UAV-BASED HIGH-RESOLUTION REMOTE SENSING AND MACHINE LEARNING FOR RISK MANAGEMENT IN HAZARD AND DISASTER AREAS

Sebastian Briechle

UAV-BASED HIGH-RESOLUTION REMOTE SENSING AND MACHINE LEARNING FOR RISK MANAGEMENT IN HAZARD AND DISASTER AREAS

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof.dr.ir. A. Veldkamp,
on account of the decision of the Doctorate Board
to be publicly defended
on Wednesday 25 August 2021 at 14.45 hrs

by

Sebastian Briechle
born on the 16th of September, 1986
in Rosenheim, Germany

This dissertation has been approved by:

Supervisor
        prof.dr.ir. M.G. Vosselman

Co-supervisor
        Prof. Dr. P. Krzystek

UNIVERSITY
OF TWENTE. | ITC

Graduation Committee:

Chair / secretary
     prof.dr. F.D. van der Meer   University of Twente

Supervisor
     prof.dr.ir. M.G. Vosselman   University of Twente

Co-supervisor
     Prof. Dr. P. Krzystek   Munich University of Applied Sciences

Committee Members
     prof.dr. V.G. Jetten   University of Twente
     prof.dr. N. Kerle   University of Twente
     prof.dr. N. Pfeifer   Vienna University of Technology
     prof.dr. I. Korpela   University of Helsinki

Acknowledgements

First of all, my deepest gratitude goes to my co-supervisor, Prof. Peter Krzystek, head of the Photogrammetry and Remote Sensing Laboratory at the Munich University of Applied Sciences. His tireless commitment over many years to advancing research at the faculty is very remarkable. I have been constantly motivated by his high level of enthusiasm. His equally challenging and supportive manner has helped me to develop professionally. His technical expertise and his openness to share this valuable knowledge have been very supportive and key to the accomplishment of my goals. It was mainly thanks to him that I received four years of financing for my position, allowing for long-term planning reliability. Together with the excellent and always up-to-date infrastructure in his laboratory, this security has enabled me to concentrate on my actual research work and to successfully complete this doctoral thesis within the planned time interval of four years. Our first joint visit to a scientific conference in Riva del Garda and, of course, the two missions conducted in the Chernobyl Exclusion Zone were unforgettable. Peter, I very much appreciate your continuous support and your open and honest manner.

Furthermore, my sincere thanks go to my supervisor at the University of Twente, Faculty ITC, Prof. George Vosselman. He kindly gave me the opportunity to join his research team as an external PhD student. Thank you very much for this leap of faith; I hope I was able to meet your expectations! Due to the long distance, our personal meetings were mostly limited to online sessions. However, in these meetings, I have always received excellent professional and, above all, strategic guidance and advice. I also appreciate the professional exchanges and the lively discussions within his research group during the PhD seminar and the PhD reading class. George, it has been an honour to conduct my PhD under your supervision. I admire your expertise and highly appreciate your pleasant, selfless manner.

This entire research project has been based on successful collaboration with Dr Norbert Molitor (Plejades GmbH). Without his years of involvement in the Chernobyl Nuclear Power Plant environment and the resulting contacts with local authorities, the missions in the ChEZ could never have taken place. I would like to take this opportunity to thank him for his sound technical input on all questions regarding the unusual study area, his enthusiastic and detailed explanations and his incredibly valuable support on site.

At this point, I would like to thank the graduation committee members for accepting to review my PhD thesis. Furthermore, I want to convey my sincere thanks to my current and former office-mates and colleagues at FK08, who created a very pleasant and motivating working

Abstract

In hazard and disaster areas, successful fact-based risk management requires precise geoinformation. Especially in areas that cover a large expanse, hazard maps and action plans – generated from high-resolution remote sensing data – play a key role for decision makers and emergency forces. In this context, unmanned aerial vehicles (UAVs) are of increasing interest, mainly because of their flexibility and the minimized risk to task forces by avoiding risky trespassing. In addition, machine learning-based methods on the data processing side enable large amounts of collected data to be converted into usable end products quickly and efficiently.

The first part of this dissertation covers the detection of buried radioactive waste sites in the Chernobyl Exclusion Zone, which was created in the aftermath of the 1986 nuclear catastrophe. In post-disaster remediation work, liquidation materials contaminated by radioactive fallout were buried in trenches and clamps. Unfortunately, most of these on-site burials were insufficiently documented and have become overgrown with dense vegetation. For the purpose of risk management, it is crucial to know their exact locations and dimensions. To avoid risky area trespassing, UAV-based gamma spectrometry surveys were performed in the past, achieving promising results in open areas. However, the results were inaccurate in forested areas. Known trenches and clamps display two major characteristics: terrain height anomalies in the decimetre range and vegetation anomalies that tend to appear in the immediate vicinity of the burials. Thus, a UAV-based method was developed to detect unknown radioactive waste sites in forested areas by combining high-resolution lidar data with multispectral (MS) imagery. First, a digital terrain model (DTM) was extracted from the lidar point cloud. Then, based on datasets from both optical sensors, 3D vegetation mapping was performed at the single-tree level. Additionally, tree-based features (e.g., density, height, and species) were derived. Subsequently, these vegetation-related features were combined with feature sets compiled from the DTM. The data were processed using a reliable feature selection method and a random forest-based classifier for the simultaneous detection of both trenches and clamps. This approach was tested in densely vegetated and thus challenging areas, where historical maps only roughly indicate the existence of radioactive waste sites. Finally, the detection results were verified by on-site test drillings that successfully confirmed the existence of previously unknown buried nuclear materials in the classified areas.

As indicated, 3D vegetation mapping can support the detection of radioactive waste sites. Furthermore, precise vegetation maps at the tree level can also assist in forest fire simulations and improve models for

estimating workers' radiation dose uptake. Thus, the second part of this PhD thesis focuses on the mapping of single trees. More precisely, the combined classification of presegmented single trees with respect to tree species and standing dead trees was extensively investigated. Therefore, deep learning-based approaches for both 2D and 3D data were developed to enable object classification based on automatically extracted features. PointNet++ was adapted to the task of classifying raw 3D point clouds of presegmented single trees. Aside from the 3D geometry, additional point attributes such as lidar echo pulse width, surface normals and MS features were also integrated into the classification process. In the final study, an image-based 2D convolutional neural network (CNN) approach called Silvi-Net was created to gather meaningful features from both airborne lidar data and MS images. Overall, the network's performance was proven using 2D and 3D datasets from two natural forest areas acquired with different sensor models and varying geometric and spectral resolutions. Using pretrained weights and recursive retraining of CNN model parameters (transfer learning), Silvi-Net showed high generalization capacity, even for datasets with a reduced number of samples. Interestingly, multiple silhouette-like side-view images rendered from the point clouds of the single trees significantly increased the overall accuracy compared to experiments only utilizing MS imagery. Using identical data, the Silvi-Net and PointNet++ results were compared, revealing the advantages and disadvantages of the approaches. All in all, Silvi-Net facilitated the generation of reliable vegetation maps that are of major importance in risk management applications, automated forest inventory and monitoring projects.

Samenvatting

In risico- en rampgebieden moet een op feiten gebaseerd risicobeheer beschikken over nauwkeurige geo-informatie. Vooral in gebieden met een grote omvang spelen gevarenkaarten en actieplannen – gegenereerd uit hoge-resolutie remote sensing gegevens – een sleutelrol voor besluitvormers en noodhulpdiensten. In deze context zijn onbemande luchtvaartuigen (UAV's) van toenemend belang, vooral vanwege hun grote flexibiliteit en het geminimaliseerde risico voor de task forces door het betreden van gevaarlijke gebieden te vermijden. Bovendien maken op machine learning gebaseerde methoden aan de kant van de gegevensverwerking het mogelijk om de enorme hoeveelheden verzamelde gegevens snel en efficiënt om te zetten in bruikbare eindproducten.

Het eerste deel van dit proefschrift behandelt de opsporing van ondergrondse locaties voor radioactief afval in de exclusieve zone van Tsjernobyl, die werd gecreëerd in de nasleep van de kernramp in 1986. Bij de saneringswerkzaamheden na de ramp werden met radioactieve neerslag besmette materialen begraven in greppels en aardwallen. Helaas waren de meeste van deze begravingen ter plaatse onvoldoende gedocumenteerd en zijn ze overwoekerd door dichte begroeiing. Voor het risicobeheer is het echter van cruciaal belang hun precieze ligging en afmetingen te kennen. Om het betreden van riskante gebieden te vermijden, werden in het verleden op UAV gebaseerde gammaspectrometrische onderzoeken uitgevoerd, die veelbelovende resultaten opleverden in open gebieden. In beboste gebieden waren de resultaten echter ongeschikt. Gebleken is dat bekende greppels en aardwallen twee belangrijke kenmerken vertonen. Ten eerste, hoogte-anomalieën van het terrein in het decimeterbereik en, ten tweede, vegetatie-anomalieën die de neiging hebben in de onmiddellijke nabijheid van de begravingen te verschijnen. Daarom werd een UAV-gebaseerde methode ontwikkeld om onbekende plaatsen met radioactief afval op te sporen in beboste gebieden, waarbij lidargegevens met hoge resolutie en multispectrale (MS) beelden werden samengevoegd. In eerste instantie wordt een digitaal terreinmodel (DTM) geëxtraheerd uit de lidar puntenwolk. Vervolgens, gebaseerd op datasets van beide optische sensoren, wordt 3D vegetatie kartering uitgevoerd. Bovendien worden op bomen gebaseerde kenmerken afgeleid, zoals boomdichtheid, boomhoogte en boomsoort. Vervolgens worden deze vegetatie-gerelateerde kenmerken gecombineerd met kenmerkenreeksen die zijn samengesteld uit het DTM. De verdere verwerking omvat een betrouwbare kenmerkselectiemethode en een random forest-gebaseerde classificator voor de gelijktijdige detectie van zowel greppels als aardwallen. De algemene aanpak werd getest in dichtbegroeide en dus moeilijke gebieden, waar historische kaarten

slechts een ruwe indicatie geven van het bestaan van locaties met radioactief afval. Tenslotte werden de detectieresultaten geverifieerd door proefboringen ter plaatse die met succes het bestaan van voorheen onbekend begraven nucleair materiaal in de geclassificeerde gebieden bevestigden.

Zoals aangegeven kan 3D-vegetatiekartering helpen bij de opsporing van locaties met radioactief afval. Bovendien kunnen nauwkeurige vegetatiekarteringen op boomniveau ook helpen bij de simulatie van bosbranden en de modellen voor de raming van de stralingsdosis die werknemers oplopen verbeteren. Het tweede deel van dit proefschrift richt zich dan ook op de kartering van individuele bomen. Om precies te zijn werd de gecombineerde classificatie van voorgesegmenteerde alleenstaande bomen met betrekking tot boomsoorten en staande dode bomen uitgebreid onderzocht. Daarom zijn deep learning-gebaseerde benaderingen voor zowel 2D als 3D data ontwikkeld, die objectclassificatie op basis van automatisch geëxtraheerde kenmerken mogelijk maken. Binnen dit onderwerp werd PointNet++ aangepast voor de taak van het classificeren van ruwe 3D puntenwolken van individuele voorgesegmenteerde bomen. Naast de 3D geometrie werden ook additionele puntattributen zoals lidar echopulsbreedte, oppervlaktenormalen, en MS kenmerken geïntegreerd in het classificatieproces. In een laatste studie wordt een beeldgebaseerde 2D convolutionele neurale netwerk (CNN) aanpak, genaamd Silvi-Net, geïntroduceerd voor het leren van zinvolle kenmerken van zowel lidar gegevens en MS beelden. De prestaties van het netwerk werden bewezen met behulp van 2D en 3D datasets van twee natuurlijke bosgebieden, verkregen met verschillende sensormodellen, en variërende geometrische en spectrale resolutie. Door gebruik te maken van vooraf getrainde gewichten en recursieve training van CNN-modelparameters (=transfer learning), vertoonde Silvi-Net een hoog generalisatievermogen, zelfs voor datasets met een beperkt aantal samples. Interessant is dat meervoudige silhouet-achtige beelden met zijaanzichten, gerenderd uit de puntenwolken van de afzonderlijke bomen, de algehele nauwkeurigheid significant verhoogden in vergelijking met experimenten die alleen gebruik maakten van MS-beelden. Gebruikmakend van identieke data, worden de Silvi-Net resultaten vergeleken met die van PointNet++, waarbij voor- en nadelen van de benaderingen aan het licht komen. Al met al vergemakkelijkt Silvi-Net het genereren van betrouwbare vegetatiekaarten die van groot belang zijn in risicobeheersingstoepassingen, geautomatiseerde bosinventarisatie en monitoringprojecten.

Table of Contents

List of figures

List of tables

List of abbreviations

ALS	Airborne laser scanning
BFNP	Bavarian Forest National Park
CAM	Class activation mapping
CBH	Crown base height
ChEZ	Chernobyl Exclusion Zone
CHM	Canopy height model
ChNPP	Chernobyl Nuclear Power Plant
CIR	Colour-infrared
CNN	Convolutional neural network
CPU	Central processing unit
CRP	Calibrated reflectance panel
CRWME	Central Radioactive Waste Management Enterprise
CV	Cross-validation
DBH	Diameter at breast height
DEM	Digital elevation model
DL	Deep learning
DLS	Downwelling light sensor
DNN	Deep neural network
DTM	Digital terrain model
EC	Laser echo characteristics
EDR	Equivalent dose rate
EPSG	European Petroleum Survey Group Geodesy
EW	Laser echo pulse width
FN	False negative
FOV	Field of view
FP	False positive
FPFH	Fast point feature histograms
GCP	Ground control point
GNSS	Global Navigation Satellite System
GPR	Ground-penetrating radar
GPS	Global Positioning System
GPU	Graphics processing unit
GSD	Ground sampling distance

ICP	Iterative closest point
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
INT	Intensity
ISPRS	International Society for Photogrammetry and Remote Sensing
Lidar	Light detection and ranging
LM	Lidar metrics
LR	Learning rate
MCARI	Modified Chlorophyll Absorption Ratio Index
ML	Machine learning
MLP	Multi-layer perceptron
MRESR	Modified Red Edge Simple Ratio
MS	Multispectral
NDVI	Normalized Difference Vegetation Index
NIR	Near-infrared
OA	Overall accuracy
PA	Producer's accuracy
RE	Red edge
REDVI	Red Edge Difference Vegetation Index
ReLU	Rectified linear unit
RENDVI	Red Edge Normalized Difference Vegetation Index
RF	Random forest
RFE	Recursive feature elimination
RGB	Red green blue
RGB-D	RGB-depth
RWTSP	Radioactive Waste Temporary Storage Place
std	Standard deviation
SVM	Support vector machine
TF	Tree features
TN	True negative
TP	True positive
UAV	Unmanned aerial vehicle
VI	Vegetation index

1 Introduction

In hazard and disaster areas, precise and reliable geoinformation is of major importance for successful risk management. In both pre- and post-disaster cases, detailed spatial information plays a key role for decision makers and emergency forces. Especially in areas that cover a large expanse, high-resolution remote sensing data support the generation of hazard maps and action plans and, thus, enable fact-based and quick decisions. In the past, various studies have been conducted using remote sensing data for the mapping and monitoring of natural hazards and man-made disasters. Global-scale hazards, such as changes in atmospheric composition (e.g., Laj et al., 2009), rising sea levels (e.g., Cazenave and Llovel, 2010) and glacier recession (e.g., Barry, 2006), have been studied thoroughly. On a more regional scale, numerous researchers have explored storm events (e.g., Friedman and Li, 2000) and mapped the damage after earthquakes (e.g., Vetrivel, 2018) or floods (e.g., Smith, 1997). Furthermore, researchers have worked on the monitoring of droughts (e.g., Kogan, 1997), coastal erosion (e.g., White and El Asmar, 1999) and landslides (e.g., Metternicht et al., 2005).

This PhD thesis mainly focuses on the aftermath of the 1986 nuclear disaster in Chernobyl, located approximately 100 km north of Kiev, Ukraine. Specifically, this research addresses the detection of unknown radioactive biomass deposits in the Chernobyl Exclusion Zone (ChEZ). Extensive investigation was performed on the deep learning (DL)-based mapping of single trees, especially with respect to the combined classification of tree species and standing dead trees. Precise 3D vegetation mapping is typically used for automated forest inventory, including the estimation of structural parameters at the tree level. In terms of risk management, vegetation maps can also support forest fire simulations and – as outlined in this PhD thesis – give indirect evidence about contaminated subsoil.

1.1. Background

On April 26, 1986, the explosion of Reactor Unit 4 at the Chernobyl Nuclear Power Plant (ChNPP) resulted in a nuclear disaster. The International Nuclear and Radiological Event Scale recorded this as a major accident, which is the highest level for this scale. The spread of radioactive materials necessitated the evacuation of the surrounding territory. To re-establish safety, the authorities created the ChEZ, which is controlled by the military and remains in effect until today. The main radioactive fallout trail (Western Trace, Figure 1) caused extraordinary contamination of the 30- to 40-year-old pine forest located 1.5–2.0 km west of the ChNPP. Because of the presence of large amounts of micron-sized nuclear fuel hot particles in this area, the trees turned reddish-brown and died. Thus, the area is referred to as the Red Forest

(Figure 2). During liquidation measures by the civil defence troops (Figure 3), the contaminated biomass was cut down and – together with the contaminated topsoil layer and construction debris – buried on-site in Radioactive Waste Temporary Storage Places (RWTSPs). Covering the burials with a clean soil layer of approximately 1 m helped to immediately reduce the dose rate on the surface. Additionally, the area was reforested to prevent wind resuspension and water erosion of the soil cover layer.
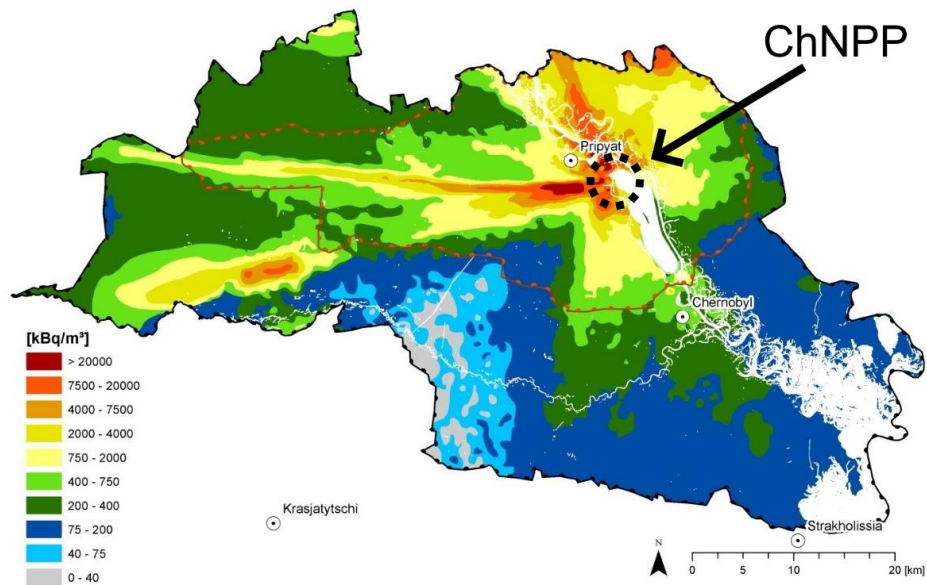
Figure 1: Surface contamination in the ChEZ with Cs-137 (as of 1997); Source: Molitor et al., 2018.

Figure 2: Red Forest. Source: chornobyl.in.ua.

Figure 3: Remediation works in the ChEZ. © A.P. Yakubchik (chornobyl.in.ua).

However, appropriate protection against radiation mobilisation into the groundwater was not established. Because radionuclides migrate into the environment via water pathways, the RWTSPs represent a serious hazard to the surrounding area. While the direct consumption of contaminated water by humans can be avoided with appropriate measures, the hazard of indirect ingestion is more challenging. Molitor et al. (2017) described the possible absorption of contaminated foodstuffs by humans as a consequence of propagation of radionuclides in the food chain. The model assumes that contaminated groundwater is absorbed by plants and grass, leading to root and leaf uptake of radionuclides. With domestic animals consuming grass and certain plants, radionuclides further migrate into foodstuffs such as meat, eggs, milk and dairy products – and can finally end up in the human organism.

1.2. Detection of radioactive waste sites

Recently, there has been increasing pressure to make the ChEZ inhabitable again. To improve safety during maintenance and facilitate the eventual remediation of burials, detailed hazard maps are needed for risk management. These maps should ideally include all of the trenches and clamps and should indicate their exact number, location, geometry and contents. Unfortunately, the liquidation measures in the aftermath of the nuclear disaster were performed under extraordinary conditions. Therefore, the exact locations of the burials were not sufficiently documented, and most current information relies on improvised records. Nevertheless, more than 700 burials have been identified over the past three decades. However, approximately 300 out of 1,000 burials have yet to be found (Molitor et al., 2017). The detection of these undiscovered radioactive waste sites in the ChEZ has been investigated using different sensor types. Bugai et al. (2005) found that experiments based on ground dose rate measurements and electromagnetic soil conductivity surveys were not expedient. Thus, the authors subsequently utilized the technique of ground-penetrating radar (GPR) and successfully elucidated the subsurface geometry of one suspected burial site in the RWTSP Red Forest. Based on new GPR surveys, Saintenoy et al. (2017) refined the position of this trench and detected a new trench in its immediate vicinity. However, ground-based methods have two major downsides. First, they are often time consuming and limited to the exploration of relatively small areas. Second, ground-based campaigns can result in considerable radiation exposure for workers.

Regarding these issues, unmanned airborne solutions are an advantageous option. After the 2011 nuclear reactor explosion in Fukushima, Towler et al. (2012) highlighted the need for unmanned aerial vehicle (UAV)-based remote sensing applications in disaster areas to map both the structural and radiation content of a post-disaster environment.

Additionally, the authors argued that their system showed high efficiency in planning operations. Moreover, MacFarlane et al. (2014) presented a UAV-based solution for the remote detection of radiation anomalies with a high spatial resolution of less than 1 m. In the ChEZ, Zabulonov et al. (2015) utilized a gamma spectrometer system and conducted UAV-based surveys. By analysing the total gamma intensity field and its local component along the flight line, local anomalies were identified. The authors demonstrated that this local inhomogeneity partly indicated conspicuous areas with unknown burials. In contaminated areas, the local gamma ray intensity significantly exceeded the level of the total background radiation. As a result, burials in non-vegetated areas were identified with a detection rate of 90%. Nevertheless, in forested areas, the detection rate dropped down to 50% because the biomass substantially distorted the measurements.

Interestingly, in-field observations have revealed two major characteristics that are typical for the burials. Because most of the trenches and clamps have been exposed to natural processes such as settlement and erosion, they are represented by elevations and settlements in the decimetre range (Figure 4). Furthermore, in some cases, vegetation anomalies tend to appear in the immediate vicinity of the waste sites (Figure 5) and can be considered possible indirect indicators.

Figure 4: Possible trench in Red Forest 2.5, showing slight height settlements in the decimetre range. © Peter Krzystek

Figure 5: Vegetation anomalies in trench areas. © Peter Bayer

1.3. Vegetation mapping

The impact of radioactive contamination on the vegetation in the ChEZ has been studied comprehensively. Undoubtedly, radioactive isotopes (mainly Cs-137 and Sr-90) still exist in the soil of contaminated areas due to their half-lives of approximately 30 years (Flynn et al., 1965). Moreover, the transportation of radionuclides in the groundwater was modelled by Bugai et al. (2012), and the biogenic migration of radionuclides from subsurface storage into plants was described by Kashparov et al. (2012). The absorption of radionuclides differs between tree species (Tikhomirov and Shcheglov, 1994), and the absorbed irradiation dose correlates with the tree mortality rate (Arkhipov et al., 1994). Interestingly, Thiry et al. (2009) demonstrated that trees growing on burial sites accumulated significantly more Cs-137 and Sr-90 in their above-ground biomass compared to trees growing a few metres away from the burial sites. Furthermore, the level of radioactive contamination correlates well with certain spectral characteristics of leaves and needles (Davids and Tyler, 2003) as well as effects occurring on the morphological level (Yoschenko et al., 2011). Overall, these studies encourage the assumption that vegetation anomalies (e.g., tree density, dominant tree species, tree shape and spectral reflectance) can be expected in highly contaminated areas such as the vicinity of trenches and clamps. More precisely, vegetation anomalies are believed to be indirect indicators of unmapped radioactive burials.

Therefore, it is crucial to investigate current remote sensing methods for vegetation mapping. In general, the generation of precise and reliable vegetation maps from high-resolution remote sensing data is a

fundamental task in forestry. For instance, forest attributes such as above-ground biomass and growing stock can be estimated based on allometric models. Over the past decade, drone-based approaches to forest applications have emerged (Guimarães, 2020) – mainly because of low material and operational costs, high flexibility regarding spatial and temporal resolution and the absence of risk to crews (Tang and Shao, 2015). Because drone-based approaches enable precise area-wide inventory of forest structural variables, tree-level approaches are of increasing interest. Furthermore, to investigate the robustness of various forest compositions to changing climatic conditions, forest managers and nature conservationists require quantitative mapping results (Overbeck and Schmidt, 2012). In this context, the classification of dead trees has become increasingly important. However, the overall performance of approaches for individual tree species classification in dense (and thus complex) temperate forests is still insufficient for practical use, as an overall accuracy (OA) of at least 90–95% is required for multi-class tasks.

1.4. Key idea and applications

Detection methods for trenches and clamps based on visual ground inspection or GPR measurements are time, effort and radiation dose intensive. These disadvantages can typically be surpassed by airborne techniques, allowing for quick area-wide mapping and avoiding risky trespassing of the study area. In the past, UAV-based gamma spectrometry surveys in the ChEZ achieved promising results in open areas. However, the results were inaccurate in forested areas.

As already mentioned, known trenches and clamps display two major characteristics: first, digital terrain model (DTM) anomalies in the decimetre range and, second, vegetation anomalies with respect to tree height, tree density, crown volume, above-ground biomass and tree species composition in their surroundings. Therefore, the key idea of this PhD thesis was to utilize UAV-based remote sensing devices in combination with machine learning (ML) methods to perform precise 3D vegetation mapping followed by the classification of areas containing buried radioactive materials (Figure 6). As sensor technologies, light detection and ranging (lidar) and multispectral (MS) imagery were identified as promising solutions. Because airborne lidar is able to penetrate vegetation, it has been widely used for DTM extraction – even in overgrown areas (Sithole and Vosselman, 2004; Gevaert et al., 2018). Furthermore, small terrain height anomalies can be detected using DTM information and standard ML techniques. For instance, cultural remains have been identified in archaeological investigations (Lasaponara et al., 2010; Bollandsås et al., 2012), even when located

under dense vegetation (Guyot et al., 2018). Additionally, lidar facilitates precise 3D mapping of forests at the tree level (Reitberger et al., 2008). In combination with MS imagery, lidar has demonstrated promising potential for the calculation of forest structural variables (Latifi and Heurich, 2019).

In the context of risk management, precise vegetation maps can assist in forest fire simulations and help to create action plans to mitigate these hazard situations more efficiently. Regarding the ChEZ study area, vegetation maps can additionally be utilized to build complex radiation models. For instance, Molitor et al. (2018) simulated workers' radiation dose uptake in the ChEZ. However, until now, their models have been rather basic and require improvement to achieve the most realistic possible results. By incorporating precise 3D models of the vegetation, simulations could even consider the individual radiation levels of single trees.
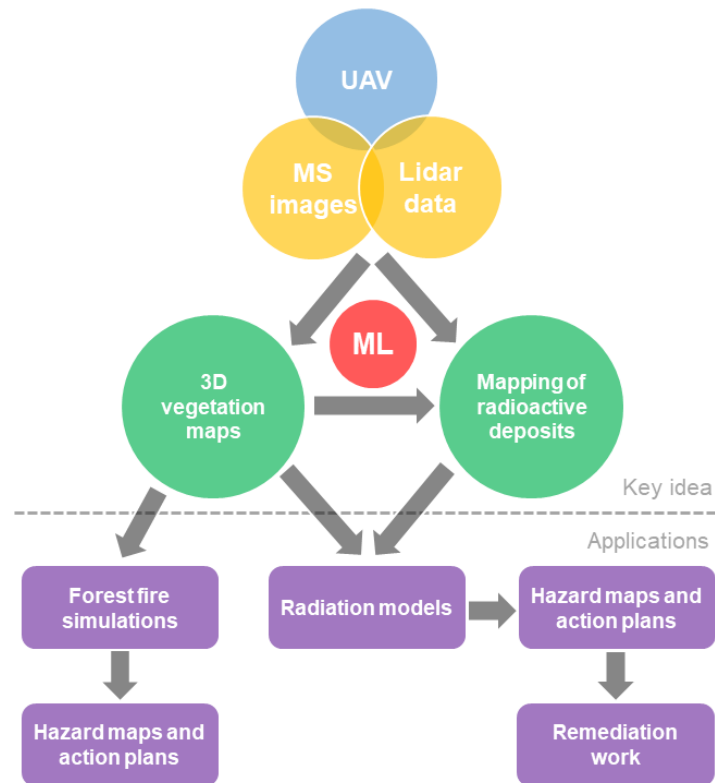
Figure 6: Key idea of this PhD thesis and possible applications for risk management.

1.5. Research background and objectives

The research reported in this PhD thesis is part of the project GeoFlyer, which is funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung). The project aims at optimizing an unmanned aircraft system for the mapping of remote hazard areas. In accordance with the project goals, this PhD thesis covers two distinct subjects. Utilizing UAV-based optical sensors and ML algorithms, the detection of radioactive waste sites in the ChEZ was investigated. Furthermore, this thesis focuses on the DL-based classification of single trees from high-resolution remote sensing data. The key objectives are presented in the following paragraphs. For more details on the individual studies, please refer to the corresponding chapters.

Initially, a pilot study (see Chapter 2) is performed based on the data of a one-week mission conducted in the ChEZ. The main objective is to investigate the feasibility of UAV-based detection of unknown radioactive waste sites using lidar data and MS imagery. The area of focus is the densely vegetated and thus challenging area in the Red Forest, where historical maps have only roughly indicated the existence of trenches. Experiments are conducted using a random forest (RF)-based classifier to explore the effectiveness of different handcrafted feature subsets. The goal is to implement an appropriate feature selection method to prevent model overfitting and achieve accurate and reliable results. In the following comprehensive study (see Chapter 3), the overall objective is to improve the detection of trenches and clamps. Therefore, a second mission is performed in the ChEZ, optimizing the experimental setup and the findings obtained in the pilot study. Regarding the data collection, simultaneous recording of lidar data and MS imagery in the same flight mission is to be enabled. Thus, an extension of the sensor setup is used to facilitate the generation of overlapping high-resolution datasets. A further goal is the implementation of a data-driven correction method for the radiometric calibration of the laser scanner. In terms of classifier performance, one main purpose is to optimize the detection process by extending the feature set. Based on the enriched feature set, the simultaneous detection of both trenches and clamps is tackled. Finally, the classification results are to be evaluated by on-site test drillings.

The second part of this PhD thesis focuses on DL-based methods for tree species mapping. To achieve this goal, a preliminary study is performed using raw point clouds (Figure 7) for tree species classification (see Chapter 4). Lidar data and MS imagery are used to adapt Point-Net++ to the classification of presegmented single trees (Figure 8) with respect to tree species and standing dead trees. Besides 3D coordinates, the objective is to integrate additional point attributes such as

lidar echo pulse width, surface normals and MS features. Additionally, data augmentation and hyperparameter tuning are to be incorporated into the network training.
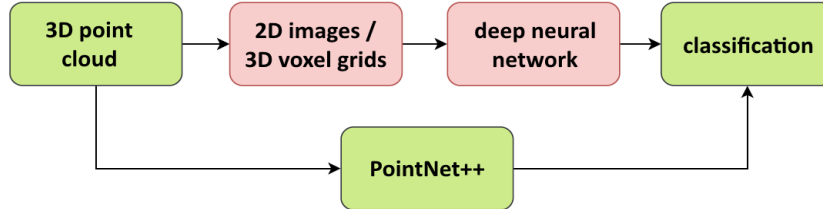
Figure 7: Basic principle of 3D deep neural networks (DNNs) such as Point-Net++, operating directly on 3D point clouds without a prior transformation into 2D images or 3D voxel grids (Briechle et al., 2019).
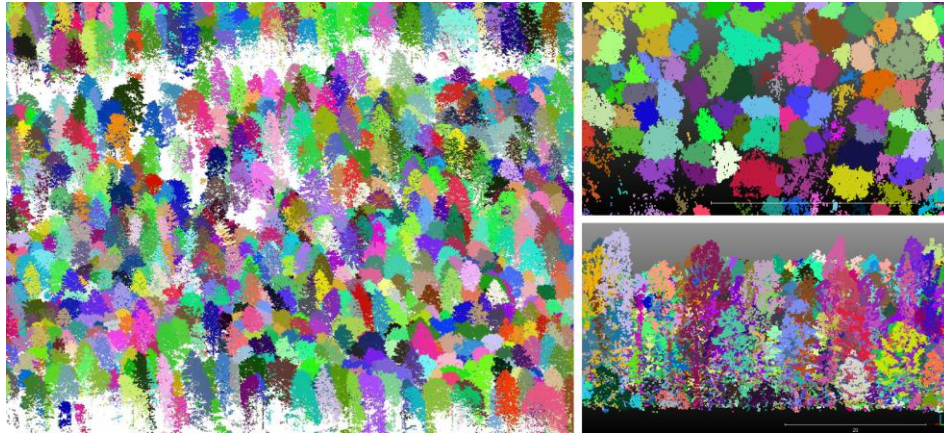
Figure 8: Segmented single trees in oblique (left), bird's eye (top right), and side view (bottom right); random colour rendering.

As mentioned above, training PointNet++ from scratch requires a huge set of training data and is, thus, rather inappropriate for relatively small datasets. In contrast, 2D convolutional neural networks (CNNs) are widely utilized in combination with transfer learning, enabling quick adaptation to new tasks – even on clearly reduced datasets. Thus, the idea is to fine tune a CNN model using pretrained weights for network initialization and recursive retraining of a subset of network layers. Consequently, the objective of the final study of this thesis (see Chapter 5) is to explore the use of 2D CNNs for the combined classification of presegmented 3D tree objects with respect to tree species and standing dead trees. The study includes experiments focusing on the number of samples and varying geometric and spectral resolutions. One main objective is the generation of appropriate 2D representations

from the lidar point clouds of single trees. Additionally, laser echo characteristics (EC) and MS information are to be integrated. The most important study aim is designing a network architecture in order to process images rendered from both sensor datasets. Finally, using identical data from two study areas, the 2D- and 3D-based DL approaches are verified and compared to reveal their advantages and disadvantages.

1.6. Structure of the thesis

This dissertation is comprised of six chapters. Enclosed by the introduction and the synthesis, Chapters 2 through 5 present peer-reviewed journal and conference papers covering both the detection of radioactive waste sites (Chapters 2 and 3) and DL-based vegetation mapping (Chapters 4 and 5). Each of these four chapters is a standalone work with its own research objectives, materials, methodology, experiments and results, discussions and conclusions. Thus, there is considerable overlap between these contributions in terms of the background, datasets and related works. Nevertheless, if a reader is only interested in a specific chapter, there is no need to consult another part of this thesis for its full comprehension.

2 UAV-based detection of unknown radioactive biomass deposits in Chernobyl's Exclusion Zone[1]

Abstract

Shortly after the explosion of the ChNPP in 1986, radioactive fall-out and contaminated trees (so-called Red Forest) were buried in the ChEZ. These days, exact locations of the buried contaminated material are needed. Moreover, 3D vegetation maps are necessary to simulate the impact of tornados and forest fires. After 30 years, some of the so-called trenches and clamps are visible. However, some of them have become overgrown and have slightly settled in the centimetre and decimetre range. This chapter presents a pipeline that comprises 3D vegetation mapping and ML methods to precisely map trenches and clamps from remote sensing data. The dataset for our experiments consists of UAV-based lidar data, MS data, and aerial gamma-spectrometry data. Depending on the study areas, OA values ranging from 95.6% to 99.0% were reached for the classification of radioactive deposits. Our first results demonstrate an accurate and reliable UAV-based detection of unknown radioactive biomass deposits in the ChEZ.

2.1. Introduction

In the first months after the Chernobyl accident large parts of contaminated material resulting from clean-up operations were buried into nine RWTSPs. On the one hand, these clean-up measures resulted in a decrease of the external exposure dose rate by a factor of about 10. Parts of the RWTSPs were also re-vegetated in order to stabilize the surface covers and to reduce the effect of re-suspension of contaminated dust particles by wind. On the other hand, however, these clean-up operations led to a higher risk of groundwater contamination. These RWTSPs were created in the vicinity of the ChNPP and contain about 1,000 excavated trenches and/or clamps in which radioactive material was buried and covered with a clean soil layer (Molitor et al., 2017). Unfortunately, a comprehensive and complete documentation of the volumes, the radionuclide inventories, the exact positions, and the exact total number of trenches and clamps is not available. In order to justify and substantiate action plans to mitigate actual and future hazards, it is extremely important to have appropriate up-to-date descriptions and evaluations on the current radiological situation and its evolution.

Since the explosion of the ChNPP a lot of research investigation from different special fields have been carried out in the ChEZ. To name just a few: Kashparov et al. (2012) present studies on the radionuclide behaviour in the environment after the disposal of radioactive waste into shallow subsurface storages. The main results describe the biogenic migration of radionuclides from subsurface storages into plants. Bugai et al. (2012) focused on the hydrogeological characterization and

groundwater transport modelling to describe the migration of radionuclides over a 16-year period (1986–2002). Besides these research studies, the European Union financed a project called "Support to radioactive waste management in Ukraine" (EU, 2014–2017). One purpose of this project was the investigation of radioactive waste burial. Partially, the RWTSPs have already been well investigated. Though, some other RWTSPs do not have such detailed information about the location and radioactivity of inventory – these are subject for ongoing investigation efforts. In 2015, airborne geophysical surveys proved that the mapping of burials, especially in low vegetated areas, is possible using high-resolution gamma field radiation measurements. The results prove that the local gamma ray intensity significantly exceeds the level of the total background radiation in contaminated areas. In vegetated areas, the biomass significantly distorts the measurements because the radiation of the vegetation can not be taken into account during the calibration process (Zabulonov et al., 2015).

The experiments with the gamma spectrometer conducted so far in the forested areas of ChEZ revealed that the interpretation of the spectrometer measurement signal is difficult and not yet solved. Thus, other sensor systems are to be envisaged supplementary to detect buried contaminated material. As these deposits can be characterized by slight height anomalies of the ground (trenches and clamps) the idea is to advantageously use an UAV-based lidar system: On the one hand, this measurement technique operates at a sufficient safety distance, which is of prime importance considering the study area. On the other hand, UAV-based lidar permits a precise 3D reconstruction of the tree landscape as well as the detection of small terrain height anomalies. It even allows the creation of 3-dimensional contamination maps by taking into account the tree specific radiation. In non-forested areas some of the known deposits are covered with low vegetation that – compared to the neighbourhood – shows anomalies w.r.t the species and the green tone. Hence, we also used a UAV-mounted high resolution MS camera to collect image data in the visible and near-infrared (NIR) spectrum that provides additional information for the classification process. To sum it up, the key idea of this research is to combine UAV-based lidar, MS images and airborne gamma spectrometer measurements to automatically map unknown buried radioactive deposits in the ChEZ.

In the following sections we address (ii) the drone systems, the sensors and the preprocessing of the data, (iii) the entire processing pipeline including the 3D vegetation mapping and the classification scheme for the radioactive burials, (iv) the experiments and results, (v) discuss our results and (vi) conclude and give an outlook to potential future research topics.

2.2. Materials

2.2.1. UAV systems

Persuading results from UAV surveys start with the selection of the appropriate sensors to perform the requested task. Nevertheless, the capability of the carrier system will decide on the performance of the survey both in terms of quality and quantity. To achieve our objectives we relied on two different UAVs. For the lidar measurements an octocopter was available that was developed by a team of the department of nuclear physics technologies of the Institute of Environment Geochemistry at the National Academy of Sciences of Ukraine. This octocopter was already used for the airborne geophysical surveys mentioned above. As carrier for the MS camera we used the Quantum Trinity VTOL (vertical take-off and landing) fixed-wing system. The hybrid concept of this VTOL drone combines the advantages of multi-copters and fixed-wing drones. Using multiple propeller arrays or tilt rotors, it can take off and land like a rotary wing drone. Once airborne, it can transition to horizontal flight and work as a fixed-wing drone.

2.2.2. Lidar data

As lidar system we used the YellowScan Mapper laser scanner mounted on the octocopter carrying the payload of 2.2 kg (excl. copter batteries) at a maximum mission time of around 20 minutes. In order to get flight trajectories with centimetre precision a Global Navigation Satellite System (GNSS) base station was set up to collect simultaneous measurements that were used in DPGS processing. Seven flights were conducted in four selected areas of ChEZ in November 2017 in leaf-off situation. The measurement rate was 18.5 kHz, the flights were realized at a speed of 4–7 m/s and an altitude of 50 m resulting in a nominal point density of 25–40 points/m² (side lap 50%). The boresight parameters of the laser scanner provided by the manufacturer were checked using data of a special calibration flight over a building and were fully verified in the subsequent strip alignment. All the strips of a flight mission were aligned to achieve a consistent lidar point cloud using the software package BayesStripAlign from BayesMap Solutions. At average, the mean discrepancies between the neighbouring strips were around 5 cm. This means that the boresight parameters had been successfully adjusted and the strip trajectories appropriately mapped. An absolute 3D georeferencing was achieved by fitting the lidar point clouds to the enclosing polygons of buildings. Finally, ground points were filtered from the lidar point clouds and were subsequently interpolated into a DTM grid with a grid size of 0.5 m.

2.2.3. MS data

To capture MS data we used the Tetracam ADC Snap camera mounted on the Quantum Trinity VTOL system. The ADC Snap camera captures light wavelengths in three bands between 520 nm and 920 nm (Green, Red, NIR). The flight duration of the Trinity drone was around 50 minutes at a mean speed of 17 m/s (flying altitude 130 m, side lap 40%, end lap 80%). All flights were carried out in fully automatic mode to avoid the direct trespassing of contaminated areas. Dense photogrammetric point clouds (15 points/m²) and orthomosaics (ground sampling distance (GSD) 8 cm) were generated from the MS images using standard structure-from-motion Software. Due to missing ground control points (GCPs) in a few inaccessible areas two photogrammetric point clouds were registered to the lidar data by manual measurements of GCPs with a standard deviation (std) of 70 cm.

2.2.4. Gamma spectrometry data

In addition to the lidar data and the MS images, measurements of an aerial gamma-ray spectrometer survey were available as selective data. The UAV-based aero-gamma spectrometry system (see Figure 9) enables high-resolution mapping of radiation contamination without any risk to human health. The spectrometer data were collected by the airborne gamma spectrometric complex "ASPEK". The on-board unit is intended to perform geo-referenced measurements of the gamma radiation spectra as well as altitude, pressure and temperature measurements. The spectrometer consists of five blocks of gamma radiation detectors based on thallium-doped sodium iodide scintillator crystals with a size of 63x63 mm (Zabulonov et al., 2015). These crystals emit light as soon as gamma rays interact with the atoms in the crystals. The intensity of the produced light can be measured and is proportional to the energy deposited by the gamma rays (Melcher, 2000). All blocks work synchronously and the signals of all detectors are summarized. The measurement system allows real time control over the unit sensitivity by turning on/off single detector blocks – depending on the required sensitivity. With a total weight of about 7.5 kg (excl. batteries) and a dimension of 300 mm x 300 mm the spectrometer can be installed on an octocopter (Zabulonov et al., 2017). The flights were performed in parallel tracks with a track-to-track distance of 100 m. The average altitude was 30 m at an average flight speed of 5 m/s and a measurement interval of 1 s. The high payload of the detector array does not allow any physical collimation. This fact results in a spectrometer measurement angle of 120 degrees with a footprint of around 100 m. The preprocessing of the spectrometry data contained the following steps: First, the total gamma intensity (*gamma_total*) was computed. In a second step, the background gamma intensity was calculated using a sliding average method. Finally, the local component of the

gamma intensity (*gamma_local*) was derived as the difference of the total gamma intensity and the background gamma intensity (Zabulonov et al., 2015). It needs to be mentioned that the measured signal is the sum of different signals, i.e. the natural radiation from space and ground, the radiation from close-range contaminated vegetation, the radiation from fallout on the ground, and from buried biomass.

