

Delft University of Technology  
Master of Science Thesis in Embedded Systems and Computer Science

# Monitoring the health of urban greenery with terrestrial low-cost, mobile sensors

Akshit Gupta





# Monitoring the health of urban greenery with terrestrial low-cost, mobile sensors

Master of Science Thesis in Embedded Systems and Computer  
Science

Embedded and Networked Systems Group  
Faculty of Electrical Engineering, Mathematics and Computer Science  
Delft University of Technology  
Mekelweg 4, 2628 CD Delft, The Netherlands

Akshit Gupta



April 2022

**Author**

Akshit Gupta

**Title**

Monitoring the health of urban greenery with terrestrial low-cost, mobile sensors

**MSc Presentation Date** 26 April 2022

**Graduation Committee**

dr. RangaRao Venkatesha Prasad	Delft University of Technology
dr. Simone Mora	Massachusetts Institute of Technology
dr. Ujwal Gadiraju	Delft University of Technology

## Abstract

Urban forests and vegetation are fundamental for developing resilient cities. Thus, the effective management and protection of urban trees and greenery are essential. Nowadays, urban trees are experiencing atypical amount of natural and human-induced stresses which affects their functionality, productivity and survival. The current methods for monitoring the health of urban trees mainly comprises of manual inspection by arborists and remote sensing. However, all these methods are riddled with various challenges involving scalability, spatio-temporal resolutions and quality of assessment. The goal of this thesis was to develop a method which can autonomously measure the health of trees on a city-wide scale with high spatio-temporal resolutions at low costs.

To achieve this goal, we first performed an in-depth survey and comparative analysis of the existing state-of-the-art techniques for tree health measurement, identified a research gap and based on this, developed a novel system to measure tree health autonomously from ground level in urban cities. The system can be deployed both in a drive-by sensing paradigm on moving vehicles such as taxis and garbage trucks or be carried by humans in a citizen science paradigm. A computer vision model developed using transfer learning and traditional image processing techniques were employed to fuse the data collected by low cost thermal and multispectral imaging sensors on the edge devices. The approach was evaluated through data collection experiments performed in Cambridge, USA. Comparison with parameters in ground truth datasets revealed several significant relationships. Thus, motivating various studies in the future along with potential large-scale deployment of this technique in cities and municipalities around the world.



# Preface

Problems at the intersection of multiple fields are challenging, while also being intriguing and rewarding. With the proliferation of IoT and AI, the way we view and comprehend urban cities is radically changing. This transformation is further accelerated in part to tackle the challenges induced by climate change and in part to improve people's lives by enabling previously impossible use-cases. For these reasons, I find contributions in this domain important as well as impactful. Managing and preserving nature is essential in the face of climate change and in this thesis, it is highlighted how low-cost sensing and deep learning techniques can be employed to help in the maintenance of existing trees in urban cities.

First, I would like to thank Dr. Simone Mora for providing me with the flexibility of exploring various domains and leading me to this topic. I am grateful to him for his guidance, feedback and motivations imparted during this thesis. I am also grateful to Dr. R. Venkatesha Prasad for his insightful and lively conversations that will extend well beyond this thesis and for handling all the logistical procedures throughout this work.

I would also like to extend my thanks to Yuki Machida for helping me with data collection in Cambridge and Sujay Narayana for arranging any needed hardware in Delft.

I would also like to deeply thank my parents for their frequent impartment of wisdom on flying high and staying grounded. Further, I would like to thank my friends for the fun and support. I am also thankful for the intense running trainings by the X running group and survivalrun trainings along with various activities by S.B.V. Slopend especially during the lockdowns. These trainings helped me to clear my mind during those long days.

I am also grateful to the MIT Senseable City Lab for hosting me in-person during this thesis. Finally, I also appreciate the monetary support provided by the Renswoude Foundation in Delft, FAST and EFL Stichting for my travel expenses and accommodation in Cambridge.

Akshit Gupta

Delft, The Netherlands

April, 2022



Sponsored by:



# Contents

<b>Preface</b>	v
<b>1 Introduction</b>	1
1.1 Context	2
1.2 Problem Statement	3
1.2.1 Requirement analysis	4
1.3 Research Contributions	5
1.4 Thesis Structure	5
<b>2 Problem Background and Preliminaries</b>	7
2.1 Monitoring the health of urban greenery	7
2.2 Inspection Strategies for Greenery health	8
2.2.1 Manual Inspection Techniques	8
2.2.2 Embedded Sensors	10
2.2.3 Imaging Based Methods	13
2.2.4 Comparison	17
2.2.5 Research gaps and influence on design	18
2.3 Metrics (NDVI and CTD)	19
2.4 Image Segmentation	20
2.4.1 IoU	20
2.4.2 Mask R-CNN	21
2.4.3 Transfer Learning	21
2.5 Takeaways	22
<b>3 Materials and Methodology</b>	23
3.1 System Design	23
3.1.1 Hardware	23
3.1.2 Software	25
3.2 System Architecture	26
3.3 Visual Image Processing Pipeline	31
3.4 Development of Custom Mask R-CNN	31
3.4.1 Motivation for a custom instance segmentation model	31
3.4.2 Dataset Creation	32
3.4.3 Training	33
3.4.4 Model Quantization	33
3.5 (Ground truth) Tree Health Dataset	34
3.6 Data Collection Experiments	35

<b>4</b>	<b>Analysis and Results</b>	<b>39</b>
4.1	Mask R-CNN analysis	39
4.1.1	Training Curve	39
4.1.2	Performance	39
4.1.3	Takeaways	42
4.2	Analysis of Tree Health	43
4.2.1	High-level analysis	43
4.2.2	Low-level health condition analysis	45
4.2.3	Takeaways	47
<b>5</b>	<b>Discussion and Future Work</b>	<b>49</b>
5.1	Discussion and limitations of our approach	49
5.1.1	Generalisation of Mask R-CNN performance	49
5.1.2	Generalisation of tree health results	49
5.1.3	Comparison with comparable earlier works	50
5.1.4	CTD analysis	51
5.1.5	Deployment: low speed citizen science data collection vs high speed vehicles	51
5.1.6	Thermal imaging for detecting internal decay in trunks	51
5.2	Future Work	52
5.2.1	Feasibility of modelling based classification	52
5.2.2	Mapping relation between UHI effect with Tree health	52
5.2.3	Follow up study with large-scale evaluation	52
5.2.4	Selective and Smart Irrigation	53
<b>6</b>	<b>Conclusions</b>	<b>55</b>
6.1	Methodology and Results	55
6.2	Research Contributions	56
6.3	Limitations	57
6.4	Future work	57
<b>7</b>	<b>Scientific Papers</b>	<b>59</b>
<b>A</b>	<b>APPENDIX</b>	<b>93</b>
A.1	Settings during Mask R-CNN training	93
A.2	Data Cleaning	93
A.3	Additional Results	95

# Chapter 1

## Introduction

Urban forests, parks and greenery increase cities' resilience due to their ability to mitigate the effects of climate change. As per the latest IPCC Report [13], global warming and climate change presents multiple risks to humans as well as nature. Thus, protecting, managing and restoring ecosystems is fundamental for climate resilient development. In cities, urban canopies and vegetation provide a wide range of ecosystem benefits such as air filtering, carbon sequestration, reduced energy consumption and decreasing local temperatures [52, 58]. Trees mainly reduce local temperatures and heat index by both shading (which decreases the direct incoming net radiation from the sun) and transpiration (which transfers the latent heat energy from the leaf to the atmosphere and in turn reduces the amount of hot air trapped in its surroundings [56]). Both these factors complement each other to enhance the cooling benefits of trees [56]. Further, trees help in mitigating the urban heat island effect (which leads to an air temperature difference as high as 10° C [56] in highly urbanized cities compared to the rural or sub-urban surrounding). Thus, improving the perceived tolerance of the citizens to the environment (Human Thermal Comfort) [56, 41].

