelwert Gradient	3,7

Tabelle 9: Ranking und Einfluss der Merkmale auf das Training für Vaihingen mit Random Forest

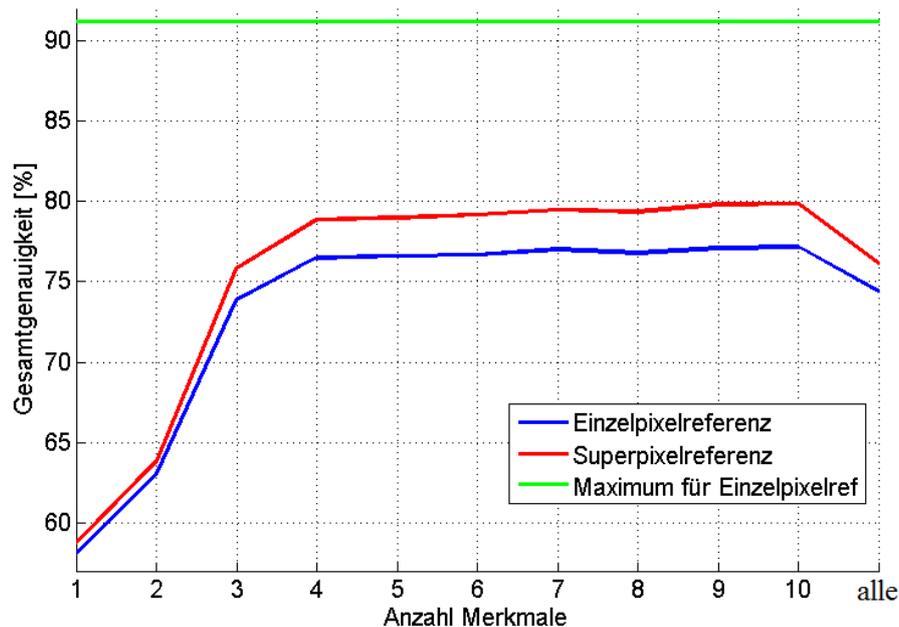


Abbildung 16: Gesamtgenauigkeiten für Vaihingen abhängig von der gewählten Anzahl an Merkmalen mit Random Forest

Es ist deutlich erkennbar, dass die Werte zunächst immer besser werden, wenn ein neues Merkmal hinzukommt. Ab einer gewissen Anzahl an Merkmalen stagniert die Übereinstimmung allerdings, und wird kaum oder gar nicht mehr besser. Diesen Effekt nennt man Sättigungseffekt. Ein neues Merkmal hat ab einer Anzahl von 4 bis 5 Merkmalen keinen erkennbaren Informationszuwachs mehr, um die Genauigkeiten weiter zu verbessern. Die Ergebnisse werden sogar wieder schlechter, wenn alle möglichen Merkmale genutzt werden. Wenn mehr Merkmale

genutzt werden, erhöht sich dementsprechend auch die Dimension des Merkmalsraums bzw. des Merkmalsvektors. Dadurch stehen zum Finden der Entscheidungsgrenzen pro Merkmal weniger Trainingsdaten zur Verfügung, da deren Gesamtanzahl die Gleiche bleibt.

Der Vergleich mit der Superpixelreferenz liefert stets Genauigkeiten, die 2-3% besser sind als die des Vergleichs mit der Einzelpixelreferenz. Allerdings ist die Gesamtgenauigkeit, die durch den Vergleich mit der Einzelpixelreferenz maximal erreichbar, ist auch auf 91,2% begrenzt. Das sich dieser Wert nicht ändert liegt daran, dass für jede Anzahl von Merkmalen für Random Forest mit genau den gleichen Parametern für den SLIC-Algorithmus gearbeitet wurde und somit immer die gleichen Superpixel entstanden sind.

Mit beiden Referenzen stagniert die Kurve ab einer Anzahl von vier bis fünf Merkmalen. Danach sind zusätzliche vier Merkmale notwendig, um die Gesamtgenauigkeit um 1% zu erhöhen. Für die Qualitäten der einzelnen Klassen verhalten sich die Kurven genauso, wie für die Gesamtgenauigkeit. Auch hier stagnieren die Qualitäten ab einer Anzahl von vier bis fünf Merkmalen. Für weitere Klassifikationen mit dem Datensatz aus Vaihingen würden aus den genannten Gründen die Merkmale Maximum der Sättigung, Mittelwert des ersten Farbkanals, Mittelwert des NDOM, Mittelwert des zweiten Farbkanals und Minimum des NDVI ausgewählt werden. In Tabelle 10 sind die Genauigkeitsmaße, die mit diesen fünf Merkmalen herauskommen zusammengefasst.

Klasse	Straße		Gebäude		hohe Veg.		niedrige Veg.	
	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.
<b>Vergleich</b>								
<i>Correctness</i> [%]	80,1	81,0	76,7	77,9	66,8	71,0	73,6	78,6
<i>Completeness</i> [%]	74,8	76,7	79,2	80,6	68,2	71,1	81,8	85,4
<i>Quality</i> [%]	62,9	64,9	64,0	65,9	50,4	54,5	62,2	68,3
<b>Gesamtgenauigkeiten</b> [%]	Einzelpixelreferenz				Superpixelreferenz			
	76,6				79,0			

Tabelle 10: Genauigkeitsmaße für Vaihingen mit 5 Merkmalen für Random Forest.

Die erreichten Qualitäten der Klassen Straße, Gebäude und hohe Vegetation liegen mit 62 bis 66% sehr nah beieinander. Nur die Klasse der niedrigen Vegetation ist mit 50 bzw 54% ein Ausreißer. Hier ist ein Blick auf die Konfusionsmatrizen, die bei dem Vergleich mit den Referenzbildern herauskommen, sinnvoll. Dort erkennt man, dass ein vergleichsweise hoher Anteil (4%) der hohen Vegetation im Referenzbild bei der Klassifikation der niedrigen Vegetation zugewiesen wurde. Andersherum (in der Referenz niedrige Vegetation, die der hohen Vegetation zugewiesen wurde) ist die Fehlklassifikation mit knapp 3% nicht ganz so hoch. Hinzu kommt, dass bei der niedrigen Vegetation der prozentuale Gesamtanteil nicht so hoch ist wie der, der anderen Klassen. Deshalb kommt es schon aufgrund der 4 bzw. 3% Fehlklassifikation zu einer viel geringeren Qualität.

Die Gesamtgenauigkeiten sind mit 76% für die Einzelpixelreferenz und 79% für die Superpixel-

referenz schon sehr hoch.

#### 4.2.2 Saudi-Arabien

Für den Datensatz aus Saudi-Arabien wird ebenfalls ein Ranking erstellt. Die 7 einflussreichsten Merkmale sind in Tabelle 11 zu sehen. Die Untersuchung erfolgt in der dargestellten Reihenfolge genau so wie für den Datensatz aus Vaihingen. Für die Erstellung der Superpixel werden die Parameter aus Kapitel 4.1 für Saudi-Arabien gewählt. Nach jedem Training und anschließender Klassifikation erfolgt wieder der Vergleich mit der Einzel- und Superpixelreferenz. Die Ergebnisse der Gesamtgenauigkeiten sind in Abb.17 graphisch dargestellt.

Nummer	Merkmal	Gewichtung in [%]
1	Standardabweichung der Sättigung	9,0
2	Mittelwert von Farbkanal 1	10,0
3	Maximum von Farbkanal 2	6,8
4	Mittelwert der Intensität	4,5
5	Mittelwert von Farbkanal 3	3,7
6	Mittelwert der Sättigung	3,5
7	Standardabweichung von Farbkanal 2	2,9

Tabelle 11: Ranking und Einfluss auf das Training der Merkmale für Saudi-Arabien

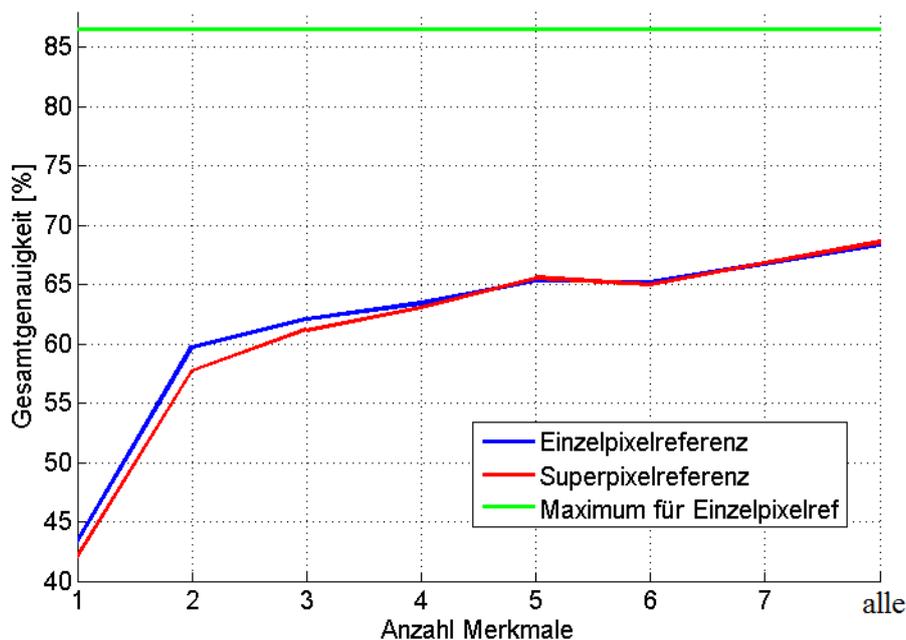
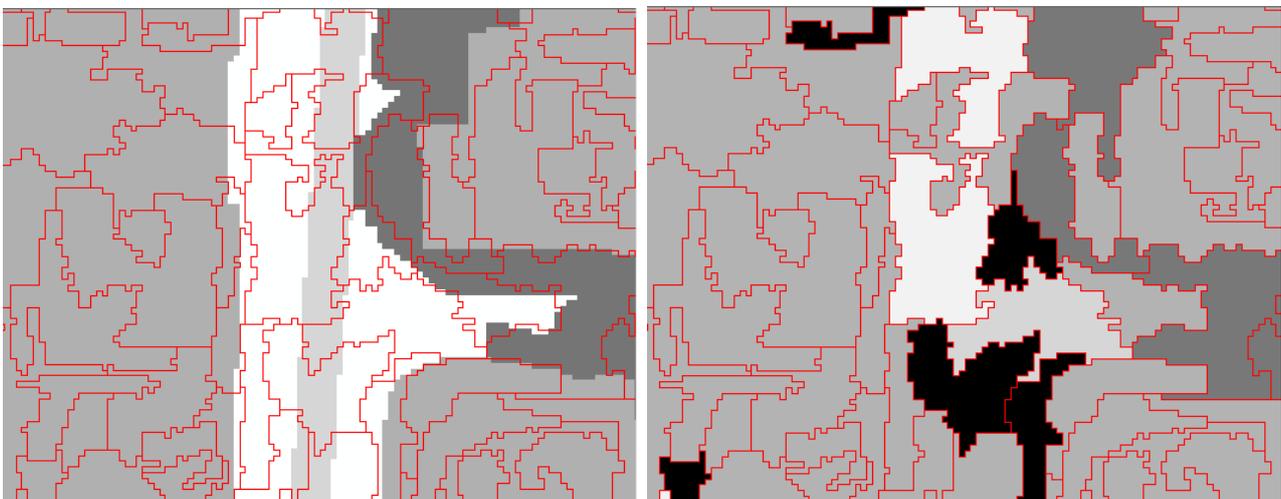


Abbildung 17: Gesamtgenauigkeiten für Saudi-Arabien in Abhängigkeit von der gewählten Anzahl an Merkmalen mit Random Forest

In Abb.17 ist eine Stagnation ab etwa dem fünften Merkmal zu erkennen. Die Gesamtgenauigkeit ist mit fünf Merkmalen nur noch 3% schlechter, als wenn alle Merkmale genutzt werden. Die Ergebnisse der Vergleiche mit der Einzel- und Superpixelreferenz unterscheiden sich kaum. Die maximal erreichbare Gesamtgenauigkeit für den Vergleich mit der Einzelpixelreferenz liegt bei 86,5% und ist damit um etwa 5% schlechter als für den Datensatz aus Vaihingen. Das ist darauf zurückzuführen, dass hier die Superpixel nicht gut an Klassengrenzen in der Einzelpixelreferenz passen. In Abb.18 sind die erstellten Superpixel für eine Anzahl von fünf Merkmalen zu sehen. Hier fällt sehr stark auf, dass viele Superpixel die Klassen Asphalt (weiß) und Straße (hellgrau) nicht trennen und beide Klassen in ihnen vertreten sind. Die anderen Klassen lassen sich im Vergleich noch sehr gut antrennen. Die Klassen Asphalt und Straße haben sehr ähnliche Merkmalsvektoren und können vom Klassifikator deshalb nur sehr schwer unterschieden werden. Deshalb kommt es zu den Superpixeln, die sich kaum an die Grenzen zwischen diesen beiden Klassen halten. In Abb18-(b) sieht man, dass in der Klassifikation sehr wenig Superpixel die Klassen Asphalt oder Straße angenommen haben. Helfen würde hier entweder, die Klassen Asphalt und Straße zusammen zu legen oder anhand von Kontextinformationen (z.B. das Straßen etwa 10 Meter breit sind) die Klassen eindeutiger voneinander zu trennen.