Figure 9: UAV-mounted gamma spectrometer (Zabulonov et al., 2017).

2.2.5. Study areas

Three different study areas are presented in this chapter (see Figure 10). Study area "Yanov Station 3.3/3.5/3.7" (size around 1 km²) was just captured with the MS camera. Most of the clamps are clearly visible because they are not covered with high vegetation. In the second study area "Red Forest 2.5" (area size 6 ha), dense vegetation (birches and pines with a height up to 21 metres) covers the ground and makes it impossible to detect trenches and clamps just using an airborne photogrammetric sensor. In this case, the lidar sensor is superior because of the penetration of the laser beams to the ground. Study area "Yanov Station 3.3" (area size 15 ha) is the former railway freight terminal, which has become partly overgrown with trees with a height up to 27 metres. In this area, no photogrammetric data are available.
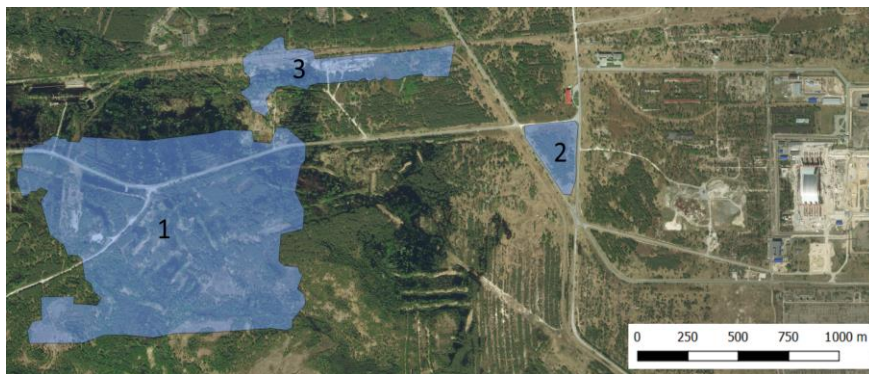
Figure 10: Locations of the three study areas in the ChEZ (source: bing map ©Microsoft Corporation).

2.3. Methods

Depending on the tree species parts of the contaminated material are accumulated in the above-ground biomass. The radioactive deposits are characterized by slight ground settlements or elevations, vegetation anomalies and a higher rate of gamma decays. Thus, the key idea is to classify the deposit areas from salient features calculated from the bare ground height, the vegetation, the gamma radiation and the reflection data acquired in the visible and NIR channels.

2.3.1. 3D vegetation mapping

Based on the normalized cut algorithm single trees are automatically segmented from the lidar data resulting in tree position, tree height, crown volume and crown base height (CBH) for each tree (Reitberger et al., 2009). The tree segmentation is evaluated by visual interpretation to find the best normalized cut threshold since no tree reference data are available. Subsequently, the single trees are classified w.r.t. to the two main tree species birch (*Betula pubescens*) and pine (*Pinus sylvestris*) using an RF classifier. The feature set (48 features) comprises radiometric and geometric features, e.g. height dependent and density dependent features, the crown shape and the mean intensity of laser points of a single tree per height layer.

Since no destructive sampling of reference trees is available we need to estimate the tree biomass solely from allometric equations using the tree height and tree species. First, we calculate the diameter at breast height (DBH) from the tree height (Widlowski et al., 2003). Next, we convert the DBH into the total above-ground biomass using the equations in Repola (2009).

2.3.2. Detection of trenches and clamps

In order to automatically detect trenches and clamps it is necessary to generate meaningful features from the available data and apply the classification based upon a well-defined evaluation strategy.

2.3.2.1. Feature description

A meaningful feature to describe terrain height anomalies is the normalized height. For every DTM grid position, the corresponding absolute terrain height is reduced by the height trend (mean height in a surrounding quadratic area of 50 x 50 m). Since contaminated areas are supposed to be characterized by vegetation anomalies multiple features are calculated to describe the vegetation. One set of features are the so-called lidar metrics (LM; Næsset, 2004). These height- and density-dependant features are calculated for eight different height layers (0–1.5 m, 1.5–5 m, 5–12 m, >12 m, >2 m, 0.5–2 m, 0.5–5 m, >0 m) for a cell size of 5 x 5 m. Finally, the metrics are resampled to the DTM

grid size by bi-linear interpolation. Another feature set describing the vegetation comprises 14 so-called tree features (TF; see Table 1), which are calculated for every grid point of the DTM. The close-up trees in a circular area are detected by applying a range search within a radius of 30 m. Available features that characterize the gamma radiation are the total intensity and the local component of the gamma spectrometry measurements. These data are available as single measurements (point distance around 5 m) along the UAV flight tracks and are bi-linearly interpolated on the NTM grid. Additionally, MS data (orthomosaic) are available for study area 1. The data of the three channels (Green, Red, NIR) are also bi-linearly interpolated on the DTM grid. All in all, the generated features can be grouped into five feature subsets (see Table 2). The total feature sets for the particular experiments are individually generated w.r.t. the characteristics of the study areas and the availability of remote sensing data.

Table 1: Definition of TF.

TF#	Definition
1	Number of trees
2	Mean tree height [m]
3	Std of tree height [m]
4	Mean CB_height [m]
5	Std of CB_height [m]
6	Pine tree ratio (# pine trees / # all trees)
7	Sum of crown volumes [m³]
8	Mean crown volume [m³]
9	Std of crown volume [m³]
10	Mean DBH [cm]
11	Std of DBH [cm]
12	Sum of above-ground biomass [kg]
13	Mean above-ground biomass [kg]
14	Std of above-ground biomass [kg]

2.3.2.2. Classification and evaluation strategy

Prior to the RF classification, training and test areas are manually defined and labelled according to their classes by visual interpretation of the datasets. Partially available RWTSP reference data also support the labelling. The training set is balanced and highly correlated redundant features are removed from the feature set to avoid a deterioration of the classification result. First, we calculate the correlation coefficient from the covariance matrix of all features. Second, we eliminate features with a correlation coefficient larger than 0.75. An RF classifier is trained including a recursive feature elimination (RFE) to optimize the feature set by discarding non-important features. The classifier is evaluated by the OA and κ values by a repeated five-fold cross-validation

(CV). Moreover, the mean decrease in accuracy (proportion of observations that are incorrectly classified by removing the feature) is used to evaluate the relevance of the individual features. The class predictions for the test data are estimated and the prediction quality is checked using the OA, precision and recall. Finally, the class probabilities beyond a certain threshold (i.e. 95%) are visualized for the particular study areas.

Table 2: Available feature subsets in the study areas.

Study area	1	2	3
Normalized height	X^2	X	X
LM	X^2	X	
TF		X	
Gamma spectrometry features	X	X	X
MS features	X		

2.4. Experiments

2.4.1. Vegetation mapping and biomass estimation

Based on the segmentation of single trees the two tree species (birch, pine) were classified using an RF classifier. The classifier was trained using a reference dataset in study area 2 consisting of 216 manually labelled, well-balanced tree species and a sample ratio of 0.7 (70% for training, 30% for test). A five-fold CV showed an OA of 95.4%, a κ value of 91.4%, a precision of 94.3%, and a recall of 97.1%. After a feature relevance assessment the mean intensity of the tree crowns turned out to be the most relevant feature.

The DBH parameters were estimated from the tree heights using allometric functions and were subsequently used to estimate the total above-ground biomass. Figure 11 and Figure 12 show the segmented trees for the study area Red Forest 2.5 coloured in dependence on the tree species and on their total above-ground biomass, respectively.

---

[2] For study area 1 (Yanov Station 3.3/3.5/3.7), the normalized height and the LM are calculated from the photogrammetric point cloud due to missing lidar data.
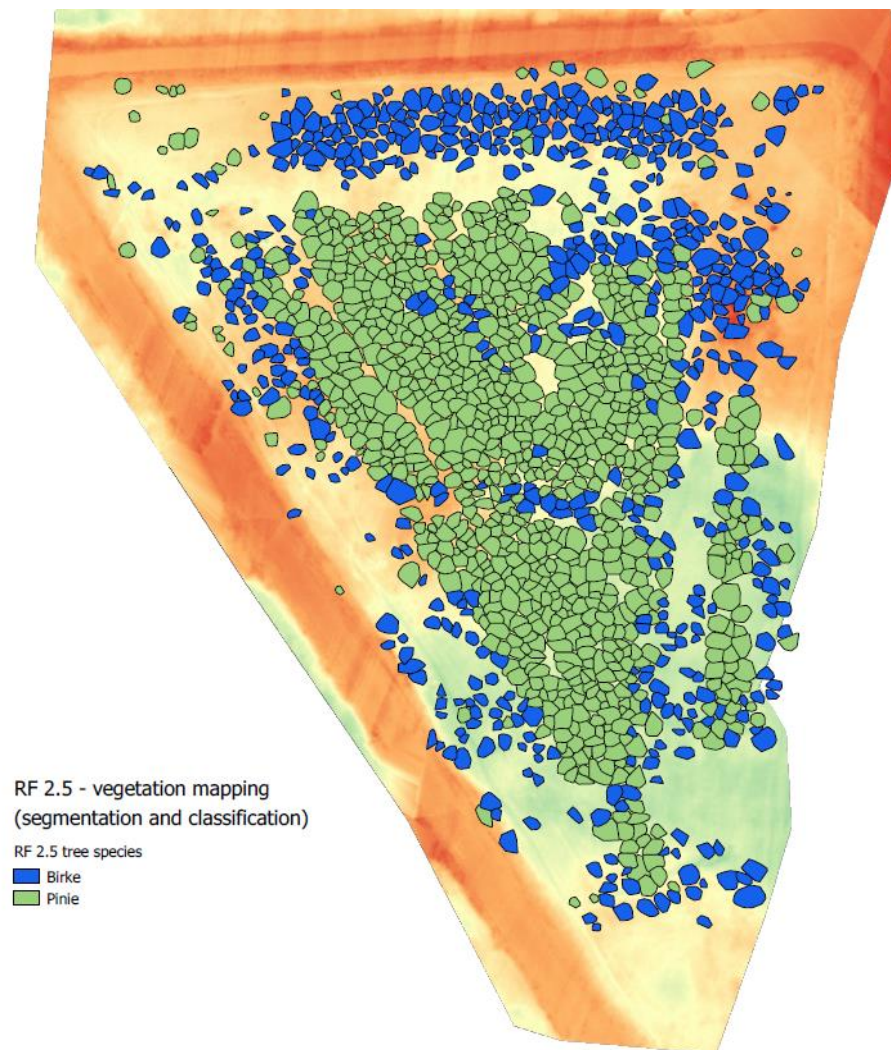
Figure 11: Study area 2: Tree species classification result.
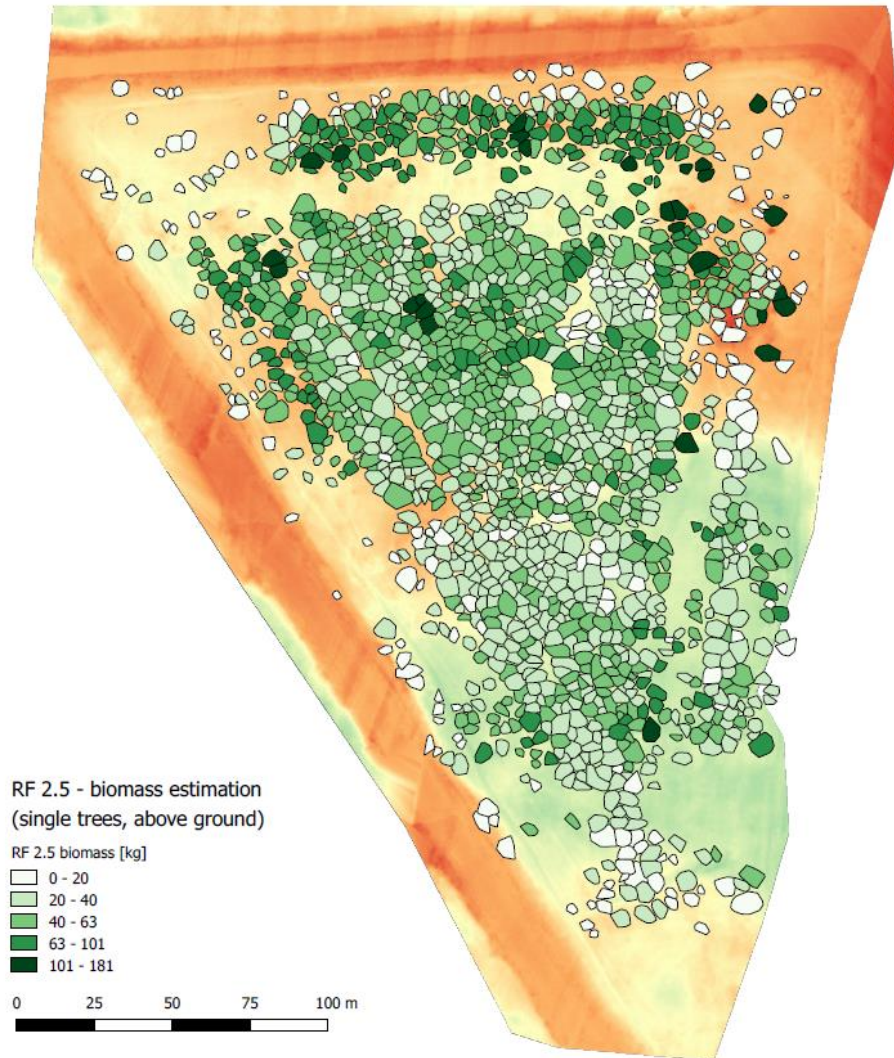
Figure 12: Study area 2: Biomass estimation result.

2.4.2. Detection of trenches and clamps

The individual feature subsets in the study areas that were used as input for the feature selection process are depicted in Table 2. Table 3 gives an overview of the number of effectively applied features in the classification process (used #feat), the class names, the number of training and test samples, the five-fold CV results of the training da-taset and the results of the class prediction of the test data (prec. = precision, rec. = recall). In the following subsections, the results of the different study areas are presented in detail.

2.4.2.1. Study area 1: Yanov Station 3.3/3.5/3.7

During the feature selection the gamma spectrometry features, the green channel as well as huge parts of the LM subset were removed. The RFE process resulted in 16 remaining features out of 22 uncorrelated features. The feature importance depending on the mean decrease in accuracy showed that – as expected – the normalized height and the NIR channel are by far the most relevant features. Class predictions and class probabilities were calculated for the test area (Figure 13, *left*). The prediction result of class probability values of more than 95% for the class "clamps" are presented in Figure 13 (*right*). The detected clamps in this study area have already been known to the ChEZ authorities. These relatively high, non-vegetated clamps can partly be seen from the non-contaminated roads and even in satellite images. Nevertheless, the results (OA=95.6%, κ =89.5%, precision=93.6%, recall=91.7%; see Table 3) show that our method is feasible to detect this kind of clamps that are solely captured by MS cameras.
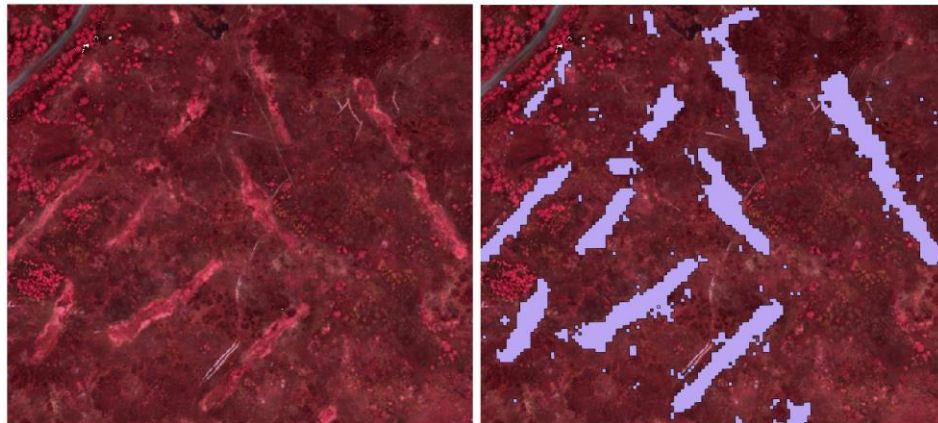
Figure 13: Study area 1: Test area (left); Classification result for class "clamp", class probability values >95% (right).

2.4.2.2. Study area 2: Red Forest 2.5

In the area Red Forest 2.5, the RFE dropped 5 of the 17 uncorrelated features. Four LM as well as the gamma local component were removed during this step. It is not surprising that the normalized height was again the most important feature. Besides eight of the LM, three TF remained in the feature set. In this study area with dense vegetation, two TF are even ranked under the top five. Obviously, vegetation anomalies occur in the trench areas since the standard deviations of the crown volume, the above-ground biomass, and the DBH turned out as prominent features. The classification result showed excellent accuracy (OA=99.0%, κ =98.0%, precision=99.4%, recall=98.6%; see Table 3). The filtered DTM as well as the prediction result of the study

area 2 are presented in Figure 14 and Figure 15. In the area Red Forest 2.5, four so far unknown trenches overgrown with quite dense vegetation were detected. The length of these trenches ranges between 50 m and 100 m with a width up to 5 m and a depth up to 1.5 m.

RF 2.5 - Digital Terrain Model (DTM)

orthometric height [m]

▣	111
▣	111.5
▣	112
▣	112.5
▣	113
▣	113.5
▣	114
▣	114.5
▣	115
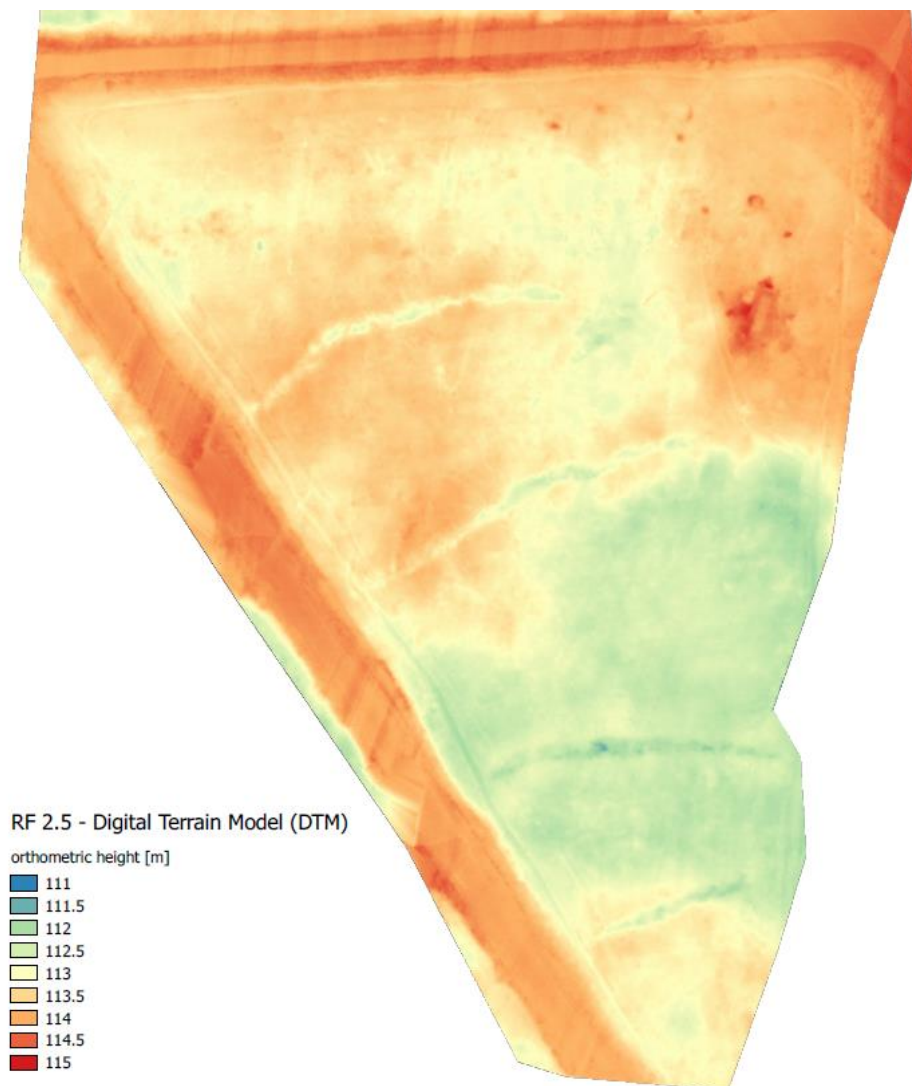
Figure 14: Study area 2: Filtered DTM.
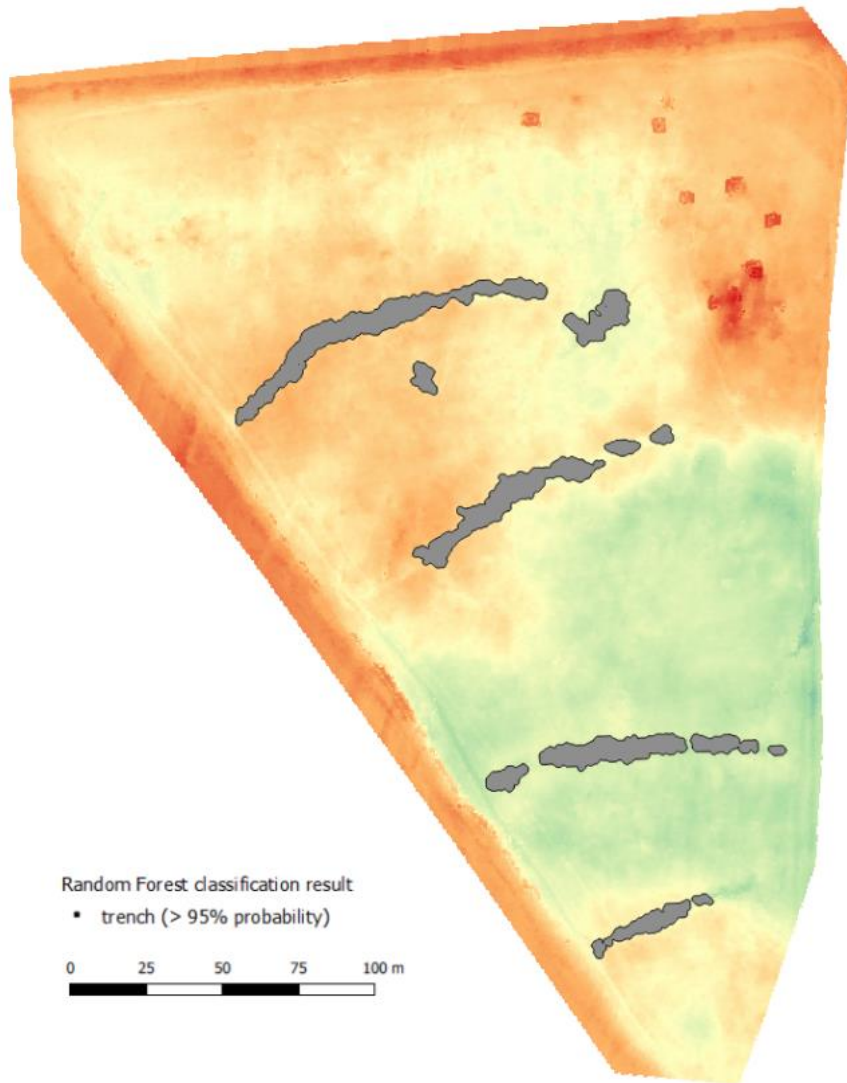
Figure 15: Study area 2: Classification result for class "trench", class probability values >95%.

2.4.2.3. Study area 3: Yanov Station 3.3

In the area Yanov Station 3.3, an initial visualization showed a potential correlation between the terrain height anomalies and the local component of the gamma spectrometer measurements. In this study area, the fairly sparse vegetation coverage motivates the usage of the bilinearly interpolated spectrometer data. The impact of the low vegetation on the measured radiation signal is assumed to be negligible. In other words, we suppose that the gamma spectrometer measurement

is a gamma ray superimposition of the buried radioactive material, the constant impact of the top soil layer, and the constant impact of the background radiation from space. Due to lacking reference data a test dataset was manually generated taking into account the information of the normalized height map – overlaid with the local component of the gamma spectrometer measurements (Figure 16, *top*). The selection of features is not necessary on a feature set consisting of only two features (*normalized height*, *gamma_local*). The local component of the gamma spectrometer measurements turned out as the five times more important feature in the RF feature importance assessment using the mean decrease in accuracy. The parameters describing the classification quality are very satisfying (OA=98.4%, κ=95.5%, precision=95.9%, recall=97.3%; see Table 3). The prediction result for the study area 3 is presented in Figure 16 (*bottom*), showing class probability values of more than 99% for the class "contaminated". In conclusion, in the partly overgrown area of Yanov Station 3.3, the contaminated areas can be automatically detected using the local component of the gamma spectrometer measurements and the normalized height as features.
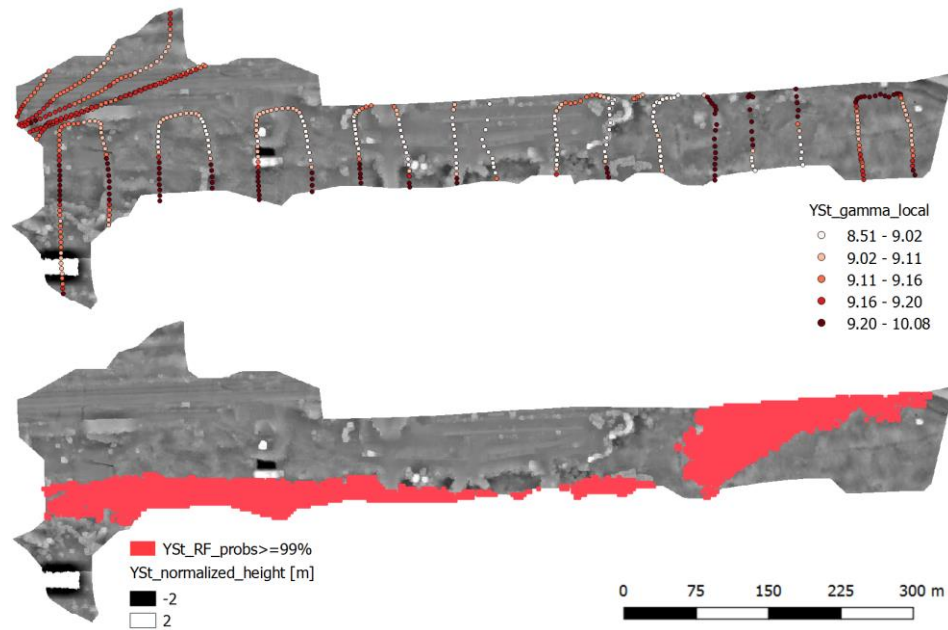
Figure 16: Study area 3: Normalized height map and local component of the gamma spectrometer measurements [μSv/h] (top); Normalized height map and classification result for class "contaminated", class probability values >99% (bottom).

Table 3: Overview of the RF classification parameters and results.

Study area	1	2	3
Used #feat	16	12	2
Class names	1: clamp 2: no clamp	1: trench 2: no trench	1: contaminated 2: non-cont.
Training samples	83643	2473	86862
Five-fold CV (training data)	OA = 98.8% κ = 94.6%	OA = 99.8% κ = 99.6%	OA = 98.5% κ = 96.6%
Test samples	52,724	2,446	120,151
OA [%]	95.6	99.0	98.4
κ [%]	89.5	98.0	95.5
precision [%]	93.6	99.4	95.9
recall [%]	91.7	98.6	97.3

2.5. Discussion

The classification accuracy of more than 95% for DTM grid points potentially located in trenches and clamps is fairly high in all study areas (see Table 3). Our concept proves that the UAV-based detection of unknown radioactive biomass deposits in the ChEZ is successful for both MS data and lidar data. The normalized height appears to be the most important feature for the selected areas. Furthermore, the results in study area 2 (Red Forest 2.5) show that the features derived from the 3D vegetation mapping have also a significant impact on the classification result in forested areas. In non-forested areas such as study area 1, the MS features (mostly NIR) are important to classify contaminated clamps.

The integration of gamma spectrometry features into the classification process is problematic since the measured radiation is a superposition of different effects. As to be expected, our results verify that bi-linearly interpolated gamma spectrometry data is suitable in study area 3. Similar results were expected for study area 1, but it looks like that the use of bi-linearly interpolated gamma spectrometry data is not applicable here since the spectrometry features significantly deteriorate the classification result. One explanation might be that the deposits are shielded by a thick soil coverage. A feature assessment in study area 2 (Red Forest 2.5) leads to the assumption that gamma spectrometry data cannot be utilized as bi-linearly interpolated data. Especially in the forested areas, a model with a higher complexity is needed to remove the radiation impact of the above-ground biomass from the gamma spectrometry data and to isolate the radiation impact generated by the buried radioactive material.

An internal technical report about the results of the UAV gamma spectrometry surveys in 2015 (Molitor, 2017) states that burials can be identified by analysing anomalies of the local component of the gamma

intensity. The probability ranges from more than 50% in vegetated areas to more than 90% in low vegetated areas. Our results demonstrate that comparable results can be achieved for low vegetated areas just using MS data (study area 1). However, in vegetated areas (study area 2), considerable improvement was reached – as expected – by taking advantage of lidar data. Thus, UAV lidar is mandatory to detect trenches and clamps in forested areas because of its capability to penetrate vegetation.

2.6. Conclusion and outlook

The experiments prove that the UAV-based lidar and MS image technology in combination with aerial gamma spectrometry surveys can successfully map unknown deposits of buried radioactive biomass in the ChEZ. Excellent RF classification results for the detection of trenches and clamps were achieved by fusing geometric features, newly developed vegetation features, gamma spectrometer measurements and MS image data. To minimize workers' dose uptake and to optimize remediation work, the existing hazard maps were verified and upgraded. Future work will focus on the generation of overlapping datasets combining lidar data and MS image data for overgrown and non-overgrown areas. This data fusion enables the advantages of two sensor technologies in different situations:

- Lidar data: accurate DTM, LM and TF in forested areas
- MS data: improved classification results, especially in non-forested areas

As far as the methodology is concerned, the feature set shall be supported by adding e.g. point feature histograms (Rusu et al., 2008), additional 3D and 2D features (Weinmann et al., 2014) and shape distributions (Osada et al., 2002) to further improve the classification result. Above all, research is intended to model the impact of the biomass to the gamma spectrometer measurements and to finally dissolve the impacts of the different radiation sources.

Acknowledgements

3 Detection of radioactive waste sites in the Chernobyl Exclusion Zone using UAV-based lidar data and multispectral imagery[3]

Abstract

The severe accident at the ChNPP in 1986 resulted in extraordinary contamination of the surrounding territory, which necessitated the creation of the ChEZ. During the accident, liquidation materials contaminated by radioactive fallout (e.g., contaminated soil and trees) were buried in RWTSPs. The exact locations of these burials were not always sufficiently documented. However, for safety management, including eventual remediation works, it is crucial to know their locations and rely on precise hazard maps. Over the past 34 years, most of these trenches and clamps have been exposed to natural processes. In addition to settlement and erosion, they have become overgrown with dense vegetation. To date, more than 700 burials have been thoroughly investigated, but a large number of burial sites (approximately 300) are still unknown. In the past, numerous burials were identified based on settlement or elevation in the decimetre range, and vegetation anomalies that tend to appear in the immediate vicinity. Nevertheless, conventional detection methods are time, effort and radiation dose intensive. Airborne gamma spectrometry and visual ground inspection of morphology and vegetation can provide useful complementary information, but it is insufficient for precisely localizing unknown burial sites in many cases. Therefore, sensor technologies, such as UAV-based lidar and MS imagery, have been identified as potential alternative solutions. This chapter presents a novel method to detect radioactive waste sites based on a set of prominent features generated from high-resolution remote sensing data in combination with an RF classifier. Initially, we generate a DTM and 3D vegetation map from the data and derive tree-based features, including tree density, tree height, and tree species. Feature subsets compiled from normalized DTM height, fast point feature histograms (FPFH), and LM are then incorporated. Next, an RF classifier is trained on reference areas defined by visual interpretation of the DTM grid. A backward feature selection strategy reduces the feature space significantly and avoids overfitting. Feature relevance assessment clearly demonstrates that the members of all feature subsets represent a final list of the most prominent features. For three representative study areas, the mean OA is 98.2% when using area-wide test data. Cohens' kappa coefficient κ ranges from 0.609 to 0.758. Additionally, we demonstrate the transferability of a trained classifier to an adjacent study area (OA = 93.6%, κ = 0.452). As expected, when utilizing the classifier on geometrically incorrect and incomplete reference data, which were generated from old maps and orthophotos based on visual inspection, the OA decreases significantly to 65.1% (κ = 0.481). Finally, detection is verified through 38 borings that successfully confirm the existence of previously unknown buried nuclear materials in classified areas. These results

demonstrate that the proposed methodology is applicable to detecting area-wide unknown radioactive biomass burials in the ChEZ.

3.1. Introduction

On April 26, 1986, an explosion of Reactor Unit 4 at the ChNPP, located approximately 100 km north of Kiev, Ukraine, was recorded as a major accident according to the International Nuclear and Radiological Event Scale. To re-establish safety, the surrounding population was evacuated from severely affected areas, and the ChEZ was created. This zone remains in effect today. Based on the extraordinary contamination caused by the spread of radioactive materials, the forest in the fallout trails adjacent to the ChNPP turned reddish-brown and died (hereafter referred to as the Red Forest). During the implementation of liquidation measures, the contaminated biomass and topsoil were buried on-site in trenches and clamps, then covered by a clean soil layer with a typical thickness of approximately 1 m. These burials were created in areas called RWTSPs. A total of nine RWTSPs were created in the vicinity of the ChNPP, consisting of approximately 1,000 trenches and clamps (Molitor et al., 2017). To stabilize the surface cover and reduce re-suspension of contaminated dust particles by wind, many parts of the RWTSPs were re-vegetated. Initially, clean-up actions decreased the external exposure dose rate significantly. However, based on the absence of appropriate barriers, burials in RWTSPs do not provide sufficient protection against radiation mobilization into groundwater. Unfortunately, accident liquidation measures were performed under extraordinary boundary conditions. Therefore, no detailed systematic documentation of the RWTSPs is available. Most current information relies on improvised records that were created during or after the liquidation measures. Thus, a comprehensive overview and precise mapping of trenches and clamps (exact number, location, geometry, contents, etc.) are necessary to improve safety during maintenance and facilitate the eventual remediation of burials.

The main objective of this study is to map unknown radioactive biomass burials in the ChEZ automatically and accurately. On-site visual inspections of known burials revealed three important facts. First, many burials can be identified based on settlement or elevation in the decimetre range or more. Second, distinctive vegetation features are also characteristic of many burials. Third, based on significant remaining contamination, a remote measurement technique is required to avoid risky trespassing of the study areas. Therefore, the key idea of this study is to utilize remote sensing devices in combination with ML methods to perform precise 3D vegetation mapping followed by the classification of areas containing buried radioactive materials. As a sen-

sor platform, a UAV can be deployed to collect data from a safe distance. Because it is able to penetrate vegetation, airborne lidar data can be utilized to extract precise DTMs, even in overgrown areas (Sithole and Vosselman, 2004; Gevaert et al., 2018). By combining DTM information with standard ML techniques, the detection of small terrain height anomalies has been performed successfully in the past. By inspecting small elevation changes, cultural remains can be identified in archaeological investigations (Lasaponara et al., 2010; Bollandsås et al., 2012), even when they are located under dense vegetation (Guyot et al., 2018). Furthermore, lidar facilitates precise 3D mapping of forests at the tree level (Reitberger et al., 2008). Because MS imagery can support tree species classification (Holmgren et al., 2008; Trier et al., 2018), we supplemented the lidar sensor with two high-resolution MS cameras to collect image data in the visible and NIR spectra.

The current chapter proposes the utilization of terrain-based and vegetation-based features in combination with an RF classifier for the area-wide detection of unknown radioactive waste sites in the ChEZ. Using UAV-based lidar data and MS imagery, we conduct a precise 3D vegetation mapping at the tree level and introduce novel tree-based features. Subsequently, we train an RF classifier to predict the class labels for each DTM grid point ("trench", "clamp", "non-contaminated"). Excellent classification results are achieved, and a backward feature selection strategy demonstrates that e.g. TF and FPFH can enhance the proposed classifier. Our results are verified by 38 borings that confirm the existence and absence of previously unknown buried nuclear materials in the classified areas (OA = 89.5%). Interestingly, at an average depth of 1–2 m, the borings hit radioactive material, exhibiting dose rates that were up to 30 times greater than those on the terrain surface. Finally, the existing hazard maps have been upgraded to minimize workers' radiation dose uptake and optimize accident liquidation.

In the following sections, we address related work, the study areas, reference data, the UAV system, sensors, and data preprocessing. Subsequently, we present the entire processing pipeline for the detection of radioactive burials and the conducted experiments, including sensitivity analysis. Finally, we discuss the results and draw conclusions from our research.

3.2. Related work

3.2.1. Risk management in the ChEZ

An EU-financed project entitled "Support to radioactive waste management in Ukraine" (European Union, 2014–2017) facilitated numerous research studies, including the investigation on radioactive waste sites.

Some RWTSPs have already been thoroughly investigated, but an estimated 300 burial sites remain unknown in terms of their exact location and geometry. These sites must be identified and characterized. Because ground dose rate measurements and electromagnetic soil conductivity surveys were not expedient, Bugai et al. (2005) successfully carried out ground-based geo-radar surveys using GPR. Using this technique, the subsurface geometry of one suspected burial site was elucidated. Based on a GPR survey, Saintenoy et al. (2017) verified the position of a trench in the RWTSP Red Forest. However, such ground-based methods necessitate considerable radiation exposure for workers. Furthermore, the GPR approach is relatively time consuming for the exploration of larger vegetated areas that are difficult to access. Zabulonov et al. (2015) conducted UAV-based geophysical surveys using a gamma spectrometer. By detecting local inhomogeneity in spectrometer data, the authors revealed that the accurate mapping of burials located in non-vegetated areas is possible using high-resolution gamma field radiation measurements (detection rate = 90%). The results demonstrated that local gamma ray intensity significantly exceeds the level of the total background radiation in contaminated areas. However, in vegetated areas, biomass substantially distorts measurements and the detection rate drops significantly to 50%. Recently, there has been increased pressure to make the ChEZ accessible again. To this end, simulations have been carried out to model workers' radiation dose uptake (Molitor et al., 2018). Nevertheless, improved models are needed to create action plans for forest maintenance and mitigate hazard situations – such as forest fires and tornadoes – more efficiently. Therefore, simulations must consider the individual radiation levels of different tree species to achieve the most realistic possible results.

3.2.2. Vegetation anomalies in the ChEZ

In the ChEZ, various studies have been carried out to explore the impact of radioactive contamination on the vegetation. Because the half-lives of Cs-137 and Sr-90 are approximately 30 years (Flynn et al., 1965), these radioactive isotopes still exist in the soil of contaminated areas and have been absorbed by the vegetation. Bugai et al. (2012) characterized the hydro-geological situation in the ChEZ and modelled groundwater transport to describe the migration of radionuclides. Furthermore, Kashparov et al. (2012) presented studies describing the biogenic migration of radionuclides from subsurface storage into plants. Moreover, Thiry et al. (2009) demonstrated that trees growing on burial sites accumulated an average of 1.7 times more Cs-137 and 5.4 times more Sr-90 in their above-ground biomass compared to trees growing off of burial sites. In a long-term observation, Arkhipov et al. (1994) explored the forest stands regarding viability and mortality

rate, re-establishment and canopy growth, and reproduction anomalies. Their results show that these parameters depend on the absorbed irradiation dose, on the age of the tree stand, and on forest composition. Moreover, Yoschenko et al. (2011) investigated over 1,100 replanted Scots pines selected from areas with strongly varying levels of radioactive contamination. The authors reported on the effects occurring on the morphological level of trees that were exposed to chronic radiation. Based on long-term studies, Tikhomirov and Shcheglov (1994) revealed that the migration of radionuclides is highly dependent on the tree species. Furthermore, the authors found out that coniferous trees are on average an order of magnitude less resistant to radioactive contamination than deciduous trees. According to Davids and Tyler (2003), spectral reflectance measurements of silver birch and Scots pine can be used to detect the effect of radionuclide contamination on the vegetation. Particularly, the authors demonstrated that certain spectral characteristics of leaves and needles correlate well with the level of radioactive contamination. All in all, vegetation anomalies (e.g., tree density, dominant tree species, tree shape, spectral reflectance of the vegetation) tend to appear in highly contaminated areas such as the vicinity of trenches and clamps.

3.2.3. Vegetation mapping

In the past, extensive research has been conducted on vegetation mapping using remote sensing data, particularly airborne laser scanning (ALS) point clouds fused with optical imagery (Latifi and Heurich, 2019). One major research focus was to establish methods that would work at the tree level. A delineation of single trees from ALS data was either performed based on a previously generated canopy height model (Pyysalo and Hyyppä, 2002; Solberg et al., 2006) or on the original 3D point cloud (Reitberger et al., 2009; Wu et al., 2016). For the classification of individual tree species, a large majority of previous studies relied on a two-step approach (Fassnacht et al., 2016). First, handcrafted feature sets describing the geometry and radiometry of single trees were extracted from the data. Second, appropriate ML classifiers were applied to categorize the single trees. For example, Yu et al. (2017) classified three tree species using MS ALS data (OA = 86%). Moreover, Shi et al. (2018a) categorized five species, fusing ALS data and hyperspectral imagery (OA = 84%). Based on the features generated from ALS data and colour-infrared (CIR) imagery, Kamińska et al. (2018) classified six tree classes (OA = 94%). Recently, Amiri et al. (2019) reported on a combined classification of tree species and standing dead trees with crowns (OA = 82%). Overall, these studies motivated the generation of handcrafted features for an RF-based classification of individual tree species in the ChEZ.

3.3. Materials

3.3.1. Study areas

In this chapter, experimental results for three different study areas with a total area of 37 ha are presented (Figure 17). Located approximately 1.5 km west of the ChNPP, these areas are situated in the RWTSP Red Forest and RWTSP Yanov Station areas. For these areas, a historical map created from the memories of so-called liquidation workers indicated the existence of possible radioactive waste sites (Figure 18). In a first UAV flight mission conducted in November of 2017, area#1 (6 ha) was surveyed using a YellowScan Mapper I laser scanner (Briechle et al., 2018). The lidar data were supplemented with MS images captured in a second mission in April of 2018. At the same time, both lidar data and MS images were collected in area#2 (8 ha) and area#3 (23 ha). With a tree density of approximately 400 trees/ha, all three study areas are densely vegetated. The main tree species are Scots pine (*Pinus sylvestris*), silver birch (*Betula pendula*), and black alder (*Alnus glutinosa*) with tree heights of up to 30 m (Bonzom et al., 2016). Overall, the forest stand was found to be dominated by Scots pine planted after the nuclear disaster, comprising approximately 50% of all trees. Based on visual interpretation of aerial imagery, we roughly estimated the distribution of pines, birches, and alders in area#1 (50/20/30), area#2 (40/30/30), and area#3 (60/20/20).
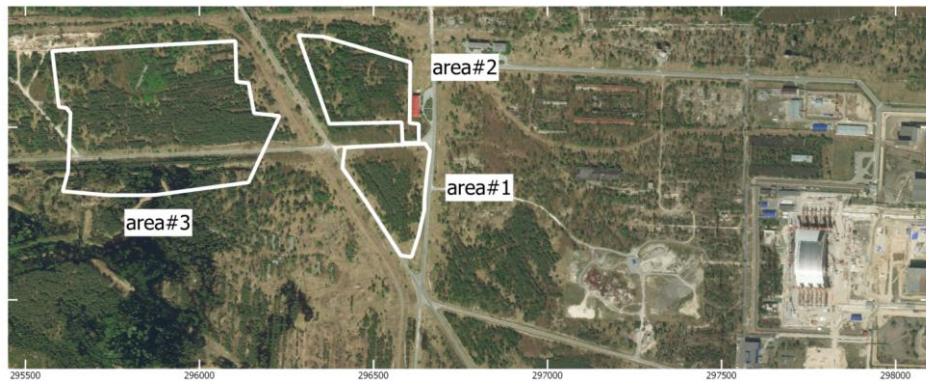
Figure 17: Overview of the study areas, located ~1.5 km west of Reactor Unit 4 at the ChNPP; Coordinate system: UTM zone 36N (EPSG 32636); Base map source: bing map ©Microsoft Corporation.
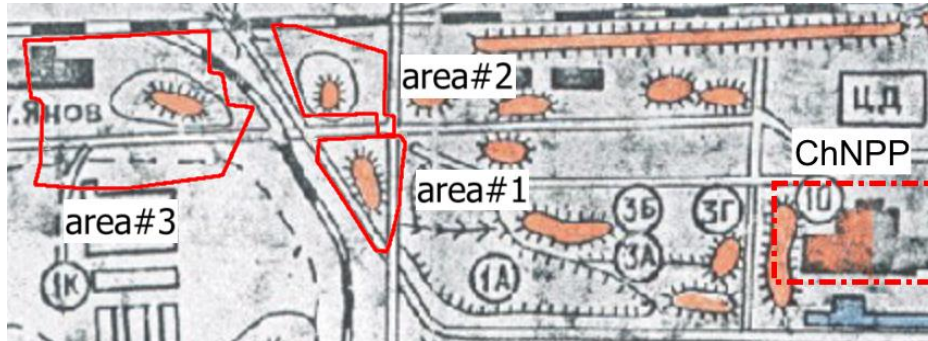
Figure 18: Scan of a historical map that was created from the memories of liquidators, roughly showing positions and extent of burials (Antropov et al., 2001); outlines of study areas in red.