However, urban trees are experiencing an ample amount of *abiotic* and *biotic* stressors that are exacerbated due to climate change [53, 75, 30]. As a result, their functionality, productivity, and survival are of increasing concern [98]. Trees with poor health cannot provide most of the aforementioned ecosystem services [54, 57]. For instance, trees with low transpiration rates do not cool the environment sufficiently and trees with low growth rate have reduced shading effect.

Greenery has been a high-maintenance asset because it is constantly changing due to natural and human-induced stresses. Globally, the total cost of inspection, maintenance and settlement of tree damages is estimated to be more than \$2 trillion USD [101]. Large trees also yield higher maintenance costs [54], yet compared to smaller trees, large trees can provide up to 8 times more ecosystem benefits [73]. Nowadays, the health of trees can be inspected by arborists (a tree expert) with good quality results, but usually at high costs, leading to an assessment that has a low spatial and temporal resolution. However, in recent years, a few technology-assisted screening methods have been developed to complement inspection by arborists. Satellite-based imaging can cover large areas although at a low spatial granularity [47], with data quality depending on the availability of clear skies [47] and low-time resolution as satellites can revisit the

same spot only at an interval of few days. Airborne sensing using Unmanned Aerial Vehicles (UAVs) or aeroplanes leads to an increased spatial granularity [47]; yet it involves high operational costs and is not suitable in highly urbanized environments due to aviation authority regulations. Further, both airborne sensing and satellite imagery can only capture the overhead view of the tree canopy. As a result, vegetation elements such as green walls, short trees or shrubs present under the tree canopy are missed [69] or even misinterpreted as part of tree canopy [1].

Recently, several projects have investigated the use of low-cost technological alternatives to monitor parameters of urban environment in cities; for instance, using Google street view (GSV) images to detect the presence of trees such as [68, 85], or using drive-by sensing to measure air pollution in an area [6]. These projects are set within the field of opportunistic sensing and are aimed at developing platforms that can be deployed and operated without the need for an established, expensive infrastructure and human operators.

Following on this trend and the critical need for managing and protecting urban ecosystems, in this thesis, we will try to answer the following research question: how can we measure terrestrially (from ground level) the health of urban trees at a low cost and in an opportunistic fashion at a high spatial and temporal resolution? In the process of answering this question, we will describe the existing state of the art techniques for tree health assessment, identify the research gaps and based on a comparative analysis, design a custom method from scratch to measure tree health from the ground level in urban cities. The method is then evaluated on a real-world dataset revealing several significant relationships. We also hope that this work serves to inspire deeper collaborations between environmental and computer scientists of the future to revolutionise the field of greenery health monitoring.

## 1.1 Context

This research work was performed in collaboration with the Senseable City Lab [19] located at Massachusetts Institute of Technology (MIT) along with the Embedded and Networked systems group [2] at Technische Universiteit Delft (TU Delft). The Senseable City lab deals with research at the interface of cities and technology. An urban platform developed in the lab called **CityScanner** [6] currently measures particulate matter, temperature and humidity in various cities around the world (USA, Sweden, Kazakhstan) with a drive-by sensing approach. Drive-by sensing is an approach where sensors are deployed on small number of everyday vehicles like taxis and garbage trucks as shown in Figure 1.2. As the vehicle on which the sensor is mounted moves, the data for the phenomenon being measured is collected. The number of devices needed to be deployed to cover the entire city is also extremely small. For instance, it was shown that just **ten** random taxis can capture data for one-third of Manhattan streets in a single day using this approach in [76]. This allows drive-by sensing to collect data at both high temporal and spatial resolution compared to stationary and airborne sensing paradigms as show in Figure 1.1.

Considering the advantages of drive-by sensing and the low expected variability of tree health within a single day, the system developed in this work will either be suitable to be deployed independently in a drive-by sensing paradigm

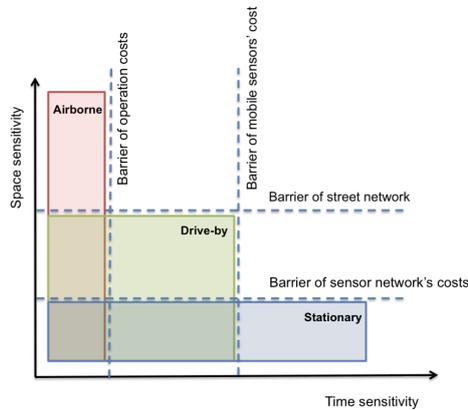


Figure 1.1: Comparison of different sensing paradigms. Stationary sensing is limited with respect to spatial resolution and airborne sensing is limited with respect to temporal resolution (As shown in [76])



Figure 1.2: A CityScanner device (in white box) deployed on top of a tuk-tuk in Stockholm [18]

similar to CityScanner or serve as an additional module to the existing City-Scanner sensing devices or be suitable to be carried by humans in a citizen science based paradigm.

## 1.2 Problem Statement

Monitoring, managing and protecting urban trees and greenery are essential for developing climate adaptive cities of the future. As discussed earlier, the current methods of monitoring urban trees include manual inspection by arborists and remote sensing with satellites and UAVs. The inspection by arborists leads to low spatio-temporal resolution of monitoring. Remote sensing based approaches lead to inspections only from the overhead viewpoint and are susceptible to misinterpreting the vegetation elements present underneath the canopy. Further, they are subjected to various other challenges for deployment in highly urbanised environments such as aviation authority regulations. The main aim of this

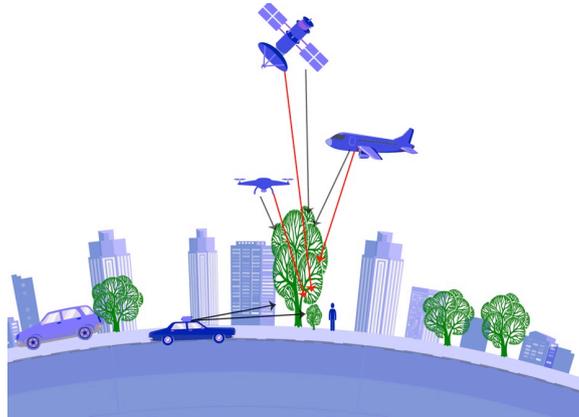


Figure 1.3: **An illustration of the challenges induced due to remote sensing in highly urbanised environments. The smaller trees (or other vegetation elements) underneath the larger trees are missed or misrepresented using strictly overhead analysis by satellites, UAVs or airborne sensing. In contrast, the arborist based approach has low scalability. Drive-by sensing based approach can be a suitable alternative.**

work is to create a system to measure the health of urban trees terrestrially at high temporal and spatial resolutions. To ensure wide scalability, the system should be low cost and suitable for capturing data with as little human intervention as possible or be completely autonomous.

### 1.2.1 Requirement analysis

A complete quantitative model for such a system does not exist and falls outside the scope of this thesis, but the system can be evaluated on a few important parameters which can serve as main requirements:

- **Functional Requirements:**

1. **Mobile sensing:** Recall from the previous section that the drive-by sensing paradigm allows us to achieve both high temporal and spatial resolution. Hence, the system should be suitable for this paradigm by either being deployable on moving vehicles or being small enough to be carried by humans as a citizen science project.

Our literature study directly led to the selection of technologies suitable for this approach and influenced the system design (see Section [2.2.5](#)).

2. **Data Privacy:** The system should not send any privacy-invasive data concerning a citizen or buildings to a centralised server. In other words, the system should not store any uniquely identifiable citizen or building information. This follows from GDPR (General Data Protection Regulation) compliance in EU (European Union) and EEA (European Economic Area).