(a) Superpixel vor dem Referenzbild

(b) Ergebnis der Klassifikation der einzelnen Superpixel

Abbildung 18: Ergebnisse der Superpixelerstellung und der Klassifikation für Saudi-Arabien mit 5 Merkmalen

In Abb.18 ist außerdem zu sehen, dass zwischen der Klasse Gebäude und den Klassen Asphalt und Straße viele Fehlklassifikationen vorkommen. Auch hier lässt sich wieder die Konfusionsmatrix hinzuziehen. Nur etwa 25% der Pixel, die in der Referenz zur Klasse Asphalt gehören, werden auch korrekt als diese klassifiziert. Um die 39% der Asphalt-Pixel werden als Gebäude und 29% als Straßen klassifiziert. Andersherum verhält es sich nicht besser: Es wurden fast genauso viele Straßenpixel dem Asphalt zugewiesen wie Asphalt korrekt der Klasse Asphalt zugewiesen

wurde. Das Problem hier ist, dass im Höhenmodell (siehe Abb.6 ) sehr viele Lücken auftreten und diese vor allem auf Gebäuden zu finden sind. Diese entstehen durch Matching-Fehler, wodurch die Höhen nicht ermittelt werden können. Das ist auch der Grund, weshalb der Wert des NDSM im Ranking sehr weit unten auftaucht und für das Training nicht in Frage kommt. Für den Klassifikator ist es sehr schwer, ohne die Information der Höhe ein Gebäude (meist mit Flachdach) von der Klasse Asphalt oder Straße zu unterscheiden. Aus diesen Gründen kommt es zu den teilweise sehr schlechten Genauigkeitsmaßen, die für die ausgewählte Variante mit 5 Merkmalen in Tabelle 12 zusammengefasst sind.

Eine erneute Erstellung des Höhenmodells und mehr Trainingsdaten, sowie eine Zusammenlegung der Klassen Straße und Asphalt könnten für die Saudi-Arabien Daten die Genauigkeiten um einiges verbessern.

Klasse	Asphalt		hohe Veg.		Gebäude		Straßen		Boden	
	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.
<b>Vergleich</b>										
<i>Correctness</i> [%]	25,4	29,5	41,3	38,8	91,1	89,7	68,7	72,5	52,9	51,6
<i>Completeness</i> [%]	37,4	40,7	60,6	59,0	70,3	71,6	69,4	69,6	74,4	71,9
<i>Quality</i> [%]	17,8	20,5	32,4	30,3	65,8	66,2	52,6	54,9	44,7	42,9
<b>Gesamtgen. [%]</b>	Einzelpixelreferenz					Superpixelreferenz				
	67,9					69,2				

Tabelle 12: Genauigkeitsmaße der Klassifikation für Saudi-Arabien mit fünf Merkmalen für Random Forest

## 4.3 Logistische Regression

### 4.3.1 Vaihingen

Die optimale Reihenfolge der Merkmale für das Training mit der logistischen Regression ist in Tabelle 13 dargestellt. Die Ergebnisse der Gesamtgenauigkeiten für die Vergleiche mit der Einzel- und Superpixelreferenz sind in Abb.19 graphisch dargestellt. Für den SLIC-Algorithmus wurden die Parameter genutzt, die in Kapitel 4.1 für Vaihingen ausgewählt wurden.

Anz. Parameter	Parameter
1	Maximum der Sättigung
2	Mittelwert des NDOM
3	Mittelwert des NDVI
4	Mittelwert von Farbkanal 1
5	Gradhist-mean
6	Energieverteilung
7	Mittelwert des zweiten Farbkanals
8	Minimum des NDVI
9	Standardabweichung von Farbkanal 2
10	Minimum des NDSM

Tabelle 13: Einfügereihenfolge der Merkmale für die logistische Regression für Vaihingen

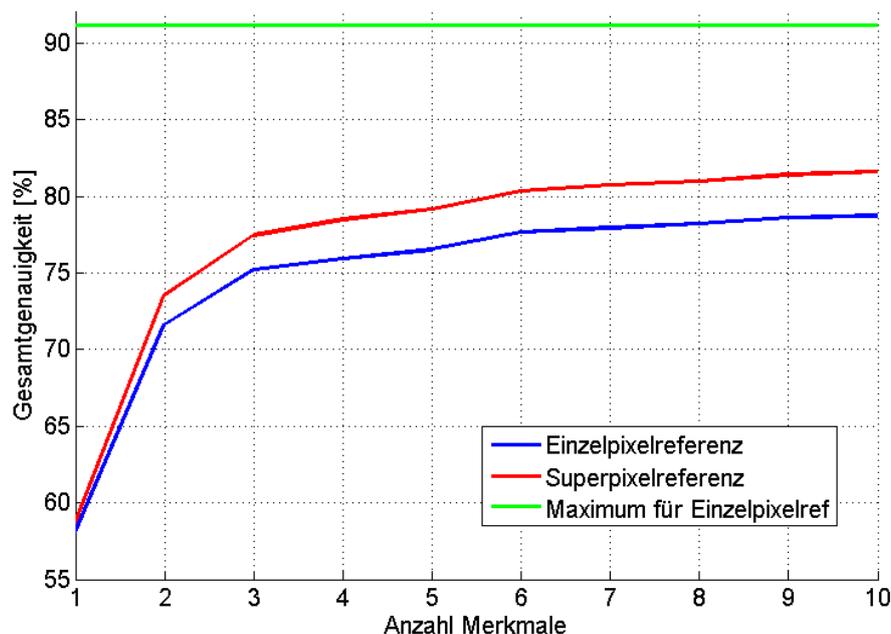


Abbildung 19: Gesamtgenauigkeiten für Vaihingen, abhängig von der gewählten Anzahl an Merkmalen mit der logistischen Regression

Wie bei der Untersuchung von Random Forest steigen die Kurven zunächst stark an, wenn ein Merkmal hinzukommt. Der Sättigungseffekt tritt ab etwa 6 Merkmalen auf. Anschließend steigt die Gesamtgenauigkeit, wenn weitere vier Merkmale hinzukommen, nur noch um 1%. Dieses Phänomen ist, genauso wie bei Random Forest, wieder der höheren Dimension des Merkmalsvektors geschuldet. Für jedes Merkmal stehen bei einer höheren Anzahl an Merkmalen weniger Trainingsprimitive zur Verfügung.

Ausgewählt werden die sechs Merkmale Maximum der Sättigung, Mittelwert des NDOM, Mittelwert des NDVI, Mittelwert des ersten Farbkanal, Mittelwert des Gradientenhistogramms sowie die Energieverteilung. Die Genauigkeitsmaße des Trainings und der Klassifikation mit diesen sechs Merkmalen sind in Tabelle 14 zusammengefasst.

Die Kurve der Gesamtgenauigkeit des Vergleichs mit der Einzelpixelreferenz hat ein Maximum bei 78,7% und die der Superpixelreferenz bei 81,6%. Die Differenz liegt also bei knapp 3%. Die maximal erreichbare Gesamtgenauigkeit für den Vergleich mit der Einzelpixelreferenz liegt bei 91,2%, womit das Ergebnis anhand der Einzelpixelreferenz im Vergleich sehr gut ist.

Klasse	Straße		Gebäude		hohe Veg.		niedrige Veg.	
	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.
<b>Vergleich</b>								
<i>Correctness</i> [%]	81,6	83,0	76,5	77,1	69,7	75,4	74,7	79,6
<i>Completeness</i> [%]	76,3	78,0	82,7	85,2	67,2	69,9	82,7	86,8
<i>Quality</i> [%]	64,7	66,9	65,9	68,2	51,7	56,5	63,3	69,6
<b>Gesamtgenauigkeiten</b> [%]	Einzelpixelreferenz				Superpixelreferenz			
	77,7				80,3			

Tabelle 14: Genauigkeitsmaße für Vaihingen mit 6 Merkmalen für die logistische Regression

Die Qualitäten der Klassen Straße, Gebäude und der niedrigen Vegetation liegen (wie auch bei Random Forest) mit 63 bis 66% für den Vergleich mit der Einzelpixelreferenz, und 67 bis 70% für den Vergleich mit der Superpixelreferenz sehr nah beieinander. Die Qualität der Klasse „hohe Vegetation“ beträgt 52 bzw. 57%. Eine Erklärung liefert auch hier die Konfusionsmatrix. Ganze 19% der als hohe Vegetation klassifizierten Superpixel gehören eigentlich zur Klasse der niedrigen Vegetation. Hinzu kommt, dass die Klasse der hohen Vegetation die wenigsten Trainings- sowie Klassifikationsprimitive besitzt und sich so die einzelne Fehlklassifikation deutlich mehr auf das Gesamtergebnis der Klasse auswirkt. Die Qualitätsmaße liegen bis zu 6% auseinander.

### 4.3.2 Saudi-Arabien

Die optimale Reihenfolge der Merkmale für das Training der logistischen Regression für den datensatz aus Saudi-Arabien ist in Tabelle 15 dargestellt. Der Vergleich der Klassifikationsergebnisse erfolgt wieder anhand der Einzel- und der Superpixelreferenz. Die Ergebnisse der Gesamtgenauigkeiten sind in Abb. 20 graphisch dargestellt. Für den SLIC-Algorithmus werden die Parameter, die in Kapitel 4.1 für die Daten aus Saudi-Arabien festgelegt wurden, genutzt.

Anz. Parameter	Parameter
1	Standardabweichung der Sättigung
2	Mittelwert von Farbkanal 1
3	Mittelwert von Farbkanal 3
4	Standardabweichung von Farbkanal 2
5	Mittelwert der Intensität
6	Mittelwert der Sättigung
7	Maximum von Farbkanal 2

Tabelle 15: Einfügereihenfolge der Merkmale für die logistische Regression mit dem zweiten Datensatz

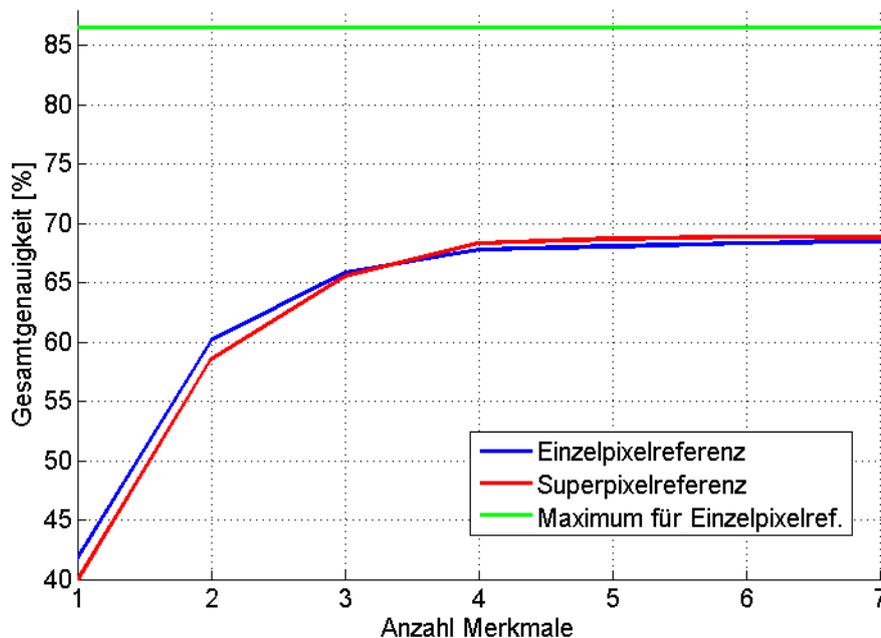


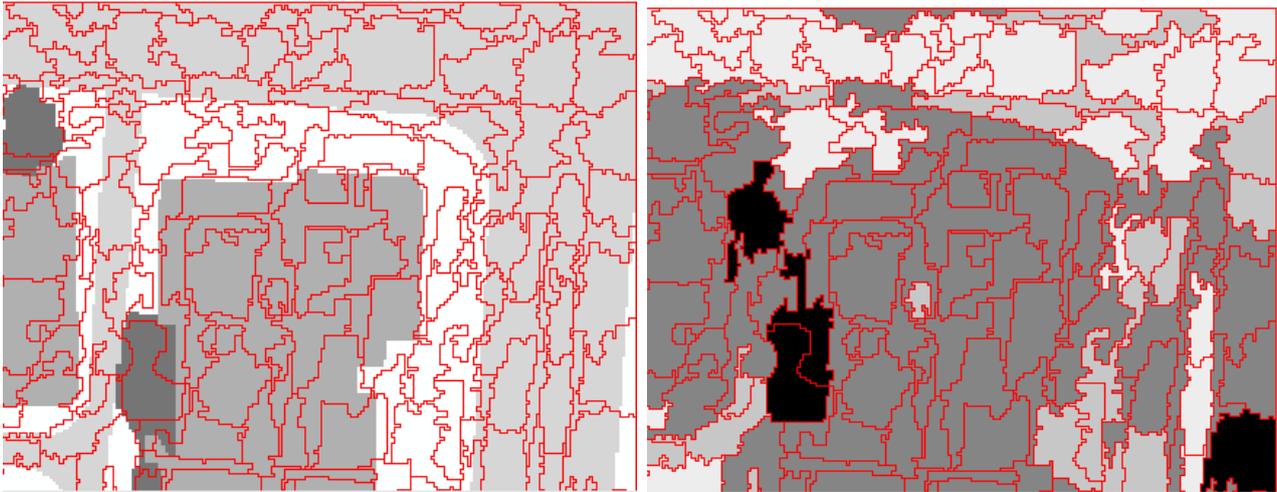
Abbildung 20: Gesamtgenauigkeiten für die logistische Regression, abhängig von der Anzahl der verwendeten Merkmale für Saudi-Arabien

Eine Stagnation ist in Abb.20 etwa ab dem vierten bis fünften Merkmal zu erkennen. Die Gesamtgenauigkeit liegt für fünf Merkmale bei 68 bzw. 69% (je nach Referenz). Anschließend steigt die Gesamtgenauigkeit um weniger als ein Prozent für die nächsten beiden Merkmale.