3.3.2. Reference data

3.3.2.1. Visual interpretation

For obvious reasons, the ground truth data for tree species classification could not be acquired by field measurements. Thus, the reference dataset was prepared manually and balanced according to the three occurring tree species (pine, birch, alder). Based on visual interpretation of the generated orthomosaics and 3D geometries of single trees, we labelled 684 tree segments (228 per tree species). Thereafter, for each study area, these reference datasets were randomly split into training and test datasets using a sample ratio of 0.8.

Reference data for trenches and clamps were only available in area#3, using RWTSP vector data. These data were digitized from old maps and orthophotos based on visual inspection in the past using a simple office scanner and uncalibrated digital cameras. Therefore, the data are geometrically incomplete and show random position offsets in the order of a few metres. This can be clearly observed by overlaying the reference data and the normalized DTM height, which was generated from the ALS data (Figure 19). Nevertheless, we used this rather inappropriate dataset in the labelling process for one of our experiments. In all other experiments, we relied on manually labelled reference data created by visual interpretation of the normalized DTM. In area#1 and area#2, the structure of burials was quite obvious and labelling was straightforward. However, in area#3, the labelling was more challenging because of fairly non-systematic DTM structures. Here, the directions and lengths of the burials were partially unclear. Although the labelling was supported by RWTSP vector data, the reference dataset in this study area must be considered less reliable. For the labelling of DTM grid pixels, we outlined polygons representing 25 trenches and eight clamps from all three study areas, covering an area of 9,116 m²

and 2,900 m², respectively. Using a 50-cm DTM, more than 48,000 samples were extracted from these burials. Additionally, samples for non-contaminated DTM pixels were selected some metres off the burials to guarantee the pureness of samples. Then, for each study area, we utilized a sample ratio of 0.5 to randomly split the reference data into training and test datasets. Subsequently, the particular test datasets were completed by considering all remaining unlabelled pixels in the corresponding study areas as non-contaminated. As a consequence, the quality of the classifier was tested area-wide and, especially in the transition zones from burials to non-contaminated areas.

Figure 19: Available RWTSP reference data and manually labelled reference data in area#3, showing possible trenches/clamps and normalized DTM height. (base map: bing map ©Microsoft Corporation).

3.3.2.2. Ground reference data

In preceding ground surveys, neither a visual interpretation of the terrain nor anomalies in the equivalent dose rate (EDR) measurements on the ground surface ($EDR_{surface}$) indicated existing trenches. However, a strong evidence was provided based on the detection methods described in this chapter. Therefore, during an evaluation campaign conducted in autumn 2018, 20 on-site borings were drilled directly into the suspected trenches of area#1 and nine in area#2. For comparison, nine additional borings were drilled at a considerable distance of at

least 15 m away from the expected trenches of area#1 (Table 6). Borings C1, C2, C3, C4 and C5 were planned at spots showing unexceptional $EDR_{surface}$ values. In contrast, the locations for borings C98, C206, C208 and C216 were selected in places showing relatively high $EDR_{surface}$ values, thus assuming they were suspicious for buried contaminated waste. In area#3, no borings have been conducted so far. In general, all borings were performed by the Central Radioactive Waste Management Enterprise (CRWME), a state organisation responsible for maintenance and management of the RWTSPs under authority of the State Agency of Ukraine for the Exclusion Zone Management. The borings were carried out using a gasoline motor drill (up to a depth of 1.5 m) and a Geolog-2 hand drill (depth 1.5–2.5 m). The EDR was measured on the ground surface and inside the bore holes using an MKS-07 "Poshuk" dosimeter-radiometer. Table 4, Table 5 and Table 6 show $EDR_{surface}$ values, maximum EDR values (EDR_{max}) and the corresponding depth for each boring. Ratio $EDR_{max}/EDR_{surface}$ is a coefficient to demonstrate the degree of contamination inside the borings compared to the ground surface.

Table 4: Borings in suspected trenches of area#1.

Boring ID	$EDR_{surface}$ [µSv/h]	EDR_{max} [µSv/h]	Depth for EDR_{max} [m]	$EDR_{max}/EDR_{surface}$
111	1.5	13.4	1.8	8.9
112	1.0	7.2	0.8	7.2
113	0.4	13.5	1.4	33.8
114	1.9	7.6	0.6	4.0
115	1.6	14.6	1.0	9.1
121	1.0	20.3	1.4	20.3
122	1.6	7.7	1.2	4.8
123	1.9	14.3	1.2	7.5
124	0.9	26.5	1.4	29.4
125	0.8	24.5	1.8	30.6
131	1.3	2.4	0.8	1.9
132	0.7	7.0	1.0	10.3
133	0.9	8.6	1.2	9.2
134	2.9	4.1	0.8	1.4
135	1.9	8.3	0.8	4.4
151	3.8	16.3	1.2	4.3
152	1.0	18.6	1.2	18.6
153	2.9	17.5	0.6	6.0
154	1.3	7.3	1.6	5.6
155	1.5	10.3	1.2	6.9

Table 5: Borings in suspected trenches of area#2.

Boring ID	EDR$_{surface}$ [µSv/h]	EDR$_{max}$ [µSv/h]	Depth for EDR$_{max}$ [m]	EDR$_{max}$/EDR$_{surface}$
211	0.7	6.0	1.6	8.6
221	1.5	10.1	1.4	6.7
222	1.0	30.2	1.6	30.2
232	2.4	10.3	0.6	4.3
233	2.9	12.3	1.0	4.2
234	0.6	12.8	1.2	21.3
242	1.5	18.8	1.0	12.5
243	4.7	14.9	1.4	3.2
244	1.3	20.2	1.6	15.5

Table 6: Borings drilled at a considerable distance away from the trenches of area#1.

Boring ID	EDR$_{surface}$ [µSv/h]	EDR$_{max}$ [µSv/h]	Depth for EDR$_{max}$ [m]	EDR$_{max}$/EDR$_{surface}$
C1	1.3	1.3	0.0	1.0
C2	3.2	3.2	0.0	1.0
C3	3.9	3.9	0.0	1.0
C4	1.0	1.0	0.0	1.0
C5	0.6	0.6	0.0	1.0
C198	10.8	10.8	0.0	1.0
C206	7.5	7.5	0.0	1.0
C208	8.0	8.0	0.0	1.0
C216	4.8	6.7	0.2	1.4

3.3.3. UAV system

UAV-based sensor systems operate at a sufficient safe distance, which is of prime importance considering the high radiation dose rates within the study area. For our flight missions, we utilized an octocopter that was developed by a team from the Department of Nuclear Physics Technologies of the Institute of Environment Geochemistry of the National Academy of Sciences of Ukraine. With a maximum payload of 5 kg, both the lidar system and the MS cameras can be carried by the octocopter simultaneously (Figure 20). Nevertheless, we relied on a separate configuration, mainly to maximize lidar mission time. Regarding safety issues, the pilots had to ensure a minimum distance from the ChNPP of 1 km. Moreover, no power lines were flown over to avoid possible electromagnetic interaction with GNSS signals and radio link. All flights were carried out in fully automatic mode to avoid the direct trespassing of contaminated areas. Advantageously, local authorities allowed flying the UAV out of visual line of sight. To avoid collisions, the relative altitude in mapping mode was set to at least 50 m in this flat terrain, guaranteeing a safety distance of approximately 20 m to the highest tree crowns. Furthermore, the safety distance for operator

at take-off and landing was approximately 15 m. Finally, all data presented in this chapter were collected during sunny, partly cloudy weather conditions at a mostly constant wind speed (2–3 m/s).
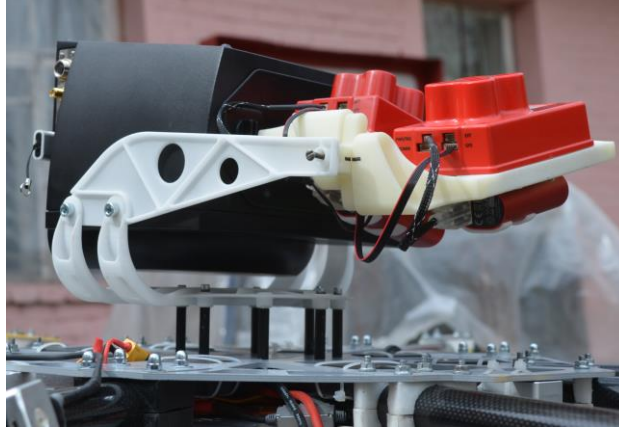
Figure 20: UAV-mounted laser scanner and MS cameras.

3.3.4. Lidar data

Lidar data were collected by a YellowScan Mapper I laser scanner (Table 7). Flights were conducted at a constant altitude of 50 m and speed of 6 m/s. Therefore, the relatively slow speed of the copter enabled a high lidar point density, which is important for a sufficient penetration rate to the ground to ensure the generation of detailed DTMs. The maximum lidar mission time with one battery set was approximately 20 minutes. GNSS waypoints were defined in the flight plans to guarantee parallel flight lines and a constant line-to-line distance of 40 m. During a one-week mission in the ChEZ in April of 2018, five leaf-off lidar flights were conducted in three selected areas. More flights were not possible due to upcoming wind speeds up to 20 m/s. A calibration flight over a building was conducted on the first day to check the boresight angles provided by the manufacturer. Simultaneously to all UAV flights, GNSS measurements were collected by a Trimble R4 base station (measurement rate = 1 s). These data were used in a differential GNSS postprocessing step using the commercial software Inertial Explorer® 8.70 (NovAtel Inc., 2017) to ensure flight trajectories with centimetre-level precision. Following the flight missions, the scan angle range of the lidar data was reduced from 100° to 70° to eliminate potentially inaccurate long-distance measurements on the edges of strips, resulting in a nominal point density of approximately 53 points/m² (effective side lap of 43%).

For the following lidar preprocessing steps, we utilized BayesStripAlign 1.3 (BayesMap Solutions LLC, 2018). The boresight angle differences

were calculated using the "calib" option and the strips of the aforementioned calibration flight. The residuals amounted to 0.02° (roll), -0.12° (pitch), and -0.12° (yaw) and were subsequently used to correct all recorded flight strips ("corr" option). Next, all strips were aligned to achieve geometrically consistent lidar point clouds. Note that the "align" option in BayesStripAlign 1.3 is based on the approach discussed in Jalobeanu and Gonçalves (2014). Overall, the mean discrepancy between adjacent strips was approximately 5 cm, which is in the range of the measurement accuracy of the lidar instrument. By fitting the ALS point clouds to the enclosing polygons of nearby buildings ("shift" option), absolute 3D georeferencing with an accuracy of a few centimetres was achieved.

Table 7: Technical specifications of the YellowScan Mapper I laser scanner (YellowScan, 2016).

Item	Value
Laser wavelength	905 nm
Pulse rate	18.5 kHz
Echoes per shot	3 (first, middle, last return)
Weight	2.1 kg (battery included)
Range resolution	4 cm
Precision[4]	10 cm
Absolute (XY) accuracy[5]	0.10 m + 1% of altitude
Absolute (Z) accuracy	0.10 m + 0.5% of altitude

3.3.5. MS imagery

We captured MS images using two MicaSense RedEdge cameras (Table 8) with a total payload of 500 g, incl. external batteries. These MS cameras capture light wavelengths in five spectral bands between 475 nm and 840 nm. The two cameras were mounted in a twisted configuration with an angle of 22.5°. This setup guaranteed a 50% side overlap of the camera foot-prints and a total field of view (FOV) of approximately 70°. Time synchronization of the cameras was realized by recording images with a frame rate of 2 s at each even Global Positioning System (GPS) second. During data collection, the altitude was 130 m, leading to a ground sample distance of 8.9 cm/pixel. A flight speed of 9 m/s accounted for a forward overlap of 79%, whereas the lateral overlap was set to 50%. To compensate for changing illumination conditions during and between the flights, we utilized both MicaSense's calibrated reflectance panel (CRP) and downwelling light sensor (DLS). Basically, these accessories are able to provide useful information for the subsequent reflectance calibration. In practice, we took close-up

---

[4] Also called reproducibility or repeatability (variation of measurements taken on the same target).
[5] Degree of conformity of a measured position to its actual (true) value.

images of the CRP before and after each flight. Moreover, the DLS was installed on top of the UAV, facing upwards with a clear view of the sky. Before each flight, we checked the calibration of the DLS magnetometer to ensure correct heading and orientation measurements.

Table 8: Technical specifications of the MicaSense RedEdge camera (MicaSense Inc, 2015).

Item	Value
Focal length	5.5 mm
FOV	47.2° horizontal, 69.7° horizontal (double-camera setup), 35.4° vertical
Imager size / resolution	4.8 mm x 3.6 mm / 1,280 x 960 pixels
Spectral bands	Blue (475 nm), Green (560 nm), Red (668 nm), Red Edge (717 nm), NIR (840 nm)

Agisoft PhotoScan Professional 1.4.1 (Agisoft LLC, 2018) was used for the following postprocessing steps. Initially, all images were aligned area-wise in a bundle adjustment (option "highest"). Here, the implemented RedEdge camera model was based on a standard frame camera model. In detail, the model specifies the transformation from point coordinates in the local camera coordinate system to the pixel coordinates in the image frame. Overall, the model comprises focal length, principal point offsets, two radial and two tangential distortion coefficients, and two skewness coefficients. In our approach, we estimated the master and slave camera models for the first flight and applied it to all other flights. The mean reprojection error for all flights was 1.3 pixels, corroborating our assumption of constant camera parameters between flights. Finally, dense photogrammetric 3D point clouds were generated with a point density of approximately 80 points/m² (option "high").

3.4. Methodology

3.4.1. Outline of the proposed method

Initially, various preprocessing steps must be performed for detecting radioactive waste sites. First, the reflectance of the 2D aerial images was calibrated, and Normalized Difference Vegetation Index (NDVI) and Red Edge Normalized Difference Vegetation Index (RENDVI) orthomosaics were created. Previous studies have proven a positive effect on the results for tree species classification by normalizing lidar intensity values (Höfle and Pfeifer, 2007; Korpela et al., 2010; Ørka et al., 2012). Thus, in our approach, the radiometric information of all lidar points was adjusted in a data-driven correction step. Based on the resulting lidar point cloud, ground points were filtered and a regular

DTM grid was computed. Next, the lidar point cloud was segmented into 3D clusters representing single trees using the normalized cut algorithm (Reitberger et al., 2009). The convex hulls of the segmented trees were then projected onto the orthomosaics to extract NDVI and RENDVI features for each segment. To supplement the geometric and radiometric features generated from the lidar point clouds of single tree segments, tree species classification was conducted based on a standard RF classifier including a feature selection step. Based on the results of tree classification, tree-based features were generated to describe possible vegetation anomalies. Next, the normalized DTM height (h_{norm}) was extracted from the DTM grid. In combination with LM and FPFH, an RF classifier was trained to distinguish burial sites (trenches and clamps) from non-contaminated areas. In an RFE step, the most significant features were identified. Finally, probability maps and enclosing 2D polygons for burial sites were generated. The entire processing pipeline is illustrated in Figure 21. In the following sections, all important steps are described in greater detail.
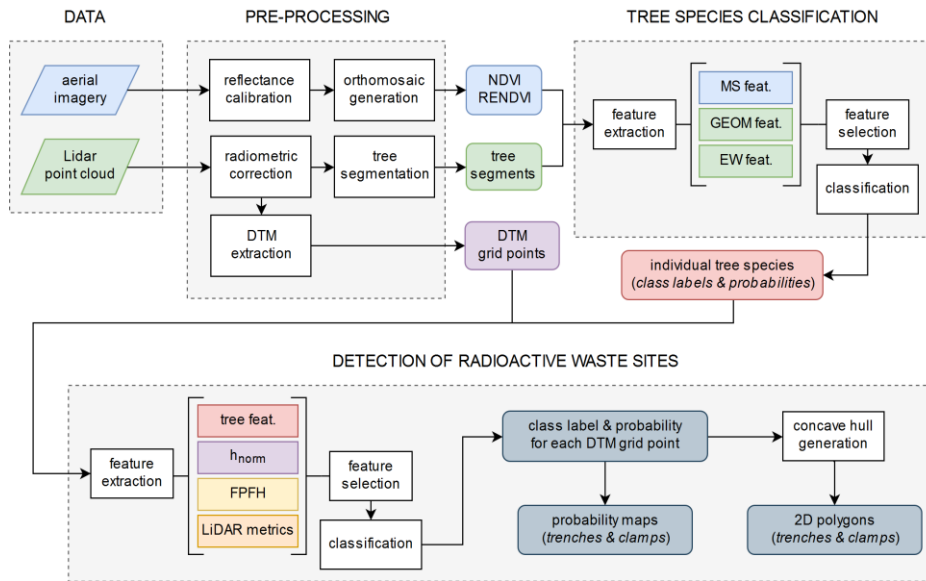
Figure 21: Overview of the proposed method for detecting radioactive waste sites.

3.4.2. Preprocessing of sensor data

3.4.2.1. Reflectance calibration of MS imagery

MS imagery was captured at changing light conditions over different days and times. Therefore, the reflectance must be calibrated to generate high-quality images. We performed this step in Agisoft PhotoScan Professional 1.4.1, integrating the spectral information from both the CRP and DLS. Initially, panel-specific absolute reflectance values for the spectral range of the camera were introduced. These calibration data were provided by the manufacturer in increments of 1 nm for each band. In the software, the transformation of raw pixel values into reflectance basically follows a two-step approach (MicaSense Inc, 2019). First, the raw pixel values of all images were converted to absolute spectral radiance using a camera-specific radiometric calibration model. Here, sensor-specific characteristics such as sensor black level and lens vignette effects were compensated. Second, a factor was estimated, allowing for a transformation from radiance to reflectance. At this point, the reflectance values from the CRP calibration data and the radiance values of the CRP images captured in the field were taken into account. Overall, the CRP-based calibration procedure assumes constant lighting conditions over the flight duration. In case CRP images were taken before and after the flight, the reflectance of MS imagery is calibrated based on linear interpolation of the CRP information. In situations with changing conditions in the middle of a flight, the DLS can help to improve reflectance calibration by applying additional corrections image-wise. Although the integration of DLS data was suitable in area#1 and area#2, the approach reached its limits in area#3. Rapidly changing cloud coverage throughout the entire flight impeded an accurate calibration of reflectance. In the event of clouds shadowing the captured image area but not the DLS, radiometric information of the images could not be corrected properly. As a consequence, the MS imagery in area#3 was not used for further investigation. Thus, in this area, tree species classification was only based on the geometric and radiometric features generated from lidar data.

3.4.2.2. Orthomosaic generation

Based on missing GPCs in overflown inaccessible areas, we registered the photogrammetric data to the georeferenced lidar data. Therefore, different steps in several software packages had to be executed for area#1 and area#2. Initially, the georeferenced lidar point cloud was transformed to the aforementioned photogrammetric dense point cloud via iterative closest point (ICP) algorithm in CloudCompare 2.8 (Cloud-Compare Development Team, 2019). Here, the root mean squared error was 0.237 m. Then, we utilized Quick Terrain Modeler (Applied Imagery, 2018) to calculate a digital elevation model (DEM) from the

transformed lidar point cloud. Next, the resulting 15-cm DEM was exported as a TIFF image and loaded into Agisoft PhotoScan Professional 1.4.1. Here, an orthomosaic with a 10-cm resolution was computed. In this orthorectification procedure, we relied on the precise lidar DEM instead of using a less accurate DEM calculated from a dense photogrammetric point cloud. Subsequently, vegetation index (VI) images were calculated. In addition to the widely used NDVI images, RENDVI images (Sims and Gamon, 2002) were created by utilizing the red edge (RE) and NIR channel. The RENDVI (Eq. 1) can be computed as

$$\text{RENDVI} = \frac{RE - NIR}{RE + NIR}. \tag{1}$$

Afterwards, the photogrammetric products must be georeferenced. The dense point clouds were transformed in CloudCompare 2.8, utilizing the particular inverse transformation matrix of the aforementioned ICP. Finally, based on the 2D components of these matrices, the orthomosaics were shifted in QGIS 2.18 (QGIS Development Team, 2017).

3.4.2.3. Radiometric correction of lidar data

In addition to the 3D coordinates of each laser point, the YellowScan Mapper I provides intensity values. The instrument is equipped with an Ibeo LUX 2010® laser unit for generating Gaussian-shaped pulses. According to the technical description, the recorded intensity values are equivalent to the widths Ew_i of the echo pulses measured at a fixed internal threshold g_0 (Figure 22). In a series of *n* Gaussian return pulses

$$g(t) = \Sigma_{i=1}^{n} A_i \cdot exp\left(-\frac{(t - t_i)^2}{\sigma_i^2}\right), \tag{2}$$

the parameter Ew_i theoretically depends on the shape of the return pulses, which is defined by the pulse width σ_i and amplitude A_i (Eq. 2). Both of these parameters are influenced by the scan angle. Assuming a flat terrain and the Lampertian law for scattering targets, the pulse width σ_i is slightly broadened towards the swath edge of a laser strip. Additionally, echo width widening leads to a reduction in the amplitude A_i with respect to the scan angle, assuming that the emitted laser power is constant (Ussyshkin et al., 2009). Furthermore, the amplitude A_i itself is dependent on the square of the distance between the sensor and target, meaning it also depends on the scan angle. In summary, we can theoretically expect a significant change in the shape of the return pulse and, the parameter Ew_i toward the swath edge. To the best of our knowledge, changes in the parameter Ew_i toward the swath edge have not been studied experimentally to date. However, many researchers have evaluated the dependency of intensity ($= A_i * \sigma_i$) and pulse width on the scan angle. Jutzi and Gross (2009) demonstrated

that lidar intensity decreases according to the cosine of the scan angle. In principle, the experiments conducted by Kaasalainen et al. (2011) using terrestrial laser scanners confirm this finding. Reitberger et al. (2007) investigated pulse width in forest areas and calculated scan-angle-dependent increases in pulse width of 15%/20° for laser points with well-defined and smooth reflecting areas (i.e., stem points). Additionally, the same study confirmed a decrease in intensity of 13%/20° corrected by the square of the laser distance for ground points.
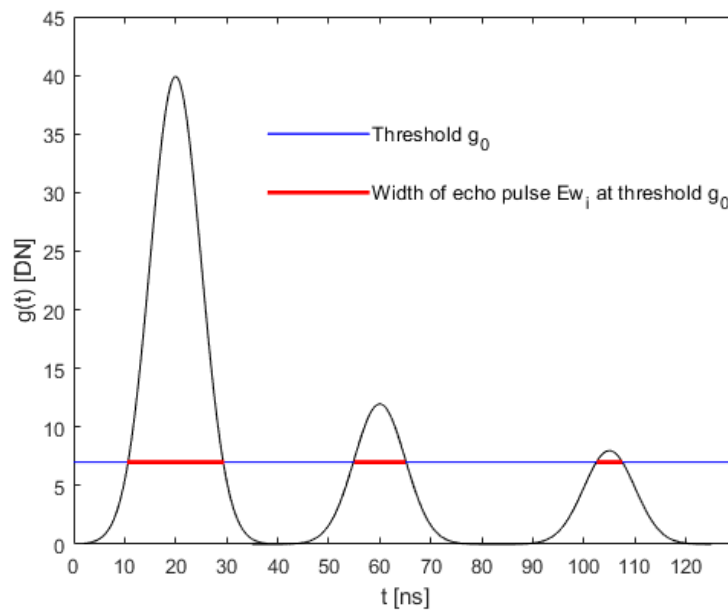
Figure 22: Definition of parameter Ew in YellowScan Mapper I (according to Ibeo Automotive Systems GmbH (2010)).

Figure 23 (*left*) illustrates the raw Ew_i values for area#1, where dark areas are a result of the scan angle effect of the raw parameter Ew_i, which was not compensated in the individual laser strips. For 3D vegetation mapping, raw Ew_i values must be corrected to avoid misclassification effects. Our data-driven method corrects the parameter Ew_i based on the incidence angle, which is approximately equal to the sum of the scan angle and roll angle. First, we filter single return points from the ground points in one reference strip and assume that the forest floor has consistent spectral reflectance properties. Second, the raw values of Ew_i for the reference strip are divided into bins of size 1° with respect to the corresponding incidence angle. Third, the mean and std value of Ew_i are calculated for each bin. Subsequently, a parabola (R² = 0.95) is fitted to the mean values of Ew_i (see Figure 24, *top*). The

vertex value of the parabola Ew_{ref} and $std(Ew_{ref})$ of the corresponding bin are used as constant parameters in the empirical correction function. The corrected values Ew_{corr}^i are computed as

$$Ew_{corr}^i = \left(Ew_{raw}^i - Ew_{fit}^i\right) * \frac{std\left(Ew_{ref}\right)}{std\left(Ew_{bin}^i\right)} + Ew_{ref}, \qquad [3]$$

where Ew_{raw}^i is the raw value of a single point, Ew_{fit}^i is the reference value of a single point on the parabola, and $std(Ew_{bin}^i)$ is the std of the corresponding bin (Eq. 3). Note that the correction function is only valid for one specific flight height because we did not perform drone flights at different altitudes. Figure 24 (*bottom*) presents the result of correction for the reference strip in area#1. The fitted red horizontal line indicates that the mean values of Ew_{corr}^i are independent of the incidence angle. Figure 23 (*right*) presents all points in area#1 coloured according to their corrected values. Apparently, the edge effects disappear and the parameters Ew_{corr}^i can be used as additional features for 3D vegetation mapping.
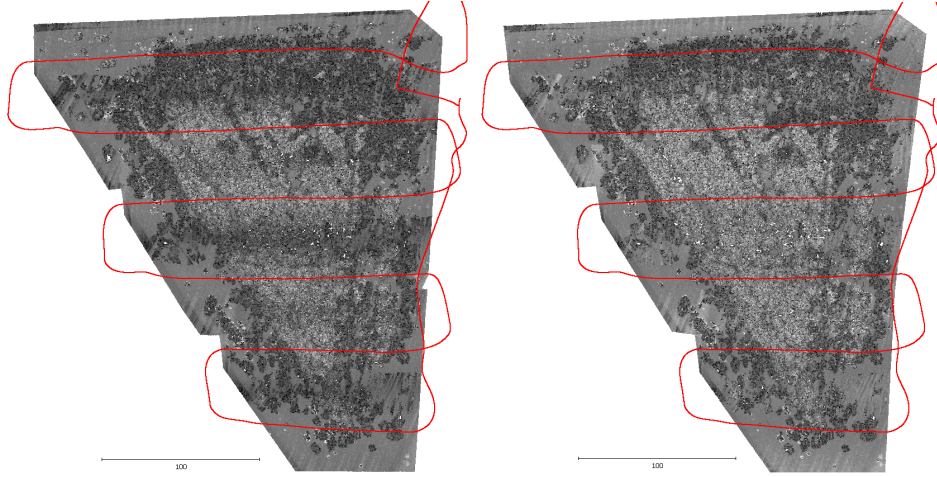
Figure 23: Ew_i values in area#1, before (left) and after (right) correction; flight lines overlay (red).
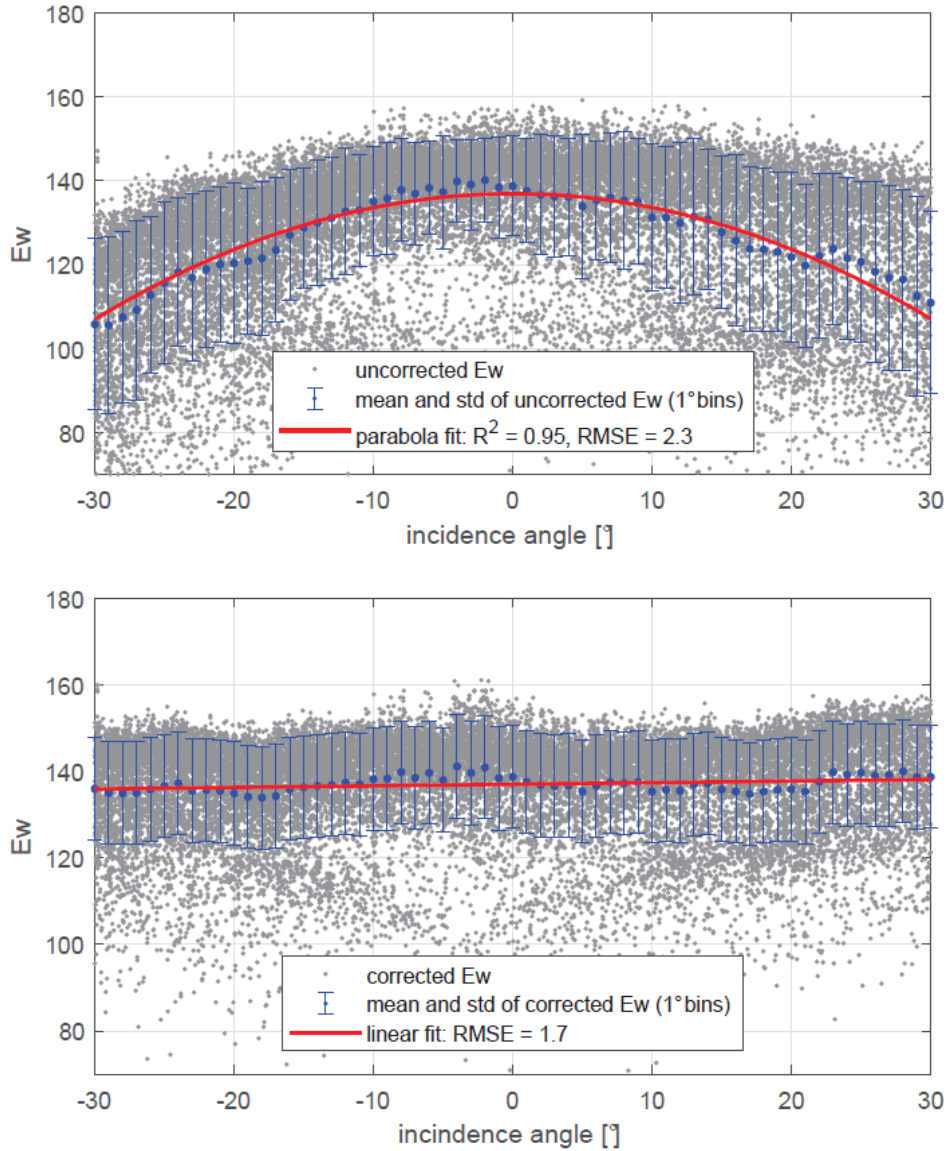
Figure 24: Ew_i values in the reference strip (area#1).

3.4.2.4. Tree segmentation

As a basis for the tree species classification performed at a later stage, single trees were segmented from the lidar point cloud utilizing the TreeFinder software package (PrimaVision Technologies GbR, 2017). The implementation is based on the normalized cut algorithm (Shi and Malik, 2000), which is a top-down method for segmenting objects over a discrete graph structure $G = (V, E)$. The vertices V represent individual

objects and the edges *E* correspond to neighbourhood topology. The input 3D lidar point cloud is split into disjointed segments to minimize the normalized cut criterion. A recursive bisection of the graph's vertices *V* into disjointed segments *A* and *B* maximizes the intra-segment similarity of objects and minimizes their inter-segment dissimilarity. The normalized cut criterion to be minimized is defined as

$$NCut(A,B) = \frac{Cut(A,B)}{Assoc(A,V)} + \frac{Cut(A,B)}{Assoc(B,V)}, \qquad [4]$$

Where $Cut(A,B) = \sum_{i \in A, j \in B} w_{ij}$ is defined as the sum of all weights between *A* and *B* segments and $Assoc(A,V) = \sum_{i \in A, j \in V} w_{ij}$ is the sum of the weights of all edges ending in *A* segments (Eq. 4).

Following the recommendations of Reitberger et al. (2009), we set the static stopping criterion of the normalized cut segmentation to 0.16. Visual inspection helped to verify that no oversegmentation or undersegmentation occurred. The result of tree segmentation is a set of polygons describing the hulls of single trees, as well as characteristic parameters, such as tree height, CBH, and crown volume.

3.4.2.5. DTM generation

In the next step, we used the ground routine based on Axelsson (2000) in the commercial software TerraSolid TerraScan™ (Soininen, 2016) to filter ground points from the ALS point cloud. The resulting ground points with a point density of approximately five points per m² were subsequently interpolated into a 50-cm DTM grid (Polewski et al., 2015b). The impact of the DTM smoothing coefficient α on detection results is discussed in Section 3.5.2.

3.4.3. Tree species classification

A classic ML approach using the RF classifier requires a set of hand-crafted features. These features are typically calculated from the available datasets, adapting as much as possible to the various tasks being addressed. In general, they are engineered manually and are mainly based on statistical terms. For tree species classification, we generated salient features for each tree segment. In total, the feature set consisted of 65 features (Table 9) derived from both lidar and MS data: 32 geometry features (*GEOM*), 14 *Ew*-based features (*EW*), and 19 MS features (*MS*), including statistics calculated from the NDVI and RENDVI. Thereafter, the feature space was reduced and an RF classifier was trained to classify tree segments as species of birch, pine, and alder. Note that the detailed classification strategy will be presented in Section 3.4.5.

Table 9: Feature set for tree species classification (65 features); GEOM#(1–20) are adopted from Næsset (2004), GEOM#(21–32) and EW#(1–14) from Reitberger (2010).

Features	Definition	Quantity
GEOM#(1–10)	Density distribution of points per height layer[6]	10
GEOM#(11–20)	Vertical distribution of the tree substance per height layer	10
GEOM#(21–30)	Mean distance of points to the segment centre	10
GEOM#(31–32)	Std of the distance of crown points to the segment centre, in x and y direction	2
EW#1	Mean *Ew* of points of a single tree	1
EW#(2–11)	Mean *Ew* of points of a single tree per height layer	10
EW#12	(Σ middle / Σ first) reflections	1
EW#13	(Σ single / Σ first) reflections	1
EW#14	(Σ first + Σ middle)/(Σ single + Σ last) reflections	1
MS#(1–8)	NDVI max/min/max-min/ mean/std/quartiles[7]	8
MS#(9–16)	RENDVI max/min/max-min/ mean/std/quartiles[8]	8
MS#(17–19)	Entries of covariance matrix of NDVI and RENDVI	3

3.4.4. Detection of radioactive waste sites

The main characteristics of radioactive burials are slight ground settlements (Figure 25) or elevations and vegetation anomalies in the surrounding areas. Therefore, we utilized the results of tree species classification and DTM extraction to prepare a feature set containing handcrafted tree-based features *TF*, normalized DTM height h_{norm}, fast point feature histograms *FPFH* and lidar metrics *LM* (Table 10).

Table 10: Feature subsets for detection of radioactive waste sites (186 features in total).

Feature subset	Definition	Quantity
TF	Tree-based features	19
h_{norm}	Normalized DTM height	1
FPFH	Fast point feature histograms	33
LM	Lidar metrics	133

---

[6] Increasing numbering from bottom (1) to top (10).
[7] 1st, 2nd, 3rd quartiles.
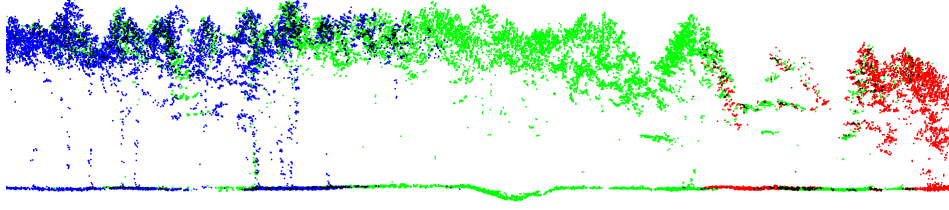
Figure 25: Cross section of lidar point cloud in area#1 showing possible trench. Points are coloured according to flight line.

Vegetation anomalies are supposed to be an indicator for the existence of radioactive burials. Therefore, we calculated 19 *TF* based upon the results of prior 3D vegetation mapping (Table 11). In general, the resolution of all generated *TF* was 5x5 m. For each grid point, nearby trees were found by applying a range search within a radius r_{tree} describing a local neighbourhood. Then, the feature set was derived based on the information of these filtered tree segments. Initially, the number of neighbouring trees resulted in *TF#1*. Next, the tree attributes tree height, CBH, and crown volume were used to generate features *TF#(2–7)* that represent the mean and std of these three attributes within the circular neighbourhood. *TF#(14–15)* were based on the area size of the tree polygons. For each filtered tree, the crown diameter D_{cr} was calculated from the crown size by assuming a circular crown shape, resulting in *TF#(16–17)*. Moreover, ratios describing the frequencies of individual tree species in each neighbourhood were defined as *TF#(8–10)*. Additionally, for each tree class, tree species classification also provided the class probabilities p_i for each tree segment. The mean p_i values of neighbouring trees were used to define the features *TF#(11–13)*. Finally, we estimated the total above-ground biomass P_a of individual trees. To this end, we utilized a transcontinental allometric model developed for the main Eurasian tree species using tree height H and crown diameter D_{cr} as regressors. We estimated P_a using the allometric function

$$lnP_a = a_0 + a_1 lnH + a_2 lnD_{cr} \qquad [5]$$

to generate features *TF#(18–19)*. Equation 5 and the coefficients a_0, a_1, and a_2 for the tree species of pine, birch, and alder were taken from Usoltsev et al. (2019).

Table 11: Definition of tree-based features TF.

TF#	Feature	Definition
1	tree_amount	Number of trees
2	tree_height_mean	Mean tree height [m]
3	tree_height_std	Std of tree height [m]
4	CBH_mean	Mean CBH [m]
5	CBH_std	Std of CBH [m]
6	crownVol_mean	Mean crown volume [m³]
7	crownVol_std	Std of crown volume [m³]
8	treeClass0_ratio	Pine tree ratio (Σ pines/ Σ trees)
9	treeClass1_ratio	Alder tree ratio (Σ alders/ Σ trees)
10	treeClass2_ratio	Birch tree ratio (Σ birches/ Σ trees)
11	pred0_mean	Mean pine probability
12	pred1_mean	Mean alder probability
13	pred2_mean	Mean birch probability
14	areaSize_mean	Mean 2D tree dimension [m²]
15	areaSize_std	Std of 2D tree dimension [m²]
16	D_{cr}_mean	Mean crown diameter [m]
17	D_{cr}_std	Std of crown diameter [m]
18	biomass_mean	Mean above-ground biomass [kg]
19	biomass_std	Std of above-ground biomass [kg]

Prior experiments have demonstrated that the normalized DTM height h_{norm} is a promising feature for describing terrain height anomalies (Briechle et al., 2018). This feature was calculated by reducing the corresponding absolute terrain height h_{DTM} for each DTM grid position by h_{trend} (Eq. 6):

$$h_{norm} = h_{DTM} - h_{trend} \qquad [6]$$

h_{trend} is the mean height in a surrounding quadratic area defined by an edge length e_{trend} and is calculated for every DTM grid point using a sliding window approach.

Because the shape of trenches and clamps is cylindrical, local shape descriptors are a promising feature set because they are able to distinguish between different geometric shapes (plane, cylinder, sphere, etc.). We calculated a set of FPFH (Rusu et al., 2009), an enhanced version of point feature histograms (Rusu et al., 2008), which were developed for real time robotics applications. Retaining most of the power of the point feature histograms, these features are invariant to 3D translations or rotations and robust to varying point densities and noisy datasets. As input data, we utilized filtered ground points from the lidar data, with a point density of approximately five points per m² in forested areas. Initially, normal vectors were computed for each ground point. Here, the neighbourhood size was defined by a radius r_N. Next, FPFH features were calculated for each point of a synthetic 2x2 m grid. In detail, geometric properties were estimated by modelling

the relationships between surface normals to characterize the local geometry in the area around each grid point. This calculation is based on the analysis of the eigenvectors and eigenvalues of the covariance matrix formed by points within a circular neighbourhood defined by a radius r_{FPFH}. Based on the angular differences between each pair of normals, a normalized multidimensional histogram was computed, leading to a total of 33 FPFH features.

The feature set was completed with 133 height- and density-dependent lidar features to characterize vegetation. Based on the raw 3D point cloud, these LM (Næsset, 2004) were computed for quadratic cells defined by a cell size e_{cell}. In detail, features were generated for 10 different height layers (0–1.5 m, 0.5–2 m, 0.5–5 m, 1.5–5 m, 5–12 m, >-0.5 m, <0 m, >0 m, >2 m, >12 m).

Note that the decisive control parameters for extracting the overall feature set were optimized in a sensitivity analysis (see Section 3.5.2), leading to a final list of best-performing parameters. Moreover, the complete feature set consisting of 186 individual features was bilinearly resampled to a grid of 50 cm resolution, which is congruent to the DTM. We used this feature set to train an RF classifier (see Section 3.4.5) for categorizing DTM grid points into the classes "trench", "clamp" or "non-contaminated". Next, polygonal objects were generated from DTM grid points, which showed a trench or clamp probability of more than 95% as concave hulls. Finally, isolated classification errors were eliminated using an area threshold of 30 m².

3.4.5. Classification strategy

In our study, we utilized RF classifiers for the classification of both tree species and radioactive waste sites. This popular supervised ML method has proven to achieve high-quality results in typical remote sensing tasks such as land cover classification (Pal, 2005; Rodriguez-Galiano et al., 2012) and tree species classification (Immitzer et al., 2012; Puissant et al., 2014). By now, this type of classifier has been implemented in various programming languages. In our experiments, we used the "randomForest" package (Liaw and Wiener, 2002) in R (R Core Team, 2018). In the preprocessing procedure, the high-dimensional feature space was gradually reduced to avoid overfitting. At the beginning, highly correlated redundant features were deleted from the feature set to avoid deterioration of classification accuracy. This process was based on the application of a threshold to feature-to-feature cross-correlation. First, a covariance matrix was calculated. Second, a cross-correlation threshold θ was used to eliminate high pair-wise redundancy. Specifically, one feature of any feature pair with a correlation coefficient exceeding θ was eliminated. Different θ values (0.85,

0.90, 0.95, 1.00) were tested to find an optimal threshold. To investigate the capability of the proposed handcrafted feature subsets, RF classifiers were trained on these reduced datasets. Next, an RF model was built using all feature subsets. Three iterations of five-fold cross-validation were used to derive average OA and κ values. The number of trees was set to 500, with a minimum size for the terminal nodes of one. Generally, trees were grown to the maximum depth. The parameter controlling size of the samples to be drawn per node was equal to the number of training samples. Furthermore, the number of features randomly sampled as candidates at each split was set to the square root of the number of total features. Finally, the optimal model was selected based on the metric "accuracy".

In general, irrelevant features can have a negative impact on model accuracy. Therefore, it is essential to perform feature selection prior to building the final model. On a smaller and more robust dataset, both overfitting and training time can be reduced. In the literature, a backward feature selection technique based on the RFE algorithm has been recommended quite often (Ma et al., 2017; Gregorutti et al., 2017). At the beginning, this algorithm trains an RF model on the overall feature set. Thereby, the relevance of individual features can be computed based on the mean decrease in accuracy. Then, the least important features can be recursively excluded from the feature space until a final best performing set is obtained. In the present study, we utilized the RFE algorithm from the caret package (Kuhn, 2008) in R and set the "rerank" parameter to "false". Thus, the variable importance was not recalculated each time features had to be removed. Based on the corresponding OA value, the resulting RFE model was compared to the RF model containing all available features. Finally, the RFE model was utilized to calculate class predictions $pred_i$ and class probabilities $prob_i$ for each sample.

In all study areas, the quality of the final RFE model was verified by comparing the predicted labels and real values of the particular test dataset. For quantitative evaluation, a confusion matrix was generated and quality metrics (OA, κ, precision, recall, F_1 score) were derived. Furthermore, we investigated the transferability of the trained classifiers. Therefore, the classifier optimized on study area#1 was applied to all labelled data of area#2.