We fulfilled this in our system by processing all the data on the main single board computer on the edge device (refer Sections 3.4 and 3.3).

- **Non-functional Requirements:**

1. **Cost:** The system should be as low cost as possible (less than \$1000) to make it widely accessible. In other words, the system should not use very expensive sensors over \$800 and should compensate for noise and lower quality hardware sensing using software-based algorithms. In our system, this was achieved using low-cost hardware and data fusion from sensors on the edge device (refer Section 3.1.1).
2. **Quality of Assessment:** The performance of the system with respect to ground truth is not the highest priority, but the ability of the system to quantify the differences amongst trees is the main priority. This follows naturally from the non-availability of terrestrial tree health datasets with daily or monthly resolutions. This is elaborated in depth in Section 4.2

## 1.3 Research Contributions

The main goal of this thesis is to provide researchers with a method to measure the health of trees terrestrially in a low cost opportunistic fashion. The main research contributions of this thesis are three-fold:

1. **Knowledge**

- 1.1. A review and comparison of the existing state-of-the-art methods to measure health of trees (refer Section 2.1). Further, we elaborate on our experience of fusing two of these methods and deploying them in a real world scenario (refer Sections 5.1.6 and 5.1.2).

2. **Technology**

- 2.1. Development of an autonomous system suitable to deployed on moving infrastructure or citizen science paradigms to measure health of urban trees terrestrially in urban cities (refer Chapter 3).
- 2.2. In the process of developing this system, we developed a custom deep learning model for instance segmentation of tree canopies (tree tops) in RGN (red, green, near infrared) images with processing on the edge (refer Section 3.4). This mask R-CNN based approach allowed us to extract useful information from low resolution thermal images.

Deployment of this method on a large scale will contribute towards protecting and managing nature effectively for cities and municipalities around the world.

## 1.4 Thesis Structure

This thesis is structured as follows: in Chapter 2, we survey the state of the art methods and briefly describe the preliminaries. This survey influences the

choice of technology for our system design. Chapter 3 deals with the design and implementation of our system. The data collection experiments performed in Cambridge, USA and ground truth tree health datasets are also described in this chapter. We analyse the data collected and evaluate the effectiveness of our method in Chapter 4. We extend the discussion of our results as well as the effectiveness of our method in Chapter 5 along with the challenges associated with our method. The areas of future work are also identified from these challenges. Finally, we conclude the thesis in Chapter 6.

## Chapter 2

# Problem Background and Preliminaries

This chapter provides background information as well as current state of the art technologies in the context of tree health monitoring in Sections 2.1 and 2.2. A comparison between these technologies is performed and a research gap is identified based on this. This is followed by Sections 2.3 and 2.4 which give an overview of metrics and brief technological theory that will serve as the basis for the upcoming chapters.

### 2.1 Monitoring the health of urban greenery

Like humans, trees have several attributes or parameters that can be monitored to check for illnesses. For instance, just like the human body temperature (attribute) is used to detect health problems which might be the result of an underlying illness; similarly, the canopy (tree top) temperature of trees is used to detect a problem which might be the result of water stress (underlying illness). This section aims to elaborate the state of the art methodologies to measure health of trees as well as to identify the attributes related to health that can be captured by these methodologies. We also highlight the existing research gaps and discuss the design choices for the development of our system.

**Materials and Methods.** All the works perused were found by performing a query search on Google Scholar and Web of Science. This was carried out between June 2021 and February 2022. The keywords were tree, tree health, vegetation health, vegetation index, mapping, monitoring, inspection, methods, techniques, sensing, sensor, mobile, thermal, infrared, thermography, drive-by, vehicle, vehicular sensor network. The keywords were combined with "AND" and "OR" operations as well as the use of brackets ("()"). Amongst the results generated, the abstracts of the works were read to select relevant papers which were further analysed. The reference list of relevant works was also taken into account and repeated references were also examined. Works older than the year 2000 were excluded from the study. After this process, we retained sixty-nine papers that were read in depth. Later, nine papers were found to be irrelevant in the specific context of the problem and were excluded. Each paper was

annotated with relevant interesting parts being digitally highlighted for later analysis.

## 2.2 Inspection Strategies for Greenery health

In this section, we first give an overview of current manual methods used to determine greenery health and then discuss some of the emerging sensing techniques for the measurement of tree health. These emerging approaches are largely based on imaging based methods. The aim in these methods is to generate and detect a representation of single or multiple attributes of a tree such as chlorophyll, canopy temperature, water transpiration etc. Herewith, each approach was further analysed to critically to determine the extent of tree health attributes measured by it.

Each of the methods reviewed is classified based on the primary and secondary sensors used and is summarized briefly in Table [2.2](#) and [2.3](#).

### 2.2.1 Manual Inspection Techniques

In general, arborists measure tree health firstly by visual inspection. However, visual inspection does not provide sufficient information about the presence of trees' damages and their extent. For instance, external symptoms of decay may be absent even in the presence of internal decay [\[67\]](#). This in turn may lead to delayed action when the tree has already reached its terminal stages and in order to avoid disease spread to other parts or trees [\[35\]](#). Hence, in order to verify the health and provide complete information about the extent of damage in trees, visual inspection is sometimes combined with other manual inspection methods like electrical resistance (involves calculating voltage difference by attaching electrodes and passing electric current through the trunk) or destructive instruments like increment borers (a wooden core from the tree is extracted and analysed).

These manual inspection methods can be broadly classified into invasive or non invasive categories and might have screening, diagnostic or evaluation purposes [\[67\]](#) [\[92\]](#). Invasive methods require drilling and penetration in the sapwood (living wood), thus creating an entry for pathogens or altering the structural integrity of. On the other hand, non-invasive methods do not need deep penetration into the sapwood. However, both invasive and non-invasive methods require intensive manual human labour with low spatial scalability as the analysis is done one tree at a time. Overall, the general rule is to start with the least damage causing method for screening the health of trees such as stress wave velocity [\[92\]](#) and then apply a more aggressive technique such as boroscope to get more information about the damage or decay [\[78\]](#).

Currently, various manual inspection methods exist and they are summarised based in Table [2.1](#). More details can be found in References [\[92\]](#), [\[67\]](#) and [\[51\]](#).

Method	Property Measured	Quality of Assessment	Cost*	Level of invasiveness
<b>Increment borer</b>	Visual inspection	Low	\$	Invasive
<b>Boroscope</b>	Internal visual inspection from inside	Moderate	\$\$	Invasive
<b>Resistograph</b>	Penetration resistance	Moderate	\$\$	Invasive
<b>Shigometer</b>	Single probe electrical resistivity	Moderate	\$	Invasive
<b>Fractometer</b>	Strength and stiffness	Moderate	\$	Invasive
<b>Stress wave velocity</b>	Single path acoustic wave velocity	Low	\$\$	Non-Invasive
<b>Electrical resistance</b>	Multi-probe electrical resistivity	Moderate	\$\$	Non-Invasive
<b>Stress wave tomography</b>	Multi-path acoustic wave velocity	High	\$\$\$	Non-Invasive
<b>Electromagnetic tomography</b>	Electromagnetic wave permittivity	High	\$\$\$	Non-Invasive
<b>Nuclear magnetic resonance (NMR)</b>	Magnetic property	High	\$\$\$\$\$	Non-Invasive
<b>Electronic nose</b>	Odour	Low	\$\$	Non-Invasive
<b>Gamma-ray computed Tomography</b>	Gamma ray transmissivity	High	\$\$\$\$\$	Non-Invasive

Table 2.1: Summary of manual techniques for detecting tree damage (\* refers to relative cost where \$ is lowest cost and \$\$\$\$\$ is highest )

