Hochschule
für Technik
Stuttgart
University of Applied Sciences

Master of Science Programme
Photogrammetry and Geoinformatics
Master Thesis
Winter Term 2022/2023

Assessing the Vitality of Urban Trees using Remote Sensing and Deep Learning

by

Javier Alejandro Vargas Gómez

Supervisors: Prof. Dr.-Ing. Michael Hahn
Prof. Dr. Thomas Esch

Assessing the Vitality of Urban Trees using Remote Sensing and Deep Learning

by

Javier Alejandro Vargas Gómez

A dissertation presented in partial fulfillment of the requirements for the degree of Master of Science in the Department of Geomatics, Computer Science and Mathematics, Stuttgart University of Applied Sciences

Declaration

The following Master thesis was prepared in my own words without any additional help. All used sources of literature are listed at the end of the thesis.

I hereby grant to Stuttgart University of Applied Sciences permission to reproduce and to distribute publicly paper and electronic copies of this document in whole and in part.

Stuttgart, 28.02.2023 Javier Alejandro Vargas Gómez

Approved by:

Prof. Dr.-Ing. Michael Hahn

Acknowledgement

I would like to acknowledge the Hochschule für Technik Stuttgart and DAAD for receiving me as student and providing me the opportunity to learn and improve my professional profile. The academic performance of the staff during the different education scenarios has been outstanding but also the personal experience of receiving classes from different Professors and interacting with national and international students has been an important aspect that I will always keep with me.

I would like to express my gratitude to my supervisors, Prof. Michael Hahn and Prof. Thomas Esch, for their support and critic observations during the research development. Their open attitude towards questions during the master courses and thesis work gave me the confidence to focus my study on relevant issues of the current times. I would also like to make a special acknowledgement to Dr. Fatemeh Alidoost for involving me and my colleagues in the Deep Learning topic and for her support during the development of the thesis project. Additionally, I would like to thank to Prof. Eberhard Gülch, Prof. Dietrich Schröder and Ms. Beate Baur for their support and observations during the early stage of the project.

The motivation and strength to never give up transmitted by my partner Alejandra was an important source for continuing everyday despite the obstacles found in the normal research process, and I want to give a very special acknowledgement to her. The rest of my family in Colombia and my father in the sky are also part of this acknowledgement and I dedicate this achievement to them, specially to my mother and father who made a big effort to educate me and shared to me the respect for the teaching profession and the love for the education.

This project would not have been developed without the support of several people. I would like to thank my PG colleagues for sharing different approaches, opinions and ease the discussion about academic topics and the life in Germany, specially to Diego Satian for his guidance and support through this work. To Leidy, Juan, Sourav, Sayantan, Shristi, Fithro, Lien, Abdalla, Johana, Andrea and Arturo, thank you so much.

The development of this thesis work would not be possible without the information provided by the Landeshauptstadt Hannover and its Department of Environment and Urban Greenspace, specially to Ms. Caroline Bank and Ms. Angelika Kreuzer. Additionally, the datasets provided by Mr. Dirk Prause from the LBEG were very important for the analysis of the information. The sources and help provided by Mr. Simon Freund and Mr. Mattias Schulz-Merkel from HFT staff were also very relevant and I also thank them for their support and good disposition.

I would like to make an additional acknowledgement to Eng. Edgar Chacon for sharing to me his professionalism, experience, and enjoyment for the activities in the forestry sector. Here, I express a great gratitude to him and his company. Finally, I would like to thank my hometown University, Universidad del Valle, and its academic program of Agricultural Engineering, for preparing me to face the life in a personal and academic way, and for giving me the most valuable tools for facing the different challenges in the master course.

Master Course Photogrammetry and Geoinformatics

Assessing the Vitality of Urban Trees using Remote Sensing and Deep Learning

Abstract

The use of Convolutional Neural Networks (CNN) has been widely implemented in forestry-related tasks as species classification, crown detection and mortality identification. The usage of several sources as images, point clouds and elevation models have generated relevant results in different forested areas, but unfortunately these studies have not been focused on urban trees. Therefore, the objective of this study is to investigate the performance of CNN for classifying the vitality of urban trees, which are increasingly affected and stressed by the Urban Heat Island Effect. Aerial and Sentinel-2 images are sampled for feeding the CNN model. The prediction of the vitality classes shows a precision of 74,69%, especially for the most represented class (healthy trees). The achieved results allow to better understand the performance of a CNN network for determining the vitality of trees in an urban context where diversity of vegetation patterns can represent a big challenge for classification tasks.

Keywords: Aerial, Satellite, Imagery, CNN, Urban Trees, Vitality.

Table of Content

Acknowledgement ..... **ii**

Abstract **1**

Table of Content **2**

List of Figures **4**

List of Tables **6**

Abbreviations **7**

1 Introduction **8**
1.1 Motivation 8
1.2 State of the art 9
1.3 Objectives 14
1.4 Report structure 14

2 Theoretical Background **15**
2.1 The Role of Urban Trees 15
2.2 Heat Stress and Vitality of Urban Trees 16
2.3 Convolutional Neural Networks and Computer Vision models 21

3 Methodology **25**
3.1 Study area 25
3.2 Data sources 26
3.2.1 Tree inventory (Baumkataster) 26
3.2.2 Multispectral images 26
3.2.3 Environmental variables 28
3.2.4 TreeSatAI Benchmark Archive 31
3.3 Materials and Tools 32
3.3.1 Virtual Machine (VM) 32
3.3.2 Vector and Raster datasets 32
3.3.3 Codes and Scripts 32
3.4 Processings and Analyses 33
3.4.1 Data processing 33
3.4.2 Filtering of tree inventory 35
3.4.3 Data labeling 38
3.4.4 Image patches sampling 39
3.4.5 Input data preparation 40

3.4.6 Training, internal validation and testing 41

4 Results and Discussion 43

4.1 Influence of Input Data 43
4.1.1 Urban trees and environmental conditions 43
4.1.2 Aerial and S2 imagery 46
4.1.3 Aerial image patches 50

4.2 Performance of Classification Model 51
4.2.1 Training and Validation 51
4.2.2 Testing 59

5 Conclusions and Recommendations 63

References 67

Appendices 76

List of Figures

Figure 1. Heat Island Effect representation. Taken from DWD (2022). 16
Figure 2. Site restoration of street tree in Hannover. Taken from Hannover Division of Environmental Protection (2017). 18
Figure 3. Spectral response of healthy and unhealthy trees. Taken from EOS (2023). 19
Figure 4. NDVI of urban trees in the capital city of Hannover. 19
Figure 5. Healthy tree (left) and non-healthy tree (right) with natural color band combination (a,b), CIR band combination (c,d), and NDVI (e,f). Source: Landesamt für Geoinformation und Landesvermessung Niedersachsen - LGLN (2022). 20
Figure 6. CNN architecture representation. Taken from Mathworks (2022). 21
Figure 7. Network architecture of ResNet (with 34 parameter layers). Adapted from He *et al.* (2015). 23
Figure 8. Architecture of ViT model. Taken from Dosovitskiy *et al.* (2020). 23
Figure 9. Location of the study area. Adapted from Geodatenzentrum (2023) & Geodatenportal Niedersachsen (2022). 25
Figure 10. Open Geo Data portal for orthophotos acquisition. Adapted from LGLN (2022). 27
Figure 11. Copernicus Open Access Hub for satellite images acquisition. Adapted from ESA (2022). 28
Figure 12. Air Temperature for the city of Hannover. Adapted from Hannover Department of Environment and Urban Green Space (2022b). 29
Figure 13. Heat Island Effect for the city of Hannover. Adapted from Environmental Planning and Management Unit of Hannover & GEONET (2017). 29
Figure 14. Soil moisture for the city of Hannover. Adapted from LBEG (2022). 31
Figure 15. Aerial orthomosaics (left) and S2 images (right) of the city of Hannover for the year 2019 in true color band combination. Source: LGLN (2022); ESA (2022). 33
Figure 16. Aerial orthomosaics (left) and S2 images (right) of the city of Hannover for the year 2022 in true color band combination. Source: LGLN (2022); ESA (2022). 34
Figure 17. NDVI of 2019 (left) and 2022 (right) for the city of Hannover. Adapted from LGLN (2022). 35
Figure 18. Selection of trees of classes 1,2 3 and 4 in areas of Heat Island Effect. Source: Department of Environment and Urban Green Space (2022ab). 36
Figure 19. Examples of selected trees for each class: H (left), LV (middle) and DV (right). 37

Figure 20. Square buffers generated for each selected tree. Source: LGLN (2022); Hannover Department of Environment and Urban Green Space (2022a). 38
Figure 21. Crown polygon digitized for each selected tree. Source: LGLN (2022). 39
Figure 22. Label file (.json) generation. 39
Figure 23. Patch generation process modeled in ArcGIS Pro (Model Builder). 40
Figure 24. Image patches generated for aerial (left) and Sentinel 2 (right) sources (File name: H__1_98616__Y2019.tif). 40
Figure 25. Number of trees of classes H (Healthy), LV (Limited Vitality), and DV (Deteriorated Vitality). 43
Figure 26. Number of trees of classes H, LV and DV per Urban Heat Island zone. 44
Figure 27. Maximum Soil Moisture per tree in areas with Low Heat Island Effect. 45
Figure 28. Minimum Soil Moisture per tree in areas with Moderated Heat Island Effect. 45
Figure 29. Aerial orthomosaics of the years 2019 (left) and 2022 (right) in CIR band combination. 47
Figure 30. S2 images of the years 2019 (left) and 2022 (right) in CIR band combination. 48
Figure 31. Histogram of RGB (left) and NIR (right) bands for the 2019 aerial orthomosaic. 48
Figure 32. Histogram of RGB (left) and NIR (right) bands for the 2022 aerial orthomosaic. 50
Figure 33. Image patch of DV, LV and H trees (left) and spectral profile (right). 50
Figure 34. Number of trees per crown area in image patches of training dataset. 53
Figure 35. Training and Validation loss curves for the experiments with only aerial images (left) and aerial with S2 images (right). 57
Figure 36. H label correctly predicted as class H (left) and incorrectly predicted as class LV (right). 60
Figure 37. LV label correctly predicted as class LV (left) and incorrectly predicted as class DV (right). 61
Figure 38. DV label correctly predicted as class DV (left) and incorrectly predicted as class LV (right). 62
Figure 39. Status of tree trunk (Favoritepark, Ludwigsburg). 65
Figure A 1. Workflow of Methodology Implementation. 79

List of Tables

Table 1. Research papers and related investigations about tree classification and monitoring..... 10
Table 2. Tree vitality classes. Taken from Hannover Department of Environment and Urban Green Space (2022a). 26
Table 3. Number of trees per vitality class and Heat Island Effect zones. Adapted from Environmental Planning and Management Unit of Hannover & GEONET (2017) and Department of Environment and Urban Green Space (2022b). 30
Table 4. Selected trees and corresponding location in the Urban Heat Islands. 37
Table 5. Dataset split. 41
Table 6. Drought tolerance of some of the tree species from Lower Saxony taken in count in the current study. Adapted from Brune (2016)..... 46
Table 7. Hyperparameters selected for the training and validation steps.. 51
Table 8. Performance of the ResNet-18 model for predicting the different health status of the training dataset (only Aerial). 52
Table 9. Performance of ResNet-18 and ViT models for predicting the different health status of the training dataset (Aerial+S2). 52
Table 10. Performance of model with only aerial inputs for predicting the different health status of trees on the testing dataset. 59
Table 11. Performance of model with aerial and S2 inputs for predicting the different health status of trees on the testing dataset. 59
Table A 1. Number of tree species per vitality class and Urban Heat Island zone..... 76
Table A 2. Soil moisture percentage per tree species, vitality class and Urban Heat Island zone. 77

Abbreviations

BCE	Binary Cross-Entropy
CHM	Canopy Height Model
CNN	Convolutional Neural Networks
DL	Deep Learning
DV	Deteriorated Vitality
DWD	Deutscher Wetterdienst
ES	Ecosystem Services
ESA	European Space Agency
FAO	Food and Agriculture Organization of the United Nations
H	Healthy
ICP Forests	United Nations International Cooperative Programme of the Executive Committee for the Convention on Long-range Transboundary Air Pollution in Europe
IPCC	Intergovernmental Panel on Climate Change
LBEG	Landesamt für Bergbau, Energie und Geologie
LGLN	Landesamt für Geoinformation und Landesvermessung Niedersachsen
LR	Learning Rate
LV	Limited Vitality
ML	Machine Learning
NIBIS	Portal für die Geodaten des Niedersächsischen Bodeninformationssystems
NASA	National Aeronautics and Space Administration
NN	Neural Networks
NDVI	Normalized Difference Vegetation Index
UNU	United Nations University

1 Introduction

1.1 Motivation

Urban heat waves have been an issue augmented yearly by Climate Change and a problem that has arranged a series of fires events and extreme temperatures in cities and rural areas, especially in Europe (European Space Agency - ESA, 2023b). Those situations have affected the welfare of human beings, mostly in recent years. Urban planning techniques have taken in count different measures to adapt the cities to the current challenges of global warming as green roofs, greening of tram tracks and buildings, shading the building facades with trees, and improving the public spaces towards cool spots (European Climate Adaptation Platform Climate-ADAPT, 2022). However, one of the first lines of defense against weather variations, the urban trees, can been affected by the weather itself but also by the urban and natural environmental conditions that surround them. Drought events can generate defoliation (Češljar, *et al.*, 2022; Gazol & Camarero, 2022) and warning levels of defoliation can be considered between 10% and 25% according to the ICP Forests (n.d, cited by Eurostat, 2022). The grey surfaces keep the heat provided by the sun and generate adverse effects on trees affecting the transpiration processes and thus their cooling potential (Schwaab *et al.*, 2021). Heat stress can affect the soil-water relation when the micropores that usually store the water are destroyed by compaction and then could limit the root growth (Jim, 2019). The negative effects on water balance can imply the occurrence of more and higher heat waves during the coming summer seasons (Umwelt Bundesamt, 2022). The role of specific trees in urban areas are related to the landscape approach for urban beautification by local administrations, but also they play an important role for temperature regulation in a joint work with soils and lower vegetation, especially when cities are experiencing warmer summers. In the capital city of Hannover, it is expected to have a rising of the number of very hot days and tropical nights, especially in the districts with block and block-edge building (Hannover Division of Environmental Protection, 2017).

For facing the weather variation issues related to Heat Island Effects and tree hydric stress, the city of Hannover implemented the “Adaptation Strategy to Climate Change for the Capital City Hannover”, a program that consisted of different strategies for fighting the Global Warming during the period 2012-2016. Among this local initiative, the strategies *Climate-adapted Vegetation* and *Preventive Soil and Groundwater Protection* were focused on the implementation of several activities as replanting, selection of tree species according to thermally burdened zones, site restoration of tree grates for root strengthening, and irrigation measures (Hannover Division of Environmental Protection, 2017). For complementing these activities, the implementation of Deep Learning (DL) techniques can contribute to detect health issues of trees regarding heat stress and then provide a timely alert to supply the irrigation and land management for the affected individuals. Neural Networks (NN) can imply a route of systematic identification of the issues related to Heat Island Effect and a path for formulating mitigation actions against the Climate Variation impact on cities. The proposed work aims to set a prediction mechanism of tree vitality that could be affected by Heat Island Effect in the city of Hannover, implementing aerial, satellite imagery and environmental information that helps to analyse the performance of the model for predicting ground truth classes. Furthermore, the result of the study also aims to provide sources and recommendations to approach different studies related to urban trees as species classification and mortality detection. The proposed work is aligned to the Sustainable Development Goals, mainly the Goals 3 (Good Health and Well-Being), 11 (Sustainable Cities and Communities), 13 (Climate Action), and 15 (Life on Land).

1.2 State of the art

The use of satellite, airborne and UAV products as Orthomosaics and LiDAR together with DL techniques have been a recent approach for complex studies in urban and rural environments. A group of recent research papers was summarized according to the focus of the current work, including some papers related to another ML techniques as well as visual methods related to the assessment of tree species, health status detection, and land surface temperature (Table 1).

Table 1. Research papers and related investigations about tree classification and monitoring.

Research Title	Authors (Year)	Methods	Findings and Results
Classification of urban tree species using multi-features derived from four-season RedEdge-MX data	Liu (2022)	The research was aimed to investigate the effectiveness of RedEdge-MX products (spectral, texture and DSM) for identifying tree species during a growing phase, using Maximum Likelihood Classification model and Random Forest Models.	In the investigation of the use of spectral, texture, and spectral+texcture+DSM, the accuracy of recognizing trees during flowering and leafing period was 52,98%, 86,66% and 86.90%, respectively.
The Auto Arborist Dataset: A Large-Scale Benchmark for Multiview Urban Forest Monitoring Under Domain Shift	Beery *et al.* (2022)	The paper introduces a large-scale dataset that contains 2,5 million (344 genera) of 512x12 pixel aerial images (5 cm) and street level images from 23 cities of North America. The dataset was used to train a ResNet-101 model for generalization tasks, making different experiments (regional, single city, and full dataset).	The Average Recall improvement of 21,3%, was achieved when training from single city approach to full dataset, and 18,3% when training on a region compared to a single city. Also, they noted that big cities tend to generalize well on the average while most of small cities had a poor performance.
TreeSatAI Benchmark Archive: A multi-sensor, multi-label dataset for tree species classification in remote sensing	Ahlswede *et al.* (2022)	A dataset composed of images from three different sources (aerial, Sentinel-1 and Sentinel-2) was used to generate 50.381 image patches for each source that contained information of 20 European tree species derived from forest administration data of Lower Saxony. The research aimed to test DL models (ResNet and Multi-Layer Perceptron-MLP) and one Machine Learning model (Light Gradient Boosting Machine-LightGBM) for classification tasks of forest species.	ResNet-18 model showed precision scores up to 79% only using RGB bands of aerial imagery. The use of Sentinel imagery slightly improved the weighted precision, mainly with the combination of ResNet and LightGBM. The implementation of Sentinel-2 and aerial images improve the performance of the classification in most of the experiments compared to the use of Sentinel-1 and both Sentinel products.
Automated remote sensing forest inventory using satellite imagery	Shtanchaev *et al.* (2021)	Using WorldView-2 satellite imagery, the training of classical Machine Learning algorithms was carried out with tree crowns embeddings generated by Autoencoders (Convolutional and Sparse) for forest inventory (tree species classification). They tried to compare the traditional CNN end-to-end classifiers (2D-CNN and VGG network).	VGG Network showed a higher accuracy and F1 score compared to the 2D-CNN. Smaller image sizes lead to better classification results.
The role of urban trees in reducing land surface temperatures in European cities	Schwaab *et al.* (2021)	Comparison of temperature differences between urban trees, treeless urban green spaces and urban fabric using topographic data, Land Surface Temperature (LST) and Land Use/Land Cover (LULC) high resolution data (Landsat and	Urban trees were related to 2-4 times higher reductions of surface temperature compared to treeless urban green spaces. Furthermore, urban trees reduced the surface temperature more than rural forests in Central European regions.