Wie für Random Forest gibt es zwischen den Ergebnissen anhand der Einzel- und Superpixelreferenz kaum Differenzen. Auch die Qualitäten der einzelnen Klassen sehen ähnlich aus wie die der Klassifikation mit Random Forest (Vgl Tabelle 16). Die Klasse Asphalt weist wieder mit Abstand die schlechtesten Werte für die Qualität auf. Zwar können die Grenzen zwischen Asphalt und Straße durch die Superpixel teilweise eingehalten werden (siehe Abb.21-(a)), aber bei der Klassifikation werden die Klassen Gebäude (dunkelgrau), Straße (hellgrau) und Asphalt

(sehr hell) sehr oft vertauscht und nicht korrekt klassifiziert. Dieses Problem wurde in Kapitel 4.2.2 schon erläutert und hat hier denselben Grund: Die sehr ähnlichen Merkmalsvektoren der Klassen Straße und Asphalt sind für den Klassifikator schwer zu unterscheiden. Hinzu kommen die Lücken im Höhenmodell, wodurch sich auch Gebäude von den Klassen Asphalt und Straße nur noch schwer unterscheiden lassen.

Zur Verbesserung würde auch hier ein besseres Höhenmodell helfen. Eine andere, einfache Methode wäre die Klassen Straße und Asphalt vorerst zusammen zu legen, wodurch die Fehlklassifikationen zwischen diesen Klassen herausfallen würden.



(a) Superpixel vor dem Referenzbild

(b) Ergebnis der Klassifikation der einzelnen Superpixel

Abbildung 21: Ergebnisse der Superpixelerstellung und der Klassifikation für Saudi-Arabien mit 5 Merkmalen mit der logistischen Regression

Für die ausgewählte Variante mit fünf Merkmalen sind die Genauigkeitsmaße in Tabelle 16 zusammengefasst.

Klasse	Asphalt		hohe Veg.		Gebäude		Straßen		Boden	
	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.	Einz.	Segm.
<b>Vergleich</b>										
<i>Correctness</i> [%]	39,2	44,3	54,5	58,8	87,4	86,2	69,8	71,1	55,8	55,2
<i>Completeness</i> [%]	43,8	46,3	61,6	65,6	77,8	79,7	70,3	71,3	70,5	67,0
<i>Quality</i> [%]	25,9	28,9	40,5	44,7	70,0	70,7	53,5	54,9	45,2	43,4
<b>Gesamtgen.</b> [%]	Einzelpixelreferenz					Superpixelreferenz				
	68,1					68,7				

Tabelle 16: Genauigkeitsmaße der Klassifikation für Saudi-Arabien mit fünf Merkmalen für die logistische Regression

#### 4.4 Vergleich Random Forest - Logistische Regression

Der Vergleich zwischen Random Forest und der Logistischen Regression erfolgt anhand des ersten Datensatzes, da hier mehr Daten zur Verfügung stehen und die Ergebnisse so aussagekräftiger sind. Verglichen werden jeweils die Ergebnisse aus Kapitel 4.2 und 4.3, da so für den jeweiligen Klassifikator die optimale Reihenfolge der Merkmale vorliegt.

In Abbildung 22 ist der Vergleich anhand der Einzelpixelreferenz dargestellt. Die Differenz zur maximal erreichbaren Gesamtgenauigkeit von 91,2% beträgt für beide Klassifikatoren 13-14%. Random Forest erreicht eine maximale Gesamtgenauigkeit von 77,1% und die logistische Regression von 78,7%.

In Abb.23 sind die Ergebnisse der Vergleiche mit der Superpixelreferenz dargestellt. Random Forest erreicht eine maximale Gesamtgenauigkeit von 79,9% und die logistische Regression von 81,6%.

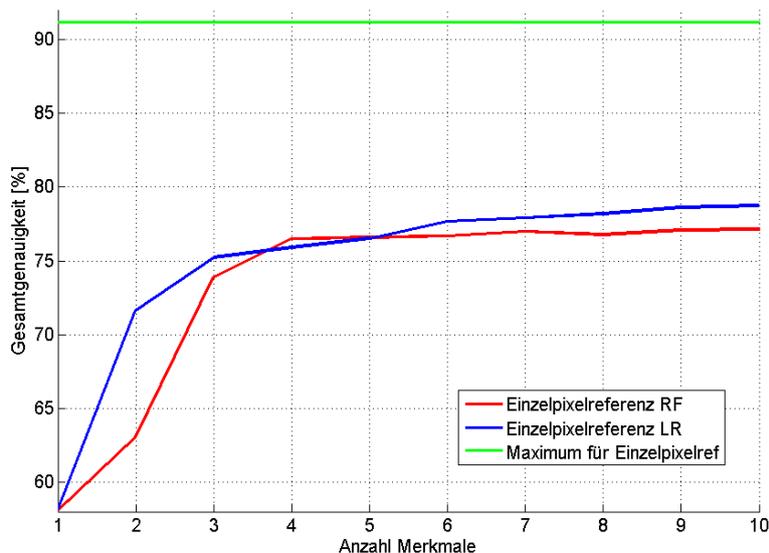


Abbildung 22: Gesamtgenauigkeiten für die logistische Regression und Random Forest abhängig von der Anzahl der verwendeten Merkmale für Vaihingen beim Vergleich mit der Einzelpixelreferenz

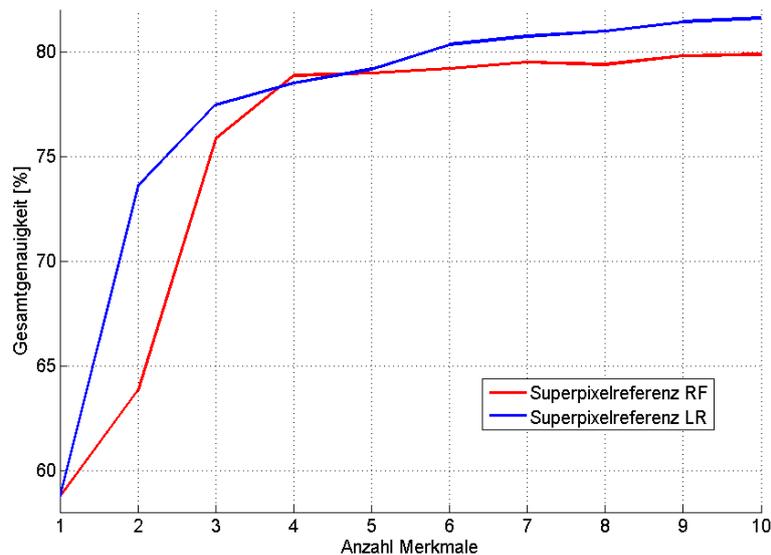


Abbildung 23: Gesamtgenauigkeiten für die logistische Regression und Random Forest abhängig von der Anzahl der verwendeten Merkmale für Vaihingen beim Vergleich mit der Superpixelreferenz

Insgesamt kommen bei der Klassifikation mit der logistischen Regression um etwa 1,5% bessere Gesamtgenauigkeiten heraus als bei Random Forest. Bei den Qualitäten der einzelnen Klassen sieht das genauso aus. Auch hier schneidet in jedem Fall die logistische Regression knapp besser ab.

Ein Grund dafür kann sein, dass Random Forest zum unabhängigen Training der 100 Entscheidungsbäume mehr Trainingsprimitive benötigt. Anhand der verwendeten acht Bilder kommen sehr häufig dieselben Primitive zum Einsatz und die Bäume bekommen teilweise die gleiche Struktur. Der Grundgedanke des Random Forest ist es aber, viele verschiedene Bäume zu erstellen, die unabhängig voneinander sind. Hier wäre eine Untersuchung mit einem größeren Umfang an Trainingsprimitiven sinnvoll, um herauszufinden ob sich das Klassifikationsergebnis Random Forest noch verbessern kann.

## 4.5 Einzelpixeluntersuchung

### 4.5.1 Random Forest

#### Vaihingen

Für den Datensatz aus Vaihingen wurden für die Superpixelklassifikation mit Random Forest die Merkmale Mittelwert der Sättigung, Mittelwert des ersten Farbkanals, Mittelwert des NDOM, Mittelwert des zweiten Farbkanals sowie das Minimum des NDVI ausgewählt. Für die Klassifikation mit Einzelpixeln besteht natürlich nicht die Möglichkeit, Mittelwerte, Minima und Maxima oder Standardabweichungen zu berechnen. Deshalb werden die Merkmale Sättigung,

erster und zweiter Farbkanal, NDVI sowie NDSM für jedes Pixel genutzt.

Das Training wird mit allen 40.000.000 vorhandenen Primitiven durchgeführt. Ein Programm-durchlauf dauert damit etwa 24 Stunden (Zum Vergleich: Ein Durchlauf mit 40x40 großen Superpixeln für den Datensatz aus Vaihingen benötigt etwa 15 bis 20 Minuten). Die Ergebnisse der Gesamtgenauigkeiten der Einzelpixelklassifikation im Vergleich zur Superpixelklassifikation sind in Abb.24 dargestellt. Die Superpixelklassifikation wurde wieder mit der Einzel- und Superpixelreferenz verglichen.

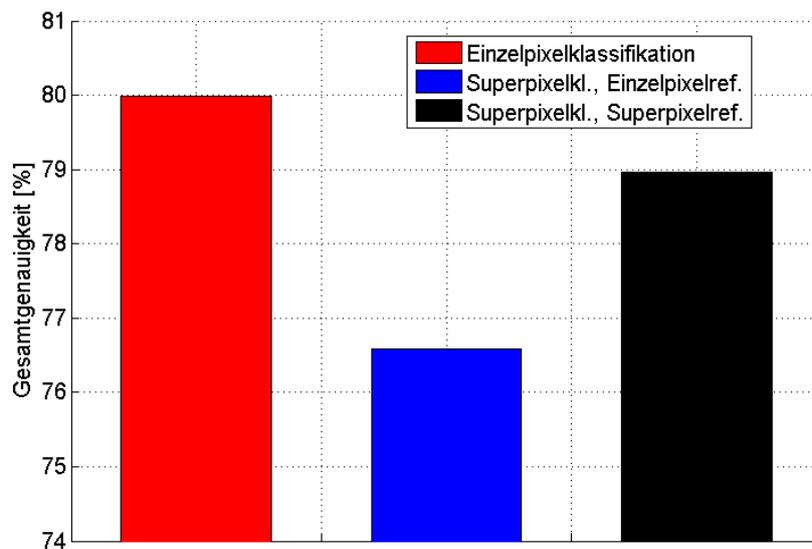


Abbildung 24: Gesamtgenauigkeiten der Einzel- und Superpixelklassifikation für Vaihingen mit Random Forest

Die Gesamtgenauigkeit der Einzelpixelklassifikation liegt bei 80%, die der Superpixelklassifikation mit Vergleich durch die Einzelpixelreferenz bei 76,6% und die mit Vergleich durch die Superpixelreferenz bei 79,0%. Die segmentbasierte Klassifikation ist also etwas schlechter als die pixelbasierte, allerdings ist der Unterschied mit 1% bzw. 3,4% gering.

Die Qualitäten der einzelnen Klassen ergeben die gleichen Ergebnisse. Eine Ausnahme bildet nur die Klasse der niedrigen Vegetation. Hier liegt die Qualität der Einzelpixelklassifikation mit 61,2% etwa ein Prozent unter der Qualität der Superpixelklassifikation mit Einzelpixelreferenz mit 62,2%. Der Vergleich anhand der Superpixelreferenz ist mit 68,3% deutlich besser.