3.5. Experiments

3.5.1. Tree species classification

RF classifiers (θ = 0.90) using all feature subsets achieved an OA of 90.2% (κ = 0.853) in area#1, 90.0% (κ = 0.850) in area#2, and 81.3% (κ = 0.720) in area#3 (Table 12). Using only individual feature subsets, the average OA values were 80.6% (geometry features), 80.1% (*Ew*-based features), and 86.7% (MS features). Here, κ ranged from 0.633 to 0.812. Depending on the study area, backward feature selection (RFE) improved the OA by 0.4% (area#3) to 2.6% (area#1, see Figure 26). Furthermore, the dimensions of feature spaces were significantly reduced (average reduction of 41.1%). On the test da-tasets, the trained classifiers achieved OA values of 88.9% (area#1), 88.1% (area#2), and 78.4% (area#3). Here, the mean F_1 scores were 0.90 (pine), 0.81 (birch), and 0.85 (alder). The seven best features for each study area were ranked according to the mean decrease in accu-racy. Features from all three feature subsets were ranked among the seven most important features. The most relevant feature was the mean *Ew* value of all laser points for a single tree (*EW#1*). Other im-portant *Ew*-based features were the mean *Ew* value of points in the two top height layers of trees (*EW#10*, *EW#11*), as well as the features describing penetration (*EW#13*, *EW#14*). Geometry features ranked in the top seven were related to crown density (*GEOM#9*), crown shape (*GEOM#31*), or the occurrence of points in the lowest height layer of trees (*GEOM#11*). The most important MS features were the second and third quartile of the NDVI values (*MS#7*, *MS#8*), as well as the second quartile of the RENDVI values (*MS#15*) inside a tree polygon.
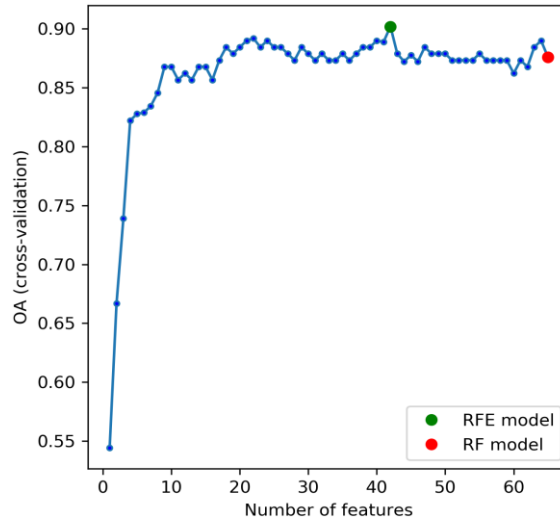
Figure 26: RFE result for tree species classification in area#1.

Table 12: Results of tree species classification.

Study area	1	2	3
Training samples	180	162	204
GEOM	83.4 (0.750)[8]	81.7 (0.725)	76.7 (0.650)
EW	85.9 (0.787)	75.6 (0.633)	78.7 (0.681)
MS	87.2 (0.812)	86.1 (0.792)	---[9]
All subsets (RF)	87.6 (0.814)	88.5 (0.828)	87.6 (0.814)
All subsets (RFE)	**90.2** (0.853)	**90.0** (0.850)	**81.3** (0.720)
Gain in OA	2.6	1.5	0.4
Feat. reduction	35.8%	46.3%	41.3%
Top 7 feat.	*EW#1*	*MS#7*	*EW#1*
	MS#8	*EW#1*	*EW#11*
	EW#10	*EW#11*	*EW#10*
	EW#11	*EW#10*	*GEOM#11*
	GEOM#9	*GEOM#9*	*EW#13*
	GEOM#31	*MS#15*	*EW#14*
	MS#15	*GEOM#31*	*GEOM#9*
Test samples	45	42	51
All subsets (RFE)	**88.9** (0.832)	**88.1** (0.821)	**78.4** (0.677)
Pine	0.88/0.94/0.91[10]	0.92/0.86/0.89	0.89/0.89/0.89
Birch	0.88/0.82/0.85	0.80/0.86/0.83	0.72/0.76/0.74
Alder	0.92/0.92/0.92	0.93/0.93/0.93	0.73/0.69/0.71

3.5.2. Sensitivity analysis

The four feature subsets used for the detection of trenches and clamps are dependent on six decisive control parameters (Table 13) that must be optimized through sensitivity analysis. During this optimization process, we used the OA value of the RF classifier as a quality measure. Different radii r_{tree} for neighbourhood definition for the selection of nearby trees were tested. The value of 5 m yielded the best performance. Therefore, the circular area size for the creation of *TF* was 78.5 m². The key parameter for the generation of h_{DTM} is the DTM smoothing coefficient *a* (Polewski et al., 2015b). Values ranging from 1 to 10 were evaluated and the optimal *a* value was determined to be three. For the computation of h_{trend}, the decisive parameter is the size of the surrounding quadratic area. Areas with a varying side length e_{trend} were tested and the best OA was achieved with e_{trend} = 5.5 m. The radii r_N and r_{FPFH} defining the neighbourhood for the computation of surface normals and FPFH generation were also optimized. Fourteen combinations of parameters were tested. The optimal values were r_N = 3.0 m and r_{FPFH} = 7.0 m. Finally, the cell size e_{cell} for the calculation of LM was set to 5x5 m.

---

[8] OA in %, κ in brackets.
[9] For area#3, no accurate MS data were available.
[10] Precision / recall / F_1 score.

Table 13: Control parameters.

Parameter	Definition	Tested values	Optimal value
r_{tree}	Radius to identify close-up trees [m]	3, 5, 10, 20, 30	5
a	DTM smoothing coefficient	1 to 10 (step size of 1)	3
e_{trend}	Edge length of quadratic area to calculate DTM trend [m]	0.5 to 45.5 m (step size of 5 m)	5.5
r_N	Radius used to compute surface normals [m]	1.5 to 5.0 m	3.0
r_{FPFH}	Radius used to generate FPFH [m]	2.0 to 10.0 m	7.0
e_{cell}	Cell size for calculation of LM [m]	2x2, 5x5, 10x10	5x5

3.5.3. Detection of radioactive waste sites

3.5.3.1. Classification results

For the detection of trenches and clamps, the mean OA of the RF classifiers (θ = 0.90) was 98.9% for the training data and 98.2% for the test data (Table 14). In general, the recall and precision for the class "non-contaminated" were 0.97 or higher in all study areas, leading to excellent F_1 scores. Furthermore, the recall values for "trench" were at least equal to 0.98, whereas the precision values ranged between 0.37 and 0.62. Apparently, numerous DTM pixels were predicted as trench pixels although labelled as "non-contaminated". Moreover, in area#3, the RF classifier reached an F_1 score of 0.80 for the class "clamp" (precision = 0.66, recall = 1.00).

A reduction of the feature sets to individual feature subsets led to mean OA values of 88.7% (*TF*), 71.6% (h_{norm}), 93.9% (*FPFH*), and 95.3% (*LM*). By including the RFE step, the feature space was considerably reduced by an average of 40.2% and the gain in OA was 0.23%. Feature relevance assessment clearly demonstrated that members of all four feature subsets are represented in the final list of the 10 most important features. The normalized DTM height h_{norm} is ranked here, as well as the features *FPFH#(16–18,27,28)*. The most important LM are features that describe the terrain surface and vegetation below 1.5 m, namely the mean value of points with a negative height *LM_mean(h<0m)* or a height below 1.5 m *LM_mean(h<1.5m)* and the minimum height of all points *LM_min(h)*. *TF* utilizing tree height (*TF#(2,3)*), mean CBH (*TF#4*), and mean crown diameter (*TF#16*) are also contained in this ranking. Furthermore, species-related features, namely alder tree ratio (*TF#9*) and mean alder probability (*TF#12*),

seem to be of importance for the waste site classification. The *TF* comprise approximately 27% of the top 10 features among all areas (eight out of 30).

Table 14: Classification results for detection of radioactive waste sites.

Study area	1	2	3
Training samples	6,222	14,310	22,872
TF	88.5 (0.841)[11]	90.7 (0.893)	86.9 (0.851)
h_{norm}	79.7 (0.590)	71.1 (0.422)	64.0 (0.460)
FPFH	96.1 (0.914)	95.4 (0.917)	90.1 (0.852)
LM	95.3 (0.910)	95.9 (0.926)	94.6 (0.923)
All subsets (RF)	98.9 (0.972)	99.4 (0.987)	97.6 (0.961)
All subsets (RFE)	**99.1** (0.982)	**99.6** (0.992)	**97.9** (0.968)
Gain in OA	0.2	0.2	0.3
Feat. reduction	43.4%	64.1%	13.2%
Top 10 feat.	*FPFH#16*	*LM_mean(h<0m)*	*FPFH#16*
	FPFH #18	*FPFH#16*	*LM_mean(h<0m)*
	LM_mean(h<0m)	*TF#3*	*LM_min(h)*
	TF#4	*FPFH#18*	*FPFH#18*
	LM_min(h)	*TF#2*	h_{norm}
	h_{norm}	*LM_mean(h<1.5m)*	*LM_mean(h<1.5m)*
	TF#12	h_{norm}	*TF#4*
	FPFH#27	*TF#12*	*FPFH#28*
	FPFH#17	*FPFH#28*	*FPFH#27*
	TF#16	*FPFH#27*	*TF#9*
Test samples	208,704	251,561	1,015,037
All subsets (RFE)	**99.1** (0.758)	**98.2** (0.744)	**97.3** (0.609)
Non-cont.	1.00/0.99/1.00[12]	1.00/0.98/0.99	1.00/0.97/0.99
Trench	0.62/0.99/0.76	0.61/0.98/0.75	0.37/0.98/0.54
Clamp	---	---	0.66/1.00/0.80

Besides the estimated class labels, we also calculated the class probabilities for each DTM pixel. From these values, probability maps (Figure 27, *right*) were created showing DTM grid points classified with more than 95% as "trench". The results are highly correlated to the DTM height (Figure 27, *left*). In the last step, polygonal objects were generated from classified DTM pixels as concave hulls. To clean up the results, we set an area threshold of 30 m² to remove small objects that were obviously generated by clusters of wrongly classified DTM pixels. Finally, an overview of all the detected radioactive burials was created (Figure 31).

---

[11] OA in %, κ in brackets.
[12] Precision / recall / F_1 score.
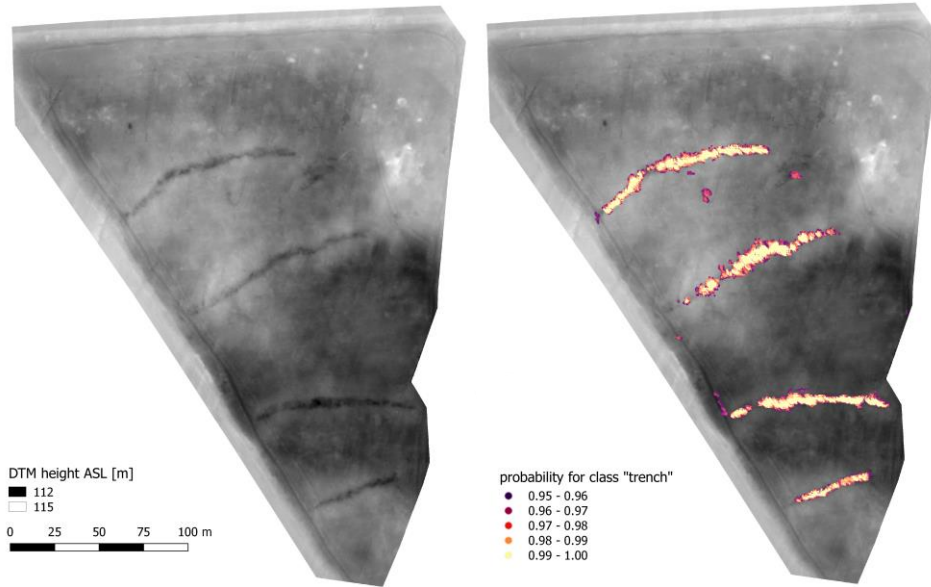
Figure 27: Classification result for area#1.

We now focus more on the spatial distribution of the classification errors. Figure 28 (*left*) shows the normalized DTM height and the manually produced polygons that were used to label the DTM pixels in area#1. Figure 28 (*right*) illustrates the classification errors on the test dataset (OA = 99.1%, κ = 0.758) and highlights the errors. On the one hand, false positives (FPs) appear in the transition zone from trench to non-contaminated areas. On the other hand, some accumulated FPs occur, and these are not correlated to the normalized DTM height in most cases. The results in area#2 (OA = 98.2%, κ = 0.744) are mostly comparable to those of area#1. In area#3 (OA = 97.2%, κ = 0.609), a large number of non-contaminated pixels were wrongly classified as trench pixels (22,491). A closer look at these grouped FPs demonstrates that they strongly resemble the typical shape of trenches (Figure 29, *left*). Moreover, they are correlated to the normalized DTM height. Furthermore, we could find a fairly high number of FPs (4,613) for the class "clamp". However, most of these errors are only located in the transition zone from clamp to non-contaminated areas (Figure 29, *right*).
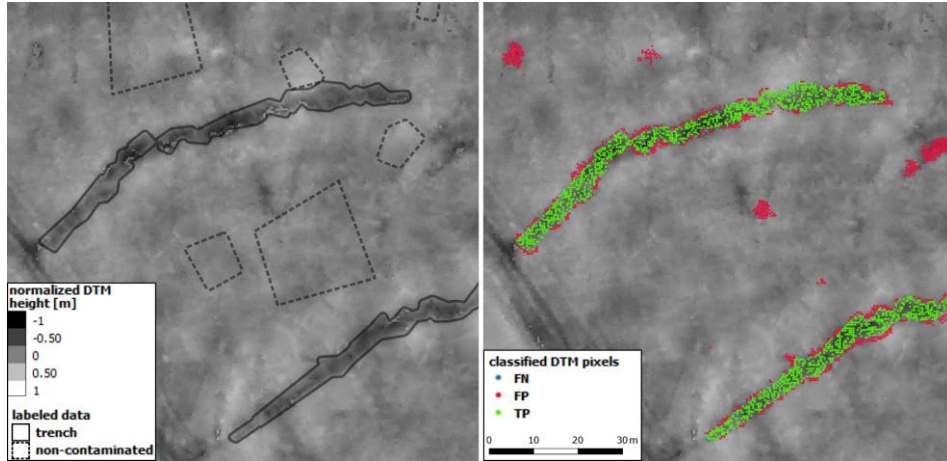
Figure 28: Classification errors for detection of trenches (section of area#1).
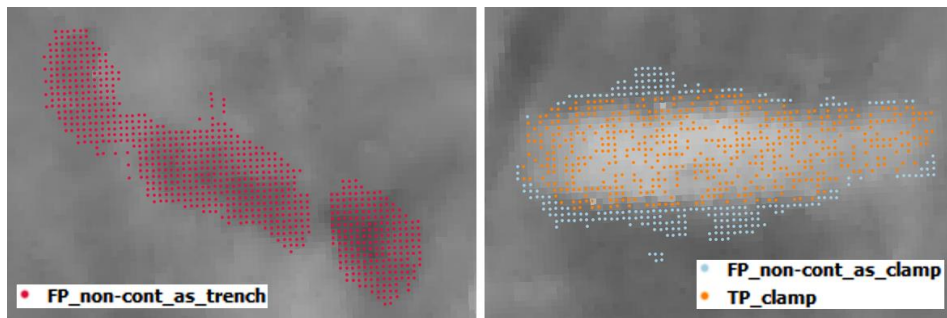
Figure 29: Classification errors (section of area#3).

3.5.3.2. Classifier generalization

In addition to classifier evaluation on manually labelled test data in the same study area, we conducted an experiment to investigate the generalization capabilities of the trained classifiers. The classifier trained for area#1 was applied to the entire dataset of area#2. If we analyse the relevant confusion matrix (Table 15), we can notice a considerable high number of false negatives (FNs) (6,440) and FPs (9,719) causing a relatively low F_1 score for the class "trench" (0.49). Note that the extremely high number of true negatives (TNs) (227,752) still led to an excellent OA (93.6%). For a section of area#2, Figure 30 (*right*) shows the spatial distribution of the classification errors. Apparently, in some areas, numerous FPs were generated in the vicinity of DTM pits (Figure 30, *left*).
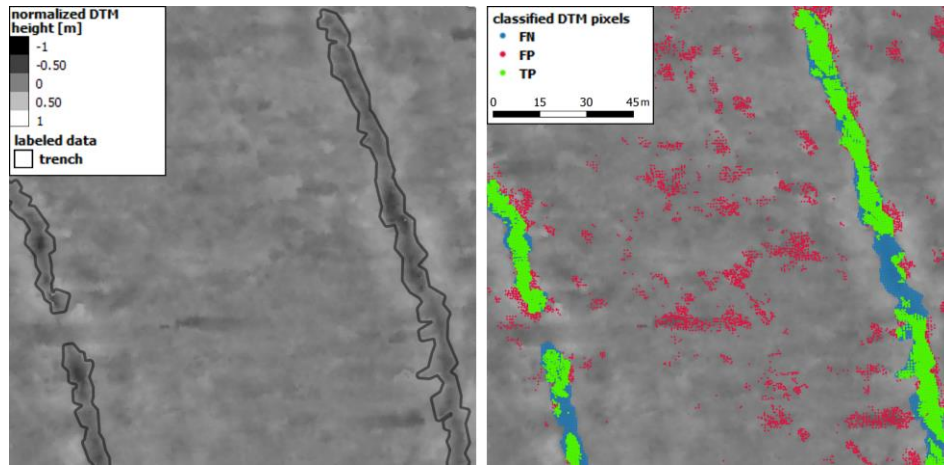
Figure 30: Generalization capabilities of the classifier (section of area#2).

Table 15: Results for applying the classifier trained for area#1 to the dataset for area#2 (OA = 93.6%).

		prediction		evaluation metrics		
		non-cont.	trench	prec.	recall	F_1
reference	non-cont.	227,752	9,719	0.97	0.96	0.97
	trench	6,440	7,650	0.44	0.54	0.49

3.5.3.3. Classifier evaluation using ground truths

First, the classifier trained on area#3 was applied to geometrically incorrect and incomplete RWTSP reference data from area#3. Here, the classification accuracy was relatively low (Table 16). Specifically, 55.7% (39,328 out of 70,598) of the DTM grid points located in the RWTSP trench areas were classified as "non-contaminated". Note that the quality of these reference data is poor because these data are based on nonprofessional generation techniques (see Section 3.3.2.1). Second, 38 borings (Figure 31) verified the existence (25) or absence (9) of buried radioactive biomass and demolition waste in 34 cases (OA = 89.5%). In four cases (borings C132, C151, C155 and C221), radioactive waste was found in areas that had been classified as "non-contaminated". However, these FNs are located only 3–9 metres off the detected trenches. All deposits occurred at an average depth of 1–2 m, exhibiting EDR values that were up to 30 times greater than those on the terrain surface (Table 4 and Table 5).
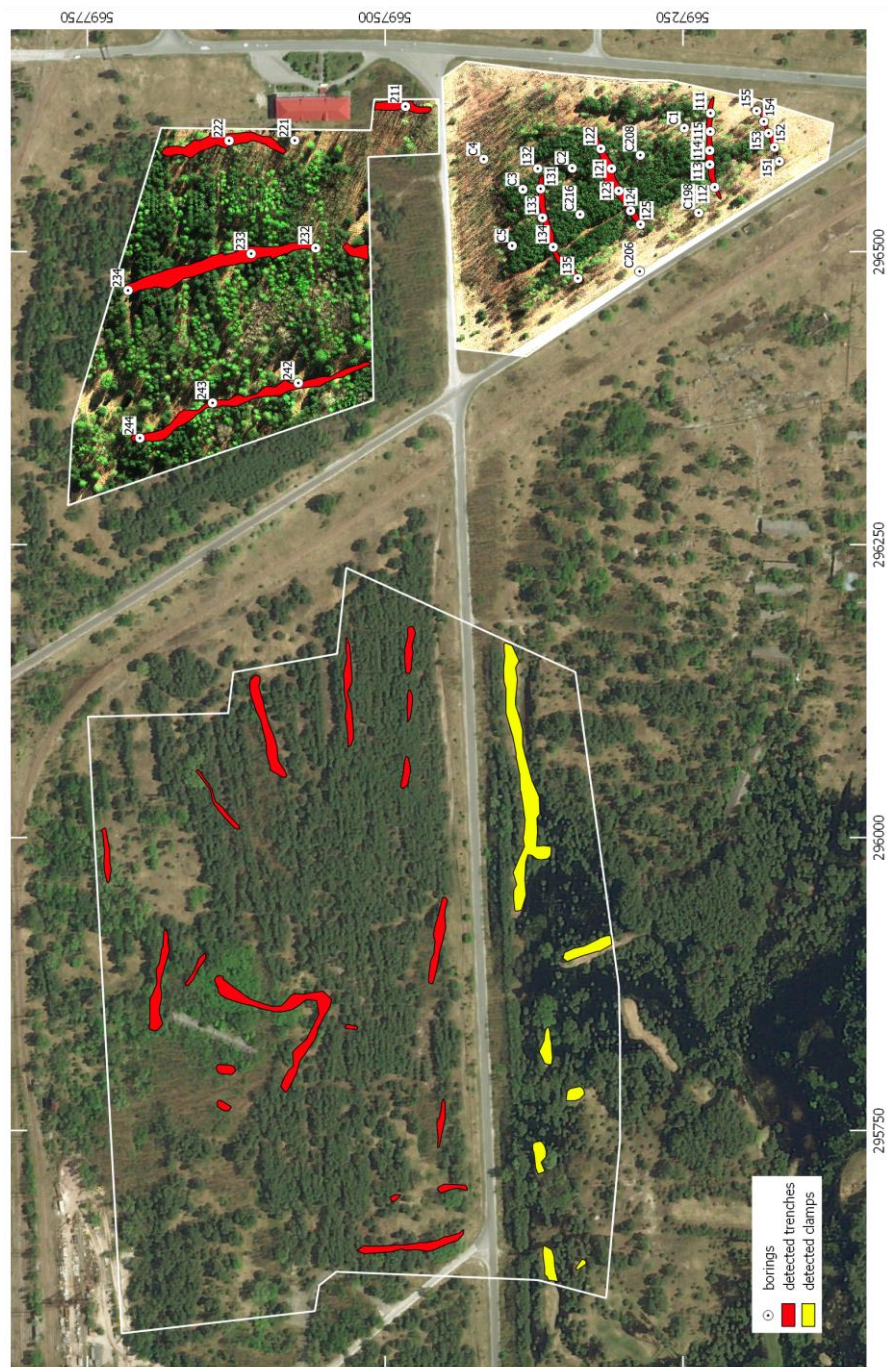
Figure 31: Detected burial sites (trenches and clamps) and conducted borings in the study areas; Coordinate system: UTM zone 36N (EPSG 32636); Base map source: bing map ©Microsoft Corporation.

Table 16: Results for applying the classifier trained for area#3 to the RWTSP reference data (OA = 65.1%).

| | | prediction | | | evaluation metrics | | |
		non-cont.	trench	clamp	prec.	recall	F_1
reference	non-cont.	39,749	111	16	0.48	1.00	0.65
	trench	39,328	28,535	2,735	0.95	0.40	0.57
	clamp	3,453	1,391	19,439	0.88	0.80	0.84

3.6. Discussion

3.6.1. Tree species classification

3.6.1.1. Classification results

In general, the results demonstrate that the three main tree species in the ChEZ (birch, alder, and pine) were successfully classified using both lidar data and MS imagery. Moreover, the low decrease in classification results on the test data compared with training data (1.3–2.9 percentage points) indicates a fairly good generalization quality. The best results were achieved in the case of the available MS data in area#1 and area#2 (mean OA = 88.5%, mean κ = 0.827). These results are highly comparable to Kamińska et al. (2018), who classified six classes of trees with an OA of 88.6% (κ = 0.851) fusing CIR imagery and leaf-off ALS data with normalized intensity values. Without MS data (area#3), the OA decreased to 78.4% (κ = 0.677). In our experiments, MS features generated from five spectral channels clearly increased the classification result by approximately 10%. This finding corresponds to Holmgren et al. (2008) and Ørka et al. (2012), who also demonstrated a significant improvement in classification results by including information from MS images. Regarding the single tree species, the overall F_1 score for pine (0.90) was generally better than for the deciduous species birch (0.81) and alder (0.85). Even without MS features (area#3), pines were classified with an F_1 score of 0.89. Obviously, MS features primarily improved the classification of birch and alder. This is because at the time of data collection, birches had already sprouted. Therefore, their characteristic spectral appearance supported the classification significantly.

3.6.1.2. Feature importance

Furthermore, feature selection led to a significant reduction of the feature space and improved the OA by 1.5% on average. According to the

feature relevance assessment, the mean *Ew* value of all laser points for a single tree was the most relevant feature. This fact substantiates the quality of our newly introduced data-driven correction method for radiometric lidar data of the YellowScan Mapper I laser scanner. Moreover, geometry- and *Ew*-based features related to the tree crown and penetration of the laser beam to the ground were also among the top seven features. This is in full accordance with Shi et al. (2018b) and Amiri et al. (2019), showing that these types of features mainly control the quality of tree species classification. In our experiments, certain quartiles of the NDVI and RENDVI values significantly enhanced the classifier. Regarding the feature subsets, the *MS* features provided better results than the lidar-related features (*GEOM*, *EW*). A classifier using only the *MS* features was only slightly worse than the classifier using the complete feature set.

3.6.2. Detection of radioactive waste sites

3.6.2.1. Classification results

Our results show that the area-wide detection of unknown radioactive burial sites in the forested areas of the ChEZ can be performed successfully using both UAV-based lidar data and MS imagery. A classification accuracy of over 97% for DTM grid points potentially located in areas with buried radioactive materials is relatively high for all study areas. By including a feature selection technique, we removed redundant and irrelevant features from the dataset and effectively avoided overfitting for a standard RF classifier. The OA values of the proposed classifiers on the test datasets are comparable to the results of a five-fold cross validation that we performed during classifier training. However, FPs in the transition zones from burials to non-contaminated areas caused a decrease in the precision values for both trenches and clamps. Manual generation of reference data based on visual DTM interpretation may certainly affect these errors.

In area#3, we realized that approximately 30% of the FPs are grouped in longish clusters and, thus, strongly resemble the typical shape of trenches. Moreover, they are correlated to the normalized DTM height. Based on the results for the class "trench" in area#1 ($F_1 = 0.76$) and area#2 ($F_1 = 0.75$), we assume that the classification accuracy for trench pixels is actually higher. Because of missing ground reference data in this study area, we can only speculate that these grouped errors are located in trenches that were not considered during manual labelling. Theoretically, if these 30% of rather obvious FP samples were taken into account in area#3, the F_1 score for "trench" would improve from 0.54 to approximately 0.72.

3.6.2.2. Feature importance

A feature relevance assessment supported the classification results and demonstrated that handcrafted features from all the subsets appear among the most important features. Clearly, local shape descriptors are very significant features for modelling the cylindrical shapes of both trenches and clamps. Two FPFH features (*FPFH#16* and *FPFH#18*) were ranked in the top four features for all study areas. Unsurprisingly, the normalized DTM height h_{norm} is a promising feature in all three study areas. Regarding the LM, we observed that the most important variables are features describing the terrain surface and low structures. In addition to the features describing the terrain geometry, the newly developed *TF* comprise approximately 27% of the top 10 features for all study areas, meaning they also had a significant impact on the classification results. Aside from the features describing the tree geometry (*TF#(2–4,16)*), species-related features such as alder tree ratio (*TF#9*) and mean alder probability (*TF#12*) are capable of further improving the proposed classifier. For example, it appears that alder trees preferably grow in the vicinity of the trenches of area#1 and area#2. In summary, tree species classification is beneficial for detecting radioactive waste sites in the ChEZ.

3.6.2.3. Transferability between study areas

To a certain degree, the capabilities of the trained classifier for trench detection in area#1 were transferable to adjacent area#2 (OA = 93.6%). Compared with the results of an RF classifier being trained directly in study area#2, the F_1 score for non-contaminated DTM pixels remained almost unchanged ($F_1 = 0.97$). Nevertheless, a considerable high number of misclassified DTM pixels caused clearly worse results for the class "trench" ($F_1 = 0.49$). When analysing the spatial distribution of these errors, we observed that numerous FPs were generated in the vicinity of DTM pits. Furthermore, comparing the feature importance of the classifiers trained in area#1 and area#2, it seems that some of the *TF* are disadvantageous for classifier generalization. Namely, the mean CBH is ranked 4th in area#1, whereas the mean and std of the tree height are in the top five features in area#2. Due to different characteristics regarding the tree geometry (crown shape, tree height), these *TF* vary between the study areas. However, these differences could not be detected in a visual interpretation of the dataset. The importance of features related to the terrain geometry is less fluctuating. In other words, these features should be focused on if a generalization of the classifier is to be considered. Unsurprisingly, when using geometrically incorrect and incomplete RWTSP reference data, the OA is relatively low (OA = 65.1%).

3.6.2.4. Verification based on ground truths

A comparison of detection results and 38 borings clearly shows that 25 true positives (TPs) and 9 TNs were verified in area#1 and area#2 (OA = 89.5%). Only four FNs occurred less than 9 m off the detected trenches, meaning the classifier falsely predicted areas as "non-contaminated". In summary, EDR measurements inside the bore holes confirmed the existence of suspected trenches that were unknown and not identified prior to the application of our methodology. Furthermore, EDR measurements revealed that the radioactive waste in the trenches was covered with a layer of approximately 0.5 m of "clean" sand. This layer shielded the active inventory. Therefore, these radioactive waste burials were not detected by standard EDR measurements.

3.6.2.5. Comparison to previous studies

To the best of our knowledge, the detection of radioactive burials using remotely sensed optical data has never been performed before. Thus, there is a lack of comparable studies. Nevertheless, our task is partially similar to certain research in the archaeology community. For example, Guyot et al. (2018) detected burial mounds based on the normalized height calculated from a 25-cm DTM. Furthermore, Lasaponara et al. (2010) demonstrated that micro-elevation changes in lidar-based DTMs support the identification of archaeological sites. Interestingly, our method significantly outperforms UAV-based gamma spectrometry surveys (Zabulonov et al., 2015), which have a detection rate of approximately 50% in vegetated areas. Moreover, compared with a historical map created from the memories of liquidation workers (Figure 18), the quality of our precise and reliable mapping approach is enormous.

3.6.3. Limitations of method

Despite fairly good results, our methodology has some limitations. Needless to say, the detection of radioactive waste sites greatly depends on the resolution of the lidar-based DTM. As stated by other studies, ground point density is the most important when detecting slight ground settlements or elevations. For example, Bollandsås et al. (2012) could show that the lidar-based DTM resolution had a significant effect on the detection success of cultural remains. Especially in densely vegetated areas, enhanced UAV lidar sensors with pulse repetition frequencies beyond 1 MHz and the realization of flight missions with cross-strips could tackle this limiting factor. Another weakness of our methodology for tree species classification is the manual preparation of training and test samples. However, because of understandable reasons, the collection of field reference data was not possible in the study area.

3.7. Conclusion

Our experiments demonstrated that area-wide detection of unknown radioactive waste sites in the ChEZ can be conducted successfully using the proposed methodology. Overall, we achieved excellent classification results fusing high-resolution UAV-based lidar data and MS imagery. A key factor of our approach is the generation of a feature set that can describe both the ground surface and nearby vegetation. Moreover, we applied an effective feature selection strategy to avoid overfitting of the RF models, resulting in remarkable generalization properties. Furthermore, we presented an efficient correction method for the radiometric data collected by the YellowScan Mapper I laser scanner. Based on geometrically and radiometrically consistent data, precise 3D vegetation mapping at the tree level enabled the generation of meaningful *TF* that supported the proposed classifier. Our results were verified based on 38 borings that confirmed the existence of previously unknown buried nuclear materials in the classified areas. Compared to the ground surface, EDR measurements revealed values up to 30 times greater in the soil layers containing buried biomass and demolition waste. Moreover, it is noteworthy that some of the parameters from 3D vegetation mapping (i.e. tree height, tree species, stem diameters, tree positions) can be used advantageously in radiologic simulations. Based on the results in Briechle et al. (2018), Molitor et al. (2018) calculated the potential exposure to external irradiation from a single tree species for a person working in the ChEZ forest. Therefore, more realistic radiologic modelling based on the method proposed in this chapter for 3D vegetation mapping should be possible. Finally, existing hazard maps were upgraded using the proposed method to minimize workers' radiation dose uptake and optimize accident liquidation.

Acknowledgements

4 Classification of tree species and standing dead trees by fusing UAV-based lidar data and multispectral imagery in the 3D deep neural network PointNet++[13]

Abstract

Knowledge of tree species mapping and of dead wood in particular is fundamental to managing our forests. Although individual tree-based approaches using lidar can successfully distinguish between deciduous and coniferous trees, the classification of multiple tree species is still limited in accuracy. Moreover, the combined mapping of standing dead trees after pest infestation is becoming increasingly important. New DL methods outperform baseline ML approaches and promise a significant accuracy gain for tree mapping. In this study, we performed a classification of multiple tree species (pine, birch, alder) and standing dead trees with crowns using the 3D DNN PointNet++ along with UAV-based lidar data and MS imagery. Aside from 3D geometry, we also integrated laser echo pulse width (EW) values and MS features into the classification process. In a preprocessing step, we generated the 3D segments of single trees using a 3D detection method. Our approach achieved an OA of 90.2% and was clearly superior to a baseline method using an RF classifier and handcrafted features (OA = 85.3%). All in all, we demonstrate that the performance of the 3D DNN is highly promising for the classification of multiple tree species and standing dead trees in practice.

4.1. Introduction

Forest inventories based on remote sensing data, particularly lidar point clouds fused with optical imagery, are the most prominent options for the inventory of forest structural variables (Latifi and Heurich, 2019). Forest attributes such as above-ground biomass and growing stock can be estimated from the spatial distribution of tree species and dead wood. Tree-level approaches utilize segmented single trees for forest inventory parameter estimations. For forest managers and nature conservationists, information about tree species, especially the classification of dead trees, is of increasing importance because forests are suffering from changing climatic conditions.

In the past, extensive research has been conducted to apply appropriate classifiers such as support vector machine (SVM), RF, or logistic regression to classify presegmented single trees with respect to tree species (Fassnacht et al., 2016) and dead trees (Yao et al., 2012). Most methods have been based on handcrafted feature sets extracted from ALS data and MS or hyperspectral imagery. Polewski (2017) successfully combined single 3D tree segments with MS aerial imagery to detect standing dead trees in a binary classification. The authors incorporated MS features generated from the covariance matrix of three image channels and classified dead trees with an OA of ~88%. Moreover, Degerickx et al. (2018) distinguished healthy (precision = 93%,

recall = 83%) from unhealthy (precision = 71%, recall = 88%) deciduous trees using ALS data and hyperspectral imagery in a regression method. Recently, Amiri et al. (2019) reported a combined classification of tree species and standing dead trees with crowns. Using a huge feature set generated from multi-wavelength lidar point clouds, four tree classes were classified with an OA of 82%. Interestingly, dead trees were only classified with 76% precision and 73% recall. However, all in all, the performance of these approaches for individual tree species classification is still not sufficient for practical use.

Currently, the utilization of high-performance DL methods as a classification tool for 3D sensed data has gained a large amount of interest in the remote sensing community. Various authors have demonstrated that standard ML concepts using, for example, SVM or RF, can be outperformed by DL-based methods (Voulodimos et al., 2018; Liu et al., 2018). One big advantage of DNNs is the automatic extraction of features as part of the training process, or so-called representation learning (LeCun et al., 2015). Griffiths and Boehm (2019) emphasized four general types of DL approaches for scene understanding from 3D sensed datasets. To utilize well-proven and efficient 2D CNNs, irregular and unordered 3D point clouds can either be transformed into RGB-depth (RGB-D) images (Zhao et al., 2018) or utilized to render multiview images (Qi et al., 2016). Furthermore, the authors discussed volumetric approaches that discretize raw 3D data, that is, as regular 3D voxel grids, and that use 3D convolutions to extract meaningful information (Zhou and Tuzel, 2018). Finally, powerful network architectures have been developed to enable a direct input of raw and unstructured point clouds without the need for a prior rasterization or voxelization. These innovative networks such as PointNet (Qi et al., 2017a), PointNet++ (Qi et al., 2017b), PointCNN (Li et al., 2018), and Super Point Graphs (Landrieu and Simonovsky, 2018) allow end-to-end classification.

To the best of our knowledge, the application of DNNs for the classification of presegmented single trees has been sparsely investigated. In urban study areas, Wegner et al. (2016) applied latest CNN-based methods to extensive datasets comprising aerial and street view images. The authors demonstrated that multiview imagery significantly improved tree detection and tree species classification, reaching close to human performance. Furthermore, Hartling et al. (2019) classified eight tree species using DenseNet (Huang et al., 2017), data from satellite imagery, and lidar data (approximately 1 point/m²) in urban study areas (OA = 83%). Moreover, Hamraz et al. (2019) generated images from ALS point clouds and made use of a CNN to classify overstory coniferous and deciduous trees in a natural forest with a cross-validated classification accuracy of 92% and 87%, respectively. So far,

using "real" 3D DNNs for vegetation mapping has not been researched sufficiently. Recently, Briechle et al. (2019) achieved promising results for adapting PointNet++ to the semantic labelling of extensive ALS point clouds, resulting in an OA = 85% for spruces and beeches.

The key idea of the current study was to adapt a 3D DNN for the classification of multiple tree species based on presegmented single tree objects. Specifically, we applied PointNet++ to a dataset composed of UAV-based lidar (including EW) and five-channel MS imagery. All in all, PointNet++ achieved excellent classification results at the single-tree level and clearly outperformed the baseline method. Furthermore, we demonstrated that MS data clearly enhanced the classification result.

In the following sections, we address the study area, sensors, data preprocessing, and reference data. Subsequently, we present the methodology for tree species classification using PointNet++ and compare it with the baseline method. Next, we demonstrate the conducted experiments and the main outcomes, including a comparison of both methods. Finally, we discuss the results referring to previous research and draw conclusions.

4.2. Materials

4.2.1. Study area

In two UAV flight missions (November of 2017 and April of 2018), both lidar data and MS images were captured in the study area ChEZ, located approximately 1.5 km west of the ChNPP (Figure 32). This densely vegetated area (37 ha) comprises approximately 400 trees/ha with tree heights of up to 30 m (Bonzom et al., 2016). The three main tree species are silver birch (*Betula pendula*), scots pine (*Pinus sylvestris*), and black alder (*Alnus glutinosa*). Moreover, standing dead trees with crowns (solely pines) can be found in the area.

Figure 32: Overview of the study area, located ~1.5 km west of the ChNPP (base map source: bing map ©Microsoft Corporation).

4.2.2. Sensors and data preprocessing

During both flight missions, an octocopter was utilized; it was developed by a team from the Department of Nuclear Physics Technologies of the Institute of Environment Geochemistry of the National Academy of Sciences of Ukraine. The copter enabled surveys, simultaneously recording with the lidar system and two MS cameras.

4.2.2.1. Lidar data

Lidar data with a nominal point density of 53 points/m² were collected in five automatic flights using a YellowScan Mapper I laser scanner at a constant altitude of 50 m. To generate a geometrically reliable 3D dataset, various postprocessing steps were conducted. First, differential postprocessing using a GNSS base station resulted in flight trajectories with centimetre-level precision. Second, the boresight angles provided by the manufacturer were checked in a calibration flight. Third, geometrically consistent lidar point clouds were generated by simultaneously aligning the flight strips (Jalobeanu and Gonçalves, 2014). Fourth, absolute 3D georeferencing was achieved by fitting the ALS point cloud to the enclosing polygons of a nearby building. Additionally, the sensor provided the intensity values for each laser point equivalent to the EW measured at a fixed internal sensor threshold. Because tree species classification can benefit from these measurements, we performed a data-driven correction step (Briechle et al., 2020). Finally, we performed single-tree segmentation using a normalized cut algorithm, resulting in single tree point clouds and enclosing tree polygons (Reitberger et al., 2009).

4.2.2.2. MS imagery

Five-band MS images (ground sample distance = 8.9 cm) were captured using two MicaSense RedEdge cameras (spectral range 475–840 nm) mounted in a twisted configuration with an angle of 22.5° (50% side overlap). Guaranteeing an extended camera footprint (FOV = 70°) equal to the lidar footprint, this setup allowed for a constant line-to-line distance for both lidar and MS sensors in a combined survey. For postprocessing the five-channel images, we utilized structure-from-motion software[14]. The processing steps included bundle adjustment (mean reprojection error of 1.3 pixels), calibration of reflectance, and the generation of dense photogrammetric 3D point clouds (80 points/m²) and 10-cm orthomosaics. Because the overflown study area is inaccessible, no GCPs were used. Therefore, photogrammetric point clouds were registered to georeferenced lidar point clouds using an ITC

---

[14] Agisoft PhotoScan Professional 1.4.1.

algorithm[15], resulting in a root mean squared error of 0.237 m (Briechle et al., 2018).

4.2.3. Reference data

Because of the high radiation dose rates within the study area, reference data were generated based on visual interpretation of 3D point clouds and MS imagery. In total, we manually labelled 1135 single tree segments assigned to the four tree classes "pine" (368 samples), "birch" (243 samples), "alder" (283 samples), and "dead tree" (241 samples), respectively.

4.3. Methodology

In the following sections, we describe the baseline method including feature engineering, classifier training and feature selection procedure. Furthermore, we give a detailed description of the classification process with the 3D DNN. Specifically, we address the preparation of dataset as well as network training, hereby focusing on hyperparameters and data augmentation.

4.3.1. Baseline method

4.3.1.1. Extraction of handcrafted features

The feature set generated from 3D lidar data (Table 17) comprised features based on the tree geometry (*GEOM*) and the laser echo characteristics (*EC*).

Table 17: 32 GEOM and 14 EC features.

Features	Definition
GEOM(1−10)[16]	Density distribution of points per height layer.
GEOM(11−20)	Vertical distribution of tree substance per height layer.
GEOM(21−30)	Mean distance of points to segment centre.
GEOM(31−32)	Std of distance from crown points to segment centre, in x and y direction.
EC1	Mean *EW* of points of a single tree.
EC(2−11)	Mean *EW* of points of a single tree per height layer.
EC12	(Σ middle / Σ first) reflections.
EC13	(Σ single / Σ first) reflections.
EC14	(Σ first + Σ middle) / (Σ single + Σ last) reflections.