Research Title	Authors (Year)	Methods	Findings and Results
		Aster imagery). The research was aimed to calibrate Generalized Additive Models for prediction of temperature differences in the mentioned covers.	
Explainable identification and mapping of trees using UAV RGB image and deep learning	Onishi, M. & Ise, T. (2021)	Using of a slope model and segmented UAV imagery (into tree crown objects) as inputs of an object-based CNN for classifying each crown image into tree types and tree species.	Classification of tree types and specific species with more than 90% of accuracy. CNN classified trees according to their shapes and leaf contrasts.
Assessing a novel modelling approach with high-resolution UAV imagery for monitoring health status in priority riparian forests.	Guerra *et al.* (2021)	It was considered four categories of tree health status (asymptomatic, dead, defoliation>50% and defoliation<50%). Multispectral UAV imagery was analysed using classical random forest and logistic regression. A set of vegetation indices (including NDVI) where considered as well. Structural data from DSM at crown level was derived from initial data.	An overall accuracy of 67% was achieved for the four categories of health status considered. The accuracy improved when were taken less categories like asymptomatic, defoliated, dead (72%), or just alive or dead (91%). The results suggested that the red band was more relevant for detecting defoliation caused by pathogens.
Deep Neural Networks with Transfer Learning for Forest Variable Estimation Using Sentinel-2 Imagery in Boreal Forest	Astola *et al.* (2021)	The research aimed to predict growing stock volume on a Boreal Forest of Finland using Deep Neural Networks. As sources, the researchers used Sentinel-2 images, topography data and Canopy Height Model (CHM), as well as a forest inventory data for model training and evaluation.	The researchers found that the solely use of RGB+NIR bands, imaging and sun angles, and topography features showed the best plot level accuracy (RMSE%=42,6%). The use of CHM and S2 images reduced the relative RMSE to 28,6%-30,7%.
Mapping Urban Tree Cover Changes Using Object-Based Convolution Neural Network (OB-CNN)	Timilsina *et al.* (2020)	An object-based CNN was used to map the urban tree cover changes between 2015 and 2016. Sources as Google Earth images and LiDAR products were introduced in the model. Samples from datasets were extracted using NDVI and CHM. The research also involved socioeconomic variables to model the tree cover changes with regression analysis.	The object-based CNN generated an overall accuracy of 98% and kappa coefficient of 0,93 for the year 2015/16. The results also demonstrated positive correlation between household income and tree cover loss.
UAV Laser Scans Allow Detection of Morphological Changes in Tree Canopy	Slavík *et al.* (2020)	Usage of multitemporal series data from UAV laser scanning for detecting the phenomenon of bending branches of dead trees during one year. The researchers implemented automatic segmentation of individual trees to observe morphological changes.	Changes detected in angle inclination of branches that affects the solar radiation interception. Branch shifting in dead trees could be related to mycorrhizal interconnections between live and dead trees. It was correctly identified an 86% of trees that displayed

Research Title	Authors (Year)	Methods	Findings and Results
			branch movement, as recorded by a human observer.
Plant Disease Classification: A Comparative Evaluation of Convolutional Neural Networks and Deep Learning Optimizers	Saleem *et al.* (2020)	A comparative evaluation for classification of plant disease was developed using several CNN architectures (AlexNet, OverFea, VGG-16, ZFNet, ResNet-50, Inception ResNet-v2, Inception-v4, MobileNet, DenseNet-121, and Xception). 26 different diseases belonging to 14 plant species were used as input.	The Xception architecture trained with the Adam optimizer yielded the highest validation accuracy (99,81%).
Individual tree species identification using Dense Convolutional Network (DenseNet) on multitemporal RGB images from UAV	Natesan *et al.* (2020)	Implementation of a deep CNN (DensNet) to classify forest tree species at the individual tree-level, using UAV products (RGB images and LiDAR) collected over three years in summer (leaf presence) and fall (no leaf presence) seasons, and capturing the variations in seasonal conditions, foliage density and greenness. The network was trained and cross-validated using the labelled dataset for learning from the tree crowns on images.	Classification accuracy over 84% for distinguishing five predominant species of coniferous trees.
Tree Species Classification and Health Status Assessment for a Mixed Broadleaf-Conifer Forest with UAS Multispectral Imaging	Abdollahnejad & Panagiotidis (2020)	DTM, DSM and orthophotos were derived from photogrammetric processed datasets and used as input for canopy spectral analysis and textural analysis as well. The study was divided in two phases. Firstly, trees were classified into two groups (broadleaf or conifer). Then, trees were classified according to type and health status creating subgroups to detail the classification.	An overall accuracy of 81,18% for the proposed method of classification and 84,71% for health status assessment. The combined use of vegetation indices and texture analysis increased the overall accuracy by 4,24%.
Detection of Fir Trees (*Abies sibirica*) Damaged by the Bark Beetle in Unmanned Aerial Vehicle Images with Deep Learning	Safonova *et al.* (2019)	Aerial images from Central Siberia were used to detect damages on forest cover by Bark Beetle using a proposed CNN. Four tree damage categories were defined: (1) healthy, (2) colonized by beetles, (3) recently died, (4) deadwood. Fifty (50) image patches of single trees for each category were manually sampled. A total number of 200 image-patches were generated and an 80%-20% split was used for training and validation, respectively.	A test accuracy of 99.7% and a loss lower than 0.001 was achieved at the 23th training epoch. The proposed model achieved a high F1 score of 92,75%, 89,86%, 89,66% and 88,89% on damage classes 1,2,3, and 4, respectively. They explained that these results can be explained by the skills of the model for distinguishing color, shape, and texture on those classes.
Forest Damage Assessment Using	Hamdi *et al.* (2019)	A modified version of U-Net architecture was used for a pixelwise classification with RGB-NIR orthophotos (0,2 m) of Bavarian forests	The model performance achieved an overall accuracy of 92% in the test dataset. They noted that some areas

Research Title	Authors (Year)	Methods	Findings and Results
Deep Learning on High Resolution Remote Sensing Data		as inputs. The research implemented the CNN in an ArcGIS environment for automatic detection and mapping damaged areas by storms.	with shadows were not detected by the network. Additionally, small and damaged areas were not labeled.
A Convolutional Neural Network Classifier Identifies Tree Species in Mixed-Conifer Forest from Hyperspectral Imagery	Fricker *et al.* (2019)	The researchers used seven dominant tree species and dead standing trees in a mixed-conifer forest in California (US) identified in aerial images, for feeding a CNN with the purpose of classifying and mapping tree species. Additionally, they used LiDAR data to identify trees with more than 5m of height for applying the classifier to the tree crown.	The model correctly classified 713 individual trees. For all the species, average F-scores was 0,87 for hyperspectral CNN model and 0,64 for the RGB-fed model.
GIS-based analysis of the tree health problems using UAV images and satellite data	Asenova, M. (2018)	UAV optical imagery was implemented as main source for visual-analytical methods. Affected areas and existing biotic, abiotic and anthropogenic factors were identified. On-site reviews were carried out for identifying pathological damages which included leaf system, branches, and stems. The damage level was assessed and the actions for limiting the pest effects were formulated.	Trees affected by pests and pathogens were recognized by the pronounced color, shape and density changes in crowns. Scot pine plantations affected by bark beetle (96%) and wind throw/wind fall (4%). Vectorization of damaged area and the required measures for protecting the forest were carried out.

As seen in the Table 1, most of the research works related to the implementation of DL approaches have been focused on forests for different tasks as species classification (Fricker *et al.*, 2019; Abdollahnejad & Panagiotidis, 2020; Natesan *et al.*, 2020; Onishi, M. & Ise, T., 2021; Shtanchaev *et al.*, 2021; Ahlswede *et al.*, 2022), variable estimation (Astola *et al.*, 2021; Onishi, M. & Ise, T., 2021), detection of change in tree properties (Slavík *et al.*, 2020), assessment of health status and damage detection (Hamdi *et al.*, 2019, Safonova *et al.*, 2019; Abdollahnejad & Panagiotidis, 2020; Guerra *et al.*, 2021). However, the study of urban trees with DL and Remote Sensing techniques has been less represented in the research field, with tasks as detection and classification of tree species (Beery *et al.*, 2022; Liu, 2022) and mapping of urban tree cover (Timilsina *et al.*, 2020). It is specially missed the works that classiffy of health status of urban trees as done with forest trees, in order to develop faster monitorings in cities that contribute to a quicker detection of non-healthy trees and reduce the number of on-site campaigns. The classification of the tree vitality through multispectral image patches aimed in this

study could provide more information about the implementation of DL techniques in urban contexts and contribute to assess the effects of the Climate Variability on trees that strenghten and speed up the response of local authorities.

1.3 Objectives

- To identify Deep Learning techniques for estimating the vitality of urban trees located in Urban Heat Islands through aerial and satellite imagery.

- To define and adapt methods for sampling multispectral aerial images that include trees with different vitality characteristics.

- To analyse the influence of spatial and spectral properties of urban trees on the model performance.

1.4 Report structure

The following work is structured as follows:

Chapter 2: Provides an overview of the *Theoretical Background* of the work, including the role of urban trees, the influence of heat stress in their vitality and the models used in the study.

Chapter 3: Specifies the sources, tools and methods that are part of the *Methodology* implementation.

Chapter 4: Includes the *Results and Discussion* where is analysed the influence of input sources in the study and the performance of the model.

Chapter 5: Presents the *Conclusions* of the study and the *Recommandations* provided for future works.

2 Theoretical Background

2.1 The Role of Urban Trees

Forests play an important role for cooling the planet Earth and provide different benefits to the human beings, producing and regulating the world's temperatures and fresh water flows through a combined effect of transpiration and shade (Ellison *et al.*, 2017). Regarding urban forests, trees provide different Ecosystem Services (ES) to the local communities. Some of these services and benefits are summarized by the Food and Agriculture Organization of the United Nations - FAO (Borelli, 2016) and the Hannover Division of Environmental Protection (2017) as follows: thermal regulation, filtering of pollutants and fine particles, rising of air moisture, CO2 absorption, water flow regulation and water quality improvement, wood and food supply for cooking and heating, improvement of mental and physical health, increasing of urban biodiversity and property monetary value. Among these services, the cooling effect and the provision of shade are especially important to counteract the Heat Island Effect and thermal stress (Locatelli, 2016; Hannover Division of Environmental Protection, 2017).

Trees have also a synergic and feedback relationship with soils. While trees offer continuously organic matter, stability, aggregation, and cover against water erosion and direct sunlight, soils offer a place to settle, stores water, contains microbiota that retain and transform compounds into available nutrients for the trees, among others. Both sources conform a strong phase of many cycles as N, C, P, K, H_2O. Soil temperature can influence the life functions and the contents of Soil Organic Carbon (SOC) that directly affects the water flow in the soil profile, and therefore in the tree hydration, nutrition, and cooling of surface (Sun & Pinker, 2004; Liu & Pu, 2019; Cates, 2022). However, there are different limitations for trees in urban environments, mainly given by poor physical properties as a very high density (soil compaction) and enhanced by insufficient soil volume (Jim, 2019) that together could address different problems as poor hydraulic conductivity, poor nutrients transportation and finally the mortality.

2.2 Heat Stress and Vitality of Urban Trees

❖ Heat Island Effect

The Heat Island Effect or Urban Heat Island is defined as an urban climate feature that is characterized by maximum temperatures on night time due to the differences in air temperature between the urban core (warmer) and the city surroundings (cooler) (Intergovernmental Panel on Climate Change - IPCC, 2022; Deutscher Wetterdienst - DWD, 2023). These differences on temperatures are influenced mainly by the building geometries, material properties of the urban surfaces, human activities as traffic and artificial cooling, density and size of urban infrastructure, sealing level and vegetation cover (Hannover Division of Environmental Protection, 2017; IPCC, 2022; DWD, 2023). The Figure 1 shows the dynamic of the Heat Island Effect and its relationship with the different climatological properties as air temperature, wind speed and sunlight exposure:

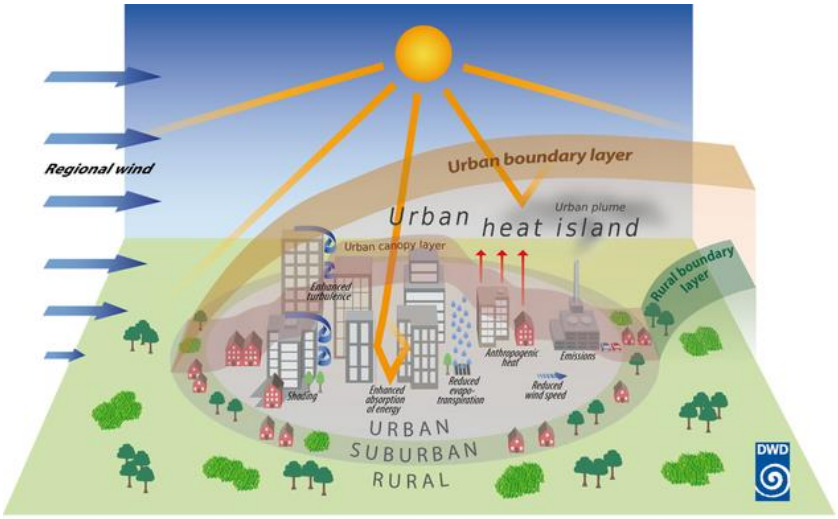

Figure 1. Heat Island Effect representation. Taken from DWD (2022).

Among the greatest impacts of the Heat Island Effect are the increasing of deaths and disease rates for all populations, especially in elder and non-healthy communities (Tan & Siri, 2023; Hannover Division of Environmental Protection, 2017). Besides, the higher temperatures reached during the night could affect the internal process of vegetation species as the biomass generation. Jing *et al.* (2016) found that high night temperatures addressed a reduction on biomass allocation

to reproduction organs, with an acceleration on ecophysiological processes and leaf growth stimulation in urban vegetation.

Urban trees are surrounded of a different environment compared to forests. While urban trees usually share the surface with buildings and houses, streets, monuments, water bodies, among others, forest trees are surrounded by communities from the same or different tree species, lakes and natural rivers (no channelized), just to mention a few. The built-up characteristics generate or modify the weather and environment dynamics, mainly about air temperature, sunlight exposure, wind speed, soil density and pollution. Those environmental conditions under the urban heat island can cause vitality loss and increase of mortality risk (Gillner *et al.* 2014). The existence of extreme or abnormal air temperatures at night affect the metabolism and vitality of trees (Roloff, 2019; Jing *et al.*, 2016).

To overcome the consequences of Climate Change, the city of Hannover implemented the “Adaptation Strategy to Climate Change for the Capital City Hannover”, a program aimed to finance the different strategies for fighting the Global Warming during the period 2012-2016. Important strategies in this initiative were called *Climate-adapted Vegetation* and *Preventive Soil and Groundwater Protection*, that were focused on the implementation of several activities as replanting, selection of tree species according to thermally burdened zones, site restoration of tree grates for root strengthening, and irrigation measures (Hannover Division of Environmental Protection, 2017). As reference, the costs for site restoration/preparation of tree roots and new planting can rise 1250 € and 3500 per tree respectively (Figure 2). In Lower Saxony, there are some species tolerant to heat stress, as *B. Pendula*, *F. excelsior*, *Q. robur*, and *A. platanoides*. But also other species could be very sensitive to drought stress as *A. glutinosa*, *P. nigra* and *S. alba* (Brune, 2016).

Figure 2. Site restoration of street tree in Hannover. Taken from Hannover Division of Environmental Protection (2017).

❖ Tree Vitality

Tree vitality can be defined as the current health status of a tree species in line with the vigor and vital energy displayed, and, according to Callow *et al.* (2018), it is related to the plant response to physiological stress. Tree vitality is usually assessed by trained arborists who visually evaluate parameters as crown size and density, leaf color and vigor, while other methods involve biochemical and physical laboratory techniques (Fite, 2008; Martinez-Trinidad *et al.*, 2010; Callow *et al.*, 2018). Crown density is often the most common parameter taken in count to identify problems in tree vitality, and defoliation is a derived observation from tree monitoring in rural and urban forestry.

The study of urban forests usually involves traditional methods of tree inventory and dendrology studies carried out by local institutions with the purpose of updating the database of trees of a city and also for developing activities of monitoring, maintenance or replacement. In the case of forest inventories, it should be taken in count the difference between Inventory and Census. A forest inventory takes in count a sample of the study area while a forest census takes in count 100% of the trees located in an entire study area, however the first one does not take in count the entire tree individuals and the second one is time demanding, even more if the areas are large or have strong slopes. Remote Sensing and DL ap-

proaches can mean a fast and adjustable way to develop forest studies and monitoring activities that contribute to the timely detection of health issues or diseases that can imply a risk for the trees, and therefore for the community that benefit from the services provided by the trees. The Normalized Difference Vegetation Index (NDVI) can provide a good source to study urban trees according to their health state as shown in the Figure 3, Figure 4, and Figure 5.

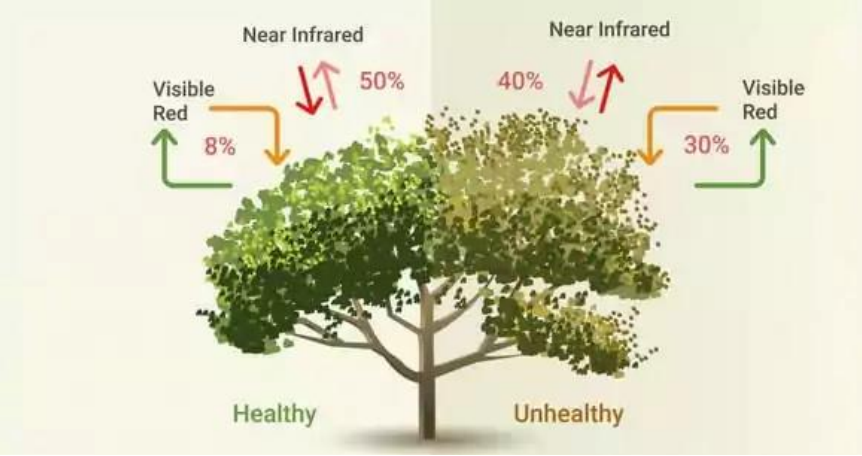

Figure 3. Spectral response of healthy and unhealthy trees. Taken from EOS (2023).

Figure 4. NDVI of urban trees in the capital city of Hannover.

Defoliated trees do not show an appropriate NDVI but the CIR combination (IR-R-G) can provide a good approach for visualize the vegetation structures that remains on trees as trunk and main branches (Figure 5d). In specific seasons, trunk and branches can be covered by non-vascular epiphytes (e.g. green moss) and NIR spectrum could be reflected of the plant easing the identification of tree crowns.

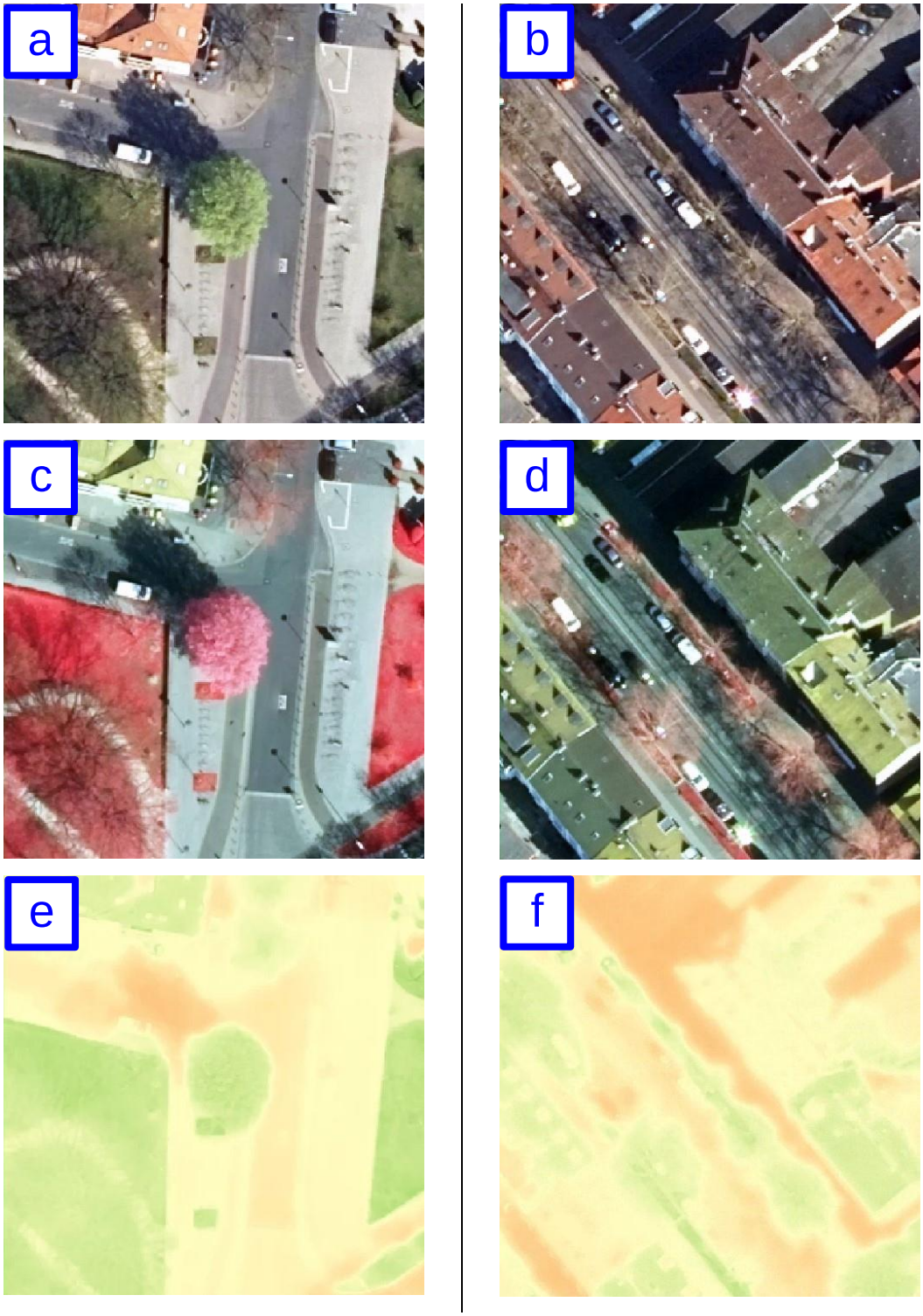

Figure 5. Healthy tree (left) and non-healthy tree (right) with natural color band combination (a,b), CIR band combination (c,d), and NDVI (e,f). Source: Landesamt für Geoinformation und Landesvermessung Niedersachsen - LGLN (2022).

2.3 Convolutional Neural Networks and Computer Vision models

❖ Convolutional Neural Networks (CNN)

CNNs or ConvNets are network architectures that learn directly from data and can contain tens or hundreds of layers that contribute to detect features of an image (Mathworks, 2023). CNNs are composed by an input layer, several hidden layers and an output layer, as shown in the following image:

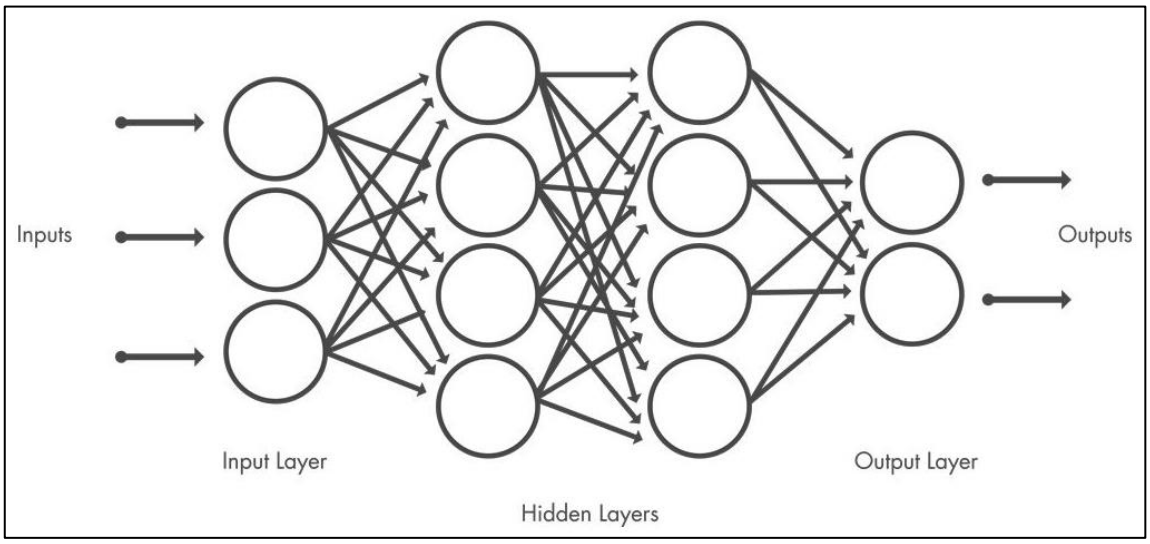

Figure 6. CNN architecture representation. Taken from Mathworks (2022).