Die Erklärung liefert ein Blick auf das Klassifikationsergebnis, das für einen Bildausschnitt in Abb.26 dargestellt ist. Rechts ist die Einzelpixelreferenz zu sehen. In der Mitte ist das Ergebnis der pixelweisen Klassifikation zu sehen. Man erkennt deutlich, dass der Klassifikator Schwierigkeiten hat, die Klassen der niedrigen und hohen Vegetation voneinander zu unterscheiden. Das kann an einem schlechten Höhenmodell in den Bereichen der Vegetation liegen (Bei der Klasse Gebäude tritt diesen Problem nicht auf). Auch die segmentbasierte Klassifikation (Abb.26, rechts) hat das gleiche Problem. Dass das Ergebnis der segmentbasierten Klassifikation mit

Superpixelreferenz so viel besser ist, liegt daran, dass schon bei der Erstellung der Superpixelreferenz viele Superpixel die eigentlich zur hohen Vegetation gehören, der niedrigen zugewiesen wurden. So ergibt sich natürlich beim Vergleich die bessere Qualität dieser Klasse.

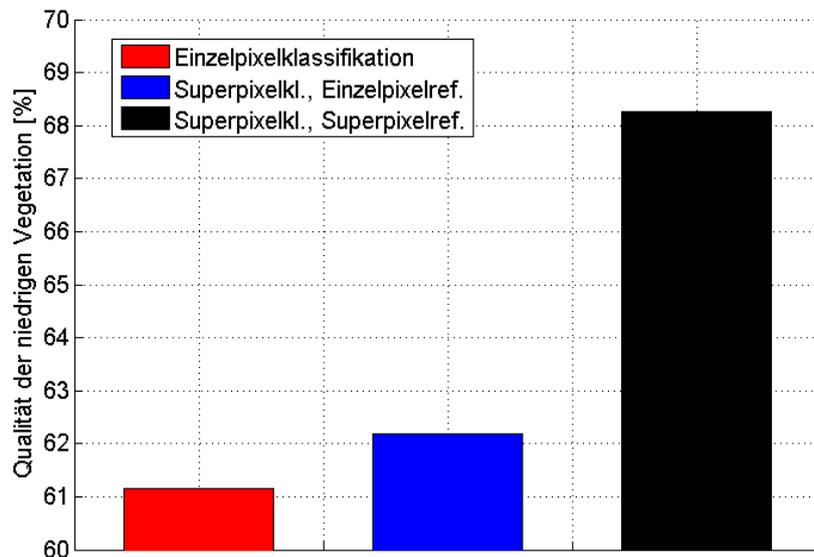


Abbildung 25: Qualität der hohen Vegetation der Einzel- und Superpixelklassifikation für Vaihingen mit Random Forest

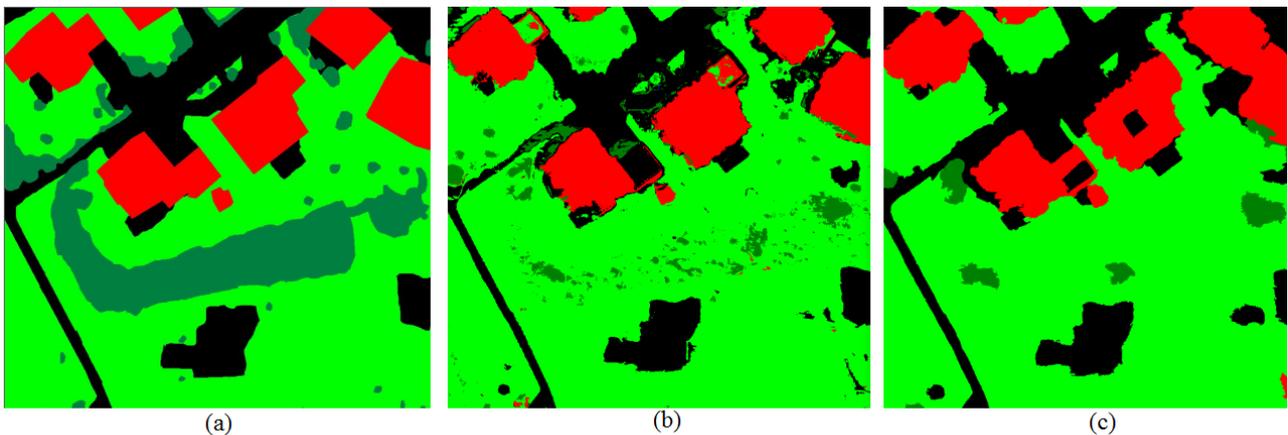


Abbildung 26: Ausschnitt der Einzelpixelreferenz (a), der Einzelpixelklassifikation (b) und der Superpixelklassifikation (c) für Vaihingen

Insgesamt liefert die segmentbasierte Klassifikation mit Random Forest annähernd so gute Ergebnisse wie eine Einzelpixelklassifikation. Der Algorithmus benötigt jedoch für die hier durchgeführte Klassifikation etwa 24 Stunden und für die mit 40x40 großen Superpixeln nur etwa 20 Minuten. Aufgrund dieser deutlich reduzierten Durchlaufzeit des Algorithmus kann für große Datenmengen so sehr viel Zeit gespart werden, während die Qualität annähernd die Gleiche

bleibt. Wichtig ist es, bei der Erstellung des Höhenmodells genau zu arbeiten und auch die Superpixelreferenz zu kontrollieren.

### Saudi-Arabien

Für das Training von Random Forest wurden für die Daten aus Saudi-Arabien die Merkmale Standardabweichung der Sättigung, Mittelwert des ersten und dritten Farbkanals, Maxima des zweiten Farbkanals und der Mittelwert der Intensität gewählt. Für die pixelbasierte Klassifikation werden nun die drei Farbkanäle, die Sättigung sowie die Intensität der Trainingsprimitive ausgewählt. Zum Training werden alle 700.000 Primitive genutzt. Die Ergebnisse der Gesamtgenauigkeiten sind in Abb.27 zu sehen. Zum einen ist die Gesamtgenauigkeit der pixelweisen Klassifikation dargestellt und zum Anderen die der Superpixelklassifikation, einmal mit Einzel- und einmal mit Superpixelreferenz.

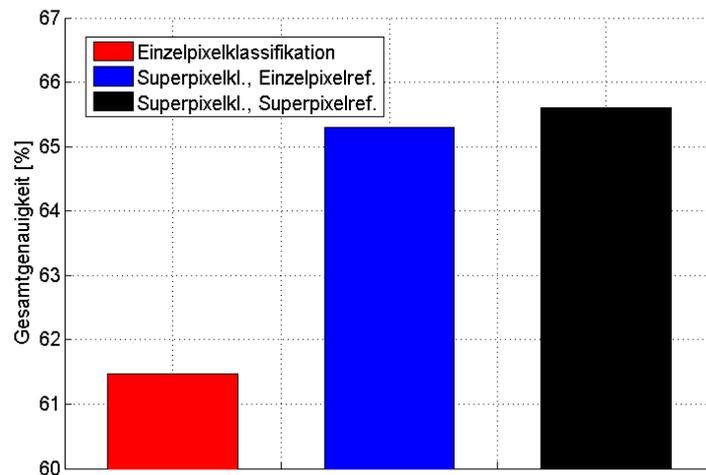


Abbildung 27: Gesamtgenauigkeiten der Einzel- und Superpixelklassifikation für Saudi-Arabien mit Random Forest

Im Gegensatz zum Datensatz aus Vaihingen schneiden hier die superpixelbasierten Klassifikationen besser ab. Die pixelbasierte Klassifikation hat eine Gesamtgenauigkeit von 61,5%. Die segmentbasierte Klassifikation hat für den Vergleich mit der Einzelpixelreferenz eine Gesamtgenauigkeit von 65,3% und für den Vergleich mit der Superpixelreferenz eine Gesamtgenauigkeit von 65,6%. Die Ergebnisse der Superpixelklassifikation liegen sehr nah zusammen und die Differenz zur Einzelpixelklassifikation beträgt etwa 4%.

In Abb.28 ist ein Bildausschnitt der Daten zu sehen. Rechts ist die Referenz mit den Klassen Asphalt (hellgrau), Gebäude (rot), Vegetation (grün), Straße (dunkelgrau) und Boden (orange) zu sehen. In der Mitte ist das Klassifikationsergebnis der pixelbasierten Klassifikation zu sehen und rechts das der Superpixelklassifikation. Man erkennt, dass im pixelbasierten Ergebnis viele Pixel dem Boden zugeordnet wurden, die eigentlich zum Gebäude gehören. Teilweise sind hier Innenhöfe in den Bildern zu sehen, die in der Referenz mit zur Klasse Gebäude zugeordnet

wurden. An dieser Stelle wäre es sinnvoll die Referenz vor einer erneuten Klassifikation zu korrigieren.

Die Klasse Vegetation hingegen hat gute Übereinstimmungen. Im rechten Bild erkennt man, dass es vor allem für die Klassen Asphalt und Straße viele Fehlklassifikationen der Superpixel auftreten. Die Straße zwischen den oberen und unteren Gebäuden im Bildausschnitt ist hier kaum noch zu erkennen. Das erklärt auch, warum die Klasse Asphalt die einzige ist, für die die pixelbasierte Klassifikation bessere Qualitäten liefert als für die Segmentbasierte (siehe Abb.29). Die Klasse Asphalt ist mit 24% auch dort die mit Abstand schlechteste, aber besser als die segmentbasierte Klassifikation mit 18 bzw. 20%. Das Problem der segmentbasierten Klassifikation entsteht schon bei der Erstellung der Superpixel. Diese grenzen die Klassen Asphalt und Straße nicht voneinander ab und so enthalten Superpixel immer Einzelpixel beider Klassen. Bei der anschließenden Klassifikation wird es für den Klassifikator dann sehr schwer bis unmöglich die richtige Klasse zu finden.

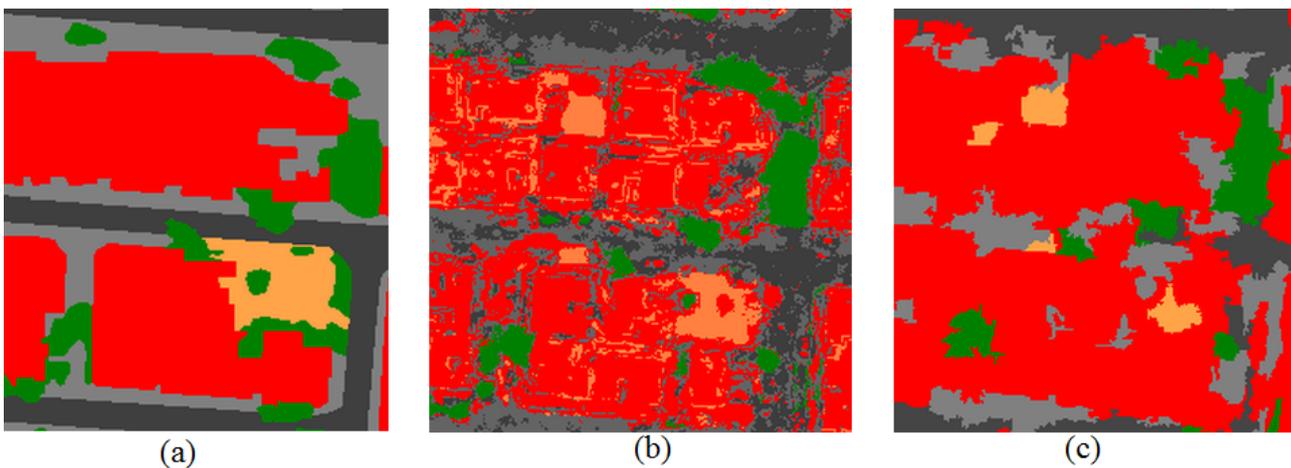


Abbildung 28: Ausschnitt der Einzelpixelreferenz (a), der Einzelpixelklassifikation (b) und der Superpixelklassifikation (c) für Saudi-Arabien

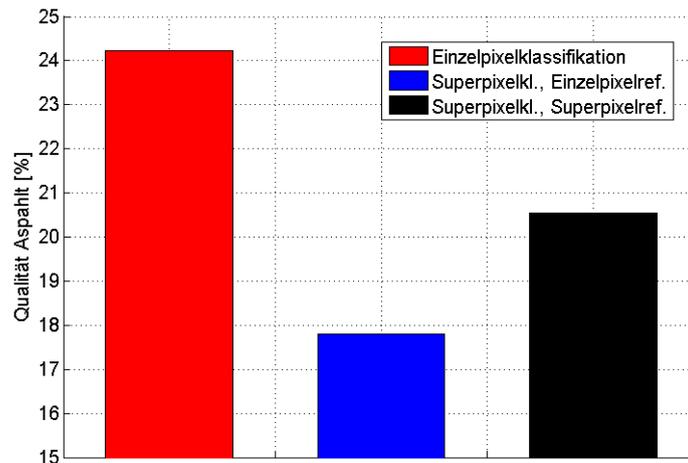


Abbildung 29: Qualität der Klasse Asphalt der Einzel- und Superpixelklassifikation für Saudi-Arabien mit Random Forest

Um mit dem Datensatz aus Saudi-Arabien aussagekräftige Ergebnisse zu erhalten, wäre es sinnvoll, das Höhenmodell zu überarbeiten, einen größeren Umfang an Trainingsdaten zu erstellen und die Einteilung der Klassen (besonders Asphalt und Straße) zu ändern. Die Ergebnisse sind so größtenteils sehr schlecht und die Ergebnisse können nicht auf andere Daten übertragen werden.

