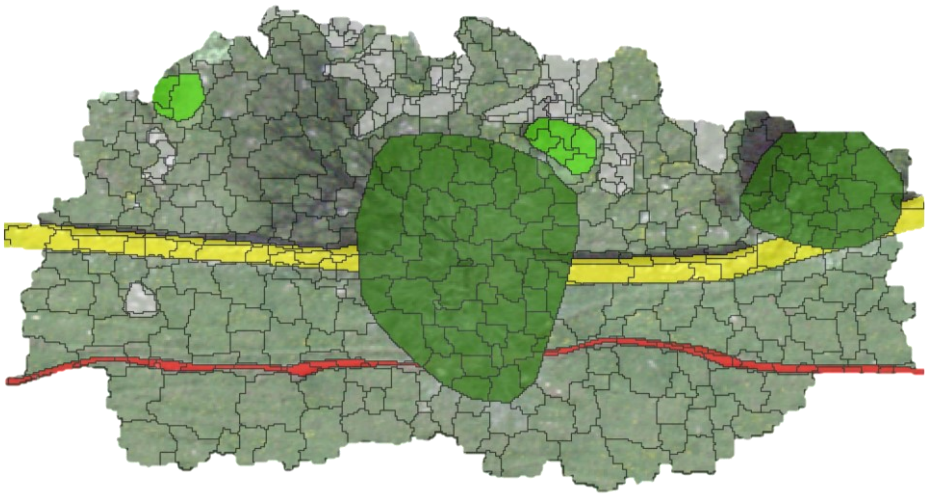
UNIVERSITÄT BERN

FRÜHLINGSSEMESTER 2022

MASTER MAJOR GEOGRAPHIE

Evaluating the Potential of High-resolution Remote Sensing Data for the Automatic Recognition of Small-scale Landscape Elements

Master Thesis



Vorgelegt von:

Simon Mägli

16-100-968

14.06.2022

Leitung: Prof. Dr. Matthias Bürgi

Betreuung: Christian Ginzler

Geographisches Institut der Universität Bern
Unit Landsysteme und Nachhaltige Ressourcennutzung (LNR)

Abstract

Small-scale landscape elements (SSLE) such as dry-stone walls, hedgerows or dead trees play an important role in terms of landscape biodiversity, landscape history as well as landscape aesthetics and recreation. In Switzerland, different monitoring programs exist that measure quantity, quality and diversity of SSLE. An example for this is the campaign "Monitoringprogramm Arten und Lebensräume Landwirtschaft" (ALL-EMA) conducted by Agroscope which records SSLE based on field surveys and aerial image interpretation. This study attempts to contribute towards contemporary SSLE monitoring frameworks by evaluating the potential of high-resolution remote sensing data for the automatic recognition of SSLE. Therefore, cost and time for SSLE inventories and monitoring could be reduced while total area coverage could be improved.

For this thesis, an approach combining object-based image analysis (OBIA) and machine learning (Random Forest) was chosen. In a first step, ten classes of SSLE were mapped across nine study areas (5.05 km^2 across the cantons of Aargau, Solothurn and Zurich) based on aerial image interpretation and field validation. Together with high-resolution LiDAR data, aerial imagery and ancillary data, ground truth data was used as input data for segmentation and feature extraction in eCognition. This resulted in two datasets (non-vegetation & vegetation classes) with a total of ~1.5 million object samples and 58 explanatory variables. The very unbalanced datasets (80% of the samples are non-SSLE objects) were subsequently cleaned up, balanced (SMOTE algorithm) and relevant features were determined using the Recursive Feature Elimination algorithm. After these preparation steps, two Random Forest models (non-vegetation & vegetation) were trained and tested using a spatial cross-tabulation approach.

Results showed that small and rare classes such as branch piles, small water bodies, clearance cairns or dry-stone walls could barely be predicted (0 – 74% producer and <6% user accuracies) while larger and more frequent classes (rocks/stones, open ground, low & high woody vegetation and dead trees) were predicted with producer accuracies between 43% and 88% and user accuracies between 1% and 80%. Main limitations were found to be very unbalanced training datasets, heterogenous remote sensing data between the different study areas and limitations resulting from the selected classification approach. However, the study also revealed a number of parameters that could be adjusted in order to improve classification results. Moreover, currently available high-resolution sensing data for Switzerland (in particular LiDAR point cloud data) can sufficiently depict many types of elements. This offers great potential for future exploration of automatic recognition of small-scale landscape elements.

Contents

List of Figures and Tables

List of Abbreviations

AG	Canton of Aargau
ALL-EMA	Monitoringprogramm Arten und Lebensräume Landwirtschaft - Espèces et milieux agricoles
ALS	Airborne Laser Scanning
ANN	Artificial Neural Networks
BFF	Biodiversitätsförderflächen
CM	Confusion Matrix
DAP	Digital Aerial Photogrammetry
DT	Decision Tree
GLCM	Grey-Level Co-Occurrence Matrix
GT	Ground Truth
KNN	K-Nearest Neighbour
LiDAR	Light Detection and Ranging
LOO	Leave-One-Out
LULC	Land Use Land Cover
MAUP	Modifiable Areal Unit Problem
MLC	Maximum Likelihood Classification
MRS	Multi-Resolution Segmentation
NDSI	Normalised Difference Soil Index
nDSM	Normalised Digital Surface Model
NDVI	Normalised Difference Vegetation Index
NDWI	Normalised Difference Water Index
OA	Overall Accuracy
OBIA	Object-Based Image Analysis
OOB	Out Of Bag
PA	Producer Accuracy
RADAR	Radio Detection and Ranging
RF	Random Forest
RFE	Recursive Feature Elimination
RS	Remote Sensing
SO	Canton of Solothurn
SSLE	Small-Scale Landscape Elements
SVM	Support Vector Machines
TLS	Terrestrial Laser Scanning
TOF	Trees Outside the Forest
UA	User Accuracy
UAV	Unmanned Aerial Vehicle
ZH	Canton of Zurich

Acknowledgements

Over the course of this master thesis, a number of people were involved without whose support the realisation of this study would not have been possible. First and foremost, I would like to express my gratitude towards Christian Ginzler and Matthias Bürgi for their dedicated mentoring, their patience and sharing their valuable expertise.

Besides them, I would like to thank the entire WSL remote sensing group for their valuable feedback. In particular, Nica Huber for the compilation of data on the presence of small-scale landscape elements, Ruedi Bösch for his help with drone applications, Bronwyn Price for sharing her expertise about eCognition Developer, Mauro Marty for his support with data conversions and Dominique Weber for giving valuable feedback about programming in R software and the application of the Random Forest algorithm. Furthermore, I appreciated the support of Raffael Bienz (data balancing in R software) and Eliane Meier (provision of ALL-EMA data).

Finally, my sincere gratitude also goes toward my partner for her unconditional support in stressful times.

1 Introduction

Small-scale landscape elements (SSLE) are defined as small structural elements of natural or anthropogenic origin that shape the character of cultural landscapes (Berner Naturschutz, 2019) but are also essential for biodiversity and many ecosystem services (Guntern et al., 2020). A cultural landscape is thereby defined as a landscape formed through the interaction of natural and anthropogenic influences (UNESCO, 1992). This study focuses on SSLEs in one part of cultural landscapes, i.e., the agricultural landscape.

Structural elements such as branch piles, dry-stone walls and small streams or ponds serve as nesting places, food source and shelter for many species but also hold historical, aesthetical and recreational values (Rey et al., 2017). Since the industrial revolution, SSLE in Switzerland have been under constant pressure through the expansion and transformation of agricultural landscapes. Traditional and structurally rich landscapes have been emptied in the course of agricultural intensification or urban expansion (Berner Naturschutz, 2019). However, over the last decades, awareness for the importance of SSLE has grown. Agricultural reforms in the 1990s managed to slow down the negative trend of the occurrence and diversity of SSLE in agricultural landscapes by creating financial incentives (Rey et al., 2017). In 2014, subsidies, so called “Landschaftsqualitätsbeiträge”, were implemented to further remunerate maintenance and care of SSLE (Rey et al., 2017).

1.1 Problem Statement

In Switzerland, cantonal and nationwide programs exist that, among other aspects, also monitor the impacts of measures supporting SSLE in agricultural landscapes. One example therefore is the campaign “Monitoringprogramm Arten und Lebensräume Landwirtschaft – Espèces et milieux agricoles” (ALL-EMA) conducted by Agroscope and WSL (Riedel et al., 2018). The program was launched 2015 with the goal of assessing the condition and change in species and habitat diversity in accordance with environmental goals for the agricultural landscape of Switzerland. Secondly, it also measures the effects of so called “Biodiversitätsförderflächen” (BFF), subsidised areas to promote biodiversity. Thirdly, it provides a data basis for current and future questions regarding species and habitat diversity in Swiss agricultural landscapes (Riedel et al., 2018).

ALL-EMA uses a set of 40 indicators to assess diversity and quality on species, habitat and BFF level. Several indicators cover SSLE: the number (*StrGamm*), diversity (*StrDiv*) and spatial heterogeneity *(StrHet)* of SSLE in 200 m^2 plots are measured within the scope of habitat surveys. Also, the proportion of study sites with woody features (*GehProz*) is measured. Furthermore, the border length of woody features neighbouring agricultural area (*GehLaeng*) is surveyed based on aerial image interpretation and the border length of water boundaries is extracted from the swissTLM3D product (swisstopo, 2022c). For the definition of sample plots, ALL-EMA uses a three-level sampling design. On a first level, a total of 170 1 km^2 study sites is chosen across Switzerland based on randomised but weighted sampling. On a second level, each site is subdivided by a grid with 50 m mesh size. The grid intersections mark the centre points for habitat surveys of which 19 are selected per site based on randomised weighted sampling. For each habitat survey, SSLE are recorded in two differently sized circle areas. In a 10m^2 circle, the proportion of open ground (low- or no-vegetation area) and shrub vegetation

(< 1 m) is measured. In a 200m^2 circle, standing and flowing water bodies, ruderal area, clearance cairns, dry-stone walls and ruins, field margins, paths as well as woody vegetation metrics such as groves, bushes, hedgerows, trees, forest edges, dissolved forest, high stem fruit trees, old trees, briars and stepped forest edges are recorded (Riedel et al., 2018).

Habitat surveys as described above are conducted by experts in the field. As such, they are time and cost intensive. With the chosen sampling design, area coverage across Switzerland is relatively low which shows the importance of applying a robust and representative sampling strategy. Moreover, some sample plots are likely not to be accessible due to impassable terrain or landowners denying access. This poses the question whether habitat surveys can be supported by automated SSLE recognition based on remote sensing data. As a result, time and cost for SSLE monitoring could be reduced while area coverage could be improved.

Over the last two decades, several studies have examined the use of automated recognition and remote sensing data for SSLE monitoring. In a publication by Völker & Müterthies (2008), operational methods for recognition, analysis and monitoring of cultural landscape parameters are presented. The authors created an automated image processing model in order to provide a standardised, transparent and objective approach to monitor landscape development. They also suggest using this approach for the controlling of agricultural subsidies targeting landscape conservation. Another example is a study conducted by Malinowski (2016) in Denmark, where LiDAR data retrieved from airborne laser scanning (ALS) and aerial imagery is used for the recognition of SSLE including hedgerows, ditches and dikes. Here, ALS data is identified as a suitable data source to monitor vegetation elements over large areas cost-efficiently.

For Switzerland, no comprehensive studies analysing automatic recognition of SSLE in agricultural landscapes are known. However, automatic recognition based on remote sensing data has become increasingly popular for vegetation structures and other landscape parameters as shown in the following examples. Weber et al. (2020) used an automatic approach for the large-scale classification of shrub forest. Pazúr et al. (2022) applied a machine learning (ML) algorithm to map the extent of cropland and grassland in Switzerland and Malkoç et al. (2021) used automated mapping for trees outside forests. These examples show that the application of automatic recognition approaches is commonplace with contemporary landscape analysis in Switzerland. It can be presumed that one reason therefore is the increased availability of high-resolution data in Switzerland. This is also the starting point of this thesis. It is assumed that with increasing spatial and temporal resolution, smaller objects such as SSLE are more likely to be recognised automatically. However, its potential has to be investigated.

1.2 Research Questions and Objectives

With those recent advances in the availability of high-resolution remote sensing data in Switzerland as well as the known potential of automated recognition for landscape objects, the general goal of this thesis is as follows:

- Evaluating the potential of high-resolution remote sensing data for the automatic recognition of small-scale landscape elements

Based on this goal, the following research questions are posed:

- RQ1: To what extent is high-resolution remote sensing data suitable for the automatic recognition of different small-scale landscape elements?

- RQ2: How do object-based image analysis and machine learning methods influence classification accuracy of different small-scale landscape elements?

In order to answer the research questions, the following objectives are set:

- Create ground truth data for selected small-scale landscape elements
- Develop and optimise a workflow for an automatic recognition of small-scale landscape elements based on remote sensing data
- Perform a plausibility analysis for the automatic classification
- Compare the classification accuracy for different small-scale landscape elements

1.3 Thesis Outline

In the following chapter, available literature about the background of this thesis is presented. Beginning with detailed information about the functions of different SSLE types, an overview of the characteristics and current state of high-resolution remote sensing data, methods of object-based image analysis (OBIA), machine learning techniques and the state of the art of automatic recognition of landscape properties is given. In chapter 3, a conceptual framework is introduced describing the chosen approach to achieve research objectives and answer research questions. Chapter 4 covers the results including mapped ground truth data as well as segmentation and classification results. Results are then discussed in chapter 5 and conclusions are drawn in chapter 6.

2 Background

2.1 Small-scale Landscape Elements

The term small-scale landscape elements (SSLE) comprises a wide variety of small structural elements in agricultural landscapes that are products of anthropogenic practices and natural processes (Beck, 1996). Over generations, many human-made structures were created for practical purposes. For example, clearance cairns are the result of stone removal from arable fields and grasslands in order to improve cultivation and crop growth (Kremer, 2015). Over time, such structures then co-formed the typical appearance of the corresponding agricultural landscape that is today associated with historical, aesthetical and recreational values (Rey et al., 2017). A study by Röser (1988) states that a landscape's recreational value is mostly based on the occurrence and diversity of boundary structures such as hedgerows or dry-stone walls. Rodewald et al. (2014) describe SSLE as essential features for the perception of different types of agricultural landscapes. For example, forest pasture landscapes are characterised by a mosaic of distinctive single trees, deadwood structures or dry-stone walls. Traditional high-stem orchards with a high proportion of deadwood give extensively farmed landscapes their typical flair. Similarly, alpine pasture landscape stand out through their high structural diversity including natural elements such as single trees, rocks and boulders or ruderal areas.

Equally important to historical, aesthetical or recreational values is the fact that SSLE simultaneously support biodiversity and ecosystem services. Many species can only exist if SSLE are available. Moreover, they are often dependent on a combination of SSLE that is spatially interconnected. Amphibians rely on both aquatic and terrestrial habitats including clearance cairns, dry-stone walls, fringe structures or woody vegetation (Guntern et al., 2020). The diversity of wild bee species is based on the occurrence of open ground areas, deadwood as well as vegetation and stone structures within a certain area (Zurbuchen & Müller, 2012). Similarly, birds such as the common redstart (Phoenicurus phoenicurus) are dependent on hollow trees that are in close distance to open ground areas with no or little vegetation (Martinez et al., 2010).

By supporting biodiversity, we also maintain or improve ecosystem services. For example, branch piles or clearance cairns provide a home for small carnivores such as weasels that can drastically reduce mouse populations in agricultural fields (Röser, 1988). Another important service is pollination. Insects such as bees, hoverflies, butterflies or beetles are essential for plant production (Guntern et al., 2020). Studies have found that the pollination of a field is improved if spatially interconnected SSLE such as hedgerows or forest edges are located close to the field (Castle et al., 2019). Furthermore, structures like hedgerows or tree rows can reduce soil erosion and provide wind protection (Röser, 1988).

In the following sections, an overview of conventional SSLE types is given. Thereby, their importance for biodiversity as well as their historical, aesthetical and recreational values (if existent) are outlined. According to Guntern et al. (2020), the umbrella term SSLE includes the categories single trees/orchards/alleys, forest edges, hedgerows and woody vegetation, deadwood structures, fringe structures, stone structures, water bodies, wetlands, open ground and ruderal area as well as others. Subsequently, common categories are presented.

2.1.1 Single Trees, Orchards and Alleys

Domestic and site-specific single trees are particularly valuable for biodiversity as they feature multiple habitat types. They serve as food source, shelter, nesting place or perching opportunities. Typical beneficiaries are birds, bats and insects but also lichens, mosses and fungi (Guntern et al., 2020). Old and prominent single trees can greatly shape the character of a landscape and contribute to its recreational value (Rodewald et al., 2014).

A special form of trees are so called pollard trees (ger: “Kopfbäume”). Through repeated capping of the trunk and consistent removal of new shoots, a characteristically head-like crown at low height is created. Over centuries pollarding was applied in order to use regrowing shoots as weaving material or food source (Braun, 2014). Today, pollarding has become rare. However, older pollard trees often have cavities and are therefore very valuable for cave-dwelling bird species, bats and insects (BirdLife Schweiz, 2003).

Traditional high-stem orchards are commonly created in loose formation across grasslands. In the process of agricultural intensification, this posed an obstruction for mechanical cultivation. Therefore, many fruit trees were removed (Kremer, 2015). Nevertheless, many people perceive high-stem orchards as an enriching feature of agricultural landscape that also contributes to recreational value (Hochstamm Suisse, 2018). High-stem orchards are among the most biodiverse habitats in central Europe (Hochstamm Suisse, 2018). Many bird species use cavities in older fruit trees for nesting and shelter (BirdLife Schweiz, 2022).

Alleys and tree rows are important links between different habitats, especially in intensively used landscapes. Traditionally, they were created for reasons of aesthetics, resource extraction as well as road traffic and landscape amelioration. Nowadays, the number of alleys is only a fraction of what it was at the beginning of the 20th century. Many trees were removed for the widening of roads. Nevertheless, alleys are still valued for their aesthetics and their practical functions such as shade or protection from snow or rainfall (Tartaro & Kunz, 2008).

2.1.2 Forest Edges

Forest edges are important transition zones, so called ecotones, between agricultural land and forests (Guntern et al., 2020). They provide habitats for species displaced from intensively used agricultural landscapes (von Büren et al., 1995). Also, forest plant species that prefer warmth and light are partially dependent on shrub belts and herbaceous fringes found in ecologically sound forest edges (Brändli et al., 2020).

Through intensification in forestry and agriculture, forest edges were (and still are) often artificially straightened resulting in an abrupt transition between forests and open lands (Imesch et al., 2015). According to current scientific knowledge, forest edges are most valuable for biodiversity if they are graded in terms of height (see Figure 1). This commonly includes a herbaceous fringe transitioning into a shrub belt which is followed by smaller trees and then taller trees marking the actual forest (Brändli et al., 2020). If forest edges are supplemented by other structural elements such as branch piles or clearance cairns, their value for biodiversity is significantly improved (Guntern et al., 2020).

Figure 1: Graded forest edge (© Simon Mägli)

2.1.3 Hedgerows and Woody Vegetation

Hedgerows are defined as linear structural elements mainly consisting of shrubs and bushes but also occasional trees (Guntern et al., 2020). If those elements are not structured linearly, they are described as woody vegetation. Both hedgerows and woody vegetation are essential for numerous species living in agricultural landscapes. This includes mammals such as bats or small carnivores, but also birds, reptiles, amphibians, insects, lichens and mosses (Guntern et al., 2020). For central Europe, it is estimated that they provide a livelihood for more than 10'000 species (Westphal, 2011). A key function of these elements is creating a corridor between different biotopes which thereby supports species mobility (AGRIDEA, 2022). Furthermore, studies indicate that they have a higher ecological value if they contain diverse woody (in particular thorny) species, but also if they include other structural elements such as branch piles (Horch & Holzgang, 2006).

Agriculture is also directly benefitting from hedgerows and woody vegetation as those elements provide shade for grazing animals as well as wind and erosion protection (AGRIDEA, 2022). Hedgerow countries are the result of parcelling processes in agricultural landscapes and therefore a key piece of regional land use history. As such, they contribute to values like identification and perception of home. Also, they can increase the visual appeal of an otherwise monotonous landscape by creating visual variation (Rodewald et al., 2014).

2.1.4 Deadwood Structures

Deadwood structures are defined as standing or lying deadwood such as branch piles, stacked firewood, tree stumps, dead trees as well as fences made of untreated wood. They provide shelter, nesting place and food source for many species (Guntern et al., 2020). For mammals such as hedgehogs and weasels, branch piles serve as hiding places which are inaccessible for their predators. Through their isolating function, they can provide protection against rain, wind and cold. Often they are used for hibernation (Stiftung WIN Wieselnetz & Agrofutura AG, 2018). Branch piles are also important for reptiles and amphibians for which they provide shelter, predator protection and food source (Sperry & Weatherhead, 2010; Indermaur & Schmidt, 2011). Furthermore, deadwood structures in general are livelihoods for deadwood fungi and insects (Guntern et al., 2020). For example, wild bees use tunnels made by beetle larvae as a place to build breeding cells (Zurbuchen & Müller, 2012).

In terms of landscape perception, dead structures are typically combined with other structural elements such as hedgerows, orchards or forest edges. As such, they shape the character of diverse and ecologically sound landscapes (Rodewald et al., 2014). However, this perception has changed over time. Branch piles and other deadwood structures that were not removed were often regarded as "untidy". Even today, this perception is still existing. Though with subsidies being paid for the creation and maintenance of such structures, awareness for their ecological value has risen (Koller et al., 2017).

2.1.5 Fringe Structures

Fringe structures are narrow strips of land alongside roads and paths, agricultural fields and also waterways (BUND (Hrsg.), 2019). Even though they might be inconspicuous, they are very valuable for biodiversity. Many insects, reptiles and amphibians are displaced from fields due to intensive agricultural practices. Those species are reliant on fringe structures that serve as refuges (Guntern et al., 2020). Also, fringe structures are corridors for species migration through their linear form and distribution in a landscape (BUND (Hrsg.), 2019). Moreover, fringe

structures alongside waterways also act as buffer against pollutants used in agriculture (Arnold et al., 2015).

Until mid-19th century, agricultural landscapes with high structural and land use diversity including a plethora of fringe structures were a common sight (Rodewald et al., 2014). However, with agricultural intensification, many of those structures disappeared as a result of land consolidations (BUND (Hrsg.), 2019). This in turn led to habitat fragmentation which limits species migration and genetical exchange within species. Experts consider this a major threat for endangered species (Frobel et al., 2018).

2.1.6 Stone Structures

The category stone structures includes dry-stone walls, clearance cairns, rocks and boulders as well as stone lenses (Guntern et al., 2020). Dry-stone walls are made of natural stone masonry without joint fillings. They are held together only by friction forces. Therefore, skilled craftsmanship is required to build long lasting walls. Dry-stone walls were traditionally built in stone-rich areas and have functioned as building walls or parcel boundaries (see Figure 2). During agricultural intensification many dry-stone walls were removed. Nowadays, they have become popular again due to their value for biodiversity (Schegk, 2015). Most importantly, they serve as heat source, wintering grounds or hiding place for reptiles, insects and small mammals such as short-tailed weasels (Mustela erminea) (Berner Naturschutz, 2019).

Figure 2: Old dry-stone wall with vegetation elements (© Simon Mägli)

Clearance cairns and stone lenses are other artificial stone structures and have similar ecological functions like dry-stone walls. Stone lenses are piles of stones where a major part of the pile lies underground and the pile itself is surrounded mostly by vegetation. If the larger part of the pile lies overground, the structure is commonly termed clearance cairn (Guntern et al., 2020). Like dry-stone walls, clearance cairns are usually created in stone-rich agricultural areas and therefore have a high historical value (Rodewald et al., 2014).

Rocks and boulders are often also partially located underground. Their overground parts can serve as habitats for lichens, mosses, ferns and small seed plants (Guntern et al., 2020).

2.1.7 Small Water Bodies

The term water bodies includes ponds, ditches, small rivers, springs as well as pools and puddles. Those structures can be permanent, semi-permanent or only temporary. Small water bodies provide habitats for many species, serve as water source for vertebrates and as hunting ground for birds and bats (Guntern et al., 2020). Also, artificial water bodies are of great value. Over centuries, ponds were created for purposes such as ice production, irrigation, peat

Figure 3: Temporary puddle: habitat for amphibians and insects (© Simon Mägli)

mining, reservoir for mills or fish production. Depending on their condition, such ponds are then inhabited by water-loving species (KARCH, 2022).

Many people underestimate the ecological function of temporary pools and puddles (see Figure 3). Often, such highly dynamic structures are perceived negatively as damages to the land and therefore removed. However, amphibians such as the natterjack toad (Epidalea calamita) are dependent on temporary small water bodies that feature little vegetation and high water temperatures. Also, this species is adjusted to the constant emerging and disappearing of such structures (KARCH, 2022).

Natural springs are sources of species-rich habitats that formed at the transition of groundwater to headwaters and terrestrial habitats. They are often inhabited by specialist species (Zollhöfer, 1997).

2.1.8 Wetlands

Small, locally restricted wetlands form in areas where water does not or only hardly drain. For agricultural landscapes, examples therefore are hollows, depressions, ruts as well as waterlogged or compacted surfaces. Like some ponds or puddles, wetlands are often highly dynamic. As such, they provide habitats for specialist plant species and other groups of organisms. A good example therefore are house martins that rely on wet areas close to their nesting places in order to gather dirt or mud for nesting (Guntern et al., 2020).

In central Europe, many wetlands have been drained in the course of agricultural expansion and intensification. This trend even has increased towards the end of the 20th century as the demand for more cropland and the efficiency of large-scale draining techniques has risen (Moser et al., 1996).

2.1.9 Open Ground and Ruderal Area

The category open ground and ruderal area includes areas with little or no vegetation (see Figure 4). Typical examples therefore are gravel paths (depending on the condition and use intensity), animal tracks, embankments, earth or sand heaps as well as gravel surfaces. Those structures provide habitats for several species. For example, lichens can grow on top and in between of pebbles (Guntern et al., 2020). Open ground areas provide nesting places for many insect species such as wild bees or wasps (Zurbuchen & Müller, 2012). In turn, those insects are food source for insectivorous species such as birds (Martinez et al., 2010).

Larger ruderal and open ground areas are typically found in extensive agricultural landscapes with old gravel tracks and fallow areas. They therefore contribute to the character and value of traditional, small-scale farmed landscapes (Rodewald et al., 2014).

Figure 4: Pits are valuable open ground and ruderal areas (© Simon Mägli)

2.2 High-resolution Remote Sensing Data

SSLE as described in chapter 2.1 are monitored through different cantonal and nationwide programs mostly based on field surveys (see chapter 1.1). In accordance with research question one (see chapter 1.2), the underlying question of this thesis is whether SSLE can be detected automatically based on high-resolution remote sensing data. In this chapter, an overview of remote sensing data, the definition of high-resolution, the state and availability of remote sensing data in general as well as the availability of high-resolution remote sensing data in Switzerland and its relevance for recognition of SSLE is given.

2.2.1 Overview of Remote Sensing

Remote sensing is defined as “the science and techniques of obtaining information about an object, land area, phenomenon, or ecosystem process acquired by a device that is not in contact with the object, area, or phenomenon under investigation” (Chawla et al., 2020, p. 298). The history of remote sensing goes back to 1858 where the first aerial photograph was taken from a balloon over Paris (Jensen, 2007). With World War 1, advantages of analog remote sensing for information collection became apparent and techniques were subsequently improved. Digital remote sensing started with the launch of the Landsat 1 satellite in 1972 (Congalton, 2010).

Remote sensing techniques are commonly categorised into active and passive systems. Active sensors measure a self-generated signal being reflected by the Earth’s surface (Horning, 2008). This category includes sensors such as RADAR (Radio Detection and Ranging), LiDAR (Light Detection and Ranging), scatterometers or laser altimeters. RADAR emits electromagnetic radiation at radio- or microwave frequencies and then measures the time between emitted and reflected or backscattered pulses with directional antennas or receivers. Thereby, the distance to an object can be determined as the pulses travel at the speed of light. LiDAR applies a similar technique, though using lasers emitting light pulses (NASA Earth Observatory, 1999). LiDAR sensors are most commonly used to retrieve digital elevation data of the Earth’s surface. Thereby, a key feature is that a single emitted light pulse can be reflected several times while penetrating objects such as tree crowns. The returned pulses then correspond to different surfaces (e.g. treetop, intermediate returns and ground) (Horning, 2008). In addition to elevation measurements, LiDAR sensors are also capable of measuring atmospheric profiles of aerosols, clouds and other atmospheric components. Thereby, laser altimeters are dedicated sensors measuring elevation data with LiDAR methods. Finally, scatterometers are high frequency microwave RADARS that are specifically designed to measure backscattered radiation. A typical use case is the retrieval of wind speed and direction over ocean surfaces (NASA Earth Observatory, 1999). Key advantages of passive systems are that they are not dependent on daylight and that they can penetrate objects such as clouds or vegetation (Horning, 2008).

In contrast to active remote sensing systems, passive sensors measure natural radiation emitted or reflected by observed objects. The most common source of radiation is reflected sunlight. Among passive remote sensors are radiometers, imaging radiometers, spectrometers and spectroradiometers. Radiometers measure the intensity of radiation in the visible, infrared or microwave portion of the electromagnetic spectrum. If a sensor includes a radiometer which is capable of scanning a two-dimensional array of pixels, it is referred to as imaging radiometer. Spectrometers are able to distinguish and analyse the spectral content of incoming

electromagnetic radiation. If those sensors are also able measure the intensity of radiation in multiple wavelength bands, they are called spectroradiometers (NASA Earth Observatory, 1999). As most passive sensors rely on sunlight as illumination source, their advantage is that they do not need their own energy source which overall results in simpler instruments. At the same time, this is also a limitation as most passive systems only function during the day. Also, radiation in the visible and infrared wavelengths cannot penetrate clouds or vegetation canopy (Horning, 2008).

Remote sensing has different characteristics that define the level of detail visible in digital images. These are commonly described as the four types of image resolution: spatial resolution, spectral characteristics, acquisition dynamics and sensor sensitivity. Spatial resolution is what most people understand by the term “resolution” and refers to the size of the of smallest discrete element in a digital image, a pixel, in ground dimensions. For example, if a satellite image has a spatial resolution of 10 m, one pixel corresponds to an area of 10x10 m on ground. The second type, spectral characteristics, describes spectral bandwidth, band placement and the number of bands. Spectral bandwidth is also referred to as spectral resolution and indicates the range of detected wavelengths in a specific image band. Band placement describes the placement of an image band in the electromagnetic spectrum and the number of bands used indicates how accurately the spectral properties of an object can be measured. Furthermore, acquisition dynamics include two components: Firstly, the time interval between the first and the second recording of a specific object (also referred to as repeat frequency of temporal resolution) and secondly, the time of acquisition. This is particularly important as many recorded objects (e.g. deciduous vegetation) are temporally dynamic and have different spectral characteristics depending on the timing of recording. Finally, the sensitivity of the sensor describes the range a sensor is able to register a signal and the potential range of digital numbers it can use to represent pixel values. The latter is often referred to as radiometric resolution (Horning, 2008).

In the case of LiDAR, recorded data is not expressed as images with pixels but as returned points which as a whole are referred to as point clouds. For each point, LiDAR systems are able to record its three-dimensional coordinates, return level (how many times a pulse was reflected up to this point), intensity (how strong the reflection was at a specific point) as well as other parameters (Hudak et al., 2009).

For remote sensing data, the term “high-resolution” generally refers to three types of resolution: spatial, spectral and temporal resolution (Benediktsson et al., 2012). For spatial resolution, the subdivision in high-, medium- and low-resolution changes over time as remote sensing instruments are progressing. Back in the 1980’s, Landsat satellite images with 60 m pixel resolution were considered as high-resolution. Nowadays, this pixel size falls into the low-resolution category, while 10-30 m is considered as medium and 30 cm – 5 m as high-resolution for satellite imagery (Earth Observing System, 2019). In terms of spectral resolution, modern systems use hyperspectral sensors which can measure reflectance in hundreds of narrow and contiguous spectral bands. High-temporal resolution generally means that data is available in shorter time intervals and over longer time periods (Benediktsson et al., 2012). In regard to LiDAR data, the technique itself is often referred to as high-resolution. However, for the resolution of elevation models the number of points per area is significant. For example, very high-resolution LiDAR drones can record up to 1500 - 2500 points per m^2 (Corte et al., 2022).

2.2.2 State and Availability of Remote Sensing Data

As mentioned before, remote sensing data has seen massive advances in resolution over the last decades. Similarly, the availability of such data has also increased significantly. Nowadays, users can choose from a wide range of open-source or chargeable products that are based on satellite, airborne or terrestrial systems (Bhunia & Shit, 2021). Since the launch of the first Landsat satellite in 1972 which marks the beginning of civilian spaceborne remote sensing, several satellite missions were started including different sensor types, time intervals and spatial resolution. Table 1 gives an overview of satellite missions from 1972 until today.