Input, output and hidden layers are required to perform different operations that alter the data properties and learn specific features. One important part is the hidden layers that, according to Mathworks (2023), can be classified in three main types of layers: convolutional layer, rectified linear unit layer (ReLu), and pooling layer. Convolutional layers compose the core building block of the CNN and are the phase where most of the computations occur (i.e. convolutional filters) (IBM, 2023a). ReLU layers allows a faster and effective training and contribute to the activation of features that will be introduced in the next layer (Mathworks, 2023). Pooling layers (or subsampling layers), develop dimensionality reductions using filters without any weights, and can be from two types: max pooling (a filter selects the pixel with the maximum value and send it to the output array) and average

pooling (a filter calculates the average value and send it to the output array) (IBM, 2023a).

CNN are often used for classification and computer vision tasks and their approach from linear algebra (matrix multiplication) contributes to identify patterns within an image (IBM, 2023a). Forest applications as tree detection, forest classification, and forest damage detection are common tasks that have been developed through different investigations and the overall accuracies reached provide good perception of DL approaches for substituting manual feature extraction that is time consuming. Some of the most relevant processes involved in the model training are data augmentation and backpropagation:

- **Data Augmentation:** Process used for increasing the data size through transformations as flip, rotation, and resizing (Awan, 2023)
- **Backpropagation:** Algorithm based on chain rule to perform a backward pass while adjusting model parameters (Kostadinov, 2023).

The neural networks implement loss functions for comparing the target and the predicted output values, providing an idea of the model performance during the data training (Yathish, 2023). Using an optimizer (algorithms that updates weights and learning rate), the losses can be reduced, and the models can improve the performance for the specified task (Doshi, 2023).

CNN, as other NN and Machine Learning (ML) models, are initiated according to some Hyperparameters that set the performance frame for the training process (IBM, 2023b). Some of the most relevant Hyperparameters are the Learning Rate (LR), the batch size, and the number of epochs. Regarding the of CNN architectures, Residual Network (ResNet) has been frequently used for forest damage and tree mortality tasks (Ahlswede *et al.*, 2022; Beery *et al.*, 2022; Saleem *et al.*, 2020). ResNet was introduced by He *et al.* (2015) and consist of shortcut connections (residual connections) inserted in the structure of a plain network that work as bypass (Daly, 2023) (Figure 7). According to this research, it basically learns residual functions with reference to layer inputs and does not learn unreferenced functions, promoting an easier optimization. He *et al.* founded that deeper ResNet architectures performed better to achieve a lower training error,

generalizing in the validation dataset, gaining accuracy from the increased depth and providing a faster convergence. The Figure 7 shows the different layer blocks and the shortcuts inserted in a plain 34-layer depth model:

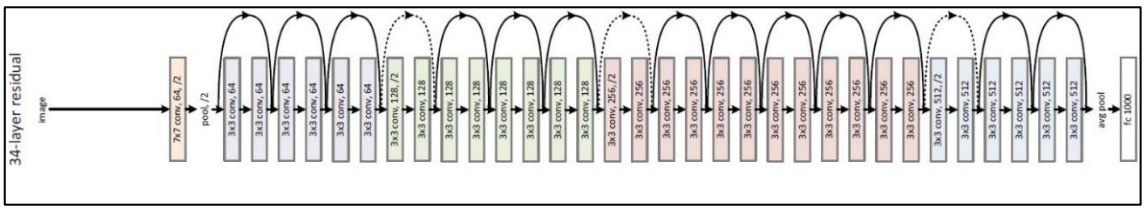

Figure 7. Network architecture of ResNet (with 34 parameter layers). Adapted from He *et al.* (2015).

Pytorch developments on ResNet architectures have provided a series of model builders under the torchvision package that could be used to instantiate the ResNet model using different options of model depth and with or without the inclusion of pre-trained weights (Pytorch, 2023).

❖ Vision Transformer

Also known as ViT, is a model part of Computer Vision branch developed by Dosovitskiy *et al.* (2020) that can be used in conjunction with CNN for classification tasks and requires fewer computational resources for pre-training compared to CNN (Boesch, 2023). The model uses split images (as fixed-size patches) that are embedded linearly with a specific position (sequence of vectors) together with an additional learnable embedding (classification token) to feed a standard Transformer encoder (Dosovitskiy *et al.*, 2020). The Figure 8 illustrates the model characteristics:

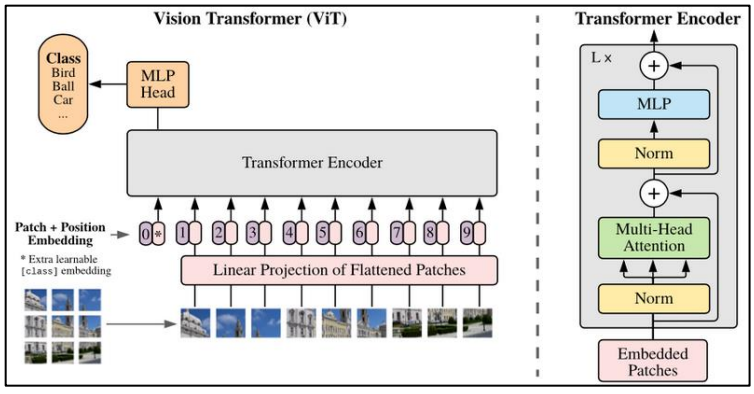

Figure 8. Architecture of ViT model. Taken from Dosovitskiy *et al.* (2020).

The models that implement late fusion techniques for fuse the final extracted features can be used to fuse Vision transformer and CNN models that are fed with images of different size (Ahlswede *et al.*, 2022).

❖ Benchmarking

Benchmarking is a term used in ML to define the practice of comparing tools to identify the best-performing technologies in an industry, and is used to measure the performance of Artificial Intelligence (AI) systems using an specific indicator through metrics that helps to compare the used tools (Dickson, 2023; Lopez, 2023). As Benchmarks examples can be named ImageNet, the NEON Tree Evaluation, and the TreeSatAI Archive. This last benchmark model will be discussed in the Methodology chapter.

3 Methodology

3.1 Study area

The capital city of Lower Saxony, Hannover, is located in the north-west of Germany at 52,3759°N – 9,3720°E (Figure 9). The city has a population of 534.094 inhabitants, surface area of 204.298 Km^2 with 11,36% of green spaces, and a total number of 91.047 trees, half of them (around 45.000 trees) make up the street lines of the city (Landeshauptstadt Hannover, 2022ab). The environmental office has developed studies about climate variability and mapped the air temperature and Heat Island Effect in the urban area. The city has implemented different strategies for mitigating the effects of Climate Change as city overheating, change of precipitation patterns and summer drought periods, which involve actions as roof greening, climate-adapted vegetation, and preventive soil and ground water protection (Hannover Division of Environmental Protection, 2017).

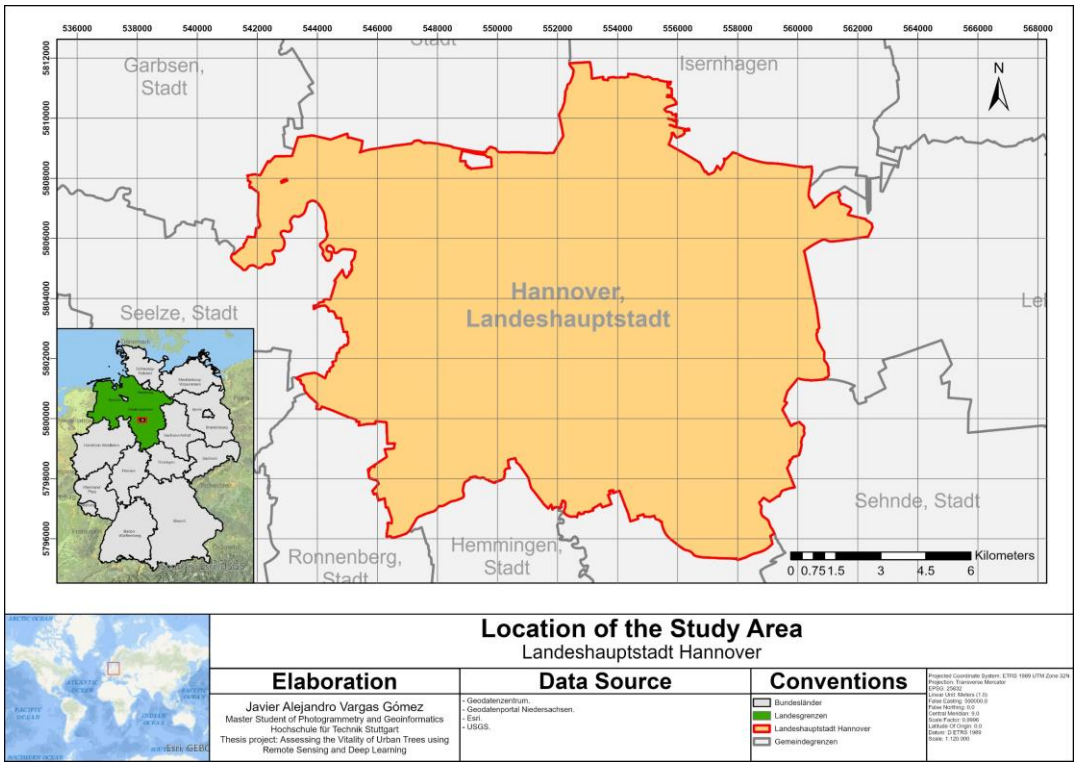

Figure 9. Location of the study area. Adapted from Geodatenzentrum (2023) & Geodatenportal Niedersachsen (2022).

3.2 Data sources

3.2.1 Tree inventory (Baumkataster)

The tree inventory of Hannover was provided by the Department of Environment and Urban Green Space (Fachbereich Umwelt und Stadtgrün, 2022a) and consisted of a database of 91.047 (382 species) with relevant information about different tree characteristics and measures as Baumhöhe (height), Gattung (genus), wissenschaftlicher Name (scientific name), gisid (unique GIS ID), Kronedurchmesser (crown diameter), Pflanzjahr (year of plantation), Baumumfang (tree girth in a 1-meter-height) and Vitalität (vitality or health status). This last field contained the most relevant information of the tree health status and therefore it was used to configurate the ground truth together with the aerial images as a starting point for the sampling step. This field is classified in six classes as shown in the Table 2:

Table 2. Tree vitality classes. Taken from Hannover Department of Environment and Urban Green Space (2022a).

Class	Description	Number of trees
0	Health level not yet defined	179
1	Healthy tree without damages or loss of vitality	49.078
2	Not totally healthy tree with a shortened but still satisfying vitality	29.561
3	Limited vitality, deadwood increasing, foliage dying up to 50%	8.748
4	Degenerating / dying tree	865
5	Dead tree	243
Not Classified	Unknown	2.373

The low availability of trees in the classes 3 and 4 made it necessary to consider additional multispectral sources as Sentinel-2 images, in order to provide more information to the model described in the *subsection 3.4.6*.

3.2.2 Multispectral images

❖ **Digital Orthophotos (DOP)**

A total number of 76 classic digital orthophoto tiles (2 Km x 2 Km) for the area of Hannover were downloaded from the LGLN Open Geo Data portal (see Figure 10).

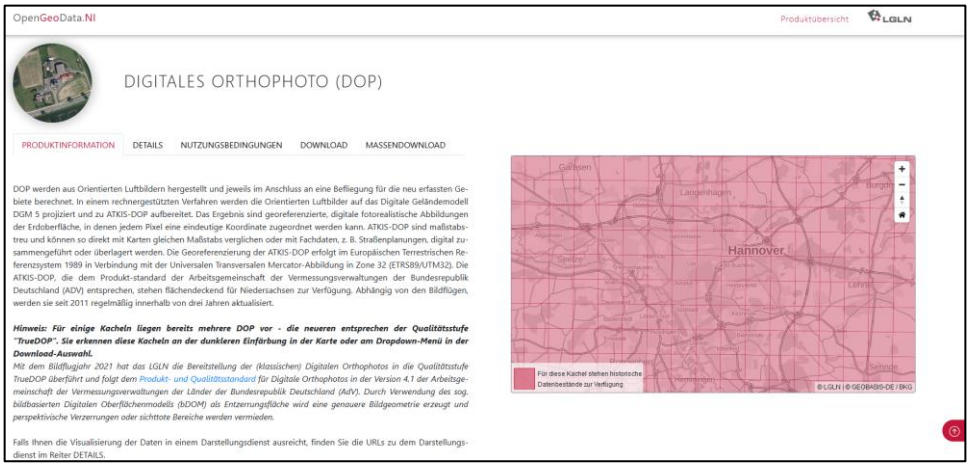

Figure 10. Open Geo Data portal for orthophotos acquisition. Adapted from LGLN (2022).

The images contained multispectral information in RGB+NIR channels in a spatial resolution of 20 cm x 20 cm for the years 2019 and 2022, and had a size of 539 MB each one. The DOP were produced from oriented aerial images which were projected into a DTM and set to the reference system ETRS89/UTM32 (LGLN, 2021).

❖ Sentinel-2 Images

Satellite imagery from the mission Sentinel-2 were downloaded covering the area of Hannover. L2A products were obtained from the Copernicus Sci Hub portal (ESA, 2022). Two (2) images tiles (100 Km x 100 Km) for the years 2019 and 2022 were obtained in closing dates of the aerial images to keep temporal consistency. The orthoimages were projected in UTM/WGS84 system and had a cloud cover percentage less than 4% and a cloud shadow percentage of 0,03% (ESA, 2022; ESA, 2023a). Each image had an approximated size of 1 GB.

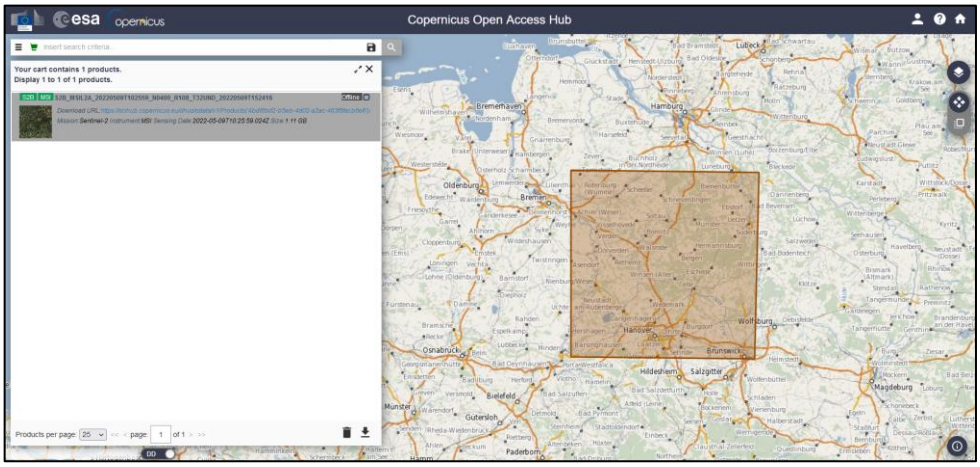

Figure 11. Copernicus Open Access Hub for satellite images acquisition. Adapted from ESA (2022).

3.2.3 Environmental variables

❖ **Air Temperature (Lufttemperatur)**

The information of Air temperature at 2 meters-height was additionally provided by the Hannover Department of Environment and Urban Green Space (2022b). It consisted of a gird with spatial distribution of average air temperatures at 4 am. Most of the warmest areas were located in the city center and in some industry areas of the outer zones of the city as shown in the Figure 12. This information was used as source for mapping of Heat Island Effect detailed in the next paragraph.

❖ **Heat Island Effect (Wärmeinseleffekt)**

The map of Urban Heat Island Effect was derived from the information of air temperature described in the last paragraph (Figure 13). The temperatures were grouped in four classes (Nicht vorhanden, Scwach, Mäßig and Stark) according to the Environmental Planning and Management Unit of Hannover (Hannover Sachgebiet Umweltplanung und -management) & GEONET (2017), and considers the air temperature in a group of built-up structures (block) within a grid cell.

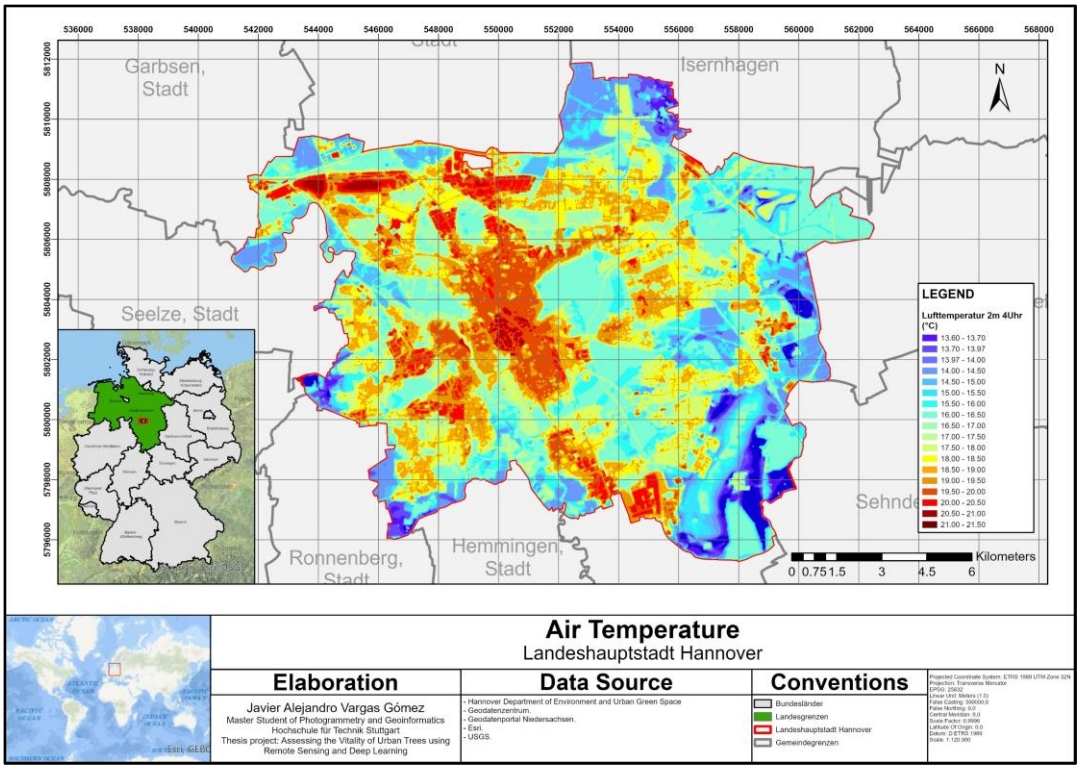

Figure 12. Air Temperature for the city of Hannover. Adapted from Hannover Department of Environment and Urban Green Space (2022b).

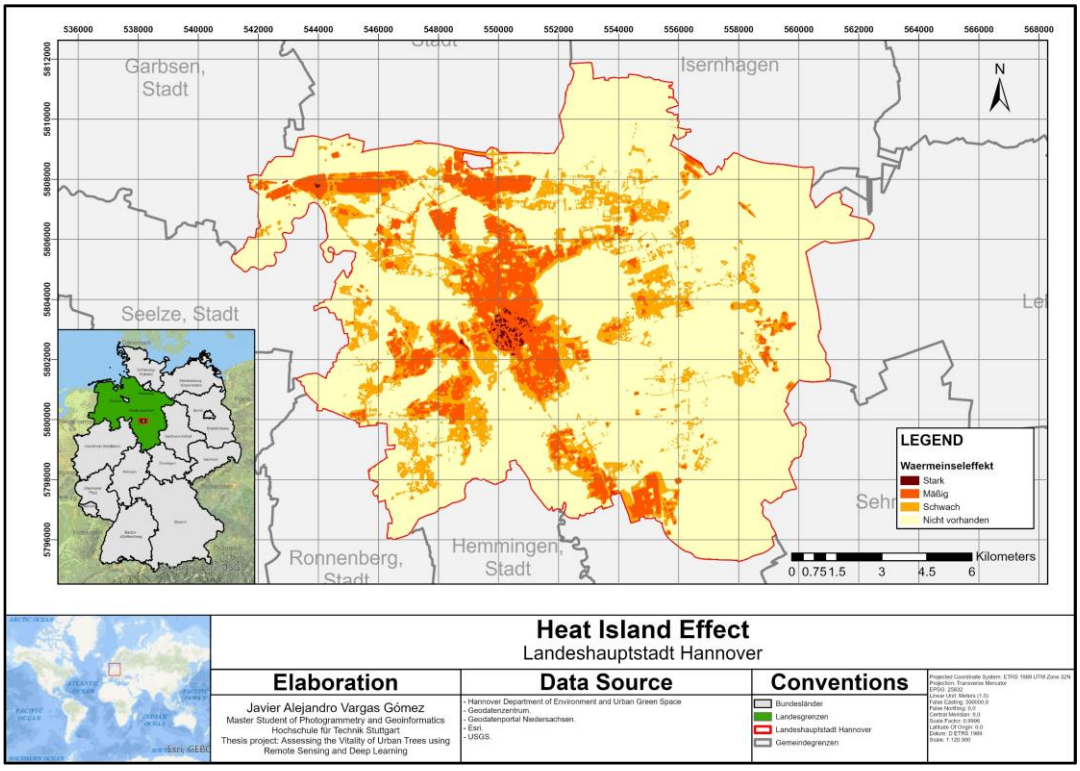

Figure 13. Heat Island Effect for the city of Hannover. Adapted from Environmental Planning and Management Unit of Hannover & GEONET (2017).

Most of the city surface corresponded to areas with no estimation of Heat Island Effect (73,92%). Among the estimated areas, Heat Island Effect with low impact (Scwach) presented the highest area (15,52%), followed by areas with moderated (Mäßig) and strong (Stark) impact that represented 10,34% and 0,22%, respectively. The area of the city center and some industries in the surroundings of the city were characterized for having moderated and strong Heat Island Effect. Those areas contained the lower density of trees (374 tree/km^2 and 52 tree/ km^2, respectively) while the highest tree density was located in areas with low Heat Island Effect (668 tree/km^2). From the 91.047 trees, 90.804 were located in the four areas defined by the Environmental Planning and Management Unit of Hannover & GEONET (2017). The Table 3 summarizes the number of trees per class of Heat Island Effect:

Table 3. Number of trees per vitality class and Heat Island Effect zones. Adapted from Environmental Planning and Management Unit of Hannover & GEONET (2017) and Department of Environment and Urban Green Space (2022b).