#### 4.5.2 Logistische Regression

##### Vaihingen

Für die logistische Regression wurden für den Datensatz aus Vaihingen die Merkmale Maximum der Sättigung, Mittelwert des NDOM, Mittelwert des NDVI, Mittelwert des ersten Farbkanals sowie der Mittelwert des Gradientenhistogramms und die Energieverteilung ausgewählt. Für die pixelbasierte Klassifikation werden die Merkmal Sättigung, NDVI, NDSM und der erste Farbkanal ausgewählt. Zum Training können nur 2,6% der Trainingsprimitive genutzt werden, was einer Anzahl von etwa 1,02 Millionen entspricht.

Die Ergebnisse der Gesamtgenauigkeiten der Einzelpixel- und Superpixelklassifikation sind in Abb.30 graphisch dargestellt. Die Superpixelklassifikation wird dabei wieder mit der Einzel- sowie der Superpixelreferenz verglichen.

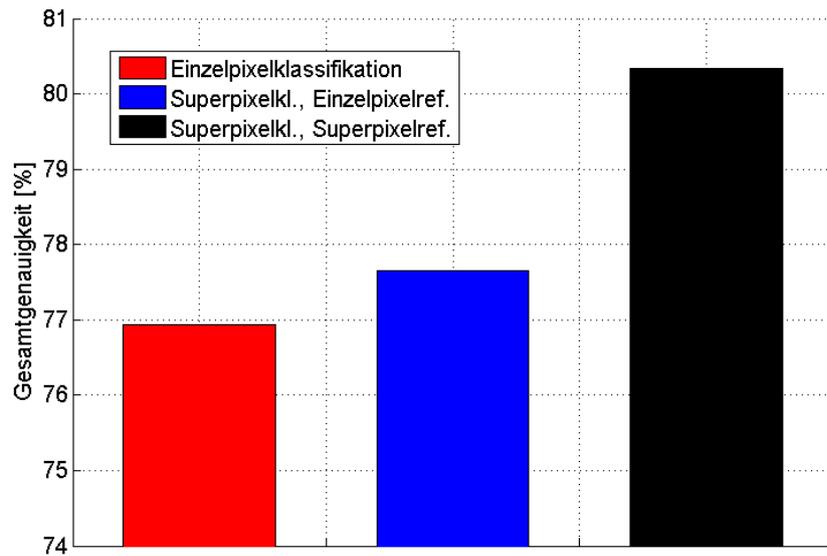


Abbildung 30: Gesamtgenauigkeiten der Einzel- und Superpixelklassifikation für Vaihingen mit der logistischen Regression

Die erreichte Gesamtgenauigkeit der pixelweisen Klassifikation liegt bei 77,0%. Die Gesamtgenauigkeit der Superpixelklassifikation mit Einzelpixelreferenz liegt bei 77,7% und die mit Superpixelreferenz bei 80,3%. Die Ergebnisse der Superpixelklassifikation sind damit knapp besser. Da die pixelbasierte Klassifikation allerdings mit nur zwei Prozent der Trainingsdaten durchgeführt wurde, ist das Ergebnis mehr als zufriedenstellend. Ein endgültiger Vergleich kann so aber leider nicht erfolgen. Vermutlich wird die pixelbasierte Klassifikation sich insgesamt noch verbessern, wenn mehr Trainingsdaten genutzt werden können.

Für den vorliegenden Fall kann die Klassifikation auf jeden Fall mit den Superpixeln erfolgen und es kommen Ergebnisse mit guten Genauigkeiten heraus.

### Saudi-Arabien

Für das Training der logistischen Regression wurden für Saudi-Arabien die Merkmale Standardabweichung der Sättigung, Mittelwert des ersten und dritten Farbkanals, Standardabweichung des zweiten Farbkanals sowie die Intensität gewählt. Für die Einzelpixelklassifikation werden also die Merkmale Sättigung, Intensität und alle drei Farbkanäle ausgewählt. Zum Training können alle 700.000 Trainingsprimitive genutzt werden. Die Ergebnisse der Gesamtgenauigkeiten sind in Abb.31 zu sehen. Der Vergleich zur Superpixelklassifikation erfolgt wieder anhand der Einzel- und Superpixelreferenz.

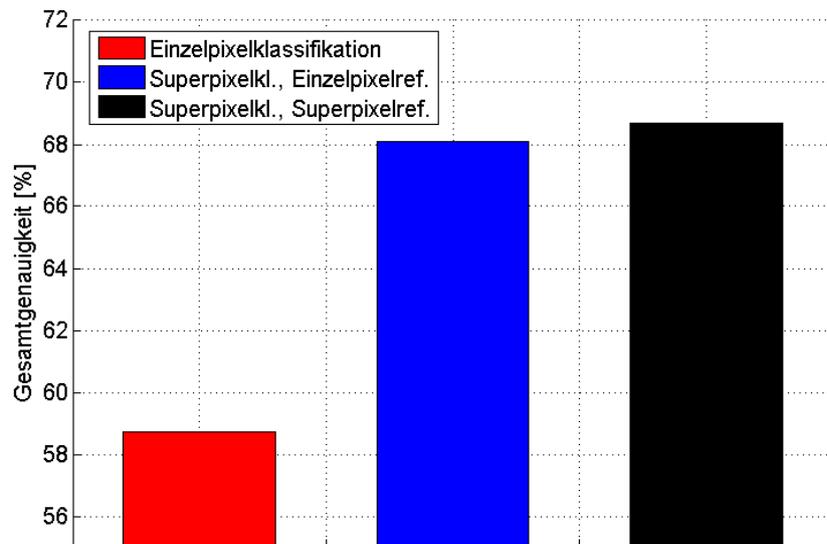


Abbildung 31: Gesamtgenauigkeiten der Einzel- und Superpixelklassifikation für Saudi-Arabien mit der logistischen Regression

Die Gesamtgenauigkeit der pixelbasierten Klassifikation liegt bei 59%. Die Gesamtgenauigkeiten der Superpixelklassifikation liegen mit der Einzelpixelreferenz bei 68% und mit der Superpixelreferenz bei 69%. Die Qualitäten der einzelnen Klassen sind ebenfalls für die Superpixelklassifikation um einiges besser.

Wie bei Random Forest werden hier ebenfalls sehr viele Pixel, die in der Referenz zu Gebäuden gehören, als Boden klassifiziert. Für die Superpixelklassifikation ist das nicht in einem solch großen Umfang der Fall. Außerdem hat die pixelweise Klassifikation, genauso wie die superpixelweise, das Problem der Unterscheidung der Klassen Asphalt und Straße.

Ebenso, wie für Random Forest kommt hier das Ergebnis heraus, dass mit den Daten aus Saudi-Arabien so nicht gut klassifiziert werden kann.

### Zusammenfassung der Ergebnisse

Für den Datensatz aus Vaihingen liegen die Genauigkeiten zwischen der pixelweisen- und der superpixelweisen Klassifikation nur wenige Prozent auseinander. Bei der Gesamtgenauigkeit beträgt die Differenz maximal 3,4%. Bei der Klassifikation durch Random Forest schneidet die pixelbasierte Klassifikation etwas besser ab, und bei der Klassifikation durch die logistische Regression schneidet die Superpixelklassifikation besser ab. Allerdings konnten bei der logistischen Regression, aufgrund der Rechenzeit und des benötigten Speichers, nur 2,6% aller vorhandener Trainingsdaten genutzt werden, wodurch dieser Vergleich nur bedingt aussagekräftig ist.

Der Datensatz aus Saudi-Arabien liefert jeweils für die Superpixelklassifikation die besseren

Ergebnisse. Diese sind für die Gesamtgenauigkeit bis zu 10% besser, als die der pixelweisen Klassifikation. Für diesen Datensatz entstehen viele Fehlklassifikationen zwischen den Klassen Straße, Asphalt und Gebäude. Zum einen liegt das an den ähnlichen Eigenschaften der Klassen Asphalt und Straße, aber auch an den vielen Lücken im Höhenmodell, die vor allem auf Dachflächen auftreten. Aus diesen Gründen sind die Ergebnisse hier wenig aussagekräftig und die Untersuchung sollte mit überarbeiteten Daten wiederholt werden.

## 5 Fazit und Ausblick

Insgesamt liefert die Superpixelklassifikation Ergebnisse, die von den Genauigkeiten nur wenige Prozent von denen der Einzelpixelklassifikation abweichen. Dabei wird die Rechenzeit des Trainings und der Klassifikation allerdings drastisch reduziert.

Schwierigkeiten bei der Superpixelklassifikation treten für Klassen auf, die prozentual sehr wenig in den vorhandenen Trainingsdaten vorkommen und so nicht ausreichend trainiert werden können. Auch Objekte, die kleiner sind als ein Superpixel werden in der Superpixelklassifikation automatisch einer umliegenden Klasse zugeordnet. Aus diesem Grund wurde für den Datensatz aus Vaihingen schon zu Beginn die Klasse „Autos“ zur Klasse „Straße“ zugeordnet.

Die Parameter für den Superpixelalgorithmus müssen zu Beginn für jeden Datensatz festgelegt werden, um die besten Ergebnisse zu erhalten.

Können geringe Genauigkeitsverluste der Klassifikationsergebnisse verkraftet werden und sind einzelne Details auf der Oberfläche für das Gesamtergebnis von untergeordneter Bedeutung, kann sehr gut mit Superpixeln gearbeitet werden.

Besonders für große Datensätze ermöglicht das Arbeiten mit Superpixeln eine deutlich schnellere Klassifikation der Geländeoberflächen. Voraussetzung für qualitativ gute Ergebnisse ist eine gute Vorbereitung. Dazu gehört ein ausreichend großer Umfang an Trainingsdaten für das Gebiet, ein Höhenmodell ohne Lücken oder andere Matching-Fehler und eine gute Auswahl an Parametern für den Superpixel-Algorithmus und auch an Merkmalen für die Klassifikatoren. Sind diese Voraussetzungen erfüllt, kann anhand der Superpixelmethode ein schnelles und qualitativ gutes Klassifikationsergebnis erwartet werden.

Um Änderungen in der Landbedeckung schnell und ohne eine Begehung feststellen zu können, eignet sich die Klassifizierung mit Superpixeln auf jeden Fall. In diesem Fall geht es um großflächige Änderungen, die auch mit Superpixeln gut erkannt werden können. Hier sollen Änderungen wie neue Häuser, Ackerflächen oder Straßen festgestellt werden. Somit würde eine einfache Methode genügen, die die klassifizierten Bilder aus zwei Epochen miteinander vergleicht und eine Veränderung angibt, wenn viele Superpixel nebeneinander eine Änderung der Landbedeckung feststellen. Geht man von einer erreichbaren Gesamtgenauigkeit von 80% für

die Superpixelklassifikation aus, so werden Änderungen für jedes Superpixel mit einer Genauigkeit von 64% detektiert. Für ein einzelnes Superpixel ist diese Genauigkeit nicht ausreichend, wenn jedoch mehrere benachbarte Superpixel eine Änderung anzeigen, kann davon ausgegangen werden, dass dieses Ergebnis stimmt.

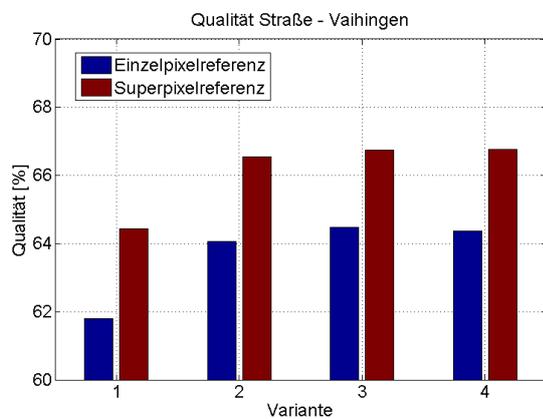
Durch diese Methode können viel Zeit und Kosten gespart werden. Die Datenerfassung kann mithilfe von Satelliten in regelmäßigen Abständen erfolgen. Die anschließende superpixelbasierte Klassifikation und der Vergleich mit der vorherigen Epoche kann automatisiert erfolgen. Zu Beginn jeder Epoche müssen ausreichend viele Trainingsdaten erstellt werden, damit der Klassifikator gut trainiert werden kann und die Genauigkeiten der Klassifikationen möglichst gut werden.