Moreover, we developed distinctive features from the five-channel orthomosaics. For this purpose, we computed five VIs from the available spectral channels. First, we calculated the NDVI, a well-known index

---

[15] CloudCompare 2.8 [GPL software].
[16] Increasing numbering from bottom (1) to top (10).

sensitive to healthy vegetation rich in chlorophyll and robust over a wide range of conditions (Rouse Jr et al., 1973):

$$\text{NDVI} = \frac{NIR - R}{NIR + R} \qquad [7]$$

Second, utilizing both RE and NIR channels, the RENDVI (Gitelson and Merzlyak, 1994) was computed. This index is a NDVI modification and has been developed for applications including forest monitoring and vegetation stress detection. RENDVI is capable of detecting small changes in canopy foliage content (Sims and Gamon, 2002):

$$\text{RENDVI} = \frac{NIR - RE}{NIR + RE} \qquad [8]$$

Third, we introduced a NDVI-inspired index. Instead of the NIR channel, the RE channel was used to generate Red Edge Difference Vegetation Index (REDVI):

$$\text{REDVI} = \frac{RE - R}{RE + R} \qquad [9]$$

Fourth, we utilized the Modified Red Edge Simple Ratio (MRESR), which is used for forest monitoring and vegetation stress detection, incorporating a correction for leaf specular reflection (Datt, 1999):

$$\text{MRESR} = \frac{NIR - B}{RE - B} \qquad [10]$$

Fifth, we included the Modified Chlorophyll Absorption Ratio Index (MCARI), a well-suited index to indicate the relative abundance of chlorophyll. Daughtry et al. (2000) introduced this index, minimizing the combined effects of soil and non-photosynthetic surfaces.

$$\text{MCARI} = \frac{RE}{R} * (0.8 * RE - R - 0.2 * G) \qquad [11]$$

We superimposed the enclosing tree polygons on the orthomosaic (Figure 33) to mask VI pixels located within the tree segments. For each of these pixels, statistical features were calculated and standardized for each object (Table 18). These resulting 60 MS features were complemented with 10 independent interchannel covariance values generated from the covariance matrix of the five VI channels. Using this feature set, an RF classifier was trained on the labelled dataset and optimized in a three-times-repeated five-fold cross-validation. Finally, we identified the five most important MS features by evaluating the feature ranking based on the mean decrease in accuracy. In descending order, these were *NDVI_skewness*, *MRESR_perc90*, *NDVI_perc90*, *RENDVI_mode*, and *MRESR_mode*.
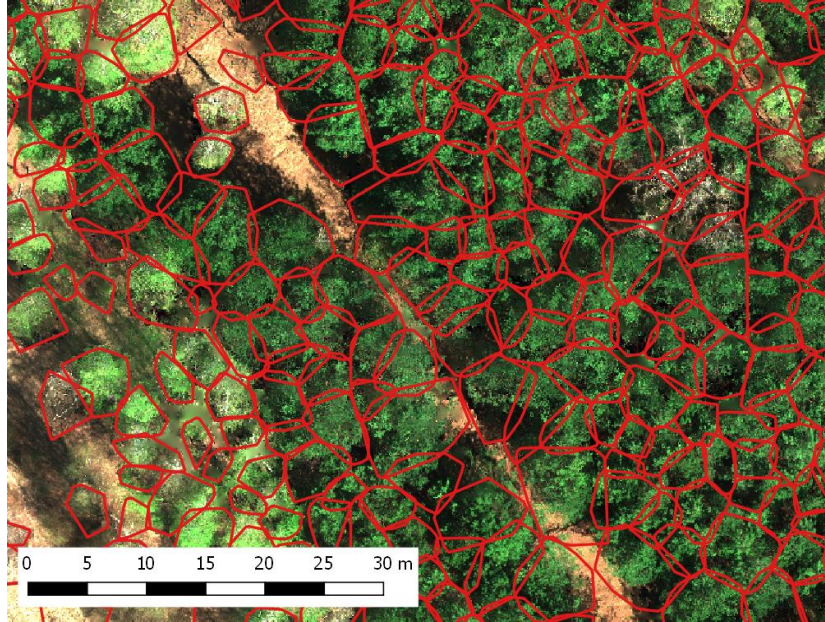
Figure 33: Superimposed tree polygons on the orthomosaic.

Table 18: Object-based statistical MS features.

Features	Definition
max, min, interval	Maximum value, minimum value, and range (max-min).
mean, std	Mean and std value.
mode	Value that appears most often.
skewness	Measure of asymmetry of the probability distribution.
kurtosis	Measure of tailedness of the probability distribution.
perc(25,50,75,90)	25th ('1st quartile'), 50th ('median'), 75th ('3rd quartile'), and 90th percentile.

4.3.1.2. Classifier training

For the baseline method, the dataset comprised 32 *GEOM* features and 14 *EC* features (see Table 17), as well as the five most important *MS* features generated from the VI orthomosaics. In a preprocessing step, highly correlated redundant features were eliminated from the feature set, here based on the application of a threshold (0.9) to feature-to-feature cross-correlation (Briechle et al., 2018). Next, an RF classifier was trained, including RFE based on Kuhn (2008) and a feature relevance assessment. Finally, the generalization quality of the RF classifier was verified by calculating classification metrics (OA, κ, precision, recall, and F_1 score) on the test dataset.

4.3.2. Classification using 3D DNN

PointNet++ is an advanced version of PointNet and incorporates hierarchical feature learning by extracting features from multiple contextual scales. Therefore, fine-grained local patterns and more general global features can be captured. In the following sections, we demonstrate the methodology for the utilization of PointNet++ to classify three tree species (pine, birch, alder) and standing dead trees using the PyTorch implementation from Wijmans (2018).

4.3.2.1. Preparation of dataset

Point sampling: For object classification, PointNet++ requires a constant number of 3D points per sample (e.g., *NUM_POINT* = 1,024; see Table 19). In practice, the distribution of points per tree is fairly heterogeneous due to the variations in the size, geometry, and species of single trees. Thus, an effective approach must meet the following conditions: First, a constant and adequate number of points per tree has to be guaranteed, and loss of information during downsampling needs to be minimized. Second, deletion of samples containing less points than *NUM_POINT* but still exceeding an acceptable number of points should be avoided. Third, synthetic generation of redundant information by extensive upsampling is not reasonable. Therefore, we introduced the two thresholds θ_1 and θ_2 in a combined sampling approach. θ_1 was utilized to randomly reduce the points per tree to a certain value. Figure 34 exemplary shows the number of remaining samples per class, in dependence of θ_1. To preserve the selected objects comprising less than θ_1 points in the dataset, we made use of a second threshold, θ_2. Trees containing at least θ_2 points were sampled up to θ_1 points using random copies of points. All in all, our procedure handled the trade-off between upsampling and downsampling, assuming that both thresholds are chosen appropriately.
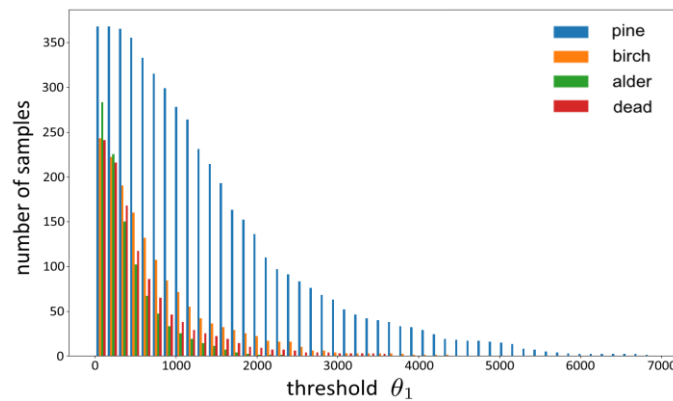
Figure 34: Number of remaining samples per tree class in dependence of threshold θ_1.

Dataset generation: Initially, the remaining samples were balanced according to the four occurring tree classes. Next, all single point clouds were standardized by subtraction of the mean x, y, and z coordinates and division by the std of x, y, and z values. Consequently, all objects were rescaled and had a mean of 0 and an std of 1. Practically, the purpose of standardization is to make the classification results independent of the geometry within each tree class, for example, the tree height and the crown width. Moreover, the *EW* values were standardized as well. Subsequently, we calculated surface normals (Figure 35) using the *estimate_normals* function from the open source library Open3D (Zhou et al., 2018). The two key arguments of the function, *radius* and *max_nn*, were set to 0.5 and 30, respectively. The parameter *radius* specifies the search radius for the neighbourhood definition, whereas *max_nn* defines the maximum number of nearest neighbours to be considered to save computation time. Next, the top five *MS* features were integrated by assigning the standardized values to each 3D point of an object (tree species, dead tree). Note that this procedure provides additional point attributes. All in all, we generated a dataset comprising raw point clouds, surface normals, echo widths per point, and five previously calculated handcrafted *MS* features.
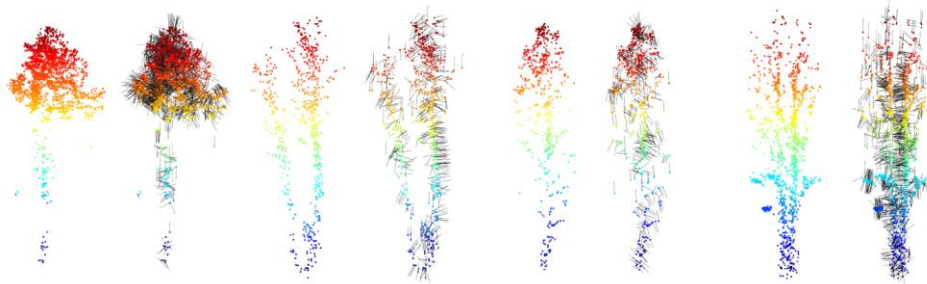
Figure 35: Samples of 3D point clouds per tree class; for each class, the samples on the right show surface normals.

4.3.2.2. Training and validation

Hyperparameters: PointNet++ is an off-the-shelf 3D DNN. Nevertheless, it is essential to consider various options to optimize network performance for specific classification tasks without model overfitting. To get a well-performing network, the most decisive PointNet++ hyperparameters were adjusted using a combination of manual search and automated grid search (Table 19). For some parameters, the default values were convenient and, therefore, remained unchanged.

Data augmentation: A popular method to avoid model overfitting on a small training dataset is the utilization of data augmentation. Furthermore, performing data augmentation during network training helps to make the neural network more robust against object variation. Before

each training epoch, we shuffled the order of samples to generate random batches. Next, we performed random transformations of the standardized 3D objects by following common practice including scaling (range = [0.80, 1.25]), rotation around vertical axis (range = [0, 2*pi]), jittering with Gaussian noise (range = ±0.05 [m]), and 3D translation of the entire point cloud (range = ±0.1 [m]). Furthermore, we set the *random input dropout* parameter to MAX_DROPOUT = 50%, thereby increasing the robustness to varying point densities and occluded object parts. Practically, the input points for each instance were randomly dropped out, generating subvolumes of the objects.

Model evaluation: For testing of the trained network, class labels were predicted on trees that were not used for the training. We compared these class predictions with the reference labels and calculated standard metrics OA, κ, precision, recall, and F_1 score. For final evaluation, we used the model showing the lowest validation loss.

Table 19: Hyperparameters and default / optimized values for PointNet++.

Hyperparameter	Value	Declaration
NUM_CLASSES	4	Number of object categories.
NUM_POINT	1,024	Number of 3D points per sample.
MAX_DROPOUT	0.5	Maximal dropout rate.
BATCH_SIZE	8	Number of samples per batch.
MAX_EPOCH	300[17]	Number of training epochs.
BASE_LR	1e-3	Initial learning rate (LR).
LR_DECAY	0.7	Initial LR decay.
BN_MOMENTUM	0.5	Initial batch norm momentum.
BNM_DECAY	0.5	Batch norm momentum decay.
OPTIMIZER	Adam	Optimization algorithm.
WEIGHT_DECAY	1e-4	L2 regularization coefficient.

4.4. Experiments and results

4.4.1. Experimental setup

The original reference dataset was prepared for object classification, performing point sampling (θ_1 = 1,024; θ_2 = 512) and class balancing (see Section 4.3.2.1). The remaining 668 samples (167 per class) were divided into 464 training and 204 test samples using a split ratio of 0.7. Note that for a fair comparison of 3D DNN and baseline method, the particular training and test datasets were identical. For network training and validation, we used an Intel Xeon Platinum 8160 central processing unit (CPU) and an Nvidia Titan V graphics processing unit (GPU) (NVIDIA Corporation, 2019) with 12 GB on Ubuntu 18.04, reach-

---

[17] No early stopping criterion was used.

ing a processing time of approximately 10 seconds per epoch. We performed classification with PointNet++ on four different datasets investigating their impact on the classification result. In more detail, the datasets represented geometry (*GEOM*, see Figure 36), geometry and surface normals (*GEOM+normals*, see Figure 37), geometry and EW values (*GEOM+EW*, see Figure 38), and all data subsets (*GEOM+EW+MS*, see Figure 39). Furthermore, we conducted comparative experiments with the previously described baseline method (RF). For validation, we compared both classifier procedures on the same test dataset.
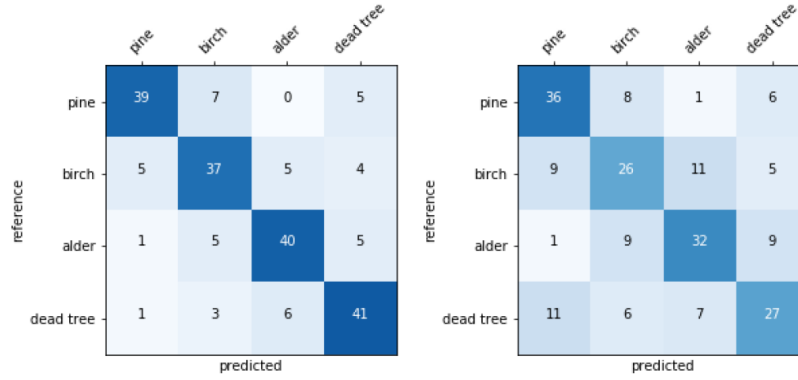
Figure 36: Confusion matrices on the test dataset using only geometry information.
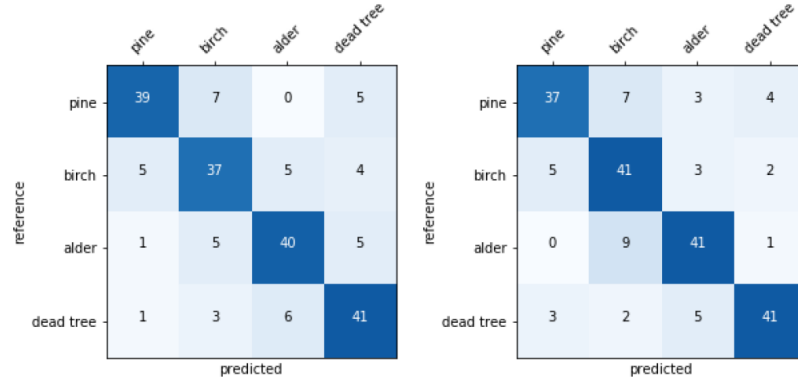
Figure 37: Confusion matrices on the test dataset using only geometry information and PointNet++ exclusive (a) and inclusive of (b) surface normals.
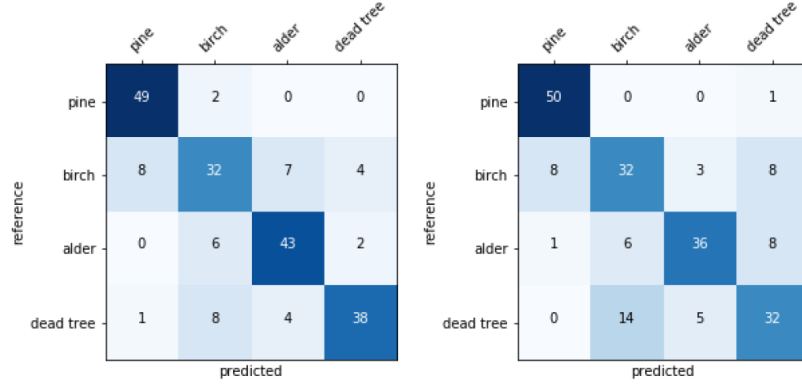
Figure 38: Confusion matrices on the test dataset using only geometry information and EW values.
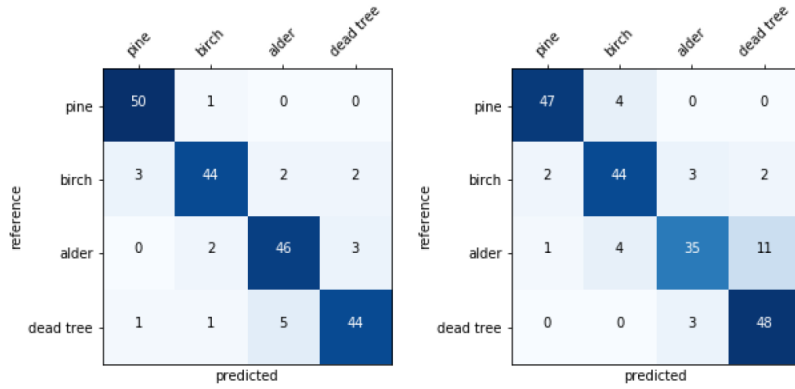
Figure 39: Confusion matrices on the test dataset using geometry information, EW values, and MS features.

4.4.2. General classification results

PointNet++ outperformed the baseline method in all experiments (Table 20). Especially, if only geometry information was used, PointNet++ and automatically extracted features led to a result that was 17.7% better than the baseline method using 32 "standard" handcrafted geometry features. Adding surface normals improved the DNN result by 1.4%. Here, no comparison to the baseline was available. Fusing geometry data with *EW* data, the OA increased by 1.0% (DNN) and 14.2% (RF), respectively. Using this feature set generated from lidar data, the DNN (OA = 79.4%) was 5.9% better than the baseline method (OA = 73.5%). Including five top MS features – namely *NDVI_skewness*, *MRESR_perc90*, *NDVI_perc90*, *RENDVI_mode*, *MRESR_mode* – the OA increased by approximately 11% for both

methods. Using all data subsets, PointNet++ (OA = 90.2%) outper-formed the baseline method (OA = 85.3%) by 4.9%.

Table 20: Classification results using different data subsets.

Feature sets	PointNet++		RF	
	OA [%]	κ	OA [%]	κ
GEOM	77.0	0.693	59.3	0.458
GEOM+normals	78.4	0.712	---	---
GEOM+EW[18]	79.4	0.725	73.5	0.647
GEOM+EW+MS[19]	**90.2**	0.869	**85.3**	0.804

4.4.3. Analysis of results using baseline method

The classification of multiple classes with the baseline method utilizing only geometry features performed fairly poor (Figure 40). Adding *EW* data increased all F_1 scores, with a major improvement of 0.24 for pine. Moreover, the top five *MS* features especially boosted the F_1 scores of birch by 0.23 and dead tree by 0.22 but could not improve alder classification. Overall, the F_1 scores ranged between 0.76 and 0.93. The feature ranking of the RF classifier clearly confirmed the importance of *MS* features for tree species classification, with all five *MS* features being ranked in the top 10 of the most important features (Table 21). Unsurprisingly, five of the *EC* features were also ranked in the top 10. These features mainly represent the interaction of the laser beam with the top layers of the tree (*EC10*, *EC11*) and penetration to the ground (*EC13*, *EC14*). Furthermore, the mean *EW* value of the laser points of a single tree (*EC1*) was ranked eighth. Finally, none of the geometry features was ranked in the top 10.

Table 21: Top 10 features using RF classifier and all data subsets.

Feature name	Feature importance[19]
NDVI_skewness	100.0
MRESR_perc90	88.6
NDVI_perc90	85.6
EC10	59.5
RENDVI_mode	54.1
EC11	52.3
EC14	39.6
EC1	39.2
EC13	36.2
MRESR_mode	33.3

---

[18] Due to the architecture of PointNet++, surface normals are mandatory when adding extra attributes such as EW values or MS features.
[19] Normalized mean decrease in accuracy.

4.4.4. Analysis of results using 3D DNN

In general, the results demonstrated that PointNet++ is an efficient 3D DNN for the classification of three tree species and dead trees using point clouds (see Figure 41). In particular, the experiments showed that the inclusion of surface normals to the geometry data improved the F_1 score for standing dead trees by 0.06. Incorporating EW values mainly led to a high F_1 value for pine (F_1 score = 0.90). Nevertheless, the F_1 score for birch decreased by 0.10 to a relatively low value of 0.65. Adding the top five MS features enhanced all F_1 scores. Interestingly, the F_1 score for birch clearly increased by 0.24. When utilizing all subsets, the F_1 scores ranged between 0.88 and 0.95.

	GEOM	GEOM+normals	GEOM+EW	GEOM+EW+MS
pine	0.67	nan	0.91	0.93
birch	0.52	nan	0.62	0.85
alder	0.63	nan	0.76	0.76
dead	0.55	nan	0.64	0.86
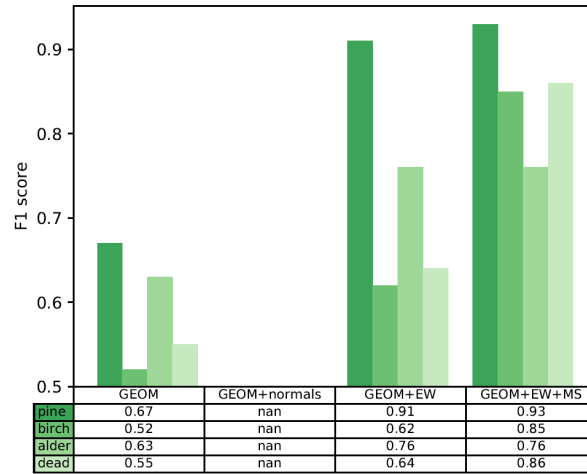
Figure 40: F_1 scores per class using RF.

	GEOM	GEOM+normals	GEOM+EW	GEOM+EW+MS
pine	0.80	0.77	0.90	0.95
birch	0.72	0.75	0.65	0.89
alder	0.78	0.80	0.82	0.88
dead	0.77	0.83	0.80	0.88
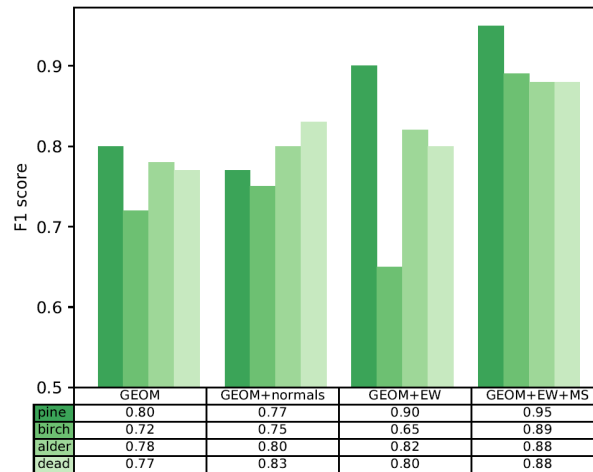
Figure 41: F_1 scores per class using PointNet++.

4.5. Discussion

The proposed framework using PointNet++ for the classification of three single tree species and standing dead trees performed fairly good. Especially, when classification was only conducted based on geometry information, the results were significantly better than those of the baseline method. Obviously, handcrafted geometry features are considerably inferior to information automatically extracted in a DNN. If we analyse the confusion matrices, we notice a higher confusion between alder and dead trees. Very likely, the tree geometry and spectral appearance of alder is similar to dead pines. Stepwise improvement of the results produced by PointNet++ was rather low when we fused surface normals and *EW* values with geometry data (1.5% and 1.0%, respectively). Interestingly, adding surface normals particularly increased the classification accuracy for dead trees. Also very important, the classification of pine, the only conifer in our study area, profited most by the *EW* values (F_1 score = 0.90), thereby confirming the findings of Reitberger et al. (2009). Furthermore, we included five MS features that were selected by the RF-based feature assessment. Embedding these features, the overall results were considerably enhanced for both methods by approximately 11% (see Table 20). Especially, the classification of birch and dead tree benefited from these *MS* features. Note that at the time of data collection, birches had already sprouted. Therefore, their characteristic spectral appearance supported the classification significantly.

Investigating the related work reveals that our approach achieves very promising and competitive results. For the classification of individual tree species, most previous studies based on classic ML approaches did not reach an acceptable accuracy level of up to 90%. Yu et al. (2017) classified three tree species using MS ALS data and an RF classifier (OA = 86%). Moreover, Shi et al. (2018a) categorized five species, fusing ALS data with hyperspectral imagery (OA = 84%). Kamińska et al. (2018) classified three tree species (spruce, pine, deciduous), each of them further categorized as "dead" or "alive". Their approach using an RF classifier and features generated from ALS data and CIR imagery reached an OA of 94%. Nevertheless, a comprehensive and, thus, fair comparison to other studies that have addressed classification of pre-segmented single trees is challenging. Collecting data using a huge variety of sensor platforms and sensor types, utilized datasets strongly differ in their spatial, spectral, and temporal resolution. Additionally, the type of study area (urban, natural, managed) and number of samples and classes fluctuate as well.

We would also like to address some limitations of PointNet++ for classification tasks. Because PointNet++ can only deal with objects com-

prising a constant number of points, point sampling including upsampling and downsampling must be performed. Thereby, information loss is unavoidable and must be minimized based on reasonable thresholds, depending on the specific point density of the dataset. Nevertheless, this disadvantage is clearly compensated by the DNN performance with its ability to automatically extract meaningful information from 3D datasets. Moreover, 3D DNNs such as PointNet++ need to be trained from scratch using a specific and fairly high number of training samples. Contrary to well-known 2D CNNs, no publicly available databases such as ImageNet (Deng et al., 2009) can be used for transfer learning and reasonable weight initialization.

4.6. Conclusion

Our experiments demonstrated that 3D DNN PointNet++ could successfully be applied to the classification of three tree species – pine, birch, and alder – and standing dead trees. Fusing UAV-based lidar data and features generated from five-channel MS imagery, we achieved an OA better than 90% at the single-tree level. Moreover, classification with PointNet++ was clearly superior to the described baseline method in all cases. All in all, our DL-based approach provided detailed and reliable 3D vegetation maps at the tree level in the study area ChEZ. In a next step, a large scale experiment in an extended forest area is intended to verify the promising results of this current study, thereby demonstrating the suitability for practical use.

Acknowledgements

5 Silvi-Net – A dual-CNN approach for combined classification of tree species and standing dead trees from remote sensing data[20]

Abstract

Forest managers and nature conservationists rely on the precise mapping of single trees from remote sensing data for efficient estimation of forest attributes. In recent years, additional quantification of dead wood in particular has garnered interest. However, tree-level approaches utilizing segmented single trees are still limited in accuracy and their application is therefore mostly restricted to research studies. Furthermore, the combined classification of presegmented single trees with respect to tree species and health status is important for practical use but has been insufficiently investigated so far. Therefore, we introduce Silvi-Net, an approach based on CNNs fusing airborne lidar data and MS images for 3D object classification. First, we segment single 3D trees from the lidar point cloud, render multiple silhouette-like side-view images, and enrich them with calibrated EC. Second, projected outlines of the segmented trees are used to crop and mask the MS orthomosaic and to generate MS image patches for each tree. Third, we independently train two ResNet-18 networks to learn meaningful features from both datasets. This optimization process is based on pre-trained CNN weights and recursive retraining of model parameters. Finally, the extracted features are fused for a final classification step based on a standard MLP and majority voting. We analysed the network's performance on data captured in two study areas, the ChEZ and the Bavarian Forest National Park (BFNP). For both study areas, the lidar point density was approximately 55 points/m² and the GSD values of the true orthophotos were 10 cm (ChEZ) and 20 cm (BFNP). In general, the trained models showed high generalization capacity on independent test data, achieving an OA of 96.1% for the classification of pines, birches, alders, and dead trees (ChEZ) – and 91.5% for coniferous, deciduous, snags, and dead trees (BFNP). Interestingly, lidar-based imagery increased the OA by 2.5% (ChEZ) and 5.9% (BFNP) compared to experiments only utilizing MS imagery. Moreover, Silvi-Net also demonstrated superior OA compared to the baseline method PointNet++ by 11.3% (ChEZ) and 2.2% (BFNP). Overall, the effectiveness of our approach was proven using 2D and 3D datasets from two natural forest areas (400–530 trees/ha), acquired with different sensor models, and varying geometric and spectral resolutions. Using the technique of transfer learning, Silvi-Net facilitates fast model convergence, even for datasets with a reduced number of samples. Consequently, operators can generate reliable maps that are of major importance in applications such as automated inventory and monitoring projects.

5.1. Introduction

In forestry, the precise and reliable mapping of tree species is a fundamental concern. The classification of dead wood in particular is of increasing importance because forests are suffering from changing climatic conditions. Furthermore, tree-level approaches are increasingly of interest in area-wide forest inventory. For instance, forest attributes such as above-ground biomass and growing stock can be estimated based on tree-specific allometric models (Chave et al., 2014). Moreover, forest managers and nature conservationists require quantitative mapping results to investigate the robustness and sustainability of various forest compositions (Overbeck and Schmidt, 2012). Besides these conventional applications of vegetation mapping, tree species information can also be advantageous in more unusual cases. For example, Briechle et al. (2020b) showed that observed vegetation anomalies are helpful for the detection of unknown radioactive waste sites in the ChEZ.

5.1.1. Conventional approaches

Traditionally, forest inventory has been based on manual field measurements. Forest managers have typically relied on sample-based procedures followed by area-wide extrapolation (McRoberts and Tomppo, 2007). Nevertheless, in situ inventory is labour intensive and, therefore, both time consuming and expensive. For a temperate forest area of around 300 km², Latifi et al. (2015) demonstrated that lidar-based data collection is 90% less expensive compared to a conventional forest inventory. Fassnacht et al. (2016) reviewed the work of various researchers who have investigated forest parameter estimation at the single-tree level using remote sensing data. Airplanes, helicopters, and innovative platforms such as UAVs equipped with lidar sensors and MS or hyperspectral cameras enable acquisition of high-resolution data from a bird's eye view. In particular, the fusion of lidar point clouds and optical multi-channel imagery is the most prominent option for the inventory of forest structural variables (Latifi and Heurich, 2019). In a preprocessing step, single trees are typically delineated from ALS data. This tree segmentation is mostly based on a canopy height model (CHM) (Pyysalo and Hyyppä, 2002; Solberg et al., 2006) or on the original 3D point cloud (Reitberger et al., 2009; Wu et al., 2016). After the segmentation process, extracted single tree objects can be classified according to tree species. Therefore, the majority of previous studies typically relied on a two-step approach. First, handcrafted feature sets describing the geometry and radiometry of single trees were generated from the remote sensing data. Second, appropriate ML classifiers, such as SVM or RF, were applied for classification. For example, Heinzel and Koch (2012) investigated different feature sets derived from full-waveform lidar data, hyperspectral data, and CIR images in

a temperate forest. Their SVM-based method could classify pine (*Pinus sylvestris*), spruce (*Picea abies*), oak (*Quercus petraea*), and beech (*Fagus sylvatica*) with an OA of 89.7%, 88.7%, 83.1%, and 90.7%, respectively. Dalponte et al. (2012) used airborne hyperspectral imagery and lidar data from a mountain area in the Southern Alps. They investigated the performance of both RF and SVM classifiers on different feature subsets generated from data with varying spatial resolutions. Overall, seven species and a "non-forest" class were classified with an OA of 83.0%. In a mixed temperate forest, Shi et al. (2018a) categorized five species by fusing ALS data with hyperspectral imagery (OA = 83.7%). The authors successfully combined plant functional traits (e.g., equivalent water thickness, leaf mass per area and leaf chlorophyll), spectral features, and LM.

Recently, the classification of dead trees has become increasingly important. Most previous studies regarded this task as a binary problem and classified tree objects into dead or living. For instance, Yao et al. (2012) utilized an SVM classifier and handcrafted features generated from full waveform lidar data (25 points/m²) captured in a mixed mountain forest in the BFNP. Based on features derived from the 3D point cloud, laser intensity and EW, their method classified dead and living trees with an OA of 73% for leaf-on trees and 71% for leaf-off trees. Polewski et al. (2015a) presented an active learning-based approach to detect standing dead trees (snags) in the BFNP. Using features from ALS point clouds and CIR imagery, manually labelled single trees were classified into dead and living with an OA of 89%. Casas et al. (2016) proposed a classification model based on single-tree ALS metrics and separated snags from living trees with an OA of 92%. In a comprehensive study, Kamińska et al. (2018) trained an RF classifier using intensity and structural variables from multi-temporal ALS data (6 points/m²) and spectral information generated from 20-cm leaf-on CIR images. Their method classified three tree species (spruce, pine, and deciduous), and further categorized them as "dead" or "alive" (OA = 94.3%). More recently, Krzystek et al. (2020) conducted a large-scale experiment in an area of 924 km² to classify single trees in the BFNP. Based on ALS data and CIR imagery, their binary classifier separated dead from living trees with an OA of 93%. In summary, the overall performance of approaches for individual tree species classification in dense (and thus complex) temperate forests is still insufficient for practical use, requiring an OA of at least 90% for multi-class tasks.

5.1.2. DL-based approaches

In recent years, utilizing high-performing DL methods as classification tools has garnered a large amount of interest, outperforming standard ML approaches in various tasks (Voulodimos et al. 2018). Presumably, the biggest advantage of these DNNs is their representation learning,

which characterizes the automatic extraction of features as part of the training process (LeCun et al., 2004). For scene understanding from irregular and unordered 3D point clouds, Griffiths and Boehm (2019) outlined four general types of DL approaches. On the one hand, the authors reviewed methods that either render multi-view images (Qi et al., 2016) or transform input data into RGB-D images (Zhao et al., 2018). Thus, proven and efficient 2D CNNs such as AlexNet (Krizhevsky et al., 2012), VGG (Simonyan and Zisserman, 2015), and ResNet (He et al., 2015) can be applied. On the other hand, the authors discussed volumetric approaches that discretize raw 3D data as regular 3D voxel grids and subsequently use 3D convolutions to extract meaningful information (Zhou and Tuzel, 2018). Recently, powerful network architectures such as PointNet++ (Qi et al., 2017b) and PointCNN (Li et al., 2018) have been developed. These 3D DNNs enable direct input of raw and unstructured point clouds without the need for prior rasterization or voxelization. Therefore, they allow end-to-end classification of 3D point clouds.

So far, the application of DL methods for the classification of presegmented single trees based on lidar data has been rarely investigated. Presumably, one reason for this research gap is the lack of large training datasets. In a natural forest (330 stems/ha), Hamraz et al. (2019) utilized a CNN to classify overstory coniferous and deciduous trees. By generating images from leaf-off and leaf-on ALS point clouds (50 points/m²), a cross-validated classification accuracy of 92% for coniferous trees and 87% for deciduous trees were reached. Overall, the CNN was up to 14% more effective than traditional learning methods using handcrafted features. In an urban study area, Hartling et al. (2019) fused data from satellite imagery and lidar data. Using Dense-Net (Huang et al., 2017), an OA of 83% in classifying eight individual tree species was achieved. Moreover, their approach was clearly superior to both RF (OA = 52%) and SVM (OA = 52%) classifiers, even with restricted training sample quantities. In a tropical wetland located in South China, Sun et al. (2019b) developed a patch-based classification algorithm for seven classes, including six individual tree classes (1388 training samples, 362 test samples). Initially, single trees were segmented by calculating a CHM from the lidar point cloud (5–8 points/m²). Then, the segment information was utilized to generate 64x64 image patches by cropping 10-cm aerial RGB images. Their most effective model was a modified version of ResNet-50, which classified image patches with an OA of 90%. In the same study area, Sun et al. (2019a) mapped 18 tree species using ALS data and high-resolution RGB images, achieving an OA of 73% at the single-tree level. Their approach involved the application of three well-known CNNs (AlexNet, VGG-16, and ResNet-50). Recently, Briechle et al. (2020a) classified three tree species (pine, birch, and alder) and standing dead pines with

crowns using PointNet++ along with UAV-based lidar data and MS imagery. Aside from 3D geometry, EW values and MS features were also integrated into the classification process. Overall, their DL-based method (OA = 90%) successfully used raw 3D data and was superior to a baseline method using an RF classifier and handcrafted features (OA = 85%).

5.1.3. Key idea and main issues

In the present chapter, the objective was to classify presegmented 3D single tree objects with respect to tree species and dead trees in a combined approach. Therefore, because of its proven outstanding performance, a CNN-based procedure was chosen. Additionally, we applied the technique of transfer learning to tree species classification. Instead of training all model parameters from scratch, this approach is based on pretrained parameters that are fine-tuned on the basis of a task-specific dataset. Especially for relatively small datasets, this procedure allows effective model adaptation, even if there is a considerable domain shift between an existing image collection and a new dataset (Prabha et al., 2020). Essentially, our approach was supported by the idea that a person would likely classify a single tree by looking at its silhouette from different angles. Our approach was further supported by a review of DL methods for 3D data, which found that systems using discrete 2D representations of 3D data typically outperform approaches based on 3D voxel representations (Ioannidou et al., 2017). We therefore wanted to investigate whether multi-view 2D images could also exceed point networks. Thus, the key idea of this study was to train a CNN fusing MS image patches and multiple side-view images generated from UAV-based and helicopter-based lidar data. In addition to the geometric information, we also incorporated calibrated laser EC into the classification pipeline. The experiments conducted examined the following research questions:

- Can this new method successfully be applied to data from two regions captured with different lidar sensors and MS cameras?
- Compared to the baseline approach using PointNet++, is there an improvement in classification accuracy when utilizing a CNN approach and multiple 2D representations of single trees?

Furthermore, we investigated some relevant practical issues:

- Is the masking of MS image patches necessary?
- Can the incorporation of laser EC improve performance?
- Which classes can be classified more accurately than others? Why are some classes particularly difficult to distinguish?

The most innovative contribution of our pipeline for single-tree classification is the fusion of MS image patches and multi-view images generated from 3D point clouds in a dual-CNN approach. Furthermore, we initialize the CNN models using pretrained weights and optimize the network parameters by recursive retraining. To visualize the networks' decisions, we use class activation mapping (CAM).

In the following sections, we address the study areas, sensors, data preprocessing, and reference data. Subsequently, we present our methodology for tree species classification and the baseline method. Then, we outline the conducted experiments and the main outcomes, including a comparison of both methods. Finally, we discuss the results in relation to previous research and draw conclusions.

5.2. Materials

5.2.1. Study areas

In this chapter, we present experiments building on datasets from two study areas. The first study area, ChEZ, is densely vegetated with a tree density of approximately 400 trees/ha. The main tree species are Scots pine (*Pinus sylvestris*), silver birch (*Betula pendula*), and black alder (*Alnus glutinosa*), with tree heights up to 30 m (Bonzom et al., 2016). Overall, the forest stand is dominated by Scots pine planted after the nuclear disaster of 1986 (Yoschenko et al., 2011), comprising approximately 50% of all trees. Based on visual interpretation of aerial imagery, we roughly estimated the distribution of pines, birches, and alders to be 50%, 20%, and 30%, respectively. The second study area, BFNP, was established in 1970 and is part of the Natura 2000 network, which was founded to protect the most endangered habitats and species in Europe. The BFNP contains protected flora and fauna of exceptional natural value (Zenáhlíková et al., 2015). The forest area is dominated by Norway spruce (*Picea abies*), European beech (*Fagus sylvatica*), silver fir (*Abies alba*), and larch (*Larix*). Furthermore, other tree species appear less frequently, such as silver birch (*Betula pendula*), sycamore maple (*Acer pseudoplatanus*), and common rowan (*Sorbus aucuparia*) (Cailleret et al., 2014). Due to bark beetle infestation, extensive areas are covered with dead wood – fallen dead trees, standing dead trees, and standing dead trees without crowns (also known as snags).

5.2.2. Data acquisition and preprocessing

In the ChEZ, we utilized an octocopter developed by a team from the Department of Nuclear Physics Technologies of the Institute of Environment Geochemistry of the National Academy of Sciences of Ukraine.

All flights were carried out in fully automatic mode using GNSS waypoints. Data collection was performed during sunny and partly cloudy weather conditions at a mostly constant wind speed (2–3 m/s). In April 2018, lidar data were collected by a YellowScan Mapper I laser scanner, resulting in a nominal point density of approximately 53 points/m². Before the data collection, a calibration flight over a building was conducted to check the boresight angles (BayesMap Solutions LLC, 2018) preset by the manufacturer. Differential GNSS postprocessing (NovAtel Inc., 2017) incorporating GNSS measurements collected by a Trimble R4 base station ensured flight trajectories with centimetre-level precision. Overall, the mean discrepancy between adjacent lidar strips was approximately 5 cm, which is in the range of the measurement accuracy of the instrument. Absolute 3D georeferencing with an accuracy of a few centimetres was achieved by fitting the ALS point cloud to the enclosing polygons of a nearby building. Moreover, the recorded lidar data were radiometrically corrected based on the data-driven method presented in Briechle et al. (2020b). Additionally, we captured MS images using two MicaSense RedEdge cameras that were mounted in a twisted configuration with an angle of approximately 23°. Compared to a FOV of 47° for a single camera setup, this setup guaranteed a 50% side overlap of the two camera footprints, thereby increasing the total FOV to approximately 70°. To compensate for changing lighting conditions during and between the flights, we utilized MicaSense's CRP and DLS. These accessories provided useful information for the subsequent reflectance calibration in Agisoft PhotoScan Professional 1.4.1 (Agisoft LLC, 2018). Next, all images were aligned in a bundle adjustment, resulting in a mean reprojection error of 1.3 pixels. Finally, 10-cm MS true orthophotos were generated using the lidar-based surface model as a reference.

In the BFNP, airborne full waveform data were acquired in June 2017 (leaf-on condition) using a Riegl LMS-Q680i instrument carried by a helicopter. The resulting average point density was 55 points/m². Additionally, a calibration flight was conducted on a nearby airfield and enabled the correction of the raw amplitude values with regard to travelling distance of the laser beam (Amiri et al., 2019). Next, georeferencing quality was checked based on in-field measurements of vertical and planimetric objects, such as flat areas and enclosed building polygons, respectively. On average, the mean 3D displacements of the lidar data were less than 10 cm. MS aerial imagery in the BFNP was also acquired in June 2017, using a Leica DMC III camera. GNSS data and Inertial Navigation System data provided initial values for the exterior camera orientation. Using the software package Agisoft PhotoScan Professional 1.4.1, the aerotriangulation was performed based on aerial images, a camera calibration model, and GCPs, leading to a sigma naught of 30% of the GSD. Next, we generated true orthophotos on

the basis of the lidar-based digital surface model. Finally, single trees were delineated from the lidar point cloud in both study areas utilizing the normalized cut algorithm presented by Reitberger et al. (2009). Following the authors' recommendations, we set the static stopping criterion of the normalized cut segmentation to 0.16. The segmentation quality was not tested quantitatively, however visual inspection helped to verify that no major oversegmentation or undersegmentation occurred. Aside from individual point clouds, the segmentation also provided projected 2D polygons for each tree. Table 22 shows an overview of study areas, sensor platforms, sensor equipment, and data acquisition parameters.

Table 22: Study areas and sensor equipment.

	ChEZ	**BFNP**
Location	51°23'N, 30°04'E	49°04'N, 13°18'E
Size of study area	37 ha	8.3 km²
Tree density	400 trees/ha	530 trees/ha
Tree height	15–30 m	15–50 m
Platform	UAV (octocopter)	Helicopter D-HFCE/AS350
Lidar sensor	YellowScan Mapper I	Riegl LMS-Q680i
Laser wavelength	905 nm	1550 nm
Echo characteristics	Pulse width	Intensity
Flight altitude	50 m	550 m
Flight speed	6 m/s	30 m/s
Point density	53 points/ m² (leaf-off)	55 points/m² (leaf on)
MS camera	MicaSense RedEdge	Leica DMC III
Focal length	5.5 mm	92 mm
MS bands	blue (B), green (G), red (R), RE, NIR	B, G, R, NIR
Flight altitude	130 m	2880 m
Flight speed	9 m/s	30 m/s
End lap/side lap (%)	79/50	80/60
GSD of orthomosaics	10 cm	20 cm

5.2.3. Reference data

Based on visual interpretation, single tree segments were manually labelled using an interactive tool. Note that incorrect segments were generally not considered in the labelling process to make our classification results independent of the segmentation quality. In the user interface, randomly chosen tree segments are displayed in 3D. The point cloud can be rotated, thereby supporting the annotator in verifying the class label. Furthermore, the corresponding 2D polygon for each segment is superimposed on the aerial image. In detail, the trees in the ChEZ were manually subdivided into the classes "pine", "birch", "al-

der", and "dead tree". In the BFNP, we labelled the trees with the categories "coniferous" (mostly spruce), "deciduous" (mostly beech and larch), "snag", and "dead tree" (Figure 42). Here, "snag" refers to a partly or completely dead tree missing a crown or most of the smaller branches (Yao et al., 2012). In contrast, trees labelled "dead tree" are dead trees with crowns. The distinction between "snag" and "dead tree" was based on the subjective perception of three different research assistants. Subsequently, the labelled samples were randomly sorted into training, validation, and test datasets (see Table 23 and Table 24). Note that we also included class balancing for both training and validation data.
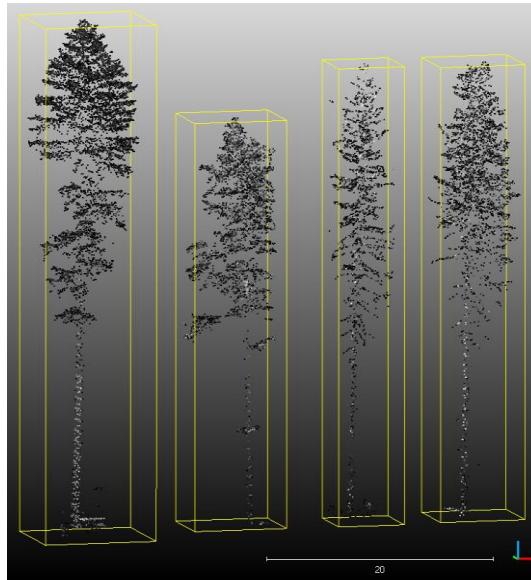
Figure 42: 3D point clouds for selected samples from BFNP dataset, coloured by normalized intensity from black (0) to white (1). From left to right: coniferous, deciduous, snag, dead tree.

Table 23: Number of samples for study area ChEZ; train/val/test split: 56%/14%/30%.

Tree class	Training samples	Validation samples	Test samples
pine	93	23	51
birch	93	23	51
alder	93	23	51
dead tree	93	23	51
Σ	372	92	204

Table 24: Number of samples for study area BFNP; train/val/test split: 51%/22%/27%.