Tree Vitality Class	Heat Island Effect (Wärmeinseleffekt)			
	Strong (Stark)	Moderated (Mäßig)	Low (Schwach)	Not available (Nicht vorhanden)
0	0	11	29	139
1	14	3.793	9.149	35.966
2	6	3.019	8.651	17.807
3	2	904	2.642	5.192
4	0	63	320	482
5	0	8	41	194
Not Classified	2	108	334	1.928

❖ Soil Moisture (Bodenkundliche Feuchtestufe)

Information of soil moisture was searched on the Soil Information Systems Portal for the Geodata of Lower-Saxony (NIBIS) and provided by the Landesamt für Bergbau, Energie und Geologie (LBEG, 2022). Most of the trees recorded in the Baumkataster were located in areas with soil moisture less than 3%. Trees with Vitalität 3 and 4 (limited and deteriorated vitality) were located mainly in areas with soil humidity between 3% and 5% (Figure 14).

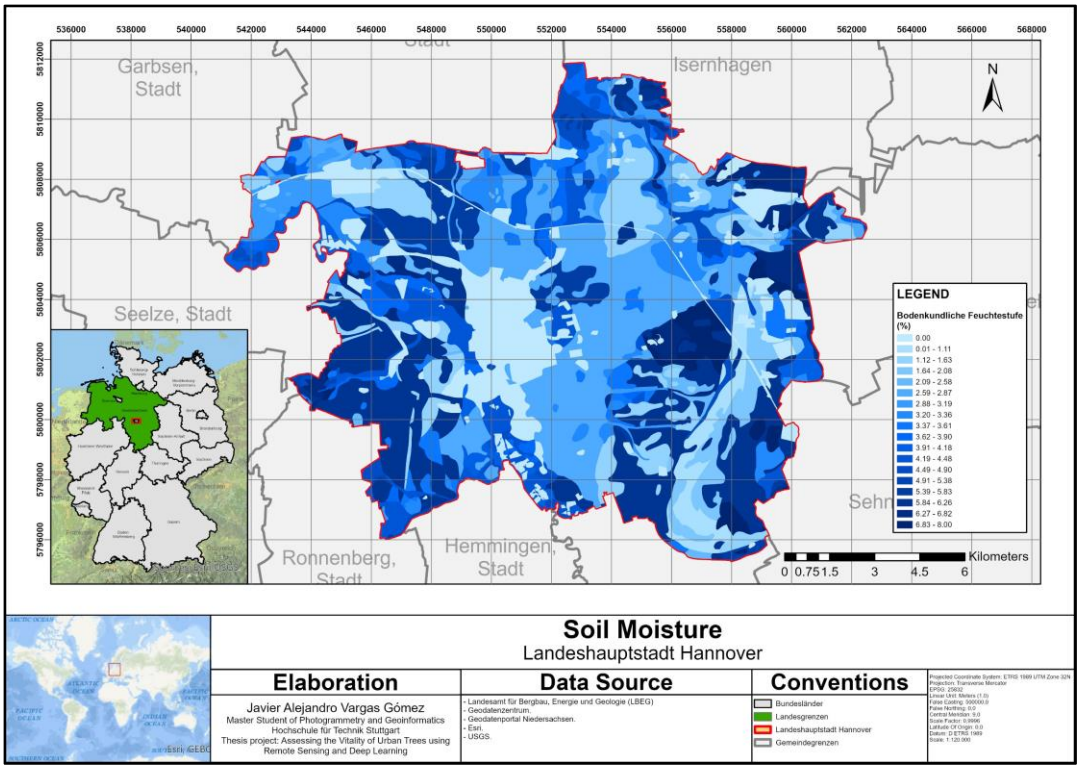

Figure 14. Soil moisture for the city of Hannover. Adapted from LBEG (2022).

3.2.4 TreeSatAI Benchmark Archive

The TreeSatAI Benchmark is a data archive developed for benchmarking related to forestry studies and consist of a dataset developed by Schulz *et al.* (2022) and python codes developed by Ahlswede *et al.* (2022). The dataset contains multi-spectral information from aerial sensors and Sentinel 1 and 2 missions for the area of Lower Saxony, multi labels for 15 tree genus and single labels for 20 species classes. The code involves pre-trained models (Resnet18, Multi-Layer perceptron-MLP, and Vision Transformer-ViT) with access to different configurations that can contain multiple or single datasets (e.g. Aerial and Sentinel-2) with options for training the model from scratch or using pre-trained weights. The research developed by Ahlswede *et al.* (2022) revealed important results regarding the tree classification performance of the models used as a Weighted Precision, F1 and mAP scores of 80,27% 71,54% and 79,50%, respectively.

As source for this research, the codes developed by Ahlswede *et al.* (2022) were adapted to the current study context, applying the sampling process to the area

of Hannover, preparing the labels of the vitality of urban trees selected from the Tree Inventory and the splitting specifications for training, validation, and testing.

3.3 Materials and Tools

3.3.1 Virtual Machine (VM)

A powerful virtual machine was used for running the model scripts. The VM consisted mainly of two processors Intel Xeon Gold 6226R, an installed RAM of 128 GB and a NVIDIA card GRID RTX8000-12Q with 12288 MB GDDR6 dedicated video memory and CUDA version 11.6.134.

3.3.2 Vector and Raster datasets

The processing of vector and raster data was done in both ArcGIS and QGIS. QGIS was mainly used for managing vector data, generating the mosaic of images downloaded from the LGLN Open Geo Data portal, and vectorizing raster inputs. ArcGIS was used for other geoprocesses as clipping/ extracting by mask, spatial joins and generating image files with the label name using the model builder. Additionally, ERDAS IMAGINE 2022 was used to generate the NDVI for the area of Hannover which guaranteed the original pixel size compared to the other software used. All (.tif) raster files were generated in order to keep the same spatial resolution of the original input images as the TreeSatAI dataset.

3.3.3 Codes and Scripts

Notepad++ and Visual Code were used for the writing of labels, calculating patch-image statistics, plotting loss curves, and adapting the different models developed by Ahlswede *et al.* (2022). Citrix Workspace was used as interface for connecting to the Virtual Machine that would run the configurations set in the models through the Prompt. Anaconda was used to managing different sources and libraries in a specific virtual environment, as Torch and Numpy.

3.4 Processings and Analyses

The Figure A 1 (Appendices) summarizes the workflow for the methodology implementation, consisting of data processing, filtering of tree inventory, data labeling, image sampling, input data preparation, training, validation, and testing.

3.4.1 Data processing

The 76 aerial orthophotos were merged in a single file (orthomosaic) which was extracted for the extent of the city, obtaining a 160 GB orthomosaic file. For visualization and mapping purposes, the city orthomosaic was also generated in a lighter ECW format (Figure 15). Satellite imagery bands (all S2 bands except B10) were stacked and clipped also for the study area keeping a spatial resolution of 10 m. The band order was sorted as NIR-R-G-B for aerial and B02-B03-B04-B08-B05-B06-B07-B08A-B11-B12-B01-B09 for Sentinel-2 (Ahlswede *et al.*, 2022). Vector data was organized and clipped to the extent of the city boundary.

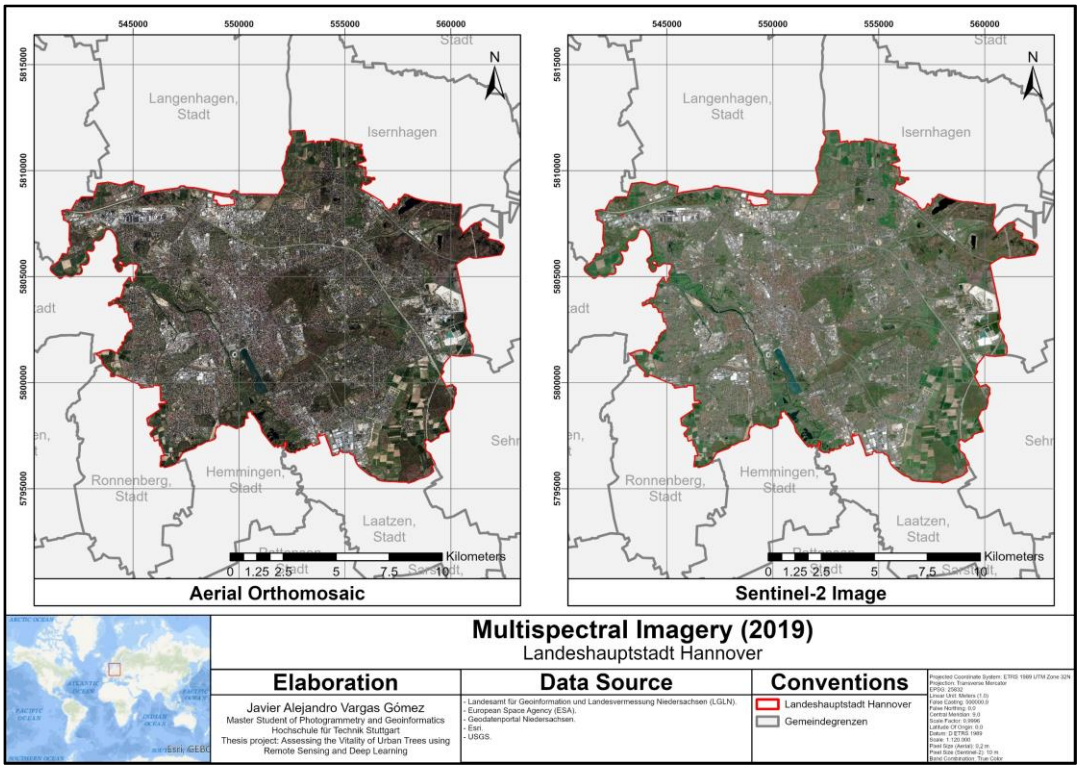

Figure 15. Aerial orthomosaics (left) and S2 images (right) of the city of Hannover for the year 2019 in true color band combination. Source: LGLN (2022); ESA (2022).

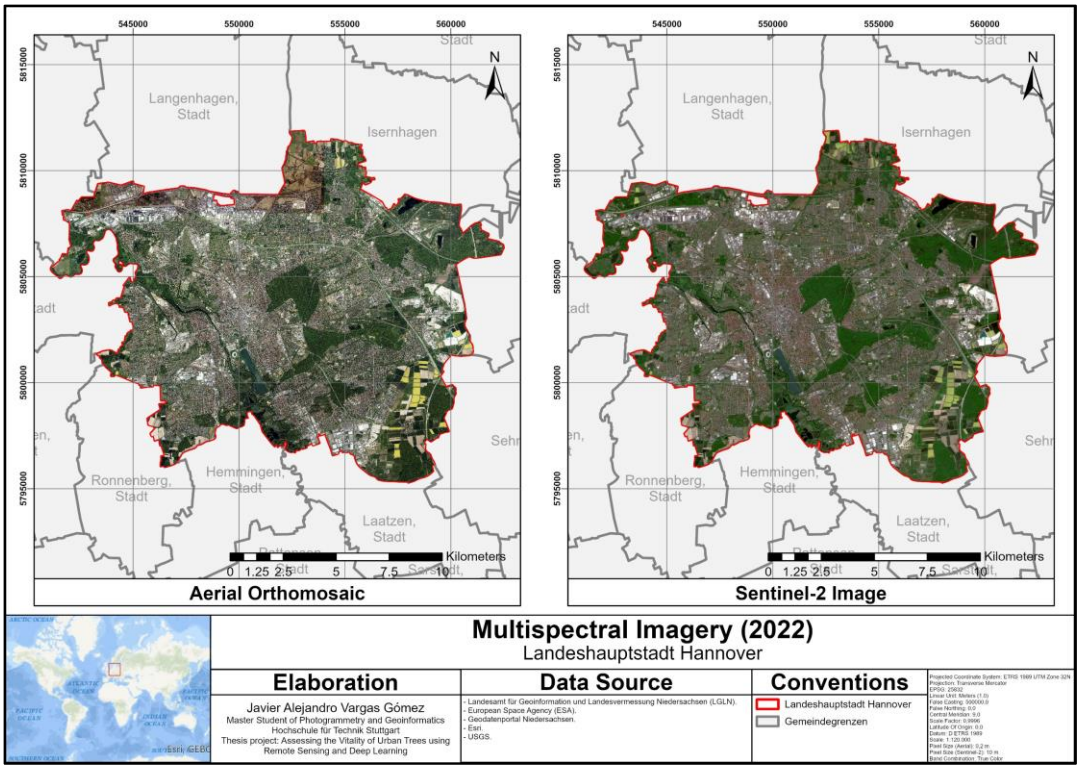

Figure 16. Aerial orthomosaics (left) and S2 images (right) of the city of Hannover for the year 2022 in true color band combination. Source: LGLN (2022); ESA (2022).

❖ Normalized Difference Vegetation Index (NDVI)

The aerial orthomosaics created in the previous step was processed to obtain a spatial distribution of the NDVI in order to distinguish healthy from unhealthy vegetation, and also for filtering processes due to the fact that not all the healthy trees presented foliage due to the beginning of the spring (this could confuse the network making it interpret a healthy and naturally defoliated tree with a non-healthy tree). The Figure 17 shows the NDVI for each year.

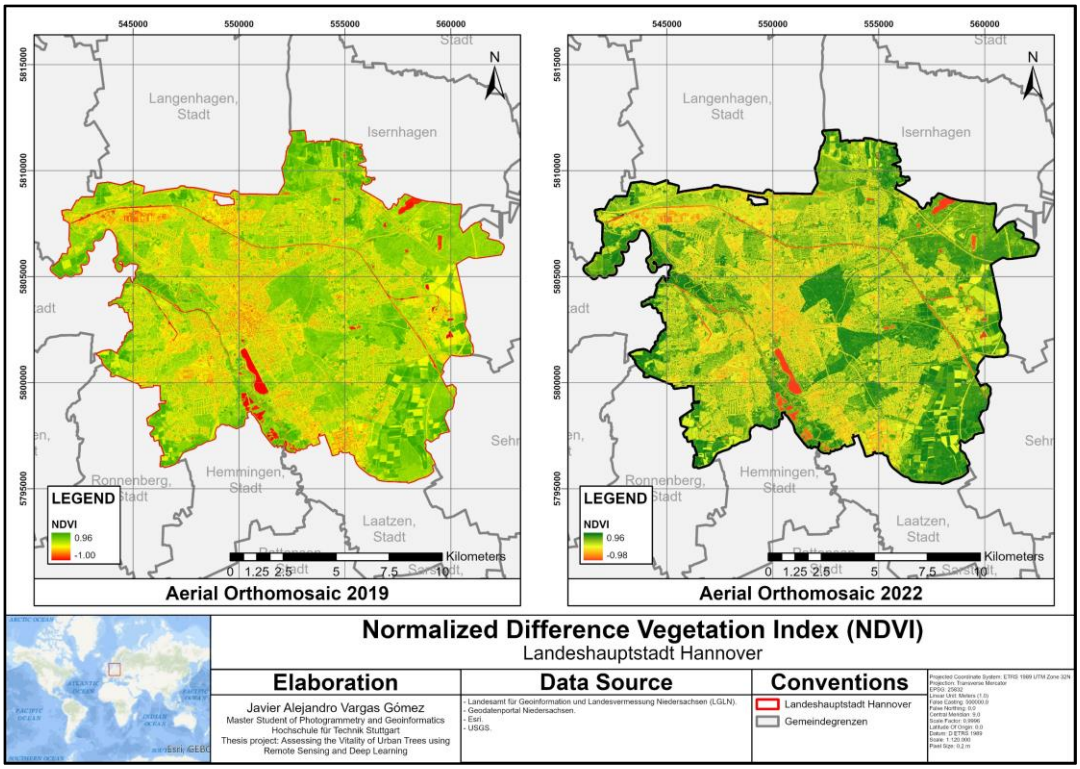

Figure 17. NDVI of 2019 (left) and 2022 (right) for the city of Hannover. Adapted from LGLN (2022).

3.4.2 Filtering of tree inventory

Initially, the trees located in areas with Heat Island Effect (low, moderate and strong) were extracted. The trees with Vitalität values 1, 2, 3 and 4 were sub-selected because they represented the healthy and non-healthy species that are part of the scope of the study (Figure 18). Trees with Vitalität 0 (health status not defined) and 5 (dead) were not considered in the process. Therefore, the field “Health Status” was created in the attribute table for classifying the trees in three categories: Healthy (Vitalität class 1 and 2), Limited Vitality (Vitalität class 3) and Deteriorated Vitality (Vitalität class 4).

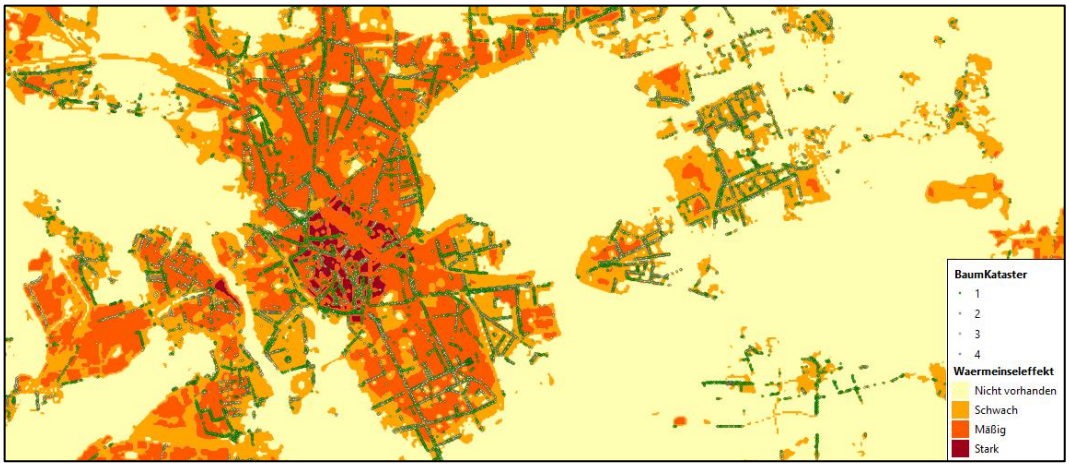

Figure 18. Selection of trees of classes 1,2 3 and 4 in areas of Heat Island Effect. Source: Department of Environment and Urban Green Space (2022ab).

Despite the very extended information of the Tree Inventory, there were still some gaps regarding the completeness of the database, therefore the tree inventory was filtered to keep even more solid and consistent information about the health status of the tree and the variables that could be related to its development (e.g. year of plantation). The crown diameter usually can help to relate the area occupied of the tree in the image patch. However, the information was only available for a reduced number of species and then it was not completely considered in the study.

For avoiding oversampling of the Vitalität classes 1 and 2 (Healthy) that represented the highest number of trees in the Baumkataster, the image patches were filtered primarily for keeping most of the trees of classes 3 and 4 (Limited and Deteriorated Vitality) without overlapping from other image patches of healthy class. Additionally, the overlap between image patches of all classes was suppressed based on the selection of trees that best represented the classes (e.g. Deteriorated Vitality trees had usually no leaves, Limited Vitality trees had just a few leave coverage and Healthy trees mainly had a canopy full of leaves and a high NDVI). For this reason, visual checkups of the tree canopy were developed using the CIR band combination of the aerial orthomosaics and maximum values of NDVI of the Healthy trees. Moreover, for including more trees of the less represented class (Deteriorated Vitality), trees with no specification of plantation year were taken in count.

Since the aerial images corresponded to different stages of spring (April 2019 and May 2022), there were more trees selected from 2019 were DV and LV trees still did not grow leaves and then they could best represent the defoliated trees. This contributed to include more partially or fully defoliated trees in the mentioned classes. Finally, 708 trees were selected from the original database: 264 Healthy trees (37,29%), 232 trees with Limited Vitality (32,77%), and 212 trees with Deteriorated Vitality (29,94%). There were not trees selected from the Stronger (Stark) Heat Island Effect zone, either for being located on shaded areas or because of the sampling process prioritized the less represented class (DV). The Table 4 summarizes the number of trees selected for the study and their location in the heat island zones:

Table 4. Selected trees and corresponding location in the Urban Heat Islands.

Tree Vitality (Vitalität)	Label Class	Heat Island Effect (Wärmeinseleffekt)		
		Strong (Stark)	Moderated (Mäßig)	Low (Schwach)
1	H	0	81 (44%)	104 (56%)
2				
3	LV	0	86 (37%)	146 (63%)
4	DV	0	42 (20%)	170 (80%)

The Figure 19 shows some examples of representative trees of each class. It is possible to see the difference in the defoliation state of the LV and DV, and the canopy vigorousness of class H.

Figure 19. Examples of selected trees for each class: H (left), LV (middle) and DV (right).

3.4.3 Data labeling

A new field was created in the attribute table of the tree inventory to label the features following the structure health status acronym__Vitalität class__gisid__Y(Flight Mission Year) addressing a file name as *DV__1__567__Y2019.tif*. A square buffer was created for each of the selected trees (point features) as shown in the image below. These polygons were used to extract the image patches required to feed the CNN and ViT models, using a model builder in ArcGIS Pro (*subsection 3.4.4*). The inclusion of the year in the filename was done according to the date of the aerial flight (2019 and 2022). A similar process was developed for the satellite images, providing the same file name in a different directory as suggested by Schulz *et al.* (2022).

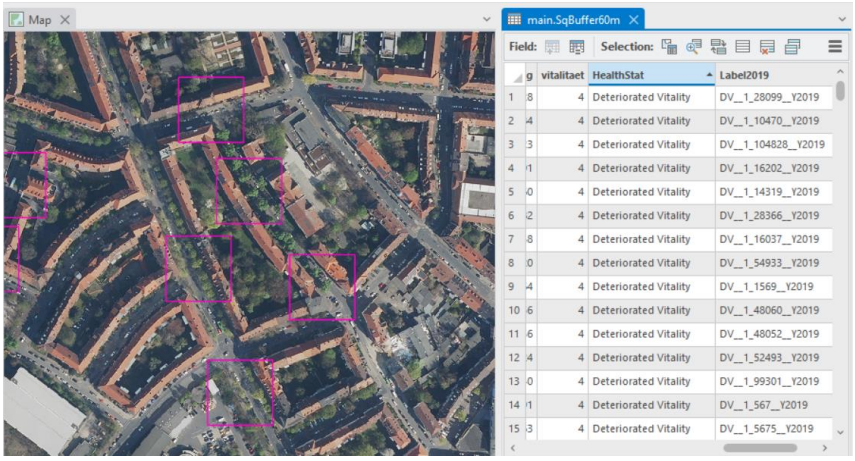

Figure 20. Square buffers generated for each selected tree. Source: LGLN (2022); Hannover Department of Environment and Urban Green Space (2022a).