Mission	Launch year	Sensors	Revisit (day)	Spatial resolution
Landsat	1972, 1975, 1978, 1982, 1984, 1993, 1999, 2013, 2020	Panchromatic and multi-spectral sensor	16	120 m, 100 m, 60 m, 30 m, 15 m
SPOT	1986, 1990, 1993, 1998, 2002, 2012	Imaging spectroradiometer	1–3	2.5 m, 5 m, 10 m, 20 m
ERS	1991, 1995	IR radiometer, microwave sounder, radiometer, SAR	3, 35, 336	26 m across track and 6–30 m along track
RADARSAT	1995, 2007, 2018	SAR	1	8–100 m, 3–100 m, 3–100 m
MODIS	1999, 2002	Imaging spectroradiometer	1	1000 m, 500 m, 250 m
IKONOS	1999	Imaging spectroradiometer	3	Panchromatic: 80 cm B, G, R, NIR: 3.2 m
QuickBird	2000, 2001	Imaging spectroradiometer	2.4–5.9	Panchromatic: 65 cm/61 cm B, G, R, NIR: 2.62 m/2.44 m
Envisat	2002	ASAR, MERIS, AATSR, RA-2, MWR, GOMOS, MIPAS, SCIAMACHY, DORIS, LRR	35	300 m, 30–150 m
GeoEye	2008	Imaging spectroradiometer	8.3	Panchromatic: 41 cm B, G, R, NIR: 1.65 m
WorldView	2007, 2009, 2014, 2016	Imaging spectroradiometer, Laser altimeter	1.7, 1.1, <1, 3	Panchromatic 0.5 m; Panchromatic and stereo images: 0.46 m multispectral: 1.84 m; Panchromatic 0.34 m and multispectral 1.36 m
Sentinel 1-6	2014, 2015, 2016, 2017, 2021	Radar and super-spectral imaging	12, 10, 27	5–20 m, 5–40 m, 10 m & 20 m & 60 m

Table 1: Overview of commonly used remote sensing satellites. Table adapted from Zhu et al. (2018)

Apart from satellite-based remote sensing, data is also available from airborne missions. This includes airborne laser scanning (ALS), terrestrial laser scanning (TLS) and digital aerial photogrammetry (DAP) (Bhunia & Shit, 2021). Such products have become increasingly popular in recent years due to rising demand for accurate and up-to-date information and simultaneous decrease in costs (Mielcarek et al., 2020). In particular, applications of unmanned aerial vehicles (UAVs) have gained in significance recently. UAVs can be equipped with a rapidly growing number of sensors and instruments which makes them appealing for applications in smaller areas that demand higher temporal availability of data (Pajares, 2015).

2.2.3 High-resolution Remote Sensing Data in Switzerland

Coming back to the initial question of automatic SSLE recognition with remote sensing data, this chapter covers state and availability of high-resolution remote sensing data in Switzerland.

In general, the availability of high-resolution data in Switzerland has improved greatly over the last two decades. A popular product is Sentinel-2 satellite imagery which is freely available since 2015. The satellites Sentinel-2A/B record data from 13 spectral bands including red, green, and near-infrared wavelengths which allows to analyse the condition of vegetation structures based on the chlorophyll content of leaves. For Switzerland, new Sentinel imagery is available every third to fifth day (depending on scene overlap) with a swath width of 290 km and spatial resolution of up to 10 m. This makes Sentinel products considerably more appealing than data from Landsat missions (Weber et al., 2018). However, 10 m pixel size is still bigger than most SSLE types. Luckily, different types of airborne data exist that show more potential for the recognition of SSLE.

On a national scale, the Federal Office of Topography (swisstopo), records 4-band (red, green, blue, near-infrared) aerial imagery. Swisstopo offers two products, SWISSIMAGE (only RGB information) and SWISSIMAGE RS (4-band information). Between 1998 and 2005, the products were created based on analogue cameras (RGB only) with a spatial resolution of 50 cm. In a second phase between 2005 and 2016, digital cameras (Leica ADS40 / ADS80) were used to create imagery with a spatial resolution of 25 cm in lowlands and the Jura region and 50 cm in alpine regions. From 2017, the products feature a spatial resolution of 10 cm and are conducted with a temporal frequency of approximately three years. Data is acquired alternately in leaf-on or leaf-off condition for most regions (swisstopo, 2022a). In addition to aerial imagery, swisstopo also conducts ALS missions. The swissSURFACE^{3D} product is a classified point cloud based on airborne LiDAR data. The data is collected over six stages between 2017 and 2023. Therefore, the product is not yet available for all regions. SwissSurface^{3D} has a minimal point density of 5 points/m^2 and an average of 15 – 20 points/m^2 (swisstopo, 2019). This level of detail allows to deliminate even very small landscape features.

Figure 5 shows the example of a dry-stone wall on a slope which is perfectly recognisable on the 10 cm SWISSIMAGE product (middle) and the swissSURFACE^{3D} raster (right) derived from airborne LiDAR data with an approximate point density of 20 points/m^2.

Figure 5: Dry-stone wall (© Simon Mägli) recognisable on SWISSIMAGE aerial imagery (middle) and swissSURFACE3D raster data (right), both © swisstopo

Besides nationwide data provided by swisstopo, different cantons also conduct their own data acquisitions. However, availability, spatial and temporal resolution vary greatly across different cantons. Among the better equipped are the cantons of Aargau, Appenzell Inner- & Ausserhoden, Bern, Freiburg, Genf, Glarus, Neuenburg, St.Gallen, Schaffhausen, Schwyz, Solothurn Thurgau, Waadt, Zug or Zurich. Those cantons provide up-to-date high-resolution remote sensing data which is also suitable for the recognition of small objects such as SSLE.

2.3 Object-based Image Analysis

A common approach for automatic object recognition with high-resolution remote sensing data is the use of object-based image analysis (OBIA) methods. This chapter covers the advancements in image analysis over the last decades, an overview of OBIA methodology as well as information about available segmentation techniques and software.

2.3.1 Advancements in Image Analysis

Since spaceborne data became available for a broader audience with Landsat-1 in 1972, the pixel-based approach has long been the dominant paradigm in the field of digital image analysis (Blaschke et al., 2014). Initially, pixel-based approaches were used to classify land cover based on spectral properties of individual pixels, therefore not accounting for the context of the pixels (Lang et al., 2019). In the 1970s, this technique was then optimised to include image texture analysis which describes the variance and spatial arrangement of neighbouring pixel values (Blaschke et al., 2014; Marceau et al., 1990).

Around the beginning of the 21st century, higher resolution imagery became increasingly available. Commercial spaceborne images such as IKONOS (< 2 m) or QuickBird (< 1 m) together with airborne imagery products offered new opportunities for image analysis. While earlier data featured a relatively coarse resolution, it was now possible to deliminate smaller objects such as individual trees or residential roofs (Hossain & Chen, 2019; Kucharczyk et al., 2020). Additionally, higher-resolution imagery was observed to have potential negative effects on pixel-based classification due higher within-class spectral variability (Blaschke et al., 2014; Hay et al., 1996; Marceau et al., 1990).

The time around the year 2000 marks the onset of geographic object-based image analysis (OBIA) as real-world geographic objects were now being represented by multiple instead of single pixels (Blaschke & Strobl, 2001) and the commercial software “eCognition” became available. Figure 6 shows the increasing amount of OBIA publications which is coupled to better availability of high-resolution data and OBIA software.

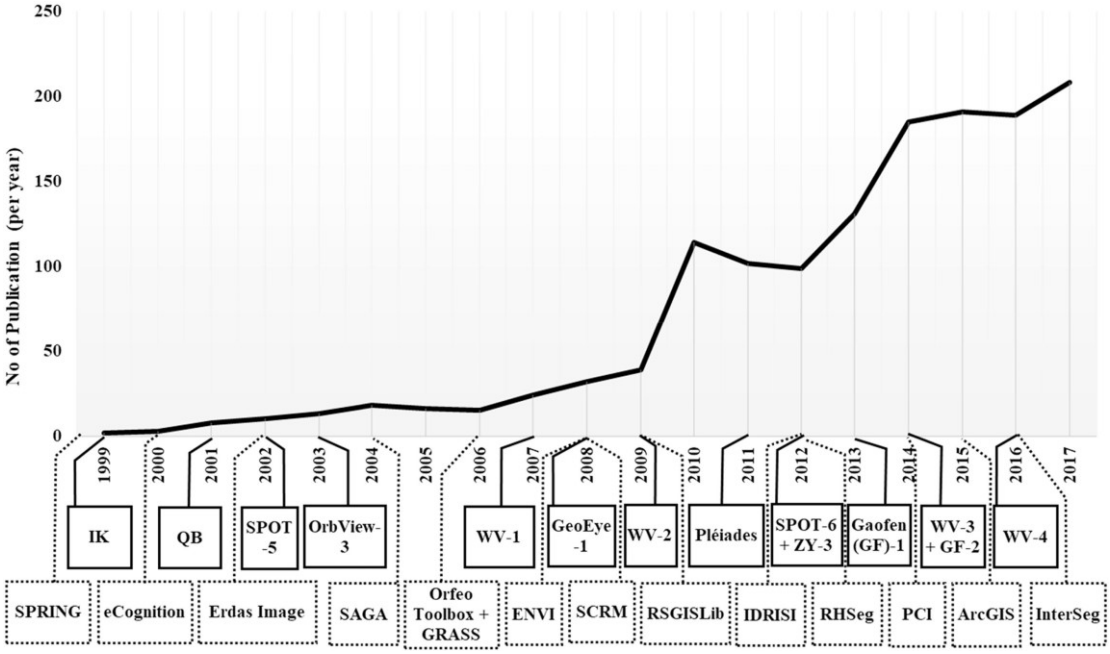


Figure 6: The amount of geographic OBIA publications over time, source: Hossain & Chen (2019)

Apart from geographical applications (often termed “GEOBIA”), OBIA methods are also used in other fields including biomedical imaging, astronomy, microscopy (Hay & Castilla, 2008). In the geographical context, OBIA is defined as an approach that subdivides images into so-called image objects which correspond to groups of neighbouring pixels that represent real-world geographical objects. Image objects are then analysed based on characteristics such as spectral, textural, geometrical and contextual features. Feature values are finally used to classify image objects into corresponding geographic object classes (Kucharczyk et al., 2020).

Over the last two decades, OBIA has established as a new paradigm for the analysis of high-resolution remote sensing data (Chen et al., 2018). Hay & Castilla (2008) name three key reasons therefore. Firstly, they argue that OBIA mimics human visual behaviour of subdividing images into multiple meaningful objects. Secondly, they mention the benefits of additional object-related information such as texture, geometry or contextual information. Thirdly, they suggest that OBIA mitigates the modifiable areal unit problem (MAUP) in remote sensing. MAUP describes the phenomenon where differing observations for a specific object can be made if data with different spatial resolution is used. Thereby OBIA helps by “shifting from arbitrary observation units (pixels) to meaningful observation units (image objects)” (Kucharczyk et al., 2020, p. 4). A further advantage of the OBIA approach is that image segmentation, the process of subdividing images into meaningful objects, can be performed on multiple scale levels (Hay et al., 2001). Thereby, high-level objects are defined by the sub-level objects they contain which facilitates exact delineation of smaller objects.

2.3.2 Overview of OBIA Methodology

According to Kucharczyk et al. (2020), the general workflow and the requirements of OBIA approaches are as follows. First of all, OBIA applications are mostly based on high-resolution imagery and so called H-resolution settings. H-resolution refers to a situation where “the geographic objects of interest are significantly larger (typically 3–5 times) than the pixels they are composed of” (Kucharczyk et al., 2020, p. 5). H-resolution situations do not imply the use of high-resolution data as large objects such as forests can also be depicted with a pixel size of e.g. 30 m (Blaschke et al., 2014).

After the acquisition of suitable input imagery, OBIA applications usually require data pre-processing. Depending on used input data, this can include atmospheric corrections, orthorectification as well as georeferencing or transforming into the desired coordinate system. Additionally, a wide range of indices can be calculated based on individual image bands of input imagery. Typical examples are vegetation indices such as the Normalised Difference Vegetation Index (NDVI) or principal component images (Grippa et al., 2017). In addition to using aerial or satellite imagery, ancillary raster elevation data such as digital terrain models (DTMs) or digital surface models (DSMs) is often used to extract characteristic information for image objects. If all required input data is compiled, the most important step is to resample all input data the same spatial resolution (Chen et al., 2018).

In a next step, a classification scheme has to be defined. This means that a legend is established where classes (objects) of interest and their hierarchical levels are defined (Blaschke et al., 2014). Thereby, it is key that the classification scheme is mutually exclusive, exhaustive, and hierarchical (Griffith & Hay, 2018). Once input data and the classification scheme are prepared, segmentation and merging algorithms are applied on the image scene (the extent of the input imagery). Thereby, neighbouring pixels are grouped into image objects representing

real-world geographic objects. An important difference exists between the terms image segment and image object. An image segment refers to the result of the segmentation algorithm. Image segments are reviewed visually to determine if they accurately delineate real-world objects of interest. If the results are considered meaningful, the term image object is used therefore (Lang, 2008). Ideally, image objects correspond one-to-one to the objects of interest. However, this is often very unlikely due to the complexity of image scenes (Castilla & Hay, 2008). Nevertheless, real-life objects should be segmented as precisely as possible because the classification accuracy calculated in a later step is directly dependent on how well an image object is characterised by spectral, textural, geometrical or contextual features (Castilla & Hay, 2008). As perfect segmentation is rarely achieved, image scenes are often over- or under-segmented. Over-segmentation refers to a situation where image segments are smaller than real-world objects of interest due to low heterogeneity between neighbouring segments. This is typically considered as a H-resolution situation. In such cases, image segments have to be merged together in order to express meaningful image objects. Contrarily, under-segmentation occurs in L-resolution situations: when the homogeneity of individual image segments is too low and therefore not enough segments are available in the scene to accurately delineate objects of interest (Castilla & Hay, 2008). Castilla & Hay (2008) also suggest that “a good segmentation is one that shows little over-segmentation and no under-segmentation” (p. 96). Further information on different segmentation technique will be presented in chapter 2.3.3.

Once segmentation has resulted in meaningful image objects, so-called features can be extracted from each image object. Features express spectral, textural, geometrical or contextual characteristics of individual image objects and are calculated based on available input data. Extracted features are then used as training inputs for a classification model. Thereby, either all extracted features or a subset thereof can be used. The latter approach is called feature reduction and aims to improve classification accuracy and computation time by removing less meaningful features (Chen et al., 2018). Feature reduction can be performed manually or through machine learning (ML) techniques. Though, it is generally recommended to apply ML algorithms as they do not assume a specific data distribution (Chen et al., 2018).

After performing feature extraction and optional feature reduction, image objects and their corresponding features are typically used as inputs for a classification model. Therefore, a sampling design has be chosen in a first step (Kucharczyk et al., 2020). The sampling design defines the minimum number of samples per class, the sampling units for testing (e.g. validation) samples as well as how training and testing samples are selected (Ye et al., 2018). Müller & Guido (2017) suggest using an equal amount of training samples per class. However, they also state that classification accuracy is greatly dependent on sample quality. Generally, they recommend using as many high quality samples as possible. In terms of sampling units, using polygon sampling units instead of pixels has become a trend in recent years (Ye et al., 2018). Finally, for the selection of samples, probabilistic methods such as simple random, stratified random or systematic sampling are typically applied (Ye et al., 2018). In addition to the general sampling design, a response design has to be defined. It refers to the approach of how segmented image object samples are labelled based on reference data. Typically, reference data is either gathered in field surveys (supported by pre-existing thematic maps or GIS data) or based on visual interpretation of H-resolution remote sensing data (Ye et al., 2018). Thereby, it is also possible to use the same data for both image segmentation and visual interpretation. However, the procedure of determining ground truth labels has to be defined unambiguously (Stehman & Foody, 2019). If field data is used in combination with remote sensing data, an important aspect is the temporal lag between reference and classification data which has to

be minimised as much as possible (Whiteside et al., 2014). A potential solution to this issue is the application of drones which has grown in popularity in recent years. Drones can greatly improve time and cost efficiency of field surveys in smaller study areas (Chen et al., 2018; Stehman & Foody, 2019).

In a next step, image objects are classified. Therefore, different methods including rule-based, supervised or unsupervised classification are available (Kucharczyk et al., 2020). Rule-based techniques refer to software such eCognition Developer (Trimble Geospatial, 2022) which allow to create predefined classification rulesets based on expert knowledge (Osmólska & Hawryło, 2018). For supervised approaches, training samples are used to train a classification algorithm. The supervised approach has become very popular over the last decade, in particular in combination with machine learning (ML) techniques. Common ML algorithms include Artificial Neural Networks (ANN), K-Nearest Neighbour (KNN), Support Vector Machines (SVM), Random Forest (RF), Maximum Likelihood Classification (MLC) or simple decision trees (DT) (Li et al., 2016; Ma, Li, et al., 2017). More in-depth information about ML and a comparison of different algorithms will be presented in chapter 2.4.

After performing classification, an accuracy assessment must be applied. The general goal of this step is to compute the thematic (and often also geometric) overlap of classified image objects in comparison to reference data. Therefore, previously segmented testing samples are often used (Kucharczyk et al., 2020). Alternatively, digitised vector reference data can also be utilised which reduces the effect of segmentation inaccuracies (Ye et al., 2018). The most frequent output of accuracy assessments are so-called confusion matrices (CM). A CM indicates common thematical accuracy statistics such as per-class user's and producer's accuracies, overall accuracy and the Kappa coefficient (Ye et al., 2018). However, the significance of the Kappa coefficient is increasingly questioned as it is often highly correlated to overall accuracy, but can also deliver misleading results (Liu et al., 2007; Pontius & Millones, 2011). Often it is also more significant to not only consider overall accuracy but also per-class accuracy statistics. This is particularly important if class frequency is heavily unbalanced (Chollet, 2018; Pontius & Millones, 2011).

2.3.3 Segmentation Techniques

In this chapter, potentials and limitations of different segmentation methods for OBIA are presented as segmentation results contribute significantly to overall classification accuracy (Hossain & Chen, 2019). According to literature, segmentation algorithms can be categorised into edge-based (Perona & Malik, 1990), region-based (Beveridge et al., 1989) and hybrid techniques (Haris et al., 1998).

Edge-based algorithms first identify edges which are defined as abrupt pixel transitions between objects (Martin et al., 2004). Therefore, edge detectors are used applying filtering, enhancement and detection procedures (Jain et al., 1995). After edge identification, edges are transformed into closed boundaries which includes removal of edges created by noisy pixels and connecting all edges to create a complete image object. Edge-based algorithms are arguably easy to implement. However, they are often missing contextual information (Hossain & Chen, 2019). The most common edge-based algorithm is called Watershed Segmentation and transforms an image into a gradient, thereby identifying objects with topographical surfaces (Mezaris et al., 2004). Despite its tendency to over-segment images, Watershed Segmentation

is widely used for natural image segmentation. However, it is often utilised in combination with other algorithms (Hossain & Chen, 2019).

While edge-based algorithms first detect edges and then fill the object comprised by them, region-based algorithms start by continuously growing an object from its centre until a boundary is reached (Y.-J. Zhang, 2006). Thereby two main operations are applied: merging and splitting (Fan et al., 2001). According to Bins et al. (1996), the basic procedure of region-based algorithms starts with obtaining an initial (under- or over-) segmentation of images, followed by merging or splitting adjacent segments depending on their similarity (or dissimilarity). These two steps are then applied iteratively until no segments have to be merged or split. For the merging, two different types of algorithms are applied: region growing and merging (Hossain & Chen, 2019). Key challenges for both approaches are seed selection (the start of the growing area) and similarity (difference between pixels) (Lucchese & Mitray, 2001). Thereby, a critical role plays the homogeneity criteria which defines whether adjacent pixels are merged in order to grow the region (Nock & Nielsen, 2004). A very popular region-merging approach is the Multi-Resolution Segmentation (MRS) technique which performs pairwise merging to form bigger objects, starting from single pixel objects. Thereby, it uses local instead of global (image scene) criteria (Hossain & Chen, 2019).

In contrast to the region growing and merging which is based on the similarity of neighbouring pixels, splitting and merging techniques start at the entire image scene splitting it into segments based on inhomogeneity criteria (Blaschke et al., 2004). Thereby, the entire image acts as seed and is continuously split into subregions until all of them become homogenous (Martin et al., 2004). A known limitation of region splitting is the tendency of image objects being too square (Cheng et al., 2001). Therefore, region splitting and region merging techniques are often combined in order to achieve homogenous regions that are as large as possible. Generally, region-based methods show better results than edge-based approaches. However, the challenge of finding appropriate seed and other parameters remains (Hossain & Chen, 2019).

Hybrid segmentation techniques were developed in order to overcome the limitations of edge-based and region-based approaches. While edge-based algorithms perform well in detecting edges but worse in creating closed segments, region-based methods are the opposite (Wang & Li, 2014). Hybrid methods therefore combine both approaches by first outlining initial segments using edge-based methods and subsequently merging them using region-based techniques (X. Zhang et al., 2014). However, hybrid approaches remain complicated which is also a reason why most software solutions do not include such algorithms (Hossain & Chen, 2019).

Overall, hundreds of segmentation algorithms have been developed but only few of them are available as tools or implemented in software solutions. In regard to segmentation software, eCognition Developer was found to be most popular being used in half of all analysed OBIA articles (Blaschke et al., 2004). Other popular commercial segmentation software includes ERDAS IMAGINE (Hexagon Geospatial, 2022), ArcGIS Spatial Analyst (ESRI, 2022c) or ENVI (Harris Geospatial, 2022). Additionally, freeware products such as GRASS GIS (GRASS Development Team, 2022), Orfeo Toolbox (CNES, 2022) or RSGISLib (Bunting et al., 2014) is also available.

2.4 Machine Learning

As mentioned in chapter 2.3.2, in recent years, machine learning (ML) approaches have become very popular within OBIA as a technique to perform feature reduction, optimise segmentation or most importantly, as a classification tool. This chapter focuses on the use for classification and gives an overview of relevant ML methods.

According to Rogan et al. (2008), ML is defined as “induction algorithms that analyze information, recognize patterns, and improve prediction accuracy through automated, repeated learning from training data (p. 2273). A key advantage of ML is that no assumptions about used data, its functional form or probability distributions have to be made beforehand. However, this also requires representative and comprehensive training data (Lary, 2022). Over the last two decades, ML has become a major area of focus within remote sensing applications.

According to Rogan et al. (2008) the popularity of ML is based on the following reasons:

- ML algorithms can deal well with multi-modal, noisy and missing data due to their non-parametric nature
- They can significantly reduce computational costs for large and complex data measurement spaces
- Categorical and continuous ancillary data can easily be integrated
- Users can evaluate the importance of individual input variables in relation to their contribution to overall classification accuracy
- They are very adaptable and can be configured to improve performance for particular problems
- They can accommodate multiple subcategories per response variable

In general, ML methods for classification problems can be grouped into two categories: supervised classification and unsupervised classification. Thereby, supervised refers to the model being trained with labelled input data. Contrarily, unsupervised algorithms attempt to find patterns without supervision and thereby cluster similar inputs into classes (Scheunders et al., 2017). Additionally, there are also hybrid approaches such as semi-supervised learning algorithms which use both labelled and unlabelled training samples (Bruzzone et al., 2006). Another hybrid approach is called Active Learning and includes an intermediate labelling of samples in order to improve classification accuracy (Crawford et al., 2013).

Based on literature indicating its great success within OBIA applications (e.g. Li et al., 2016), only supervised classification is applied in this thesis. In the following chapters, the most common supervised classification algorithms will be briefly presented.

2.4.1 Support Vector Machines

Support Vector Machines (SVM) is an algorithm that focuses on "training samples that are closest in feature space to the optimal boundary between the classes" (Maxwell et al., 2018, p. 2789). These samples are then defined as "support vectors", hence the name of the algorithm (Cardoso-Fernandes et al., 2020). In this process, SVM attempts to find an optimal boundary (also called optimal hyperplane), thereby maximising the separation or margin between support vectors (see Figure 7). In principle, SVM can only be applied to distinguish two classes. However, this issue can be circumvented by iterating through all possible combinations of classes. This in turn leads to a considerable increase in computation time if many classes are analysed (Maxwell et al., 2018).

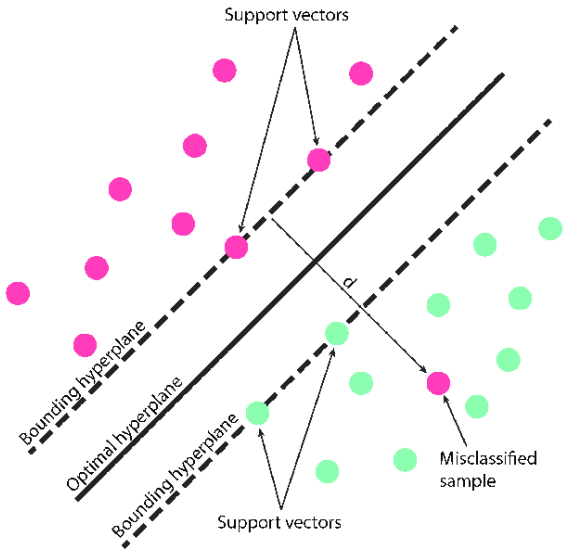


Figure 7: Optimal hyperplane und support vectors used with the SVM method. Source: Cardoso-Fernandes et al., 2020

2.4.2 Decision trees

The decision tree (DT) approach includes a reversive split of input data based on a series of decisions (Pal & Mather, 2003). Following the analogy of a real-world tree, branches symbolise possible paths after a decision was made (e.g. if a value is below or above a certain threshold). Ultimately, the tree's leaves then represent sample classes (if the algorithm is used for regression problems, leaves represent continuous variables instead of classes). Thereby, splitting criteria is defined automatically by the model itself (Maxwell et al., 2018).

DTs are generally easy to retrace. Due to their simplicity, computational cost is also very low. On the downside, DTs are prone to generating non-optimal or overfitting models. A common measure to improve the latter is pruning by which one or more splits (e.g. branches) is removed (Maxwell et al., 2018). Thereby, accuracy for the used input data is reduced but accuracy for samples outside of the training dataset is improved (Pal & Mather, 2003).

2.4.3 Random Forest

Random Forest (RF) overcomes the weaknesses of single DTs by using a large number of them. As such, it is categorised as a so-called ensemble classifier (Breiman, 2001). RF uses the majority vote of all trees in order to assign the final classification of a sample. Therefore, non-optimal or overfitting results of a single DT are balanced by the results of other trees. This mechanism is further enhanced by training each DT with a randomly generated subset of training samples that is only used for this specific tree. By only using a subset, an individual DT therefore produces worse results. However, this leads to less correlated DTs and more reliable results overall. Another characteristic of RF is so-called out of bag (OOB) data which refers to input samples not used for training. It is generally used as independent validation data for accuracy assessment. Furthermore, pruning is not necessary as the presence of multiple trees can mitigate overfitting issues. However, due to the large number of trees, individual DT cannot be visualised for better understanding (Breiman, 2001).

2.4.4 K-Nearest Neighbour

The K-Nearest Neighbour (KNN) method differs from classifiers such as RF or SVM as no classification model is established. KNN instead applies a technique where every unknown sample is directly compared against the training data (Maselli et al., 2005). Thereby, the "unknown sample is assigned to the most common class of the k training samples that are nearest in the feature space to the unknown sample" (Maxwell et al., 2018, p. 2791). The k number of training samples used for the vote can thereby be defined as parameter. K is usually a small, positive integer (Li et al., 2016). Also, higher k values lead to more generalised results (Maxwell et al., 2018).

2.4.5 Artificial Neural Networks

Artificial Neural Networks (ANN), often also termed Neural Networks (or Nets), are an analogy of human brains with axons and interconnections through synapses (Atkinson & Tatnall, 1997). Instead of axons, ANN use "neurons" organised in layers (see Figure 8). Thereby, input and output layers are defined including a neuron for each input variable and output class. Between input and output layers several hidden layers are typically installed. Every neuron in a layer is then connected to all neurons in adjacent layers. Thereby, a weighting is assigned manually for each connection. Together with a non-linear activation function, the weighting values are typically set in a trial-and-error process. Therefore, the training of ANN is generally very time-consuming. Furthermore, it can also result in non-optimal or overfitting classifications. However, ANN also have great potential in describing complex decision boundaries (Maxwell et al., 2018).

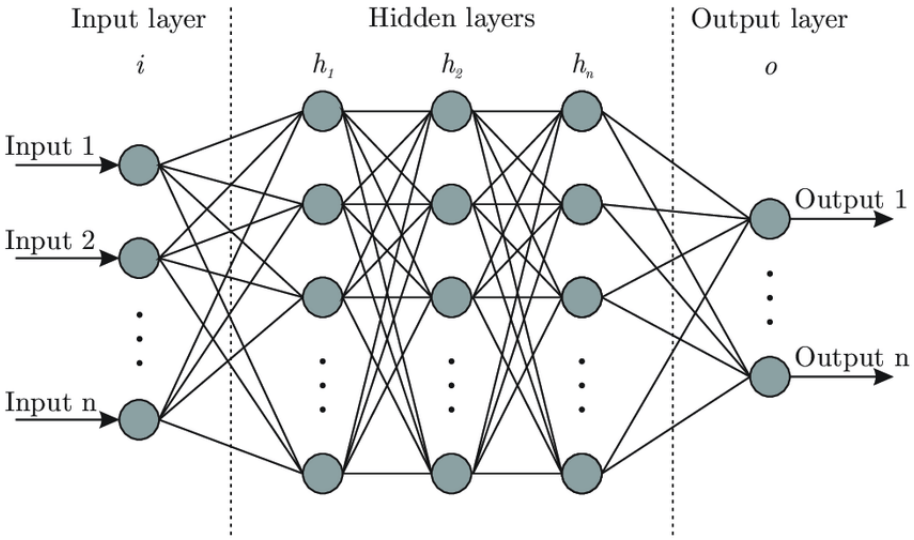


Figure 8: Artificial neural network architecture. Source: Bre et al., 2018

In regard to OBIA applications, supervised ML techniques have become the dominant approach. In a review of over 200 case studies that applied OBIA for land-cover classification, ANN approaches where most popular (29%), followed by SVM (25%), RF (20%), DT (15%) and other classifiers (Ma, Li, et al., 2017). Thereby, studies using the RF classifier achieved the highest mean overall accuracy, followed in descending order by SVM, DT, ANN and others (Ma, Li, et al., 2017). Similar results are presented in a study by Li et al. (2016) using drone imagery at an agricultural study area. Here, RF has resulted in highest accuracies while KNN provided the worst results.

2.5 Automatic Recognition of Landscape Properties

While chapter 2.3 and 2.4 give an overview on OBIA and ML methods, this chapter provides literature references for automatic recognition of landscapes properties based on OBIA and ML methods. By concentrating on landscape parameters instead of SSLE specifically, the focus is intentionally widened. The reason therefore is that only limited literature about automatic SSLE recognition was found. However, many studies apply similar methods to other landscape properties such as land use and land cover (LULC) detection. Therefore, these studies are also reviewed. For guidance, literature was analysed in respect to the following questions:

- What input data is used with which resolution?
- Which segmentation methods are used?
- What (if any) ML approach is applied?
- How good is the classification accuracy?

2.5.1 Automatic Recognition of SSLE

As mentioned before, only a limited number of studies was found analysing automatic recognition specifically of SSLE. One of those studies was conducted by Völker & Müterthies (2008). They extracted hedgerows based on an OBIA approach using coloured-infrared (CIR) orthophotos with a spatial resolution of 1 m and ancillary vector data. Segmentation and classification were performed within eCognition Developer software. For the segmentation, ancillary vector data was used to delineate field borders. However, no further details about the segmentation algorithm and setting are mentioned. For classification a hierarchical rule-set was established using NDVI and NIR features as thresholds. Finally, the accuracy assessment indicates a 100% producer and 70% user accuracy for extracted hedgerows. Contrarily, non-hedgerows sample were classified with 76.9% producer and 100% user accuracy. The authors therefore suggest that this automated approach represents a standardised and objective method for the continuation of nature conservation cadastres. For example, verifications within agricultural direct payment frameworks could be performed on the basis of automated classification of aerial imagery. However, they imply that results could differ substantially if other types of SSLE are analysed.

In a study by Hou & Walz (2013), automatic recognition of vegetation structures including groves, tree rows and forest edge transitions was evaluated. Thereby, the goal was to develop data analysis methods for a landscape structure monitoring framework that includes third dimension (height) data. As input data, multispectral RapidEye satellite imagery (level 3A) with a pixel size of 5 m, ancillary topographic spatial vector data as well as ALS-based digital elevation and surface models (both 1 m resolution) were used. Also, a normalised digital surface model (nDSM) was derived from elevation data. Subsequently, the authors applied a hierarchical multi-scale image segmentation approach in eCognition Developer in order to adapt to the different sizes of relevant objects. On the first hierarchy level, RapidEye data was used to classify major land use types (forest, grass, farmland, settlement, water body and bare rock). On a lower, more detailed level nDSM data was then used to subdivide forest and farmland area into several subclasses including groves, tree rows and forest edge transition zones. Therefore, a pixel-based object resizing method was used. Segmentation and classification was based on features including NDVI, NDWI, nDSM height information, a newly created vegetation index (REVI) which is based on red-edge band data as well as geometry characteristics (length/width) of image objects. Unfortunately, the study does not provide an accuracy

assessment on the included SSLE types due to lack of reference data. However, they conclude that the pixel-based object resizing method combined with nDSM data is capable of detecting more detailed landscape elements such as groves, tree rows and forest edge transitions.

Vegetation elements in rural landscapes were also analysed by Bolyn et al. (2019). This study focuses on the automatic detection of trees outside the forest (TOF) and classifies them into seven classes: isolated trees, aligned trees, agglomerated trees, hedges, groves, shrubs and others. The authors used leaf-off LiDAR data (0.8 points/m^2) from which a canopy height model (1 m pixel size) was computed. Additionally, leaf-on aerial imagery (25 cm) was utilised to calculate a NDVI raster layer with 1 m pixel size. Using a combination of freeware including R software and Orfeo Toolbox, elevated objects from the ground were first extracted based on height data. Subsequently, an unsupervised classification K-means clustering was applied in order to separate vegetation from buildings based on NDVI values. In a next step, TOF objects were separated from forests based on gap criteria and subsequently polygonised. TOF polygons where then classified based geometric and spatial features. Finally, classification accuracy was assessed based on a comparison with data visually interpreted by an expert. Results indicate an overall accuracy of 78% with per-class user and producer accuracies ranging between 35% and 97%. Therefore, the authors conclude that automated TOF classification is possible with high-resolution remote sensing data.

While the studies presented before concentrate on vegetation structures, Vanhuysse et al. (2014) focus on OBIA-based detection of dry-stone walls. Therefore, aerial imagery (RGB) with a spatial resolution of 15 cm is used. In a first approach, a rule-set in eCognition Developer was developed including edge detection filters, the contrast split algorithm as well as criteria related to geometric and relational features. However, results showed a high number of false positives. Therefore, a second approach was chosen using the multi-resolution segmentation (MRS) algorithm in eCognition Developer. Subsequently, 13 object features including spectral information (layer means, brightness, skewness, intensity, redness index), geometry (length/width, asymmetry, border index, density) and relations to neighbouring- and sub-objects (mean difference to darker neighbours, mean density of sub-objects) were extracted and used as input features for multiple ML models in R software. ML models included SVM, RF, KNN and Recursive Partitioning (RPart). Thereby, SVM produced best results (mean Kappa: 0.79) followed by RF (0.76) and RPart (0.67) and KNN (0.56). Based on these results the authors conclude that semi-automated object-based detection of dry-stone walls can provide satisfactory results.