## 2.2.2 Embedded Sensors

Embedded methods involves the use of static sensors which are directly attached to the trees or are placed near the trees to measure a property at regular intervals. Usually, they generate data at high temporal resolution with little or no human labour involved. The property measured by these methods can vary from detection of sudden vibrations to measure of water uptake and transpiration. Potamitis et al. [79] explores the use of an accelerometer-based sensor which is attached to the tree and constantly monitors for the presence of movement of insects or larvae in the internal part of the tree by transmitting the internal vibrations of the trunk. This approach is deemed to be cost-efficient as it reduces repeated visits by a human to examine the tree for presence of insects at regular intervals. Shabandri et al. [86] uses various sensors to detect a multitude of physical phenomenon such as sudden tree movements, availability of enough sunlight, enough soil moisture etc and sends the gathered data to a centralised server to shows real time alerts in an app. In [97] and [20], sensors are installed in the surrounding of the tree which are used as input features for developing a machine learning based algorithm to classify between healthy and unhealthy trees. In both these works, various features such as air temperature, humidity, soil humidity and soil acidity are fed specifically into a radial basis function based neural network in case of [97] and logistic regression in case of [20]. In the follow-up work [94], the same approach is extended by transferring the sensor data to a centralised server using NB-IoT. In a similar approach, Wu et al. [101] use the bifurcation of features into static and dynamic categories and feed them to a neural network to output a new index namely Urban Tree Health Index (UTHI). The static features are composed of parameters such as tree species, age, rooting area etc. whereas dynamic features consists of parameters such as air temperature, air humidity, soil moisture, tilt angle etc. The dynamic features are fed into a recurrent based neural network (RNN) layer and the static features are fused with the output of the RNN into fully connected neural network layers. In all these works [97] [20] [94] [101], the ground truth tree health data is obtained using manual inspection. In other works, Borges et al. [35] propose an Electrical Impedance Spectroscopy (EIS system) to assess the physiological stress in the trees. In this work, a pair of electrodes are placed in the trunk of the trees at diametric positions and an alternating current or a voltage with varying frequencies is passed and the resulting impedances are measured. It is found that there is a strong correlation with the ratio of impedances at predefined frequencies with the physiological stress in the trees. In order to measure tree growth, dendrometers can also be used [61] which can detect long term seasonal growth patterns, daily cycles of water uptake and shorter patterns like swelling after rainfall and subsequent drying [45, 83]. Further, these dendrometers can be combined with sap flow sensors to estimate both transpiration as well as hydration at tree level and detect early signs of tree mortality [80]. More details can also be found in [89] where a review of various static sensors used to measure tree health is also given. In general, all these methods require the installation of a sensor either on or near the trees. This leads to a high cost of deployment on a per tree basis with the benefits of reduced human labour thereafter. Further, with the development of new types of sensors, new tree properties can be measured in future.

Works	Primary Sensors Used	Auxillary Sensors	Inspection Level	Metric Computed	Deployment Type	Real World Evaluation
[38] and [39]	Thermal Camera	NA	Manual	Cool zones	Handheld	over 2000 trees selected over 10 year
[78]	Thermal Camera	NA	Manual	Temperature gradient, cool zones	Handheld	2 sample trees
[87]	Thermal Camera	NIR modified Camera, LiDAR	Manual	tree temperature rise and disease progress	UAV	40 sample trees
[71]	Thermal Camera	NA	Manual	tree temperature rise and disease progress	Handheld	15 sample trees
[36]	Thermal Camera	NA	Manual	Temperature gradient (Internal Defects)	Handheld	48 trees
[88]	Thermal Camera	NA	Manual	Temperature gradient (Internal Defects)	Handheld	20 sample trees
[47]	Thermal Camera	Visible Camera	Automatic	TWSI from thermal data, LAIe from RGB data	Car	172 trees on 2.52km trees
[74]	Thermal Camera	NA	Manual	CWSI	Handheld	5 sample trees
[60] and [32]	Thermal Camera	Visible Camera	Automatic	Canopy Temperature (Water Stress)	Handheld	44 images (trees not mentioned)
[64]	NDVI sensor, IR thermometer,	Thermal Camera, visible camera, multispectral camera	Manual	NDVI, canopy temperature correlation	Car	20 sample trees
[65]	Thermal Camera	NA	Automatic	Detection of structural defect in trunk	Handheld	8 sample trees
[33] and [93]	NDVI Sensors	LiDAR	Automatic	NDVI value	Mobile Robot	Lab environment
[46]	Airborne Hyperspectral camera	Airborne LiDAR	Automatic	NDVI value	Stationary	175 trees, setup ambiguous
[95]	RGB Camera	NA	Automatic	Estimated NDVI value after modifying camera	Handheld	controlled indoor setup
[27]	Custom Hyperspectral camera	NA	Automatic	NDVI, NDRE, LAI	Handheld	NA
[63]	Satellite based NDVI	Satellite based LST	Manual	Relation NDVI and LST	Satellite	Entire Mongolia: Coarse Resolution 8962 Trees
[103]	Multispectral Remote Sensing Data	NA	Manual	NDVI	NA	
	Satellite based NDVI	Satellite based LST	Manual	Relation NDVI and LST	Satellite	Entire Mongolia: Coarse Resolution 118 trees
[44]	Airborne Hyperspectral Camera	Airborne LiDAR	Automatic	LAI and NDVI	Airplane	

Table 2.2: A summary of emerging approaches analysed for detecting tree health

Works	Primary Sensors Used	Auxillary Sensors	Inspection Level	Metric Computed	Deployment Type	Real World Evaluation
[102]	LiDAR: Vehicle borne and airborne LiDAR	Infrared Camera	Automatic	Custom health classification in 4 levels	Helicopter and Car	220 trees
[84] [81]	Array of fixed light sensors GSV Images	NA Smartphone	Automatic Automatic	LAI LAI	Handheld Stationary, Handheld	3 sample trees Exact number not mentioned NA
[3]		NA	Automatic	Greenery View Index (No health assessment)	Mobile	
[96]	Baidu Street View Images	NA	Automatic	Panoramic Greenery View Index (No health assessment)	Mobile	Sanya, China
[70]	GSV Images	NA	Automatic	Shading effect of trees (No health assessment)	Mobile	Downtown area of Boston
[79]	Accelerometer based sensor detecting vibrations	NA	Manual	Presence of insects	Stationary, Attached to tree	11 sample trees
[35]	electrical impedance spectroscopy	NA	Manual	Presence of disease and hydration stress	Stationary, Attached to tree	24 sample trees
[86]	Temperature, moisture, carbon sensor	Air quality sensor, light-dependent resistor, vibration sensor	Manual	Real time readings for alerts	Stationary, Attached to tree	NA
[97] [20]	Sensors to measure surrounding ambient features	NA	Automatic	Binary health classification	Stationary, Attached around the tree	100 and 14 sample trees respectively
[20] [94]	Sensors to measure surrounding ambient features	NA	Automatic	Binary health classification	Stationary, Attached around the tree	14 and 15 sample trees respectively
[101]	Sensors to measure surrounding ambient features	Information like location, age, site condition etc.	Automatic	Five point health classification	Stationary, Attached around the tree	1418 trees

Table 2.3: Contd: A summary of emerging approaches analysed for detecting tree health

### 2.2.3 Imaging Based Methods

This section goes over imaging based techniques to measure tree health. We go over thermal imaging, hyper spectral imaging and LiDAR based methods. The works are classified into each section based on the primary sensor used. A subset of the works reviewed were found to use a combination of multiple approaches and they are classified as multi-sensory methods. The methods in this section allow analysis of multiple trees using a single sensing equipment. This makes them suitable for low-cost sensing at the expense of varying quality of sensing. In most of the works surveyed, it is observed that the data analysis is usually done manually and deterministically by a human.