## Literatur

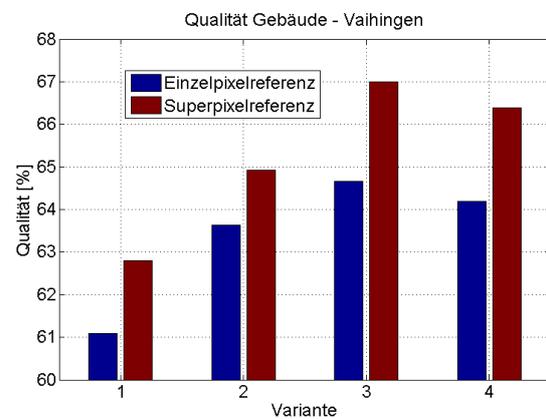
- [Achanta et al., 2010] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2010). Slic superpixels. Technical report, Ecole Polytechnique Fedrale de Lausanne (EPFL).
- [Albertz and Wiggenhagen, 2009] Albertz, J. and Wiggenhagen, M. (2009). *Taschenbuch zur Photogrammetrie und Fernerkundung*, volume 5. Wichmann.
- [Alobeid et al., 2011] Alobeid, A., Jacobsen, K., Heipke, C., and Alrajhi, M. (2011). Building monitoring with differential dsms. *ISPRS Hannover Workshop*, 67.
- [Bach et al., 2013] Bach, E., Fröhlich, B., Walde, I., Hese, S., Schmullius, C., and Denzler, J. (2013). Landbedeckungsklassifikation von satellitenaufnahmen unter verwendung von kontextinformation mit minimaler menschlicher interaktion. *DGPF (Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.) Tagungsband*, 22.
- [Bishop, 2006] Bishop, C. (2006). *Pattern Recognition and Machine Learning*, volume 11. Springer.
- [Breiman, 2001] Breiman, L. (2001). Random forest. *Machine Learning*, 45.
- [Haala and Brenner, 1999] Haala, N. and Brenner, C. (1999). Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54:130–137.
- [Haralick, 1979] Haralick, R. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67.
- [Heipke, 2016] Heipke, C. (2016). Bildanalyse 2 - nichtprobabilistische diskriminative klassifikation.
- [Rodehorst, 1997] Rodehorst, V. (1997). Digitale bildanalyse in der architekturphotogrammetrie. Technical report, Deutsche Forschungsgemeinschaft.
- [Rottensteiner, 2015] Rottensteiner, F. (2015). *Handbuch der Geodäsie - Photogrammetrie und Fernerkundung*, chapter Kontextbasierte Ansätze in der Bildanalyse.
- [Rutzinger et al., 2009] Rutzinger, M., Rottensteiner, F., and 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.
- [Sohn et al., 2013] Sohn, G., Gerke, M., Wegner, J., and Rottensteiner, F. (2013). Isprs test project on urban classification and 3d building reconstruction. Technical report, ISPRS - International Society for Photogrammetry and Remote Sensing.

- [Stoffels, 2009] Stoffels, J. (2009). Einsatz einer lokal adaptiven klassifikationsstrategie zur satellitengestützten waldinventur in einem heterogenen mittelgebirgsraum.
- [Weidner, 1997] Weidner, U. (1997). *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, chapter Digital Surface Models for Building Extraction. Birkhäuser Verlag Basel.

## 6 Anhang

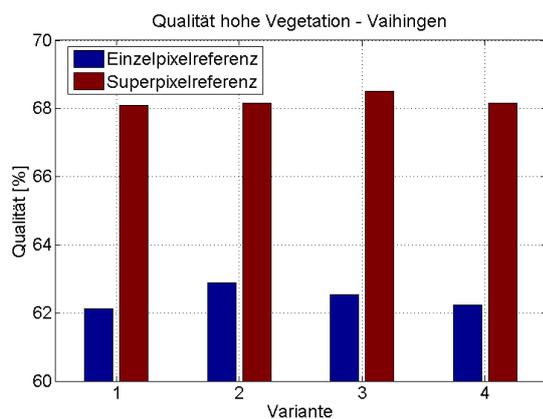


(a)

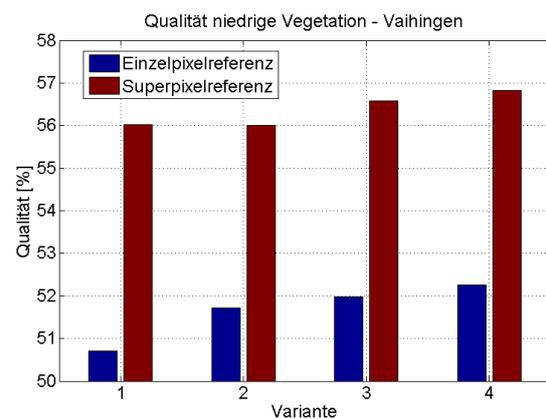


(b)

Abbildung 32: Qualitäten der Varianten für den SLIC-Algorithmus mit dem ersten Datensatz und Random Forest

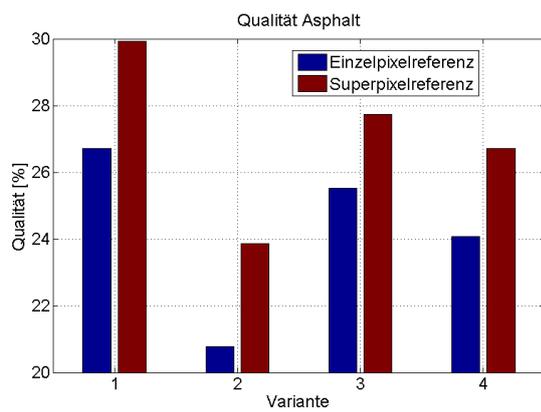


(a)

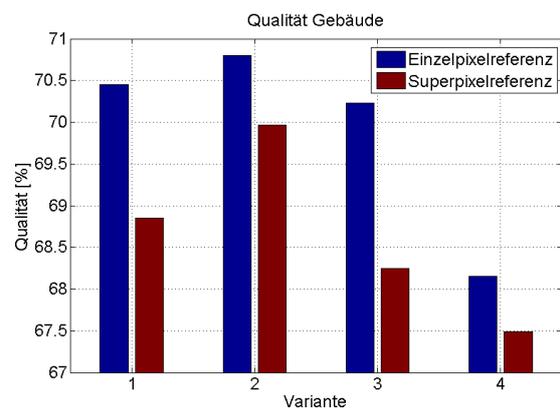


(b)

Abbildung 33: Qualitäten der Varianten für den SLIC-Algorithmus mit dem ersten Datensatz und Random Forest

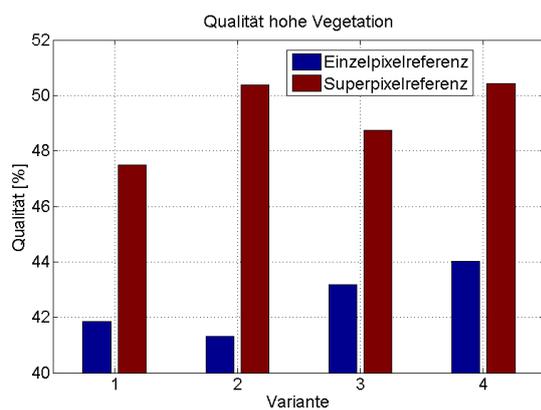


(a)

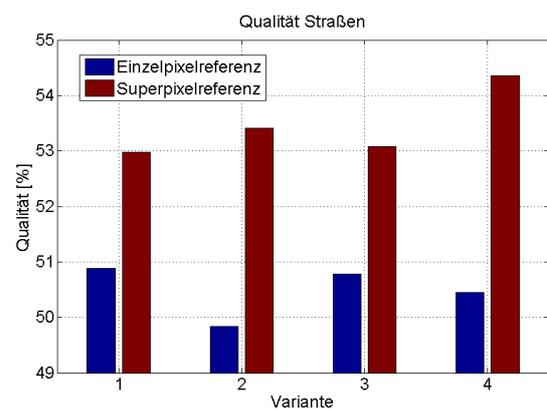


(b)

Abbildung 34: Qualitäten der Varianten für den SLIC-Algorithmus mit dem zweiten Datensatz und Random Forest



(a)



(b)

Abbildung 35: Qualitäten der Varianten für den SLIC-Algorithmus mit dem zweiten Datensatz und Random Forest

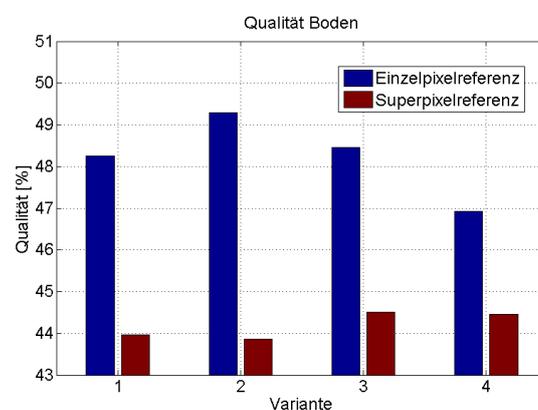
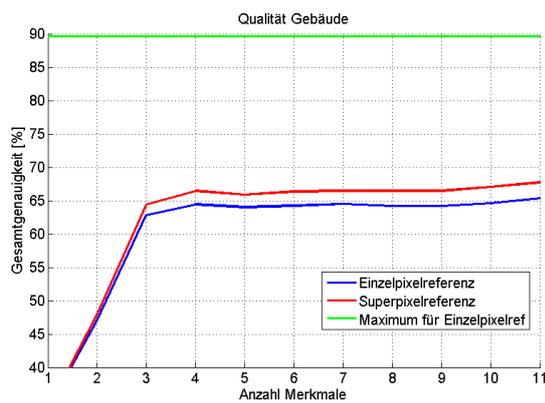
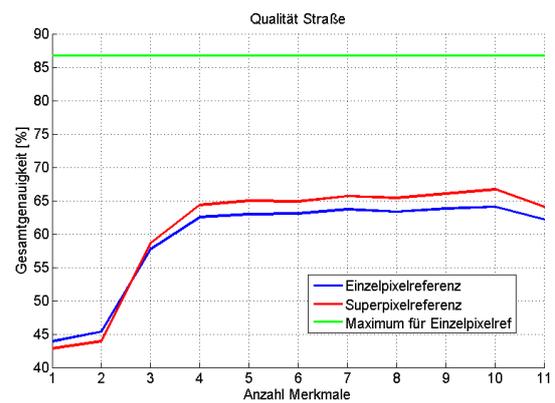


Abbildung 36: Qualitäten der Varianten für den SLIC-Algorithmus mit dem zweiten Datensatz und Random Forest

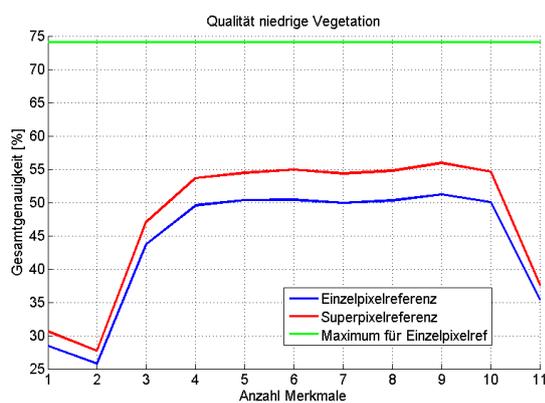


(a)

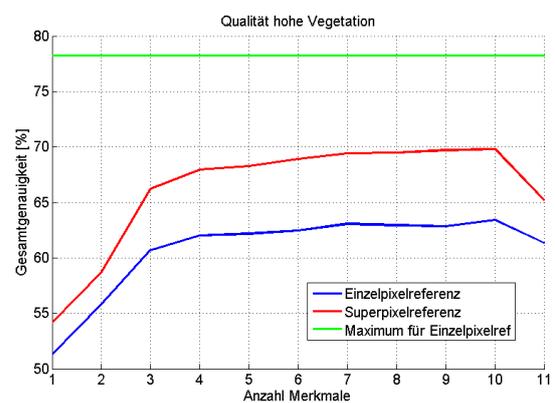


(b)

Abbildung 37: Qualitäten für Random Forest abhängig von der Anzahl der Merkmale für den ersten Datensatz

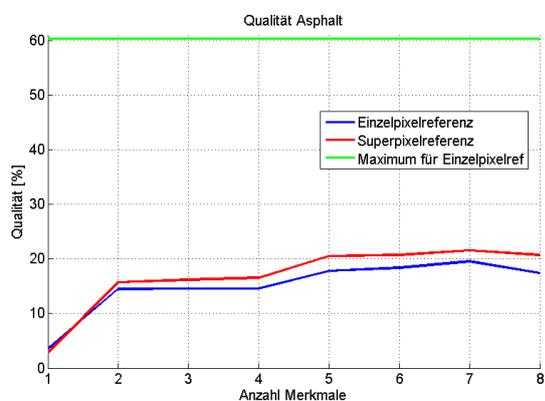


(a)