The area of each tree crown was digitized in ArcGIS Pro using both natural color and CIR combinations (Figure 21).

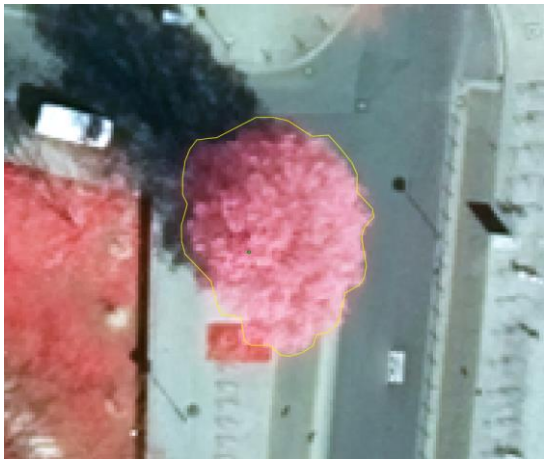

Figure 21. Crown polygon digitized for each selected tree. Source: LGLN (2022).

For creating the source file that the model would use in the training process, a label file (.json) was created through a python script taking in count the label base name, the health status class, and the area percentage occupied by the tree in the image patch, as shown in the Figure 22.

Figure 22. Label file (.json) generation.

3.4.4 Image patches sampling

A total number of 708 image patches (60 m x 60 m) were generated from the aerial orthomosaics using an ArcGIS model builder as shown in the Figure 23. Sentinel-2 images from the same temporality were also sampled to increase the size of the dataset due to the low number of trees selected among both years.

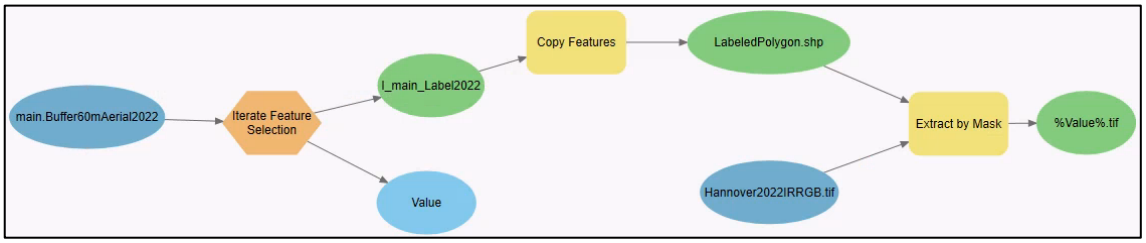

Figure 23. Patch generation process modeled in ArcGIS Pro (Model Builder).

The resulting patches of each source were like the ones showed in the Figure 24. Aerial and Sentinel-2 image patches consisted of 300x300 pixels and 6x6 pixels, respectively:

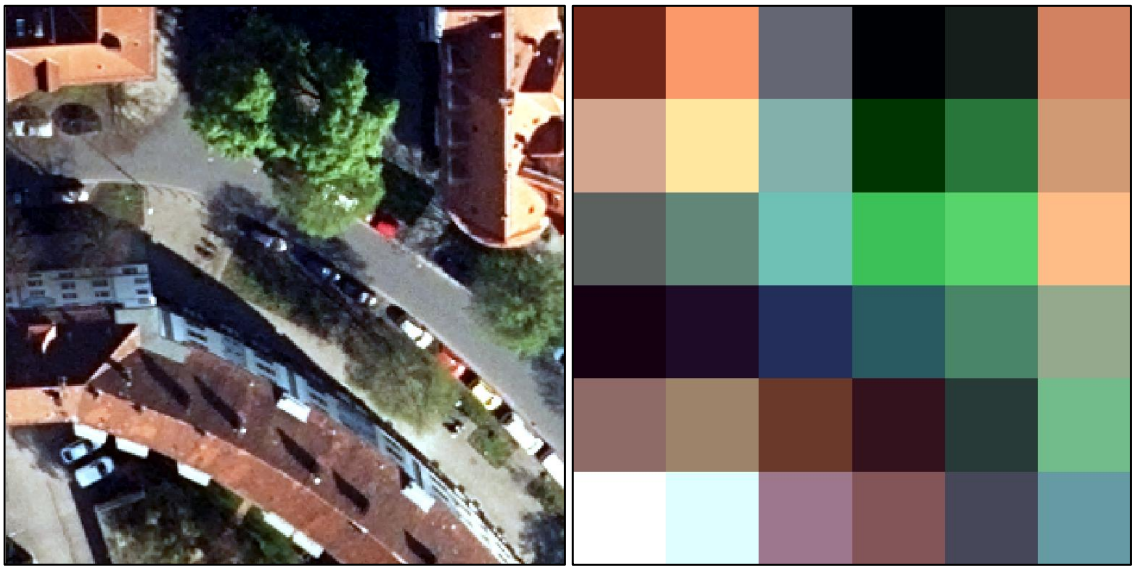

Figure 24. Image patches generated for aerial (left) and Sentinel 2 (right) sources (File name: H__1_98616__Y2019.tif).

3.4.5 Input data preparation

The preparation of the input dataset and scripts was done following the dataset structure suggested by Schulz *et al.* (2022). It consisted of organizing the folder for each image source, the folder that included the JSON file (labels), and the dataset split (training, validation, and testing) as .lst files. The data split was defined in a proportion of 90% for training, 5% for validation and 5% for testing (Table 5).

Table 5. Dataset split.

Label Class	Training split	Validation split	Testing split
DV	191	11	10
LV	209	12	11
H	238	12	14

Additionally, band means and standard deviations were calculated for each patch from each source and were used for the normalizations in the dataloaders. As the classes were imbalanced due to the characteristics of the tree inventory (just 0,95% of the inventory were trees with vitality class 4), a calculation of the Inverse Number of Samples (INS) for weights calculation suggested by Shrivastava (2022) was done in order to introduce some contributions to the class with the lowest number of samples in the training dataset (DV):

```
## Determine the weights using the Inverse of Number of Samples (INS)
no_of_classes=3
samples_per_cls=(191, 209, 238)##DV-LV-H

def get_weights_inverse_num_ofsamples(no_of_classes, samples_per_cls, power = 1):
    weights_for_samples = 1.0/np.array(np.power(samples_per_cls, power))
    weights_for_samples = weights_for_samples/np.sum(weights_for_samples)*no_of_classes
    return weights_for_samples

x = get_weights_inverse_num_ofsamples(no_of_classes, samples_per_cls)
```

3.4.6 Training, internal validation and testing

❖ Model training and validation

The training step aimed to compare two configurations based on the input data, as made by Ahlswede *et al.* (2022). Firstly, ResNet-18 model was trained using only a dataset of aerial images. Secondly, for increasing the information per tree in the model and overcome the issue of the low number of samples, the ViT model was fed with Sentinel-2 image patches, using late fusion with the final features from the ResNet-18 model fed with aerial images.

The training and validation process was started using the command prompt, specifying the number of CUDA used by the Virtual Machine, the model that will be used from the dataloders, the name of the weights file that will be stored for the testing step, the log file for storing the metric results of training and validation,

and the label file path that will be used across the different stages of the model (Ahlswede *et al.*, 2022). The training script would call different sub-scripts as follows:

- **Augmentations**: It defined the flippings and rotations of input images.
- **Dataloaders:** It defined the builder functions for loading the datasets of both sources and specifies the transformations in data (i.e. normalization of aerial data for introducing it as input of the ResNet model). Each one of the S2 image bands were organized as a grid for being used as input in the ViT model and then each band would be a token for the model (please see Figure 8).
- **Loss Function:** It set the *BCE With Logits Loss* function that would use the weights calculated for reducing the influence of class imbalance.
- **Optimizer:** It set the Adam optimizer that would update the weights and LR, and reduce the loss.
- **Hyperparameters:** It specified or called the hyperparameters of the network as base and maximum LR, scheduler, batch size and number of epochs.
- **Model Trainer:** Class called in the training script for using the functions that would retrieve the input data from dataloaders and use it to calculate the predictions, losses, and model metrics. It also set the backpropagation line for update the errors with the optimizer.

❖ Model testing

Analogically as the training step, the evaluation process was started using the command prompt, specifying the number of CUDA used by the Virtual Machine, the model that will be used from the dataloders, the file path of the weights file saved during the training, and the label file path. The testing step consisted of processing a batch of 35 images to obtain the predictions which would be introduced in a sigmoid function in order to set the predictions in an interval from 0 to 1. The script developed by Ahlswede *et al.* (2022) was adapted to print the predicted labels and analyse the performance of the model.

4 Results and Discussion

4.1 Influence of Input Data

4.1.1 Urban trees and environmental conditions

When analysing the trees selected in the study (Table A 1 in Appendices), the species *A. platanoides*, *Q. robur*, and *T. intermedia* were the most frequent with 150, 113, and 48 individuals, respectively. The class H was was mainly represented by the species *A. platanoides* and *T. intermedia*, while classes LV and DV were represented mostly by *Q. robur* and *A. platanoides* (Figure 25).

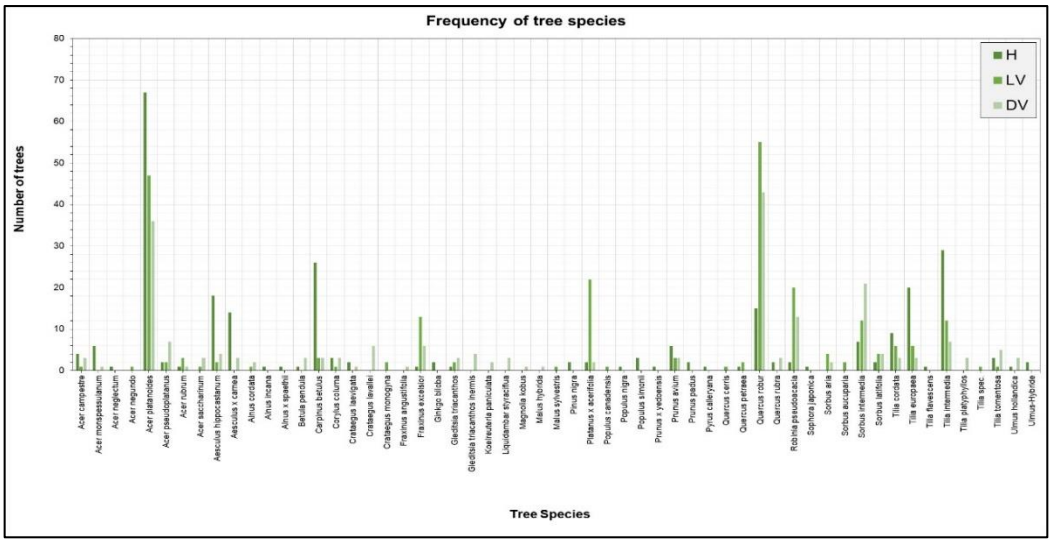

Figure 25. Number of trees of classes H (Healthy), LV (Limited Vitality), and DV (Deteriorated Vitality).

Regarding the location of the selected trees in the urban heat islands, it is possible to see that most of the species were located in the Low zone (464). The species *A. platanoides* was the most frequent in both Moderated (53) and Low zones (97), followed by *P. x acerifolia* (21) and *Q. robur* (93), respectively, as shown in the selection of the most frequent species in the Figure 26. The presence of *A. platanoides* can be due to its characteristics as tolerance to high evapotranspiration or prolonged drought phenomena as well as for different services provided as shade and biomass (Caudullo & de Rigo, 2016). According to Pretzsch *et al.*

(2018), the species *P. x acerifolia* is characterized for increasing their crown projection area and increasing the cooling capacity proportionally to the age compared to other species as *T. cordata*, *R. pseudoacacia*, and *A. hippocastanum*, which is important for urban planning and then can be relevant for providing shade and cooling the urban surface of cities. Regarding *Q. robur*, it is also important and valuable as park and roadside trees due to its provision of shade (Eaton *et al.*, 2016).

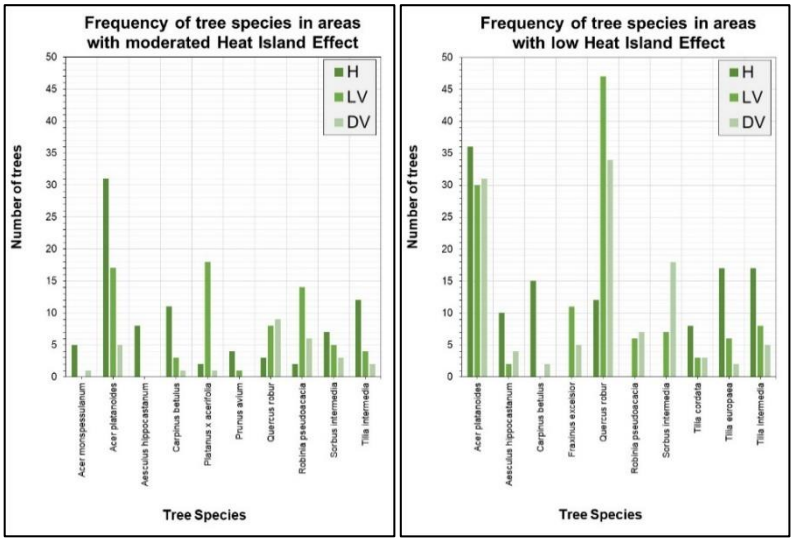

Figure 26. Number of trees of classes H, LV and DV per Urban Heat Island zone.

The Table A 2 (Appendices) present the species, vitality class, Heat Island Effect zone and soil moisture level during spring time. It shows that, in general, most of the trees with the highest soil moisture were located in areas with Low Heat Island Effect, with species as *A. platanoides* (6,64%) in classes LV and H, and *S. intermedia* (6,33%) in class DV (Figure 27). On the other hand, there were more trees located in areas of moderated Heat Island Effect over dried soils or with very low moisture content, that is the case of *S. latifolia* in classes LV and H, and *G. triacanthos inermis* in class DV (Figure 28).

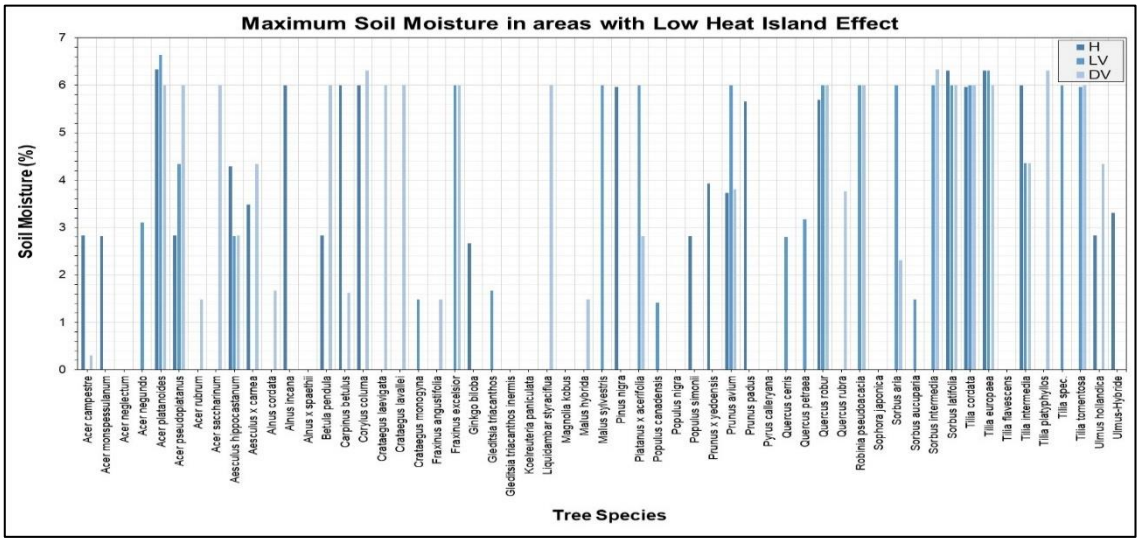

Figure 27. Maximum Soil Moisture per tree in areas with Low Heat Island Effect.

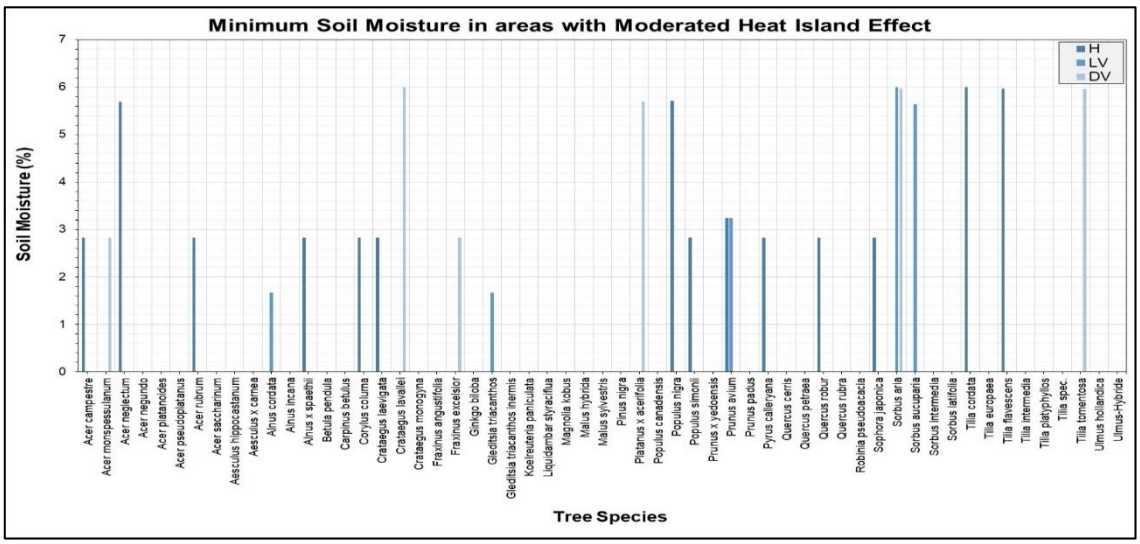

Figure 28. Minimum Soil Moisture per tree in areas with Moderated Heat Island Effect.

According to Brune (2016), the species *A. platanoides* is moderately tolerant to drought stress in the area of Lower Saxony, as well as other species of the selected dataset of trees as *F. excelsior* and *Q. robur*. In a similar way, Selig (2020) classified *Q. robur* and *T. cordata* as species with high tolerance to drought. This could mean that some of the most common species in the study present tolerance to low soil moisture levels and canopies with provision of shade while they were located in areas with moderated Heat Island Effect. On the other hand, Brune (2016) also found that other species that were taken in count in the current study can be very sensitive (*P. nigra*), moderately sensitive (*A. pseudoplatanus* and *A.*

hippocastanum), and very tolerant (*B. pendula*). The Table 6 illustrates the overall assessment of drought tolerance of some of the tree species of the current study according to Brune (2016):

Table 6. Drought tolerance of some of the tree species from Lower Saxony taken in count in the current study. Adapted from Brune (2016).

Tree Species	Overall Assessment	Number of Species
A. platanoides	Moderately tolerant	150
A. pseudoplatanus	Moderately sensitive	11
A. hippocastanum	Moderately sensitive	24
B. pendula	Very tolerant	4
F. excelsior	Moderately tolerant	20
P. nigra	Very sensitive	1
Q. robur	Moderately tolerant	113

The tolerance to drought phenomena is strongly related to the heat stress that trees can show because the soil plays an important role for water and nutrient storage, and this availability of water and nutrients during dry periods can represent a difference between life and death. For this reason, it is recommended to introduce very and moderately tolerant tree species in areas where Heat Island Effect is strong or moderate, mainly in the surroundings of the Hauptbahnhof. These actions should be developed together with urban landscape approaches as façade greening and improvement of the soil volume for roots as suggested in the Action Fields 3, 4 and 5 of the “Adaptation Strategy to Climate Change for the Capital City Hannover” (Hannover Division of Environmental Protection, 2017).

4.1.2 Aerial and S2 imagery

The aerial sources clipped for the area of Hannover were visually analysed for understanding the dynamics of vegetation between epochs. The beginning of the spring showed in the Figure 15 indicated that most of the vegetation probably had not leaves and flowers, or they were starting to grow as part of their biological cycle. In contrast, the Figure 16 (right) shows more vegetation and a different growing state that can be easily noticed in the municipal forest *Eilenriede* located in the middle of the city and known for being the largest urban forest in Europe (Tourismus Marketing Niedersachsen, 2023). It is possible to see the land cover dynamics in the surrounding area of both images where industries (white/gray

roofs) and agricultural fields exist. The CIR combination for infrared response of vegetation let to understand much more about the distribution of this cover as well as water bodies presence. The Figure 29 (left) and Figure 30 (left) shows the vegetation cover that absorbed the light in a less rate than the vegetation showed in the images of the right, probably due to the physiological state of vegetation that is more vigorous in 2022 (end of spring). It is also possible to see the water bodies as the Maschsee, the Leine river, and the Stichkanal Hannover-Linden, as well as a concentration of artificial lakes in the south of the city. The natural water bodies could be providing moisture to the air and strengthening the fresh air flux to the city center.