#### Thermal Imaging

Thermal imaging based on IR (InfraRed) radiation emitted from biological materials is one of the most emerging technological approaches for tree health monitoring. It has been mainly used for either detecting cavities in tree trunks or measuring the water stress in trees.

Early works in this field [38, 39] relied on experts manually reviewing thermal images to identify conditions such as cavities and zones of decay in the bark or branches of a tree (wooden part). The areas with cavities appear cooler in thermal footprint than the rest of the bark surface. However, the extent as well the cause of the decay can hardly be estimated using thermal images alone. Moreover, to provide reliable results, the tree surface has to be shielded from direct sunlight since sun-exposed part may show higher temperature in the thermal images; hiding potential damages. Further, since water absorbs the IR radiation, the part which is examined cannot be wet. Moreover, the surface of the tree being examined has to be free of moss or other vegetation.

Leong et al. [67] argued that while thermal imaging is a good technique to screen trees for possible damage (binary classification), more advanced techniques are needed for an accurate health inspection. A recent review on thermal imaging techniques for tree health assessment by Vidal and Pitarma [92] found that there is no generalised temperature gradient pattern in terms of shape and size, along the bark to detect damage across various species of trees. Although the area near a deteriorated tissue might have a lower temperature, different conditions can generate different temperature gradients along the trunk of two different trees even if they belong to the same species. The main works analysed by the authors are also summarised concisely in Figure A.6 in Appendix (As shown from [92]).

Pitarma et al. [78] give an overview of the intricacies as well as the complexities of thermal imaging to detect the health of branches and the trunk taking into account the atmospheric temperature, the exposure to sun and the thermal environment of observations. The method used in this work again relies on the temperature differences between various parts of the tree. Specifically, it is illustrated that even if a part of the tree is exposed to sun, inferences can still be made about the health of a branch by comparing its temperature to atmospheric temperature. However, these inferences require expert knowledge.

On a separate note, while most of the previous works focus on identifying colder parts in the tree's trunk as an indicator of poor health, others [87][71] argue that the early onset of diseases in trees is in-turn associated with temper-

ature increase. For instance, Smigaj et al. [87] used a thermal camera mounted on a UAV to detect sub-degree temperature rise in the leaf and canopy temperature of trees affected by *red band needle blight*; a common disease. The infection level of each tree was manually assessed at ground level for validation. Similarly, Majdak et al. [71] found that infected trees have higher trunk temperature than un-infected trees and the differences are more noticeable on warm and sunny days than on cold and cloudy days. In a study that leveraged a drive-by sensing approach, Fuentes et al. [47] mounted both thermal and visible imaging cameras on top of moving vehicles to monitor the green infrastructure of Melbourne, Australia, at a tree-by-tree scale. The tree growth was estimated in terms of Leaf Area Index (LAI) and Tree Water Stress Index (TWSI). The LAI was measured by applying computer vision algorithms to the RGB images of tree canopy to segment leaves from the rest of the environment. The TWSI was measured using the canopy temperatures obtained by the thermal camera. The approach was also deemed privacy preserving as the cameras were mounted in an upward looking fashion facing towards the sky and the images are captured above the pedestrian level. Other work employing thermal cameras for water stress detection include [74], [88] and [64].

Finally, Kwok et al. [65] proposed a machine learning based method to automatically extract abnormal tree parts potentially containing cavities from thermal images installed at static positions. The method used k-mean clustering and Sobel gradient filter to identify potential cavities and the evaluation was performed on a dataset specifically created for this study and consisted of four tree species along with manual verification by an arborist. Similarly, Jiménez-Bello et al. [60] developed an automated method to calculate plant water stress by using a fusion of RGB (visible) and thermal imaging. It was found that the type of tree under examination had an important influence in determining the correlation with ground truth water stress results. Interestingly the process was fully automated using unsupervised classification in RGB images. In the follow study [32], further intricacies involved in usage of thermal imaging to detect water stress are enlightened. It was seen that the size of leaves may have a significant effect on the performance of thermal imaging. Further, the absolute values of CWSI and canopy temperature do not yield accurate correlation with actual water stress and thus, relative comparison with control trees is required. On a separate note, Burcham et al. [36] suggest that thermal imaging does not provide accurate results about the internal condition of trees containing decay and cavities and can only be used to detect superficial bark surface damage damage like detached bark or mechanical damage.

### Multispectral/Hyperspectral Imaging

In hyperspectral and multispectral imaging, various bands in the electromagnetic spectrum are captured which may or may not be in the visible band range. This captured data is then used to calculate various vegetation indexes, the most popular of which is NDVI. NDVI stands for Normalized Difference Vegetation Index. It relies on the ratio between visible red band and the near infrared (NIR) signals reflected by the vegetation and is widely used for vegetation health assessment. The NDVI index relies on the property of the chlorophyll present in the leaves which absorb red light and the cell structure of the leaves which reflects NIR. Higher NDVI values symbolise healthy photosynthetic capacity

while lower values symbolise poor health or presence of stress in trees or absence of vegetation. This technique, while already being measured using satellites and drones can also be measured terrestrially to give an estimate of tree health. This terrestrial measurement can be done using either active NDVI sensors which have their own energy source or passive NDVI sensors such as hyperspectral or multispectral cameras. Both of these types of sensors are readily available in the market from various manufacturers [100] [4]. Huang et al. [59] give a comprehensive review on the effectiveness of NDVI as a measure of the health of vegetation, but they also argue for the need of calibration of NDVI sensors. Further, Bahe et al. [31] suggest evidence that NDVI values can give accurate stress detection results when comparing data within one species and not across diverse species.

In recent works, Bietresato et al. [33] uses a mobile robot with NDVI and LiDAR sensors to detect health of five plants in a controlled environment. LiDAR is used to model the plant volume whereas NDVI sensors by OptRx [100] (popular in literature) are used to measure plant health and discriminate between vegetation, flower pots and background. The preceding work is continued in [93] and the vegetation thickness and NDVI index are combined to give a diagnostic matrix to give region wise vegetation index map. The approach of combining NDVI and LiDAR is also used in [46] where individual tree are identified from point clouds using both hyperspectral and depth information. Then, for health analysis, geometric parameters of the trees such as height, inclination and crown diameter are determined using LiDAR data and physiological parameters to be determined from hyper-spectral data are left to be explored in a further study. In [95], in order to reduce the cost associated with NDVI sensors, the NDVI index is calculated by modifying a regular RGB camera to remove the NIR rejection filters. In [27], the authors provide the design and implementation of a handheld generic sensor that can be configured to give various characteristics of a plant such as NDVI, LAI and nutrient requirements. The maximum distance of the target under measurement from the sensor in this study is 4m.

As discussed earlier, NDVI is also measured using satellites or other airborne approaches, e.g. employing aeroplanes or UAVs. In [63], the NDVI index calculated using satellite data is correlated with Land Surface Temperature (LST). Degerickx et al. [44] utilised LiDAR data for individual tree segmentation and hyperspectral imaging from an aeroplane to detect tree health. The authors in [103] used multispectral remote sensing data to measure health of trees at both pixel level and whole tree level. Lausch et al. [66] gave an overview of various remote sensing strategies available for sensing forest health and advocate the fusion of terrestrial data along with remote sensing based approach to achieve a better health assessment.