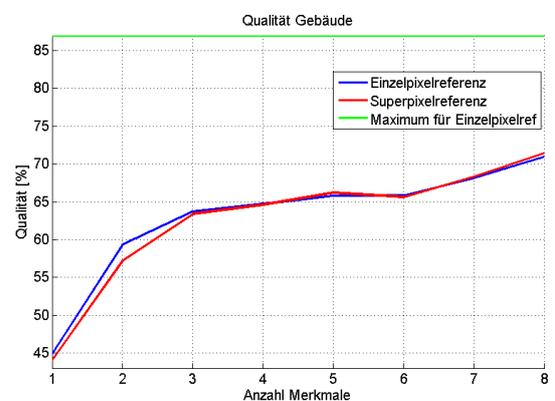


(b)

Abbildung 38: Qualitäten für Random Forest abhängig von der Anzahl der Merkmale für den ersten Datensatz



(a)



(b)

Abbildung 39: Qualitäten für Random Forest abhängig von der Anzahl der Merkmale für den zweiten Datensatz

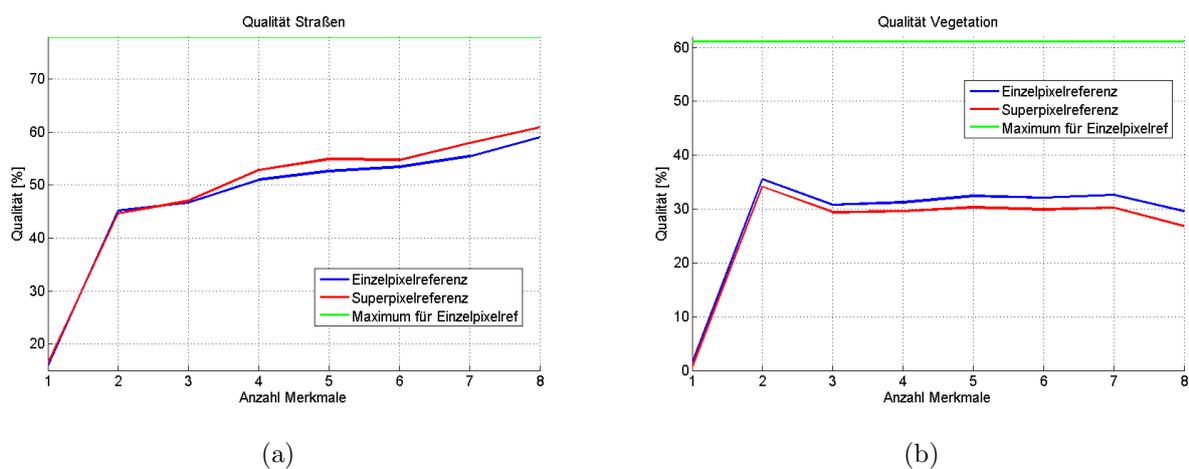


Abbildung 40: Qualitäten für Random Forest abhängig von der Anzahl der Merkmale für den zweiten Datensatz

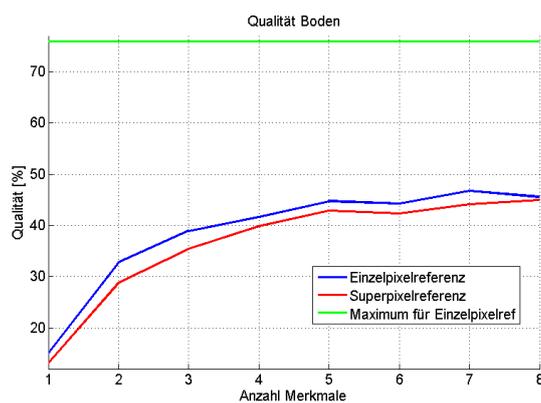


Abbildung 41: Qualitäten für Random Forest abhängig von der Anzahl der Merkmale für den zweiten Datensatz

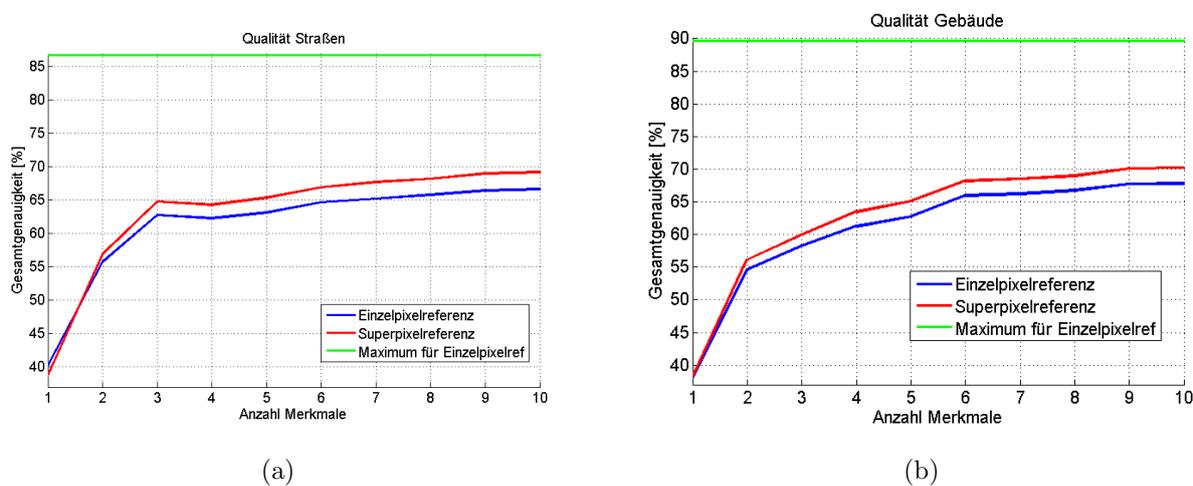
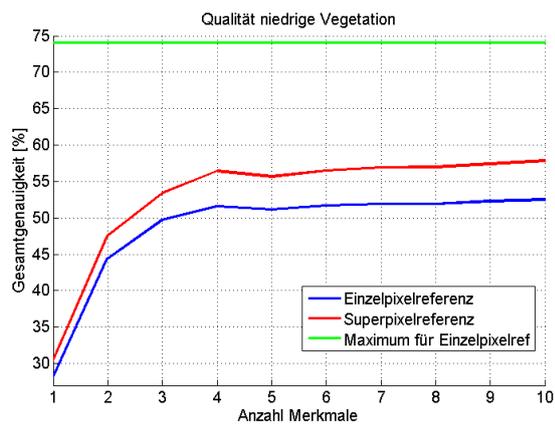
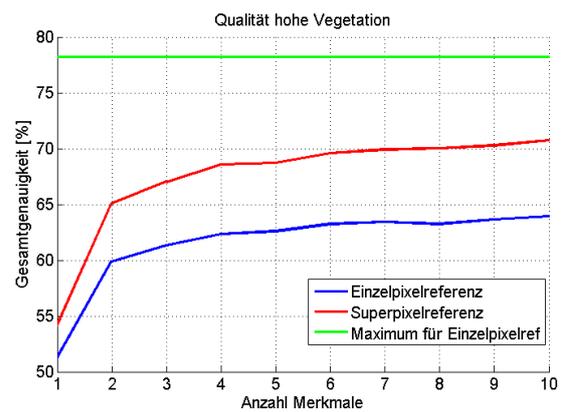


Abbildung 42: Qualitäten für die logistische Regression abhängig von der Anzahl der Merkmale für den ersten Datensatz

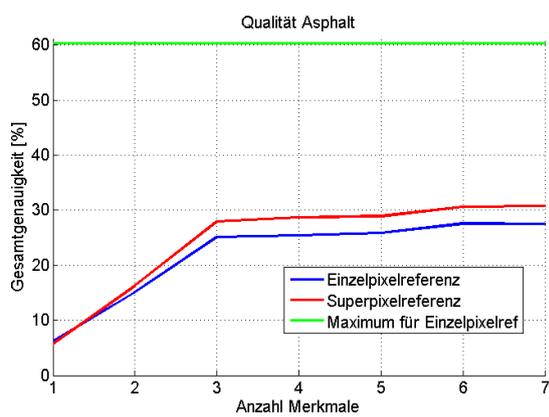


(a)

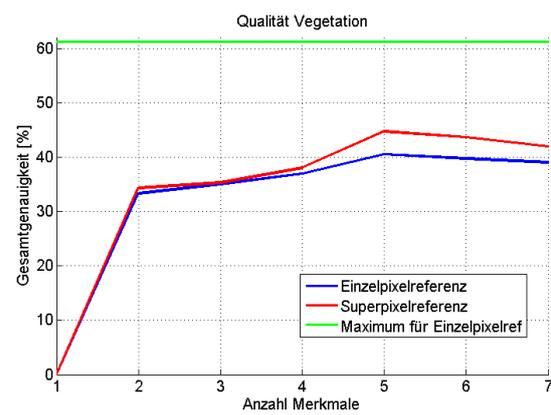


(b)

Abbildung 43: Qualitäten für die logistische Regression abhängig von der Anzahl der Merkmale für den ersten Datensatz

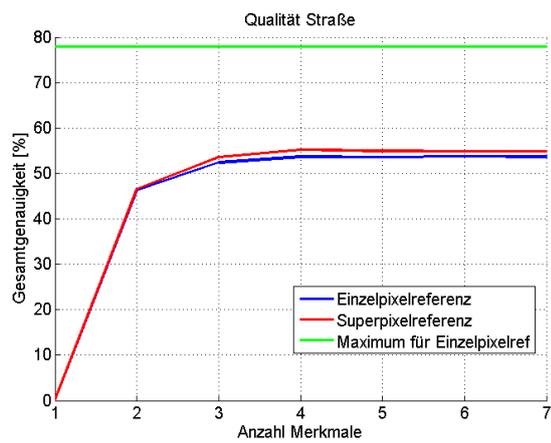


(a)

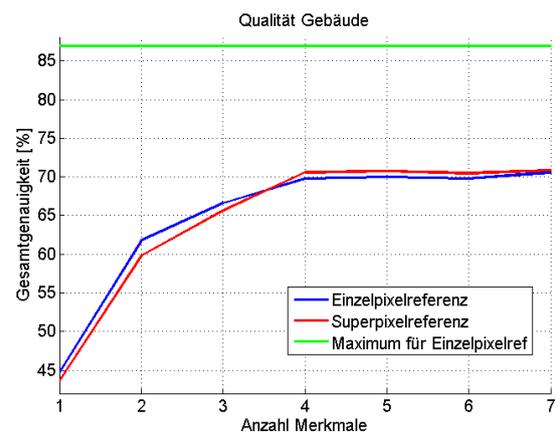


(b)

Abbildung 44: Qualitäten für die logistische Regression abhängig von der Anzahl der Merkmale für den zweiten Datensatz



(a)



(b)

Abbildung 45: Qualitäten für die logistische Regression abhängig von der Anzahl der Merkmale für den zweiten Datensatz

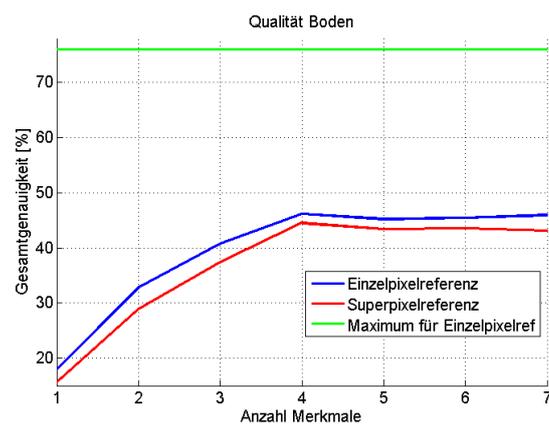


Abbildung 46: Qualitäten für die logistische Regression abhängig von der Anzahl der Merkmale für den zweiten Datensatz

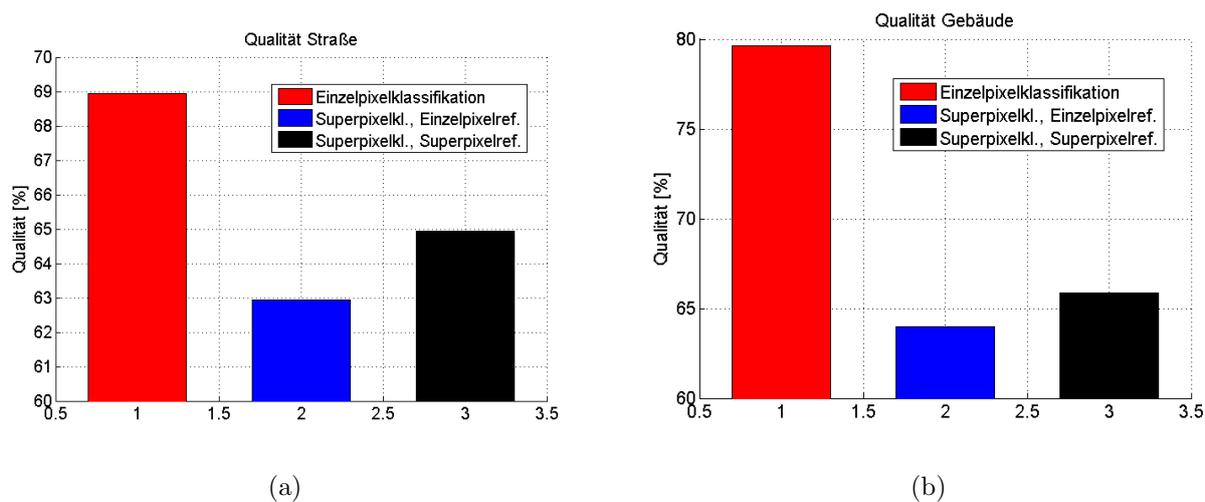


Abbildung 47: Qualitäten der Einzel- und Superpixelklassifikation für Random Forest für den ersten Datensatz