2.5.2 Automatic Recognition of Other Landscape Properties

Apart from detection of specific SSLE types, object-based recognition is widely used in the field of landscape analysis. As mentioned previously, land use and land cover (LULC) classification is the most popular topic for OBIA applications. In a Swedish master thesis by Jia (2015), land cover was classified through object-based decision tree classification based on orthophoto (4-band, 25 cm) and LiDAR data (0.5 - 1 point/m^2, rasterised to 1 m). Thereby, MRS and decision classification was applied in eCognition Developer using features including vegetation indices (RVI, NDVI), a water index (NDWI) as well as other spectral and textural indices. Results showed 88.6% (LiDAR data only) and 89.2% (orthophoto only) classification accuracy for the land cover classes forest, openland, water and building. The integrated classification (LiDAR and orthophoto combined) even resulted in 95.2% accuracy. In addition to the decision tree classification, the support vector machine (SVM) algorithm was also tested.

This machine learning method achieved 89.2% (orthophoto), 77.1% (LiDAR) and 97,3% (integrated) classification accuracy. The study concludes that object-based land cover classification based on machine learning is highly effective as it can exploit spectral and spatial features from input data in an efficient manner and classifies images with high accuracy.

In a study by Kavzoglu et al. (2019) analysing automatic mapping of agricultural crop types, the combination of OBIA with different ML algorithms was evaluated. Here, WorldView-2 satellite imagery (pan-sharpened to 50 cm pixel size) was used to classify 11 LULC categories (water, agriculture, corn, hazelnut, road, soil, building, concrete, barren, shadow and forest). Thereby, segmentation was conducted based on MRS and four classification algorithms (RF, DT, KNN and Canonical Correlation Forest (CFF)) were applied. Results showed highest overall accuracy achieved by CFF (94%), followed by RF (91%), KNN (90%) and DT (87%). The authors conclude that all classifiers performed well but ensemble classifiers such as CFF and RF are likely to result in higher accuracies.

Crop monitoring was also analysed in a study by Lebourgeois et al. (2017), combining OBIA with the RF algorithm. In this study, multi-source and multi-temporal satellite imagery (spatial resolution: 0.5 – 15 m) was used. They applied the MRS algorithm in eCognition Developer and extracted several features including reflectance values (mean and standard deviation in green, red, NIR and short wave infrared), spectral indices (mean and standard deviation of NDVI, brightness index, NDWI and MNDWI), textural indices (mean and standard deviation of correlation, contrast, dissimilarity, entropy and variance Haralick indices) as well as ancillary features such as mean and standard deviation of altitude, slope and object size. Subsequently, a RF model was trained and optimised using feature reduction techniques. Overall, and accuracy of 91.7% and 64.4% for the cropland and crop subclass levels respectively was achieved. Furthermore, spectral variables were to be most important, with a better score for spectral indices over reflectance features.

To sum up, studies applying OBIA and ML for LULC analysis show very good results overall, achieving above 90% classification accuracy in many cases. Thereby, MRS was found to be popular for segmentation problems and RF is often used for classification. In regard to the topic of this thesis, the question is now how well SSLE can be detected automatically based on such methods. The main challenge is thereby, that many SSLE types are considerably smaller than most LULC classes. While larger SSLE vegetation elements such as hedges or trees were recognised relatively well (see chapter 2.5.1), classification of smaller elements such as branch piles or clearance cairns is assumed to be less accurate. This is already indicated by the study analysing automatic recognition of dry stone walls (Vanhuysse et al., 2014), which achieved slightly worse results than the reviewed study about LULC classification. However, this assumption still has to be verified.

In the following chapter, the conceptual framework and methodical approach of this thesis is presented. Thereby, used input data as well as chosen segmentation and classification methods are introduced.

3 Materials and Methods

3.1 Conceptual Framework and Method Overview

The general workflow (see Figure 9) of this study includes the creation of ground truth (GT) data, based on visual interpretation of remote sensing data and subsequent field validation. In combination with ancillary vector data, remote sensing data serves as input for object-based image analysis (OBIA) where training and testing samples are created using an image segmentation algorithm and describing variables (features) are extracted for each sample.

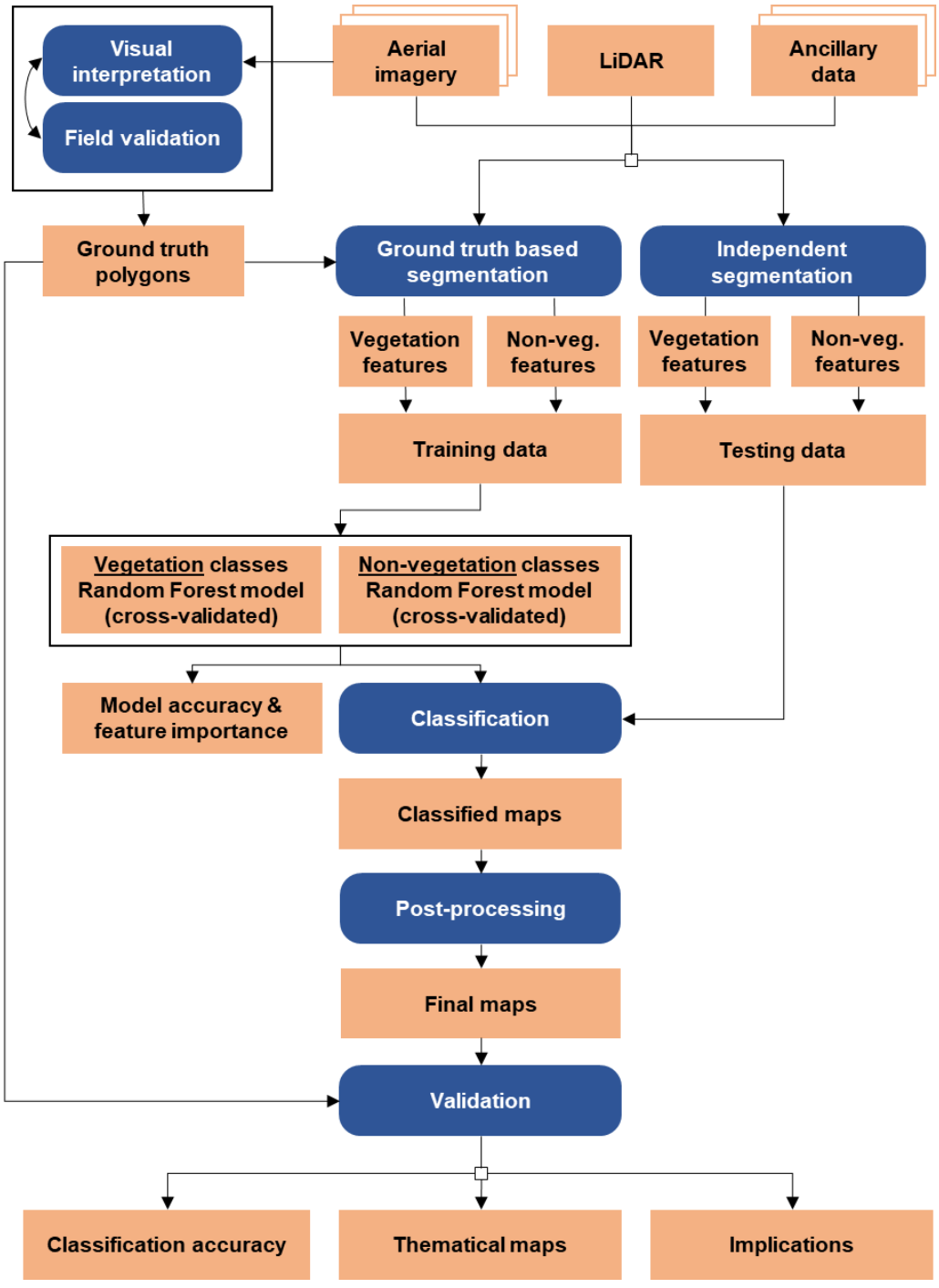


Figure 9: Conceptual framework and methodical overview. Squares symbolise products, rounded boxes represent methodological steps

Samples are then used as input for two machine learning models (Random Forest) that predict the SSLE class of testing samples. After a simple post-processing, classification accuracy of the model predictions is assessed based on a spatial cross-tabulation approach. Finally, the results and implications of the chosen approach are discussed in reference to the current state of research.

3.2 Study Areas

3.2.1 Overview

The study area consists of nine plots with a size between 0.24 and 1.06 km^2 and a total of 5.05 km^2. The plots are situated in the cantons of Solothurn, Aargau and Zurich and include the biogeographical regions Jura and Central Plateau (see Figure 10).

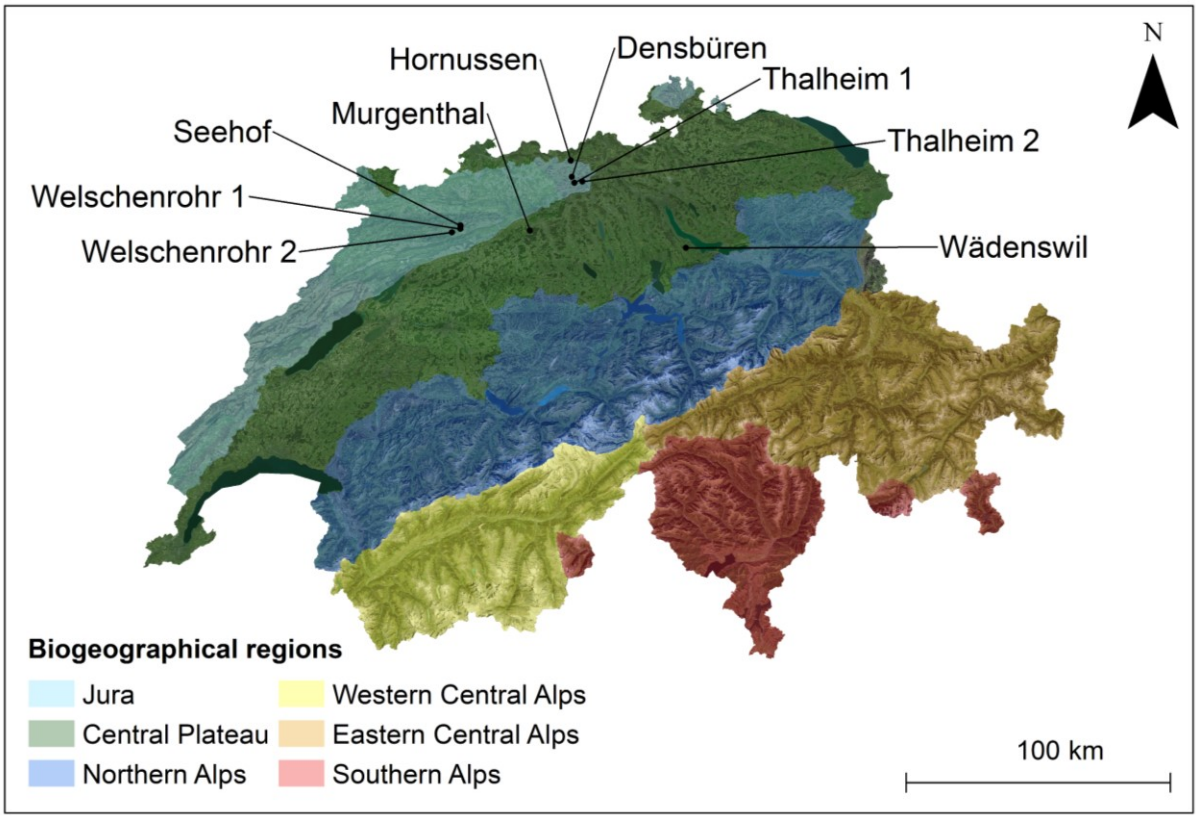


Figure 10: Study areas and biogeographical regions (© swisstopo, biogeographical regions © BAFU)

In line with the ALL-EMA program (chapter 1.1), the study area is restricted to the agricultural landscape. Therefore, settlement (reference: “Siedlung”, swissTLMRegio © swisstopo) and forest areas (reference: “Wald”, swissTLM3D © swisstopo) were masked out. Additionally, areas which are clearly not used for farming (see Appendix 1) were excluded manually.

Also, a 10 m buffer was added to the remaining agricultural plots in order to incorporate SSLE located at the edge of the forest. This is the case for many branch piles, dry-stone walls and dead trees.

3.2.2 Selection Criteria

The selection of appropriate study areas in Switzerland was based on following criteria:

- Availability of appropriate remote sensing data
- Occurrence and diversity of SSLE
- Overlap with plots from the ALL-EMA (see chapter 1.1) program
- Biogeographical diversity

Remote sensing data available for this study varied in up-to-dateness and spatial resolution. Generally, areas where data is not older than five years (2016+) and datasets featuring a high spatial resolution (<50 cm for aerial imagery, >10 points/m^2 for LiDAR point clouds) were chosen.

Secondly, appropriate study areas should boast high occurrence and diversity of SSLE. In a preliminary examination, vector datasets provided by different cantons were searched for SSLE locations. However, available datasets often included only certain types of SSLE. For this reason, ALL-EMA data provided by Agroscope was consulted. As mentioned in chapter 1.1, the ALL-EMA program records SSLE based on 170 1 km^2 plots across Switzerland. This allowed to pick plots with high occurrence and diversity of SSLE. Also, a comparison between automatically classified areas and ALL-EMA surveys is possible. Initially, the idea was that study areas should include all biogeographical regions. However, due to the scope of the thesis and the data available, only two of those regions are represented in this study.

3.2.3 Characteristics of Study Areas

In the following sections, the characteristics of each of the nine study areas are briefly described to give a short overview.

Densbüren

The study area Densbüren (see Figure 11) is situated in the Jura region of canton of Aargau and spans an area of 0.69 km^2. It is characterised by hilly terrain and small-scale grassland agriculture. As such, it records a relatively high presence of SSLE mainly consisting of hedgerows and single trees. The study area also includes the amphibian conservation area “Feret” in the south, which features multiple ponds, branch piles and clearance cairns.

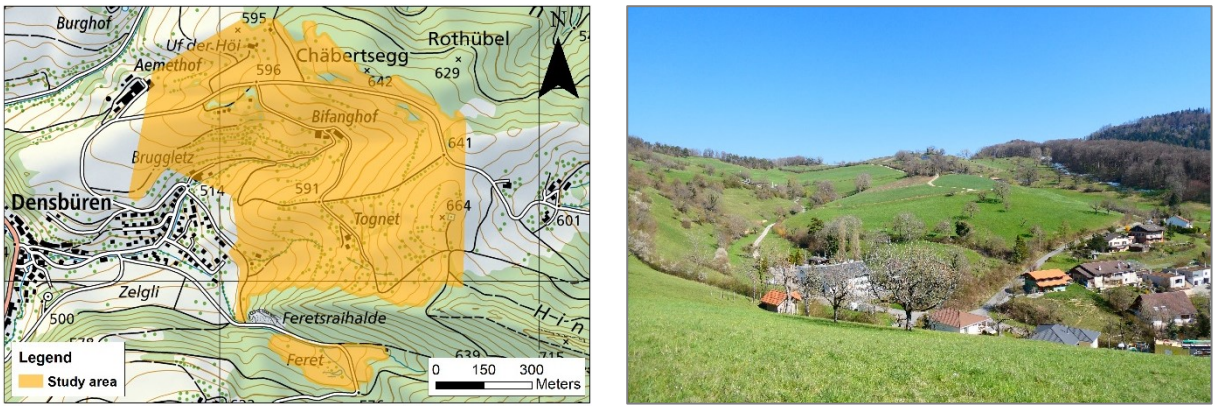


Figure 11: Left: study area Densbüren (© swisstopo), right: view to the east (© Simon Mägli)

Hornussen

The study region Hornussen (see Figure 12) is located at the edge of the Aargau Jura region. The plot lies on a plateau north of the village of Hornussen and is characterised by medium sized arable fields and a few grassland patches. As such, agricultural practices are rather intensive and the occurrence of SSLE is relatively low. The size of the study area is 0.47 km^2

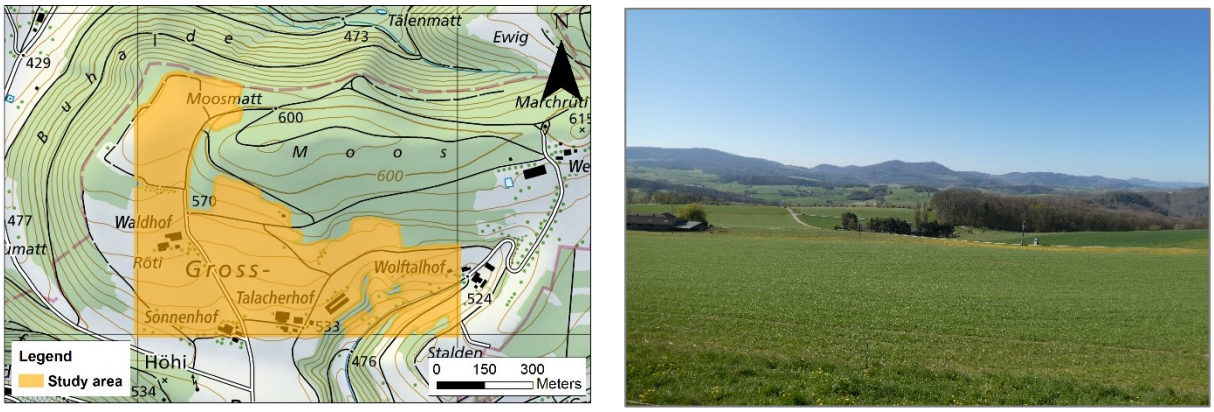


Figure 12: Left: study area Hornussen (© swisstopo), right: view to the south (© Simon Mägli)

Murgenthal

Murgenthal is the largest area analysed in this thesis spanning 1.06 km^2. The plot (Figure 13) lies within the Central Plateau region close to the river Aare. The agricultural landscape can be described as open, medium sized arable fields with occasional hedgerows. Despite high agricultural intensity, several SSLE such as dead trees, hedgerows and single trees can be found. Additionally, many branch piles are located in the vicinity of farmyards or at the edge of residential areas.

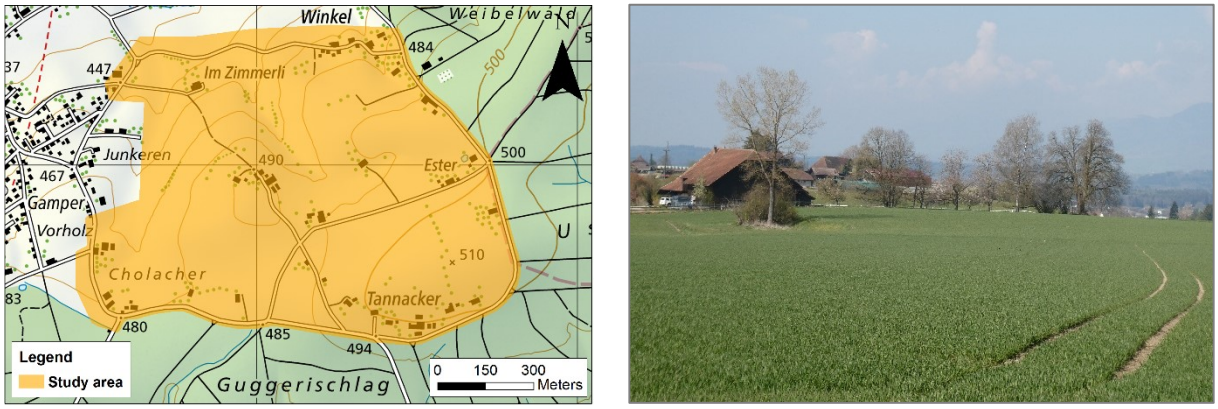


Figure 13: Left: study area Murgenthal (© swisstopo), right: view to the north (© Simon Mägli)

Thalheim 1

Thalheim 1 is another area within the Aargau Jura region with its typical hill crests. The plot (see Figure 14) spans 0.63 km^2 and is located right next to the well-known “Staffelegg”. It features mostly medium sized arable fields but also some smaller grassland areas. The occurrence of SSLE is rather mediocre. However, a few fields stand out with well-maintained branch piles and hedgerows.

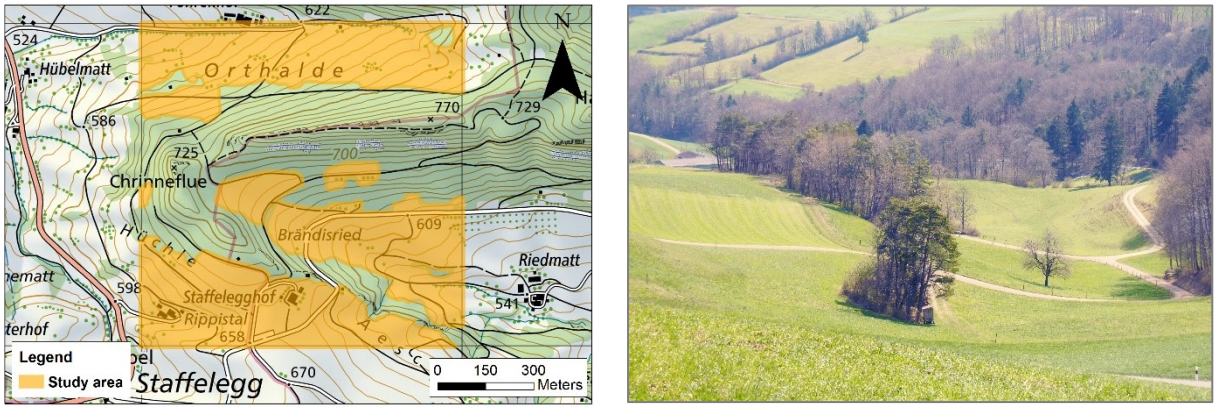


Figure 14: Left: study area Thalheim 1 (© swisstopo), right: view to the south-east (© Simon Mägli)

Thalheim 2

The second Thalheim region (see Figure 15) is located to the north of the village and has slightly different characteristics than Thalheim 1. The area spans only 0.33 km^2 and farming here is less intensive due to steeper gradients. This leads to small-scale grassland fields which boast many single trees, hedgerows and even dry-stone walls.

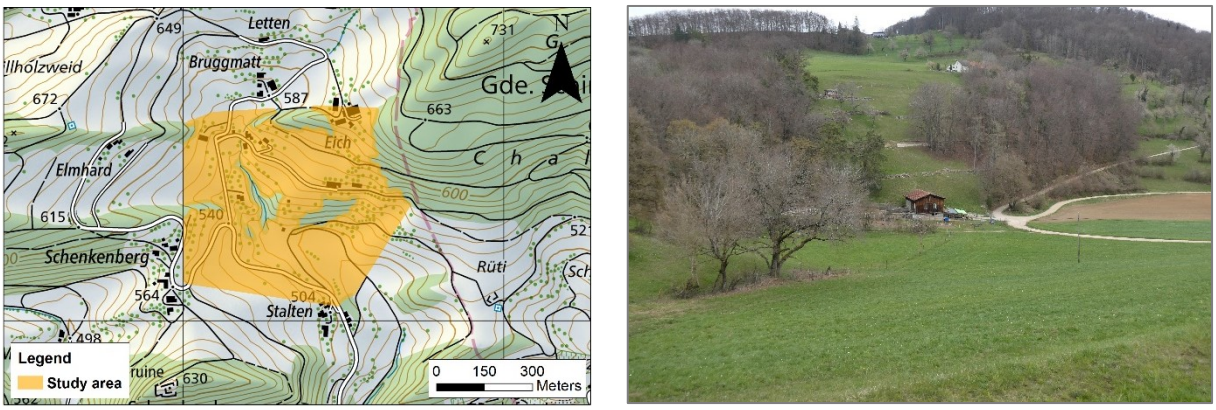


Figure 15: Left: study area Thalheim 2 (© swisstopo), right: view to the north (© Simon Mägli)

Seehof

Spanning 0.49 km^2, the Seehof study area (Figure 16) is situated in the Jura region, right at the border between the cantons of Bern and Solothurn (most of the study area lies on Bernese territory). This border is marked by the distinctive dry-stone wall visible in Figure 16. In addition to dry-stone walls, the rough terrain also features a high number of clearance cairns and single trees. Agriculture in this area only consists of small-scale grassland farming.

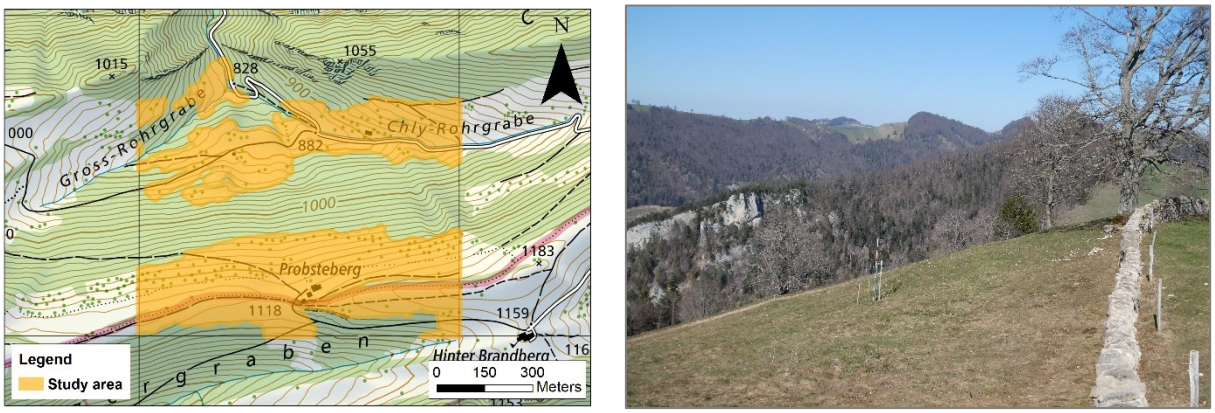


Figure 16: Study area Seehof (© swisstopo), right: view to the north-east (© Simon Mägli)

Welschenrohr 1

The Welschenrohr 1 area (see Figure 17) has a size of 0.4 km^2 and is located right next to the Seehof area. The northern part of the plot lies on top of the second Jura mountain chain, while the second part is down in the valley. The upper part can be characterised as small-scale grasslands with lots of rocky structures and low woody vegetation. The lower part is a very extensive summer grazing pasture including a plethora of low woody vegetation, rocks and stones as well some small water bodies.

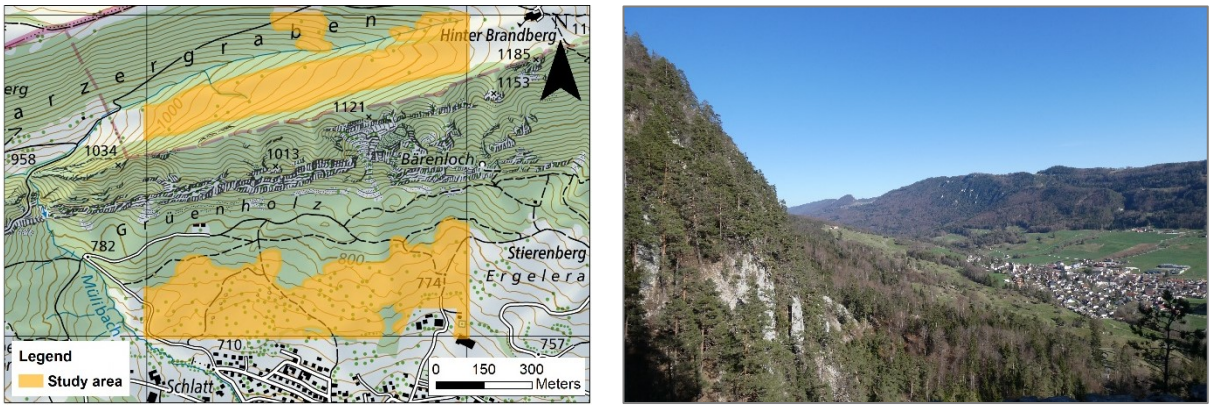


Figure 17: Study area Welschenrohr 1 (© swisstopo), right: lower part of the area (© Simon Mägli)

Welschenrohr 2

The other Welschenrohr area (see Figure 18) is located to the west of the village and on top of the second Jura mountain chain. It is also used as a summer grazing pasture and spans an area of only 0.24 km^2. It is characterised by distinctive dry-stone walls which traditionally mark the edge of the forest or borders between fields. Besides, many single trees and rocky features are present on the plot.

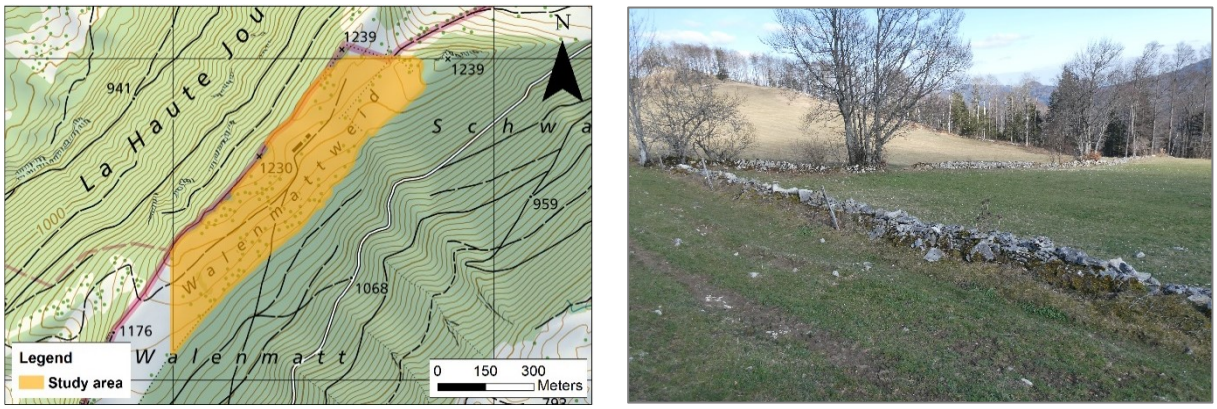


Figure 18: Study area Welschenrohr 2 (© swisstopo), right: old dry-stone walls (© Simon Mägli)

Wädenswil

The Wädenswil plot (see Figure 19) is part of the Central Plateau region close the Lake Zurich. It spans an area of 0.73 km^2 and includes features that cannot be found in any of the other areas: drainage ditches. Due the wet conditions, agricultural use is limited and most of the fields are used as grassland only. Furthermore, one part of the plot is marked as nature conservation area and it includes several branch piles, dead trees and clearance cairns.

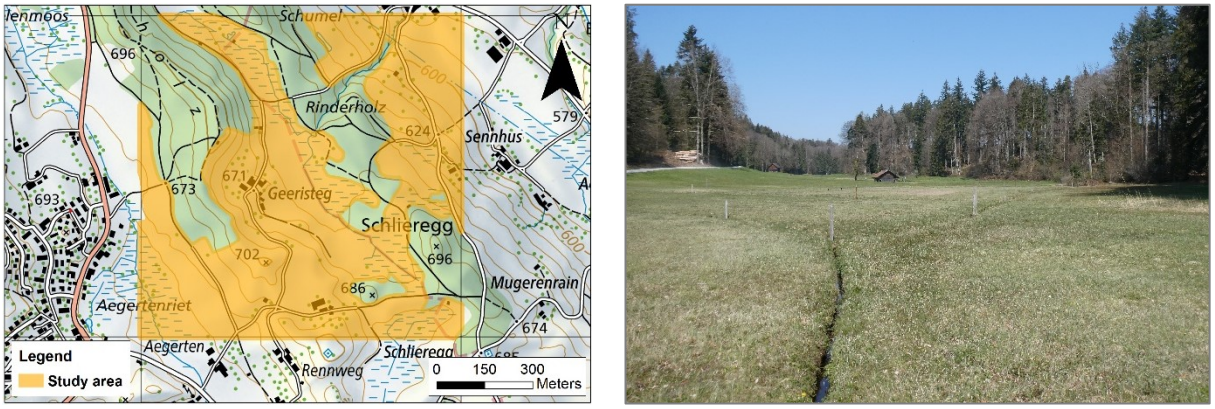


Figure 19: Study area Wädenswil (© swisstopo), right: marshland with typical ditches (© Simon Mägli)

3.3 Data Basis

For this study, a combination of LiDAR point cloud data, aerial imagery and ancillary vector and raster data was used (see Table 2). LiDAR data is provided by cantons (for study areas in the cantons of Aargau and Solothurn) and in the case of the Wädenswil area, the swissSURFACE^{3D} product by swisstopo is used. The LiDAR point clouds are mostly available in leaf-off condition and have an average point density of 17.5 to 30.8 points/m^2.

	Name	Year	Condition	Spatial res.	Source
LiDAR	LiDAR AG	2019	leaf-off	Ø 30.8 p/m^2	Kanton Aargau (2021a)
	LiDAR SO	2019	leaf-off/on	Ø 16 p/m^2	Kanton Solothurn (2021a)
	swissSURFACE3D (ZH)	2018	leaf-off	Ø 17.5 p/m^2	swisstopo (2019)
Aerial imagery	SWISSIMAGE RS (AG) (SO) (ZH)	2018 2018 2019	leaf-on leaf-off/on leaf-off	10 cm	swisstopo (2022b)
	Orthophoto RS canton AG	2019 2020	leaf-off leaf-on	25 cm 20 cm	Kanton Aargau (2021b) Kanton Aargau (2021c)
	Orthophoto RS SO[1]	2016	leaf-off/on	12.5 cm	Kanton Solothurn (2021b)
	Orthophoto RS ZH	2018 2020	leaf-on leaf-on	10 cm 5 cm	Kanton Zürich (2021a) Kanton Zürich (2021b)
Ancillary	swissTLM3D (vector)	2020	-	-	swisstopo (2022c)
	Vegetation Height Model NFI (raster)	2019	-	50 cm	Ginzler (2021)

Table 2: Overview of used remote sensing and ancillary data.

For the aerial imagery, a set of four band orthophotos was used for each study area. The idea was to use three orthophotos in different conditions and from different years per study area. However, for some areas only one or two orthophotos were available (Table 3).