## LiDAR

LiDAR stands for Light Detection and Ranging. It is used to determine variable distances to an object under consideration or contour of surfaces by targeting it with a pulsed laser and measuring the time for the reflected light to return to the receiver. In [72], the authors used a mobile LiDAR system in a two-stage process. In the first stage, the species of the tree is identified by first determining if the tree is coniferous or deciduous and then examining the

branching behaviour. In the second stage, the health of the tree is calculated by analysing the point density of the tree which involves an estimation of the leaves surrounding each branch in the tree. Wu et al. [102] compare LiDAR based airborne laser scanning (ALS) and LiDAR ground based mobile laser scanning (MLS) for tree detection, tree species classification and vitality classification. It is found that while ALS in general, gives better performance to MLS, the combination of both ALS and MLS surpasses the performance achieved by only either of them. For tree health alone, more useful features were extracted from MLS compared to ALS. Degerickx et al. [44] derive Leaf Area Index (LAI) for tree health quantification using laser penetration metric of LiDAR. It is found that specifically for LAI, LiDAR performed better than hyperspectral data. Similarly, [84] also uses low cost LiDAR sensor (\$129) for ground based LAI measurement to quantify health of oil palm trees using intensity data.

### Multi-sensory Approaches

The works discussed in this section use a combination of previously discussed methods. This enables them to measure more than one parameter related to tree health than is possible with a single sensor or overcome the drawbacks of a single sensor.

Kim and Glenn [64] used a multi modal system comprising of thermal camera, IR thermometers, multi spectral camera and NDVI sensors to detect plant water stress. The use of multiple sensors is based on the hypothesis that water stress causes: 1. leaf temperature changes which are identified using thermal camera and IR thermometers 2. leaf color changes which are identified using multispectral camera and NDVI sensors. The sensors were mounted on a mobile vehicle and evaluation was performed in a controlled environment consisting of irrigated and non irrigated trees. It was found that the array IR thermometer can serve as low cost alternative to thermal camera, at the expense of bulkiness. Further, the NDVI sensor was found to perform better than the multispectral camera to distinguish between dry and irrigated trees. Qu [81] proposed two methods to detect tree health based on the leaf area index (LAI). The first approach uses fixed light sensors installed above and below the canopy to measure received and transmitted solar radiation and quantify the LAI. The second approach uses handheld device containing two smartphones which determine the LAI by applying object segmentation on the captured image. Both the methods were evaluated in a controlled environment and the measured LAI values were compared against a commercially available handheld device (LAI-2000) which is used to measure LAI.

Further, Google and Baidu panoramic street view images have also been used to quantify the extent of urban greenery. Li et al. [70] quantified the shading effect of urban trees using panoramic google street view (GSV) images. In [104], the authors used custom hemispherical view panoramic photos and employed image processing on the resulting photos to give a quantification of urban greenery at pedestrian level. The resulting quantification of urban greenery called panoramic green cover index (PGCI) was also correlated with land surface temperatures and NDVI of the urban areas. A similar approach is used in [3] and [96] where the authors used panoramic images from Google and Baidu respectively. However, these street view based approaches have only been used to quantify the spread of urban greenery rather than it's health.

## 2.2.4 Comparison

A comparison of the surveyed methods in terms of working mechanism, quality of tree health assessment, cost (considering evaluation of multiple trees on cite-wide scale) and level of invasiveness is shown in Table 2.4.

Approach	Working Mechanism	Quality of Assessment	Cost*	Level of invasiveness
<b>Manual Inspection Methods</b>	Depends on Method	Varying on method, generally high	\$\$\$\$\$	Depends on the method
<b>Static Embedded Sensors</b>	Depends on Method	Lower than manual methods	\$\$\$	Depends on method
<b>NDVI</b>	Properties of Chlorophyll	High quality Quantitative Value	\$\$	low
<b>Thermal Imaging</b>	Cavities, temperature gradient and water stress	Cavities, temp. gradient: Mostly Binary Classification, Water Stress: Quantitative Value	\$	low
<b>LiDAR</b>	Uses laser penetration metrics for parameters like LAI and leaf density	Low quality quantitative Value	\$ to \$\$\$	low
<b>Street view based methods</b>	Image processing to detect greenery	No health assesment, only extent of greenery	\$	low
<b>Remote Sensing</b>	Satellite Imagery	Top level view only	\$	low

Table 2.4: **A comparison of approaches to analyse tree health** (*\* refers to relative cost where \$ is the lowest cost and \$\$\$\$\$ is the highest cost for large-scale evaluation of multiple trees (> 50)*)

While manual methods like nuclear wave resonance or stress wave tomography have the high quality of assessment, they are time-consuming and infeasible in terms of cost for analysing each and every tree in the urban area. Further, the amount of labour involved is also huge. Terrestrial static methods again lead to high cost for analysing each and every tree due to sensors costs with the benefit of reduced human labour. The quality of analysis is also lower than manual methods. On the other hand, although infrared thermography has been used

to compute TWSI [47] [74], most of the perused literature [92] [39] uses it for detection of presence or absence of decay only (binary classification). NDVI sensors give more quantitative information about the health of trees, however, commercially available NDVI sensors (OptRx) are more expensive than thermal cameras. Sometimes, the NDVI is also susceptible to reach the maximum value due to the chlorophyll content representing peak greenness. Hence, any health issue may become difficult to detect until the problem progresses enough to reduce the NDVI value from the peak maximum value [48]. LiDAR based detection methods as discussed in [44] [84] are mostly used to calculate the Leaf Area index (LAI) and canopy density. However, the LAI value varies with respect to commercial handheld LAI measurement devices and canopy density varies for different species. Street view based methods [70] [104] [3] [96] while cost-effective are only able to quantify the extent of urban greenery rather than its health.

### 2.2.5 Research gaps and influence on design

From the literature surveyed above, it is clear that the manual inspection methods are infeasible for large-scale deployment due to both time and cost constraints. Further, some of them are invasive as discussed in Section 2.2.1. While, satellite and airborne based remote sensing approaches can cover large areas but with low resolution per pixel [47], they are only able to analyse the vegetation from an overhead view. Further, the presence of background materials such as grass or shrubs within the same pixel affects the reliability of observations [97]. In contrast, ground based sensing can look at vegetation elements in a more holistic manner. Based on the emerging technologies perused, the use of thermal imaging to analyse attributes of tree such as water stress index and cavities in trees seemed promising and suitable for our problem statement. Also recall from Section 2.2.3, hyperpectral/multispectral imaging can also generate a number of vegetation indexes, the most popular of which is NDVI. Contradictory studies were seen on the usage of LiDAR with some works such as [84] claiming no increase in health detection performance with its addition.

It is seen that most of the works that measured tree health from ground level except two [47] [64] used manual judgment and processing by humans on the data collected and employed handheld data collection. This makes them inefficient to be deployed on large scale with data collection at regular time intervals. For the two comparable related works [47] [64], it is found that in [47], the authors only calculate the water stress index on a moving vehicle without validating the same with any ground-truth dataset. Further, the experiment was controlled by mounting the camera in an upward facing fashion to only inspect tall trees. Whereas in [64], while the authors did use a ground truth dataset for validation, the processing of both thermal and hyperspectral images was performed manually using human intervention.

In essence, although there are ample works which have tried to measure tree health from ground level (terrestrial), measuring this health autonomously without human intervention and scaling it to be suitable for city-wide evaluations is an open field of research. Given the current advances in the field of computer vision and availability of low cost multispectral & thermal cameras along with the proliferation of microcontrollers and single-board computers, we employed a combination of these technologies in the developed system. We em-

ployed traditional image processing techniques along with a custom computer vision model to automatically generate the tree health attributes which was later validated with a ground truth dataset.

## 2.3 Metrics (NDVI and CTD)

In this section, we will briefly discuss the two tree health metrics namely NDVI (Normalised Difference Vegetation Index) and CTD (Canopy Temperature Difference) used in this thesis. Each object in the physical world emits or reflects electromagnetic (EM) radiation. The electromagnetic spectrum is the range of frequencies of EM radiation. A visible imaging, thermal imaging or multispectral imaging sensor is sensitive only to particular wavelengths of these EM radiations. These captured wavelengths are then stacked over to form different channels of a pixel and the combination of these pixels form an image. For the context of our work, the information captured from different spectral bands is used to measure two different attributes of tree health which are derived from the specific technology (thermal and multispectral imaging) incorporated in the system.