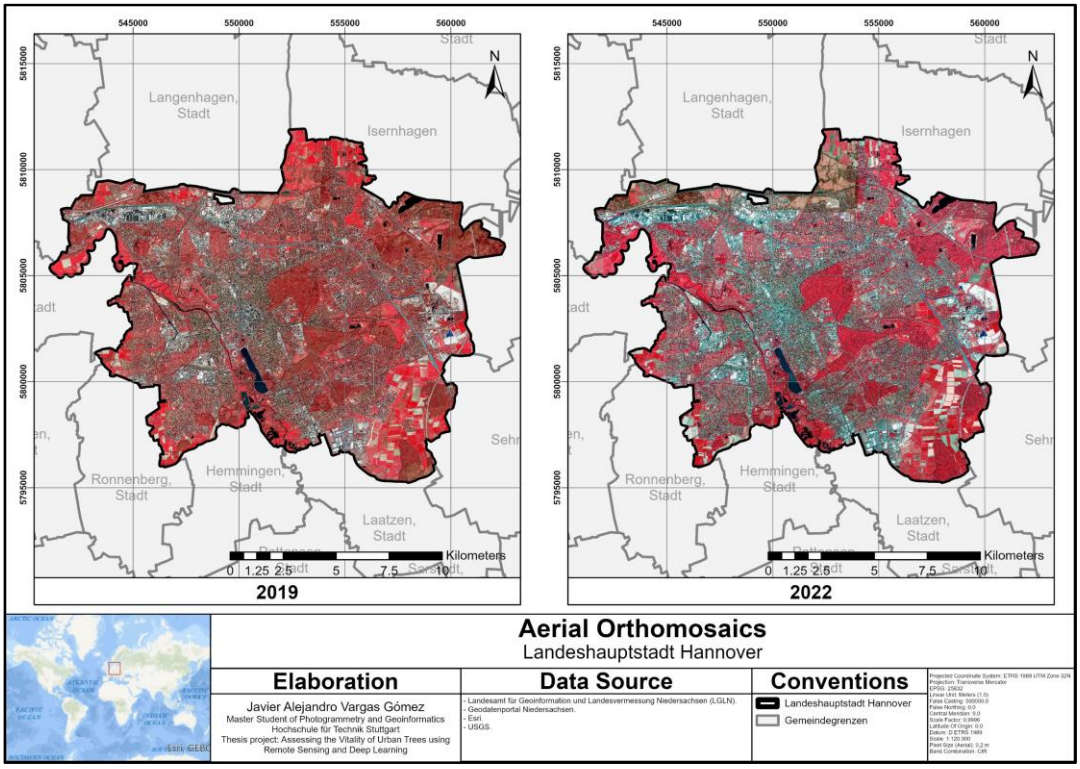

Figure 29. Aerial orthomosaics of the years 2019 (left) and 2022 (right) in CIR band combination.

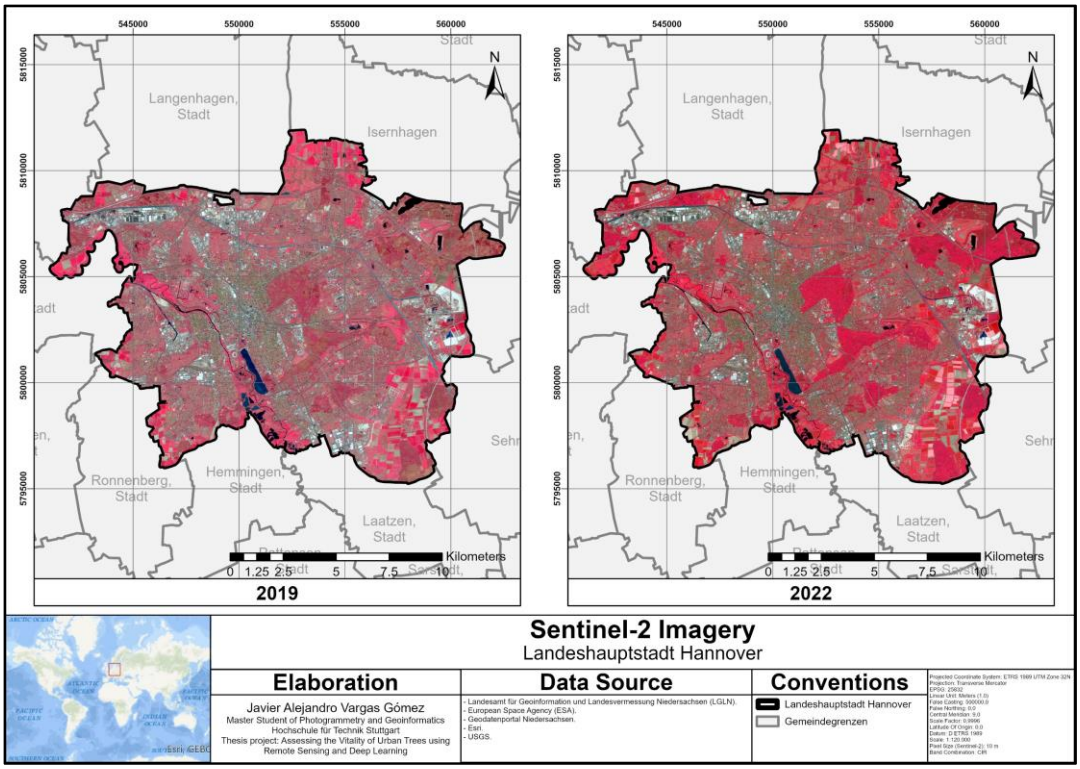

Figure 30. S2 images of the years 2019 (left) and 2022 (right) in CIR band combination.

The histograms of each aerial source for each year (Figure 31 and Figure 32) were analysed for identifying the distribution and behavior of the RGB+NIR band values for 8-Bit Images, mainly about the changes of vegetation response across the two epochs.

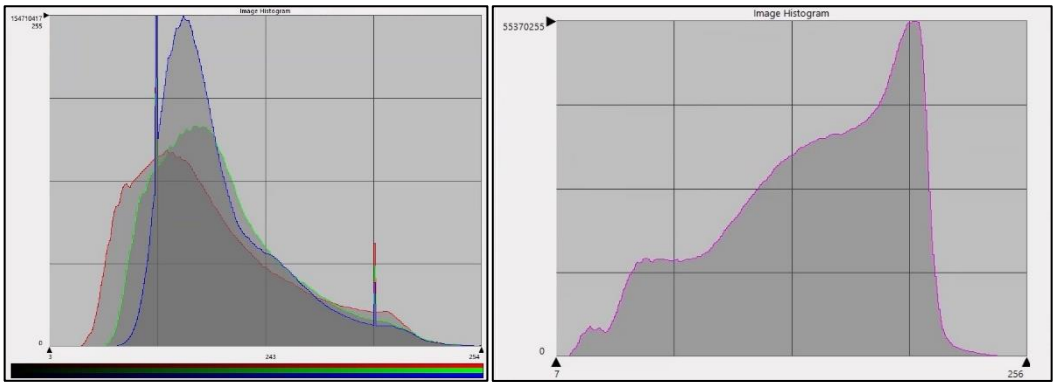

Figure 31. Histogram of RGB (left) and NIR (right) bands for the 2019 aerial orthomosaic.

For the year 2019 (Figure 31) existed a higher response of blue band compared to red and green ones. Blue light can be reflected by the water bodies but also by human made covers as roads. It can be also related to the lower amount of green pixels that correspond mainly to both conifer species or from deciduous

species that start to grow leaves in an early stage of the spring, and then the pavement light is partially reflected. The red band behavior represents the amount of light that was partially absorbed by defoliated vegetation (plant structures as trunk and branches) that reflects approximately an 8% of healthy vegetation and the 30% on non-heathy vegetation (please see Figure 3), as well as reflected by bare soil fields, house roofs, vehicles, and some mineral and synthetic materials like brick dust in tennis fields and rubber of athletics running tracks, respectively. Regarding the NIR band (Figure 31, right), it is possible to see a higher amount of pixels recorded from healthy or vigorous vegetation (which reflects 50% of light) and non-healthy vegetation (which reflects 40% of light) (EOS, 2023). Additionally, tree structures as trunk and branches could show higher pixel values (please see Figure 33 in *subsection 4.1.3*) which allowed to differentiate much better the tree crown for manual digitizing even in the beginning of the spring and on defoliated trees. The presence of darker tones in the RGB and NIR histograms can be explained either by the shadows reflected by buildings and houses or by the presence of water bodies.

The histograms of the 2022 orthomosaic (Figure 32) shows a different behavior, specially set by the increase of vegetation biomass (either of urban vegetation or crops with increased leaf area and new plant structures) as well as the energy reflected during a season of higher sun exposure (end of spring). The green response shown in the histogram of the Figure 32 (left) is higher compared to 2019 histogram and is mainly due to leaves and flowers growing, thus there were more green pixels. This behavior is also shown in the histogram of the infrared band (Figure 32, right) where the total number of pixels is higher and considering more darker tones. The different pixel values are also due to the different flight mission days of the images in the north of the city.

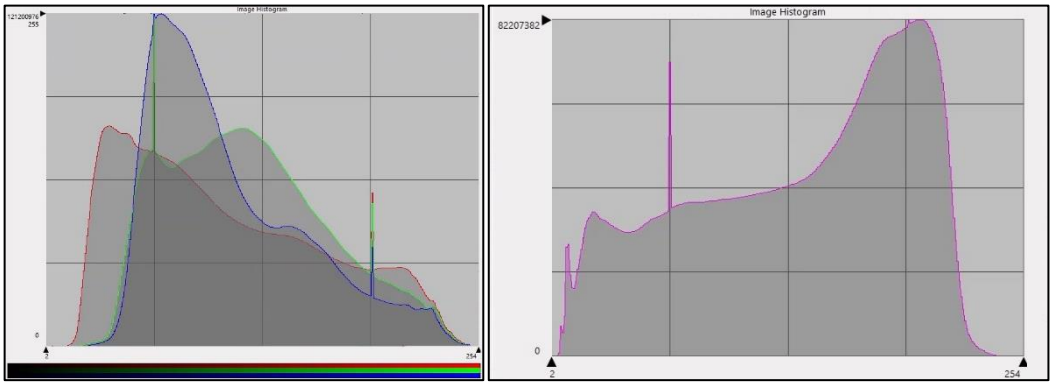

Figure 32. Histogram of RGB (left) and NIR (right) bands for the 2022 aerial orthomosaic.

4.1.3 Aerial image patches

The aerial patches that contained information about the trees of different health status presented a particular behavior due to their defoliation state or because of the physiological condition of their leaves. Therefore, in order to understand the spectral behavior of these trees, a plot of the spectral profile from the patches of each class were analysed using the Multispectral tools from ERDAS IMAGINE. The Figure 33, show the image patches (as CIR band combination) and the spectral profile of selected trees from the classes DV, LV, and H.

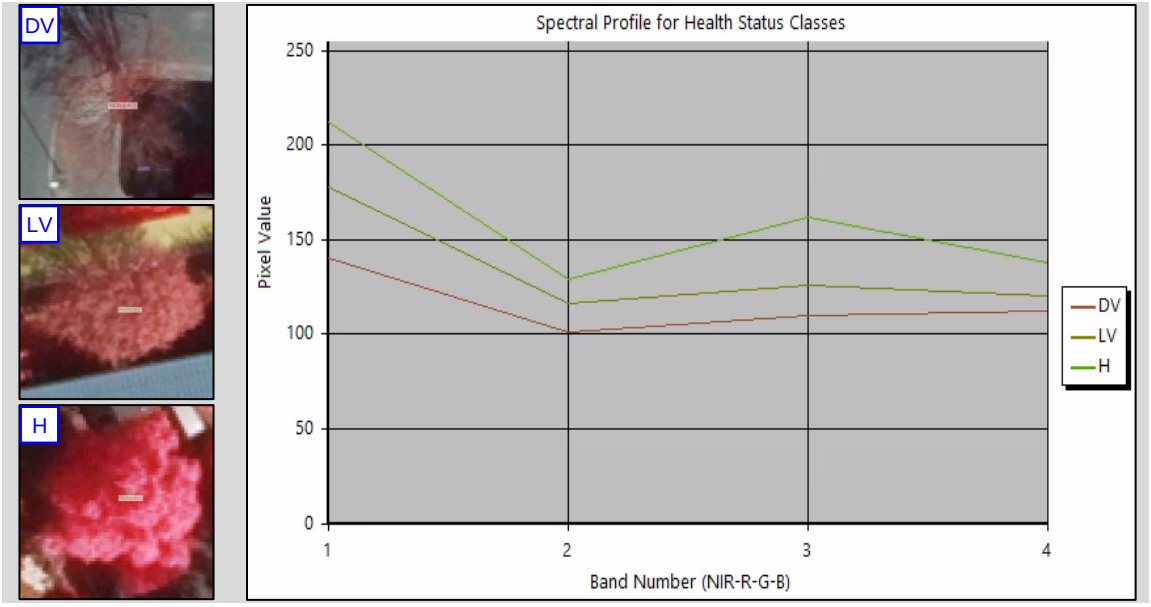

Figure 33. Image patch of DV, LV and H trees (left) and spectral profile (right).

The Figure 33 indicates that each band behaved in a similar way than the others, with the infrared spectrum showing the highest response from leaves and tree structures (trunk and branches) even from class DV. The more the tree has a

vigorous state, the higher is the value of the pixel in the NIR channel (EOS, 2023). Due to the absorption of red spectrum by the vegetation, the three classes showed the lowest values in the red band. Additionally, there was a greater difference between NIR values and red values of the class H compared to the unhealthy classes. The green band showed the highest difference of values among healthy and unhealthy trees because the green is highly reflected from leaves (NASA, 2023). Regarding the blue band, it is probable that the visibility of pavement or urban structures trough the canopy of defoliated trees generated such response.

4.2 Performance of Classification Model

4.2.1 Training and Validation

The model for each experiment (aerial and aerial/S2) was trained using different combinations of hyperparameter values and network parameters that allowed to initiate the classification. Invariable parameters as LR Scheduler (Cyclic), Optimizer (Adam) and Loss Function (BCE With Logits Loss) were kept as set by Ahlswede *et al.* (2022). The Table 7 shows a summary of the hyperparameters selected.

Table 7. Hyperparameters selected for the training and validation steps.

LR	Batch Size	Epochs	Base LR
0,005	32	200	0,001

The selection of the learning rate and epochs was done due to the low number of samples and the characteristics of the image patches (urban context with at least two covers as urban built areas and vegetation), and also after several tries it was decided to reduce the cycle size of the learning (Base LR from 0,0005 to 0,001) due to the behavior of the validation curve after half of epochs (please see Figure 35). It contributed to keep the validation losses in lower values and avoiding drastic rises from the training trend. The cyclic scheduler supported a continue increasing learning from the beginning until the first half of the cycle and, according to Smith (2017), it allowed to eliminate the search of the best global LR value and schedule. The Adam optimizer have been used in several experiments for

classification, patch-based and semantic segmentation applied to forest damage and tree mortality topics (Hamedianfar *et al.*, 2023). Due to the size of the dataset, the optimization of weights during the backpropagation was probably useful to reduce the propagated errors. The Table 8 and Table 9 summarize the metrics that result from the training and validation processes for each experiment.

Table 8. Performance of the ResNet-18 model for predicting the different health status of the training dataset (only Aerial).

Training Loss	Validation Loss	Health Status Class	Per class			Weighted		
			Precision	Recall	F1	Precision	Recall	F1
0,0167	0,3519	DV	0,5102	0,5555	0,5319	0,6389	0,6500	0,6430
		LV	0,5208	0,4545	0,4854			
		H	0,8437	0,9000	0,8709			

Table 9. Performance of ResNet-18 and ViT models for predicting the different health status of the training dataset (Aerial+S2).

Training Loss	Validation Loss	Health Status Class	Per class			Weighted		
			Precision	Recall	F1	Precision	Recall	F1
0,0528	0,2214	DV	0,5454	0,5454	0,5454	0,6673	0,6562	0,6607
		LV	0,5454	0,6000	0,5714			
		H	0,9000	0,8181	0,8571			

The weighted precision for the prediction of the tree classes was about 63,89% (only aerial) and 66,73% (aerial+S2), showing an overall low performance during the training and a higher precision of the fused model that could be related to the higher number of spectral features provided by the S2 images. As part of the expected results, the class with the highest number of samples obtained the best precision, while the other less represented classes obtained a lower precision. This result can be explained either for the higher number of samples of the Heathy class or the spatial and spectral characteristics of these trees. Spatially, there were more H trees with larger crowns that occupied more space in the image patch (Figure 34), and the infrared, red and green values were higher for this class as shown before in Figure 33. In addition, it is possible that the normalization done in the dataloaders step could benefit more the prediction of labels from class H and the amount of false positives were kept in a lower proportion.

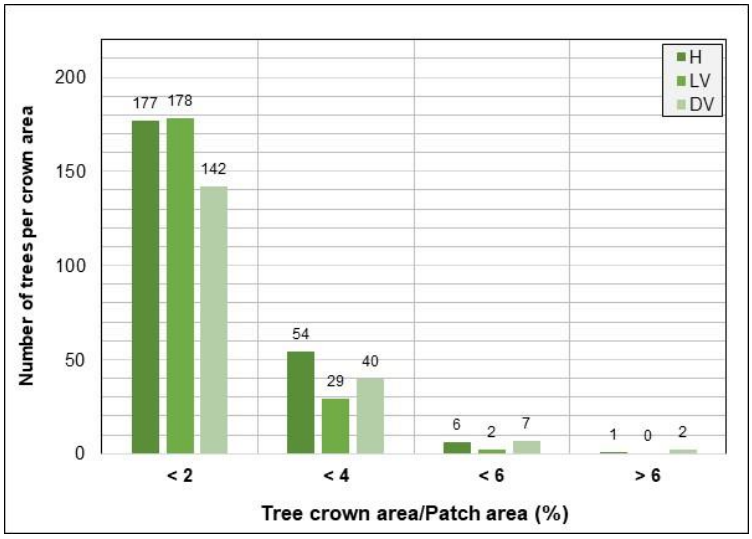

Figure 34. Number of trees per crown area in image patches of training dataset.

The weighted recall, that considered the support for each class or the number of actual occurrences (frequency) of the class in the training dataset (Leung, 2023), probably indicates that most of the half of the predictions corresponded to the correct class in both experiment trainings. When analysing the closeness between precision and recall values, it is possible to affirm that the model classification was not highly accurate in both experiments, but it correctly classified a considerable proportion of the dataset with an F1 score very close to both precision and recall. The first experiment showed recall values higher than precisions in the classes DV and H, which could mean that the model is generating more false negatives than false positives. On the other hand, in the experiment with combined sources, the recall values of class H could imply that the model correctly predicted much more class H patches (true positives) and just a part of the dataset was incorrectly classified as non-healthy or non-H class (false negatives), showing that Healthy class was well distinguished by the model and its spatial and spectral properties were kind of unique for the learning. In contrast, the recall values for classes LV and DV (Table 9) could mean that the model identified a similar proportion of true positives and false negatives, but the recall of class LV probably mean that the model identified more features in these image patches compared to the class DV despite that the dataset contained more samples of DV trees with slightly bigger crowns. This could mean that a significant difference

in crown size could force the model to learn more features of a specific class and predict it correctly.

Additionally, the closeness between DV and LV spectral profiles could cause that the network did not learn the specific features of each class as done with the class H. In most of the dataset, there were less trees in classes LV and DV with tree crown area occupying less than 6% of the whole patch, and just two trees in the DV class had an area higher than 6% of the input image. Therefore, most of the trees from unhealthy classes had a reduced area and the learning process of the model could be biased due to the DV and LV features. Most of the DV trees were completely or almost completely defoliated and urban built-up pixels were partially visible, and this specific characteristic could confuse the network. However, as the complete training dataset had trees with crown that occupied less than 7% of the whole image patch, it could be possible that the spatial features were not as relevant as the spectral information of the image bands, contrary to the results of Ahlswede *et al.* (2022) where the spatial content of sources was more relevant that the spectral content for the task of species classification. An advantage of the TreeSatAI dataset was that the image patches were sampled from forest plots data of forest administration, it means a continue cover of trees up to three species that occupied the whole image patch (multi-class), and then there were continue and larger patterns of spatial and spectral features that can be learned by the network in order to classify a tree. Here, image patches of the same size (60 x 60 m) mainly consisted of urban cover with a partial and minor percentage of area covered by urban trees up to 7,4% of the whole patch. Therefore, the number of features of the image patches were probably less. However, they provided a higher variability of values, mainly about the spectral side (urban vegetation, built-up structures, cars, etc.). Forest canopy is very dense and diverse (Safonova *et al.*, 2019) and this reality is partially transferred also to the urban context where exist isolated trees and community of trees that, despite the maintenance activities developed by local administrations (e.g. pruning and fertilization), still preserves much diversity in the crown shape and the spectral characteristics of leaves.

The results of the model performance were comparable with the results achieved by Guerra *et al.* (2021) with the use of UAV images, considering dead trees as one of the classes, and suggesting that the red band was the most relevant for detecting defoliation. However, the NIR band could be more relevant in this study, especially for the predictions of the Healthy class, due to the characteristics seen during the sampling process and also in relation to the results achieved by Ahlswede *et al.* (2022) where the NIR band influenced positively the training performance from scratch. Furthermore, the use of images from two epochs of the spring could be useful because it provided a larger set of spatial and spectral features that the network could identify than if only used images from one epoch, which coincide with the research developed by Natesan *et al.* (2020). Urban trees vitality can vary from time to time due to resistance patterns of specific species (Brune, 2016) as well as different treatments made by the local environment departments for improving their vitality. Therefore, the sample of temporal images for vitality classification tasks is strongly dependent of well updated tree inventories that, as possible, contains information of the observations made in previous years about tree vitality with its respective timeframe, and the characteristics of the health status observed (e.g. status of bark, leaves and roots). In the current study, several image patches were dismissed due to the lack of a timeframe for vitality observations, and regarding Neural Networks, the more the amount of inputs too feed the model, the better will be the performance because it provides more opportunities to the algorithms for understanding and learning the input features (Brownlee, 2022).

The presence of shadows could be a factor that affected the model performance. While most of the trees under shadows were removed from the dataset, either because they fell completely behind a house or big tree shadow, it could not be completely tackled because it would reduce substantially the inputs of the model. As mentioned by Hamdi *et al.* (2019), the model prediction can miss areas of the patch covered by shadows, and the normalization of data carried on during the dataloaders step could dismiss some pixel values of tree canopy. In this order of ideas, the textural properties could be a good integration to the model input for dismissing the shadow influence on images and, according to Abdollahnejad & Panagiotidis (2020), it can be of great importance for classifying non-defoliated

tree species due to the textural characteristics of healthy tree crowns. However, for the current study, the class H already got an acceptable precision on class prediction and, according to these authors, the distinguishment of species in tree classification tasks could not be achieved on dead trees, what could imply that texture sources could not be useful for distinguishing DV and LV trees. In contrast, Safonova *et al.* (2019) founded a very low loss value (<0.001) for health status classification in forest affected by Bark Beetle, explaining that their results could be due to the model characteristics for distinguishing texture, color and shape of the four assessed classes. Regarding the current study, the availability of textural information was limited to the spatial resolution in the case of Sentinel-1 sources and also due to the previous and not so favorable results in a similar study location (Ahlswede *et al.*, 2022), but probably TerraSAR-X or UAV textural images could provide much more information about canopy and bark structure in different vitality states and during different times of the day, especially during the heat waves at night that characterizes the Heat Island Effect.