Study area	LiDAR	Ortho 1	Ortho 2	Ortho 3
Densbüren	2019	2020	2019	2018
Hornussen	2019	2020	2019	2018
Murgenthal	2019	2020	2019	2018
Seehof	2019	2018	2016	
Thalheim 1	2019	2020	2019	2018
Thalheim 2	2019	2020	2019	2018
Welschenrohr 1	2019		2018	
Welschenrohr 2	2019		2018	
Wädenswil	2018	2020	2018	2019

Table 3: Condition (green = leaf-on, orange = leaf-off) and date of LiDAR data and aerial imagery

[1] For Solothurn, the 4-band ortho was created combining the available RGB and CIR products

Orthophotos were provided by cantons or by swisstopo (SWISSIMAGE RS product) and have a spatial resolution between 5 and 25 cm. The data basis was complemented with ancillary vector and raster data. From the swissTLM3D product, the feature classes TLM_FLIESSGEWAESSER, TLM_STEHENDES_GEWAESSER and TLM_STRASSE were used. Also, the Vegetation Height Model by the National Forest Inventory (NFI) with a spatial resolution of 50 cm was utilised.

3.4 Software

Below, the key software used in this study is listed and briefly described.

- ArcMap

  ArcMap is a well-known proprietary GIS software within the ArcGIS Desktop suite (ESRI, 2022b). In this study, it was used for data preparation, ground truth mapping, spatial analysis and accuracy assessment as well as map creation. Used version: 10.8

- LAStools

  LAStools is a collection of highly efficient, batch-scriptable, multicore command line tools for rapid processing of LiDAR data (Isenburg, 2021). It was used for preparation and exploration of LiDAR point cloud data. Used version: 210418

- eCognition Developer

  eCognition Developer is a powerful software for object-based image analysis. It allows users to quickly extract accurate and reliable geoinformation from any kind of geospatial data (Trimble Geospatial, 2022). In this study, it was used for image segmentation and feature extraction. Used version: 10.0.0

- RStudio

  RStudio is an open-source integrated software environment for R (RStudio, 2022). In this study, it was used to train a Random Forest model with features extracted in eCognition Developer. Used version: 10.0.0. The following R packages were used:

  - caret – v. 6.0-90, author: Kuhn (2008)
  - DataExplorer - v. 0.8.2, author: Cui (2020)
  - dplyr – v. 1.0.7, authors: Wickham et al. (2022)
  - randomForest – v. 4.6-14, authors: Liaw & Wiener (2002)
  - UBL – v. 0.0.7, authors: Branco et al. (2016)

3.5 Ground Truth Mapping

3.5.1 Class Definitions of SSLE

As there was no comprehensive and georeferenced SSLE data available for this study, SSLE ground truth (GT) data was created from scratch. In a first step, a set of SSLE classes was selected based on their presence in agricultural landscapes and their occurrence in the chosen study areas. The chosen classes were subsequently defined in reference to the ALL-EMA program and other expert literature (Table 4).

Class	Definition	Code	Reference
Branch piles	Min. area: 0.5 m^2, man-made, consist of multiple branches, no wood stacks, bark piles count	1	Buholzer et al., 2021
Small water bodies	Area: 1 – 1000 m^2	2	adapted from Buholzer et al., 2021[2]
Rock, stone, debris, boulders	Min. area: 2 m^2, at least fist size, max distance between two rocks within the same area: 1 m	3	adapted from Buholzer et al., 2021[3]
Clearance cairns	Min. area: 0.5 m^2, min. height: min. 0.5 m	4	adapted from Buholzer et al., 2021[4]
Dry-stone walls	Dry-stone walls	5	Buholzer et al., 2021
Open ground	Min. area: 0.5 m^2, connected low- or no-vegetation area such as patchy grassland, dirt roads, embankments or cow trails	6	adapted from Buholzer et al., 2021[5] Guntern et al., 2020
Hedgerows	Length: > 25 m, width: < 25 m, shrub proportion: > 60%, gap between crown edge: < 5 m	7	Riedel et al., 2018
Low woody vegetation	Bushes and low trees, height: 1-3 m	8	adapted from Riedel et al., 2018[6]
High woody vegetation	High bushes and trees, height: > 3 m	9	adapted from Riedel et al., 2018[6]
Dead trees	Standing dead trees, height: > 3 m	10	Guntern et al., 2020[7]

Table 4: Studied SSLE classes and their definitions

[2] Puddles were not considered due to high temporal variability. Max. area was increased to obtain more samples.
[3] The max. distance criteria was added to assist interpretation based on aerial imagery.
[4] The min. height criteria was added to distinguish class 3 and 4 based on LiDAR data.
[5] Open ground on arable fields is not included due temporal variability.
[6] Breast Height Diameter (BHM) was not considered due to the scope of the thesis.
[7] Only standing dead trees were considered.

3.5.2 Visual Interpretation of Remote Sensing Data

The SSLE GT data was created based on visual interpretation of aerial imagery and LiDAR data. For each element, a polygon was drawn by hand using ArcMap. The chosen SSLE classes were split into two major groups: non-vegetation (class 1 to 6) and vegetation (class 7 to 10). This was necessary to account for the fact that in reality non-vegetation elements can be overlapped by vegetation elements. However, in order to extract training features based on segmented image objects, GT polygons in a specific dataset must not overlap each other. Therefore, two separate GT datasets were created for non-vegetation and vegetation elements. As you can see in Figure 20, rocks and stones (class 3) are vertically covered by high woody vegetation (class 9). So in this case, class 3 polygons were created for the non-vegetation and class 9 polygons for the vegetation dataset.

Figure 20: Left: rocks covered by trees (© Simon Mägli), top right: polygons classified as rocks (class 3) in grey, bottom right: polygons classified as high woody vegetation (class 9) in dark green (© swisstopo)

For the visual interpretation of a study area, the latest available orthophoto was consulted first. If elements with height criteria (e.g. low woody vegetation is 1-3 m in height) were observed, LiDAR point cloud data and the VHM dataset were consulted in order to determine the height. If available, orthophotos in leaf-off condition were consulted in order to detect elements covered by vegetation. However, a key challenge was the temporal differences between datasets of a study area. For example, a tree can be visible in last year's orthophoto but not in this year's LiDAR point cloud, because it was cut down. Therefore, the rule was set that an element has to be present in either the latest available orthophoto or in the LiDAR point cloud in order to be counted. However, the issue of inconsistent information across multiple datasets remains. This will be further discussed in chapter 5.1.

After visual interpretation of remote sensing data, GT data was validated in-situ. The main goals were to clarify cases where elements could not be identified unambiguously and to remove false positives. Additionally, photographs were taken in order to document the study areas. The combination of visual interpretation of remote sensing data and in-situ validation is a common approach for the generation of reference data for OBIA (Ye et al., 2018).

3.6 Image Segmentation & Feature Extraction

In a second step, mapped GT data was imported to eCognition Developer together with vector data defining the study area, orthophotos, LiDAR and ancillary data for image segmentation and subsequent feature extraction. Image segmentation is the process of grouping neighbouring pixels into meaningful objects, so called image objects (Kucharczyk et al., 2020). For each image object a number of descriptive variables, so called features, can then be extracted from input datasets. The eCognition Developer software offers the great advantage of combining different input data types (LiDAR point cloud data is automatically rasterised) and automatically resampling all raster data to the highest input resolution. This is a basic requirement for OBIA applications (Chen et al., 2018).

3.6.1 Ground Truth Based vs. Independent Segmentation

For each of the nine study areas, a separate project was created within eCognition Developer. For each project, the first step was to restrict the working area to the study area geometry. The product was then quadrupled creating four equal “maps” (in eCognition Developer, a map is a sub-element of a project). The reason for this was to obtain the following four outputs:

Segmentation	Output	Description
GT-based segmentation	1) [StudyArea].NonVeg	Segmentation is based on geometry and features of non-vegetation GT data
	2) [StudyArea].Veg	Segmentation is based on geometry and features of vegetation GT data
GT- independent segmentation	3) [StudyArea].TestNon-Veg	Segmentation is independent of GT geometry. Parameters are optimised for non-vegetation classes
	4) [StudyArea].TestVeg	Segmentation is independent of GT geometry. Parameters are optimised for vegetation classes

Table 5: The four outputs for image segmentation and feature extraction (GT = ground truth)

The GT-based segmentation was performed to create training samples for the non-vegetation and vegetation group respectively. Therefore, a study area was first separated into GT image objects and their surrounding area using the geometry of mapped GT polygons (see Figure 21, a & b). This offers the advantage of maintaining the exact delineation of SSLE. The surrounding area was then further segmented to create training samples for non-SSLE (defined as “class 0”) areas. Therefore, the multi-resolution segmentation (MRS) algorithm (see chapter 2.3.3) was used (Figure 21, c & d).

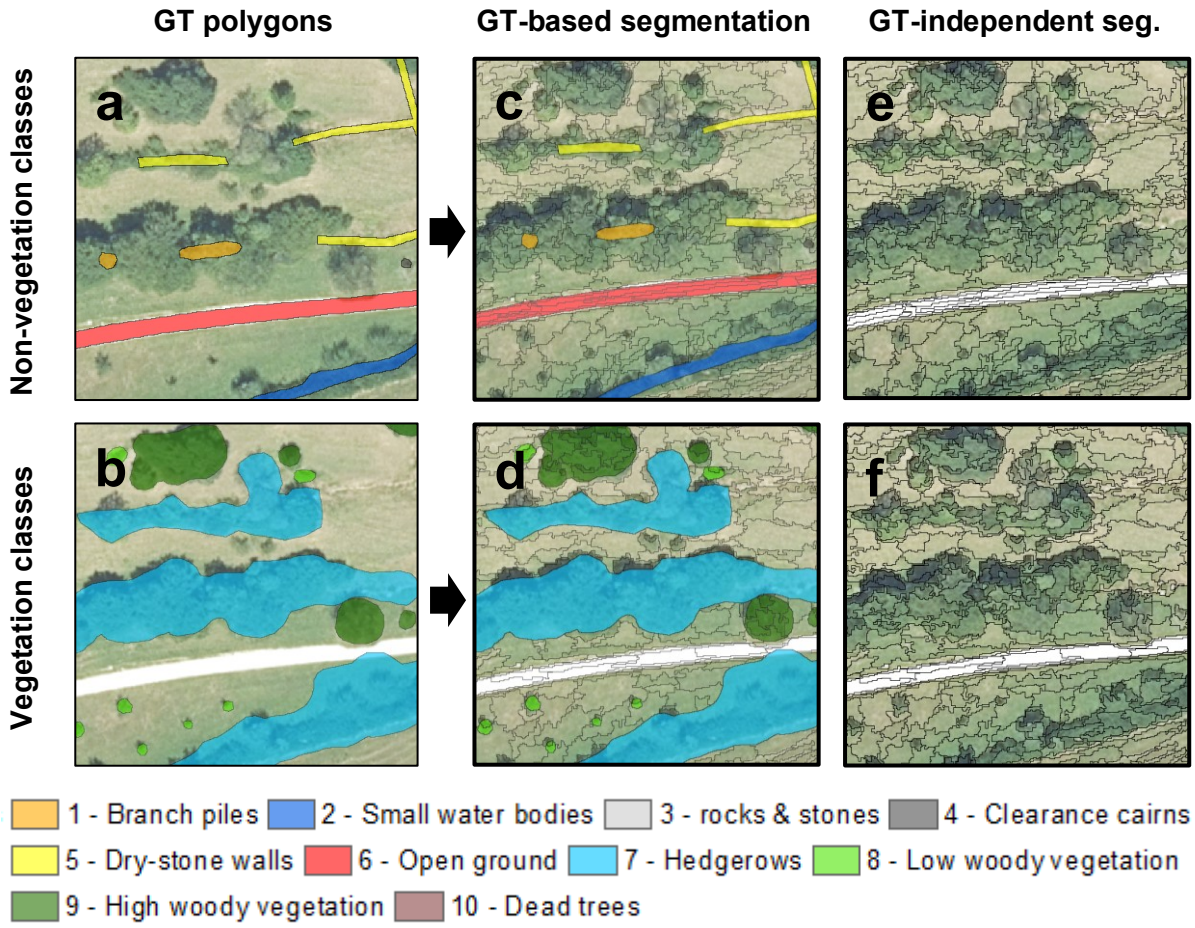


Figure 21: a) Non-veg. GT polygons, b) Veg. GT polygons c) Non-veg. training samples, d) Veg. training samples, e) Non-veg. testing samples, f) Veg. testing samples (© swisstopo)

MRS is an iterative algorithm that assembles multiple objects to create larger objects. The algorithm groups objects (starting with single-pixel objects) into pairs until a specified scale parameter (weighted with shape and compactness parameters) is locally exceeded (Trimble Geospatial, 2022). MRS was chosen due to its broad use and success in OBIA (e.g. Hou & Walz, 2014).

As some GT image objects were very large in size (e.g. large, connected gravel road polygons) and therefore unlikely to be recognised as a single element during automatic recognition, MRS was also applied to subdivide GT image objects. However, for some classes geometrical characteristics are essential (e.g. dry-stone walls can only be distinguished from clearance cairns by their linear geometry). Therefore, only image objects of the following classes were further segmented: 3 (rocks & stones), 6 (open ground), 8 (low woody vegetation), 9 (high woody vegetation) and 10 (dead trees). Furthermore, long and contiguous dry-stone walls were manually cut into several polygons in order to increase the total number of samples for this class.

The GT-independent segmentation was performed to create testing samples (see Figure 21, e & f). It was applied on the same study areas. However, no GT data was used here and image objects are solely based on the MRS algorithm. In order to ensure that training and testing data does not originate from the same area when training the machine learning model, a spatial cross-validation approach was used. This will be elaborated in chapter 3.7.4.

3.6.2 Segmentation

As already mentioned, several parameters can be adjusted when using the MRS algorithm. In this study, appropriate parameters settings were defined in a trial-and-error approach. The general goal was to delineate the smallest elements (e.g. branch piles could be 0.5 m^2) while not subdividing larger elements (e.g. a tree crown) into small and meaningless segments. Naturally, some trade-offs had to be made.

For most study areas, the following parameter settings were used: scale parameter: 150, shape: 0.9, compactness: 0.9. For the study areas Welschenrohr 1 and 2 as well as Seehof, less input data was available (Table 3). As this resulted in poorer segmentation results, the scale parameter was set to 20 in order to obtain similar results.

Apart from parameter settings, segmentation results are dependent on which features are incorporated and how those are weighted. In this study, the selection of features was determined in reference to literature and after exploring the characteristics of GT image objects.

Resulting from this, reflectance features containing mean and standard deviation values of multi-temporal, four band aerial imagery were calculated (Table 6).

Type	Feature	Data source	Reference
Reflectance	Mean_red1	Orthophoto 1	Lebourgeois et al., 2017 Husson et al., 2016
	StdDev_red1		
	Mean_green1		
	StdDev_green1		
	Mean_blue1		
	StdDev_blue1		
	Mean_nir1		
	StdDev_nir1		
	Mean_red2	Orthophoto 2	
	StdDev_red2		
	Mean_green2		
	StdDev_green2		
	Mean_blue2		
	StdDev_blue2		
	Mean_nir2		
	StdDev_nir2		
	Mean_red3	Orthophoto 3	
	StdDev_red3		
	Mean_green3		
	StdDev_green3		
	Mean_blue3		
	StdDev_blue3		
	Mean_nir3		
	StdDev_nir3		

Table 6: Used reflectance features

From the same datasets, spectral indices including NDVI (Normalised Difference Vegetation Index), NDSI (Normalised Difference Soil Index) and NDWI (Normalised Difference Water Index) values as well as the texture indices GLCM (Grey-Level Co-Occurrence Matrix) correlation and standard deviation were calculated (Table 7).

From LiDAR data, the mean number of returns and the mean, standard deviation and maximum Z value (height above ground) was used. From the geometry of image objects, the curvature to length ratio, length of the main line (sum of all distances between the nodes of the main line of an image object), area, width, compactness and density was calculated (Table 7).

Type	Feature	Data source	Reference
Spectral indices	Mean_ndvi1	Orthophoto 1	Lebourgeois et al., 2017
	StdDev_ndvi1		
	Mean_ndsi1		
	StdDev_ndsi1		
	Mean_ndwi1		
	StdDev_ndwi1		
	Mean_ndvi2	Orthophoto 2	
	StdDev_ndvi2		
	Mean_ndsi2		
	StdDev_ndsi2		
	Mean_ndwi2		
	StdDev_ndwi2		
	Mean_ndvi3	Orthophoto 3	
	StdDev_ndvi3		
	Mean_ndsi3		
	StdDev_ndsi3		
	Mean_ndwi3		
	StdDev_ndwi3		
Texture indices	GLCM_Corr	Orthophoto 2	Husson et al., 2016 Lebourgeois et al., 2017
	GLCM_StdDev		
ALS indices	Mean_NoR	LiDAR	Malinowski, 2016
	Mean_Z		
	StdDev_Z		
	Max_Z		
Geometry	Curv_Length	Segmentation results	Sheeren et al., 2009
	Length_main		
	Area		
	Width		
	Compactness		
	Density		
Ancillary indices	Mean_intInd	LiDAR	Demirbaş Çağlayan et al., 2020 Bolyn et al., 2019
	StdDev_intInd		
	Mean_vecInd	swissTLM3D	
	StdDev_vecInd		
	Mean_vhm	Vegetation Height Model	
	StdDev_vhm		

Table 7: Used spectral, texture, ALS, geometry and ancillary features

Additionally, six ancillary indices were used (Table 7). This includes mean and standard deviation of the vegetation height model (VHM), an intensity indicator and a vector indicator. The intensity indicator is based on the observation that SSLE covered by vegetation (e.g. branch piles) can often be distinguished from its surroundings by the intensity information of LiDAR points (Figure 22). Therefore, an intensity threshold was set in order to remove undesired points. Also, points above 2 m were removed. From the remaining points, a raster was then calculated and used as a feature.

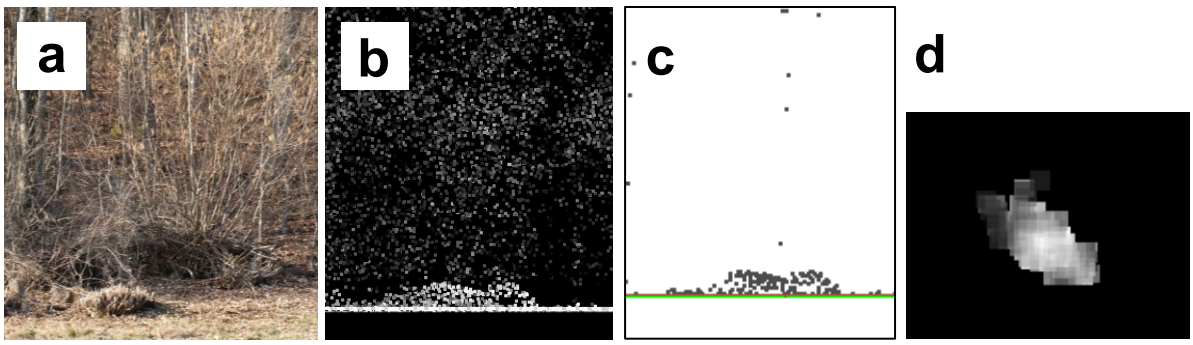


Figure 22: a) Branch pile covered by vegetation, b) branch pile distinguishable by LiDAR intensity, c) result after removing points based on thresholds, d) resulting intensity indicator raster (from above) (© Kanton Aargau)

The vector indicator is based on swissTLM3D vector data (swisstopo, 2021c). Here, a buffer was applied around the feature classes TLM_FLIESSGEWAESSER, TLM_STEHENDES_GEWAESSER and TLM_STRASSE. A binary raster was then calculated indicating whether a pixel is within this buffer.

The selection and weighting of the features for the segmentation was explored in a trial-and-error process. Results differed between non-vegetation and vegetation outputs. For non-vegetation, reflectance and spectral features as well as the mean Z feature, the intensity indicator and the vector indicator were used. Features based on the orthophoto 2 (leaf-off), the intensity and vector indicator and the mean Z feature were weighted threefold. For vegetation outputs, spectral features, NDVI features, VHM features as well as the mean Z feature were used. The latter was weighted fivefold while all orthophoto 1 (leaf-on) features and the VHM features were weighted threefold.

3.6.3 Feature Extraction

After performing image segmentation, feature values were extracted as CSV (comma-separated value) tables. For non-vegetation outputs, a total of 58 features (all features except vhm indices) were extracted per image object. For vegetation outputs, the intensity and vector indicator indices were omitted resulting in 56 extracted features per image object. Due to the high number of features, the extraction was very computationally intensive. In particular, the calculation of GLCM texture features was very time-consuming. In addition to features, pixel coordinate values for the centroid of each image object were also extracted for accuracy assessment later on. Therefore, image objects and their respective pixel coordinates were also exported as shapefiles.

3.7 Classification

Resulting CSV feature tables were subsequently imported into R software as input samples for two Random Forest (RF) classification models: a non-vegetation model (classes 1-6) and a vegetation model (classes 7-10). RF (see chapter 2.4.3) is a powerful machine learning algorithm with broad use in object-based image analysis (OBIA) problems (e.g. Husson et al., 2016). It was coined by Breiman (2001) and is defined as an ensemble of individual decision trees where the majority vote of individual tree predictions is used as model prediction. RF was chosen for this study due to its robustness against highly correlating features and its capability to manage a large amount of data and variables (e.g. Lebourgeois et al., 2017).

3.7.1 Pre-processing

CSV input data was cleaned for multiple reasons. In a first step, feature values were standardised due to varying bit depth resulting from different orthophoto inputs. Standardisation is the process of rescaling data values so that the mean of all values is 0 while the standard deviation is 1. Therefore, the dplyr package (Wickham et al., 2022) was used. In a second step, incorrect and missing values were cleaned up. Also, two features (Mean_ndwi2 and StdDev_ndwi2) were removed due to processing errors.

Furthermore, class 7 (hedgerows) samples were discarded due to a semantical overlap with other vegetation classes. More specifically this means, that hedgerows contain elements of low and high woody vegetation (class 8 and 9) as well as dead trees (class 10). The criterion that separates hedgerows from other vegetation classes is primarily the geometry (see chapter 3.5.1). However, with the chosen segmentation approach, testing samples do not represent the geometrical characteristics of hedgerows. Thus, hedgerow samples were discarded.

3.7.2 Data Balancing

Unbalanced training data is a common issue in machine learning problems. A frequent starting point is that majority classes heavily overweigh minority classes which then leads to a biased model and therefore reduces overall classification accuracy (Mellor et al., 2015). However, different techniques exist in order to balance training data. Undersampling describes the process of removing majority class samples (Freeman et al., 2012) while oversampling includes an artificial replication of minority class samples (Ling & Li, 1998). Also, methods combining both under- and oversampling exist (N. V. Chawla et al., 2002).

In this study, a combined under- and oversampling approach using the SMOTE algorithm from the UBL package (Branco et al., 2016) was applied. SMOTE is a well-known balancing algorithm developed by Chawla et al. (2002). Thereby, new samples of the minority classes are artificially generated using the nearest neighbours of these cases. Also, the majority class samples are undersampled which leads to a more balanced dataset (Branco et al., 2002).

For both the non-vegetation and vegetation model, class distribution was heavily unbalanced as class 0 (no SSLE) contained as much as 91% (non-vegetation model) and 73% (vegetation model) of all samples (see chapter 4.2). However, based on a trial-and-error process, a slightly different approach was chosen for each model. For the non-vegetation model, the number of samples per class was first randomly reduced to a maximum of 3000 in order to improve computation time. In a second step, all classes were under- and oversampled using SMOTE which

led to an equal distribution of 1692 samples per class. The decision to use an equal amount of samples per class is in line with recommendations in expert literature (Müller & Guido, 2017). For the vegetation model, the same approach was attempted initially. However, due to poor classification results, a different approach was chosen which includes undersampling classes to a maximum of 30'000 samples per class).

3.7.3 Feature Reduction

Even though RF generally deals well with a large number of input features, it is recommended to perform a feature reduction to minimise dimensionality (number of features) and thus also multi-collinearity (correlating features) and computational cost (Demirbaş Çağlayan et al., 2020; Ma, Fu, et al., 2017). Thereby, redundant features are eliminated. Various feature reduction approaches exist. Generally, two groups can be distinguished (Cunningham, 2008): feature selection and feature extraction approaches. Feature selection describes the process of selecting a subset of predictor variables while feature extraction algorithms create new predictor variables based on several existing predictors (Fassnacht et al., 2016). Feature selection is usually more suited if high feature redundancy is expected (Cunningham, 2008). Therefore, the recursive feature elimination (RFE) algorithm, a feature selection method, was chosen for this study.

RFE was applied using the caret package (Kuhn, 2008). It performs a backward selection of features based on feature importance ranking. The least important features are sequentially eliminated until all remaining features contribute significantly to the model (Kuhn, 2008). RFE was applied individually for the non-vegetation and vegetation model. However, the same parameter settings were used for both models. This includes repeated cross-validation with five repeats and ten folds.

3.7.4 Random Forest Classification

After performing pre-processing, data balancing and feature reduction, two RF classification models (non-vegetation and vegetation) were created using the randomForest package (Liaw & Wiener, 2002) in R software. All RF parameter settings were set to default. Also, based on feature selection results and expert recommendations (see chapter 4.3.1), the 20 most important features were used for each model.

Classification models generally have to deal with the issue of overfitting which occurs if the model is trained to much on a set of input samples but performs worse if samples outside of the initial input dataset are used. Therefore, a frequent measure is to split input data into training and testing (validation) samples (Fassnacht et al., 2016). Different techniques exist to data splitting. Common approaches include simple data splitting (often 70% training, 30% testing data), *k*-fold cross-validation and bootstrap-resampling (Fassnacht et al., 2016). *K*-fold cross-validation describes the process where data is randomly split into *k* equal partitions (folds). Subsequently, the technique iterates *k* times so that in each iteration a different partition is used as testing data while the remaining parts are used as training data. Bootstrap-resampling is a method where over multiple iterations, training samples are selected randomly with replacement. This means that some samples are selected several times while others are not selected. The latter are then used as testing samples (Fassnacht et al., 2016). The RF algorithm uses an internal bootstrap-resampling to generate a forest of decision trees. Samples

not included in bootstrap samples are referred to as out of bag (OOB) samples. Model performance (accuracy) is then assessed on all OOB samples (approximately one third of the total dataset) over multiple iterations (Lebourgeois et al., 2017).

In this study, the internal RF bootstrap-resampling approach was applied in a first run using all available samples as input data for each model. However, results showed unrealistically high classification accuracies (see chapter 4.3.2) which were likely caused by the massive oversampling of certain classes. Therefore, a second validation approach was chosen using a leave-one-out (LOO) cross-validation technique. LOO is a special case of *k*-fold cross-validation where the number of folds equals the number of samples and for each fold, one sample is used as testing data (Fassnacht et al., 2016). In this study, the LOO principle was applied to the nine study regions. Therefore, nine folds were performed excluding samples of one study area for each fold. Furthermore, samples based on GT-independent segmentation (Figure 21, e & f) were used as testing samples in order to incorporate the effect of segmentation accuracy on overall classification accuracy.

For validation of study areas Welschenrohr 1 and 2, features based on orthophoto 1 and 3 (chapter 3.6.2) were ignored as these datasets were missing for the two study areas.

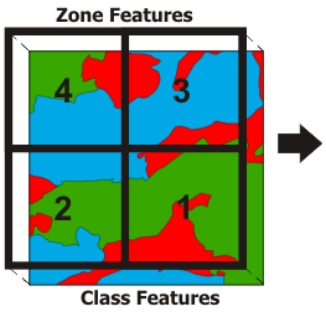
Subsequently, classification was performed for each study area including class prediction and respective prediction probability per image object. The resulting dataset was exported as CSV table and joined in ArcMap to the shapefiles previously exported from eCognition Developer. Thereby, predicted image objects were linked to their geometry using the centroid pixel coordinates of an image object.

3.8 Post-Processing

In ArcMap, classified image objects were subsequently post-processed in order to improve classification accuracy. Post-processing included reclassification of class 10 (dead trees) image objects as class 0 (no SSLE) based on the Max_Z (max. height above ground) feature. This was a very quick but effective improvement for false negatives of this class. A similar procedure could have been applied to other classes. However, this was skipped due to the scope of the thesis.

3.9 Accuracy Assessment

In reference to approaches in previous studies (Gilbert, 2020; Guerrero et al., 2019), classification accuracy of the LOO cross-validation was assessed based on a spatial intersection approach. Thereby, the accuracy of the chosen segmentation approach is also evaluated. Therefore, the Tabulate Intersection tool in ArcMap was used. This tool "computes the intersection between two feature classes and cross-tabulates the area, length or count on the intersecting features" (ESRI, 2022a). As shown in Figure 23, zone and class features are defined initially. The tool then calculates a table with the chosen output variables. For this study, the GT image objects (Figure 24, full opacity) were used as zone features while predicted image objects (Figure 24, reduced opacity) represent class features.



Output

Zone	Color	AREA	PERCENTAGE
1	Blue	0.7	7.9
1	Green	6.3	67.5
1	Red	2.3	24.6
2	Blue	3.4	37.1
2	Green	4.1	44.3
2	Red	1.7	18.7
3	Blue	5.2	56.1
3	Green	1.3	13.9
3	Red	2.8	30.0
4	Blue	5.5	59.3
4	Green	2.0	21.5
4	Red	1.8	19.1

Figure 23: Function of Tabulate Intersection tool (© ESRI)

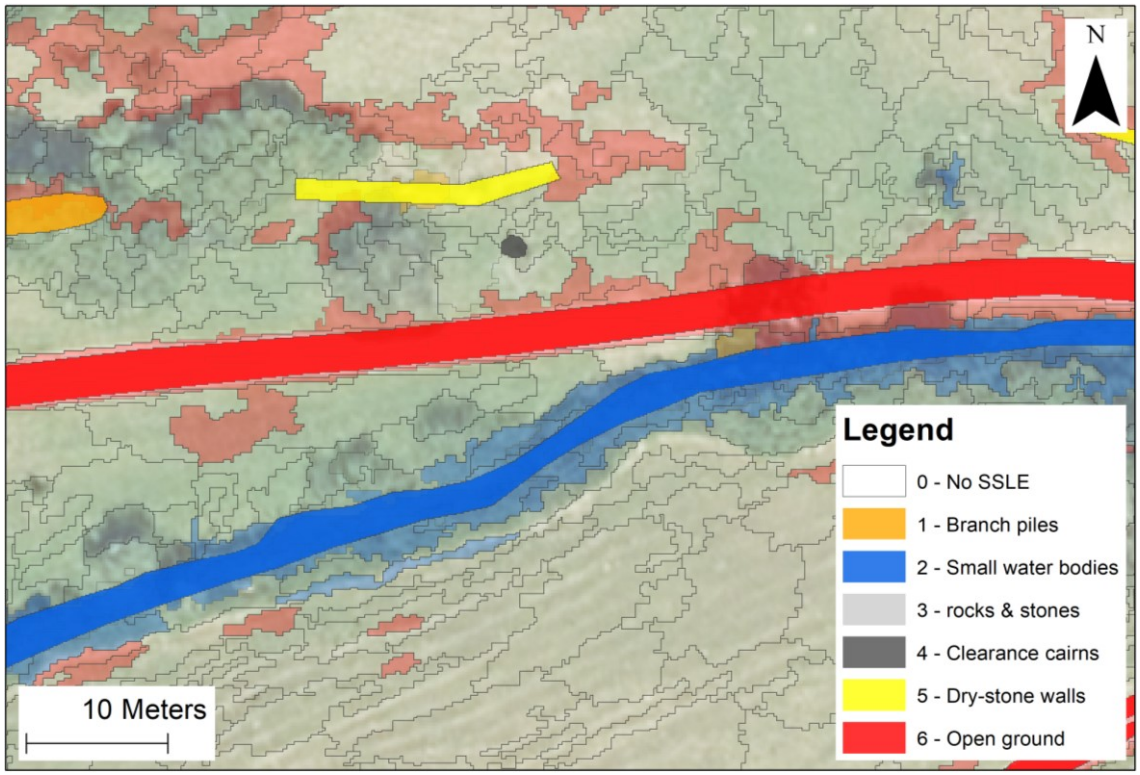


Figure 24: Spatial intersection between GT (full opacity) and predicted image objects (reduced opacity) for the assessment of classification accuracy (© Kanton Aargau)

Thereby, the tabulation indicates for each class the area of the GT polygons that is classified correctly (producer accuracy). Contrarily, it also shows the area of classified image objects that correctly represents GT data (user accuracy).

In the following chapter, the results of this study are presented. This includes an overview of mapped GT data, examples of the chosen image segmentation approach, an overview of the extracted image objects (training & testing samples) as well as feature selection results and RF model accuracy calculated in R software. Finally, the overall assessment of the classification accuracy is presented.

4 Results

4.1 Ground Truth Mapping

Across the nine study areas, a total of 6265 GT polygons were mapped manually. Thereby, class frequency varied substantially (Figure 25). The most frequent element is high woody vegetation (class 9) with more than 2000 samples, followed by rocks and stones (class 3), open ground (class 6) and low woody vegetation (class 9). Substantially less samples were collected for branch piles (class 1), small water bodies (class 2), dry-stone walls (class 5), hedgerows (class 7) and dead trees (class 10). The least mapped class was clearance cairns (class 4) with only 31 cases.