Tree class	Training samples	Validation samples	Test samples
coniferous	345	149	259
deciduous	345	149	202
snag	345	149	139
dead tree	345	149	145
Σ	1380	596	745

5.3. Methodology

5.3.1. Outline of the proposed method

In general, our network architecture is inspired by DualNet (Hou et al., 2017), a DNN which includes two parallel CNNs and a subsequent aggregation of complementary features in a final classifier. For a better understanding of the overall processing pipeline, important steps of Silvi-Net are illustrated in Figure 43.
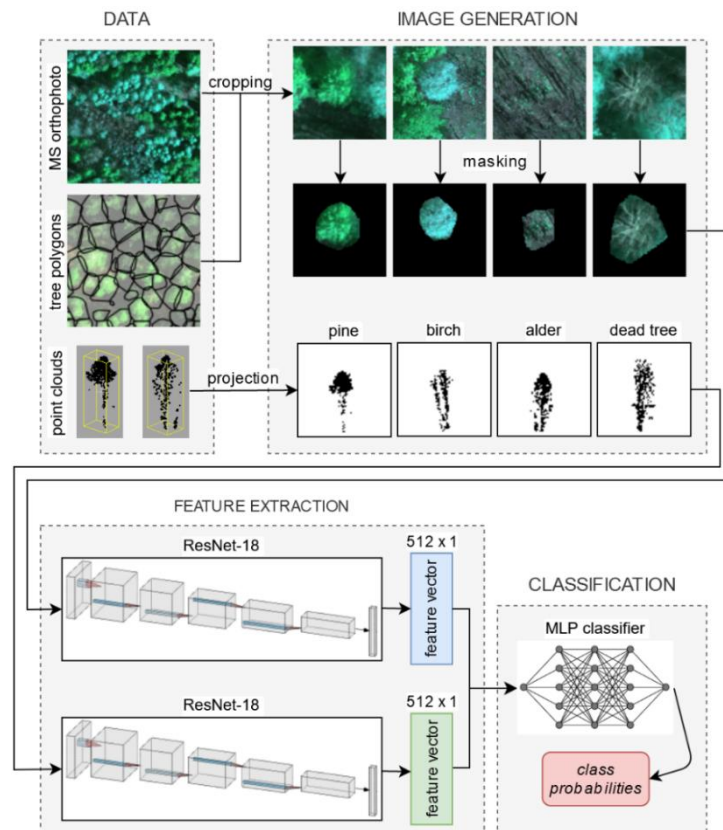
Figure 43: Outline of the proposed method, Silvi-Net$_{single}$.

Initially, 2D representations of the single trees were created in an image generation process. For each tree, an MS image patch was cropped to place the tree crown in the image centre. In this step, we utilized the polygon outlines generated by the lidar-based tree segmentation. To maintain the relative dimensions of the crowns, image patches with the same quadratic size were produced. Thus, we ensured that even the largest tree crowns were included in the images in their entirety. Outlines of the projected 2D tree polygons were used to mask pixels not corresponding to the actual tree. In addition to the MS images, we rendered multiple side-view images from the segmented 3D lidar point clouds of single trees, representing the trees' silhouettes. Optionally, these images were enriched with laser EC. Basically, we created two types of image sets – one with 12 individual images per tree, and one with an image collage comprised of all 12 side-view images per tree. In the following sections, the approaches utilizing these datasets are referred to as Silvi-Net$_{single}$, and Silvi-Net$_{collages}$ respectively. After the image generation process, features were automatically extracted using two independently trained ResNet-18 models, optimized for both the MS and side-view images. Here, we applied the idea of transfer learning and pretrained weights. To visualize the model's decisions, we produced CAM images and superimposed them on the input images. Overall, we generated 512 features per side-view image or image collage, and additional 512 features from each MS image. Next, the feature vectors were fused and fed into a standard multi-layer perceptron (MLP) that was trained to estimate class probabilities for each sample. For Silvi-Net$_{single}$, the classification led to 12 predicted labels per tree. Therefore, we introduced Silvi-Net$_{majVot}$, an additional evaluation strategy applying majority voting to these 12 predictions. The idea was to outvote individual, falsely classified side-view images and to obtain one label per tree. In the following sections, all of the steps of our approach are described in greater detail.

5.3.2. Generation of image patches

First, we present the methodology for image generation from MS orthomosaics. Because CNN-based image classification supported by transfer learning typically utilizes three-channel imagery, we reduced the five-channel images in the ChEZ area. More specifically, we transformed the B, G, and R channels into a single grey scale channel by calculating the mean value of these three channels for each pixel. Next, we added the RE and NIR channels, resulting in a three-channel image. Because of the difference in sensors, an alternative procedure was conducted for the BFNP data. In this study area, we removed the blue channel from the raw imagery and utilized the resulting CIR images. For both study areas, we normalized the three image channels inde-

pendently to values between 0 and 1. For each tree segment, the corresponding polygon was projected onto the orthomosaic. Then, a cropped image patch was produced covering a predefined quadratic region around the polygon centre. The image size resulted from the maximum crown dimension (ChEZ: 10 m x 10 m, BFNP: 12 m x 12 m) and the pixel size of the orthomosaics (ChEZ: 10 cm, BFNP: 20 cm). Thus, all tree crowns fit within the image dimensions. Ultimately, this process led to images sized 100 x 100 pixels for ChEZ and 60 x 60 pixels for BFNP. Optionally, pixels outside the tree polygon were masked out and set equal to 0. Thus, *MS_unmasked* and *MS* datasets were prepared on the basis of the ChEZ dataset (Figure 44) and the BFNP dataset (Figure 45). In total, 668 MS patches were generated in the ChEZ area, and 2,721 in the BFNP area – each with a masked and an unmasked version.
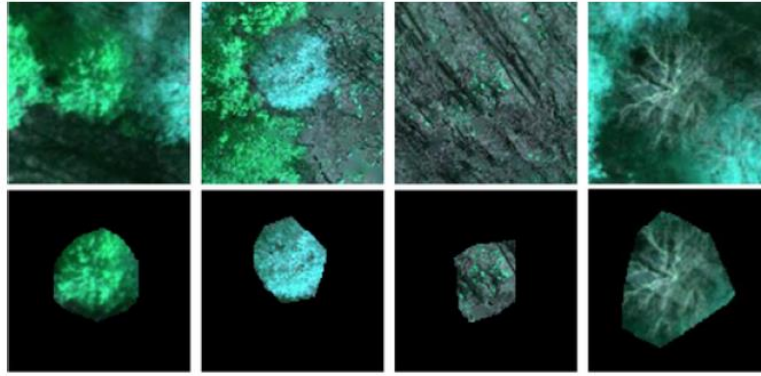
Figure 44: MS_unmasked images (first row) and MS images (second row) generated from MS orthomosaics in the ChEZ study area; false-colour images (RGB, RE, NIR). Image size corresponds to 10 m × 10 m.
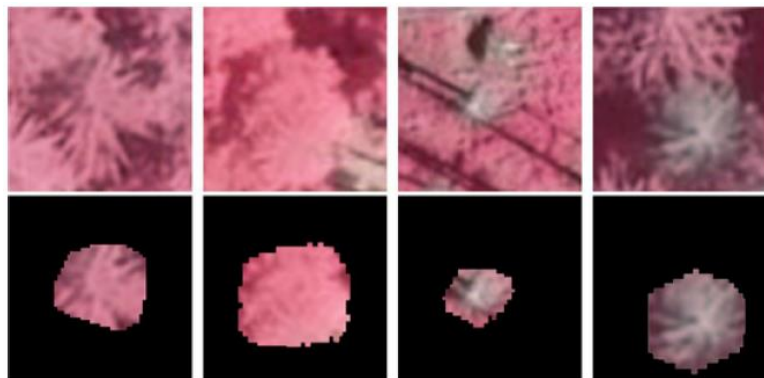
Figure 45: MS_unmasked images (first row) and MS images (second row) generated from MS orthomosaics in the BFNP study area; CIR images (G, R, NIR). Image size corresponds to 12 m × 12 m.

Additionally, we prepared two different types of images from the 3D lidar point clouds – one with 12 individual images per tree and one with an image collage comprised of all 12 side-view images per tree. In a first step, the point clouds were rotated in constant steps around the z axis to simulate multi-view positions. After visual interpretation, we decided to set the rotation angle to multiples of 30°, leading to multi-view image stacks of 12 images per tree. This was deemed an acceptable balance between information loss and redundancy. Next, we rendered binary silhouette-like images for both study areas (see Figure 46 a-d and Figure 47 a-d) by projecting the 3D data onto a virtual vertical raster. The image resolution was set to 10 cm per pixel. Because the images should completely cover even the largest trees, the image size – 260 x 260 px in the ChEZ and 500 x 500 px in the BFNP – was determined by the maximum tree height in the corresponding study area. Note that the image size in the BFNP was much larger than the required input size of ResNet-18. The average tree height was 16.5 m (std = 1.2 m) in the ChEZ and 28.1 m (std = 6.1 m) in the BFNP, respectively. As a consequence, 90% of all trees covered at least half of the image height in the ChEZ. With 70%, this ratio was clearly lower in the BFNP. In case of working with data from forests with an even larger range of tree sizes, we assume that the image size should be calculated in a different way. Otherwise, decreasing results are likely to appear because of significant loss of detail for the average and smaller trees. Furthermore, we wanted to analyse the impact of EC on our classification method. Therefore, we included EW values in the ChEZ dataset and intensity (INT) values in the BFNP dataset. Incorporating the normalized EW and INT values, we also generated 8-bit grayscale images. These two image datasets are referred to as GEOM (binary images) and GEOM_EC (grayscale images), respectively. Overall, the preprocessing of samples for Silvi-Net$_{single}$ led to 8,016 GEOM and GEOM_EC images each in the ChEZ area (training: 4,464; validation: 1,104; test: 2,448) and 32,652 GEOM and GEOM_EC images each in the BFNP (training: 16,560, validation: 7,152, test: 8,940). Subsequently, we rendered image collages utilized in the approach Silvi-Net$_{collages}$, including all 12 views per tree in one image. Therefore, assuming that only the middle third of the quadratic side-view images contains useful information, we cut out these essential image parts. Next, we randomly arranged them as matrices comprising two rows with six images each (see Figure 46 e-h and Figure 47 e-h). Thus, unlike the previously created single-view images, the number of samples was equivalent to the number of tree objects.
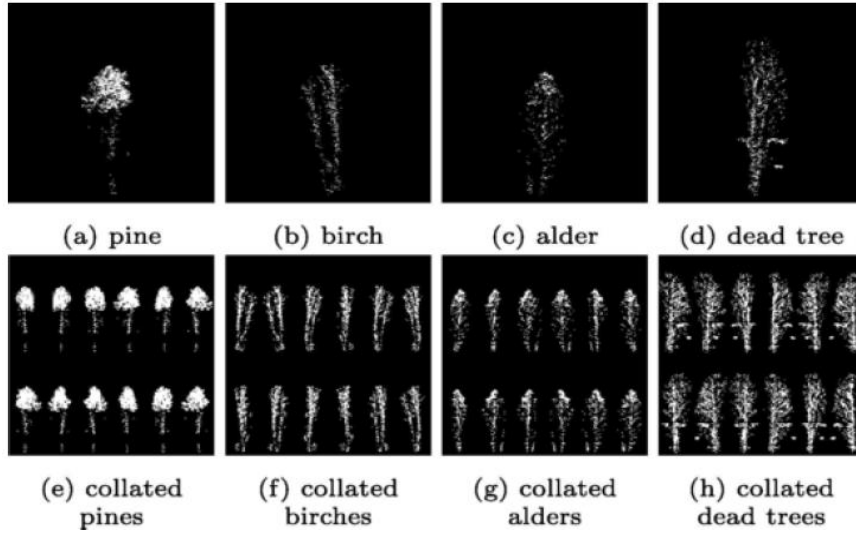
Figure 46: GEOM images generated from UAV-based lidar data in the ChEZ study area. First row (a-d): single images, image size corresponds to 26 m × 26 m. Second row: collated multi-view images. Image size corresponds to 52 m × 52 m.
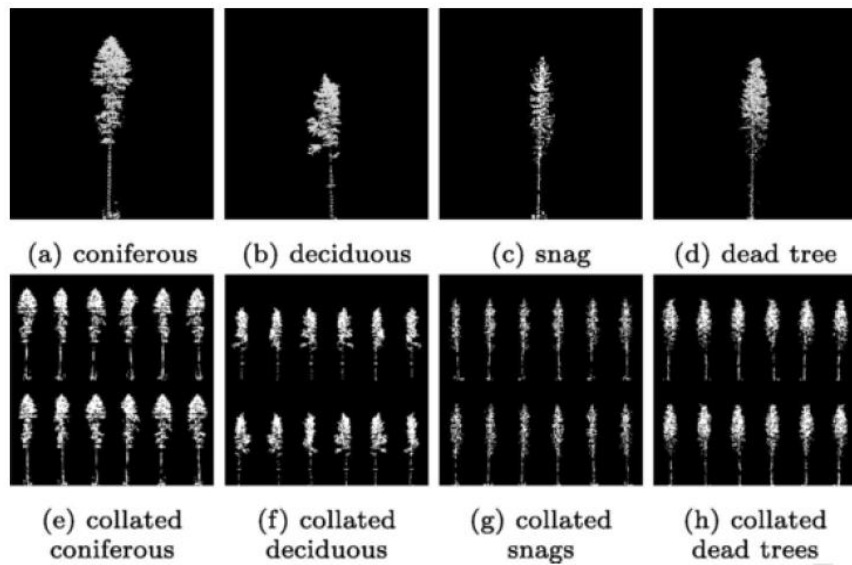
Figure 47: GEOM images generated from ALS data in study area BFNP; First row (a-d): single images, image size is corresponding to 50 m × 50 m; Second row: collated multi-view images, image size is corresponding to 100 m × 100 m.

5.3.3. CNN-based feature extraction

In our approach, automatic extraction of features was performed using a standard CNN. Our decision was motivated by the fact that CNNs are well-established neural networks for image-based deep supervised learning that are capable of achieving excellent results in the fields of pattern recognition and ML (Schmidhuber, 2015). Moreover, CNNs are especially designed to process sensor data represented as multiple 2D arrays. By considering local and global stationary properties, CNNs have achieved state-of-the-art results in popular image classification tasks, such as the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) (Russakovsky et al., 2015). In general, CNNs are DNNs consisting of numerous stacked layers. The first part, also known as the feature extractor, is mainly comprised of convolutional blocks – sequences of convolutional layers, activation layers, and pooling layers. Thereby, convolutional layers utilize filter kernels to extract low-level image features. Moreover, the kernel depth is equal to the number of image channels. The output of a convolutional layer is a feature map with one channel per filter kernel. Activation layers, such as the rectified linear unit (ReLU), account for non-linear effects. Practically speaking, ReLU sets negative values to the value 0 and reduces the problem of vanishing gradients – an effect that occurs with DNNs. Additionally, ReLU layers are computationally inexpensive and enable faster model convergence. Pooling layers essentially subsample the feature maps, using common methods such as average pooling and maximum pooling. The second part of a CNN is the actual classifier. Here, the final output feature maps are flattened into a one-dimensional vector, followed by fully connected classification layers (LeCun et al., 2015). Overall, CNNs include a huge set of model parameters – weights and biases – that need to be estimated. To reduce model overfitting, regularization techniques such as dropout and batch normalization are often included in typical CNN architectures. By adding dropout layers, co-adaptation of neurons can be prevented. Moreover, this technique approximates the idea of ensemble models and allows a higher LR. However, it also usually leads to slower model training. Additionally, batch normalization layers can help improve model stability and quality (Ioffe and Szegedy, 2015). Practically speaking, these layers apply channel-wise normalization of the feature maps and result in faster model convergence.

5.3.3.1. CNN architecture

In our classification pipeline, we utilized two standard ResNet-18 models (He et al., 2015) implemented with the PyTorch framework, version 1.1.0 (Paszke et al., 2019), which is an optimized tensor library for DL using GPUs and CPUs. With their proposed idea of residual blocks, the developers of ResNet successfully minimized the problem of vanishing

gradients. At this time, the problem was that the training accuracy of multi-layer CNNs dropped as the number of layers increased. Therefore, the authors proposed to use a reference to the previous layer to compute the output at a given layer. As a result of these skip connections (also termed shortcuts), the training of much deeper CNNs was facilitated. Moreover, these deep residual networks can achieve improved accuracy due to considerably increased depth. In 2015, He et al. (2015) won the ILSVRC using ResNet ensembles with a depth of up to 152 layers. Our decision to use ResNet-18 was motivated by preliminary studies testing different CNN architectures included in the "models" subpackage of the "torchvision 0.3.0" module in PyTorch. Here, both VGG-16 (Simonyan and Zisserman, 2015) and Densenet-121 (Huang et al., 2017) performed significantly worse than ResNet-18 when initialized using weights pretrained on the ImageNet dataset (Deng et al., 2009). Adopting the deeper version, ResNet-50, did not improve the results. Presumably, our dataset was too small to retrain all 23.6 million ResNet-50 parameters in an effective way. In contrast, we were able to robustly retrain all 11.2 million ResNet-18 parameters and optimize the network for our task. In this way, we achieved a suitable trade-off between network depth and dataset size. Figure 48 shows the architecture of ResNet-18 in a simplified way, including dimensions of tensors and filters and final feature vector. The feature extractor of ResNet-18 consists of four residual blocks. Each block is comprised of a stack of two basic blocks of 2–3 convolutional layers, followed by batch normalization and ReLU layers. Finally, an average pooling layer extracts 512 features per 224 x 224 x 3 input image.
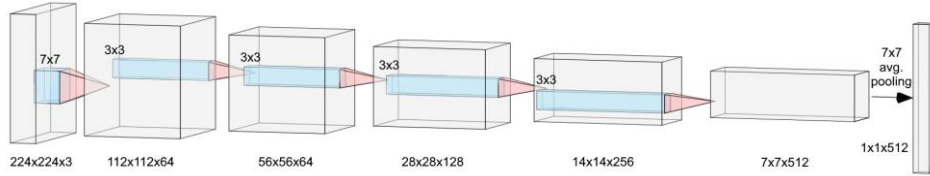
Figure 48: Simplified ResNet-18 architecture, created with a neural network drawing tool (LeNail, 2019).

5.3.3.2. CNN training

We utilized two separate ResNet-18 models optimized for the classification of *GEOM*/*GEOM_EC* images and *MS* images, respectively. At the beginning, all images were loaded and resampled to the required image size of 224 x 224 pixels. Next, the images in the range [0, 255] were converted to floating tensors in the range [0.0, 1.0]. Then, the three channels of these image tensors were standardized separately using the mean (0.485, 0.456, 0.406) and std (0.229, 0.224, 0.225) values of ImageNet. For each channel, the mean of the data was 0 and the

std was 1. Next, data augmentation was performed on the training data. This method to artificially increase the number of training samples is helpful to avoid overfitting and usually results in better generalization properties of the trained model (Goodfellow et al., 2016). In our approach, we applied a combination of random affine transformation and random horizontal flip to both training and validation data. Moreover, we randomly flipped the MS images vertically. Note that this transformation was not performed for the lidar-based side-view images to maintain the vertical orientation of trees. The affine transformation of image coordinates *x* and *y* into new image coordinates *x'* and *y'* can be described as a sequence of rotation, shearing, scaling, and translation (Eq. 12):

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \end{bmatrix} + \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \cdot \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \qquad [12]$$

where α is the rotation angle, *s* is the shearing parameter, s_x and s_y are the scale parameters for both coordinate axes, and t_x and t_y are the components of the 2D translation vector. In or approach, we allowed a relative maximum image translation of ±10% in horizontal and vertical directions and scaling parameters s_x and s_y in the interval of [0.80, 1.25]. The shearing parameter *s* was set to 0. For the rotation angle α, the range defining the maximum random value was set depending on the image type: ±20° for *GEOM* and *GEOM_EC* images and ±180° for *MS* images.

For model training, we utilized the concept of transfer learning. Numerous researchers have shown that DL-based models are able to learn features that – to a certain extent – transfer well across datasets (Hu et al., 2015; Shin et al., 2016). Instead of starting with random parameter values, models can be initialized with weights optimized for extensive and standardized databases such as ImageNet. Although the ImageNet dataset includes 1,000 object classes and is clearly different from our tree data, we assumed that it would be adaptable for the task of tree species classification. Besides relatively quick convergence, effective fine tuning of DNNs typically requires much less samples compared to training from scratch (Ng et al., 2015). Therefore, we initialized the ResNet-18 models by utilizing pretrained ImageNet weights. Moreover, we set the maximum number of epochs to 100 and implemented an early stopping criterion defined as 10 epochs with no improvement in validation loss. More precisely, the criterion for model evaluation was based on a cross-entropy loss function. For each image batch (*batch_size* = 32), we calculated the loss with shared class weights (Eq. 13) and averaged all losses per epoch.

$$loss(x, class) = -log \frac{\exp(x[class])}{\Sigma_j \exp(x[j])} \qquad [13]$$

The model hyperparameters were optimized using an Adam optimizer (Kingma and Ba, 2015). Here, we relied on the default values of PyTorch implementation. Every seven epochs, an exponential LR scheduler decayed the initial LR of 0.001 by a factor of gamma = 0.1 (Eq. 14):

$$LR_{i+1} = LR_i \cdot (1 - \gamma) \qquad [14]$$

Overall, ResNet-18 is comprised of approximately 11.2 million trainable parameters. In our classification pipeline, the crucial factor for the generalization of well-performing models was a systematic recalculation of the model parameters for each dataset. The procedure was as follows: First, we set all parameters of the feature extractor to be invariable and only retrained the 16,548 parameters of the fully connected layers. Second, we iteratively "unfreezed" the trainable parameters of the four residual blocks. Starting with the deepest block (8.4 million parameters), the number of trainable parameters increased to 10.5 million, 11.0 million, and finally 11.2 million. Third, for each dataset, the model showing the lowest cross-entropy loss on the validation dataset was stored. Finally, this best performing model was set to evaluation mode and was subsequently used to extract 512 features per image. Thus, the optimized ResNet-18 models were practically utilized as automatic feature extractors for the training, validation, and test datasets. In our implementation, we registered a so-called "forward hook" to enable feature extraction from the average pooling layer. In PyTorch, this step was performed utilizing the *register_forward_hook* function in the "*nn*" (neural network) package.

5.3.4. MLP-based tree classification

In our classification pipeline, optimized ResNet-18 models were used as automatic feature extractors. To perform tree species classification utilizing 2D representations rendered from both airborne lidar and MS data, we combined the feature sets generated from the side-view images (*GEOM* and *GEOM_EC*) and *MS* images. Next, we inputted the fused feature vectors comprising 1,024 features and the corresponding class labels to a standard MLP classifier. An MLP is a non-parametric neural network classifier with shallow structures containing only a few feature representation levels. Typically, an MLP is composed of interconnected nodes in multiple layers (namely input, hidden, and output layers), with each layer fully connected to both the preceding and succeeding layers (Del Frate et al., 2007). Moreover, the outputs of each node are weighted units followed by a nonlinear activation function (Pacifici et al., 2009). In summary, in a feed-forward manner, an MLP maps a set of input features onto a set of labels (Atkinson and Tatnall, 1997). In our method, we utilized the "MLPClassifier" class from the

"sklearn" module "neural_network" (Pedregosa et al., 2011). More specifically, we implemented an MLP with three hidden layers composed of seven neurons each, and set the hyperparameters to the default values. The particular types of feature vectors generated from the MS images and side-view images, were weighted 50% each. The MLP classifier was trained using the combined feature set calculated from the training and validation datasets. Finally, we evaluated the MLP on the independent test datasets and derived confusion matrices and standard metrics from the TP, TN, FP, and FN values. We calculated the OA (Eq. 15), precision (Eq. 16, recall (Eq. 17), and F_1 score (Eq. 18).

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

$$precision = \frac{TP}{TP + FP} \tag{16}$$

$$recall = \frac{TP}{TP + FN} \tag{17}$$

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{18}$$

5.3.5. Baseline method (PointNet++)

For the classification of 3D objects such as trees, it is unknown whether working with rendered multiple 2D images or raw 3D point clouds is favourable. Thus, we compared our new CNN-based method to the approach presented in Briechle et al. (2020a), using a PyTorch implementation of PointNet++ (Wijmans, 2018) for object classification. In the following sections, the most important steps of this baseline method will be explained, including data preparation, network training, and validation.

5.3.5.1. Preparation of dataset

Before training the network, the 3D dataset had to be prepared appropriately. Typically, PointNet++ can only manage a constant number of 3D points per sample. Therefore, we applied a combined sampling approach to achieve balance between upsampling and downsampling of data, resulting in 1,024 points per tree. Moreover, as generally proposed when working with DNNs, data was standardized. Next, we calculated the surface normals for all 3D points using the "estimate_normals" function from the open source library Open3D (Zhou et al., 2018). Then, handcrafted MS features were generated and integrated into the dataset. Here, we relied on a selection of statistical MS features computed with different VIs. Depending on the available spectral channels, the number of VIs differed between the study areas. In the BFNP, we derived the NDVI (Rouse et al., 1973) from the CIR images. In the

ChEZ area, the five-channel orthomosaics also enabled the calculation of the RENDVI (Gitelson and Merzlyak, 1994), the REDVI (Briechle et al., 2020a), the MRESR (Datt, 1999), and the MCARI (Daughtry et al., 2000). Projections of the tree polygons were utilized to filter VI pixels belonging to a single tree. Subsequently, we computed 12 object-based statistical features from these pixels for each VI – the maximum value (max), minimum value (min), range (max-min), mean value, std, mode[21], skewness[22], kurtosis[23], as well as the 25th ("1st quartile"), 50th ("median"), 75th ("3rd quartile"), and 90th percentile (perc).

To make the classifier more robust and to avoid overfitting, the feature space was reduced to the five most important MS features. Here, we relied on an RF-based feature selection technique, which has been recommended in the literature (Ma et al., 2017; Gregorutti et al., 2017), to generate a ranking of all input features according to their relative importance on the prediction. Then, for each study area, the five most decisive features were selected: *NDVI_skewness*, *MRESR_perc90*, *NDVI_perc90*, *RENDVI_mode*, and *MRESR_mode* in the ChEZ, and *NDVI_perc25*, *NDVI_skewness*, *NDVI_range*, *NDVI_mean*, and *NDVI_min* in the BFNP. Afterwards, the values of these top five MS features were standardized and assigned to each 3D point of each object, resulting in additional point attributes. Overall, the final dataset comprised 12 attributes per 3D point: the 3D coordinates and surface normals, one EC value, and five handcrafted MS features.

5.3.5.2. Training and validation

To successfully adapt PointNet++ for the task of tree species classification, we optimized the most decisive hyperparameters of the neural network (Table 25). Therefore, we used a combination of manual search and automated grid search. During model training, we also performed data augmentation to avoid model overfitting and to build generalizable models. Because the trained final model should be robust against object variation, we implemented random transformations of the 3D objects, including scaling in the range [0.80, 1.25], rotation around the vertical axis with an angle of the range [0, 2*pi], jittering with Gaussian noise (±0.05 m), and 3D translation of the entire point cloud by ±0.1 m. Moreover, setting the random input dropout parameter MAX_DROPOUT to 50% increased the robustness against varying point densities and occluded object parts. Finally, we evaluated the model showing the lowest validation loss on the test dataset and generated classification metrics (OA, precision, recall, F_1 score).

---

[21] Most frequent value.
[22] Measure of the asymmetry of the probability distribution.
[23] Measure of the tailedness of the probability distribution.

Table 25: Hyperparameter settings for PointNet++.

Hyperparameter	Value	Declaration
NUM_CLASSES	4	Number of object categories.
NUM_POINT	1,024	Number of points per sample.
BATCH_SIZE	8	Number of samples per batch.
MAX_EPOCH	100	Maximum number of training epochs.
MAX_DROPOUT	0.5	Maximum dropout rate.
OPTIMIZER	Adam	Optimization algorithm.
BASE_LR	1e-3	Initial LR.
LR_DECAY	0.7	Initial learning decay.
BN_MOMENTUM	0.5	Initial momentum for batch normalization.
BNM_DECAY	0.5	Decay of batch normalization momentum.
WEIGHT_DECAY	1e-4	L2 regularization coefficient.

5.4. Experiments

For both study areas, we conducted experiments based on different input datasets. Initially, we utilized sets of binary side-view images only (*GEOM*). To analyse the impact of laser EC on classification results, we trained Silvi-Net with *GEOM_EC* images. Subsequently, we classified single tree objects using only masked (*MS*) and unmasked (*MS_unmasked*) MS images. Finally, we fused automatically extracted features from both lidar-based image sets (*GEOM*, respectively *GEOM_EC*) and *MS* images for classification (*GEOM+MS*, respectively *GEOM_EC+MS*). In all experiments, we explored three different evaluation strategies – Silvi-Net$_{single}$, Silvi-Net$_{collages}$, and Silvi-Net$_{majVot}$. Furthermore, we integrated CAM technique into our pipeline to better understand the model's decisions on new independent data. Demystifying CNNs' status as "black box" systems, CAM can help to highlight class-specific, distinctive image regions (Zhou et al., 2016). To generate CAM images, the predicted class score in the range [0, 1] was mapped back to the final convolutional layer. In detail, CAM can be described as the dot product of the extracted weights from the final layer and the feature map. In our ResNet-based approach, the resulting CAM images sized 7 x 7 px were bilinearly upsampled and superimposed on the input images sized 224 x 224 px. In the following sections, classification results are presented for Silvi-Net and compared to those of the baseline, PointNet++.

5.4.1. Masking MS data

Initially, we investigated whether masking MS image patches would improve classification results. Therefore, classification was performed with both MS_unmasked images and MS images. In general, we observed a positive impact when masking MS image patches utilizing single tree polygons for both study areas. The gain in OA was 1.4% in the ChEZ and 2.6% in the BFNP (Table 26). These relative values represent

3 of 204 test samples (ChEZ), respectively 19 of 745 test samples (BFNP). Furthermore, MS images yielded F_1 scores between 0.90 and 0.99 in the ChEZ. Here, masking improved the results for pine and birch. In particular, pine trees were classified almost perfectly (F_1 score = 0.99). In our second dataset (BFNP), masking pixels located in the surrounding area boosted the classification of snags and dead trees. Nevertheless, the F_1 scores for snags (0.78) and dead trees (0.76) were still relatively low. Remarkably, classification based on CIR imagery led to reasonable accuracy for the coniferous (F_1 score = 0.90) and deciduous (F_1 score = 0.92) classes in this study area. By superimposing MS_unmasked images with CAM images, it can be observed that in most cases the neural network automatically identified the crucial tree crowns in the image centre (Figure 49). However, in some cases, neighbouring tree pixels in unmasked images affected the results. As a consequence, classification errors were produced because the CNN occasionally focused on nearby trees from different classes (Figure 50). Thus, we relied on masked MS images in the following experiments.

Table 26: Silvi-Net results using only MS image patches (test dataset).

Study area	Image type	OA	F_1 scores per class
ChEZ	*MS_unmasked*	0.922	0.95/0.89/0.91/0.93
	MS	**0.936**	0.99/0.92/0.90/0.93
BFNP	*MS_unmasked*	0.830	0.90/0.95/0.65/0.71
	MS	**0.856**	0.90/0.92/0.78/0.76

(a) pine   (b) birch   (c) alder   (d) dead tree

(e) coniferous   (f) deciduous   (g) snag   (h) dead tree
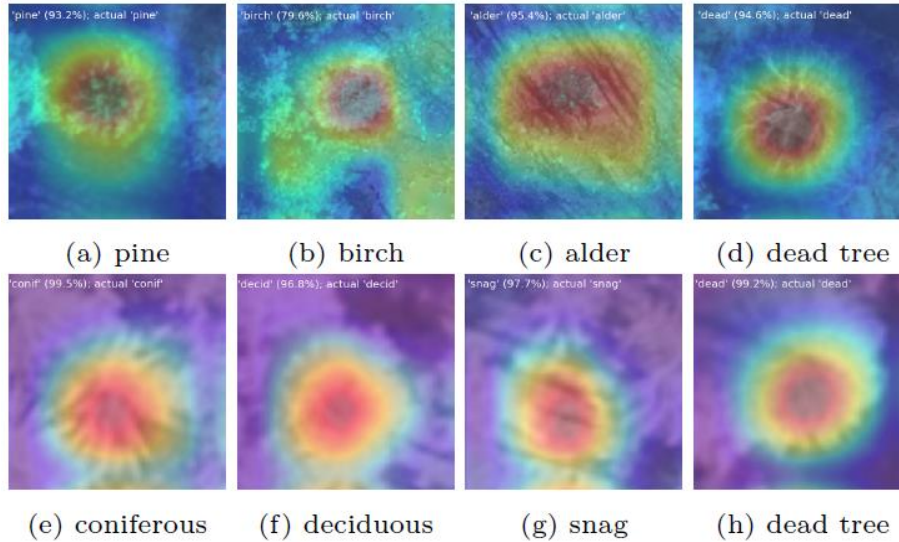
Figure 49: Examples for correct classification of MS_unmasked images in the ChEZ (a-d) and BFNP (e-h); CAM overlay.

(a) actual: coniferous

(b) actual: deciduous

(c) actual: dead tree

(d) predicted: dead tree

(e) predicted: coniferous

(f) predicted: coniferous
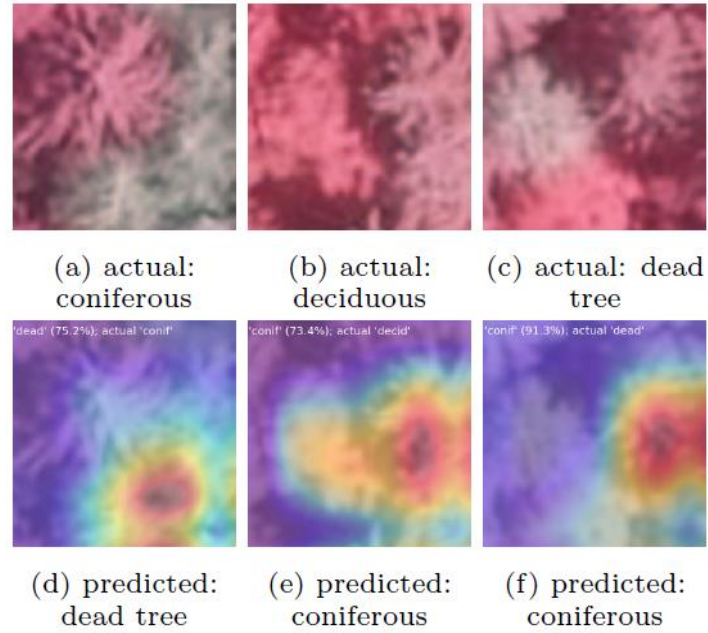
Figure 50: Examples for incorrect classification of MS_unmasked images in the BFNP; CAM overlay.

5.4.2. Results for ChEZ

The task of classifying pine, birch, alder, and dead trees in the ChEZ was generally performed best by Silvi-Net$_{majVot}$ (Table 27). Moreover, this approach outperformed baseline method PointNet++ (OA = 84.8%) by 11.3%, reaching an OA of 96.1%. Compared to the results based on only *MS* images (OA = 93.6%; Table 26), incorporating geometry information and EC improved the results by 2.5% in this study area. Note that the discrepancy between results for validation data and test data was less than 3% in all experiments. This demonstrates the high generalization capacity of Silvi-Net.

Table 27: Results for Silvi-Net and PointNet++ on ChEZ test dataset. For each feature subset, the highest OA is displayed in bold letters, and the highest F_1 scores per class are underlined.

Method	Dataset	OA	F_1 scores per class
Silvi-Net$_{single}$	*GEOM*	0.732	0.81/0.64/0.71/0.77
	GEOM_EC	0.765	0.89/0.69/0.67/0.78
	GEOM + MS	0.912	0.94/0.89/0.90/0.92
	GEOM_EC + MS	0.937	0.98/0.92/0.92/0.93
Silvi-Net$_{collages}$	*GEOM*	0.721	0.82/0.60/0.71/0.74
	GEOM_EC	0.716	0.82/0.62/0.71/0.72
	GEOM + MS	**0.951**	0.96/0.91/<u>0.97</u>/0.96
	GEOM_EC + MS	0.917	0.98/0.87/0.90/0.92
Silvi-Net$_{majVot}$	*GEOM*	**0.775**	<u>0.87</u>/<u>0.71</u>/0.72/<u>0.80</u>
	GEOM_EC	**0.804**	<u>0.93</u>/<u>0.74</u>/0.71/<u>0.82</u>
	GEOM + MS	**0.951**	<u>0.97</u>/<u>0.92</u>/0.95/<u>0.96</u>
	GEOM_EC + MS	**0.961**	<u>0.99</u>/<u>0.93</u>/<u>0.95</u>/<u>0.97</u>
PointNet++	*GEOM*	0.755	0.83/0.67/<u>0.73</u>/0.79
	GEOM_EC	0.779	0.92/0.69/<u>0.74</u>/0.78
	GEOM + MS	0.821	0.81/0.78/0.81/0.88
	GEOM_EC + MS	0.848	0.89/0.80/0.83/0.87

Now, we want to focus on more detailed results regarding single data subsets. When using only side-view images of the point clouds (*GEOM* images), the classification results (OA = 77.5%) were 2.0% better than PointNet++ when classifying raw 3D point clouds of the single trees (OA = 75.5%). Although geometric information was partially reduced during image generation, the F_1 scores were higher for all classes except alders. Classification based on *GEOM_EC* images (OA = 80.4%) was superior to *GEOM* images. By incorporating EC, the gain in OA was 2.9%. Specifically, pine as the only coniferous tree in the dataset benefited most. For both experiments based on side-view imagery generated from 3D point clouds, confusion was biggest between birch and alder (Figure 51a and Figure 51b). When combining automatically extracted features from *GEOM* images and *MS* images, the classification

results clearly increased (OA = 95.1%). By integrating MS information, the F_1 score raised by more than 0.20 for the two deciduous species alder and birch. Consequentially, confusion between these two classes was almost completely resolved (Figure 51c). Moreover, Silvi-Net$_{majVot}$ performed 13.0% better than PointNet++ based on raw 3D point clouds enriched with the top five handcrafted MS features (OA = 82.1%). Furthermore, it is noteworthy that for the *GEOM+MS* experiment, Silvi-Net$_{collages}$ was equal to Silvi-Net$_{majVot}$. Fusing *GEOM_EC* images and *MS* images yielded the best results. Here, the OA for Silvi-Net$_{majVot}$ reached 96.1%, with F_1 scores ranging between 0.93 (birch) and 0.99 (pine), and a minor remaining confusion between birch and alder (Figure 51d).

(a) *GEOM*

(b) *GEOM_EC*

(c) *GEOM+MS*

(d) *GEOM_EC+MS*

Figure 51: Confusion matrices for Silvi-Net$_{majVot}$ (ChEZ test data). Subfigures a-d show results for different feature sets.

5.4.3. Results for BFNP

When studying BFNP, general results for the classification of single trees into coniferous, deciduous, snag, and dead tree were partly different from those obtained in the ChEZ. Particularly, in two of four experiments, Silvi-Net$_{majVot}$ was slightly inferior to the baseline Point-Net++. However, when fusing side-view images and MS imagery, Silvi-Net$_{majVot}$ was the method of choice (OA = 91.5%; Table 28), exceeding the baseline method (OA = 89.3%) by 2.2%. Furthermore, our experiments show that embedding *GEOM_EC* images helped improve results by 5.9%, in contrast to an OA of 85.6% based on only *MS* images (Table 26). Moreover, the difference between validation data and test data was again in the range of a few percentage points, showing that Silvi-Net generalized well.

Table 28: Results for Silvi-Net and PointNet++ on BFNP test dataset. For each feature subset, the highest OA is displayed in bold letters, and the highest F_1 scores per class are underlined.

Method	Dataset	OA	F_1 scores per class
Silvi-Net$_{single}$	*GEOM*	0.811	0.80/0.95/0.86/0.60
	GEOM_EC	0.808	0.80/0.96/0.84/0.59
	GEOM + MS	**0.911**	0.94/<u>0.99</u>/0.86/<u>0.80</u>
	GEOM_EC + MS	0.899	0.93/0.98/0.86/0.76
Silvi-Net$_{collages}$	*GEOM*	0.744	0.68/0.96/0.84/0.51
	GEOM_EC	0.800	0.81/0.96/0.81/0.59
	GEOM + MS	0.909	<u>0.95</u>/<u>0.99</u>/0.84/0.79
	GEOM_EC + MS	0.905	0.95/<u>0.99</u>/0.82/0.79
Silvi-Net$_{majVot}$	*GEOM*	0.847	<u>0.86</u>/0.96/0.85/0.67
	GEOM_EC	0.835	0.85/<u>0.97</u>/0.84/0.60
	GEOM + MS	**0.911**	0.94/<u>0.99</u>/0.85/<u>0.80</u>
	GEOM_EC + MS	**0.915**	<u>0.96</u>/<u>0.99</u>/0.86/0.79
PointNet++	*GEOM*	**0.857**	0.85/0.97/0.88/0.72
	GEOM_EC	**0.867**	0.88/0.95/0.87/0.76
	GEOM + MS	0.882	0.90/0.97/<u>0.89</u>/0.77
	GEOM_EC + MS	0.893	0.92/0.97/<u>0.88</u>/<u>0.80</u>

For a more detailed analysis, we want to draw attention to the experiments examining the impact of single data subsets. Using only geometry data, PointNet++ (OA = 85.7%) performed slightly better than our ResNet-based approach and majority voting (OA = 84.7%). Here, the results generated by Silvi-Net$_{majVot}$ showed a considerable confusion between coniferous and dead trees (Figure 52a). In detail, the low F_1 score for dead trees (0.67) was mainly due to the fact that 16.6% (24/145) of dead trees were classified as coniferous, and 13.5% (35/259) vice versa. Nevertheless, deciduous trees were classified almost perfectly (F_1 score = 0.96). Moreover, Silvi-Net$_{majVot}$ was clearly

superior to the approaches based on single or collated imagery. When utilizing GEOM_EC images, we observed that the incorporation of EC was not advantageous for tree classification in this study area, especially since the F_1 score of dead trees dropped by 0.07 (Figure 52b). All other F_1 scores remained almost unchanged (±0.01). Surprisingly, the baseline method using PointNet++ (OA = 86.7%) benefited from EC by 1.0%. When combining geometry data and masked MS data (GEOM+MS), Silvi-Net$_{majVot}$ (OA = 91.1%) performed 2.9% better than the baseline method (OA = 88.2%). Here, incorporating MS images clearly enhanced the results by 6.4% and especially improved the confusion between coniferous trees and dead trees (Figure 52c). Note that the classification of snags was not improved by MS data. Using EC (GEOM_EC+MS) slightly improved Silvi-Net$_{majVot}$ (0.4%), reaching the best result in this study area (OA = 91.5%). However, we observed an unsolved moderate confusion between dead trees and snags (Figure 52d), with 13.1% (19/145) of dead trees being classified as snags and 12.9% (18/139) vice versa.
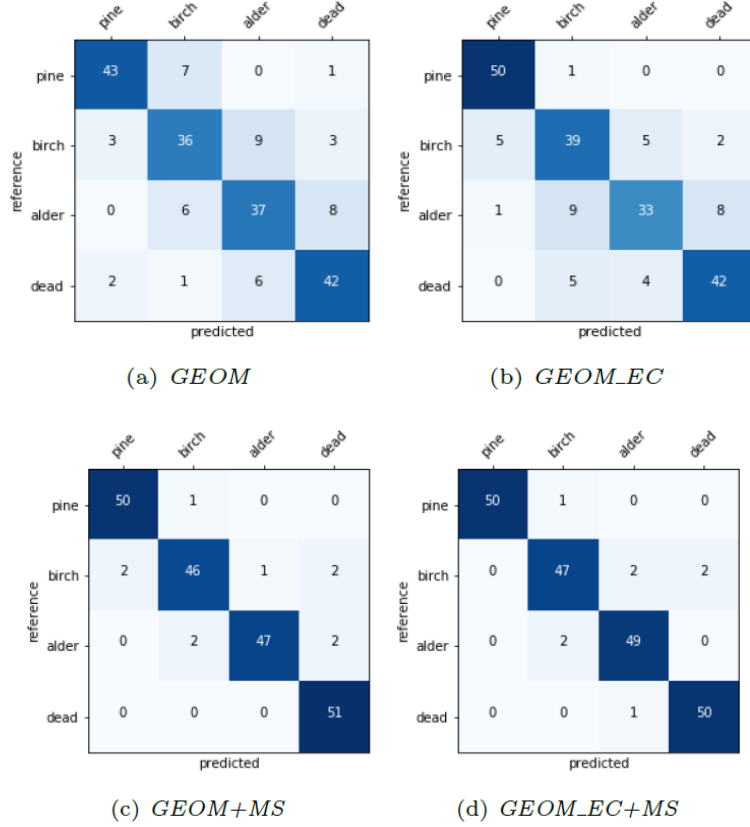

(a) *GEOM*

(b) *GEOM_EC*

(c) *GEOM+MS*

(d) *GEOM_EC+MS*
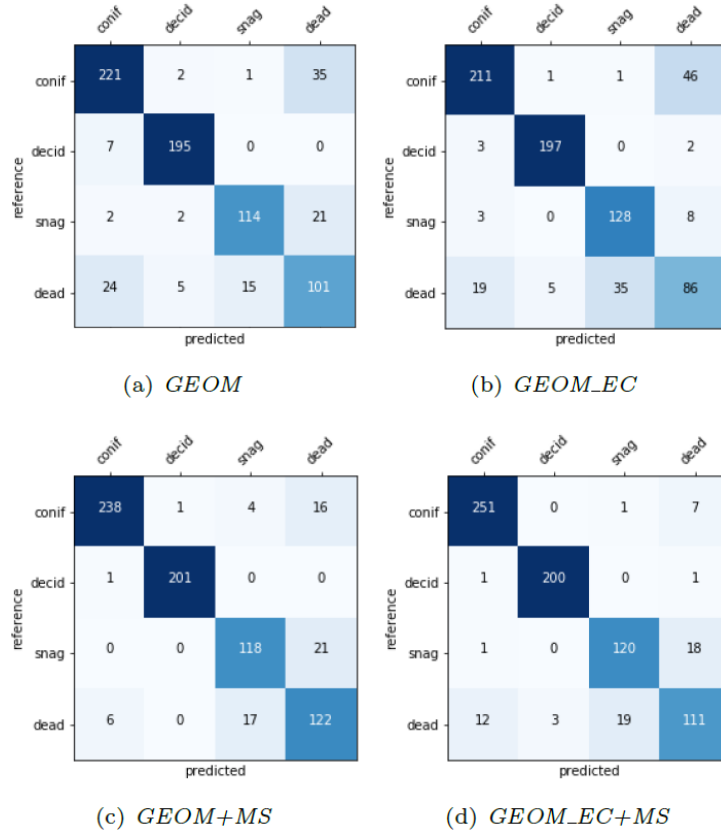
Figure 52: Confusion matrices for Silvi-Net$_{majVot}$ (BFNP test data). Subfigures show results for different feature sets.