1. **NDVI:** The multispectral sensor used in our system captures spectral reflectance (radiation) of trees at different wavelengths or across different bands, the details of which are elaborated in Chapter 3. Interestingly, combining these radiations or reflectances from different spectral bands mathematically leads to a Vegetation Index (VI). In literature, based on the application, there exists various vegetation indices such as Green Normalized Difference Vegetation Index (GNDVI) or Normalized Difference Red Edge (NDRE) [26]. One of the most popular vegetation indexes used to measure tree health is called **Normalised Difference Vegetation Index (NDVI)**. Recall from Section 2.2.3, the properties of chlorophyll and leaf cell structure in the trees on which NDVI is physically based on.

Mathematically, NDVI is calculated as:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (2.1)$$

where *NIR* and *Red* are near-infrared reflectance and visible red reflectance from the leaves of the tree respectively. As per (2.1), NDVI lies in the range [-1,1] and higher values indicate higher photosynthetic activity (or health of trees in our case). In case of calibrated NDVI values, 0.3 is used as the cutoff to distinguish between vegetation and non-vegetation elements.

2. **CTD:** The thermal imaging sensor used in our system can be used to measure the tree canopy temperature. Similar to the various vegetation indices, various thermal stress indices exist in literature [88]. Canopy Temperature Difference (CTD) is one of the most simplest of these indices. CTD is based on the property of the leaf stomata closing under water stress, thus decreasing the rate of transpiration and thereby, decreasing the rate of cooling. This leads to an increase in canopy temperature and a reduction of photosynthetic activity. This index has been found to have a

high correlation with water stress in trees [88, 32]. Mathematically, CTD is calculated as:

$$CTD = T_{canopy} - T_{air}, \quad (2.2)$$

where  $T_{canopy}$  and  $T_{air}$  are canopy temperature and air temperature respectively. Well irrigated trees without water stress will have lower values of CTD and trees under water stress have higher values of CTD (in hot weather conditions).

## 2.4 Image Segmentation

Image segmentation is the task of segmenting objects of interest (parts of tree canopy in our case) in an image by combining parts of objects together that belong to the same label (or class). When performed using deep learning, it can also be considered as a pixel-level classification problem. However, image segmentation can also be performed using traditional deterministic image processing algorithms like thresholding or watersheds. As shown in Section 3.3, we used a combination of both traditional image processing and deep-learning based techniques.

Image segmentation can be of two types depending on the complexity:

- **Semantic Segmentation (Less complex):** In this type of image segmentation, each pixel is classified into different objects without any differentiation between multiple instances of the same object. Hence, similar objects are treated as a single label from the image level.
- **Instance Segmentation (More complex):** Here, each pixel is classified into different objects while also recognising the different object instances. Hence, each object instance is treated as a separate label from the image level.

A visual demonstration of these two types of segmentation is shown in Figure 2.1

### 2.4.1 IoU

The performance of image segmentation models is measured by a metric called Intersection over Union (IoU).

Mathematically, IoU is calculated as:

$$IoU = \frac{\text{Area of Overlap between } A \text{ and } B (A \cap B)}{\text{Area of Union between } A \text{ and } B (A \cup B)} \quad (2.3)$$

where  $A$  and  $B$  are predicted mask and ground truth mask (segmentation) for an object instance in the image. IoU lies in the range [0,1] with 1 indicating perfect overlapping segmentations (masks) and 0 indicating no overlap. Higher the IoU, better is the model performance.

Usually, to measure the performance of image segmentation models, the average precision (AP) is calculated at various IoU. This average precision is actually calculated by computing the area under the precision-recall curve for each object label in the image and then taking the mean for each label across all the images in the dataset [9].

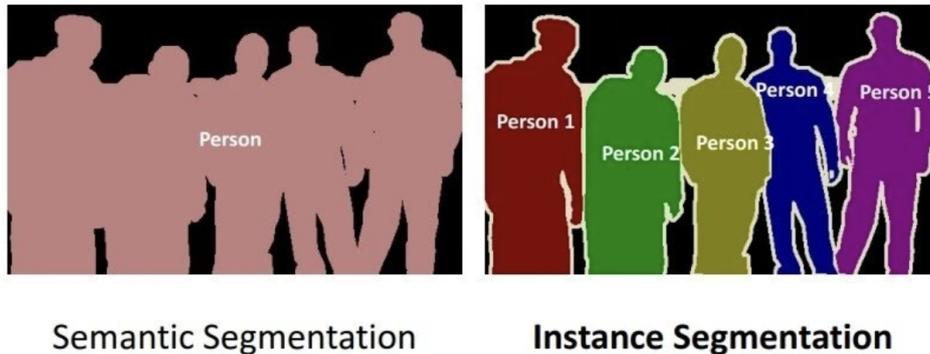


Figure 2.1: **Difference between Semantic and Instance Segmentation:** In the left image, all instances of humans (object) are classified with single label namely person. In the right image, each instance of object (human) is classified with separate labels

### 2.4.2 Mask R-CNN

Mask R-CNN is a state-of-the-art deep learning model to perform instance segmentation. It is an extension of Faster R-CNN with the addition of another branch for predicting the object mask (region of interest) in parallel with the existing branch for bounding box recognition in Faster R-CNN [55]. Faster R-CNN is an enhancement over Fast R-CNN which builds on top of R-CNN. A detailed explanation of the working of Mask R-CNN is outside the scope of this work because it involves the understanding of R-CNN [50], Fast R-CNN [49] and Faster R-CNN [82] in that order. Further, this work does not propose any modifications to the Mask R-CNN architecture, rather uses it to extract useful information from low resolution thermal imaging sensor and identify tree canopies in the image.

Mask R-CNN is a computationally heavy model compared to object detection and semantic segmentation models both in terms of training and inference. In our system, due to data privacy requirements and the need for detection of multiple trees in an image, the custom Mask R-CNN model was optimised to run on the edge device.

### 2.4.3 Transfer Learning

Transfer learning is a technique in deep learning where a model pre-trained on a large dataset related to solving one task is used as the starting point for solving another related task. The main advantage of using transfer learning to solve the another related task is, it allows to train models with good performance even with very small datasets as well as saves time and resources compared to training a model from scratch. For instance, in the field of computer vision, there are large scale open-source datasets like CIFAR-10, COCO and ImageNet containing thousands to millions of images. Employing transfer learning for image segmentation means that we can use a model pre-trained for image segmentation on large scale dataset like COCO (which does not have object labels of

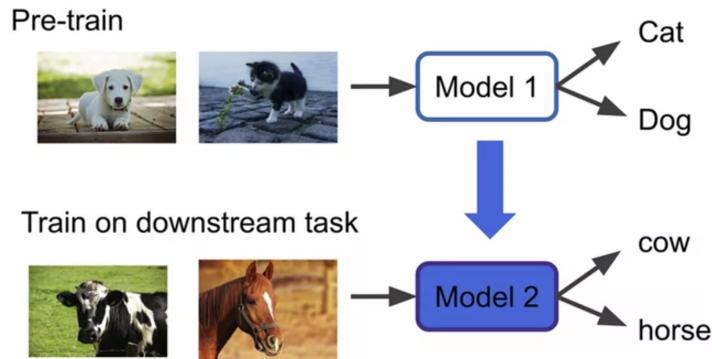


Figure 2.2: A model (Model 1) trained to classify cats and dogs can be reused and used as starting point to develop a model (Model 2) to classify cows and horses. This is because Model 1 already contains layers to detect low level features (feature extractors) like eyes, nose, ears etc. and these low level feature extractors can be reused in Model 2 without any change in weights (As shown in [28]).

interest) and use the same model weights & layers as a starting point to retrain the model on small dataset (which contains object labels of interest) collected by us. While transfer learning is domain agnostic, it is specially popular in the field of computer vision.