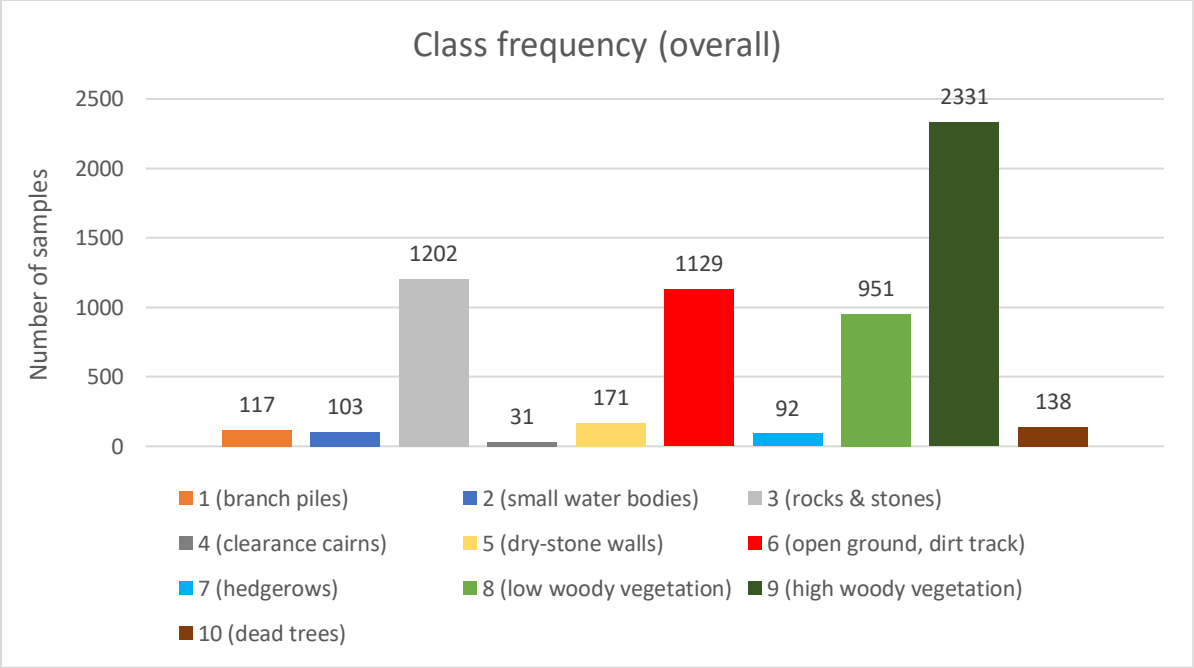


Figure 25: Overall class frequency of mapped GT truth polygons

Among the study areas, the occurrence of different classes is very heterogenous (Figure 26). For the Solothurn areas Welschenrohr 1 and 2 and Seehof, almost no branch piles were mapped. However, these areas contain a disproportionate number of rocks and stones samples. Most of the dry-stone wall samples were also found in these areas. It can be assumed that the unbalanced distribution of these classes did have an effect on classification accuracy as less and older input data was available for these regions. Furthermore, the distribution of low and high woody vegetation and dead trees is rather homogenous across all study areas with lower numbers found in study areas Thalheim 1 and Hornussen (smaller study areas). Also, open ground samples are spread across all areas. Hedgerows were mostly found in the Thalheim areas, Murgenthal and Densbüren. Small water bodies were detected in Wädenswil (drainage ditches), Seehof (small stream) and Densbüren (small streams and amphibian conservation area with ponds). Out of the small amount of clearance cairn samples found, most were located in the Densbüren and Welschenrohr 2 area.

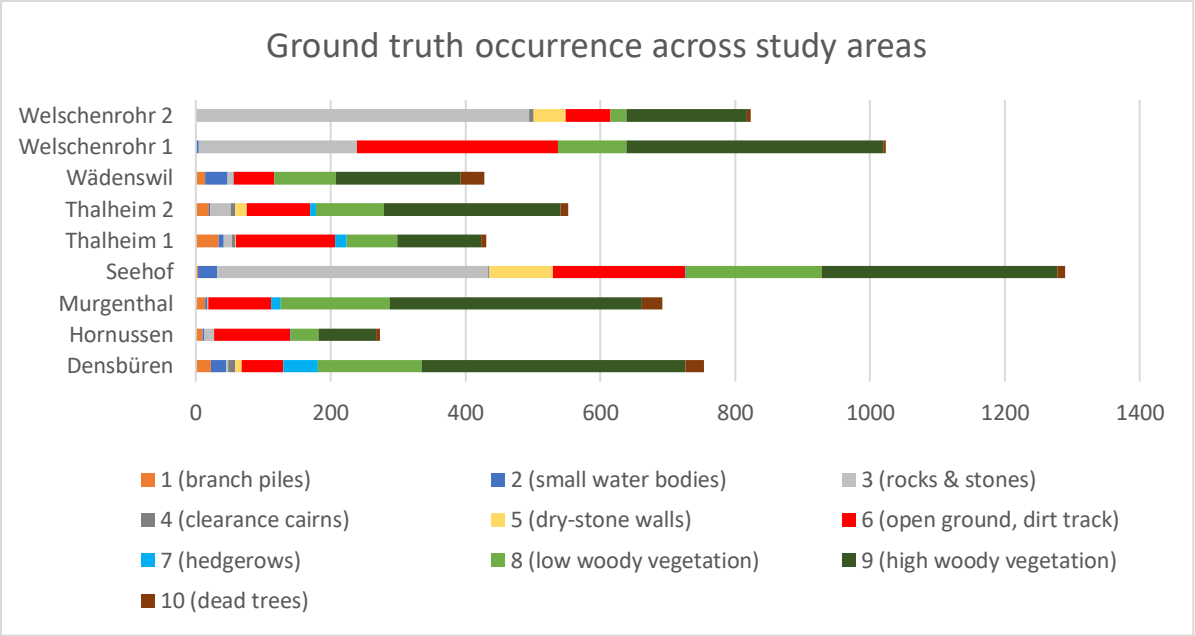


Figure 26: Occurrence of GT classes across the study areas

4.2 Image Segmentation & Feature Extraction

As mentioned in chapter 3.6, mapped GT data was imported to eCognition Developer together with vector data defining the study area, orthophotos, LiDAR and ancillary data for image segmentation and subsequent feature extraction. Based on these inputs, four outputs were generated using four different segmentation approaches (Table 5): segmentation based on geometry and features of non-vegetation (class 1 - 6) GT data (output 1), segmentation based on geometry and features of vegetation (class 7 – 10) GT data (output 2), segmentation independent of GT geometry and parameters optimised for non-vegetation classes (output 3) as well as segmentation independent of GT geometry and parameters optimised for vegetation classes (output 4). Outputs 1 and 2 were used as training samples, outputs 3 and 4 as testing (validation) samples.

On the following pages, only selected examples of segmented areas are presented as entire study areas cannot be shown due to readability reasons. The examples illustrate how training and testing samples are generated. Furthermore, potentials and limitations of the chosen segmentation approach are visualised.

Figure 27 to Figure 30 show segmentation results from the study area Thalheim 2. We can see that the geometry of dry-stone walls (yellow) is generally captured well by the segmentation algorithm. However, elements covered by vegetation are not recognised sufficiently (Figure 27 and Figure 28). For the vegetation classes (Figure 29 and Figure 30), elements are generally delineated very well with some deviations at the edge of tree crowns and bushes. Also, results of the non-vegetation and vegetation testing outputs (Figure 28 and Figure 30) differ due to parameters settings being adjusted to the targeted elements. For the study area Seehof (Figure 31 to Figure 34), image objects show different geometrical characteristics compared to Thalheim 2 as parameters had to be adjusted due to only two orthophotos being available. This was also the case for Welschenrohr 1 and 2. As a result, image objects in these areas are slightly smaller and more compact than others. However, elements are generally delineated well for both non-vegetation and vegetation classes.

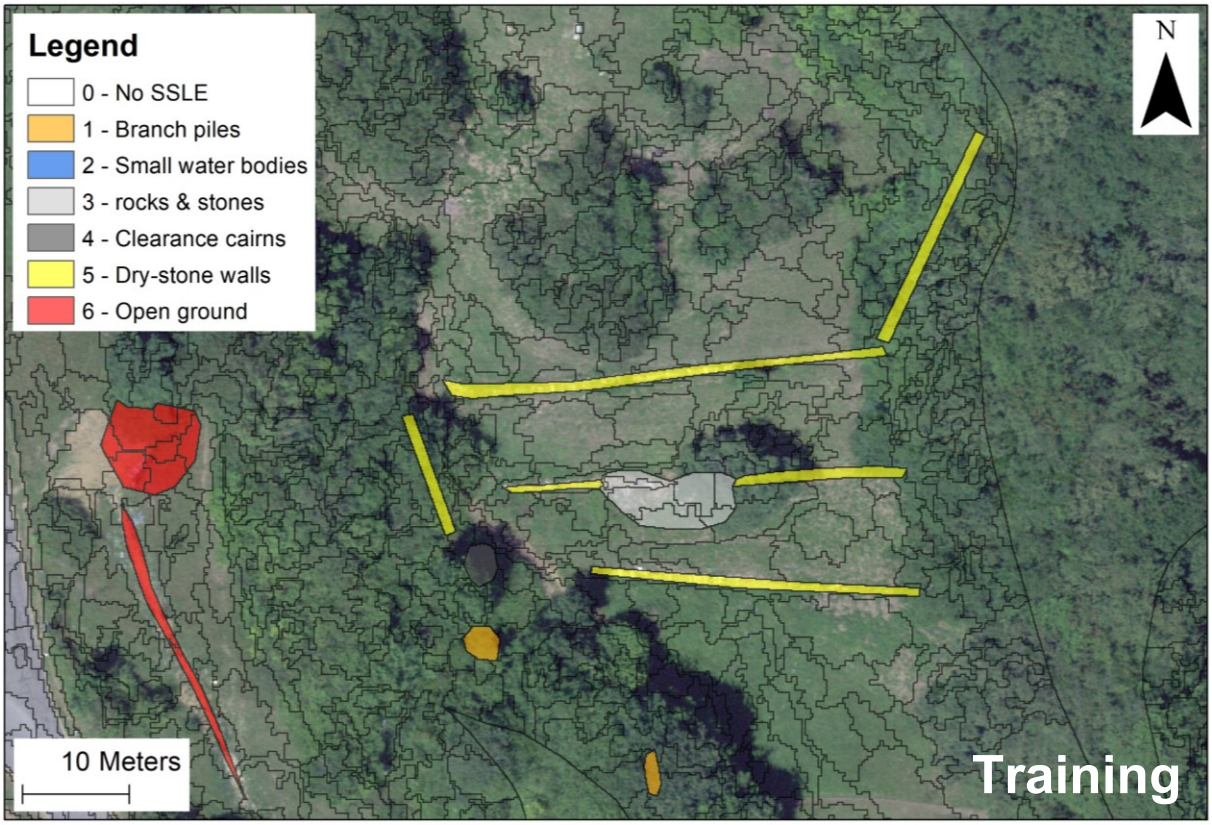


Figure 27: Training samples (non-vegetation), study area Thalheim 2 (© swisstopo)

Figure 28: Testing samples (non-vegetation), study area Thalheim 2 (© swisstopo)

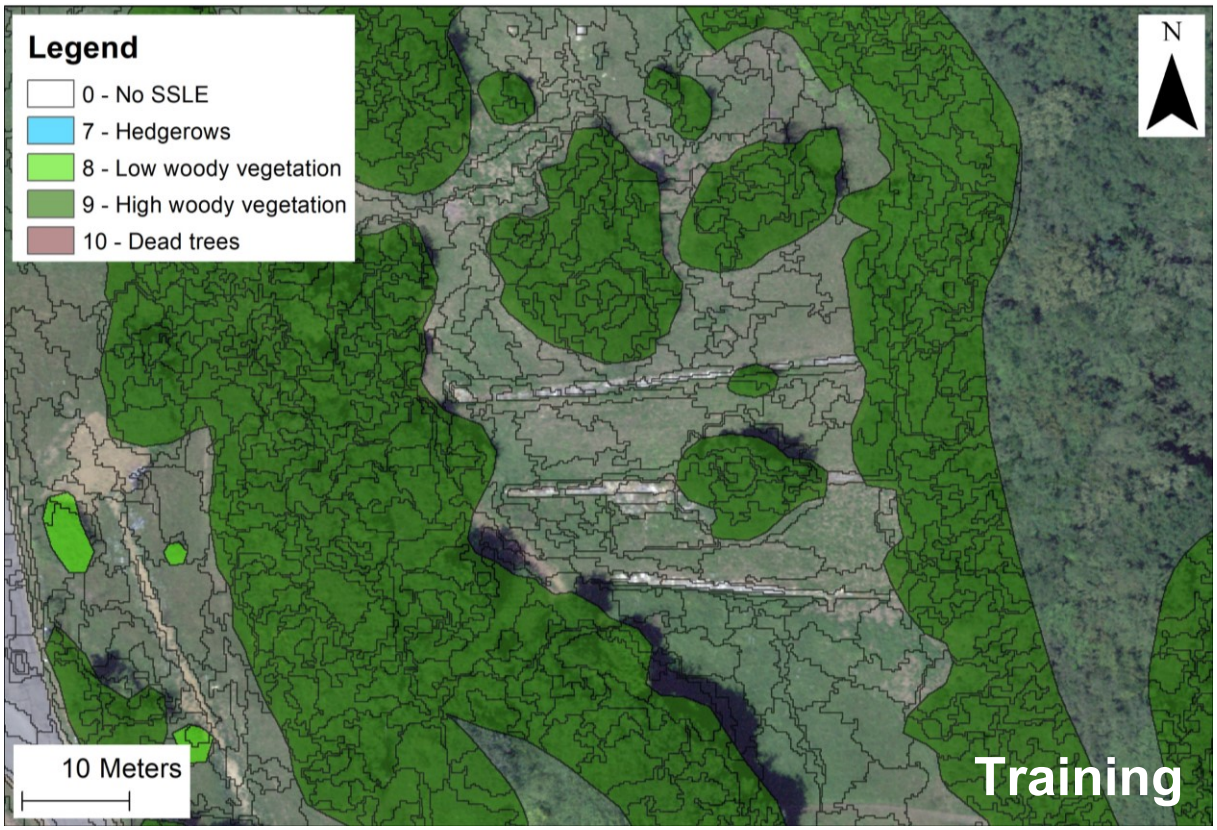


Figure 29: Training samples (vegetation), study area Thalheim 2 (© swisstopo)

Figure 30: Testing samples (vegetation), study area Thalheim 2 (© swisstopo)

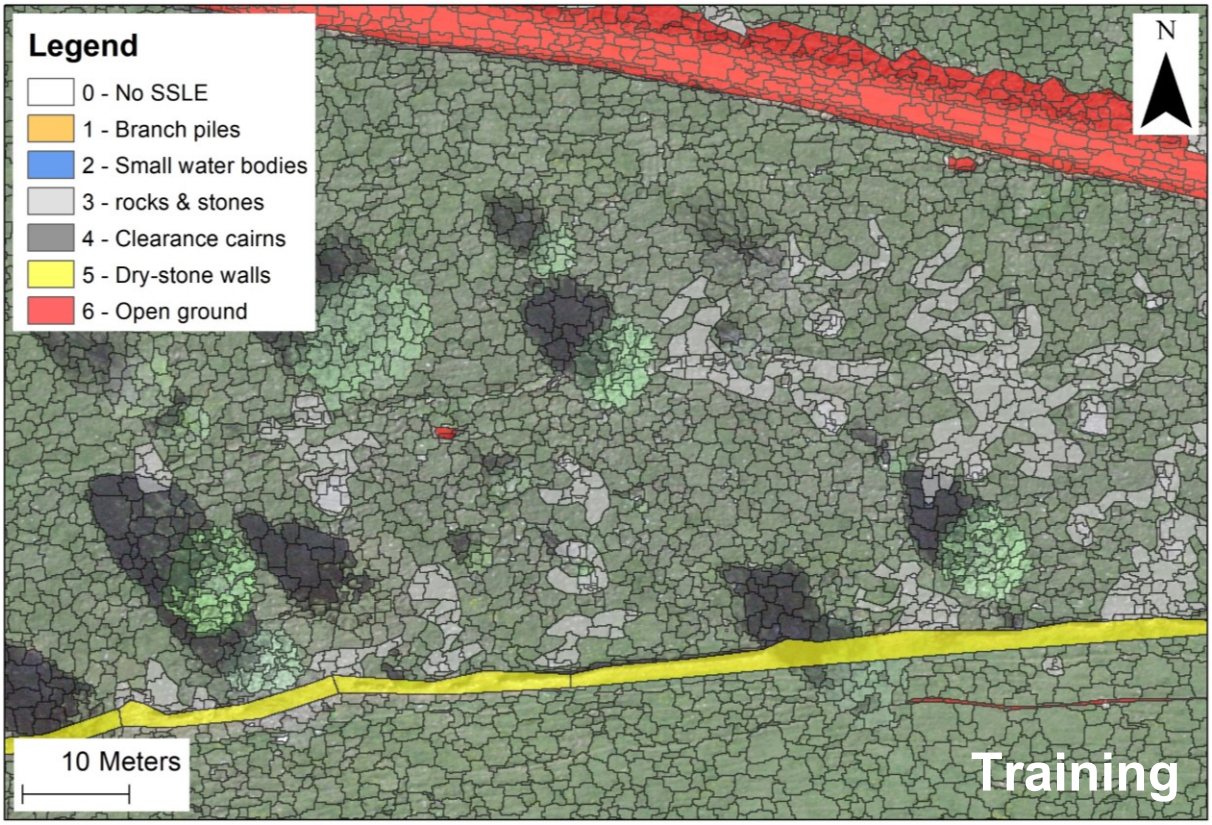


Figure 31: Training samples (non-vegetation), study area Seehof (© swisstopo)

Figure 32: Testing samples (non-vegetation), study area Seehof (© swisstopo)

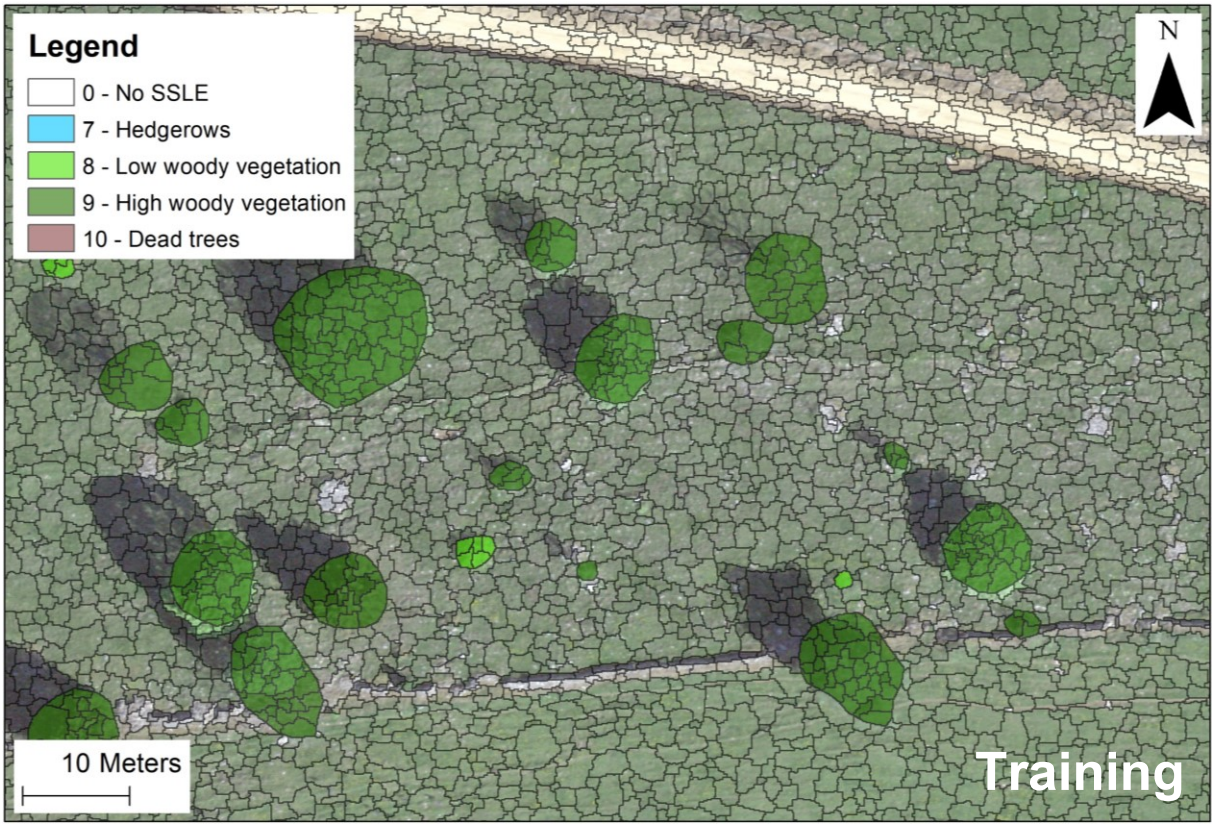


Figure 33: Training samples (vegetation), study area Seehof (© swisstopo)

Figure 34: Testing samples (vegetation), study area Seehof (© swisstopo)

Across all study areas, a total of 784'561 non-vegetation training samples (Figure 35) was generated with 58 features extracted per sample. The vast majority of samples was attributed to class 0 (no SSLE). Branch piles, clearance cairns and dry-stone walls show the same number of samples as after GT polygon creation, because these image objects were not subdivided (chapter 3.6.1). Contrarily, large and connected elements however were sub-segmented resulting in a total of 2'523 samples of small water bodies, 36'808 rocks and stones samples as well as 28'139 open ground samples.

For vegetation class elements (Figure 36), a total of 668'473 samples was created including 489'060 class 0 samples. The number of hedgerow samples remained at 92 (no sub-segmentation) whereas high woody vegetation (4'330 samples), high woody vegetation (148'634 samples) and dead tree polygons (26'357) were further subdivided.

As mentioned in chapter 3.7.2, a strongly unbalanced class distribution often has negative effects on machine learning classification. Therefore, a data balancing technique was applied in order to balance samples and improve overall computational cost by reducing the total amount of samples.

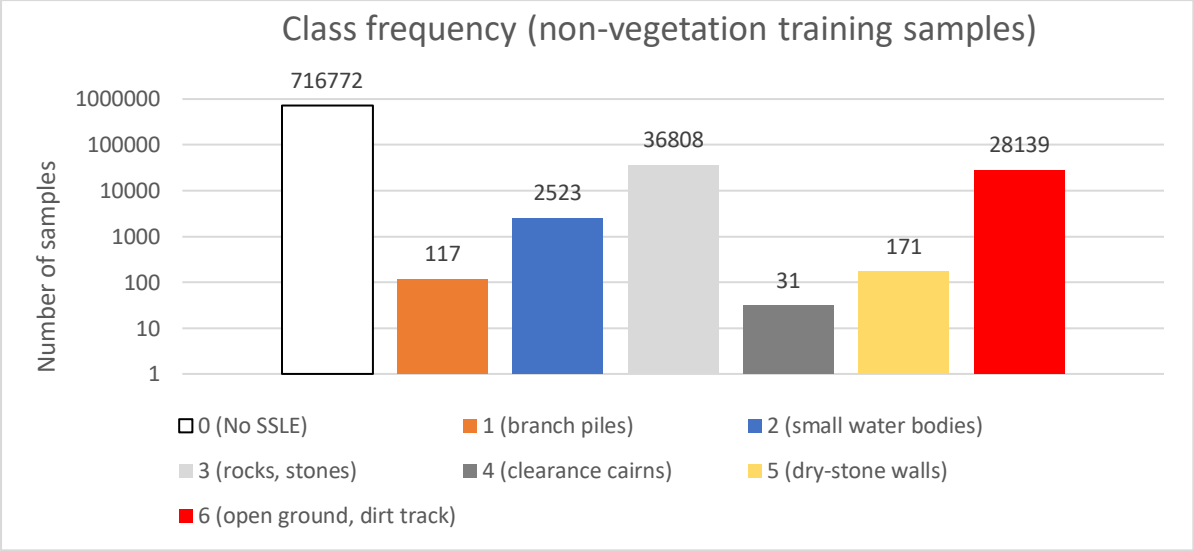


Figure 35: Class frequency of segmented training samples (non-vegetation)

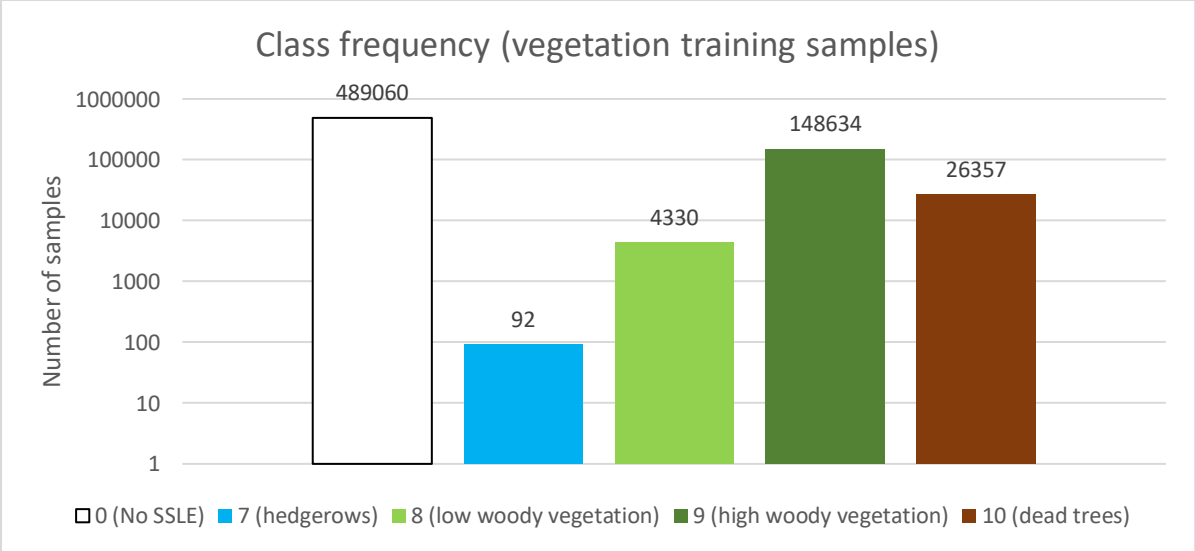


Figure 36: Class frequency of segmented training samples (vegetation)

4.3 Classification

4.3.1 Feature reduction

As described in chapter 3.7.3, the recursive feature elimination (RFE) algorithm was applied in R software to reduce the number of features used to train the classification model. The main goals of the feature reduction were thereby to minimise computational cost and correlating features. RFE iteratively eliminates the least important feature and therefore ranks the features based on their importance for the model (Kuhn, 2008). One result of this process is a performance profile (Figure 37) which indicates the model accuracy in relation to a subset of features (Kuhn, 2008). For the non-vegetation model, the 10 most important features contribute to approximately 84% model accuracy, while the top 20 features lead to 88%. Including all 56 features only improves the accuracy to about 89%. Similar results were achieved for the vegetation model. The top 10 features contribute to 85% accuracy, the top 20 to approximately 87% and the total of 54 features to 91% accuracy. Based on these performance profiles, the decision was made to incorporate only the top 20 features for each of the models, as the increase in performance is insignificant with more features. This decision is in line with expert literature recommending the use of 15 – 25 features for random forest applications (Ma, Fu, et al., 2017). Therefore, computational cost and feature collinearity can be reduced.

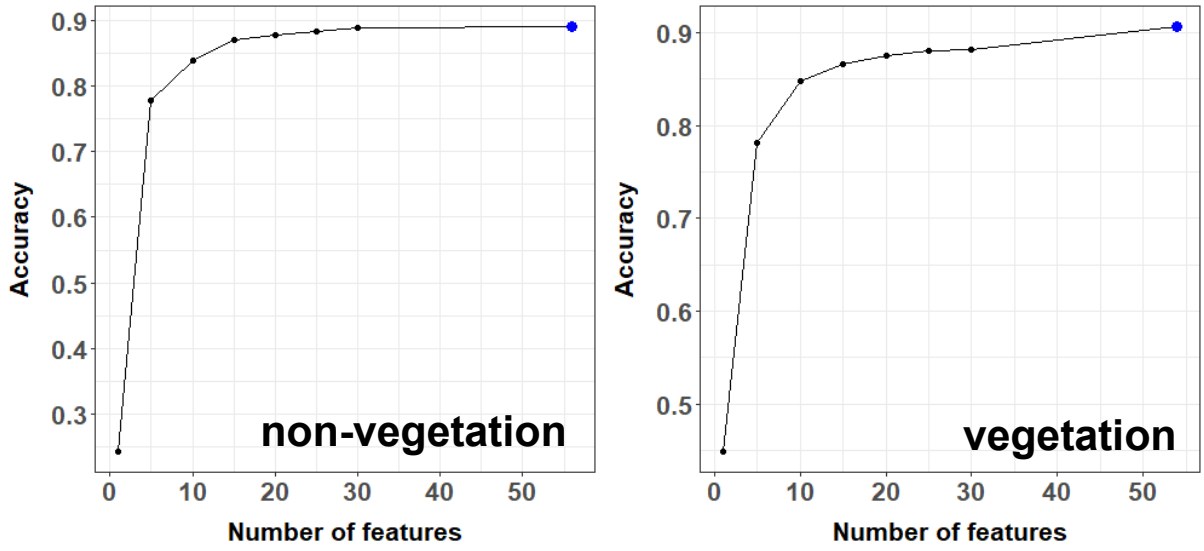


Figure 37: Model accuracy in relation to number of features used (based on RFE algorithm)

Besides performance profiles, RFE also outputs the importance score of each feature based on the ranking. Figure 38 shows the top 20 features for the non-vegetation model. The Mean_vecInd, an ancillary feature, is ranked first with slightly higher importance value than the Density feature (geometry). Considerably less important are all following features including reflectance indices (StdDev_blue2, Mean_nir2, StdDev_green2, Mean_red2, Mean_blue2, Mean_green2 and Mean_blue1), spectral indices (Mean_ndvi2, StdDev_ndsi2, Mean_ndvi1, Mean_ndwi1 and StdDev_ndvi2) as well as geometry features (Curv_Length, Width, Length_main, Compactness and Area) and one texture indices (GLCM_Corr).

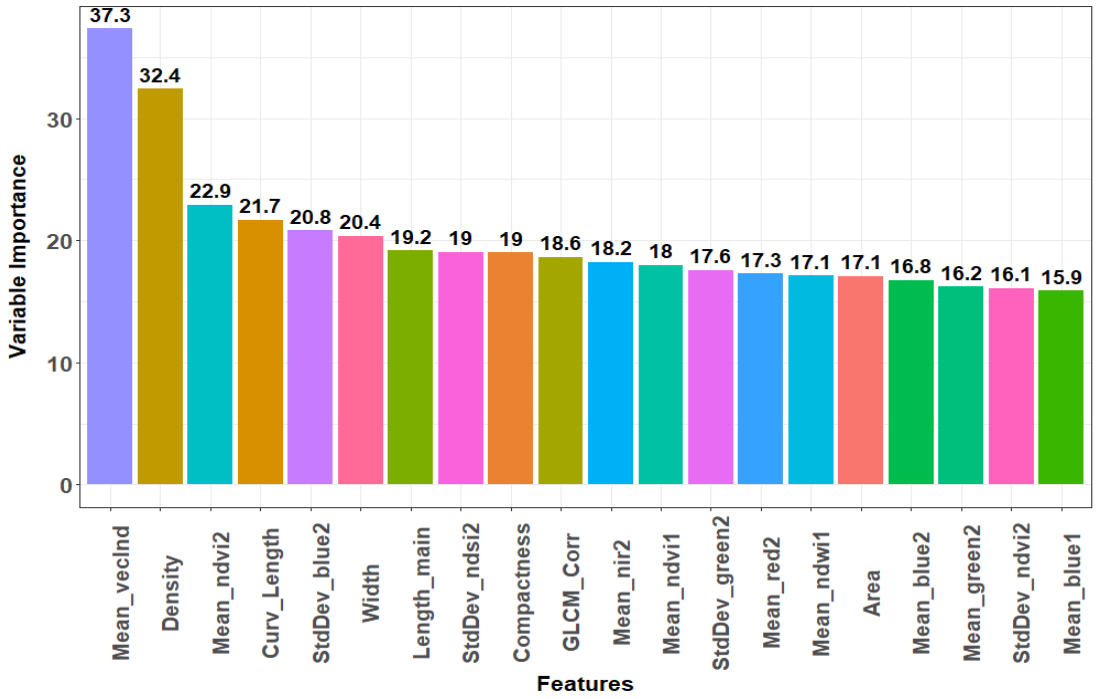


Figure 38: Importance of the top 20 features for the non-vegetation model

For the vegetation model, the top 20 features differ a lot from the non-vegetation model (Figure 39). The most important feature is Mean_vhm, followed by the ALS indices Max_Z, Mean_Z, StdDev_Z and Mean_NoR. Apart from that, reflectance features (Mean_nir2, Mean_blue2, Mean_red2, Mean_nir1, Mean_green2, StdDev_nir2 and StdDev_blue2), spectral indices (Mean_ndsi2, Mean_ndvi2 and Mean_ndvi1), geometrical features (Density, Width and

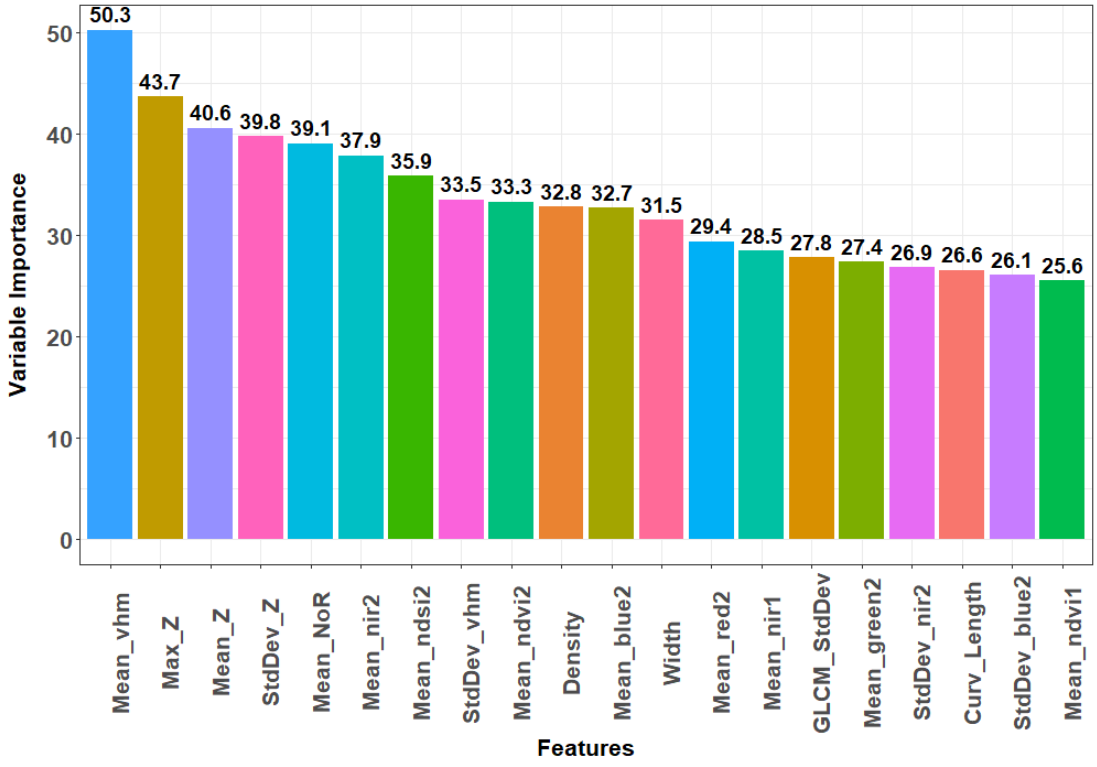


Figure 39: Importance of the top 20 features for the vegetation model

Curv_Length) as well as the ancillary feature StdDev_vhm and the texture indices GLCM_StdDev are also included in the top 20 list. The complete feature ranking for both models can be found in Appendix 2.