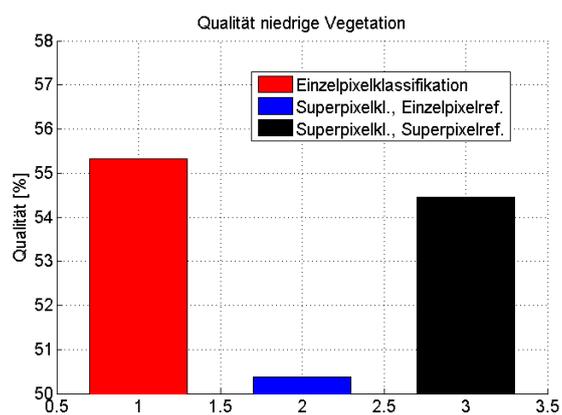


Abbildung 48: Qualitäten der Einzel- und Superpixelklassifikation für Random Forest für den ersten Datensatz

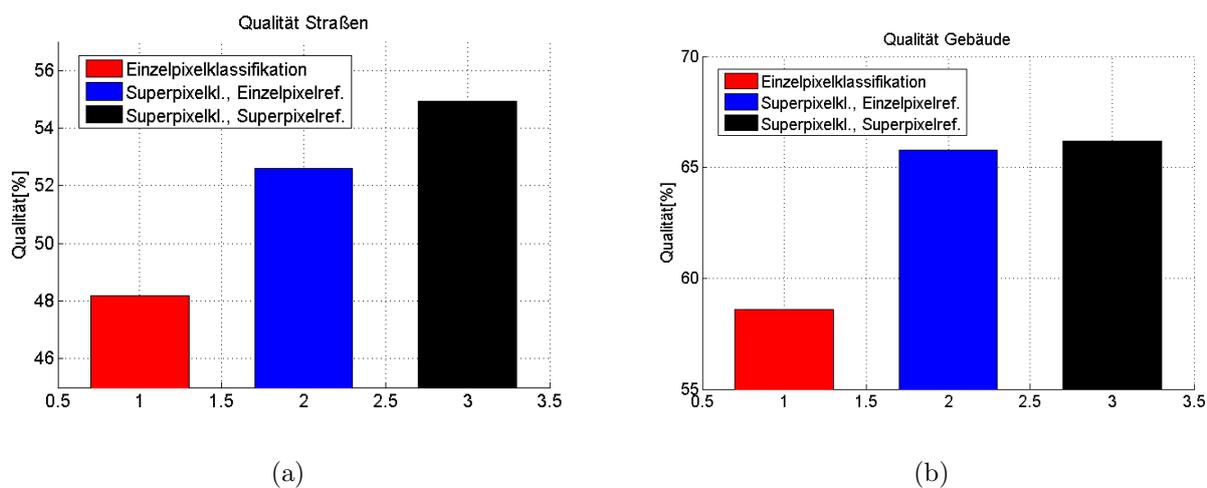


Abbildung 49: Qualitäten der Einzel- und Superpixelklassifikation für Random Forest für den zweiten Datensatz

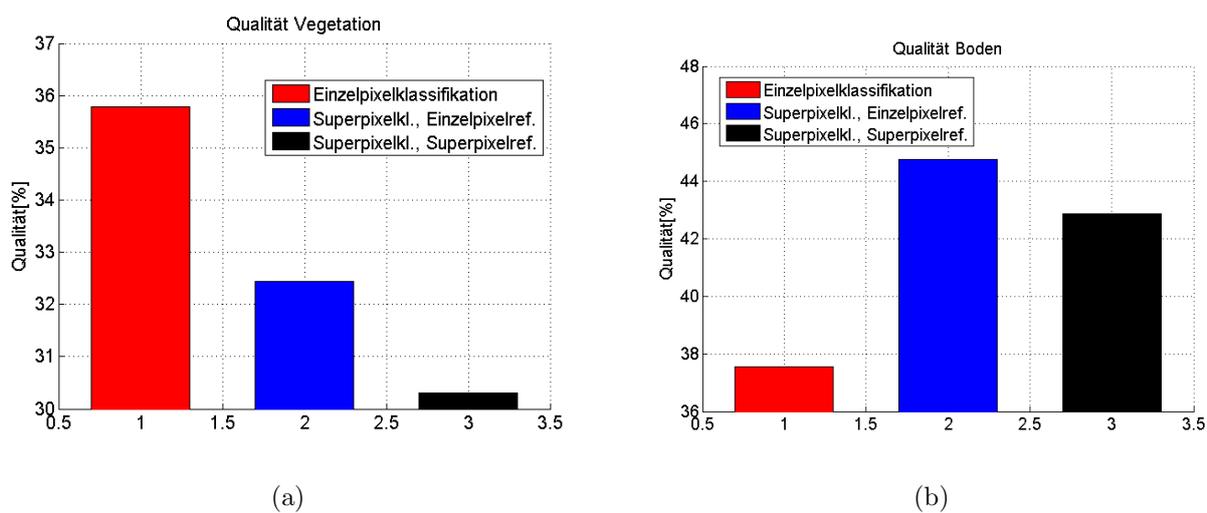
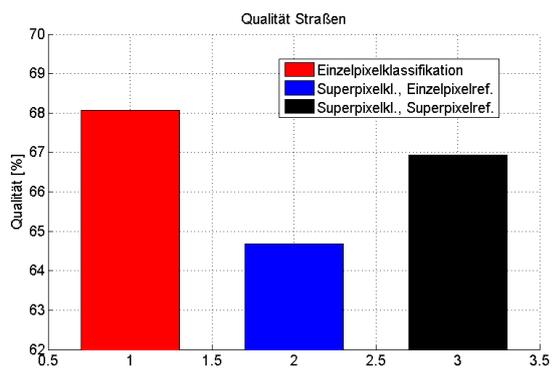
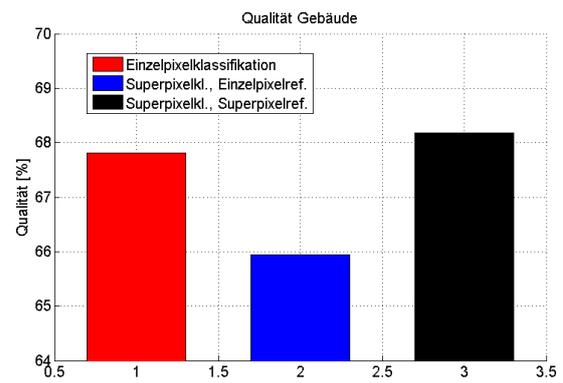


Abbildung 50: Qualitäten der Einzel- und Superpixelklassifikation für Random Forest für den zweiten Datensatz

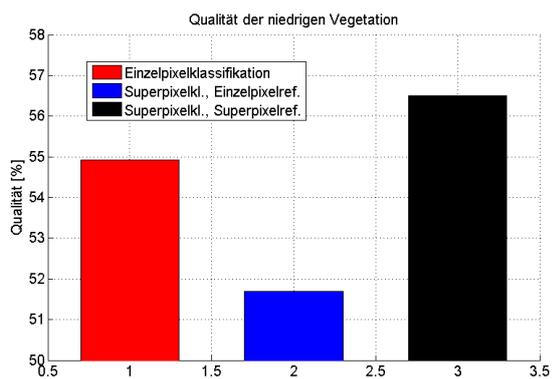


(a)

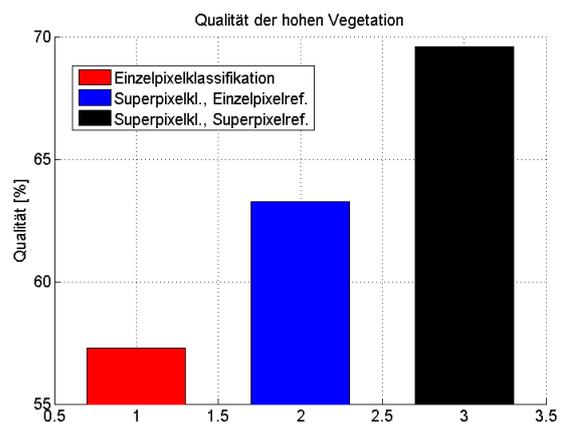


(b)

Abbildung 51: Qualitäten der Einzel- und Superpixelklassifikation für die logistische Regression für den ersten Datensatz

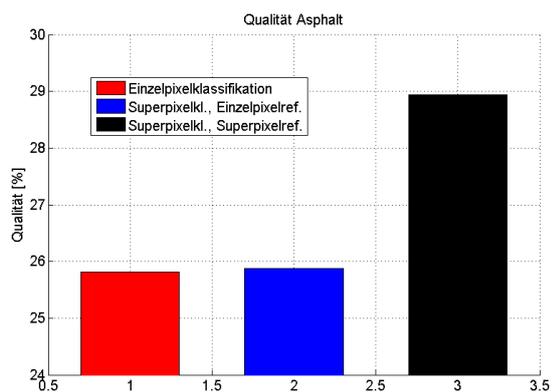


(a)

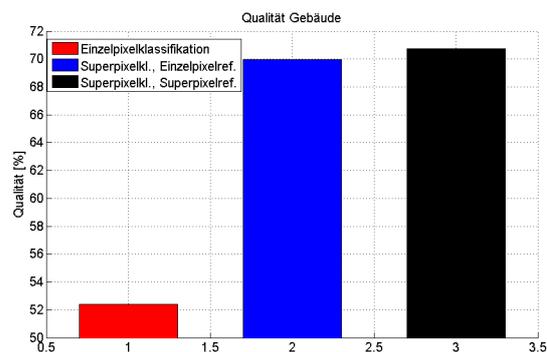


(b)

Abbildung 52: Qualitäten der Einzel- und Superpixelklassifikation für die logistische Regression für den ersten Datensatz

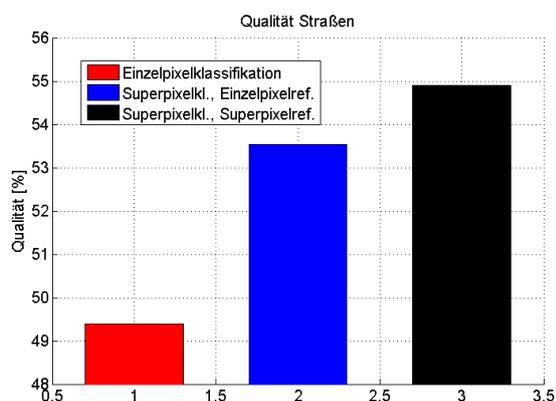


(a)

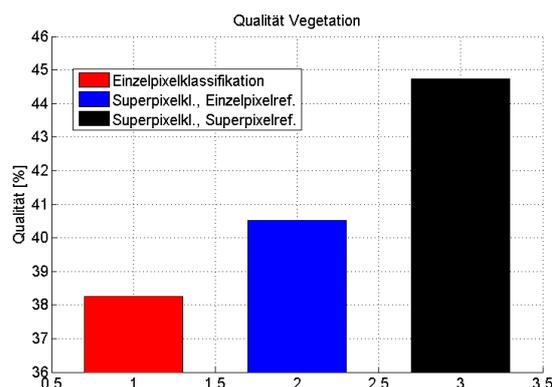


(b)

Abbildung 53: Qualitäten der Einzel- und Superpixelklassifikation für die logistische Regression für den zweiten Datensatz



(a)



(b)

Abbildung 54: Qualitäten der Einzel- und Superpixelklassifikation für die logistische Regression für den zweiten Datensatz

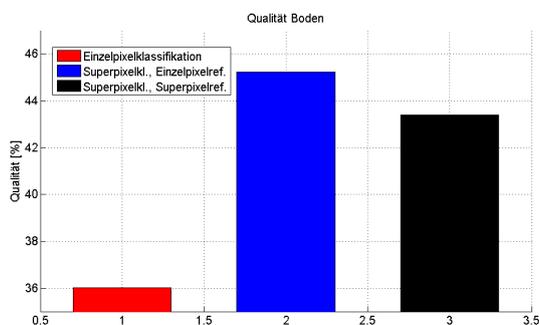


Abbildung 55: Qualitäten der Einzel- und Superpixelklassifikation für die logistische Regression für den zweiten Datensatz