In addition to the texture source, the CHM could also provide an additional contribution to the model performance. According to Timilsina *et al.* (2020), the use of CHM, Google Earth images and LiDAR sources could generate very good results of overall accuracies for mapping tree cover in urban contexts. Contrary to this, Astola *et al.* (2021) found that the use of CHM together with S2 images reduced the model accuracy. However, the use of these models along aerial sources could be a starting point to estimate the performance of CNN models together with sources of high spatial resolution that are available for use in Germany.

The learning process monitored through the curves represented in the Figure 35 shows a specific behavior for both steps of Training and Validation. The training loss curve showed that the model learnt different features from the input images, reaching values of 0,01 and 0,05 for the models with only aerial inputs and aerial with S2, respectively. The number of samples taken in count could be related to the stability of the training and the reduction of misclassification during the prediction. Additionally, the consecutive breaks in the training loss curves could be related to the LR scheduler that increases the learning rate during the first half of

the training cycle, then the model could learn most of the spatial and spectral features during the first half of the cycle, followed by the second half with less consecutive breaks. It also could be related to the backpropagation that reduced the propagated errors and less features to learn in the second cycle of the training process. On the other hand, the validation loss curve mainly showed the influence of the low number of samples used in both models, with notorious trend drops and high divergence from the training loss in the case of the ResNet model trained only with aerial sources, and showing more stability in the case of the fused-source model. However, it was noticed that after the first half cycle of the training, the validation curve does not follow the same trend than the training curve, especially shown in the Figure 35 (left), which could mean either the validation samples were not enough for validating the model or the samples did not represent good enough features compared to the training dataset.

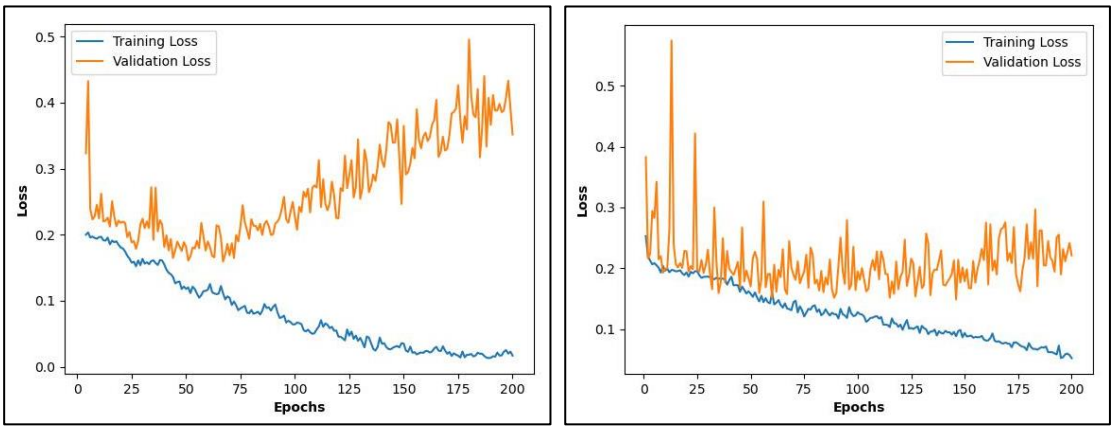

Figure 35. Training and Validation loss curves for the experiments with only aerial images (left) and aerial with S2 images (right).

Due to the variability of tree physical characteristics (e.g. crown, distribution of branches, and shape of leaves) and their distribution in the city as isolated trees or as part of a community of trees, it is probable that the validation dataset could not contain many features compared to the training dataset, and this aspect was a challenging topic to cope during the sampling process. The combination of aerial and S2 sources could provide more spectral information to the fused model, especially providing data in the red-edge, NIR and SWIR channels. However, it could not be enough to keep the validation curve close to the trend of the training loss.

According to Saxena (2023), the BCE loss function penalizes the probabilities based on the distance from the expected value, comparing each of the predicted probabilities of the actual class output for being 1 or 0 (e.g. DV or not DV). Therefore, the training process was also influenced by the loss function set in the model (BCEWithLogitsLoss) because it received the weight classes that had been previously balanced (argument). Then, it is probable that the function penalizes much more the probabilities for getting a patch that represented the class H than classes LV or DV, because the number of spatial features (bigger crowns) and spectral properties (higher values of green and NIR) learned by batch could affect positively the way that the model generate the probabilities to predict the healthy class, while the lack of high spectral values in the infrared and green channels, the crown characteristics of non-healthy trees and the less number of representative features from classes DV and LV could make the model to calculate lower predicted probabilities per epoch and the model could not penalize these values as well as with the class H, even with the optimization.

The behavior of the curve in the Figure 35 (right) was compared to the results achieved by Saleem *et al.* (2020), where the validation loss curve of a ResNet-50 model generated a similar distancing from the training curve and achieved similar training and validation losses at the end of the learning. Another similar behavior of the validation loss curves of the current study, specially from the experiment with only aerial inputs (Figure 35, left), was achieved by Hamdi *et al.* (2019), Natesan *et al.* (2020) and Shtanchaev *et al.* (2021), where the loss curves were getting farer from the training curve after an overfitting phase and was explained by a failure in the generalization of the model. This explanation could be also applicable to the current results due to the size of the dataset and the attempt of the model to make predictions on noisy data (Goyal, 2023) as the shadows projected from the trees that can alter the pixel values of some crowns. In contrast, Safonova *et al.* (2019) obtained a validation loss curve that followed the trend of the training curve and avoided to fail in the generalization, despite it presented an increase in the difference between both curves at epoch 23, and presented a similar difference at the end of the training.

4.2.2 Testing

The models trained previously were tested using a reduced number of samples (only 35 images per source). The final weights resulted from training processes were used for the classification during the testing. The following tables summarize the results of the testing step for both experiments:

Table 10. Performance of model with only aerial inputs for predicting the different health status of trees on the testing dataset.

Vitality Class	Per class			Weighted		
	Precision	Recall	F1	Precision	Recall	F1
DV	0,4166	0,5000	0,4545	0,6250	0,6176	0,6024
LV	0,6666	0,3636	0,4706			
H	0,7500	0,9231	0,8276			

Table 11. Performance of model with aerial and S2 inputs for predicting the different health status of trees on the testing dataset.

Vitality Class	Per class			Weighted		
	Precision	Recall	F1	Precision	Recall	F1
DV	0,7143	0,5000	0,5882	0,7469	0,6857	0,7109
LV	0,5454	0, 5454	0, 5454			
H	0,9286	0,9286	0,9286			

The weighted precision of the first experiment shown in the Table 10 was lower than the precision shown for the experiment of combined sources in the Table 11. This weighted precision achieved by the model with late fusion (74,69%) was satisfactory despite the number of samples used for feeding the model, with a recall and F1 score a bit lower that describes the background of the availability of samples (mainly for the classes DV and LV) and a higher number of false positives than false negatives for unhealthy classes. This result could be linked to the use of S2 patches that contributed to predict much better the studied classes, and also to specific spatial and spectral features of the aerial dataset. The per class precisions followed the trend of the training, with precisions proportional to the vitality class. However, the experiment with only aerial inputs yielded recall values higher than the precisions for classes DV and H, which could mean that the model was generating more false negatives than false positives in these classes.

Regarding the number of samples used during the testing in the experiment with both image sources, it was expected a similar output from the model compared to the training, especially in the class H where the precision reached more than 90%. However, the precision in the prediction of class DV was higher than in the class LV, which could be explained for the spectral characteristics of the patches used in this step and probably for the spatial distribution of the tree structures in the path. Even with this result, the recall and F1 score for the class DV (Table 11) kept much lower indicating that approximately the half of the DV testing data was incorrectly predicted (false negative) and also that the network could not distinguish properly the DV samples from LV patches.

The resulting classified outputs of the experiment that showed the highest weighted precision (aerial+S2) were analysed in order to understand the possible reasons of such results. Initially, it is possible that the presence of communities of healthy trees in the H patches could provide more information to learn regarding higher pixel values and more extended tree cover (Figure 36, left). Additionally, the presence of green areas as grass gardens and city parks with high infrared/green reflection could address the model to a better performance for predicting this class. On the other hand, the size of the crowns seems to influence the model performance probably for the number of pixel and mean values that represent the class (Figure 36, right).

Figure 36. H label correctly predicted as class H (left) and incorrectly predicted as class LV (right).

Regarding the class LV, the model did not perform as well as the class H (precision 54,54%) probably due to the diversity of trees and green spaces located in the patch. The model performed properly on patches with isolated or partially isolated trees which pixel values and crown shape represented properly the class LV (Figure 37, left), but some patches were incorrectly predicted (false negatives) probably due to the presence of other green surfaces (i.e. grass garden areas, living fences, and green roofs) with higher or lower response to NIR and green spectrum, which could make the model to classify the patch as another class (Figure 37, right).

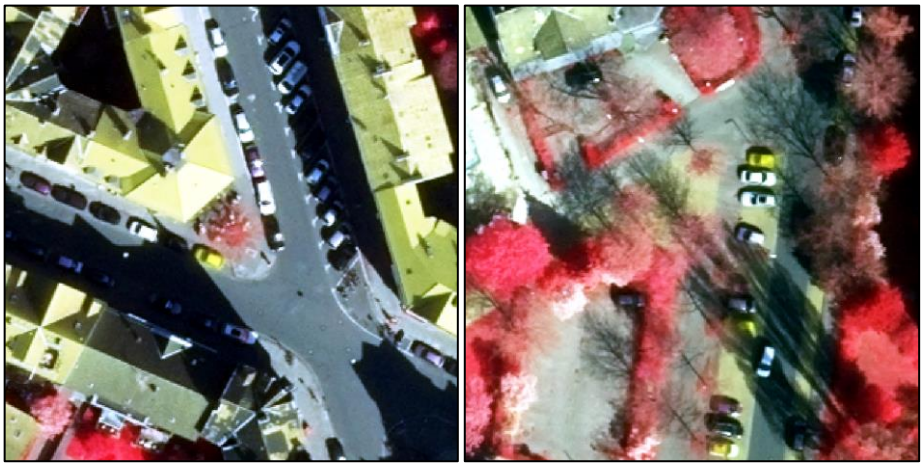

Figure 37. LV label correctly predicted as class LV (left) and incorrectly predicted as class DV (right).

On the other hand, the model predicted the class DV with a precision of 71,43% which was unexpected due to the number of samples of the dataset. An explanation for this result could came from the spatial properties of the image patch because the correctly classified patch presented a canopy over a built up surface despite the presence of other minor green areas (Figure 38, left), contrary to the wrongly classified patch that contained more diversity of tree canopy shapes and trees in community with higher spectral responses (Figure 38, right). It is probable that the normalized values of the patch showed in the Figure 38 (right) were higher than the patch showed in the Figure 38 (left), and could be more related to a class LV. However, it was expected that the size of the tree crown and the presence of visible branches in the figure of the right could be enough information for the network to learn these characteristics. This observation was one of the

most challenging aspects to tackle in the current study, because the urban trees are usually surrounded by other trees as part of a living community for providing continue shadow and cooling the environment. Additionally, the S2 images could provide more information of the spectral behavior of the vegetation in a wider range of the infrared spectrum under a lower spatial resolution. In contrast, the response of the urban surface could be providing some noise to the learning phase.

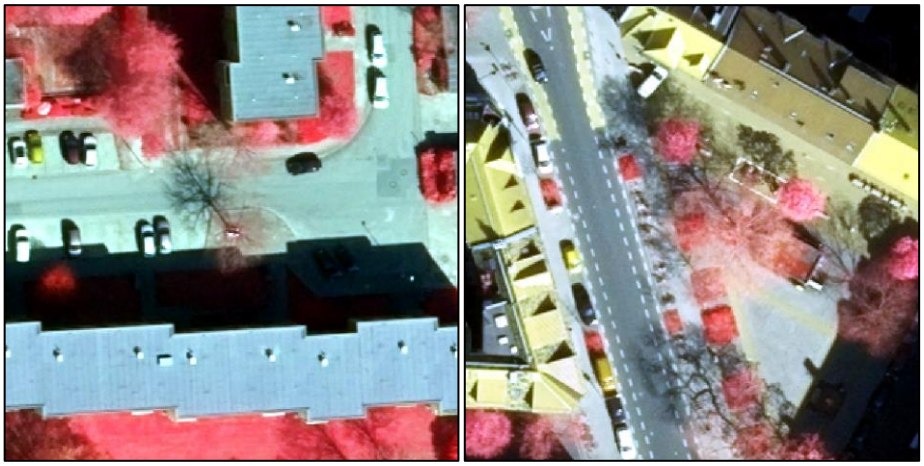

Figure 38. DV label correctly predicted as class DV (left) and incorrectly predicted as class LV (right).

The results achieved for the classes studied are related to the results of the research made by Ahlswede *et al.* (2022), where the most frequently-occurring classes were correctly predicted and the other less represented classes were not well predicted at all, with some exceptions for a specific class of tree that probably had easily identifiable features that the model learned and differentiated from the other classes. Regarding this last suggestion made by these authors, the current study showed that even the less represented class could present some features in the testing dataset that the model distinguished properly to set it apart from the classes H and LV.

5 Conclusions and Recommendations

The study of the application of Deep Learning techniques on aerial and Sentinel-2 image datasets showed a better performance for the classification task compared to the use of only aerial images. This work could contribute to extend the frame of the DL studies on forest and urban tree tasks and speed-up the monitoring of tree species by local administrations. However, some differences related to the urban context and the focus of the study should be taken in count, especially due to the extent of vegetation cover in image patches and the diversity of features in the built-up surface.

The training step yielded a weighted precision of 63,89% for the experiment of only aerial and 66,73% for the experiment of combined sources, indicating an overall low performance of the learning process, strongly affected by the number of samples introduced in the model. However, the training curves achieved low loss values which meant a good learning process through the epochs but with differences with the validation curves which could be a symptom of less penalization on the probabilities of classes LV and DV, and also of the size and the noise of the validation dataset that could cause an issue in the generalization.

The prediction of the H class achieved the highest precision for both experiments, especially due to the number of samples, spatial features as larger crowns, location within tree communities, and spectral characteristics as higher values on green and infrared channels. The low precision in the classification of classes LV and DV could be strongly related to the spatial properties of these trees as smaller crowns, lower pixel values in the green and infrared channels, and characteristic pixels of built-up areas within the canopy. However, it is probable that significant differences in the crown size could affect the model performance.

The testing step showed a weighted precision of 74,69% for the experiment with aerial and S2 inputs that was satisfactory for the process and the conditions of the samples collected. The prediction of classification for the classes H, DV, and LV was 92,86%, 71,43% and 54,54% respectively. The results achieved during the testing process of this experiment were a reflex of the spatial and spectral information in the image patches of class H, but also of the similarity of features

in both unhealthy classes that made it difficult for the model to differentiate the image patches, generating incorrect predictions. Additionally, the presence of communities of trees and green surfaces below or surrounding the canopy could generate noise in the model.

The use of aerial and Sentinel-2 sources from two different epochs introduced a higher variability in the dataset of the model. The variation between epochs of canopy density and diversity provided by the aerial image patches could contribute to introduce a wider range of features that the model could identify. The satellite image patches could provide more information in the red-edge, NIR and SWIR bands for distinguishing types of vegetation affected by heat stress.

Finally, it is necessary to take some considerations before the sampling process for the task of tree vitality classification through DL models. These considerations are related to the use of GIS and remote sensing techniques that will determine the quality of the samples used to feed a model. In the current study, the preparation of data was a crucial step to feed the model with the best representative samples of the studied classes. For obtaining patches that correctly represented the vitality of each tree, it was necessary to have at least a tree inventory that contained information of health status and that could constitute a first approximation for sampling the images. Additionally, it would be highly recommended to consider more information from the tree inventory database, as the epoch of the vitality assessment, the year of plantation, height, and crown area, in order to keep a temporal record of observations in situ that can be related to temporal datasets of images. It could strengthen the sampling process and increase considerably the number of good quality samples. Abdollahnejad & Panagiotidis (2020) considered that the use of thermal sensors could improve the classification of different health status, in addition to the use of UAV paired with multispectral sensors that can reduce the operational costs during the monitoring of forested areas on small to medium scales. Moreover, when a proper database filtering has been done to obtain the tree IDs for including in the study, the UAV flight missions during different epochs or in several days of lately spring, summer, or early fall could conduct to a consistent and numerous dataset.

Another important source that has been used in different investigations is the texture information from SAR instruments (Safonova *et al.*, 2019; Abdollahnejad & Panagiotidis, 2020; Ahlswede *et al.*, 2022; Liu, 2022). Despite the positive or not so positive results of the studies that involve texture as input data for the models, it always provides relevant information that is strongly related to the defoliation patterns of vegetation that also are linked to the health status of trees. Furthermore, there would be no limitations related to the day light or shadow noise for recording images due to the properties of SAR instruments. This could be complemented with *in situ* observations of the state of the tree trunk and roots, or even complementing the dataset with close range images as starting point to train the model, because even trees with a good and healthy canopies can present problems in other structures as shown in the Figure 39.

Figure 39. Status of tree trunk (Favoritepark, Ludwigsburg).

The use of NDVI and CHM (Timilsina *et al.*, 2020) along aerial images could add valuable information about the data stored in the sources because it could contribute to set a threshold for the dataset features. In this way, the model could learn properly what it needs to learn and dismiss features that are not required to learn for further prediction as, for example, built-up surfaces and garden grasses. Additionally, the use of software for feature extraction could optimize the crown area calculation of the image patches in very large datasets, at least for non-defoliated trees.

The use of techniques for data augmentation as SMOTE (Synthetic Minority Over-sampling Technique) could be very helpful to increase the size of the dataset regarding less represented classes. It could contribute to have a wider variety of features for the learning process. Hence, an increase in the size of the dataset could allow the implementation of deeper models that can bring a better performance for image classification as shown by He *et al.* (2015).

The considerations about the specific field of application are also very relevant to keep the consistency of the study. In this research, the focus and knowledge of urban trees were important to understand the datasets used, the dynamics of urban trees and the possible explanations of the results obtained. The difference between urban trees and forests should be also kept in mind due to specific properties of each type of tree and its surrounding environment, but also there are some characteristics as canopy density and diversity that are partially shared with urban trees and that could influence the way that the network learns the features. In this aspect, another suggestion could be to work with tree communities that share similar characteristics about health status. Despite not all urban trees are grouped in communities, it is very common to find this kind of distribution in cities and then the area occupied by a community of trees from the same class on an image patch could augment the size and amount of the features that the network is aimed to learn.

References

Abdollahnejad, A., & Panagiotidis, D. (2020). Tree Species Classification and Health Status Assessment for a Mixed Broadleaf-Conifer Forest with UAS Multispectral Imaging. *Remote Sensing*, *12*(22), 3722. https://doi.org/10.3390/rs12223722

ADAPT. (2022, June 14). *Stuttgart: combating the Heat Island Effect and poor air quality with ventilation corridors and green-blue infrastructure*. https://climate-adapt.eea.europa.eu/metadata/case-studies/stuttgart-combating-the-heat-island-effect-and-poor-air-quality-with-green-ventilation-corridors

Ahlswede, S., Schulz, C., Gava, C., Helber, P., Bischke, B., Förster, M., Arias, F., Hees, J., Demir, B., & Kleinschmit, B. (2022). TreeSatAI Benchmark Archive: A multi-sensor, multi-label dataset for tree species classification in remote sensing. *Earth Syst. Sci. Data Discuss*, *[preprint].* https://doi.org/10.5194/essd-2022-312

Asenova, M. (Ed.) (2018). *GIS-based analysis of the tree health problems using UAV images and satellite data*. SGEM.

Astola, H., Seitsonen, L., Halme, E., Molinier, M., & Lönnqvist, A. (2021). Deep Neural Networks with Transfer Learning for Forest Variable Estimation Using Sentinel-2 Imagery in Boreal Forest. *Remote Sensing*, *13*(12), 2392. https://doi.org/10.3390/rs13122392

Awan, A. A. (2023, January 14). A Complete Guide to Data Augmentation. *DataCamp*. https://www.datacamp.com/tutorial/complete-guide-data-augmentation

Beery, S., Wu, G., Edwards, T., Pavetic, F., Majewski, B., Mukherjee, S., Chan, S., Morgan, J., Rathod, V., & Huang, J. (2022). *The Auto Arborist Dataset: A Large-Scale Benchmark for Multiview Urban Forest Monitoring Under Domain Shift.* https://doi.org/10.1109/CVPR52688.2022.02061

Boesch, G. (2023, January 25). Vision Transformers (ViT) in Image Recognition – 2022 Guide. *Viso.Ai*. https://viso.ai/deep-learning/vision-transformer-vit/

Borelli, S. (2016). *Benefits of Urban Trees.* FAO.

Brownlee, J. (2023, February 4). Impact of Dataset Size on Deep Learning Model Skill And Performance Estimates. *Machine Learning Mastery*. https://machinelearningmastery.com/impact-of-dataset-size-on-deep-learning-model-skill-and-performance-estimates/

Brune, M. (2016). *UrbanTrees under Climate Change: Potential impacts of dry spells and heat waves in three German regions in the 2050s*. Report 24. Climate Service Center Germany.

Callow, D., May, P., & Johnstone, D. (2018). Tree Vitality Assessment in Urban Landscapes. *Forests*, *9*(5), 279. https://doi.org/10.3390/f9050279

Cates, A. (2022, September 2). *The connection between soil organic matter and soil water.* University of Minesota Extension. https://blog-crop-news.extension.umn.edu/2020/03/the-connection-between-soil-organic.html

Caudullo, G., & Rigo, D. de. (2016). *Acer platanoides in Europe: distribution, habitat, usage and threats* (European Atlas of Forest Tree Species). Luxembourg. European Commission.