4.3.2 Model Accuracy

A common approach to assess the accuracy of ML classification results is the use of confusion matrices (CM) indicating per-class user and producer accuracies, overall accuracy as well as other figures. Thereby, producer accuracy (PA) refers to the probability that a sample in a given class was classified correctly. Contrarily, user accuracy (UA) refers to the probability that a sample predicted to be in a specific class really is that class. Furthermore, overall accuracy (OA) is calculated by dividing the sum of all samples classified correctly by the total number of samples (Ye et al., 2018).

As mentioned in chapter 3.7.4, in a first approach, classification accuracy was assessed with the internal cross-validation of RF models. Thereby, samples that are not used for the creation of individual decision trees, so-called out of bag (OOB) samples, are used as validation samples. For the resulting CM, several accuracy figures were computed in addition to OA, PA and UA. This includes precision, recall and F1-score on both, per-class and macro level. These figures are commonly used if input data has an unbalanced class distribution which reduces the significance of OA, PA and UA. Thereby, "F1-score is the harmonic mean of precision and recall, where precision measures the accuracy of positive predictions and recall measures how many positive samples are correctly detected by the classifier" (Cardoso-Fernandes et al., 2020, p. 8). The commonly used Kappa coefficient was not computed based on expert recommendations indicating that it is often highly correlated to overall accuracy, but can also deliver misleading results (Liu et al., 2007; Pontius & Millones, 2011).

Table 8 shows the resulting confusion matrix for the non-vegetation model. With an OA of 88% and similar macro precision, recall and F1-scores, it can be assumed that the model performs well. However, accuracies for individual classes vary. Among the very well predicted classes

Reference	Prediction								
	0	1	2	3	4	5	6	Sum	PA
0	1'195	8	42	342	0	0	105	1'692	**0.71**
1	0	1'687	0	0	5	0	0	1'692	**1.00**
2	79	8	1'477	81	0	1	46	1'692	**0.87**
3	244	1	8	1'306	0	0	133	1'692	**0.77**
4	0	6	0	0	1'685	0	0	1'691	**1.00**
5	0	0	0	0	0	1'692	0	1'692	**1.00**
6	105	5	84	172	1	1	1'324	1'692	**0.78**
Sum	1'623	1'715	1'611	1'901	1'691	1'694	1'608	11'843	
UA	**0.74**	**0.98**	**0.92**	**0.69**	**1.00**	**1.00**	**0.82**	**Overall Accuracy 0.88**	
Precision	0.74	0.98	0.92	0.69	1.00	1.00	0.82	**Macro Precision 0.88**	
Recall	0.71	1.00	0.87	0.77	1.00	1.00	0.78	**Macro Recall 0.86**	
F1-Score	0.72	0.99	0.89	0.73	1.00	1.00	0.80	**Macro F1-Score 0.88**	

Table 8: Model accuracy (non-vegetation model) based on RF OOB accuracy, illustrated as confusion matrix. Class 0: No SSLE, 1: branch piles, 2: small water bodies, 3: rocks & stones, 4: clearance cairns, 5: dry-stone walls, 6: open ground

is class 1 (branch piles) with 100% (rounded) PA and 74% UA. Even better predicted were class 4 (clearance cairns) and 5 (dry-stone walls), both with 100% (rounded) PA and UA respectively. Notably, class 1, 4 and 5 were massively oversampled during data preparation (see chapter 3.7.2). Among the classes with lower accuracy values are class 0 (no SSLE), class 3 (rocks and stones) as well as class 6 (open ground). For class 3, 14% were misclassified as no SSLE while 8% were falsely considered as open ground. For class 6, 6% were misclassified as no SSLE, 5% as small water bodies and 10% as rocks and stones. For class 2 (small water bodies), 87% PA and 92% UA were recorded. Here, larger misclassifications are assigned to class 0 (5%), class 3 (5%) and class 6 (3%).

Table 9 shows the confusion matrix for the vegetation model. Here, OA as well as other macro values also amount to around 88%. Interestingly, per-class accuracies are relatively homogenous across all classes. Class 10 (dead trees) shows lowest accuracies (88% PA, 80% UA). Larger misclassifications (7%) occurred with class 9 (woody vegetation above 3 m). Contrarily, class 9 was misclassified mainly as dead trees (12%). Furthermore, class 0 (no SSLE) was also misclassified as dead trees in 10% of the cases.

Reference	Prediction					
	0	8	9	10	Sum	PA
0	19'247	583	636	2'206	22'672	**0.85**
8	247	22'067	181	176	22'671	**0.97**
9	417	795	18'773	2'687	22'672	**0.83**
10	627	466	1621	19'958	22'672	**0.88**
Sum	20'538	23'911	21'211	25'027	90'687	
UA	**0.94**	**0.92**	**0.89**	**0.80**	**Overall Accuracy 0.88**	
Precision	0.94	0.92	0.89	0.80	**Macro Precision 0.89**	
Recall	0.85	0.97	0.83	0.88	**Macro Recall 0.88**	
F1-Score	0.89	0.95	0.86	0.84	**Macro F1-Score 0.88**	

Table 9: Model accuracy (vegetation model) based on RF OOB accuracy, illustrated as confusion matrix. Class 0: No SSLE, 8: low woody vegetation, 9: high woody vegetation, 10: dead trees

4.3.3 Classification accuracy

As mentioned in chapter 3.7.4, the model validation based on RF's OOB samples as presented in the previous section shows unrealistically high accuracy values. Therefore, a second validation approach was chosen applying a leave-one-out (LOO) cross-validation technique. In order to assess the accuracy of this classification, a spatial intersection approach was chosen using the Tabulate Intersection tool in ArcMap (see chapter 3.9). Thereby, the tabulation indicates for each class the area of GT polygons that is classified correctly (producer accuracy). Contrarily, it also shows the area of predicted image objects that correctly represent GT data (user accuracy).

Table 10 shows the resulting tabulation for the non-vegetation model which is also illustrated as confusion matrix. Thereby, non-decimal numbers represent the calculated area (m^2) instead of number of samples. NA values refers to image objects that could not be matched to GT samples in ArcMap. Results indicate that OA is similar compared to the OOB confusion matrix. However, macro precision, recall and F1-Score as well as per-class PA and UA are considerably lower. The reason therefore is that almost 97% is non-SSLE area (class 0). As the majority of this class was classified correctly, this has a massive effect on OA.

Reference	Prediction									
	0	1	2	3	4	5	6	NA	Sum	PA
0	4'297'853	17'077	109'389	235'469	1'652	260	228'500	611	4'890'811	**0.88**
1	978	46	64	17	2	0	82	0	1'188	**0.04**
2	2'139	43	7'806	161	0	0	317	132	10'597	**0.74**
3	16'923	56	209	32'259	60	13	3'755	10	53'284	**0.61**
4	105	2	9	85	0	0	126	0	327	**0.00**
5	1'806	19	38	2'356	3	8	751	2	4'983	**0.00**
6	18'454	202	7'470	5'566	80	6	61'270	9	93'057	**0.66**
Sum	4'338'257	17'444	124'985	275'912	1'798	287	294'800	764	5'054'247	
UA	**0.99**	**0.00**	**0.06**	**0.12**	**0.00**	**0.03**	**0.21**	**0.00**		**OA 0.87**
Precision	0.99	0.00	0.06	0.12	0.00	0.03	0.21			**Macro Precision 0.20**
Recall	0.88	0.04	0.74	0.61	0.00	0.00	0.66			**Macro Recall 0.42**
F1-Score	0.93	0.00	0.12	0.20	0.00	0.00	0.32			**Macro F1-Score 0.22**

Table 10: Classification accuracy (non-vegetation model) based on spatial cross-tabulation (m^2). Class 0: No SSLE, 1: branch piles, 2: small water bodies, 3: rocks & stones, 4: clearance cairns, 5: dry-stone walls, 6: open ground

Reference	Prediction							
	0	7	8	9	10	NA	Sum	PA
0	4'050'044		16'523	137'987	49'633	283	4'254'470	**0.95**
7	6'252		2'052	22'506	12'241	2	43'053	
8	3'265		5'881	2'044	684	0	11'873	**0.50**
9	51'425		5'313	649'635	35'047	28	741'449	**0.88**
10	436		12	1'440	1'426	0	3'315	**0.43**
Sum	4'111'422		29'782	813'613	99'031	313	5'011'107	
UA	**0.99**		**0.20**	**0.80**	**0.01**	**0.00**		**OA 0.97**
Precision	0.99		0.20	0.80	0.01			**Macro Precision 0.50**
Recall	0.95		0.50	0.88	0.43			**Macro Recall 0.69**
F1-Score	0.97		0.28	0.84	0.03			**Macro F1-Score 0.53**

Table 11: Classification accuracy (vegetation model) based on spatial cross-tabulation (m^2). Class 7 (orange) was ignored for classification. Class 0: No SSLE, 7: hedgerows, 8: low woody vegetation, 9: high woody vegetation, 10: dead trees

Several classes were basically not detected. This includes class 1 (branch piles) where less than 4% of the area was classified correctly. Most branch piles were misclassified as non-SSLE (class 0, 83%), open ground (class 6, 7%) or small water bodies (class 2, 5%). Even worse results were achieved for clearance cairns (class 4) where no sample was classified correctly. They were mostly misclassified as non-SSLE (83%), rocks & stones (class 3, 26%) or open ground (class 6, 39%). Similarly low accuracy (<1%) was obtained for dry-stone walls (class 5). Here, most misclassifications occurred in combination with non-SSLE (36%), rocks & stones (47%) or open ground (15%). Better classified were small water bodies where 74% of the area was recognised correctly. However, UA is 6% which that this class was greatly overclassified. Overclassification of small water bodies occurred mainly in non-SSLE area. For rocks & stones, similar results were achieved including 61% PA and 12% UA. Here, misclassifications also occurred mainly in non-SSLE areas. Finally, open ground was classified with 66% PA and 21% UA. Similar to small water bodies and rocks and stones, open ground was also often mapped in non-SSLE areas.

Table 11 shows the cross-tabulated results for the vegetation model which achieved 97% OA, 50% macro precision, 69% macro recall and 53% macro F1-Score. Per-class accuracies were considerably better than those for non-vegetation classes. Class 8 (low wood vegetation) achieved 50% PA and 20% UA. It was mostly misclassified as non-SSLE (27%) or high woody vegetation (class 9, 17%). High woody vegetation achieved very good results with 88% PA and 80% UA. Highest misclassification occurred with non-SSLE area (7%) and dead trees (class 10, 5%). Dead trees show lowest accuracy values with 43% PA and only 1% UA. This means that much area was wrongly overclassified as dead trees. This includes non-SSLE area (50%) and high woody vegetation (35%). Finally, class 7 (hedgerows) were not further considered in the cross-tabulation as class 7 was not used for the RF models (see chapter 3.7.1).

On the following pages, classification results from the two example areas in Thalheim and Seehof are illustrated with thematical maps (Figure 40 - Figure 47). These illustrations serve as a basis for further discussion (see chapter 5.2).

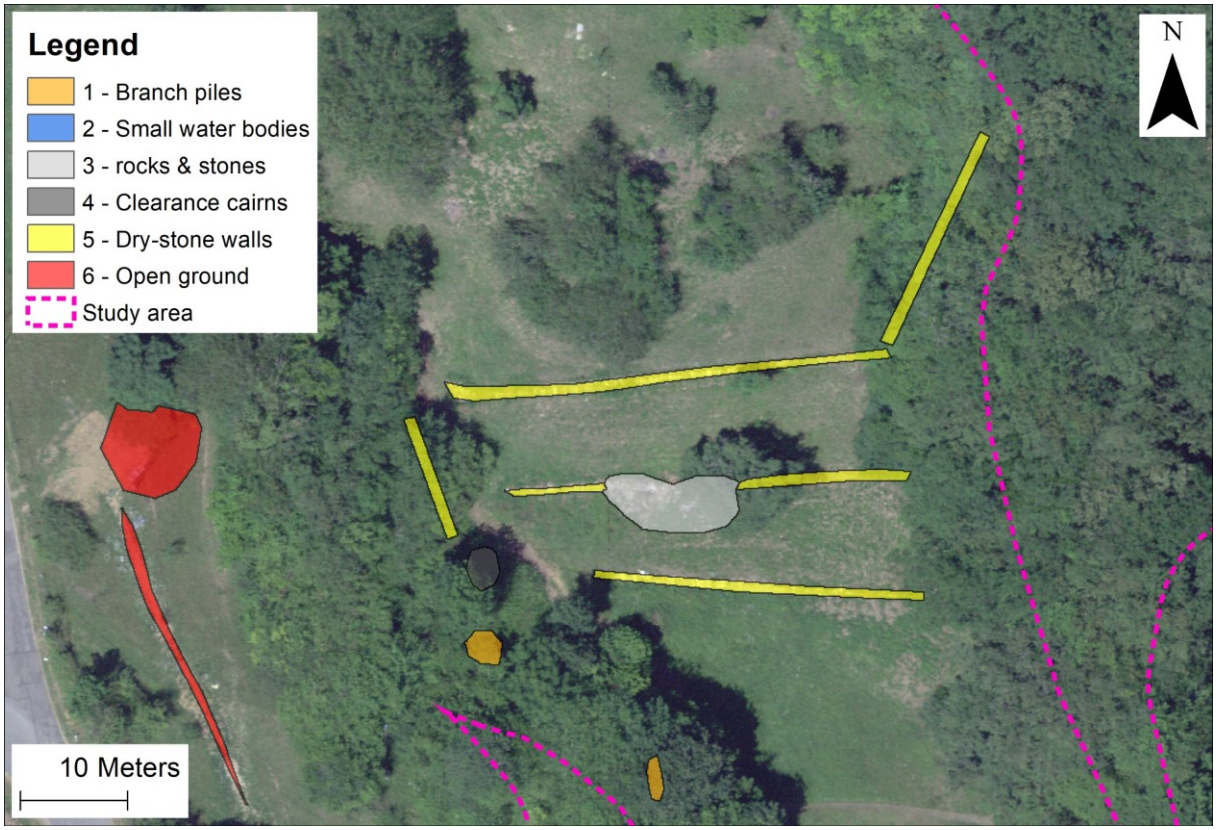


Figure 40: GT samples (non-vegetation), study area Thalheim 2 (© swisstopo)

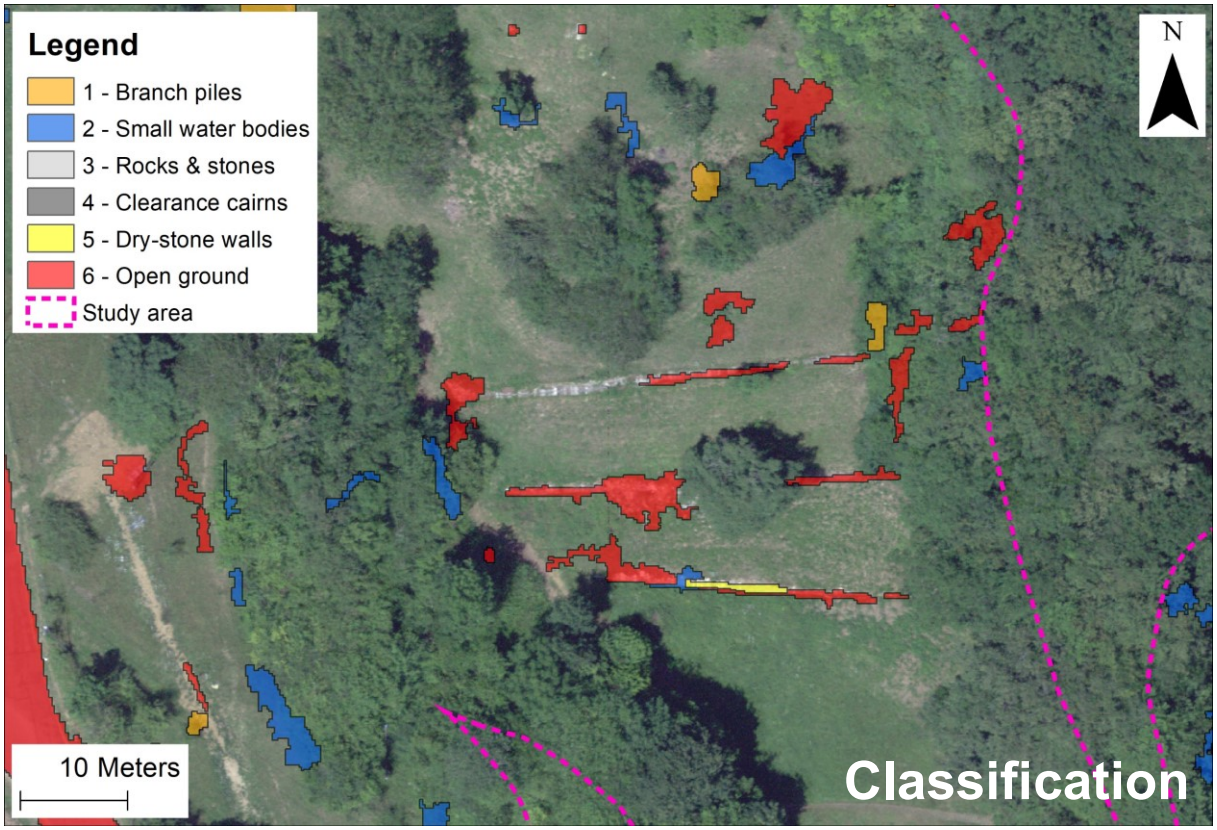


Figure 41: Classification results (non-vegetation), study area Thalheim 2 (© swisstopo)

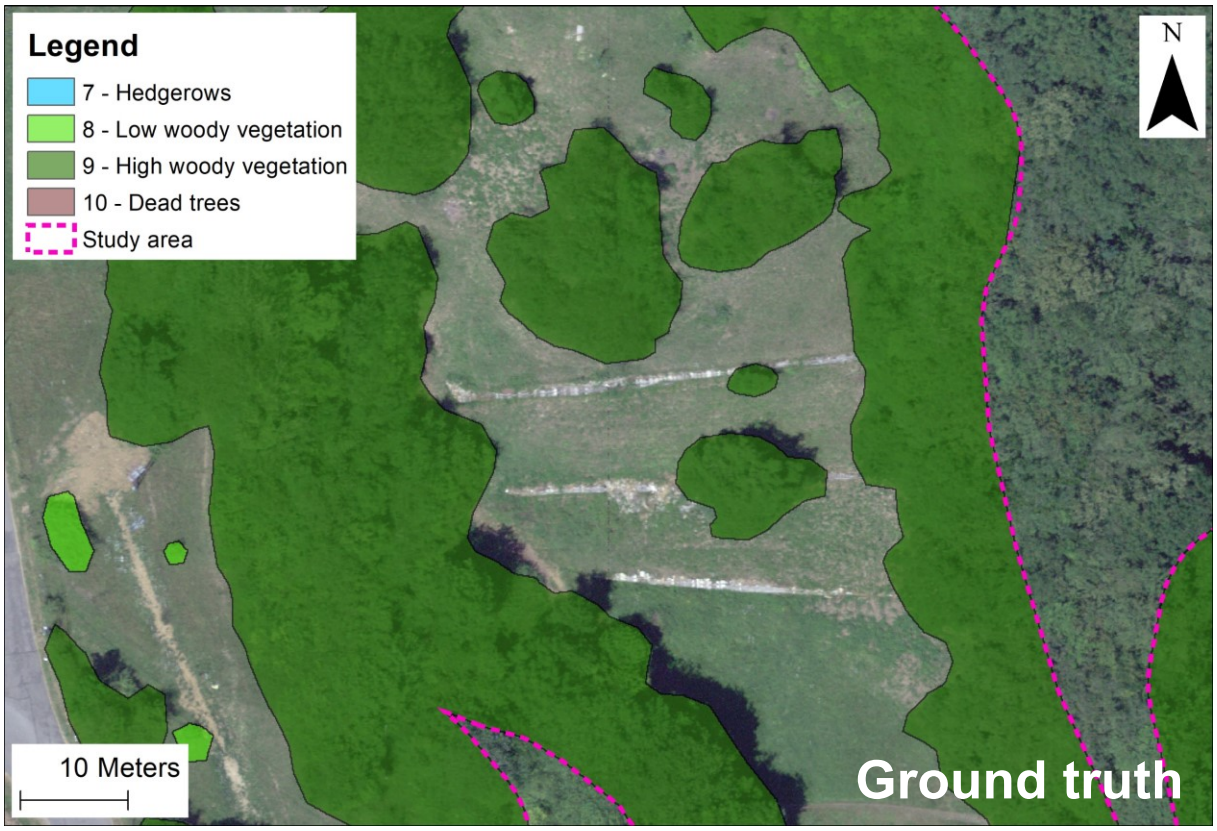


Figure 42: GT samples (vegetation), study area Thalheim 2 (© swisstopo)

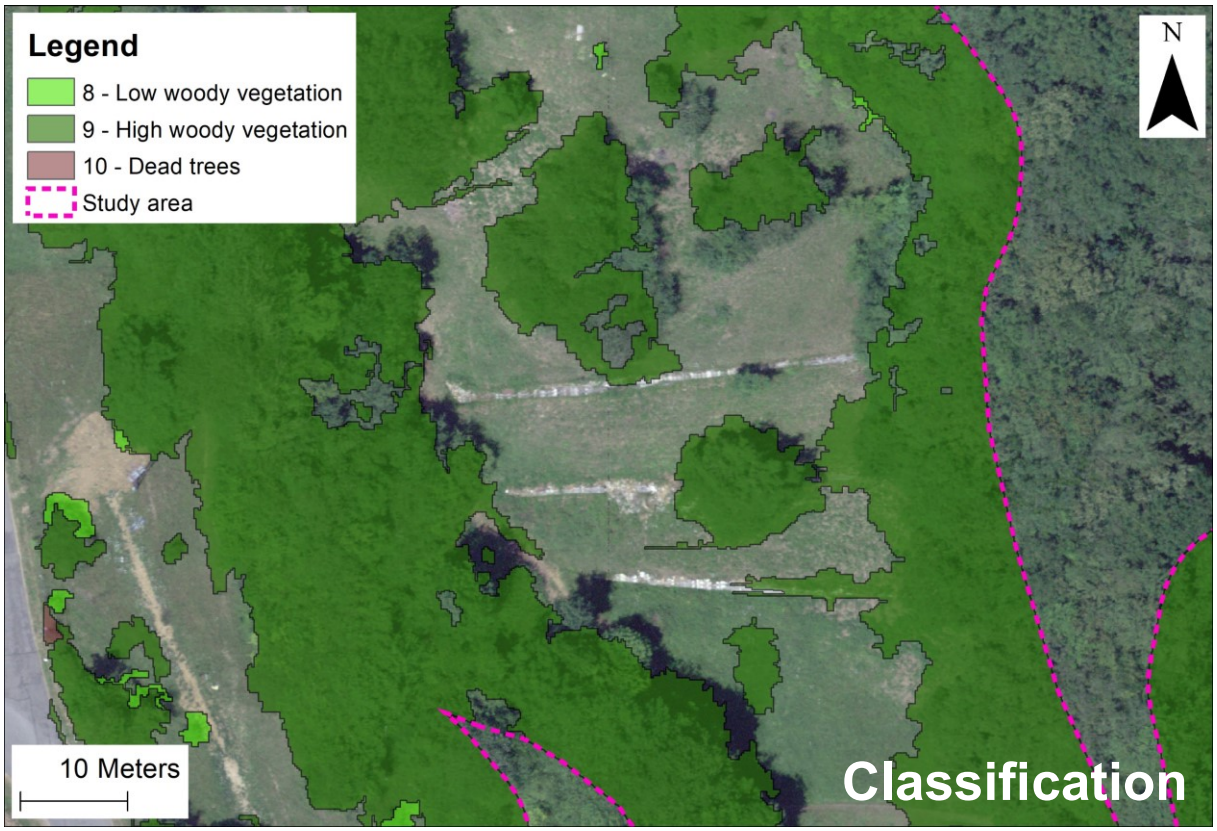


Figure 43: Classification results (vegetation), study area Thalheim 2 (© swisstopo)

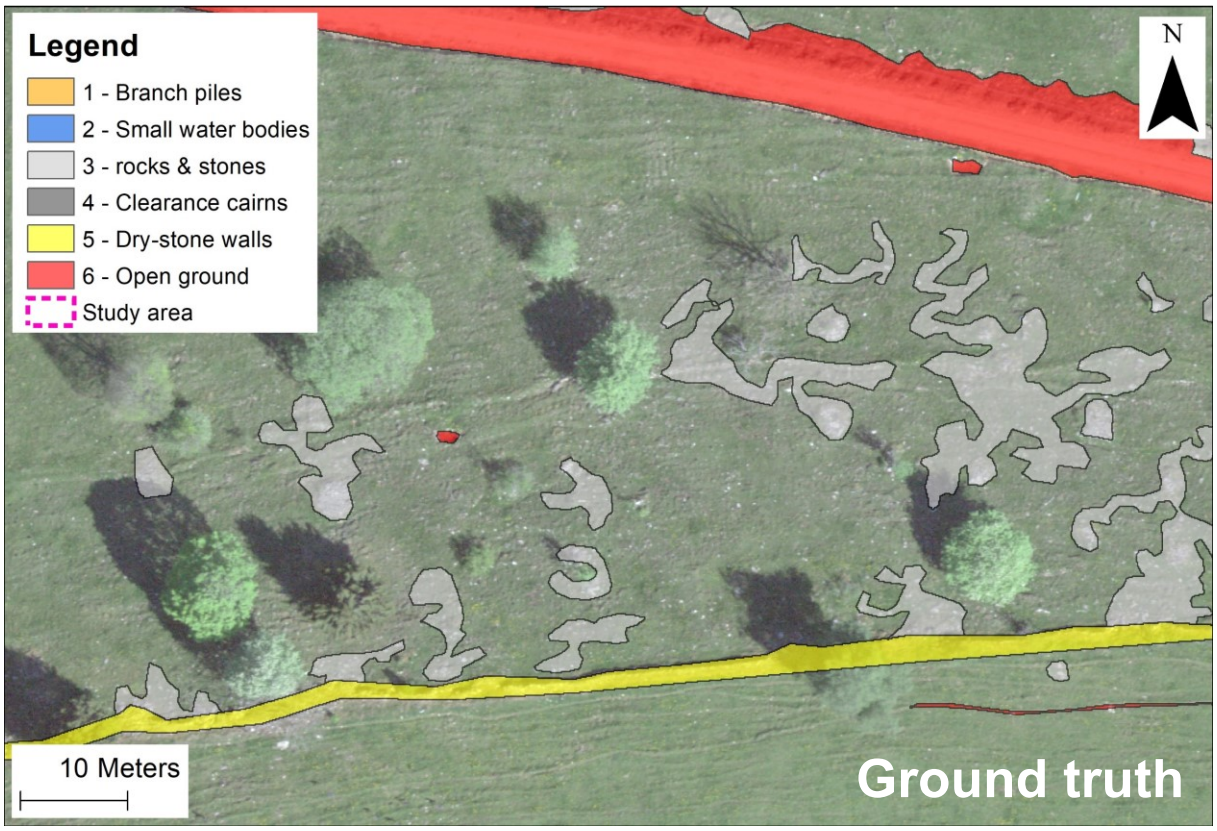


Figure 44: GT samples (non-vegetation), study area Seehof (© swisstopo)

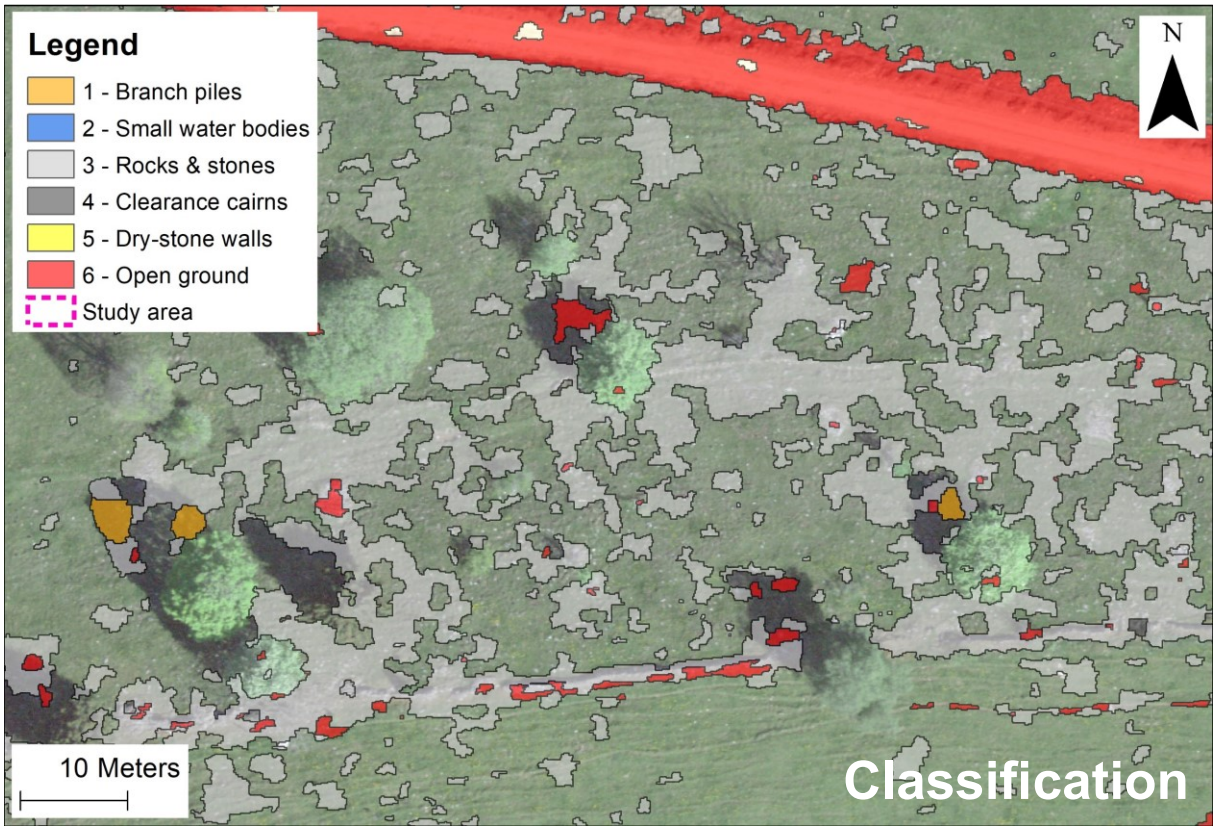


Figure 45: Classification results (non-vegetation), study area Seehof (© swisstopo)

Figure 46: GT samples (vegetation), study area Seehof (© swisstopo)

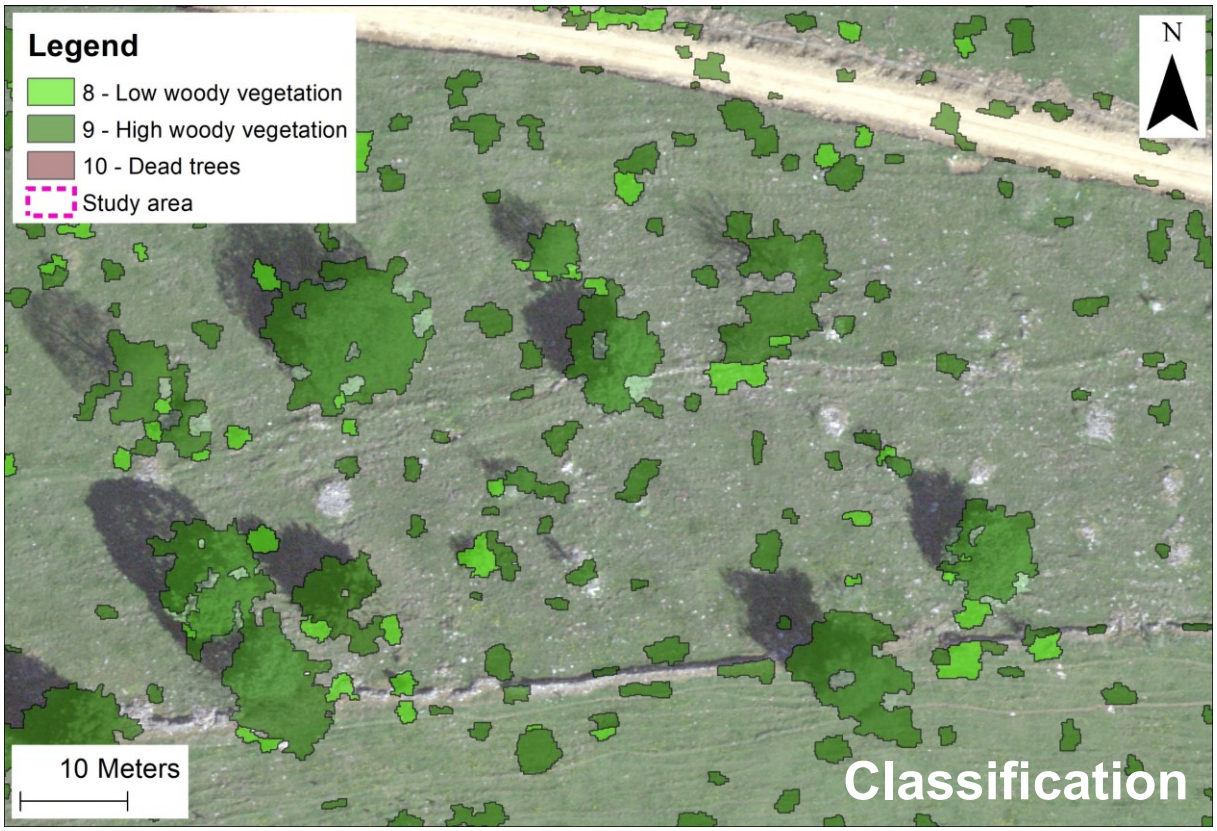


Figure 47: Classification results (vegetation), study area Seehof (© swisstopo)

5 Discussion

In this chapter, the chosen approach and the findings of this study are discussed. Thereby, potentials and limitations are listed. The following chapter covers implications of the chosen ground truth mapping approach. Subsequently, classification results for all selected small-scale landscape element (SSLE) classes, the suitability of high-resolution remote sensing data therefore as well as the implications of the chosen object-based image analysis (OBIA) methods are discussed. Finally, an outlook for future studies is presented.