5.5. Discussion

5.5.1. Main results

Overall, the newly introduced methodology for the classification of single tree species and standing dead trees was successfully applied in two study areas. In general, we achieved an OA of 96.1% using the ChEZ dataset (Figure 53) and 91.5% using the BFNP dataset (Figure 54). Note that the datasets vary in terms of forest types and sensor models, as well as geometric and spectral resolution. Therefore, the superior results in the ChEZ are mostly due to the fact that both the ground resolution and the number of spectral channels in MS images are much higher. As a result, MS images in this study area contain more extractable information for tree classification. Compared to Point-Net++, our approach yielded OA values that were 11.3% (ChEZ) and 2.2% (BFNP) better. Here, the clear lead in the ChEZ was presumably an effect of the relatively small dataset. In this study area, the network parameters of PointNet++ could not be perfectly trained from scratch. In contrast, Silvi-Net was able to deal with a reduced number of samples. Using transfer learning, model parameters were successfully retrained. In summary, the crucial factor for successful performance in our approach was the fusion of lidar data and MS images.

	GEOM	GEOM_EC	GEOM+MS	GEOM_EC+MS
pine	0.87	0.93	0.97	0.99
birch	0.71	0.74	0.92	0.93
alder	0.72	0.71	0.95	0.95
dead	0.80	0.82	0.96	0.97
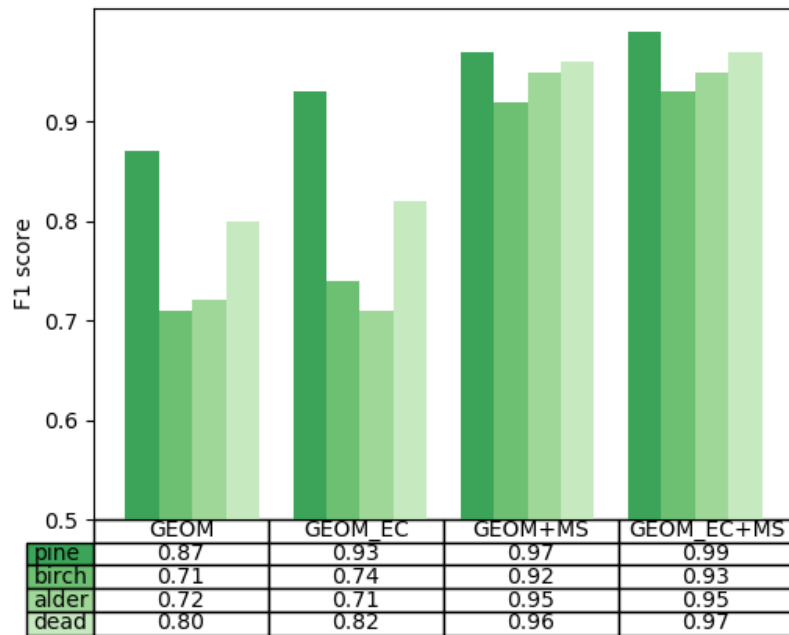
Figure 53: F_1 scores per class for Silvi-Net_majVot in the ChEZ, using different feature sets.

	GEOM	GEOM_EC	GEOM+MS	GEOM_EC+MS
coniferous	0.86	0.85	0.94	0.96
deciduous	0.96	0.97	0.99	0.99
snag	0.85	0.84	0.85	0.86
dead	0.67	0.60	0.80	0.79
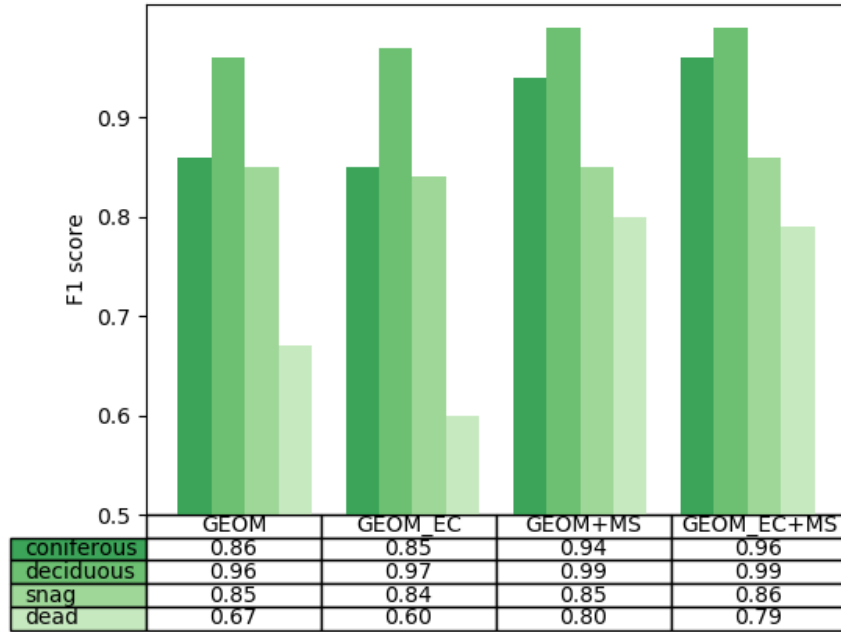
Figure 54: F_1 scores per class for Silvi-Net_majVot in the BFNP, using different feature sets.

5.5.2. Detailed results

In total, we pursued three different classification strategies, differing in the way of dealing with the 2D representations of the 3D point clouds. Overall, the conducted experiments revealed that Silvi-Net$_{majVot}$ generally performed better than Silvi-Net$_{single}$, respectively Silvi-Net$_{collages}$. Both Silvi-Net$_{single}$ and Silvi-Net$_{majVot}$ preserved most 3D information contained in the point cloud and information loss was limited to the subsampling process in the dataloader. In contrast, generating collages produced overall poorer image resolution (Silvi-Net$_{collages}$). The final size of a single tree in a collated image was 50% smaller than the tree size in a single view. Nevertheless, CAM overlays of collated *GEOM_EC* images in the ChEZ (Figure 55) and in the BFNP (Figure 56) demonstrate that Silvi-Net$_{collages}$ still identified most decisive image regions. The advantage of collated images definitely was to process all 12 views of a single tree in one sample and to enable end-to-end classification. Finally, Silvi-Net$_{majVot}$ handled the trade-off between view number and image resolution, and single misclassified samples were outvoted (see Figure 57 and Figure 58).

Let us now focus on the different classes. In the ChEZ, classification of pine, birch, alder, and dead tree was already high when using only *MS* images (OA = 93.6%). Here, automatically extracted features from the 10-cm five-channel MS imagery seemed sufficient to classify single

trees. Note that the MS-based classification results are still considerably dependent on previously conducted lidar-based tree segmentation. When incorporating geometry information and EC, the gain in OA was 2.5%, resulting in a remarkable OA of 96.1%. When only binary side-view images of point clouds were available (*GEOM* images), the OA reached a respectable 77.5%. Moreover, incorporating laser EC improved the results (OA = 80.4%) by 2.9%. Apparently, when using only lidar-based information, pine as the only conifer in the dataset was classified almost perfectly (F_1 score = 0.93). Note that confusion was biggest between birch and alder without MS imagery. However, by integrating MS information, the confusion between these two classes was almost completely resolved. Due to their relatively similar shape, these two deciduous species were difficult to differentiate when classification was only based on geometric properties and EC. Finally, it is notable that features automatically extracted from the *MS* images clearly improved the classification of dead trees, increasing the F_1 score from 0.82 to 0.97. Here, the spectral properties included in the infrared channels (NIR, RE) enhanced the separation of dead and living trees.

(a) pine   (b) birch   (c) alder   (d) dead   (e) birch as alder   (f) alder as birch
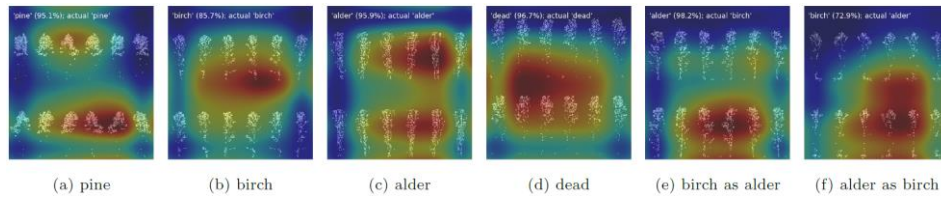
Figure 55: CAM overlays on collated GEOM_EC images in the ChEZ; examples for correct classification (a-d) and misclassification (e-f).

(a) coniferous   (b) deciduous   (c) snag   (d) dead   (e) snag as dead   (f) dead as snag
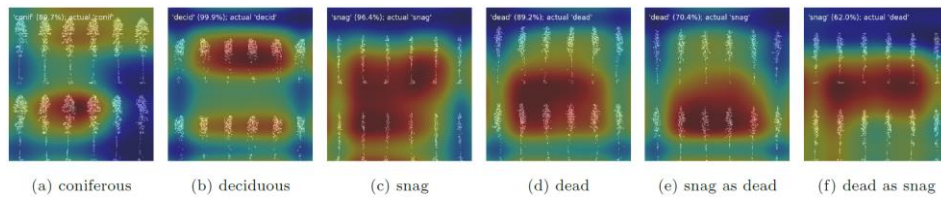
Figure 56: CAM overlays on collated GEOM_EC images in the BFNP; examples for correct classification (a-d) and misclassification (e-f).

In the BFNP, the general results for the classification of single trees into coniferous, deciduous, snag, and dead were partly different from those obtained in the ChEZ. Here, when using only lidar-based data (*GEOM_EC*), PointNet++ (OA = 86.7%) performed 3.2% better than Silvi-Net (OA = 83.5%). In this study area, information loss through generating multiple 2D representations of raw 3D point clouds exceeded the advantages of applying pretrained CNNs. The side-view images were still sufficient to classify deciduous trees almost perfectly (F_1

score = 0.97). However, confusion between dead trees and coniferous trees was considerable because most dead trees with crowns are dead coniferous trees. Thus, these two classes do not differ much in their geometric shape and, therefore, could not be separated well. Surprisingly, the incorporation of EC in side-view images negatively affected the Silvi-Net results, whereas the baseline method profited from this information by 1.0%. Specifically, numerous dead trees were classified as snags and, hence, the F_1 score for dead trees dropped by 0.07. When MS information was included in the classification process (*GEOM_EC+MS*), Silvi-Net reached an OA of 91.5%, exceeding Point-Net++ (OA = 89.3%) by 2.2%. Interestingly, compared to the results based on only MS images (OA = 85.6%), the incorporation of lidar-based side-view images improved the OA by 5.9%. Unsurprisingly, MS information reduced the confusion between coniferous and dead trees. Presumably, the NIR channel was decisive when differentiating these two classes. However, the impact on the F_1 score for snags was negligible. A plausible reason for that is that snags were insufficiently represented from a bird's eye view. We noticed an unsolved moderate confusion between dead trees and snags induced by the manual labelling process. Since the transition between these two classes representing different stages of a dying tree is fluent, some dead trees were erroneously assigned during visual inspection. Without the subdivision into snag and dead trees with crowns, we assume that our approach would have generated better results for a combined class of dead trees in general. Overall, we want to emphasize that Silvi-Net achieved remarkable results for the classification of coniferous (F_1 score = 0.96) and deciduous trees (F_1 score = 0.99), and is ready for practical use.
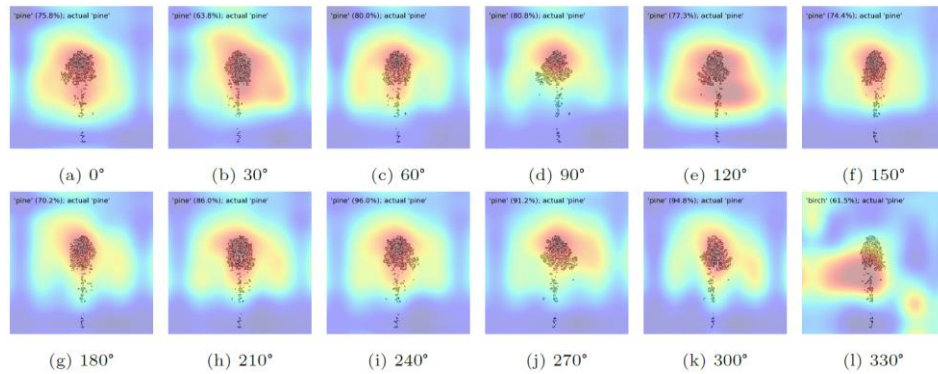
Figure 57: Majority voting of 11 true (a-k) and 1 false (l) predictions leads to correct final prediction; CAM overlays on GEOM_EC images for exemplary pine tree in the ChEZ.
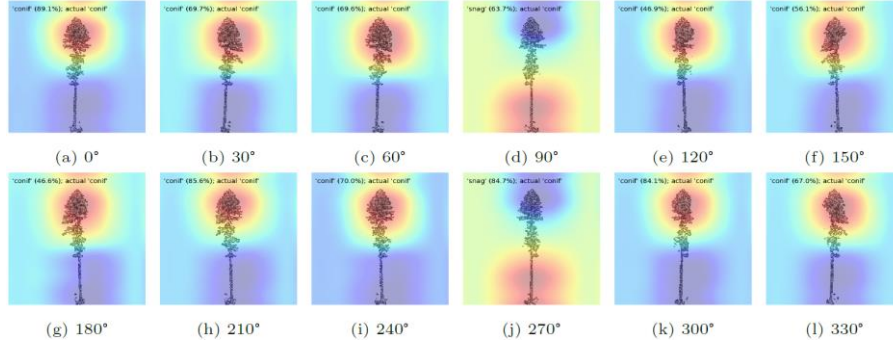
Figure 58: Majority voting of 10 true (a-c, e-i, k-l) and 2 false (d and j) pre-dictions leads to correct final prediction; CAM overlays on GEOM_EC images for exemplary coniferous tree in the BFNP.

5.5.3. Practical issues

Our experiments clearly demonstrated that masking MS image patches with single tree polygons has a positive impact on network perfor-mance. The gain in OA for independent test data was 1.4% in the ChEZ and 2.6% in the BFNP. In particular, the classification of snags and dead trees was clearly improved in the BFNP. Moreover, CAM images of falsely classified *MS_unmasked* samples revealed that, in some cases, ResNet-18 ignored the crucial tree crowns in the image centre, focusing instead on nearby trees from different classes. Note that these misclassifications only occurred in some demanding scenarios with high stand density and, thus, complex tree canopies or even crown overlap. Consequently, masking of aerial image patches is even more important in these challenging situations. In summary, we would definitely rec-ommend using masked MS images for classification. Nevertheless, from a practical point of view, we want to point out that an adequate quality of both tree segmentation and data registration is essential for successful lidar-based masking.

We conducted experiments using 2D representations of 3D point clouds and found that embedding EC slightly improved the OA of Silvi-Net (1.0%) in the ChEZ and in the BFNP (0.4%). However, regarding the single tree classes, we did not notice a significant change in results. To visualize the network's decisions, we plotted CAM overlays of exem-plary side-view *GEOM_EC* images for both correct classification and misclassification. Generally, ResNet-18 identified tree crowns as the most decisive regions in the ChEZ dataset (Figure 59), but in some cases (e.g., Figure 59c), stem information was crucial. Figure 59g clearly shows that a protruding branch falsely led to the prediction "dead tree". In the BFNP, Silvi-Net correctly classified 83.5% of the trees, with the CAM images demonstrating that the neural network was

attentive to either the crown or stem parts. However, for 123 out of 745 samples (16.5%), the predictions were wrong, such as the coniferous sample in Figure 60e being classified as a dead tree due to its obvious similarity to one (Figure 60d). When we look at Figure 60g and Figure 60h, we can understand the confusion between snags and dead trees, but some incorrect predictions were implausible, such as confusion between coniferous and deciduous samples (e.g., Figure 60f).
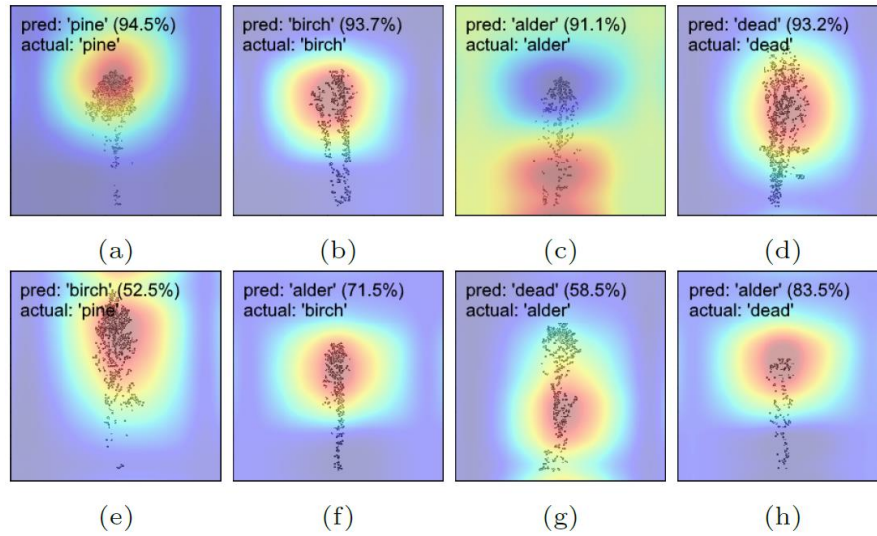
Figure 59: Examples of correct (a-d) and incorrect (e-h) classification of GEOM_EC images in the ChEZ; CAM overlay.
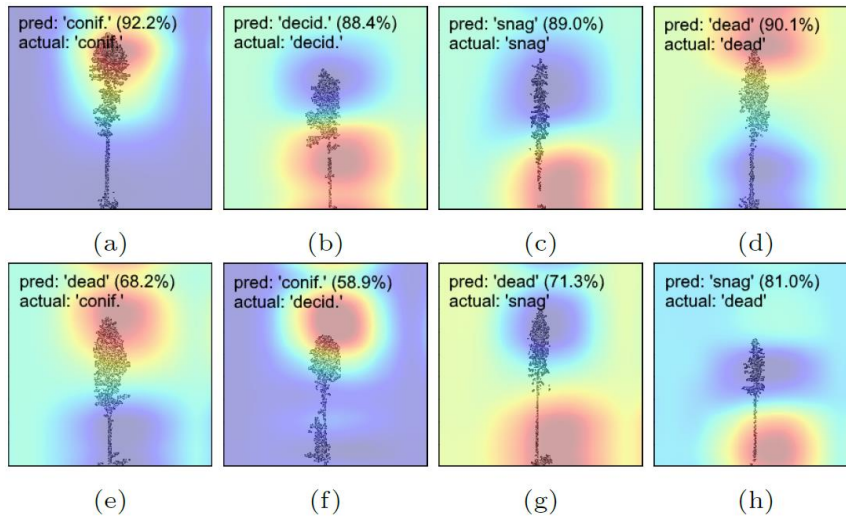
Figure 60: Examples of correct (a-d) and incorrect (e-h) classification of GEOM_EC images in the BFNP; CAM overlay.

5.5.4. Evaluation of Silvi-Net

Overall, we can name numerous advantages for our CNN-based approach to tree species classification, but we want to point out that Silvi-Net enables a comfortable fusion of 2D and 3D data captured by different sensor types. We successfully combined information comprising object geometry, laser EC for each 3D point, and reflectance in the visible and NIR spectra. Undeniably, the automatic extraction of meaningful features from previously generated 2D representations is the key factor. The technique of transfer learning using pretrained weights also facilitates fast model convergence, even for relatively small datasets. Despite these clear advantages, we would also like to address the limitations of our approach. When generating multiple side-view images by projecting 3D point clouds, some information is lost, and all images undergo resolution reduction when placed in the dataloader. Compared to PointNet++'s performance with raw 3D point clouds, information loss is considerable but unavoidable. Furthermore, we want to make clear that a well-performing upstream segmentation of single trees is mandatory for Silvi-Net to work well. In our study, we used almost perfectly delineated single trees generated by the normalized cut segmentation algorithm by manually labelling optimal segments, thereby minimizing the effect of undersegmentation or oversegmentation. However, from a practical point of view, many tree segmentation techniques will cause issues in forests with an even higher stand density and more complexity of the canopy.

5.5.5. Comparison to related work

Investigating related work indicates that Silvi-Net achieves promising and competitive results. Yet, it is challenging to provide a comprehensive and fair comparison to other studies that have addressed object-based classification of individual standing dead trees and snags. On the one hand, utilized datasets strongly differ in spatial, spectral, and temporal resolution. On the other hand, the type of study area (urban, natural, managed) and number of samples and classes fluctuate. Using binary classifiers, Krzystek et al. (2020) classified standing dead trees, snags, and living trees in the BFNP. Overall, their approach separated standing dead trees (F_1 score = 0.92; 310 test samples) from living trees (F_1 score = 0.89; 761 test samples) with an OA of 93%. Snags (76 test samples) were differentiated from living trees (1513 test samples) with an OA of 96%. Interestingly, living trees were classified with an F_1 score of 0.97, whereas the F_1 score for snags was relatively low (0.61). A comparison of multiple classifiers optimized for specific binary tasks to our holistic approach is unfeasible.

To the best of our knowledge, only a few studies have analysed the combined classification of single tree species and dead wood. On an

imbalanced test dataset, Kamińska et al. (2018) reached an OA of 94.3% for the classification of three tree species (spruce, pine, deciduous), each of them further categorized as "dead" or "alive". Instead of F_1 scores, the authors listed producer's accuracy (PA) for the single classes, which is equal to recall. In detail, spruce (146 test samples, PA = 92.4%), pine (148 test samples, PA = 94.1%), deciduous (209 test samples, PA = 99.5%), dead spruce (118 test samples, PA = 90.2%), and dead deciduous (18 test samples, PA = 94%) were classified with high accuracy. However, the PA for dead pine (13 test samples) only reached 69.2%. Recently, Amiri et al. (2019) reported a combined classification of tree species in the BFNP, namely spruce (F_1 score = 0.94), beech (F_1 score = 0.85), fir (F_1 score = 0.59), and dead spruce (F_1 score = 0.74). Based on a huge feature set generated from multi-wavelength ALS data (200 points/m²), the classifier obtained an OA of 82.1%. In summary, Silvi-Net is clearly better than the RF-based approach presented in Amiri et al. (2019). However, a reasonable comparison to Kamińska et al. (2018) is not possible, because data resolution and classification task differ too much. This denotes an urgent need for objective benchmark forest area datasets comprised of annotated high-resolution lidar data and MS or hyperspectral imagery.

5.6. Conclusions

In this work, we have presented Silvi-Net, a dual CNN-based approach for the combined classification of presegmented 3D tree objects with respect to tree species and dead wood in particular. We achieved results superior to those of the baseline method, PointNet++, especially for datasets with a reduced number of samples. Our approach proved to work with data from two natural forests with similar stand density (400–530 trees/ha). Furthermore, lidar data and MS imagery were acquired with different sensor models and, thus, varying geometric and spectral resolutions. The trained models showed high generalization capacity on independent test data. The innovative contribution of our study is the fusion of MS image patches and multiple side-view images, rendered from 3D lidar data, in a CNN-based approach. Compared to experiments conducted using only MS images, the fusion of lidar-based side-view images increased the OA by 2.5% in the ChEZ and 5.9% in the BFNP.

We automatically extracted features using two independently trained ResNet-18 networks, and utilized a standard MLP and majority voting for final object classification. Our optimization process is based on the pretrained weights and recursive retraining of CNN model parameters. For practice, we suggest a combination of high-density lidar data and multi-channel high-resolution MS images. Our results proved that lidar

data are of special importance for both the tree segmentation and classification. Because snags are insufficiently represented in bird's eye images, their classification benefited most from the lidar data. By contrast, the NIR channels of MS images allow the enhancement of dead and living tree definitions as well as tree species classification. Because of its positive impact on the network performance, we also recommend masking MS image patches and embedding calibrated laser EC into the classification process.

In future work, the challenge will be to reliably classify ten or more individual tree species and structurally complex forests. This objective can be supported by improved optical sensors providing high-quality lidar point clouds and high-resolution multi-channel images. In addition, off-the-shelf CNNs and transfer learning can be applied to the specific task of tree species classification, even for relatively small datasets. An interesting task for future work would be the application of panoptic segmentation to forest datasets. This fully DL-based method enables combined delineation and classification of single objects, utilizing prominent image-based neural networks, such as Mask R-CNN (He et al., 2017). Consequently, precise and reliable mapping results could contribute to automatic forest inventory, and support monitoring projects investigating the robustness and sustainability of different forest compositions.

Acknowledgements

6 Synthesis

6.1. Overview of the findings

Initially, a pilot study (see Chapter 2) demonstrated the feasibility of the proposed approach for the detection of unknown radioactive biomass deposits in the ChEZ – even in forested areas with considerably smaller ground point density. Based on the extracted DTM and the conducted 3D vegetation mapping, appropriate features were handcrafted to describe both DTM and vegetation anomalies. Using these feature sets, binary RF classifiers were successfully trained and tested for the classification of radioactive deposits. A follow-up comprehensive study (see Chapter 3) demonstrated the optimization of the approach for detecting radioactive waste sites. Based on improved data collection and data preprocessing, the methodology and the experimental setup were extended. Regarding data preprocessing, a data-driven method for the correction of the radiometric data collected by the laser scanner was implemented. Furthermore, the feature set was clearly extended, and the feature selection strategy was optimized to avoid overfitting of the classifiers. Overall, the results revealed that the proposed methodology is applicable for accurately detecting area-wide unknown radioactive biomass deposits in the ChEZ. The approach is capable of working in forested areas and simultaneously maps both trenches and clamps. The trained classifier demonstrated remarkable generalization properties and enabled the transferability to adjacent study areas. Most importantly, the final results were verified by on-site test drillings that successfully confirmed the existence of previously unknown highly radioactive buried biomass and demolition waste in the classified areas.

The second part of this PhD thesis focused on the application of 2D and 3D DL methods for the combined classification of multiple tree species and standing dead trees. Using raw point clouds of presegmented single trees in the ChEZ, the study referred to in Chapter 4 demonstrated that PointNet++ can be adapted to this task. The developed pipeline integrates additional point attributes and demonstrated high performance classification of three tree species and standing dead trees. In a final study (see Chapter 5), a 2D CNN architecture was designed for the classification of presegmented 3D tree objects. Basically, Silvi-Net is a dual CNN-based approach fusing image tiles rendered from both airborne lidar data and MS orthomosaics. Based on transfer learning and recursive retraining of model parameters, Silvi-Net generally showed fast model convergence, even for relatively small datasets. In general, Silvi-Net demonstrated high generalization capacity and is capable of working with data from different study areas, sensor models and geometric and spectral resolutions. In temperate forests, Silvi-Net achieved superior results compared to the baseline methods (PointNet++ and RF) and is suitable for the combined classification of tree species and dead wood in practice.

6.2. Scope of application

For long-term environmental protection, there is a need to reassess the hazardous waste burials in the ChEZ. Of a total of 1,000 suspected burials, approximately 580 have been investigated, characterised and assessed (Molitor et al., 2017). Based on a comprehensive safety assessment, the authors concluded that out of the inspected burials, 30 represent a safety issue currently (20 mSv/year for a worker) and two are expected to in the long-term (1 mSv/year for an inadvertent settler). To localise the as-yet unmapped burials in further suspicious areas, the UAV-based remote sensing method presented in this PhD thesis is planned to be utilized. In general, UAV-based systematic investigation of the ChEZ can provide a solid baseline for long-term management to justify restrictions or controlled lifting of burials (Molitor et al., 2017). For this purpose, the waste management infrastructure in the ChEZ includes disposal facilities for low-level waste. Furthermore, Molitor et al. (2018) showed that detailed information about radioactive waste sites can support radiologic modelling. Combined with precise 3D vegetation maps at the single-tree level, workers' radiation dose uptake in the field can be estimated and minimized.

This PhD thesis demonstrates that detected vegetation anomalies can indicate radioactively contaminated subsoil. Due to the extensive development of industry and agriculture, a more frequent scenario that could be addressed by these results is the contamination of soil via chemical substances. For instance, heavy metals can cause dangerous environmental conditions for flora and fauna. Thus, capturing the concentrations and spatial distribution of the contamination over large areas is important for risk management. Various studies have shown that the degree of soil contamination interrelates with certain spectral properties of the surrounding vegetation (e.g., Dunagan et al., 2007, Ren et al., 2010). Scafutto et al. (2016) even proposed that spectral signatures of the vegetation might be used to map subsurface oil reservoirs or indicate damaged pipelines. Overall, researchers have encouraged the use of accurate vegetation maps generated from high-resolution multi-channel cameras for the detection of chemically contaminated soil.

A visual analysis of studies focusing on the topic of risk management published in top remote sensing journals (Figure 61) demonstrated that links to terms such as "forest", "fire" and "drought" are most frequent. In terms of forest fires, previously generated vegetation maps are extremely useful. Structural forest parameters at the single-tree level – such as tree position, tree height, canopy cover, crown diameter and above-ground biomass – are especially useful for fire risk assessment (Morsdorf et al., 2004). Because of the climate change-related increase in droughts (Allen et al., 2010) and the existence of large,

highly fire-prone patches, Evangeliou et al. (2015) forecasted a high risk of Cs-137 displacement through future fire events in the ChEZ. Against the background of the 2020 wildfires and the associated concerns about health effects from resuspended radionuclides, the creation of advanced fire behaviour models in the ChEZ has become a major area of focus (Evangeliou and Eckhardt, 2020).
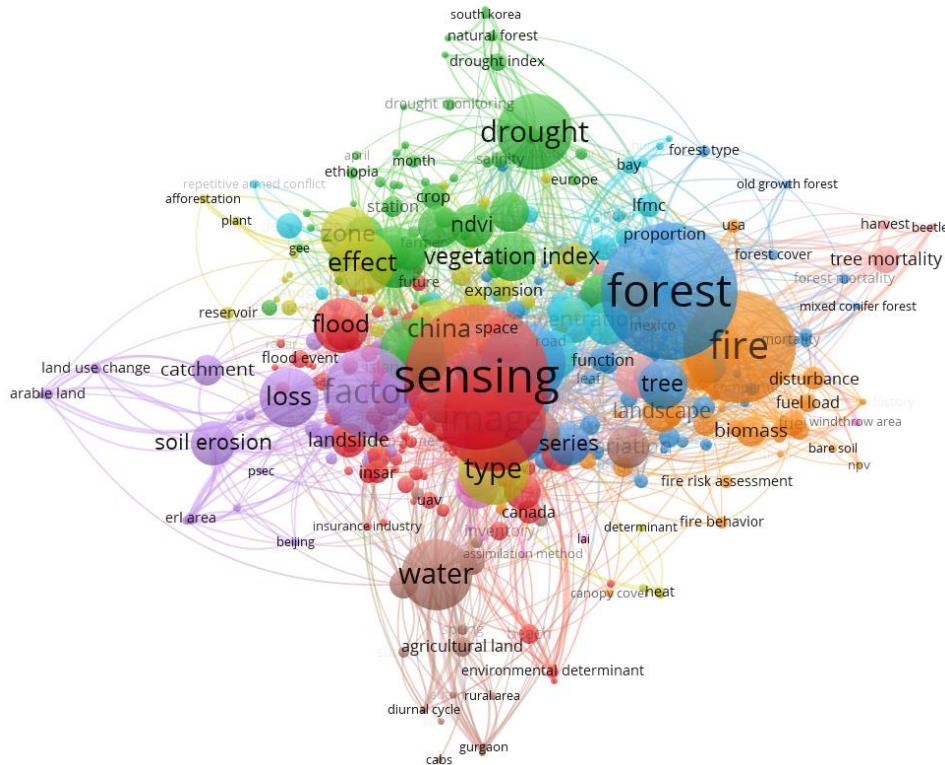
Figure 61: Co-occurrence network of the term "risk management", constructed based on titles and abstracts of 735 scientific publications in top ranked remote sensing journals; created with text mining and visualisation software VOSviewer (van Eck and Waltman, 2010).

In general, the most obvious application of vegetation mapping at the single-tree level is automated forest inventory. Particularly, the fusion of lidar point clouds and MS imagery is the most striking option for the inventory of forest structural variables (Latifi and Heurich, 2019). However, it is still common practice to perform sample-based field measurements followed by an area-wide extrapolation. Obviously, this procedure is very labour intensive and, therefore, both time consuming and expensive. In contrast, DL-based approaches such as Silvi-Net display enormous potential regarding both effectiveness and accuracy and could usher in an era of digital forestry. Furthermore, agile UAVs

equipped with lightweight but powerful optical sensors enable compre-hensive data collection at low operational costs and offer high flexibil-ity. In the near future, it may even be possible to conduct UAV-based forestry operations in real time (Guimarães et al., 2020).

Figure 62: Dead spruces after bark beetle outbreak in the Harz Mountains, Germany (Behrens, 2020).

In recent years, climate change-related effects on forests have become a major challenge for forest ecologists and resource managers. For in-stance, rising temperatures have made certain areas less suitable for heat-sensitive species (Netherer and Schopf, 2010). Furthermore, drought-related tree mortality (Guarín and Taylor, 2005) and a tem-perature-correlated increase in infestations of herbivorous insects such as bark beetle (Bale et al., 2002) have caused serious environmental and economic damage (Figure 62). Here, monitoring projects investi-gating the robustness and sustainability of different species could help to contrive a more resistant forest composition for the future (Churchill et al., 2013). Moreover, Overbeck and Schmidt (2012) recommended specific forest management strategies such as an admixture of broad-leaved tree species. In daily practice, pest infestations need to be de-tected at an early stage to minimize damage. Interestingly, Klouček et al. (2019) demonstrated that the detection of various stages of bark beetle infestation across seasons is possible. By monitoring infested trees using UAV-mounted low-cost RGB and modified NIR sensors, the authors achieved OA values of 78%-96%. When assessing the extent of tree mortality, the mapping of dead wood at the single-tree level can assist in selective logging and forest restoration. Therefore, appro-priate tree-level approaches are urgently needed. Silvi-Net is suitable

for mapping dead trees in forested areas that have been attacked by insects. Another issue regarding tree mortality is the potential risk to infrastructure and people posed by standing dead trees near communication and traffic routes (Stereńczak et al., 2017). Here, up-to-date vegetation maps can support risk assessment and help to coordinate necessary measures in a time and cost effective way.

6.3. Reflections, open issues and outlook

For both of the primary tasks addressed in this PhD thesis, the key factors were the fusion of high-resolution airborne data, thorough data preprocessing – including appropriate calibration – and the subsequent application of ML methods. Therefore, combined knowledge from the fields of geodesy (e.g., coordinate systems and transformations, sensor accuracies and calibration as well as mission planning), computer vision (e.g., processing of 2D and 3D data) and data science (e.g., manual feature extraction, feature selection and DL methods) was required. The highlights of this PhD thesis include the following:

- Detection of buried radioactive waste sites through DTM and vegetation anomalies using optical UAV-based sensors and
- Combined classification of tree species and standing dead trees via a 2D and 3D DL-based approach using automatically extracted features from airborne lidar data and MS imagery.

Nevertheless, there are some issues that remain unsolved, especially in terms of vegetation mapping. Concerning the delineation of single trees from ALS data, the experiments in this thesis relied on the application of the "old but gold" normalized cut algorithm. However, this graph-cut clustering technique leads to under- or oversegmentation of trees in structurally complex forests with overlapping canopies. For normalized cut segmentation, Amiri et al. (2018) demonstrated that an adaptive stopping criterion is capable of delineating tree crowns more accurately than applying a constant threshold value. Furthermore, Dersch et al. (2021) introduced a novel approach for single-tree segmentation, integrating automatic stem detection into normalized cut segmentation. The authors used high-resolution ALS data (>200 points/m²) from dense mixed temperate forest plots (~1000 stems/ha) and showed that their approach clearly improved the OA. Just recently, Braga et al. (2020) confirmed that DL-based instance segmentation using Mask R-CNN can be utilized to perform single-tree crown delineation. Interestingly, due to the lack of a sufficient amount of training data, the authors produced synthetic images from a set of hand-delineated tree crowns, which included tree crown overlap and varying canopy densities. Using patches of 128 x 128 pixels from pansharpened 50-cm RGB WorldView-2 imagery, the authors achieved an

F_1 score of 0.86 for single-tree detection in tropical forests. Neverthe-less, a quantitative area-wide evaluation of the trained model was not conducted in their study.

Figure 63: RIEGL miniVUX-1UAV® laser scanner (left; laser pulse repetition rate: 100 kHz) and MicaSense RedEdge-MX Dual® camera (right; 10 spectral channels).

Figure 64: Exemplary point cloud of a single deciduous tree in leaf-off situation, comprising approximately 32,000 points, captured using a RIEGL miniVUX-1UAV® laser scanner; greyscale colouring with respect to the laser intensity.

In this work, the research topic of object-based tree species classification is covered extensively. However, there are some remaining issues that prevent large-scale use of Silvi-Net in practice. Most importantly, it should be determined if Silvi-Net can be extended to the task of classifying 5 to 10 species in complex mixed forests. Therefore, a more adequate database needs to be created in the future. Large-scale aerial surveys should be performed to ensure a sufficient number of samples from a large variety of species. Using the latest UAV-compatible lidar and MS sensors with improved spatial and spectral resolution (Figure 63), extremely detailed information regarding tree structure (Figure 64) and spectral response can be captured.

To make classification results "waterproof", reliable reference data specifying the single tree species should be generated by forestry experts. Labelling could be realized either by manual annotation in the field or, more efficiently, by visual interpretation of the 2D and 3D datasets. This would require a large investment of time and money. Nevertheless, the goal should be to demonstrate that Silvi-Net can be utilized to determine the most common tree species in a specific region. This verification would allow it to be widely used in practice. Another issue to be addressed in future research is the transferability of trained networks between multi-temporal forest datasets. Here, appropriate transfer learning and network fine tuning could be the key to success. Because understory trees are invisible in the orthomosaics, successful application of Silvi-Net is currently limited to dominant trees. This restriction could be removed, provided that understory trees are segmented properly. For these smaller trees, the pipeline would have to be adapted so that the model's prediction is highly dependent on images in side view rather than bird's eye view. Note that the successful application of Silvi-Net is highly dependent on the quality of the upstream segmentation of single trees. Thus, future work should focus on the application of DL-based approaches to forest datasets for combined delineation and classification of single trees. It is conceivable that the current two-step approach presented in this thesis could be replaced by an end-to-end learning approach, for instance based on Mask R-CNN (Kattenborn et al., 2021). More precisely, panoptic segmentation using high-resolution UAV-based imagery and dense airborne lidar point clouds in complex heterogeneous forests with dense and overlapping canopies presents an interesting and demanding research topic.

Bibliography

Agisoft LLC (2018). Agisoft PhotoScan Professional 1.4.1. https://www.agisoft.com/. Accessed: 2021-01-12.

Allen, C., Macalady, A., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D., Hogg, E., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S., Semerci, A., and Cobb, N. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. Forest Ecology and Management, 259, 660-684. doi:10.1016/j.foreco.2009.09.001.

Amiri, N., Krzystek, P., Heurich, M., and Skidmore, A. (2019). Classification of tree species as well as standing dead trees using triple wavelength ALS in a temperate forest. Remote Sensing, 11. doi:10.3390/rs11222614.

Amiri, N., Polewski, P., Heurich, M., Krzystek, P., and Skidmore, A. (2018). Adaptive stopping criterion for topdown segmentation of ALS point clouds in temperate coniferous forests. ISPRS Journal of Photogrammetry and Remote Sensing, 141, 265-274. doi:10.1016/j.isprsjprs.2018.05.006.

Antropov, V., Bugai, D., Dutton, L., Gerchikov, M., Kennett, E., Ledenev, A., Novikov, A., Rudko, V., and Ziegenhagen, J. (2001). Review and Analysis of Solid Long-lived and High Level Radioactive Waste arising at the Chernobyl Nuclear Power Plant and the Restricted Zone. Technical Report. doi:10.13140/RG.2.2.31635.17442/1.

Applied Imagery (2018). Quick Terrain Modeler v8.0.7. http://www.appliedimagery.com. Accessed: 2021-01-12.

Arkhipov, N., Kuchma, N., Askbrant, S., Pasternak, P., and Musica, V. (1994). Acute and long-term effects of irradiation on pine (Pinus silvestris) stands post-Chernobyl. Science of the Total Environment, 157, 383-386. doi:10.1016/0048-9697(94)90601-7.

Atkinson, P., and Tatnall, A. (1997). Introduction neural networks in remote sensing. International Journal of Remote Sensing, 18, 699-709. doi:10.1080/014311697 218700.

Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives, 33, 110-117. https://www.isprs.org/proceed ings/XXXIII/congress/part4/111_XXXIII-part4.pdf. Accessed: 2021-01-12.

Bale, J., Masters, G., Hodkinson, I., Awmack, C., Bezemer, T., Brown, V., Butterfield, J., Buse, A., Coulson, J., Farrar, J., Good, J., Harrington, R., Hartley, S., Jones, T., Lindroth, R., Press, M., Symrnioudis, I., Watt, A., and Whittaker, J. (2002). Herbivory in global climate change research: Direct effects of rising temperature on insect herbivores. Global Change Biology, 8, 1-16. doi:10.1046/j.1365-2486.2002.00451.x.

Barry, R. (2006). The status of research on glaciers and global glacier recession: A review. Progress in Physical Geography, 30, 285-306. doi:10.1191/0309133306pp 478ra.

BayesMap Solutions LLC (2018). BayesStripAlign 2.0. http://bayesmap.com/products/ bayesstripalign/. Accessed: 2021-01-12.

Behrens, C. (2020). https://pixabay.com/de/photos/wald-kiefer-nadel-holz-bunte-5468 689/. Accessed: 2021-01-12.

Bollandsås, O., Risbøl, O., Ene, L., Nesbakken, A., Gobakken, T., and Næsset, E. (2012). Using airborne smallfootprint laser scanner data for detection of cultural remains in

forests: An experimental study of the effects of pulse density and DTM smoothing. Journal of Archaeological Science, 39, 2733-2743. doi:10.1016/j.jas.2012.04.026.