Češljar, G., Jovanović, F., Brašanac-Bosanac, L., Đorđević, I., Mitrović, S., Eremija, S., Ćirković-Mitrović, T., & Lučić, A. (2022). Impact of an Extremely Dry Period on Tree Defoliation and Tree Mortality in Serbia. *Plants (Basel, Switzerland)*, *11*(10). https://doi.org/10.3390/plants11101286

Daly, C. (2023, February 9). *Matlab deep learning/ResNet-18: Repo for ResNet-18.* GitHub. https://github.com/matlab-deep-learning/resnet-18

Dickson, B. (2023, January 10). *Why we must rethink AI benchmarks*. https://bdtechtalks.com/2021/12/06/ai-benchmarks-limitations/

Doshi, S. (2023, January 13). Various Optimization Algorithms For Training Neural Network. *Towards Data Science*. https://towardsdatascience.com/optimizers-for-training-neural-network-59450d71caf6

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., &

Houlsby, N. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.

DWD. (2023, January 10). *Urban heat islands*. https://www.dwd.de/EN/research/climateenvironment/climate_impact/urbanism/urban_heat_island/urban-heatisland_node.html

Eaton, E., Caudullo, G., Oliveira, S., de Rigo, D. (2016). *Quercus robur and Quercus petraea in Europe: distribution, habitat, usage and threats.* (European Atlas of Forest Tree Species). Luxembourg. European Commission.

Ellison, D., Morris, C. E., Locatelli, B., Sheil, D., Cohen, J., Murdiyarso, D., Gutierrez, V., van Noordwijk, M., Creed, I. F., Pokorny, J., Gaveau, D., Spracklen, D. V., Tobella, A. B., Ilstedt, U., Teuling, A. J., Gebrehiwot, S. G., Sands, D. C., Muys, B., Verbist, B., . . . Sullivan, C. A. (2017). Trees, forests and water: Cool insights for a hot world. *Global Environmental Change, 43*, 51–61. https://doi.org/10.1016/j.gloenvcha.2017.01.002

Environmental Planning and Management Unit of Hannover & GEONET. (2017). *Analyse der klimaökologischen Funktionen und Prozesse für das Stadtgebiet von Hannover*. Hannover. Landeshauptstadt Hannover.

EOS. (2023, January 10). *NDVI FAQs: Frequently Asked Questions About The Index.* EOS Data Analytics. https://eos.com/blog/ndvi-faq-all-you-need-to-know-about-ndvi/

ESA. (2022, October 17). *Sci Open Access Hub*. https://scihub.copernicus.eu/dhus/#/home

ESA. (2023a, January 10). *Data Products.* ESA. https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2/data-products

ESA. (2023b, February 22). *Heatwaves and climate change.* ESA. https://climate.esa.int/en/news-events/heatwaves-and-climate-change/

Eurostat. (2022, September 29). *Glossary:Degree of defoliation*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Degree_of_defoliation

Fite, K. (2008). *Impacts of Root Invigoration and its Individual Components on the Performance of red maple (Acer rubrum)* [Ph.D. Thesis]. Clemson University, Clemson, SC, USA.

Fricker, G. A., Ventura, J. D., Wolf, J. A., North, M. P., Davis, F. W., & Franklin, J. (2019). A Convolutional Neural Network Classifier Identifies Tree Species in Mixed-Conifer Forest from Hyperspectral Imagery. *Remote Sensing*, *11*(19), 2326. https://doi.org/10.3390/rs11192326

Gazol, A., & Camarero, J. J. (2022). Compound climate events increase tree drought mortality across European forests. *The Science of the Total Environment*, *816*, 151604. https://doi.org/10.1016/j.scitotenv.2021.151604

Geodatenportal Niedersachsen. (2022, October 14). Geodatensuche Niedersachsen. https://geoportal.geodaten.niedersachsen.de/harvest/srv/ger/catalog.search#/search?resultType=details&sortBy=relevance&from=1&to=20

Geodatenzentrum. (2023, February 23). Open Data. https://gdz.bkg.bund.de/index.php/default/open-data.html

Gillner, S., Bräuning, A., & Roloff, A. (2014). Dendrochronological analysis of urban trees: climatic response and impact of drought on frequently used tree species. *Trees*, *28*(4), 1079–1093. https://doi.org/10.1007/s00468-014-1019-9

Goyal, C. (2023, February 5). Complete Guide to Prevent Overfitting in Neural Networks (Part-1). *Analytics Vidhya*. https://www.analyticsvidhya.com/blog/2021/06/complete-guide-to-prevent-overfitting-in-neural-networks-part-1/

Guerra-Hernández, J., Díaz-Varela, R. A., Ávarez-González, J. G., & Rodríguez-González, P. M. (2021). Assessing a novel modelling approach with high resolution UAV imagery for monitoring health status in priority riparian forests. *Forest Ecosystems*, *8*(1). https://doi.org/10.1186/s40663-021-00342-8

Hamdi, Z. M., Brandmeier, M., & Straub, C. (2019). Forest Damage Assessment Using Deep Learning on High Resolution Remote Sensing Data. *Remote Sensing*, *11*(17), 1976. https://doi.org/10.3390/rs11171976

Hamedianfar, A., Mohamedou, C., Kangas, A., & Vauhkonen, J. (2022). Deep learning for forest inventory and planning: a critical review on the remote sensing approaches so far and prospects for further applications. *Forestry: An International Journal of Forest Research*, *95*(4), 451–465. https://doi.org/10.1093/forestry/cpac002

Hannover Department of Environment and Urban Green Space. (2022a). *Baumkataster* [Shapefile].

Hannover Department of Environment and Urban Green Space. (2022b). *Lufttemperatur in 2m über Grund um 4 Uhr (°C)* [ESRI ArcInfo Grid].

Hannover Division of Environmental Protection. (2017). *Living with Climate Change - Hannover Adapts* (Publication Series Municipal Environmental Protection No. 53). Hannover. Landeshauptstadt Hannover.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition.* Microsoft.

IBM. (2023a, January 10). *What are Convolutional Neural Networks?* | *IBM.* IBM. https://www.ibm.com/topics/convolutional-neural-networks

IBM. (2023b, January 13). *Hyperparameter tuning with Deep Learning Impact.* IBM. https://www.ibm.com/docs/en/wmla/1.2.3?topic=features-hyperparameter-tuning

IPCC. (2022). *Regional Fact Sheet. Urban areas.* IPCC.

Jim, C. Y. (2019). Soil volume restrictions and urban soil design for trees in confined planting sites. *Journal of Landscape Architecture*, *14*(1), 84–91. https://doi.org/10.1080/18626033.2019.1623552

Jing, P., Wang, D., Zhu, C., & Chen, J. (2016). Plant Physiological, Morphological and Yield-Related Responses to Night Temperature Changes across Different Species and Plant Functional Types. *Frontiers in Plant Science*, *7*, 1774. https://doi.org/10.3389/fpls.2016.01774

Kostadinov, S. (2023, January 14). Understanding Backpropagation Algorithm. *Towards Data Science*. https://towardsdatascience.com/understanding-back-propagation-algorithm-7bb3aa2f95fd

Landeshauptstadt Hannover. (2023a, October 20). *Bevölkerung.* Landeshauptstadt Hannover. https://www.hannover.de/Leben-in-der-Region-Hannover/Politik/Wahlen-Statistik/Statistikstellen-von-Stadt-und-Region/Statistikstelle-der-Landeshauptstadt-Hannover/Hannover-kompakt/Bev%C3%B6lkerung

Landeshauptstadt Hannover. (2023b, October 20). *Hannover - The green city*. https://www.visit-hannover.com/en/Press/Press-Information-sorted-by-topic/Hannover-The-green-city

LBEG. (2022). *Bodenkundliche Feuchtestufe.* LBEG [Shapefile].

Leung, K. (2022, January 4). Micro, Macro & Weighted Averages of F1 Score, Clearly Explained. *Towards Data Science*. https://towardsdatascience.com/micro-macro-weighted-averages-of-f1-score-clearly-explained-b603420b292f

LGLN. (2021). *Digitale Orthophoto (DOP)*. Hannover. Landesamt für Geoinformation und Landesvermessung Niedersachsen.

LGLN. (2022, October 17). *Digitales Orthophoto (DOP).* LGLN. https://opengeodata.lgln.niedersachsen.de/#dop

Liu, H. (2022). Classification of urban tree species using multi-features derived from four-season RedEdge-MX data. *Computers and Electronics in Agriculture*, *194*, 106794. https://doi.org/10.1016/j.compag.2022.106794

Liu, J., & Pu, Z. (2019). Does Soil Moisture Have an Influence on Near-Surface Temperature? *Journal of Geophysical Research: Atmospheres*, *124*(12), 6444–6466. https://doi.org/10.1029/2018JD029750

Locatelli, B. (2016). *Routledge handbook of ecosystem services*. London, New York.

Lopez, R. (2023, January 10). *How to benchmark the performance of machine learning platforms: data capacity, training speed, inference speed and model precision | Neural Designer*. https://www.neuraldesigner.com/blog/how-to-benchmark-the-performance-of-machine-learning-platforms

Martinez-Trinidad, T., Watson, W. T., Arnold, M. A., Lombardini, L., & Appel, D. N. (2010). Comparing various techniques to measure tree vitality of live oaks. *Urban Forestry & Urban Greening*, *9*(3), 199–203. https://doi.org/10.1016/j.ufug.2010.02.003

Mathworks. (2023, January 10). *What Is a Convolutional Neural Network? | 3 things you need to know.* Mathworks. https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html

Natesan, S., Armenakis, C., & Vepakomma, U. (2020). Individual tree species identification using Dense Convolutional Network (DenseNet) on multitemporal RGB images from UAV. *Journal of Unmanned Vehicle Systems*, *8*(4), 310–333. https://doi.org/10.1139/juvs-2020-0014

NASA. (2023, January 23). *Reflected Near-Infrared Waves | Science Mission Directorate.* NASA. https://science.nasa.gov/ems/08_nearinfraredwaves

Onishi, M., & Ise, T. (2021). Explainable identification and mapping of trees using UAV RGB image and deep learning. *Scientific Reports*, *11*(1), 903. https://doi.org/10.1038/s41598-020-79653-9

Pretzsch, H., Rötzer, T., & Reischl, A. (2018). *Stadtbäume im Klimawandel II: Wuchsverhalten, Umweltleistungen und Perspektiven.* TU München.

Pytorch. (2023, January 10). *ResNet — Torchvision main documentation.* PyTorch Foundation. https://pytorch.org/vision/main/models/resnet.html

Roloff, A. (2019). *Baumpflege: Baumbiologische Grundlagen und Anwendung* (3. Erweiterte Auflage). Stuttgart: Verlag Eugen Ulmer.

Safonova, A., Tabik, S., Alcaraz-Segura, D., Rubtsov, A., Maglinets, Y., & Herrera, F. (2019). Detection of Fir Trees (*Abies sibirica*) Damaged by the Bark

Beetle in Unmanned Aerial Vehicle Images with Deep Learning. *Remote Sensing*, *11*(6), 643. https://doi.org/10.3390/rs11060643

Saleem, M. H., Potgieter, J., & Arif, K. M. (2020). Plant Disease Classification: A Comparative Evaluation of Convolutional Neural Networks and Deep Learning Optimizers. *Plants (Basel, Switzerland)*, *9*(10). https://doi.org/10.3390/plants9101319

Saxena, S. (2023, February 4). Binary Cross Entropy/Log Loss for Binary Classification. *Analytics Vidhya*. https://www.analyticsvidhya.com/blog/2021/03/binary-cross-entropy-log-loss-for-binary-classification/

Schulz, C., Ahlswede, S., Gava, C., Helber, P., Bischke, B., Arias, F., Förster, M., Hees, J., Demir, B., & Kleinschmit, B. (2022). *TreesatAI Benchmark Archive for Deep Learning in Forest Applications: (1.0.1) [Data set].* https://doi.org/10.5281/ZENODO.6778154

Schwaab, J., Meier, R., Mussetti, G., Seneviratne, S., Bürgi, C., & Davin, E. L. (2021). The role of urban trees in reducing land surface temperatures in European cities. *Nature Communications*, *12*(1), 6763. https://doi.org/10.1038/s41467-021-26768-w

Selig, M. (2020). *Untersuchung zur Anpassung verschiedener Populationen von Quercus robur L. und Tilia cordata Mill. anhand physiologischer und biochemischer Reaktionen auf Frost und Trockenheit* [Ph.D.]. Leibniz Universität Hannover, Germany, Hannover.

Shrivastava, I. (2022, November 30). Handling Class Imbalance by Introducing Sample Weighting in the Loss Function. *GumGum Tech Blog*. https://medium.com/gumgum-tech/handling-class-imbalance-by-introducing-sample-weighting-in-the-loss-function-3bdebd8203b4

Shtanchaev, A., Bille, A., Sutyrina, O., & Elelimy, S. *Automated Remote Sensing Forest Inventory Using Satellite Imagery*. Skolkovo Institute of Science and Technology; Ulm University.

Slavík, M., Kuželka, K., Modlinger, R., Tomášková, I., & Surový, P. (2020). UAV Laser Scans Allow Detection of Morphological Changes in Tree Canopy. *Remote Sensing*, *12*(22), 3829. https://doi.org/10.3390/rs12223829

Smith, L. (2017). *Cyclical Learning Rates for Training Neural Networks*. U.S. Naval Research Laboratory.

Sun, D., & Pinker, R. T. (2004). Case Study of Soil Moisture Effect on Land Surface Temperature Retrieval. *IEEE Geoscience and Remote Sensing Letters*, *1*(2), 127–130. https://doi.org/10.1109/LGRS.2004.824749

Tan, D., & Siri, J. (2023, January 9). *Heat is an Urban Killer - United Nations University.* UNU. https://unu.edu/publications/articles/heat-is-an-urban-killer.html

Timilsina, S., Aryal, J., & Kirkpatrick, J. B. (2020). Mapping Urban Tree Cover Changes Using Object-Based Convolution Neural Network (OB-CNN). *Remote Sensing*, *12*(18), 3017. https://doi.org/10.3390/rs12183017

Tourismus Marketing Niedersachsen. (2023, January 22). *Europe's biggest urban forest.* Tourismus Marketing Niedersachsen GmbH. https://www.niedersachsen-tourism.com/experiences/culture-in-niedersachsen/worth-seeing/europes-biggest-urban-forest

Umwelt Bundesamt. (2022, June 14). *BAU-I-1 + 2 Heat stress in urban environments and summer Heat Island Effect*. https://www.umweltbundesamt.de/en/topics/climate-energy/climate-impacts-adaptation/impacts-of-climate-change/monitoring-report-2019/indicators-of-climate-change-impacts-adaptation/cluster-building-industry/bau-i-1-2-heat-stress-in-urban-environments-summer

Yathish, V. (2023, January 13). Loss Functions and Their Use In Neural Networks - *Towards Data Science. Towards Data Science*. https://towardsdatascience.com/loss-functions-and-their-use-in-neural-networks-a470e703f1e9

Appendices

Table A 1. Number of tree species per vitality class and Urban Heat Island zone.

Tree Species (Scientific name)	Number of Trees					
	Heat Island Effect (Wärmeinseleffekt)					
	Moderated (Mäßig)			Low (Schwach)		
	H	LV	DV	H	LV	DV
Acer campestre	3	1	0	1	0	3
Acer monspessulanum	5	0	1	1	0	0
Acer neglectum	1	0	0	0	0	0
Acer negundo	0	0	0	0	1	0
Acer platanoides	31	17	5	36	30	31
Acer pseudoplatanus	0	0	2	2	2	5
Acer rubrum	1	3	0	0	0	1
Acer saccharinum	0	0	0	0	1	3
Aesculus hippocastanum	8	0	0	10	2	4
Aesculus x carnea	4	0	0	10	0	3
Alnus cordata	0	1	0	0	0	2
Alnus incana	0	0	0	1	0	0
Alnus x spaethii	1	0	0	0	0	0
Betula pendula	0	0	1	1	0	2
Carpinus betulus	11	3	1	15	0	2
Corylus colurna	1	1	1	2	0	2
Crataegus laevigata	2	0	0	0	0	1
Crataegus lavallei	0	0	1	0	0	5
Crataegus monogyna	0	0	0	0	2	0
Fraxinus angustifolia	0	0	0	0	0	1
Fraxinus excelsior	1	2	1	0	11	5
Ginkgo biloba	1	0	0	1	0	0
Gleditsia triacanthos	1	1	0	0	1	3
Gleditsia triacanthos inermis	0	0	2	0	0	2
Koelreuteria paniculata	0	0	2	0	0	0
Liquidambar styraciflua	0	0	0	0	0	3
Magnolia kobus	0	0	1	0	0	0
Malus hybrida	0	0	0	0	0	1
Malus sylvestris	0	0	0	0	1	0
Pinus nigra	0	0	0	2	0	0
Platanus x acerifolia	2	18	1	0	4	1
Populus canadensis	0	0	0	0	1	0
Populus nigra	1	0	0	0	0	0
Populus simonii	2	0	0	1	0	0
Prunus x yedoensis	0	0	0	1	0	0
Prunus avium	4	1	0	2	2	3
Prunus padus	0	0	0	2	0	0
Pyrus calleryana	1	0	0	0	0	0
Quercus cerris	0	0	0	0	1	0
Quercus petraea	1	0	0	0	2	0
Quercus robur	3	8	9	12	47	34
Quercus rubra	2	0	0	0	0	3
Robinia pseudoacacia	2	14	6	0	6	7
Sophora japonica	1	0	0	0	0	0
Sorbus aria	0	1	1	0	3	1
Sorbus aucuparia	0	1	0	0	1	0
Sorbus intermedia	7	5	3	0	7	18
Sorbus latifolia	1	2	0	1	2	4

Tree Species (Scientific name)	Number of Trees					
	Heat Island Effect (Wärmeinseleffekt)					
	Moderated (Mäßig)			Low (Schwach)		
	H	LV	DV	H	LV	DV
Tilia cordata	1	3	0	8	3	3
Tilia europaea	3	0	1	17	6	2
Tilia flavescens	1	0	0	0	0	0
Tilia intermedia	12	4	2	17	8	5
Tilia platyphyllos	0	0	0	0	0	3
Tilia spec.	0	0	0	0	1	0
Tilia tomentosa	0	0	1	3	1	4
Ulmus hollandica	0	0	0	1	0	3
Ulmus-Hybride	1	0	0	1	0	0

Table A 2. Soil moisture percentage per tree species, vitality class and Urban Heat Island zone.

Tree Species (Scientific name)	Soil Moisture (%)					
	Heat Island Effect (Wärmeinseleffekt)					
	Moderated (Mäßig)			Low (Schwach)		
	H	LV	DV	H	LV	DV
Acer campestre	2,83-3,24	0		2,83		0-0,31
Acer monspessulanum	0-2,82		2,82	2,82		
Acer neglectum	5,69					
Acer negundo	0-5,69				3,11	
Acer platanoides		0-4,34	0-3,24	0-6,33	0-6,64	0-6
Acer pseudoplatanus			0-3,24	2,82-2,83	3,32-4,34	1,39-6
Acer rubrum	2,82	0				1,48
Acer saccharinum					0	2,82-6
Aesculus hippocastanum	0-6			0-4,29	1,35-2,82	2,83
Aesculus x carnea	0-2,82			2,63-3,48		1,47-4,34
Alnus cordata		1,67				0-1,67
Alnus incana				6		
Alnus x spaethii	2,82					
Betula pendula			0	2,83		1,48-6
Carpinus betulus	0-2,82	0-5,97	0	1,47-6		1,42-1,62
Corylus colurna	2,82	0	0	3,11-6		6-6,31
Crataegus laevigata	2,82					6
Crataegus lavallei			6			0-6
Crataegus monogyna					1,39-1,48	
Fraxinus angustifolia						1,48
Fraxinus excelsior	0	0-2,82	2,82		1,35-6	0-6
Ginkgo biloba	0			2,66		
Gleditsia triacanthos	0	1,67			1,67	0
Gleditsia triacanthos inermis			0			0
Koelreuteria paniculata			0			
Liquidambar styraciflua						2,66-6
Magnolia kobus			0			
Malus hybrida						1,48
Malus sylvestris					6	
Pinus nigra				3,24-5,97		
Platanus x acerifolia	0	0-5,69	5,69		0-6	2,82
Populus canadensis					1,41	
Populus nigra	5,71					

Tree Species (Scientific name)	Soil Moisture (%)					
	Heat Island Effect (Wärmeinseleffekt)					
	Moderated (Mäßig)			Low (Schwach)		
	H	LV	DV	H	LV	DV
Populus simonii	2,82			2,82		
Prunus x yedoensis				3,93		
Prunus avium	3,24-3,81	3,24		2,63-3,73	2,83-6	1,48-3,81
Prunus padus				5,65		
Pyrus calleryana	2,82					
Quercus cerris					2,8	
Quercus petraea	0				1,96-3,17	
Quercus robur	2,82-3,1	0-6	0-6	0-5,69	0-6	0-6
Quercus rubra	0-2,82					2,63-3,76
Robinia pseudoacacia	0	0-6	0-6		0-6	0-6
Sophora japonica	2,82					
Sorbus aria	0-2,82	6	5,97		6	2,31
Sorbus aucuparia		5,64			1,48	
Sorbus intermedia	0-2,82	0-2,82	0-2,82		1,48-6	0-6,33
Sorbus latifolia	0	0		6,31	1,39-6	4,34-6
Tilia cordata	6	0		0-5,97	1,42-6	1,42-6
Tilia europaea	0-2,63		0	0-6,31	1,48-6,31	2,82-6
Tilia flavescens	5,97					
Tilia intermedia	0-6	0-6	0-2,82	0-6	0-4,36	2,82-4,36
Tilia platyphyllos						2,82-6,31
Tilia spec.					6	
Tilia tomentosa			5,96	0	5,97	2,82-6
Ulmus hollandica				2,83		1,38-4,34
Ulmus-Hybride	0			3,31		

Figure A 1. Workflow of Methodology Implementation.