5.1 Implications of Ground Truth Mapping

Ground truth (GT) mapping in this study was performed due to lack of sufficient and homogenous reference data. Ten SSLE classes of interest were selected and subsequently mapped based on visual interpretation of remote sensing (RS) data and additional in-situ validation. A key aspect thereby is semantics: How do classes have to be defined in order to be distinguishable 1) visually and on the basis of remote sensing data and 2) by classification methods? As these criteria were not fully met for certain common SSLE found in literature, definitions had to be adjusted accordingly. However, differentiation between certain classes still remained challenging in some cases. One example therefore are classes 3 (rock, stone, debris, boulders), 4 (clearance cairns) and 5 (dry-stone walls). Figure 48 shows a case where assumably, an old dry-stone wall had collapsed now resembling a series of clearance cairns or rocks and stones. In this case, the object was mapped as class 3 as the height criteria for clearance cairns (0.5 m) was not fully met. However, this interpretation is debatable. In such cases, the LiDAR point cloud was consulted for height information during visual interpretation of RS data. Naturally, this includes the limitation that not all elements are depicted by LiDAR points, particularly if an element is covered by vegetation. Therefore, in-situ validation is very valuable. However, a key issue between visual interpretation and in-situ validation is the temporal lag. For this study, the field validation was conducted in spring 2021 while the most recent RS dataset originated from 12 months earlier (the least recent was from 2016). Therefore, some elements have disappeared in between (e.g. cut down trees), changed their characteristics (e.g. tree growth: trees that can be classified as either low or high woody vegetation depending on what input dataset is consulted) or also, new elements can appear over time (e.g. recently constructed branch piles or clearance cairns). Ultimately, such temporal inconsistencies lead to issues where parts of the input data inaccurately represent the element of interest. This will be further discussed in chapter 5.3.

Figure 48: Collapsed dry-stone wall: What is it today? Dry-stone wall, a series of clearance cairns or just rocks? (© Simon Mägli)

Ground truth data was collected across a total of nine study areas which are situated in the cantons of Aargau, Solothurn and Zurich. The areas represented the biogeographical regions “Jura” and “Central Plateau”. Agricultural use ranged from very extensively used grassland to high-intensity arable land. This is an acceptable diversity of agricultural landscapes. However, some regions including alpine regions and specialised farming areas such as vineyards were not represented. This would be a requirement if the classification model should be applied for

the entirety of Switzerland. Nevertheless, a considerable number of elements was mapped across all study areas. In total, 6265 samples were recorded with varying frequency across different study areas and also, different classes (see chapter 4.1). Well represented classes (>900 samples) are class 3 (rock, stone, boulder, debris), class 6 (open ground), class 8 (low woody vegetation) and class 9 (high woody vegetation). Unfortunately, relatively few (<200) samples were recorded for class 1 (branch piles), class 2 (small water bodies), class 5 (dry-stone walls), class 7 (hedgerows) and class 10 (dead trees). An insufficient number of samples (31) was found for class 4 (clearance cairns). Due to this uneven distribution, data balancing was necessary which had assumably major effects on classification accuracy (further discussed in chapter 5.4). Therefore, for future studies it is suggested to seek more homogenous class distribution.

Finally, a key challenge that occurred during GT mapping was the fact that certain elements could overlap each other in real-world. For example, rock elements could be covered be trees (see Figure 20). This posed an issue insofar that with the chosen segmentation approach, GT samples must be unambiguously assignable to a specific class. Therefore, two GT datasets (vegetation and non-vegetation) were created in this study. An alternative approach would be to apply multi-level segmentation where vegetation elements are segmented on a super-level and non-vegetation elements are delineated on a sub-level.

In the following chapter, classification results for all selected types of SSLE are discussed.

5.2 Classification Accuracy of Small-scale Landscape Elements

Overall, classification results for the ten selected SSLE classes are unsatisfactory. However, the classification accuracy varies substantially depending on class and study area (see Table 10 and Table 11). Generally, we can see a trend where larger, connected elements such as woody vegetation or open ground areas achieved higher accuracy values whereas smaller, isolated elements such as branch piles or clearance cairns were barely recognised (<7%).

Best classification results were achieved for class 9 (high woody vegetation) with 88% producer accuracy (PA) and 80% user accuracy (UA). This result was expected as class 9 features the highest number of training samples (~150'000) and also highest area coverage (0.74 km^2). Additionally, high woody vegetation is present in all study areas and has characteristics that make it very distinct from other elements. This includes the height above ground as well as spectral features such as the NDVI. Figure 42 and Figure 43 show that the majority of high woody vegetation was classified correctly. In particular, coherent structures including forest edges and groves were recognised well. However, transitions between woody vegetation and openland are often classified poorly. We can see that there are often artefacts in both openland and woody vegetation area. A solution therefore would be to apply a region-growing or -merging algorithm based on probability values. Interestingly, classification accuracy for the Seehof study area (see Figure 47) was much lower as many openland samples were classified as high woody vegetation. For this, there are three possible explanations. First of all, for Solothurn areas including Seehof, Welschenrohr 1 and 2, less and also more out-of-date input data was available (see chapter 3.3). Therefore, segmentation parameters had to be adjusted in order to achieve image objects which are similar to those in other study areas. Secondly, the bit depth of reflectance values for Solothurn areas differed a lot from other study areas. Even though reflectance values across all study areas were subsequently scaled applying data standardisation per individual study area, a certain inaccuracy can still be expected for the Solothurn regions. Thirdly, it is possible that misclassification occurred based on the Mean_VHM (average of the Vegetation Height Model) and Max_Z feature (max. height above ground) which were essential for the classification of vegetation elements (see Figure 39). The

third explanation is assumably the most likely as an inspection of LiDAR point cloud data for this area has shown that there are indeed many artefactual points (several meters above ground) in the openland area. However, also a combination of the three factors mentioned above is possible. Overall, achieved results for class 9 are similar to results from other studies analysing vegetation. Bolyn et al. (2019) reported 78% overall accuracy (OA) for different classes of trees outside forests (TOF). Malkoç et al. (2021) even achieved 95% overall accuracy for countrywide mapping of TOF across Switzerland. Furthermore, Völker & Müterthies (2008) reported 100% PA and 70% UA for linear woody vegetation structures.

The second best classified element is class 6 (open ground) which achieved 66% PA and 21% UA. Notably, non-SSLE samples were often mistaken as open ground area. This misclassification is obvious as a lot of the non-SSLE area consists of highly dynamic surfaces such as arable land. Additionally, open ground area by itself is often highly dynamic. Figure 40 shows that elements with reflectance values similar to those of open ground are often mistaken as such. This occurred with dry-stone walls or rocky surfaces. Furthermore, asphalted road was often also mistaken as open ground. Contrarily, gravel roads (see Figure 44 and Figure 45) were generally recognised well. It can be assumed that this is also based on the fact that gravel roads are large, coherent surfaces. Unfortunately, no comparable classification results from other studies were found for open ground.

Class 6 is followed by class 8 (low woody vegetation) which achieved 50% PA and 20% UA. Although class 8 has similar characteristics to class 9, only 4330 training samples covering a total area of 0.1 km^2 were available. This could be a reason for the considerably lower classification accuracy. As expected, low woody vegetation was often mistaken as high woody vegetation (17%) as the only criterium to distinguish those classes is the height information (Max_Z feature). Apart from high woody vegetation, class 8 was often misclassified as non-SSLE area. Conversely, 55% of the area classified as class 8 was in fact non-SSLE area. Figure 42 and Figure 43 underline the fact that class 8 was often mistaken as class 9. Naturally, class 8 is affected by the temporal lag between GT and input data characteristics (see chapter 5.1).

Another class that showed mediocre results was class 10 (dead trees) with 43% PA and only 1% UA. This means that dead trees were massively overclassified. Most notably, openland non-SSLE area was misclassified as dead trees. This effect was already greatly reduced through post-processing (see chapter 3.8). However, the result is still unsatisfactory. The presumed reason for the overclassification of dead trees is the impact of spectral features which might be similar to those of openland. However, these should have been easily distinguishable based on height information. Interestingly, most of the areas mistaken as dead trees are assigned to one specific study area (Densbüren). This indicates potential computation mistakes or differing characteristic of the input data for this study area.

Massive overclassification also occurred with class 2 (open water bodies). This class achieved 74% PA but only 6% UA. This is perfectly visible in Figure 40 and Figure 41 where numerous openland, woody vegetation and even dry-stone wall samples were misclassified as small water bodies. Several factors could have contributed to this issue. First of all, class 2 samples were extensively oversampled during data balancing (see chapter 3.7.2). From 103 original GT samples, 1692 training samples were artificially created. Secondly, many original class 2 GT samples contained “misleading” information insofar that water bodies were often covered by vegetation (see Figure 49). As a result, spectral characteristics of vegetation instead of water surfaces are passed on.

Figure 49: Class 2 sample (blue) being covered by vegetation (© swisstopo)

Therefore, it might be beneficial to only use training data from water bodies that are not covered by vegetation. However, this might also lead to many small water bodies not being detectable at all.

Classes that were barely detected at all were class 1 (branch piles), class 4 (clearance cairns) and class 5 (dry-stone walls). Branch piles achieved 4% PA and <1% UA. 99% of the area classified as branch piles was in fact non-SSLE area. Conversely, most branch pile area (82%) was classified as non-SSLE. The key challenge for branch piles was the fact that most of them were covered by vegetation. This results in false spectral information as explained above for small water bodies. Unfortunately, the use of ancillary features based on LiDAR point cloud intensity values (Mean_intInd and StdDev_intInd) which were specifically designed for the recognition of branch piles (see Figure 22) was not sufficient. Figure 40 and Figure 41 as well as Figure 44 and Figure 45 show several examples of branch piles falsely classified in open-land or forest area. High LiDAR intensity values might be the reason for those misclassifications. For the Seehof examples, another possibility is that small stone piles were mistaken as branch piles. Finally, the issue of extensive oversampling is also applicable for branch piles.

Another class that was almost completely misclassified is class 5 (dry-stone walls) with <1% PA and 3% UA. Naturally, most dry-stone wall samples were classified either as rocks and stones (47%) or as non-SSLE (36%). Contrarily, 90% of the mapped dry-stone walls are in fact non-SSLE area. Apart from that, 15% of dry-stone wall area was misclassified as open ground (class 6). Figure 40 and Figure 41 shows one of the few cases where a dry-stone wall sample was actually classified correctly. Generally, we can see here that the geometry of dry-stone walls is well recognised. However, the samples are mostly classified as open ground. Similarly, the dry-stone wall visible in Figure 44 and Figure 45 is misclassified as open ground or rocks and stones. It can be assumed, that the model has mainly captured the spectral features of dry-stone walls. This would explain why classes with similar spectral characteristics were chosen instead. Apart from that, this would also mean that the height and geometry information on dry-stone walls was disregarded by the model. A study by Vanhuysse et al. (2014) has also classified dry-stone walls based on a OBIA and ML approach. Based on aerial imagery (15 cm), they extracted 13 features including spectral information (layer means, brightness, skewness, intensity, redness index), geometry (length/width, asymmetry, border index, density) and relations to neighbouring- and sub-objects (mean difference to darker neighbours, mean density of sub-objects). Subsequently, four ML models (SVM, RF, RPart and KNN) were trained and tested. The results of the SVM and RF classifier proved to be sufficient for automatic recognition of dry-stone walls indicating a Kappa value of 0.79 (SVM) and 0.76 (RF). Therefore, it can be presumed that dry-stone walls can also be recognised based only aerial imagery only. However, it might be beneficial to use imagery with a spatial resolution of 15 cm or higher.

Ultimately, the one class that was not classified correctly in any cases is simultaneously the one with the least mapped GT samples: class 4 (clearance cairns). The majority of class 4 was mapped as non-SSLE (32%), rocks and stones (26%) or open ground (39%). Conversely, it was mostly non-SSLE area (92%) that was mistaken as clearance cairns. It is assumed that the main factor for the incapability of recognising clearance cairns was the spectral similarity to other rocky features. Therefore, higher focus should be set on the geometrical characteristics (mostly rounded and small) of clearance cairns. However, an accurate segmentation technique is therefore required.

As mentioned in chapter 3.7.1, class 7 (hedgerows) were not further analysed due to a semantical overlap with other vegetation classes. Therefore, the definition of this class must be adjusted in future studies.

5.3 Suitability of High-resolution Data

The main goal of this study was to evaluate the potential of high-resolution remote sensing data for the automatic recognition of small-scale landscape elements. As discussed in the previous chapter, all classes except class 9 (high woody vegetation) were recognised insufficiently by the chosen ML model. However, it is argued that the reason therefore is not the lacking suitability of high-resolution remote sensing data, but rather the methods that were applied in this study. Table 12 shows the estimated potential for manual recognition of SSLE, based on visual evaluation of datasets used for this study. This includes aerial imagery (leaf-on/off, 4-band, 5 – 25 cm) and LiDAR point cloud data (leaf-on/off, 17.5 – 30.8 points/m^2).

Type of SSLE	Ortho, leaf-on	Ortho, leaf-off	LiDAR, leaf-on	LiDAR, leaf-off
Branch piles	only if vegetation-free	only if vegetation-free	only partially	
Small water bodies	only if vegetation-free	only if vegetation-free		
Rock, stone, debris, boulders	only if vegetation-free	only if vegetation-free	only partially	
Clearance cairns	only if vegetation-free	only if vegetation-free	only partially	
Dry-stone walls	only if vegetation-free	only if vegetation-free	only partially	
Open ground	only if vegetation-free	only if vegetation-free	only partially	only partially
Hedgerows				
Low woody vegetation				
High woody vegetation				
Dead trees				

Table 12: Estimated potential of high-resolution data used in this study for manual recognition of SSLE classes (green = high potential, orange = partially, red = not possible)

Thereby “only if vegetation-free” refers to limitations if objects are covered by vegetation. Furthermore, “only partially” refers to the partial (does not work in every case) potential of depicting objects covered by vegetation, based on LiDAR intensity data and height above ground. Furthermore, while this study focused on the properties of standing dead trees, Figure 50 shows the potential of LiDAR data for the detection of lying deadwood which is clearly visible on the forest floor.

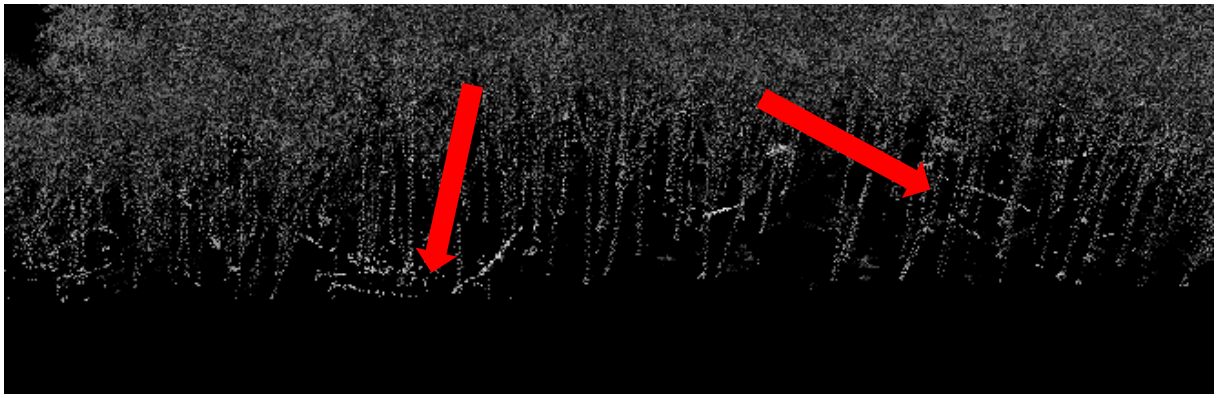


Figure 50: Example of lying deadwood recognisable with LiDAR data (© Kanton Aargau)

According to Kucharczyk et al. (2020), the essential aspect for successful object recognition is the use of H-resolution data. In the case of aerial imagery this means that the objects of interest have to be significantly larger (typically 3–5 times) than the pixels they are composed of. In regard to this, LiDAR data with very high pixel density can achieve accurate depictions of very small objects.

5.4 Evaluation of Applied Object-based Image Analysis Methods

Apart from classification accuracy statistics and a suitability evaluation of high-resolution remote sensing for automatic SSLE recognition, this study also focused on what impact OBIA and ML methods have on classification accuracy. Originally, the plan was to compare multiple segmentation and classification algorithms. However, due to time constraints, only one approach was ultimately tested. In this chapter, the potentials and limitations of the chosen approach are discussed and compared with expert literature.

In the following sections, best practice recommendations by Kucharczyk et al. (2020) (see Table 13) are compared with the methods applied in this study. Thereby, potentials and limitations of the chosen approach are discussed.

Methodological Section	Requirements and Recommendations
Image and Ancillary Data Pre-Processing	• All images must be co-registered and have the same spatial resolution.
Classification Design	• The classification scheme should be mutually exclusive, exhaustive, and hierarchical.
Segmentation and Merging	• Performing segmentation at multiple hierarchical levels may be useful for capturing image-objects of varying size and class. • We recommend performing segmentation parameter optimization using a statistical approach.
Feature Extraction and Feature Space Reduction	• We recommend performing feature space reduction to reduce the number of features used for classification. • The recommended number of features for random forest and support vector machine classification is 15–25 and 10–20. This number may vary for other classifiers.
Image-Object Classification	• Larger training sample sizes generally result in higher overall accuracy (assuming the samples are of high quality). • When possible, each training class should have an equal number of samples. • Learning curves can be constructed to assess the effect of training set size on overall accuracy. • The ideal test set will contain as many high-quality samples as possible (and necessary), with an equal number of samples per class. • We recommend the list-frame approach in conjunction with random sampling for choosing training and test samples, which avoids sampling bias due to image-object size. • We recommend using machine learning classifiers, particularly random forest and support vector machine due to their high reported accuracies.
Accuracy Assessment	• We recommend using a confusion matrix to evaluate per-class user's and producer's accuracies and overall accuracy. • We recommend evaluating segmentation accuracy; however, there is a lack of methodological consensus regarding this component.

Table 13: Best practices for OBIA. Summary adapted from Kucharczyk et al. (2020)

A basic recommendation of Kucharczyk et al. (2020) is that all input data must be co-registered and have the same spatial resolution. For this thesis the eCognition Developer software was used. Therefore, this pre-processing step was conducted automatically.

Furthermore, the authors suggest a mutually exclusive, exhaustive and hierarchical classification scheme. As discussed in chapter 5.1, certain analysed classes are very similar to each other. Therefore, it is essential to have a clear criterium that allows to distinguish them. This requirement was generally fulfilled. However, for class 7 (hedgerows) class definitions were not sufficient. As per definition (see Table 4), hedgerows differ from other woody vegetation structures only by their geometrical properties. However, those properties could not be extracted with the chosen segmentation method. Consequently, the hedgerow class was discarded.

In terms of using multiple hierarchical levels for segmentation, a slightly different approach was chosen in this study. Instead of creating hierarchical levels, segmentation was performed for two separate image scenes: one contained only non-vegetation, the other only vegetation samples. Thereby the two segmented datasets are not hierarchical which means that extracted non-vegetation samples are not sub-objects of vegetation samples. Up to now, a better method for the separation of non-vegetation and vegetation samples is still not known. However, the chosen approach could be optimised by means of applying a merging or region-growing algorithm which reduces the issue of oversegmentation. Thereby classification for e.g. forest-openland transitions could be improved. Additionally, a segmentation parameter optimisation tool could be used instead of a trial-and-error process. Furthermore, inaccurate segmentation results in this study were arguably also caused by very heterogenous input data. As mentioned in chapter 3.6.2, segmentation parameters had to be adjusted a lot for the Solothurn areas. The reason therefore was that only one (Welschenrohr areas) resp. two (Seehof) orthophotos were available for those study areas. Furthermore, input data across different study areas was also very heterogenous in terms of spatial resolution. It is assumed that this also had an effect on segmentation accuracy. A possible optimisation therefore is the use of one (more) homogenous dataset such as the nationwide SWISSIMAGE product (swisstopo, 2022a).

Further recommendations by Kucharczyk et al. (2020) include the use of feature space reduction algorithms in order to remove irrelevant and misleading features used for classification. For this thesis, the Recursive Feature Elimination method was therefore applied. Also the recommendation of using between 15 and 25 features for RF classifications was adopted.

For image-object classification, the authors suggest using training datasets as large as possible while maintaining high quality samples and equal class distribution. In this thesis, certain classes only contained very few samples. For this reason, a data balancing algorithm was applied which artificially oversampled minority class samples. However, it is argued that those samples featured relatively low quality (e.g. elements covered by vegetation resulting in misleading spectral information). As a result, “low quality” was simultaneously upscaled leading to lower classification accuracy for these classes. In regard to model validation, Kucharczyk et al. (2020) recommend using a high number of testing samples which are also available in good quality. For this thesis, a leave-one-out cross-validation method was applied thereby making use of all available samples as training samples. This arguably leads to more inaccurate validation results. Therefore, for future studies, it is recommended to exclude a certain percentage (30-50%) of samples before training and subsequently use them for testing. In regard to the applied ML algorithm, Random Forest (RF) is a robust a widely accepted technique. In this study, RF was applied using default parameter settings. Naturally, there are many options to influence classification outcomes. Most notably, the “mtry” and “ntree” parameters could be evaluated in a trial-and-error process. Apart from that, other ML algorithms could be tested.

Popular and successful alternatives are Support Vector Machines (SVM) or Artificial Neural Networks (ANN).

Finally, confusion matrices are recommended as solid tools for accuracy assessments. In particular, per-class user and producer accuracy should be analysed apart from overall accuracy. Additionally, segmentation accuracy could also be evaluated.

5.5 Outlook

Although the chosen approach for this thesis resulted in unsatisfactory classification results, several key observations were made which could help to improve classification accuracy. Based on the observations, the following suggestions are made:

- Create a binary classification model for each element of interest. Thereby, the model could be tuned specifically for the chosen element.
- Create ground truth data based on image segmentation results. Thereby, segmentation is optimised first which leads to more relevant image-objects.
- In addition to nation- or canton-wide aerial imagery, perform UAV applications to obtain very high-resolution and up-to-date input data.
- In order to improve overall methodical reproducibility, replace analysis performed with commercial software solutions with open-source tools.

6 Conclusion

This study evaluated the potential of high-resolution remote sensing data for the automatic recognition of small-scale landscape elements. Thereby, two research questions were set. Research question one reads as follows:

- RQ1: To what extent is high-resolution remote sensing data suitable for the automatic recognition of different small-scale landscape elements?

In order to answer this question, ten small-scale landscape elements were selected. This included: branch piles, small water bodies, rock/stone/debris/boulders, clearance cairns, dry-stone walls, open ground, hedgerows, low woody vegetation (1 – 3 m in height), high woody vegetation (> 3 m) and dead trees. Selected small-scale landscape elements were subsequently mapped across nine study areas. Mapped ground truth then served as input data for an automated classification approach based object-based image analysis and machine learning methods. Results showed that small and rare classes such as branch piles, small water bodies, clearance cairns or dry-stone walls could barely be predicted (0 – 74% producer and <6% user accuracies) while larger and more frequent classes (rock/stone/debris/boulders, open ground, low & high woody vegetation and dead trees) were predicted with producer accuracies between 43% and 88% and user accuracies between 1% and 80%. In particular, high woody vegetation was the only class that achieved good results (88% producer accuracy and 80% user accuracy). However, in general it can be concluded that small-scale landscape elements could not be recognised automatically in a sufficient manner. Nevertheless, currently available high-resolution remote sensing data for Switzerland is argued to be sufficient for the recognition of small-scale landscape elements. In a preceding visual analysis of high-resolution aerial imagery (spatial resolution 5 – 25 cm) and LiDAR point cloud data (17.5 – 30.8 points/m^2), most small-scale landscape elements analysed in this study could be detected by eye. Thereby, the important criterium is the availability of H-resolution data. This means that the objects of interest have to be significantly larger (typically 3–5 times) than the pixels they are composed of. This is the case for all analysed small-scale landscape elements. Furthermore, LiDAR point cloud is arguably a promising data source for detection of objects covered by vegetation.

Additionally, impacts of the chosen methods on classification accuracy were also analysed. Therefore the second research question reads as follows:

- RQ2: How do object-based image analysis and machine learning methods influence classification accuracy of different small-scale landscape elements?

Originally, the idea was to compare multiple segmentation and classification algorithms. However, due to the scope of this thesis, only one approach was ultimately tested. This included multi-resolution segmentation and a classification based on the Random Forest machine learning algorithm. Thereby, the segmentation algorithm performed well but improvement are possible in regard to region-growing and merging. As mentioned above, results from the Random Forest classification were insufficient. Main limitations were arguably very unbalanced training datasets that showed low sample quality for certain classes.

Nevertheless, several key observations were made during this study which could serve as basis for future studies. In particular, refinements in segmentation and Random Forest parameters should be evaluated. Also, the use of drone imagery could lead to promising results.

7 Bibliography

AGRIDEA. (2022). *Förderung der Biodiversität in der Schweizer Landwirtschaft*. Retrieved 31 May 2022, from http://www.bff-spb.ch/.

Arnold, B., Reiss, J., Schneider, A., & Büro am Fluss e.V. (2015). *Gewässerrandstreifen in Baden-Württemberg*. WBW Fortbildungsgesellschaft für Gewässerentwicklung mbH.

Atkinson, P. M., & Tatnall, A. R. L. (1997). Introduction neural networks in remote sensing. *International Journal of Remote Sensing*, *18*(4), 699–709. https://doi.org/10.1080/014311697218700

Beck, R. (1996). *Naturlandschaft - Kulturlandschaft: die Veränderung der Landschaften nach der Nutzbarmachung durch den Menschen*. Ecomed-Verlag-Ges.

Benediktsson, J. A., Chanussot, J., & Moon, W. M. (2012). Very High-resolution remote sensing: Challenges and opportunities [point of view]. *Proceedings of the IEEE*, *100*(6), 1907–1910. https://doi.org/10.1109/JPROC.2012.2190811

Berner Naturschutz. (2019). *Kleinstrukturen*. Retrieved 31 May 2022, from https://www.weu.be.ch/content/dam/weu/dokumente/lanat/de/umwelt/naturschutz/bauen-planen/gehoelze/Berner-Naturschutz-Kleinstrukturen-d.pdf.

Beveridge, J. R., Griffith, J., Kohler, R. R., Hanson, A. R., & Riseman, E. M. (1989). Segmenting images using localized histograms and region merging. *International Journal of Computer Vision*, *2*(3), 311–347. https://doi.org/10.1007/BF00158168

Bhunia, G. S., & Shit, P. K. (2021). Recent development and future challenges of geospatial approaches for enhancing forest inventories. In *Forest Resources Resilience and Conflicts*. Elsevier Inc. https://doi.org/10.1016/b978-0-12-822931-6.00001-0

Bins, L. S. A., Fonseca, L. M. G., Erthal, G. J., & Ii, F. M. (1996). Satellite Imagery Segmentation: a region growing approach. In *Anais VIII Simposia Brasilerio Sensoriamenta Remoto* (pp. 667–680).

BirdLife Schweiz. (2003). *Kleinstrukturen-Praxismerkblatt 5: Kopfweiden*. Retrieved 31 May 2022, from https://www.birdlife.ch/de/content/kleinstrukturen-praxismerkblatt-5-kopfweiden.

BirdLife Schweiz. (2022). *Hochstamm-Obstgärten*. Retrieved 31 May 2022, from https://www.birdlife.ch/de/content/hochstamm-obstgaerten.

Blaschke, T., Charles, B., & Pekkarinen, A. (2004). Remote sensing image analysis: including the spatial domain. In S. M. de Jong & F. D. (Eds. . van der Meer (Eds.), *Remote Sensing Image Analysis: Including the Spatial Domain* (pp. 211–236). Springer Netherlands. https://doi.org/10.1017/S0032247400010123

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., & Tiede, D. (2014). Geographic Object-Based Image Analysis - Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, *87*, 180–191. https://doi.org/10.1016/j.isprsjprs.2013.09.014

Blaschke, T., & Strobl, J. (2001). What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. *GIS – Zeitschrift Für Geoinformationssysteme*, *14*(6), 12–17.

Bolyn, C., Lejeune, P., Michez, A., & Latte, N. (2019). Automated classification of trees outside forest for supporting operational management in rural landscapes. *Remote*

Sensing, *11*(10). https://doi.org/10.3390/rs11101146

Branco, P., Ribeiro, R. P., & Torgo, L. (2016). UBL: an R package for utility-based learning. *ArXiv Preprint ArXiv:1604.08079*. https://cran.r-project.org/package=UBL

Brändli, U.-B., Abegg, M., & Allgaier Leuch, B. (2020). *Schweizerisches Landesforstinventar. Ergebnisse der vierten Erhebung 2009–2017.* Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft WSL, Bundesamt für Umwelt, Birmensdorf, Bern.

Braun, B. (2014). *Kopfweiden. Handbuch Naturschutz und Landschaftspflege*. 1–8.

Bre, F., Gimenez, J. M., & Fachinotti, V. D. (2018). Prediction of wind pressure coefficients on building surfaces using artificial neural networks. *Energy and Buildings*, *158*(November), 1429–1441. https://doi.org/10.1016/j.enbuild.2017.11.045

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

Bruzzone, L., Chi, M., & Marconcini, M. (2006). A novel transductive SVM for semisupervised classification of remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(11), 3363–3372. https://doi.org/10.1109/TGRS.2006.877950

Buholzer, S., Indermaur, A., & Riedel, S. (2021). *Handbuch für die Felddatenerhebung ALL-EMA. Version 3.0*. *118*, 1–46.

BUND (Hrsg.). (2019). *Wegraine und Gewässerrandstreifen als Teil des kommunalen Biotopverbundes. Ein Analyseleitfaden zur Kartierung und ökologischen Aufwertung landwirtschaftlich übernutzter Saumbiotope. Praxisleitfaden*.

Bunting, P., Clewley, D., Lucas, R. M., & Gillingham, S. (2014). The Remote Sensing and GIS Software Library (RSGISLib). *Computers and Geosciences*, *62*, 216–226. https://doi.org/10.1016/j.cageo.2013.08.007

Cardoso-Fernandes, J., Teodoro, A. C., Lima, A., & Roda-Robles, E. (2020). Semi-automatization of support vector machines to map lithium (Li) bearing pegmatites. *Remote Sensing*, *12*(14). https://doi.org/10.3390/rs12142319

Castilla, G., & Hay, G. J. (2008). Image objects and geographic objects. In T. Blaschke, S. Lang, & G. J. (Eds. . Hay (Eds.), *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications* (pp. 91–110). Springer.

Castle, D., Grass, I., & Westphal, C. (2019). Fruit quantity and quality of strawberries benefit from enhanced pollinator abundance at hedgerows in agricultural landscapes. *Agriculture, Ecosystems and Environment*, *275*, 14–22.

Chawla, A., Kumar, A., Warghat, A., Singh, S., Bhushan, S., Sharma, R. K., Bhattacharya, A., & Kumar, S. (2020). Approaches for conservation and improvement of Himalayan plant genetic resources. *Advancement in Crop Improvement Techniques*, 297–317. https://doi.org/10.1016/b978-0-12-818581-0.00018-8

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.

Chen, G., Weng, Q., Hay, G. J., & He, Y. (2018). Geographic object-based image analysis (GEOBIA): emerging trends and future opportunities. *GIScience and Remote Sensing*, *55*(2), 159–182. https://doi.org/10.1080/15481603.2018.1426092

Cheng, H. D., Jiang, X. H., Sun, Y., & Wang, J. (2001). Color image segmentation: Advances and prospects. *Pattern Recognition*, *34*(12), 2259–2281. https://doi.org/10.1016/S0031-3203(00)00149-7

Chollet, F. (2018). *Deep Learning with Python*. Manning Publications Co.

CNES. (2022). *Orfeo Toolbox*. Retrieved 31 May 2022, from https://www.orfeo-toolbox.org/.

Congalton, R. (2010). Remote sensing: an overview. *GIScience & Remote Sensing*, *47*(4), 443–459.

Corte, A. P. D., Neto, E. M. d. C., Rex, F. E., Souza, D., Behling, A., Mohan, M., Sanquetta, M. N. I., Silva, C. A., Klauberg, C., Sanquetta, C. R., Veras, H. F. P., de Almeida, D. R. A., Prata, G., Zambrano, A. M. A., Trautenmüller, J. W., Moraes, A. de, Karasinski, M. A., & Broadbent, E. N. (2022). High-Density UAV-LiDAR in an Integrated Crop-Livestock-Forest System: Sampling Forest Inventory or Forest Inventory Based on Individual Tree Detection (ITD). *Drones*, *6*(2). https://doi.org/10.3390/drones6020048

Crawford, M. M., Tuia, D., & Yang, H. L. (2013). *Active Learning: Any Value for Classification of Remotely Sensed Data?* 1–31.

Cui, B. (2020). *DataExplorer: automate data exploration and treatment. R package version 0.8. 1.* Retrieved 31 May 2022, from https://cran.r-project.org/package=DataExplorer.

Cunningham, P. (2008). Dimension reduction. *Machine Learning Techniques for Multimedia*, 91–112.