Bonzom, J.-M., Hättenschwiler, S., Lecomte-Pradines, C., Chauvet, E., Gaschak, S., Beaugelin-Seiller, K., Della-Vedova, C., Dubourg, N., Maksimenko, A., Garnier-Laplace, J., and Adam-Guillermin, C. (2016). Effects of radionuclide contamination on leaf litter decomposition in the Chernobyl Exclusion Zone. Science of the Total Environment, 562, 596-603. doi:10.1016/j.scitotenv.2016.04.006.

Braga, J., Peripato, V., Dalagnol, R., Ferreira, M., Tarabalka, Y., Aragão, L., de Campos Velho, H., Shiguemori, E., and Wagner, F. (2020). Tree crown delineation algorithm based on a convolutional neural network. Remote Sensing, 12. doi:10.3390/RS12081288.

Briechle, S., Krzystek, P., and Vosselman, G. (2019). Semantic labeling of als point clouds for tree species mapping using the deep neural network pointnet++. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives, 42, 951-955. doi:10.5194/isprs-archives-XLII-2-W13-951-2019.

Briechle, S., Krzystek, P., and Vosselman, G. (2020a). Classification of tree species and standing dead trees by fusing UAV-based lidar data and multispectral imagery in the 3D deep neural network PointNet++. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, V-2-2020, 203-210. doi:10.5194/isprs-annals-V-2-2020-203-2020.

Briechle, S., Krzystek, P., and Vosselman, G. (2021). Silvi-Net – A dual-CNN approach for combined classification of tree species and standing dead trees from remote sensing data. International Journal of Applied Earth Observation and Geoinformation, 98. doi:10.1016/j.jag.2020.102292.

Briechle, S., Molitor, N., Krzystek, P., and Vosselman, G. (2020b). Detection of radioactive waste sites in the Chornobyl Exclusion Zone using UAV-based lidar data and multispectral imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 167, 345-362. doi:10.1016/j.isprsjprs.2020.06.015.

Briechle, S., Sizov, A., Tretyak, O., Antropov, V., Molitor, N., and Krzystek, P. (2018). UAV-based detection of unknown radioactive biomass deposits in Chernobyl's Exclusion Zone. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives, 42, 163-169. doi:10.5194/isprs-archives-XLII-2-163-2018.

Bugai, D., Kashparov, V., Dewiére, L., Khomutinin, Y., Levchuk, S., and Yoschenko, V. (2005). Characterization of subsurface geometry and radioactivity distribution in the trench containing Chernobyl clean-up wastes. Environmental Geology, 47, 869-881. doi:10.1007/s00254-004-1218-6.

Bugai, D., Skalskyy, A., Dzhepo, S., Kubko, Y., Kashparov, V., Van Meir, N., Stammose, D., Simonucci, C., and Martin-Garin, A. (2012). Radionuclide migration at experimental polygon at Red Forest waste site in Chernobyl zone. Part 2: Hydrogeological characterization and groundwater transport modeling. Applied Geochemistry, 27, 1359-1374. doi:10.1016/j.apgeochem.2011.09.028.

Cailleret, M., Heurich, M., and Bugmann, H. (2014). Reduction in browsing intensity may not compensate climate change effects on tree species composition in the Bavarian Forest National Park. Forest Ecology and Management, 328, 179-192. doi:10.1016/j.foreco.2014.05.030.

Casas, A., García, M., Siegel, R., Koltunov, A., Ramírez, C., and Ustin, S. (2016). Burned forest characterization at single-tree level with airborne laser scanning for assessing

wildlife habitat. Remote Sensing of Environment, 175, 231-241. doi:10.1016/j.rse. 2015.12.044.

Cazenave, A., and Llovel, W. (2010). Contemporary sea level rise. Annual Review of Marine Science, 2, 145-173. doi:10.1146/annurev-marine-120308-081105.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M., Delitti, W., Du-que, A., Eid, T., Fearnside, P., Goodman, R., Henry, M., Martínez-Yrízar, A., Mugasha, W., Muller-Landau, H., Mencuccini, M., Nelson, B., Ngomanda, A., Nogueira, E., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C., Saldarriaga, J., and Vieilledent, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biology, 20, 3177-3190. doi:10.1111/gcb.12629.

Churchill, D., Larson, A., Dahlgreen, M., Franklin, J., Hessburg, P., and Lutz, J. (2013). Restoring forest resilience: From reference spatial patterns to silvicultural prescriptions and monitoring. Forest Ecology and Management, 291, 442-457. doi:10.1016/j.foreco.2012.11.007.

Dalponte, M., Bruzzone, L., and Gianelle, D. (2012). Tree species classification in the southern alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and lidar data. Remote Sensing of Environment, 123, 258-270. doi:10.1016/j.rse.2012.03.013.

Datt, B. (1999). A new reflectance index for remote sensing of chlorophyll content in higher plants: Tests using eucalyptus leaves. Journal of Plant Physiology, 154, 30-36. doi:10.1016/S0176-1617(99)80314-9.

Daughtry, C., Walthall, C., Kim, M., De Colstoun, E., and McMurtrey III, J. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sensing of Environment, 74, 229-239. doi:10.1016/S0034-4257(00)00113-9.

Davids, C., and Tyler, A. (2003). Detecting contamination-induced tree stress within the Chernobyl Exclusion Zone. Remote Sensing of Environment, 85, 30-38. doi:10.1016/S0034-4257(02)00184-0.

Degerickx, J., Roberts, D., McFadden, J., Hermy, M., and Somers, B. (2018). Urban tree health assessment using airborne hyperspectral and lidar imagery. International Journal of Applied Earth Observation and Geoinformation, 73, 26-38. doi:10.1016/j.jag.2018.05.021.

Del Frate, F., Pacifici, F., Schiavon, G., and Solimini, C. (2007). Use of neural networks for automatic classification from high-resolution images. IEEE Transactions on Geoscience and Remote Sensing, 45, 800-809. doi:10.1109/TGRS.2007.892009.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 248-255. doi: 10.1109/CVPR.2009.5206848.

Dersch, S., Heurich, M., Krueger, N., and Krzystek, P. (2021). Combining graph-cut clustering with objectbased stem detection for tree segmentation in highly dense airborne lidar point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 172, 207-222. doi:10.1016/j.isprsjprs.2020.11.016.

Dunagan, S., Gilmore, M., and Varekamp, J. (2007). Effects of mercury on visible/near-infrared reflectance spectra of mustard spinach plants (brassica rapa p.). Environmental Pollution, 148, 301-311. doi:10.1016/j.envpol.2006.10.023.

van Eck, N., and Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics, 84, 523-538. doi:10.1007/s11192-009-0146-3.

European Union (2014-2017). Support to radioactive waste management in Ukraine. https://nuclear.jrc.ec.europa.eu/tipins/contracts/support-radioactive-waste-man agement-ukraine. Accessed: 2021-01-12.

Evangeliou, N., Balkanski, Y., Cozic, A., Hao, W., Mouillot, F., Thonicke, K., Paugam, R., Zibtsev, S., Mousseau, T., Wang, R., Poulter, B., Petkov, A., Yue, C., Cadule, P., Koffi, B., Kaiser, J., Møller, A., and Classen, A. (2015). Fire evolution in the radioactive forests of Ukraine and Belarus: Future risks for the population and the environment. Ecological Monographs, 85, 49-72. doi:10.1890/14-1227.1.

Evangeliou, N., and Eckhardt, S. (2020). Uncovering transport, deposition and impact of radionuclides released after the early spring 2020 wildfires in the Chernobyl Exclusion Zone. Scientific Reports, 10. doi:10.1038/s41598-020-67620-3.

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

Flynn, K.F., Glendenin, L.E., Harkness, A.L. and Steinberg, E.P. (1965). Half-lives of 90Sr and 137Cs. Journal of Inorganic and Nuclear Chemistry, 27, 21-23. doi:10.1016/0022-1902(65)80184-1.

Friedman, K., and Li, X. (2000). Monitoring hurricanes over the ocean with wide swath SAR. Johns Hopkins APL Technical Digest (Applied Physics Laboratory), 21, 80-85.

Gevaert, C., Persello, C., Nex, F., and Vosselman, G. (2018). A deep learning approach to DTM extraction from imagery using rule-based training labels. ISPRS Journal of Photogrammetry and Remote Sensing, 142, 106-123. doi:10.1016/j.isprsjprs.2018.06.001.

Gitelson, A., and Merzlyak, M. N. (1994). Spectral reflectance changes associated with autumn senescence of Aesculus hippocastanum L. and Acer platanoides L. leaves. Spectral features and relation to chlorophyll estimation. Journal of Plant Physiology, 143, 286-292. doi:10.1016/S0176-1617(11)81633-0.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep Learning. MIT Press. http://www.deeplearningbook.org. Accessed: 2021-01-12.

Gregorutti, B., Michel, B., and Saint-Pierre, P. (2017). Correlation and variable importance in random forests. Statistics and Computing, 27, 659-678. doi:10.1007/s11222-016-9646-1.

Griffiths, D., and Boehm, J. (2019). A review on deep learning techniques for 3D sensed data classification. Remote Sensing, 11. doi:10.3390/rs11121499.

Guarín, A., and Taylor, A. (2005). Drought triggered tree mortality in mixed conifer forests in Yosemite National Park, California, USA. Forest Ecology and Management, 218, 229-244. doi:10.1016/j.foreco.2005.07.014.

Guimarães, N., Pádua, L., Marques, P., Silva, N., Peres, E., and Sousa, J. (2020). Forestry remote sensing from unmanned aerial vehicles: A review focusing on the data, processing and potentialities. Remote Sensing, 12. doi:10.3390/rs12061046.

Guyot, A., Hubert-Moy, L., and Lorho, T. (2018). Detecting Neolithic burial mounds from lidar-derived elevation data using a multi-scale approach and machine learning techniques. Remote Sensing, 10. doi:10.3390/rs10020225.

Hamraz, H., Jacobs, N., Contreras, M., and Clark, C. (2019). Deep learning for conifer/deciduous classification of airborne lidar 3D point clouds representing individual trees. ISPRS Journal of Photogrammetry and Remote Sensing, 158, 219-230. doi:10.1016/j.isprsjprs.2019.10.011.

Hartling, S., Sagan, V., Sidike, P., Maimaitijiang, M., and Carron, J. (2019). Urban tree species classification using a WorldView-2/3 and lidar data fusion approach and deep learning. Sensors (Switzerland), 19. doi:10.3390/s19061284.

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask R-CNN. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2980-2988. doi:10.1109/ICCV.2017.322.

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition. CoRR, abs/1512.03385. http://arxiv.org/abs/1512.03385. Accessed: 2021-01-12.

Heinzel, J., and Koch, B. (2012). Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. International Journal of Applied Earth Observation and Geoinformation, 18, 101-110. doi:10.1016/j.jag.2012.01.025.

Höfle, B., and Pfeifer, N. (2007). Correction of laser scanning intensity data: Data and model-driven approaches. ISPRS Journal of Photogrammetry and Remote Sensing, 62, 415-433. doi:10.1016/j.isprsjprs.2007.05.008.

Holmgren, J., Persson, A., and Söderman, U. (2008). Species identification of individual trees by combining high resolution lidar data with multi-spectral images. International Journal of Remote Sensing, 29, 1537-1552. doi:10.1080/01431160701736471.

Hou, S., Liu, X., and Wang, Z. (2017). DualNet: Learn complementary features for image recognition. Proceedings of the IEEE International Conference on Computer Vision (CVPR), pp. 502-510. doi:10.1109/ICCV.2017.62.

Hu, F., Xia, G.-S., Hu, J., and Zhang, L. (2015). Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. Remote Sensing, 7, 14680-14707. doi:10.3390/rs71114680.

Huang, G., Liu, Z., Van Der Maaten, L., andWeinberger, K. Q. (2017). Densely connected convolutional networks. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4700-4708. doi:10.1109/CVPR.2017.243.

Ibeo Automotive Systems GmbH (2010). Operating manual ibeo LUX 2010® laser scanner v1.6. Unpublished internal technical report. Page 20.

Immitzer, M., Atzberger, C., and Koukal, T. (2012). Tree species classification with random forest using very high spatial resolution 8-band worldview-2 satellite data. Remote Sensing, 4, 2661-2693. doi:10.3390/rs4092661.

Ioannidou, A., Chatzilari, E., Nikolopoulos, S., and Kompatsiaris, I. (2017). Deep learning advances in computer vision with 3D data: A survey. ACM Computing Surveys, 50. doi:10.1145/3042064.

Ioffe, S., and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. https://arxiv.org/pdf/1502.03167.pdf. Accessed: 2021-01-12.

Jalobeanu, A., and Gonçalves, G. R. (2014). Automated probabilistic lidar swath registration. In AGU Fall Meeting Abstracts. doi:10.13140/2.1.3126.6883.

Jutzi, B., and Gross, H. (2009). Normalization of lidar intensity data based on range and surface incidence angle. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38, 213-218. https://pdfs.semanticscholar.org/91fc/1c40a8168bf6060642a01ff456a36e42ebe0.pdf. Accessed: 2021-01-12.

Kaasalainen, S., Jaakkola, A., Kaasalainen, M., Krooks, A., and Kukko, A. (2011). Analysis of incidence angle and distance effects on terrestrial laser scanner intensity: Search for correction methods. Remote Sensing, 3, 2207-2221. doi:10.3390/rs3102207.

Kamińska, A., Lisiewicz, M., Stereńczak, K., Kraszewski, B., and Sadkowski, R. (2018). Species-related single dead tree detection using multi-temporal ALS data and CIR imagery. Remote Sensing of Environment, 219, 31-43. doi:10.1016/j.rse.2018.10.005.

Kashparov, V., Yoschenko, V., Levchuk, S., Bugai, D., Van Meir, N., Simonucci, C., and Martin-Garin, A. (2012). Radionuclide migration in the experimental polygon of the Red Forest waste site in the Chernobyl zone – Part 1: Characterization of the waste trench, fuel particle transformation processes in soils, biogenic fluxes and effects on biota. Applied Geochemistry, 27, 1348-1358. doi:10.1016/j.apgeochem.2011.11.004.

Kattenborn, T., Leitloff, J., Schiefer, F., and Hinz, S. (2021). Review on convolutional neural networks (CNN) in vegetation remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 173, 24-49. doi:10.1016/j.isprsjprs.2020.12.010.

Kingma, D., and Ba, J. (2015). Adam: A method for stochastic optimization. In: 3rd International Conference on Learning Representations, ICLR 2015 – Conference Track Proceedings. https://arxiv.org/pdf/1412.6980.pdf. Accessed: 2021-01-12.

Klouček, T., Komárek, J., Surový, P., Hrach, K., Janata, P., and Vašíček, B. (2019). The use of UAV mounted sensors for precise detection of bark beetle infestation. Remote Sensing, 11. doi:10.3390/rs11131561.

Kogan, F. (1997). Global drought watch from space. Bulletin of the American Meteorological Society, 78, 621-636. doi:10.1175/1520-0477(1997)078<0621:GDWFS>2.0.CO;2.

Korpela, I., Ørka, H., Hyyppä, J., Heikkinen, V., and Tokola, T. (2010). Range and AGC normalization in airborne discrete-return lidar intensity data for forest canopies. ISPRS Journal of Photogrammetry and Remote Sensing, 65, 369-379. doi:10.1016/j.isprsjprs.2010.04.003.

Krizhevsky, A., Sutskever, I., and Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 2, 1097-1105. http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf. Accessed: 2021-01-12.

Krzystek, P., Serebryanyk, A., Schnörr, C., Červenka, J., and Heurich, M. (2020). Large-scale mapping of tree species and dead trees in Sumava National Park and Bavarian Forest National Park using lidar and multispectral imagery. Remote Sensing, 12. doi:10.3390/rs12040661.

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

Laj, P., Klausen, J., Bilde, M., Plaß-Duelmer, C., Pappalardo, G., Clerbaux, C., Baltensperger, U., Hjorth, J., Simpson, D., Reimann, S., Coheur, P.-F., Richter, A., De Mazière, M., Rudich, Y., McFiggans, G., Torseth, K., Wiedensohler, A., Morin, S., Schulz, M., Allan, J., Attié, J.-L., Barnes, I., Birmili, W., Cammas, J., Dommen, J., Dorn, H.-P., Fowler, D., Fuzzi, S., Glasius, M., Granier, C., Hermann, M., Isaksen, I., Kinne, S., Koren, I., Madonna, F., Maione, M., Massling, A., Moehler, O., Mona, L., Monks, P., Müller, D., Müller, T., Orphal, J., Peuch, V.-H., Stratmann, F., Tanré, D., Tyndall, G., Abo Riziq, A., Van Roozendael, M., Villani, P., Wehner, B., Wex, H., and Zardini, A. (2009). Measuring atmospheric composition change. Atmospheric Environment, 43, 5351-5414. doi:10.1016/j.atmosenv.2009.08.020.

Landrieu, L., and Simonovsky, M. (2018). Large-scale point cloud semantic segmentation with superpoint graphs. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4558-4567. doi:10.1109/CVPR.2018.00479.

Lasaponara, R., Coluzzi, R., Gizzi, F., and Masini, N. (2010). On the lidar contribution for the archaeological and geomorphological study of a deserted medieval village in Southern Italy. Journal of Geophysics and Engineering, 7, 155-163. doi:10.1088/1742-2132/7/2/S01.

Latifi, H., Fassnacht, F., Müller, J., Tharani, A., Dech, S., and Heurich, M. (2015). Forest inventories by lidar data: A comparison of single tree segmentation and metric-based methods for inventories of a heterogeneous temperate forest. International Journal of Applied Earth Observation and Geoinformation, 42, 162-174. doi:10.1016/j.jag.2015.06.008.

Latifi, H., and Heurich, M. (2019). Multi-scale remote sensing-assisted forest inventory: A glimpse of the state-of-the-art and future prospects. Remote Sensing, 11. doi:10.3390/rs11111260.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature, 521, 436-444. doi:10.1038/nature14539.

LeCun, Y., Huang, F., and Bottou, L. (2004). Learning methods for generic object recognition with invariance to pose and lighting. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2, 97-104. doi:10.1109/CVPR.2004.1315150.

LeNail, A. (2019). NN-SVG: Publication-ready neural network architecture schematics. Journal of Open Source Software, 4, 747. http://alexlenail.me/NN-SVG/AlexNet.html. Accessed: 2021-01-12.

Li, Y., Bu, R., Sun, M., Wu, W., Di, X., and Chen, B. (2018). PointCNN: Convolution on X-transformed points. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Eds.), Advances in Neural Information Processing Systems 31, pp. 820-830. Curran Associates, Inc. http://papers.nips.cc/paper/7362-pointcnn-convolution-on-x-transformed-points.pdf. Accessed: 2021-01-12.

Liaw, A., and Wiener, M. (2002). Classification and regression by randomforest. R News, 2, 18-22. https://CRAN.R-project.org/doc/Rnews/. Accessed: 2021-01-12.

Liu, T., Abd-Elrahman, A., Morton, J., and Wilhelm, V. (2018). Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. GIScience and Remote Sensing, 55, 243-264. doi:10.1080/15481603.2018.1426091.

Ma, L., Fu, T., Blaschke, T., Li, M., Tiede, D., Zhou, Z., Ma, X., and 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. doi:10.3390/ijgi6020051.

MacFarlane, J., Payton, O., Keatley, A., Scott, G., Pullin, H., Crane, R., Smilion, M., Popescu, I., Curlea, V., and Scott, T. (2014). Lightweight aerial vehicles for monitoring, assessment and mapping of radiation anomalies. Journal of Environmental Radioactivity, 136, 127-130. doi:10.1016/j.jenvrad.2014.05.008.

McRoberts, R., and Tomppo, E. (2007). Remote sensing support for national forest inventories. Remote Sensing of Environment, 110, 412-419. doi:10.1016/j.rse.2006.09.034.

Melcher, C. L. (2000). Scintillation crystals for PET. Journal of nuclear medicine, 41, 1051-1055. https://jnm.snmjournals.org/content/jnumed/41/6/1051.full.pdf. Accessed: 2021-01-12.

Metternicht, G., Hurni, L., and Gogu, R. (2005). Remote sensing of landslides: An analysis of the potential contribution to geo-spatial systems for hazard assessment in mountainous environments. Remote Sensing of Environment, 98, 284-303. doi:10.1016/j.rse.2005.08.004.

MicaSense Inc (2015). MicaSense RedEdge 3 Multispectral Camera User Manual, Rev 06. https://support.micasense.com/hc/en-us/article_attachments/204648307/RedEdge_User_Manual_06.pdf. Accessed: 2021-01-12.

MicaSense Inc (2019). Use of calibrated reflectance panels for rededge data. https://support.micasense.com/hc/en-us/articles/115000765514-Use-of-Calibrated-Reflectance-Panels-For-RedEdge-Data. Accessed: 2021-01-12.

Molitor, N. (2017). Technical Report Task D3: Annex 16: Complementary Airborne Geophysical Survey. Unpublished internal technical report.

Molitor, N., Thierfeldt, S., Haneke, K., Nitzsche, O., Bugai, D., Sizov, A., and Drace, Z. (2017). Recent safety assessment findings on management of legacy wastes from Chornobyl accident. In: International Conference on Nuclear Decommissioning and Environment Recovery, INUDECO. https://www.researchgate.net/publication/317400817_Recent_safety_assessment_findings_on_management_of_legacy_wastes_from_Chornobyl_accident. Accessed: 2021-01-12.

Molitor, N., Zabulonov, Y., Krzystek, P., Siebold, K., Haneke, K., Nitzsche, O., Groot J., d., Antropov, V., Tretyak, O., Sizov, A., Bugai, D., Kovach, V., Bayer, P., Drace, Z., and Bachmaier, B. (2018). New developments in airborne geophysical survey technologies and their application for investigation and radiological assessment of highly contaminated areas and legacy radioactive waste storages in the Chornobyl Exclusion Zone. In: International Conference on Nuclear Decommissioning and Environment Recovery, INUDECO. https://www.researchgate.net/publication/327704022_New_developments_in_airborne_geophysical_survey_technologies_and_their_application_for_investigation_and_radiological_assessment_of_highly_contaminated_areas_and_legacy_radioactive_waste_storages_in_t. Accessed: 2021-01-12.

Morsdorf, F., Meier, E., Kötz, B., Itten, K., Dobbertin, M., and Allgöwer, B. (2004). Lidar-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. Remote Sensing of Environment, 92, 353-362. doi:10.1016/j.rse.2004.05.013.

Næsset, E. (2004). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. Scandinavian Journal of Forest Research, 19, 164-179. doi:10.1080/02827580310019257.

Netherer, S., and Schopf, A. (2010). Potential effects of climate change on insect herbivores in European forests – general aspects and the pine processionary moth as specific example. Forest Ecology and Management, 259, 831-838. doi:10.1016/j.foreco. 2009.07.034.

Ng, H.-W., Nguyen, V., Vonikakis, V., and Winkler, S. (2015). Deep learning for emotion recognition on small datasets using transfer learning. ICMI 2015 – Proceedings of the 2015 ACM International Conference on Multimodal Interaction, pp. 443-449. doi:10.1145/2818346.2830593.

NovAtel Inc. (2017). Inertial Explorer 8.70 – GNSS and inertial post-processing software. https://www.novatel.com/products/software/inertial-explorer/. Accessed: 2021-01-12.

NVIDIA Corporation (2019). Nvidia Titan V – Nvidia's supercomputing GPU architecture. https://www.nvidia.com/en-us/titan/titan-v/. Accessed: 2021-01-12.

Osada, R., Funkhouser, T., Chazelle, B., and Dobkin, D. (2002). Shape distributions. ACM Transactions on Graphics (TOG), 21, 807-832. doi:10.1145/571647.571648.

Overbeck, M., and Schmidt, M. (2012). Modelling infestation risk of Norway spruce by Ips typographus (L.) in the Lower Saxon Harz Mountains (Germany). Forest Ecology and Management, 266, 115-125. doi:10.1016/j.foreco.2011.11.011.

Pacifici, F., Chini, M., and Emery, W. (2009). A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. Remote Sensing of Environment, 113, 1276-1292. doi:10.1016/j.rse.2009.02.014.

Pal, M. (2005). Random forest classifier for remote sensing classification. International Journal of Remote Sensing, 26, 217-222. doi:10.1080/01431160412331269698.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (Eds.), Advances in Neural Information Processing Systems 32, pp. 8024-8035. Curran Associates, Inc. http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf. Accessed: 2021-01-12.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825-2830. https://arxiv.org/pdf/1201.0490.pdf. Accessed: 2021-01-12.

Polewski, P. (2017). Reconstruction of standing and fallen single dead trees in forested areas from LiDAR data and aerial imagery. Ph.D. thesis Technische Universität München. http://publikationen.badw.de/de/044724403/pdf/CC%20BY. Accessed: 2021-01-12.

Polewski, P., Yao, W., Heurich, M., Krzystek, P., and Stilla, U. (2015a). Active learning approach to detecting standing dead trees from ALS point clouds combined with aerial infrared imagery. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2015, 10-18. doi:10.1109/CVPRW.2015.7301378.

Polewski, P., Yao, W., Heurich, M., Krzystek, P., and Stilla, U. (2015b). Detection of fallen trees in ALS point clouds using a Normalized Cut approach trained by simulation. ISPRS Journal of Photogrammetry and Remote Sensing, 105, 252-271. doi:10.1016/j.isprsjprs.2015.01.010.

Prabha, R., Tom, M., Rothermel, M., Baltsavias, E., Leal-Taixe, L., and Schindler, K. (2020). Lake ice monitoring with webcams and crowd-sourced images. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, V-2-2020, 549-556. doi:10.5194/isprs-annals-V-2-2020-549-2020.

Primavision Technologies GbR (2017). 3D tree segmentation from point clouds (lidar, DSM) for forest inventory. http://primavision-tec.de/products/prod_tree-finder. Accessed: 2021-01-12.

Puissant, A., Rougiera, S., and Stumpf, A. (2014). Object-oriented mapping of urban trees using random forest classifiers. International Journal of Applied Earth Observation and Geoinformation, 26, 235-245. doi:10.1016/j.jag.2013.07.002.

Pyysalo, U., and Hyyppä, H. (2002). Reconstructing tree crowns from laser scanner data for feature extraction. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives, 34. https://pdfs.seman ticscholar.org/47b8/6e912bcde3c1c56c1cf71781737e12df90f4.pdf. Accessed: 2021-01-12.

QGIS Development Team (2017). QGIS 2.18.3. https://qgis.org. Accessed: 2021-01-12.

Qi, C., Su, H., Mo, K., and Guibas, L. (2017a). PointNet: Deep learning on point sets for 3D classification and segmentation. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2017-January, pp. 77-85. doi:10.1109/CVPR.2017.16.

Qi, C., Su, H., Niebner, M., Dai, A., Yan, M., and Guibas, L. (2016). Volumetric and multi-view CNNs for object classification on 3D data. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2016, 5648-5656. doi:10.1109/CVPR.2016.609.

Qi, C., Yi, L., Su, H., and Guibas, L. (2017b). PointNet++: Deep hierarchical feature learning on point sets in a metric space. Advances in Neural Information Processing Systems, 2017, 5100-5109. https://arxiv.org/pdf/1706.02413.pdf. Accessed: 2021-01-12.

R Core Team (2018). R: A language and environment for statistical computing. R foundation for statistical computing. Vienna, Austria. https://www.R-project.org/. Accessed: 2021-01-12.

Reitberger, J. (2010). 3D-Segmentierung von Einzelbäumen und Baumartenklassifikation aus Daten flugzeuggetragener Full Waveform Laserscanner. Ph.D. thesis Technische Universität München. https://mediatum.ub.tum.de/doc/963556/963556.pdf. Accessed: 2021-01-12.

Reitberger, J., Krzystek, P., and Stilla, U. (2007). Combined tree segmentation and stem detection using full waveform lidar data. ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007, September 12-14, 2007, Espoo, Finland, 36, 332-337. https://www.isprs.org/proceedings/XXXVI/3-W52/final_papers/Reitberger_2007 .pdf. Accessed: 2021-01-12.

Reitberger, J., Krzystek, P., and Stilla, U. (2008). 3D segmentation and classification of single trees with full waveform lidar data. Proceedings of SilviLaser 2008, 8, 17-19. https://www.pf.bgu.tum.de/pub/2008/reitberger_co_stilla_silvi08_pap.pdf. Accessed: 2021-01-12.

Reitberger, J., Schnörr, C., Krzystek, P., and Stilla, U. (2009). 3D segmentation of single trees exploiting full waveform lidar data. ISPRS Journal of Photogrammetry and Remote Sensing, 64, 561-574. doi:10.1016/j.isprsjprs.2009.04.002.

Ren, H.-Y., Zhuang, D.-F., Pan, J.-J., Shi, X.-Z., Shi, R.-H., and Wang, H.-J. (2010). Study on canopy spectral characteristics of paddy polluted by heavy metals. Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy and Spectral Analysis, 30, 430-434. doi:10.3964/j.issn.1000-0593(2010)02-0430-05.

Repola, J. (2009). Biomass functions for Scots pine, Norway spruce and birch in Finland. Silva Fennica, 43, 625-647. https://www.silvafennica.fi/pdf/article184.pdf. Accessed: 2021-01-12.

Ørka, H., Gobakken, T., Næsset, E., Ene, L., and Lien, V. (2012). Simultaneously acquired airborne laser scanning and multispectral imagery for individual tree species identification. Canadian Journal of Remote Sensing, 38, 125-138. doi:10.5589/m12-021.

Rodriguez-Galiano, V., Ghimire, B., Rogan, J., Chica-Olmo, M., and Rigol-Sanchez, J. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing, 67, 93-104. doi:10.1016/j.isprsjprs.2011.11.002.

Rouse Jr, J., Haas, R., Schell, J., and Deering, D. (1973). Monitoring vegetation systems in the great plains with ERTS. Third ERTS Symposium, NASA, SP-351, 309-317. https://ntrs.nasa.gov/api/citations/19740022614/downloads/19740022614.pdf. Accessed: 2021-01-12.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A., and Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115, 211-252. doi:10.1007/s11263-015-0816-y.

Rusu, R. B., Blodow, N., and Beetz, M. (2009). Fast point feature histograms (FPFH) for 3D registration. In IEEE International Conference on Robotics and Automation (ICRA), pp. 3212-3217. doi:10.1109/ROBOT.2009.5152473.

Rusu, R. B., Marton, Z. C., Blodow, N., and Beetz, M. (2008). Learning informative point classes for the acquisition of object model maps. In 2008 10th International Conference on Control, Automation, Robotics and Vision (ICARCV), pp. 643-650. doi:10.1109/ICARCV.2008.4795593.

Saintenoy, A., Courbet, C., Nguyen, L., Léger, E., and Bugai, D. (2017). Results from two GPR surveys (2002 and 2015) in the Chernobyl Exclusion Zone. In 2017 9th International Workshop on Advanced Ground Penetrating Radar (IWAGPR), pp. 1-6. IEEE. doi:10.1109/IWAGPR.2017.7996094.

Scafutto, R., de Souza Filho, C., and Rivard, B. (2016). Characterization of mineral substrates impregnated with crude oils using proximal infrared hyperspectral imaging. Remote Sensing of Environment, 179, 116-130. doi:10.1016/j.rse.2016.03.033.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85-117. doi:10.1016/j.neunet.2014.09.003.

Shi, J., and Malik, J. (2000). Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, 888-905. doi:10.1109/34.868688.

Shi, Y., Skidmore, A., Wang, T., Holzwarth, S., Heiden, U., Pinnel, N., Zhu, X., and Heurich, M. (2018a). Tree species classification using plant functional traits from lidar and hyperspectral data. International Journal of Applied Earth Observation and Geoinformation, 73, 207-219. doi:10.1016/j.jag.2018.06.018.

Shi, Y., Wang, T., Skidmore, A., and Heurich, M. (2018b). Important lidar metrics for discriminating forest tree species in Central Europe. ISPRS Journal of Photogrammetry and Remote Sensing, 137, 163-174. doi:10.1016/j.isprsjprs.2018.02.002.

Shin, H.-C., Roth, H., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., and Summers, R. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Transactions on Medical Imaging, 35, 1285-1298. doi:10.1109/TMI.2016.2528162.

Simonyan, K., and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations (ICLR) – Conference Track Proceedings. https://arxiv.org/pdf/1409.1556.pdf. Accessed: 2021-01-12.

Sims, D. A., and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Remote Sensing of Environment, 81, 337-354. doi:10.1016/S0034-4257(02)00010-X.

Sithole, G., and Vosselman, G. (2004). Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 59, 85-101. doi:10.1016/j.isprsjprs. 2004.05.004.

Smith, L. (1997). Satellite remote sensing of river inundation area, stage, and discharge: A review. Hydrological Processes, 11, 1427-1439. doi:10.1002/(sici)1099-1085(199708)11:10<1427::aid-hyp473>3.0.co;2-s.

Soininen, A. (2016). TerraScan user's guide. https://www.terrasolid.com/download/ tscan.pdf. Accessed: 2021-01-12.

Solberg, S., Næsset, E., and Bollandsås, O. (2006). Single tree segmentation using airborne laser scanner data in a structurally heterogeneous spruce forest. Photogrammetric Engineering and Remote Sensing, 72, 1369-1378. doi:10.14358/PERS.72.12. 1369.

Stereńczak, K., Kraszewski, B., Mielcarek, M., and Piasecka, Ż. (2017). Inventory of standing dead trees in the surroundings of communication routes – the contribution of remote sensing to potential risk assessments. Forest Ecology and Management, 402, 76-91. doi:10.1016/j.foreco.2017.07.018.

Sun, Y., Huang, J., Ao, Z., Lao, D., and Xin, Q. (2019a). Deep learning approaches for the mapping of tree species diversity in a tropical wetland using airborne lidar and high-spatial-resolution remote sensing images. Forests, 10. doi:10.3390/F10111047.

Sun, Y., Xin, Q., Huang, J., Huang, B., and Zhang, H. (2019b). Characterizing tree species of a tropical wetland in southern china at the individual tree level based on convolutional neural network. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12, 4415-4425. doi:10.1109/JSTARS.2019.2950721.

Tang, L., and Shao, G. (2015). Drone remote sensing for forestry research and practices. Journal of Forestry Research, 26, 791-797. doi:10.1007/s11676-015-0088-y.

Thiry, Y., Colle, C., Yoschenko, V., Levchuk, S., Van Hees, M., Hurtevent, P., and Kashparov, V. (2009). Impact of Scots pine (Pinus sylvestris L.) plantings on long term 137Cs and 90Sr recycling from a waste burial site in the Chernobyl Red Forest. Journal of Environmental Radioactivity, 100, 1062-1068. doi:10.1016/j.jenvrad.2009. 05.003.

Tikhomirov, F., and Shcheglov, A. (1994). Main investigation results on the forest radioecology in the Kyshtym and Chernobyl accident zones. Science of the Total Environment, 157, 45-57. doi:10.1016/0048-9697(94)90564-9.

Towler, J., Krawiec, B., and Kochersberger, K. (2012). Radiation mapping in post-disaster environments using an autonomous helicopter. Remote Sensing, 4, 1995-2015. doi:10.3390/rs4071995.

Trier, Ø. D., Salberg, A.-B., Kermit, M., Rudjord, Ø., Gobakken, T., Næsset, E., and Aarsten, D. (2018). Tree species classification in Norway from airborne hyperspectral and airborne laser scanning data. European Journal of Remote Sensing, 51, 336-351. doi:10.1080/22797254.2018.1434424.

Usoltsev, V., Shobairi, S., and Chasovskikh, V. (2019). Comparing of allometric models of single-tree biomass intended for airborne laser sensing and terrestrial taxation of carbon pool in the forests of Eurasia. Natural Resource Modeling, 32. doi:10.1111/ nrm.12187.

Ussyshkin, R., Ravi, R., Ilnicki, M., and Pokorny, M. (2009). Mitigating the impact of the laser footprint size on airborne lidar data accuracy. American Society for Photogrammetry and Remote Sensing Annual Conference 2009 (ASPRS), 2, 854-862.

https://pdfs.semanticscholar.org/55e5/d6812d2e833c78efbd0a213ee276510
cf08c.pdf. Accessed: 2021-01-12.

Vetrivel, A., Gerke, M., Kerle, N., Nex, F., and Vosselman, G. (2018). Disaster damage
detection through synergistic use of deep learning and 3D point cloud features de-
rived from very high resolution oblique aerial images, and multiple-kernel-learning.
ISPRS Journal of Photogrammetry and Remote Sensing, 140, 45-59. doi:10.1016/
j.isprsjprs.2017.03.001.

Voulodimos, A., Doulamis, N., Doulamis, A., and Protopapadakis, E. (2018). Deep learn-
ing for computer vision: A brief review. Computational Intelligence and Neuroscience,
2018. doi:10.1155/2018/7068349.

Wegner, J., Branson, S., Hall, D., Schindler, K., and Perona, P. (2016). Cataloging public
objects using aerial and street-level images – urban trees. Proceedings of the IEEE
Computer Society Conference on Computer Vision and Pattern Recognition (CVPR),
2016-December, pp. 6014-6023. doi:10.1109/CVPR.2016.647.

Weinmann, M., Jutzi, B., and Mallet, C. (2014). Semantic 3D scene interpretation: a
framework combining optimal neighborhood size selection with relevant features. IS-
PRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sci-
ences, 2, 181-188. doi:10.5194/isprsannals-II-3-181-2014.

White, K., and El Asmar, H. (1999). Monitoring changing position of coastlines using
Thematic Mapper imagery, an example from the Nile Delta. Geomorphology, 29, 93-
105. doi:10.1016/S0169-555X(99)00008-2.

Widlowski, J., Verstraete, M., Pinty, B., and Gobron, N. (2003). Allometric Relationships
of Selected European Tree Species. Office for Official Publications of the European
Communities. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC26286/
EUR%2020855%20EN.pdf. Accessed: 2021-01-12.

Wijmans, E. (2018). PointNet++ PyTorch. https://github.com/erikwijmans/Pointnet2_Py
Torch. Accessed: 2021-01-12.

Wu, B., Yu, B., Wu, Q., Huang, Y., Chen, Z., and Wu, J. (2016). Individual tree crown
delineation using localized contour tree method and airborne lidar data in coniferous
forests. International Journal of Applied Earth Observation and Geoinformation, 52,
82-94. doi:10.1016/j.jag.2016.06.003.

Yao, W., Krzystek, P., and Heurich, M. (2012). Identifying standing dead trees in forest
areas based on 3D single tree detection from full waveform lidar data. ISPRS Annals
of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 1, 359-
364. doi:10.5194/isprsannals-I-7-359-2012.

YellowScan (2016). YellowScan Mapper – The lightweight LiDAR solution for UAVs.
https://www.microgeo.it/public/userfiles/Yellowscan-Mapper-datasheet.pdf.
Accessed: 2021-01-12.

Yoschenko, V., Kashparov, V., Melnychuk, M., Levchuk, S., Bondar, Y., Lazarev, M., Yo-
schenko, M., Farfán, E., and Jannik, G. (2011). Chronic irradiation of Scots pine trees
(Pinus Sylvestris) in the Chernobyl Exclusion Zone: Dosimetry and radiobiological
effects. Health Physics, 101, 393-408. doi:10.1097/HP.0b013e3182118094.

Yu, X., Hyyppä, J., Litkey, P., Kaartinen, H., Vastaranta, M., and Holopainen, M. (2017).
Single-sensor solution to tree species classification using multispectral airborne laser
scanning. Remote Sensing, 9. doi:10.3390/rs9020108.

Zabulonov, Y., Burtnyak, V., and Odukalets, L. (2017). System for effective remote con-
trol and monitoring of radiation situation based on unmanned aerial vehicle. Science
and Innovation, 13, 40-45. doi:10.15407/scine13.04.040.

Zabulonov, Y., Burtnyak, V., and Zolkin, I. (2015). Airborne gamma spectrometric survey in the Chernobyl exclusion zone based on oktokopter UAV type. Problems of Atomic Science and Technology, 99, 163-167. https://vant.kipt.kharkov.ua/ARTICLE/VANT_2015_5/article_2015_5_163.pdf. Accessed: 2021-01-12.

Zenáhlíková, J., Červenka, J., Čížková, P., Bečka, P., Starý, M., Marek, P., Křenová, Z., and Svoboda, M. (2015). The Biomonitoring project – monitoring of forest ecosystems in non-intervention areas of the Sumava National Park. Silva Gabreta, 21, 95-104. https://pdfs.semanticscholar.org/8070/5c4e3a6771212622a97c46ba96267246cd55.pdf. Accessed: 2021-01-12.

Zhao, R., Pang, M., and Wang, J. (2018). Classifying airborne lidar point clouds via deep features learned by a multi-scale convolutional neural network. International Journal of Geographical Information Science, 32, 960-979. doi:10.1080/13658816.2018.1431840.

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A. (2016). Learning deep features for discriminative localization. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2016, 2921-2929. doi:10.1109/CVPR.2016.319.

Zhou, Q.-Y., Park, J., and Koltun, V. (2018). Open3D: A modern library for 3D data processing. https://arxiv.org/pdf/1801.09847.pdf. Accessed: 2021-01-12.

Zhou, Y., and Tuzel, O. (2018). VoxelNet: End-to-end learning for point cloud based 3D object detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4490-4499. doi:10.1109/CVPR.2018.00472.

Curriculum Vitae

Sebastian Briechle was born on 16.09.1986 in Rosenheim, Germany, where he spent a happy childhood. After grammar school, he went to Australia in 2006 on a one-year work and travel visa. In 2007, he decided to study geodesy and geoinformation at the Technical University of Munich (TUM), where he received his BSc (2011) and MSc (2014) – including a specialisation in engineering surveying. From 2014 to 2017, he worked as a research engineer at the German Electron Synchrotron (DESY) in Hamburg, Germany. As part of his work, he developed a high-precision, laser-based measurement system for the European X-ray free-electron laser (XFEL) facility, a 3.4-km-long linear accelerator commissioned in 2017. Using this measurement system, the horizontal and vertical refraction effects in the accelerator tunnel were successfully determined and ensured the sub-millimetre accurate adjustment of the components. Since 2017, he worked as a research assistant at the Faculty of Geoinformation at Munich University of Applied Sciences (MUAS). As part of Peter Krzystek's team, his research focuses on developing and testing deep learning-based methods for segmenting and classifying objects from aerial images and 3D point clouds. His work also includes the calibration of laser scanners and MS cameras as well as accuracy studies. As an external PhD student in George Vosselman's group, he joined the Faculty of Geoinformation and Earth Observation (ITC) at the University of Twente (UT), Enschede, Holland, in 2018. Sebastian lives in Munich with his wife and young son.

Author's publications

Peer-reviewed papers

Briechle, S., Krzystek, P., and Vosselman, G. (**2021**). Silvi-Net – A dual-CNN approach for combined classification of tree species and standing dead trees from remote sensing data. International Journal of Applied Earth Observation and Geoinformation, 98, 2021, doi:10.1016/j.jag.2020.102292.

Briechle, S., Molitor, N., Krzystek, P., and Vosselman, G. (**2020**). Detection of radioactive waste sites in the Chornobyl Exclusion Zone using UAV-based lidar data and multispectral imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 167, 345-362. doi:10.1016/j.isprsjprs.2020.06.015.

Briechle, S., Krzystek, P., and Vosselman, G. (**2020**). Classification of tree species and standing dead trees by fusing UAV-based lidar data and multispectral imagery in the 3D deep neural network Point-Net++. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, V-2-2020 , 203-210. doi:10.5194/isprs-annals-V-2-2020-203-2020.

Decking et al. (**2020**). A MHz-repetition-rate hard X-ray free-electron laser driven by a superconducting linear accelerator. Nature Photonics, 14(6), 391-397. doi:10.1038/s41566-020-0607-z.

Conference proceedings

Krzystek, P., Molitor, N., Briechle, S., Bugai, D., Antropov, V., Tretiak, O., and Zabulonov, Y. (**2020**). UAV-based mapping of vegetation and contamination in the Chornobyl Exclusion Zone – new developments and future prospects. International Conference on Nuclear Decommissioning and Environment Recovery, INUDECO 2020.

Briechle, S., Krzystek, P., and Vosselman, G. (**2019**). Semantic labeling of ALS point clouds for tree species mapping using the deep neural network PointNet++. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(2), 951-955. doi:10.5194/isprs-archives-XLII-2-W13-951-2019.

Briechle, S., Sizov, A., Tretyak, O., Antropov, V., Molitor, N., and Krzystek, P. (**2018**). UAV-based detection of unknown radioactive biomass deposits in Chernobyl's Exclusion Zone. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(2), 163-169. doi:10.5194/isprs-archives-XLII-2-163-2018.