Demirbaş Çağlayan, S., Leloglu, U. M., Ginzler, C., Psomas, A., Zeydanlı, U. S., Bilgin, C. C., & Waser, L. T. (2020). Species level classification of Mediterranean sparse forests-maquis formations using Sentinel-2 imagery. *Geocarto International*, 1–20. https://doi.org/10.1080/10106049.2020.1783581

Earth Observing System. (2019). *What Is Spatial Resolution Of Satellite Imagery Data?* Retrieved 31 May 2022, from https://eos.com/blog/satellite-data-what-spatial-resolution-is-enough-for-you/.

ESRI. (2022a). *ArcGIS. Tabulate Intersection Tool*. Retrieved 29 May 2022, from https://desktop.arcgis.com/en/arcmap/latest/tools/analysis-toolbox/tabulate-intersection.htm.

ESRI. (2022b). *ArcGIS Desktop*. Retrieved 31 May 2022, from https://www.esri.com/en-us/arcgis/products/arcgis-desktop/overview.

ESRI. (2022c). *ArcGIS Spatial Analyst*. Retrieved 31 May 2022, from https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/an-overview-of-the-segmentation-and-classification-tools.htm.

Fan, J., Yau, D. K. Y., Elmagarmid, A. K., & Aref, W. G. (2001). Automatic image segmentation by integrating color-edge extraction and seeded region growing. *IEEE Transactions on Image Processing*, *10*(10), 1454–1466. https://doi.org/10.1109/83.951532

Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., & Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, *186*, 64–87. https://doi.org/10.1016/j.rse.2016.08.013

Freeman, E. A., Moisen, G. G., & Frescino, T. S. (2012). Evaluating effectiveness of down-sampling for stratified designs and unbalanced prevalence in Random Forest models of tree species distributions in Nevada. *Ecological Modelling*, *233*, 1–10.

Frobel, K., Wessel, M., & Klein, D. (2018). *Handbuch Biotopverbund Deutschland - Vom Konzept bis zur Umsetzung einer Grünen Infrastruktur*. B.U.N.D. e.V.

Gilbert, C. (2020). *Analysis of Interdunal Wetlands and Ecosystem Dynamics using UAS and OBIA in Ludington State Park , Michigan*. Master's Theses. 5173.

Ginzler. (2021). *Vegetation Height Model NFI. National Forest Inventory (NFI)*.

https://doi.org/doi:10.16904/1000001.1

GRASS Development Team. (2022). *GRASS GIS*. Retrieved 31 May 2022, from https://grass.osgeo.org/.

Griffith, D. C., & Hay, G. J. (2018). Integrating geobia, machine learning, and volunteered geographic information to map vegetation over rooftops. *ISPRS International Journal of Geo-Information*, *7*(12). https://doi.org/10.3390/ijgi7120462

Grippa, T., Lennert, M., Beaumont, B., Vanhuysse, S., Stephenne, N., & Wolff, E. (2017). An open-source semi-automated processing chain for urban object-based classification. *Remote Sensing*, *9*(4). https://doi.org/10.3390/rs9040358

Guerrero, M. K. M. R., Vivar, J. A. M., Ramos, R. V., & Tamondong, A. M. (2019). Assessment of Seagrass Percent Cover and Water Quality Using UAV Images and Field Measurements in Bolinao, Pangasinan. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(4/W19), 233–240. https://doi.org/10.5194/isprs-archives-XLII-4-W19-233-2019

Guntern, J., Pauli, D., & Klaus, G. (2020). Biodiversitätsfördernde Strukturen im Landwirtschaftsgebiet. Bedeutung, Entwicklung und Stossrichtungen für die Förderung. *Forum Biodiversität Schweiz (SCNAT), Bern*, 90. https://biodiversity.scnat.ch/publications/other_publications/uuid/i/f278cef9-b02b-51e1-8962-554847c00423-Biodiversitätsfördernde_Strukturen_im_Landwirtschaftsgebiet

Haris, K., Efstratiadis, S. N., Maglaveras, N., & Katsaggelos, A. K. (1998). Hybrid image segmentation using watersheds and fast region merging. *IEEE Transactions on Image Processing*, *7*(12), 1684–1699. https://doi.org/10.1109/83.730380

Harris Geospatial. (2022). *ENVI*. Retrieved 31 May 2022, from https://www.l3harrisgeospatial.com/Software-Technology/ENVI.

Hay, G. J., & Castilla, G. (2008). Geographic object-based image analysis (GEOBIA): A new name for a new discipline. In T. Blaschke, S. Lang, & G. J. (Eds. . Hay (Eds.), *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications* (pp. 75–89). Springer.

Hay, G. J., Marceau, D. J., Dubé, P., & Bocuhard, A. (2001). A multiscale framework for landscape analysis: Object-specific analysis and upscaling. *Landscape Ecology*, *16*, 471–490. file:///Users/DavidRoth/Documents/Library.papers3/Articles/2001/Hay/Landscape Ecol 2001 Hay.pdf%5Cnpapers3://publication/uuid/6CF8A57C-FBBD-43F5-BF70-D76DC444CAF7

Hay, G. J., Niemann, K. O., & McLean, G. F. (1996). An object-specific image-texture analysis of H-resolution forest imagery. *Remote Sensing of Environment*, *55*(2), 108–122. https://doi.org/10.1016/0034-4257(95)00189-1

Hexagon Geospatial. (2022). *ERDAS IMAGINE*. Retrieved 31 May 2022, from https://www.hexagongeospatial.com/products/power-portfolio/erdas-imagine.

Hochstamm Suisse. (2018). *Hochstamm Suisse: Früchte, Geschichte, Bedeutung und Biodiversität*. http://www.hochstamm-suisse.ch/home.html

Horch, P., & Holzgang, O. (2006). Hecken und Heckenbrüter: Erkenntnisse aus den drei Inventaren 1979, 1988 und 1999 im Kanton Thurgau. *Der Ornithologische Beobachter*, *103*, 96–101.

Horning, N. (2008). Remote sensing. In *Encyclopedia of Ecology* (pp. 2986–2994). Academic Press. https://doi.org/https://doi.org/10.1016/B978-008045405-4.00237-8

Hossain, M. D., & Chen, D. (2019). Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. *ISPRS Journal of Photogrammetry and Remote Sensing*, *150*(February), 115–134. https://doi.org/10.1016/j.isprsjprs.2019.02.009

Hou, W., & Walz, U. (2013). Enhanced analysis of landscape structure: Inclusion of transition zones and small-scale landscape elements. *Ecological Indicators*, *31*, 15–24. https://doi.org/10.1016/j.ecolind.2012.11.014

Hou, W., & Walz, U. (2014). Extraction of small biotopes and ecotones from multi-temporal RapidEye data and a high-resolution normalized digital surface model. *International Journal of Remote Sensing*, *35*(20), 7245–7262. https://doi.org/10.1080/01431161.2014.967890

Hudak, A. T., Evans, J. S., & Smith, A. M. S. (2009). LiDAR utility for natural resource managers. *Remote Sensing*, *1*(4), 934–951. https://doi.org/10.3390/rs1040934

Husson, E., Ecke, F., & Reese, H. (2016). Comparison of manual mapping and automated object-based image analysis of non-submerged aquatic vegetation from very-high-resolution UAS images. *Remote Sensing*, *8*(9), 1–18. https://doi.org/10.3390/rs8090724

Imesch, N., Stadler, B., Bolliger, M., & Schneider, O. (2015). Biodiversität im Wald: Ziele und Massnahmen. Vollzugshilfe zur Erhaltung und Förderung der biologischen Vielfalt im Schweizer Wald. *Bundesamt Für Umwelt, Bern. Umwelt-Vollzug*, *Nr. 1503*, 186. https://www.bafu.admin.ch/bafu/de/home/themen/biodiversitaet/publikationen-studien/publikationen/ziele-und-massnahmen-wald.html

Indermaur, L., & Schmidt, B. (2011). Quantitative recommendations for amphibian terrestrial habitat conservation derived from habitat selection behavior. *Ecological Applications*, *21*, 2548–2554.

Isenburg, M. (2021). *LAStools – efficient tools for LiDAR processing. Version 210418*. Retrieved 31 May 2022, from https://rapidlasso.de/.

Jain, R., Kasturi, R., & Schunck, B. G. (1995). *Machine Vision*. McGraw-Hill.

Jensen, J. (2007). *Remote Sensing of Environment: An Earth Resource Perspective* (2nd ed.).

Jia, Y. (2015). *Object-based Land Cover Classification with Orthophoto and LIDAR Data*.

Kanton Aargau. (2021a). *LiDAR-Punktdaten 2019 [Daten des Kantons Aargau]*. Retrieved 31 May 2022, from https://www.ag.ch/geoportal/geodatenshop/Datendokumentation.aspx?Datensatzelement=5725.

Kanton Aargau. (2021b). *Orthofotos 2019 [Daten des Kantons Aargau]*. Retrieved 31 May 2022, from https://www.ag.ch/geoportal/geodatenshop/Datendokumentation.aspx?Datensatzelement=5584.

Kanton Aargau. (2021c). *Orthofotos 2020 [Daten des Kantons Aargau]*. Retrieved 31 May 2022, from https://www.ag.ch/geoportal/geodatenshop/Datendokumentation.aspx?Datensatzelement=5906.

Kanton Solothurn. (2021a). *LiDAR 2019*. Retrieved 31 May 2022, from https://geoweb.so.ch/geodaten/index.php.

Kanton Solothurn. (2021b). *Orthofoto 2016 RGB & CIR*. Retrieved 31 May 2022, from https://geoweb.so.ch/geodaten/index.php.

Kanton Zürich. (2021a). *Orthofoto Sommer RGB/Infrarot 2018*. Retrieved 31 May 2022, from https://geolion.zh.ch/geodatensatz/show?giszhnr=493.

Kanton Zürich. (2021b). *Orthofoto Sommer RGB/Infrarot 2020*. Retrieved 29 May 2022, from https://geolion.zh.ch/geodatensatz/show?giszhnr=527.

KARCH. (2022). *Lebensräume*. Retrieved 31 May 2022, from http://www.karch.ch/karch/de/home/lebensraume.html.

Kavzoglu, T., Colkesen, I., & Tonbul, H. (2019). Agricultural Crop Type Mapping Using Object-Based Image Analysis with Advanced Ensemble Learning Algorithms. *40th Asian Conference on Remote Sensing - ACRS2019*.

Koller, O., Ineichen, S., & Schüpbach, W. (2017). Asthaufen – Lebensraum für eine vielfältige Fauna. *Vierteljahrsschrift NGZH*, *162*, 8–10.

Kremer, B. P. (2015). *Kulturlandschaften lesen: vielfältige Lebensräume erkennen und verstehen*. Haupt Verlag.

Kucharczyk, M., Hay, G. J., Ghaffarian, S., & Hugenholtz, C. H. (2020). Geographic object-based image analysis: A primer and future directions. *Remote Sensing*, *12*(12), 1–33. https://doi.org/10.3390/rs12122012

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*, 1–26. http://dx.doi.org/10.18637/jss.v028.i05

Lang, S. (2008). Object-based image analysis for remote sensing applications: Modeling reality—Dealing with complexity. In T. Blaschke, S. Lang, & G. J. (Eds. . Hay (Eds.), *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications* (pp. 3–27). Springer.

Lang, S., Hay, G. J., Baraldi, A., Tiede, D., & Blaschke, T. (2019). GEOBIA achievements and spatial opportunities in the era of big earth observation data. *ISPRS International Journal of Geo-Information*, *8*(11). https://doi.org/10.3390/ijgi8110474

Lary, D. J. (2022). Machine Learning for Environmental Sensing. In *Treatise on Geomorphology* (Second Edi, Vol. 1). Elsevier. https://doi.org/10.1016/b978-0-12-818234-5.00149-8

Lebourgeois, V., Dupuy, S., Vintrou, É., Ameline, M., Butler, S., & Bégué, A. (2017). A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM). *Remote Sensing*, *9*(3), 1–20. https://doi.org/10.3390/rs9030259

Li, M., Ma, L., Blaschke, T., Cheng, L., & Tiede, D. (2016). A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments. *International Journal of Applied Earth Observation and Geoinformation*, *49*, 87–98. https://doi.org/10.1016/j.jag.2016.01.011

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22. https://cran.r-project.org/doc/Rnews/. https://cran.r-project.org/package=randomForest

Ling, C. X., & Li, C. (1998). Data mining for direct marketing: Problems and solutions. *Kdd*, *98*, 73–79.

Liu, C., Frazier, P., & Kumar, L. (2007). Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment*, *107*(4), 606–616. https://doi.org/10.1016/j.rse.2006.10.010

Lucchese, L., & Mitray, S. K. (2001). Color image segmentation: A state-of-the-art survey.

Proceedings of the Indian National Science Academy (INSA-A). Delhi, Indian: Natl Sci Acad, *67*, 207–221. http://ultra.sdk.free.fr/docs/Image-Processing/filters/Color Image Segmentation-A State-of-the-Art Survey.pdf%5Cnhttp://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.84.4896

Ma, L., Fu, T., Blaschke, T., Li, M., Tiede, D., Zhou, Z., Ma, X., & Chen, D. (2017). Evaluation of feature selection methods for object-based land cover mapping of unmanned aerial vehicle imagery using random forest and support vector machine classifiers. *ISPRS International Journal of Geo-Information*, *6*(2). https://doi.org/10.3390/ijgi6020051

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017). A review of supervised object-based land-cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *130*, 277–293. https://doi.org/10.1016/j.isprsjprs.2017.06.001

Malinowski, R. (2016). *Research and development in the use of LIDAR elevation data for mapping of landscape elements including hedgerows , ditches and dikes*. https://pure.au.dk/ws/files/104726052/R_D_in_use_of_LIDAR_elevation_data_for_mapping_of_landscape_elements.pdf

Malkoç, E., Rüetschi, M., Ginzler, C., & Waser, L. T. (2021). Countrywide mapping of trees outside forests based on remote sensing data in Switzerland. *International Journal of Applied Earth Observation and Geoinformation*, *100*(February), 102336. https://doi.org/10.1016/j.jag.2021.102336

Marceau, D. J., Howarth, P. J., Dubois, J. M. M., & Gratton, D. J. (1990). Evaluation of the Grey-Level Co-Occurrence Matrix Method for Land-Cover Classification Using SPOT Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *28*(4), 513–519. https://doi.org/10.1109/TGRS.1990.572937

Martin, D. R., Fowlkes, C. C., & Malik, J. (2004). Learning to detect natural image boundaries using brightness and texture. *Advances in Neural Information Processing Systems*, *26*(5), 530–549.

Martinez, N., Jenni, L., & Wyss, E. (2010). Habitat structure versus food abundance: the importance of sparse vegetation for the common redstart Phoenicurus phoenicurus. *Journal of Ornithology*, *151*, 297–307.

Maselli, F., Chirici, G., Bottai, L., Corona, P., & Marchetti, M. (2005). Estimation of Mediterranean forest attributes by the application of k-NN procedures to multitemporal Landsat ETM+ images. *International Journal of Remote Sensing*, *26*(17), 3781–3796. https://doi.org/10.1080/01431160500166433

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

Mellor, A., Boukir, S., Haywood, A., & Jones, S. (2015). Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. *ISPRS Journal of Photogrammetry and Remote Sensing*, *105*, 155–168.

Mezaris, V., Kompatsiaris, I., & Strintzis, M. G. (2004). Still image segmentation tools for object-based multimedia applications. *International Journal of Pattern Recognition and Artificial Intelligence*, *18*(4), 701–725. https://doi.org/10.1142/S0218001404003393

Mielcarek, M., Kamińska, A., & Stereńczak, K. (2020). Digital aerial photogrammetry (DAP) and airborne laser scanning (ALS) as sources of information about tree height: Comparisons of the accuracy of remote sensing methods for tree height estimation. *Remote Sensing*, *12*(11). https://doi.org/10.3390/rs12111808

Moser, M., Prentice, C., & Frazier, S. (1996). A global overview of wetland loss and degradation. *Proceedings to the 6th Meeting of the Conference of Contracting Parties of the Ramsar Convention, Vol 10*.

Müller, A. C., & Guido, S. (2017). *Introduction to Machine Learning with Python*. O'Reilly Media.

NASA Earth Observatory. (1999). *Remote Sensing Methods*. Retrieved 31 May 2022, from https://earthobservatory.nasa.gov/features/RemoteSensing/remote_08.php.

Nock, R., & Nielsen, F. (2004). Statistical region growing. *IEEE Trans. Pattern Anal. Mach. Intell.*, *26*, 1452–1458.

Osmólska, A., & Hawryło, P. (2018). Using a GEOBIA framework for integrating different data sources and classification methods in context of land use/land cover mapping. *Geod. Cartogr*, *67*, 99–116.

Pajares, G. (2015). Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs). *Photogrammetric Engineering and Remote Sensing*, *81*(4), 281–329. https://doi.org/10.14358/PERS.81.4.281

Pal, M., & Mather, P. M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, *86*(4), 554–565. https://doi.org/10.1016/S0034-4257(03)00132-9

Pazúr, R., Huber, N., Weber, D., Ginzler, C., & Price, B. (2022). A national extent map of cropland and grassland for Switzerland based on Sentinel-2 data. *Earth System Science Data*, *14*(1), 295–305. https://doi.org/10.5194/essd-14-295-2022

Perona, P., & Malik, J. (1990). Scale-Space and Edge Detection Using Anisotropic Diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*(7), 629–639. https://doi.org/10.1109/34.56205

Pontius, R. G., & Millones, M. (2011). Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, *32*(15), 4407–4429. https://doi.org/10.1080/01431161.2011.552923

Rey, L., Hunziker, M., Stremlow, M., Arn, D., Rudaz, G., & Kienast, F. (2017). *Wandel der Landschaft: Erkenntnisse aus dem Monitoringprogramm Landschaftsbeobachtung Schweiz (LABES)*.

Riedel, S., Meier, E., Buholzer, S., Herzog, F., Indermauer, A., Lüscher, G., Walter, T., Winizki, J., Hofer, G., Ecker, K., & Ginzler, C. (2018). Methodenbericht ALL-EMA Arten und Lebensräume Landwirtschaft – Espèces et milieux agricoles. In *Agroscope Science* (Issue 57).

Rodewald, R., Schwyzer, Y., & Liechti, K. (2014). *Katalog der charakteristischen Kulturlandschaften der Schweiz. Grundlage zur Ermittlung von Landschaftsentwicklungszielen*. https://www.sl-fp.ch/de/stiftung-landschaftsschutz-schweiz/dokumentation/katalog-charakteristische-kulturlandschaften-53.html

Rogan, J., Franklin, J., Stow, D., Miller, J., Woodcock, C., & Roberts, D. (2008). Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. *Remote Sensing of Environment*, *112*(5), 2272–2283. https://doi.org/10.1016/j.rse.2007.10.004

Röser, B. (1988). *Saum- und Kleinbiotope. Ökologische Funktion, wirtschaftliche Bedeutung und Schutzwürdigkeit in Agrarlandschaften*. ecomed.

Schegk, I. (2015). Trockenmauerwerk - Renaissance einer traditionellen Bautechnik. In W. Jäger (Ed.), *Mauerwerk - Kalender: Bemessung, Bauen im Bestand* (2015th ed.). Ernst

& Sohn GmbH & Co. KG.

Scheunders, P., Tuia, D., & Moser, G. (2017). Contributions of machine learning to remote sensing data analysis. In *Comprehensive Remote Sensing* (Vols 1–9). Elsevier. https://doi.org/10.1016/B978-0-12-409548-9.10343-4

Sheeren, D., Bastin, N., Ouin, A., Ladet, S., Balent, G., & Lacombe, J. P. (2009). Discriminating small wooded elements in rural landscape from aerial photography: A hybrid pixel/object-based analysis approach. *International Journal of Remote Sensing*, *30*(19), 4979–4990. https://doi.org/10.1080/01431160903022928

Sperry, J. H., & Weatherhead, P. J. (2010). Ratsnakes and brush piles: Intended and unintended consequences of improving habitat for wildlife? *American Midland Naturalist*, *163*, 311–317.

Stehman, S. V., & Foody, G. M. (2019). Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment*, *231*(June). https://doi.org/10.1016/j.rse.2019.05.018

Stiftung WIN Wieselnetz & Agrofutura AG. (2018). *Fördermassnahmen für Wiesel im Landwirtschaftsgebiet. Ein Ansatz zur Erhaltung der Biodiversität und zur Reduktion von Wühlmausschäden im Wiesland. Faktenblatt und Anleitung*.

swisstopo. (2019). *swissSURFACE3D - Die klassifizierte Punktwolke der Schweiz*. Retrieved 31 May 2022, from https://www.swisstopo.admin.ch/de/geodata/height/surface3d.html.

swisstopo. (2021). *swissTLMRegio - Das kleinmassstäbliche digitale Landschaftsmodell der Schweiz*. Retrieved 31 May 2022, from https://www.swisstopo.admin.ch/de/geodata/landscape/tlmregio.html.

swisstopo. (2022a). *SWISSIMAGE - Das digitale Orthomosaik der Schweiz*. Retrieved 31 May 2022, from https://www.swisstopo.admin.ch/de/geodata/images/ortho/swissimage10.html.

swisstopo. (2022b). *SWISSIMAGE RS - Orthofotos der Schweiz für Anwendungen in der Fernerkundung*. Retrieved 31 May 2022, from https://www.swisstopo.admin.ch/de/geodata/images/ortho/swissimage-rs.html.

swisstopo. (2022c). *swissTLM3D - Das grossmassstäbliche Topografische Landschaftsmodell der Schweiz*. Retrieved 31 May 2022, from https://www.swisstopo.admin.ch/de/geodata/landscape/tlm3d.html.

Tartaro, P., & Kunz, S. (2008). *Bestand und Bedeutung von Alleen und Alleenlandschaften in der Schweiz*. Bern: Stiftung Landschaftsschutz Schweiz.

Trimble Geospatial. (2022). *eCognition Developer*. Retrieved 31 May 2022, from https://geospatial.trimble.com/products-and-solutions/trimble-ecognition.

UNESCO. (1992). *Cultural landscape*.

Vanhuysse, S., Hölbling, D., Friedl, B., Hansona, E., Krtalić, A., Hagenlocher, M., Racetin, I., & Wolff, E. (2014). Object-based image analysis for detecting indicators of mine presence to support suspected hazardous area re-delineation. *South-Eastern European Journal of Earth Observation and Geomatics*, *3*(2S).

Völker, A., & Müterthies, A. (2008). Landschaftsökologische Modellierung und automatisierte Erfassung von Landschaftselementen für das Monitoring und die Bewertung einer nachhaltigen Kulturlandschaft. *Publikationen Der Deutschen Gesellschaft Für Photogrammetrie , Fernerkundung Und Geoinformation e.V.*, *17*, 161–170.

von Büren, D., Diez, C., Bader, L., & Budde, A. (1995). *Waldrand – Artenreiches Grenzland* (No. 14; SBN-Merkbl.).

Wang, M., & Li, R. (2014). Segmentation of high spatial resolution remote sensing imagery based on hard-boundary constraint and two-stage merging. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(9), 5712–5725. https://doi.org/10.1109/TGRS.2013.2292053

Weber, D., Ginzler, C., Flückiger, S., & Rosset, C. (2018). Potenzial von Sentinel-2-Satellitendaten für Anwendungen im Waldbereich. *Schweizerische Zeitschrift Fur Forstwesen*, *169*(1), 26–34. https://doi.org/10.3188/szf.2018.0026

Weber, D., Rüetschi, M., Small, D., & Ginzler, C. (2020). Large-scale classification of shrub forest with remote sensing data. *Schweizerische Zeitschrift Fur Forstwesen*, *171*(2), 51–59. https://doi.org/10.3188/szf.2020.0051

Westphal, U. (2011). *Hecken - Lebensräume in Garten und Landschaft: Ökologie, Artenvielfalt, Praxis*. Pala-Verlag.

Whiteside, T. G., Maier, S. W., & Boggs, G. S. (2014). International Journal of Applied Earth Observation and Geoinformation Area-based and location-based validation of classified image objects. *International Journal of Applied Earth Observations and Geoinformation*, *28*, 117–130. http://dx.doi.org/10.1016/j.jag.2013.11.009

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org

Ye, S., Pontius, R. G., & Rakshit, R. (2018). A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, *141*, 137–147. https://doi.org/10.1016/j.isprsjprs.2018.04.002

Zhang, X., Xiao, P., Feng, X., Wang, J., & Wang, Z. (2014). Hybrid region merging method for segmentation of high-resolution remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *98*, 19–28. https://doi.org/10.1016/j.isprsjprs.2014.09.011

Zhang, Y.-J. (2006). An Overview of Image and Video Segmentation in the Last 40 Years. In Y.-J. (Eds. . Zhang (Ed.), *Advances in Image and Video Segmentation* (pp. 1–15). IRM Press.

Zhu, L., Suomalainen, J., Liu, J., Hyyppä, J., Kaartinen, H., & Haggren, H. (2018). A review: Remote sensing sensors. In *Multi-purposeful application of geospatial data* (pp. 19–42).

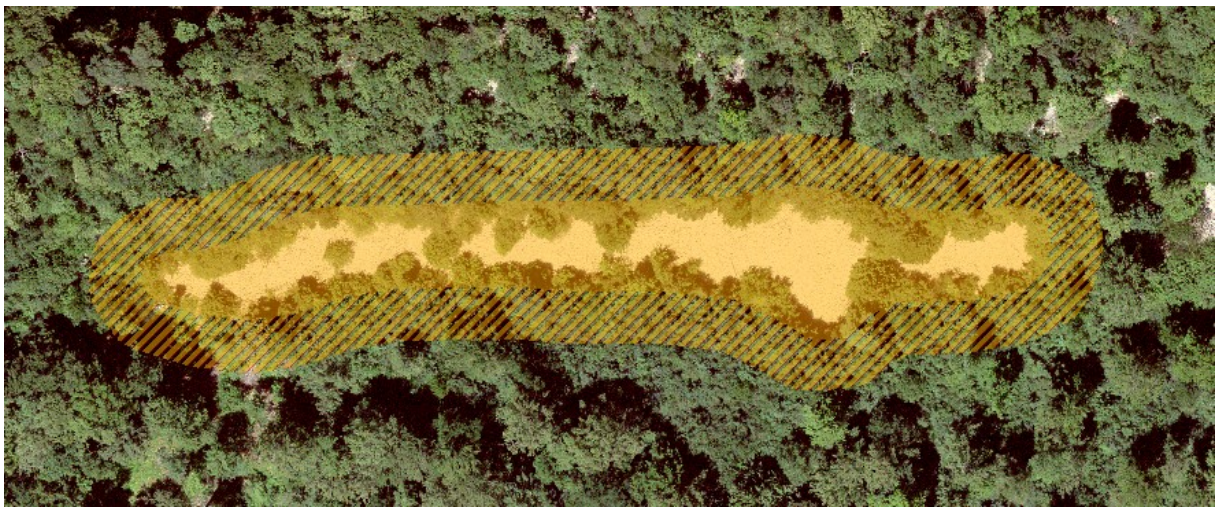
Zollhöfer, J. M. (1997). Quellen die unbekannten Biotope: erfassen, bewerten, schützen. In *Page Bristol-Schriftenreihe Band 6*. Bristol-Stiftung.

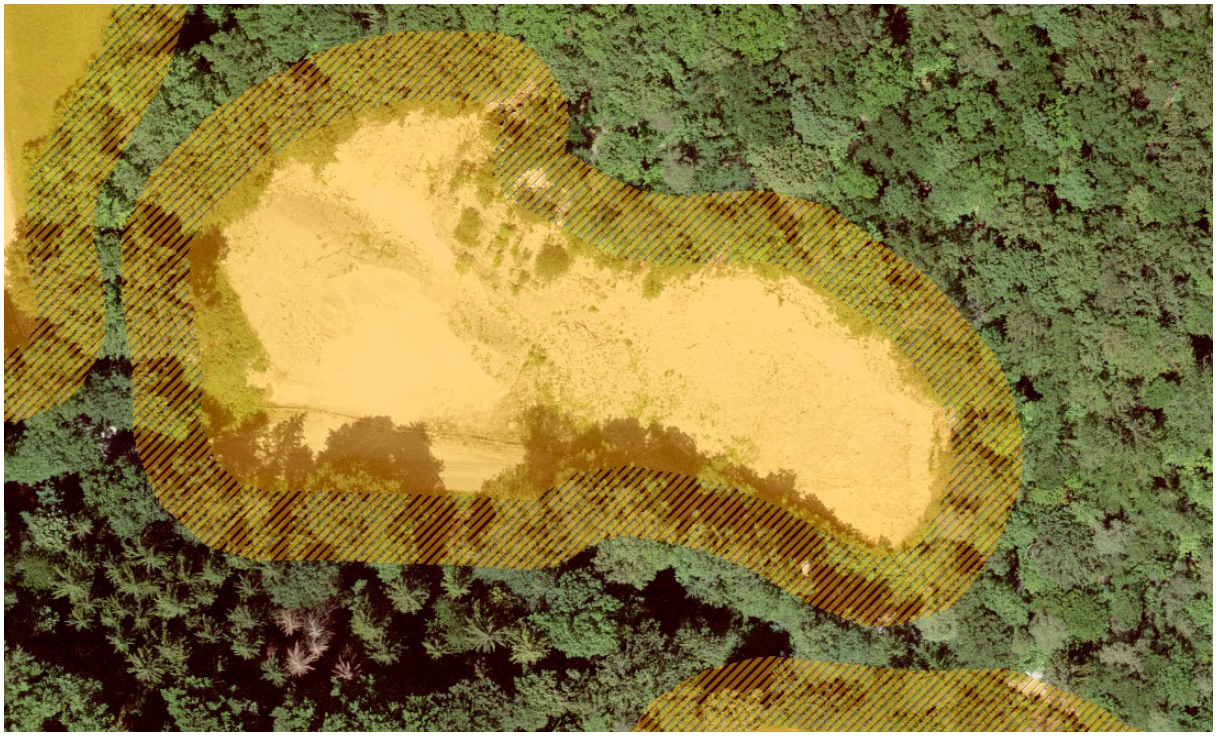
Zurbuchen, A., & Müller, A. (2012). *Wildbienenschutz - von der Wissenschaft zur Praxis*. Bristol Stiftung, Zürich; Haupt, Bern, Stuttgart, Wien.

8 Appendix

Appendix 1: Examples of Manual Adjustments for Study Area Definition

Following examples illustrate how areas obviously not used for farming where manually removed from the study area.





Appendix 2: Complete RFE feature ranking

(Non-vegetation model)

Rank	Feature
1	Mean_vecInd
2	Density
3	Mean_ndvi2
4	Curv_Length
5	StdDev_blue2
6	Width
7	Length_main
8	StdDev_ndsi2
9	Compactness
10	GLCM_Corr
11	Mean_nir2
12	Mean_ndvi1
13	StdDev_green2
14	Mean_red2
15	Mean_ndwi1
16	Area
17	Mean_blue2
18	Mean_green2
19	StdDev_ndvi2
20	Mean_blue1
21	StdDev_red2
22	Mean_ndsi2
23	Mean_red1
24	GLCM_StdDev
25	Mean_green1
26	StdDev_blue1
27	Mean_ndsi1
28	StdDev_Z
29	Max_Z
30	StdDev_nir1
31	StdDev_nir2
32	StdDev_ndsi1
33	StdDev_intInd
34	Mean_intInd
35	Mean_Z
36	StdDev_ndvi1
37	StdDev_red1
38	Mean_nir1
39	Mean_blue3
40	StdDev_ndwi1
41	StdDev_green1
42	StdDev_vecInd
43	Mean_NoR
44	Mean_ndvi3
45	Mean_red3
46	Mean_ndsi3
47	Mean_green3
48	StdDev_ndwi3
49	Mean_nir3
50	StdDev_blue3
51	StdDev_ndvi3
52	StdDev_ndsi3
53	StdDev_green3
54	StdDev_red3
55	Mean_ndwi3
56	StdDev_nir3

(Vegetation model)

Rank	Feature
1	Mean_vhm
2	Max_Z
3	Mean_Z
4	StdDev_Z
5	Mean_NoR
6	Mean_nir2
7	Mean_ndsi2
8	StdDev_vhm
9	Mean_ndvi2
10	Density
11	Mean_blue2
12	Width
13	Mean_red2
14	Mean_nir1
15	GLCM_StdDev
16	Mean_green2
17	StdDev_nir2
18	Curv_Length
19	StdDev_blue2
20	StdDev_ndvi2
21	Mean_ndvi1
22	StdDev_red2
23	Mean_blue1
24	Compactness
25	StdDev_ndsi2
26	StdDev_blue1
27	StdDev_green2
28	Mean_red1
29	Area
30	StdDev_red1
31	Length_main
32	Mean_green1
33	Mean_ndwi1
34	StdDev_green1
35	Mean_ndsi1
36	StdDev_ndvi1
37	StdDev_nir1
38	StdDev_ndwi3
39	GLCM_Corr
40	Mean_ndsi3
41	StdDev_ndwi1
42	StdDev_blue3
43	Mean_nir3
44	Mean_ndvi3
45	StdDev_ndvi3
46	StdDev_ndsi3
47	StdDev_nir3
48	StdDev_ndsi1
49	Mean_blue3
50	StdDev_red3
51	Mean_red3
52	Mean_ndwi3
53	StdDev_green3
54	Mean_green3

Erklärung

gemäss Art. 30 RSL Phil.-nat. 18

Name/Vorname:	Mägli Simon
Matrikelnummer:	16-100-968
Studiengang:	Master of Science in Geography

Bachelor ☐ Master ☑ Dissertation ☐

Titel der Arbeit:	Evaluating the Potential of High-resolution Remote Sensing Data for the Automatic Recognition of Small-scale Landscape Elements
LeiterIn der Arbeit:	Prof. Dr. Matthias Bürgi, Christian Ginzler

Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe r des Gesetzes vom 5. September 1996 über die Universität zum Entzug des auf Grund dieser Arbeit verliehenen Titels berechtigt ist.
Für die Zwecke der Begutachtung und der Überprüfung der Einhaltung der Selbständigkeitserklärung bzw. der Reglemente betreffend Plagiate erteile ich der Universität Bern das Recht, die dazu erforderlichen Personendaten zu bearbeiten und Nutzungshandlungen vorzunehmen, insbesondere die schriftliche Arbeit zu vervielfältigen und dauerhaft in einer Datenbank zu speichern sowie diese zur Überprüfung von Arbeiten Dritter zu verwenden oder hierzu zur Verfügung zu stellen.

Bern, 14.06.2022

Ort/Datum

Unterschrift