For the context of this work, as discussed more in Section 3.4, transfer learning was used for the task of instance segmentation of tree canopies using a Mask R-CNN model pretrained on COCO dataset [8].

## 2.5 Takeaways

In this chapter, we surveyed the existing works related to monitoring the health of trees. We found that despite ample research on this topic, there has not been prior research on the measuring tree health health autonomously without human intervention and scaling it to be suitable for deployment on large scale. Based on the existing research, we also identified the technologies that were suitable to be incorporated into our system. Further, we briefly discussed the preliminaries about image segmentation and the metrics that will be used in the remainder of this thesis.

## Chapter 3

# Materials and Methodology

In this chapter, we elaborate on the hardware and software components used in our system along with their implementation in Sections 3.1 and 3.2. A visual image processing pipeline is also demonstrated in Section 3.3. The implementation of custom Mask R-CNN is elaborated in detail in Section 3.4. Finally, we elaborate on the ground truth dataset used for our evaluation and the data collection experiments performed in Cambridge in Sections 3.5 and 3.6.

### 3.1 System Design

As described in the Section 2.2.5, we incorporated thermal and multispectral imaging sensors in our system. We omitted LiDAR in the current system because of unavailability of low cost LiDAR module with high distance range as well as the conflicting results about the usefulness of LiDAR in the works discussed in the previous Chapter 2. In this section, we will elaborate on the hardware components and the software choices employed to develop the system along with reasoning for choosing them over other available alternatives.

#### 3.1.1 Hardware

The following indicates all the hardware components employed to develop the system:

1. **FLIR Lepton 3.5 (Thermal Imaging)**: This is a thermal sensor with an uncooled microbolometer and captures images with a resolution of 160x120 and FOV (field of view) of 57°. It captures long wave infrared radiation in the spectral range of 8  $\mu\text{m}$  to 14  $\mu\text{m}$  and can measure temperatures in the range of -10 to 140°C (High Gain mode). The sensor supports radiometric calibration which allows it to measure the true temperature of the surface being observed if certain conditions are met as per FLIR and elaborated in [12]. Nevertheless, from previous works [88] utilising these types of sensors as well as during our experiments, such micro thermal sensors employing uncooled bolometers are susceptible to non uniform noise and temperature instability when powered on before giving stable outputs. Hence, it is imperative to wait for some minutes before collecting data with this sensor. An alternative to this module was

a thermal sensor array such as [16] which has better noise performance. However, FLIR Lepton was chosen as it has a higher thermal resolution than a thermal sensor array and lower costs (\$150) compared to other thermal camera modules. More details about this module can be found at [15].

2. **MAPIR Survey 3W RGN (Multispectral Imaging) with GPS**

**Receiver:** This is a compact low cost (\$400) multi-spectral imaging system which captures RGN (Red, Green and Near Infrared) images with sensitivity to wavelengths of Red 660nm, Green 550nm and near infrared 850nm. It has a resolution of 12 Megapixel (4000x3000) and FOV (field of view) of 87°. It was ideally developed for mounting on drones and UAVs and provides a PWM (Pulse width modulation) interface through the HDMI port to issue basic control commands such as taking an image or mounting/unmounting the memory card. Further, it supports an external GPS module manufactured by MAPIR to geotag the images taken. Although, a multispectral module developed for connecting to an embedded system would have been preferred instead of a full fledged imaging system like MAPIR Survey, at the time of this work, we were unable to find any multispectral modules available in the market. In fact, MAPIR has an upcoming multispectral module array [14] which is scheduled to start shipping in Q3 2022. More details about this component can be found at [22].

3. **OpenMV cam H7 with Lepton adapter module:**

OpenMV cam H7 is a small low-power microcontroller board with support for running computer vision algorithms on the edge and interfaces easily with FLIR Lepton 3.5 using the lepton adapter module. It uses the popular STM32 series [21] of microprocessors and supports running Tensorflow Lite [25] models. Further, it provides a Remote Procedure Call (RPC) interface to send/receive control commands and data from another microcontroller or single board computer. It was chosen considering the imperative requirement of a module to integrate the FLIR Lepton 3.5 into the system and its low cost (\$65+\$15) along with the tensorflow lite support. Further, this microcontroller was also successfully used earlier to integrate FLIR Lepton into the CityScanner project. More details about this module can be found at [17].

4. **Raspberry Pi 4:**

This is a single board computer that serves as the brain of our system. It was chosen because of its low cost (\$35), flexibility to run various frameworks in multiple languages such as OpenCV and Tensorflow in Python or C++. Further, it is very flexible in terms of hardware expansions with various modules available to add extra capabilities to it. For instance, a dedicated tensor processing unit (TPU) to speed up the running of the deep learning models [10]. More details about this component can be found at [77].

All of these components are shown in Figure 3.1. In essence, the total hardware related cost of our system is  $\$150+400+65+15+35 = \$665$

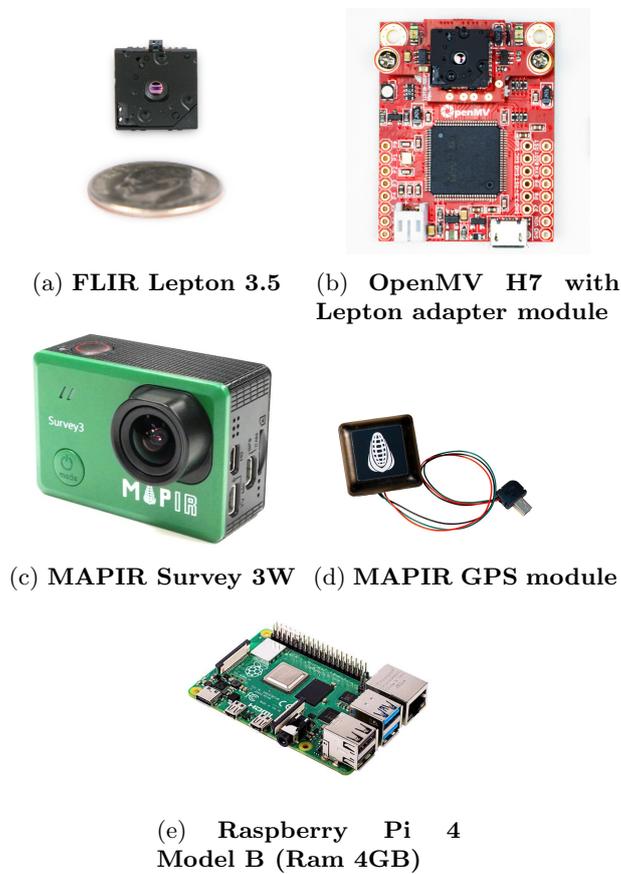


Figure 3.1: **Hardware Components used in the system**

### 3.1.2 Software

The following indicates all the software based choices that were used to develop the system:

1. Language: Python 3 was selected as the core language to develop the system. The idea behind choosing Python instead of a low level language like C or C++ was the flexibility offered by Python in terms of support for various deep learning and image processing frameworks such as OpenCV, Tensorflow and PyTorch along with direct integration with micropython based microcontrollers such as OpenMV cam H7.
2. Software Frameworks: Tensorflow [24] was used as the deep learning framework of choice instead of pytorch. The main reasoning behind this is the support for easy conversion of Tensorflow models into Tensorflow Lite [25] by quantizing them and enabling them to run on the edge. In addition, OpenCV was used as the image processing library.
3. Communication with Hardware: The communication between Raspberry Pi and FLIR Lepton 3.5 was implemented using the RPC interface over

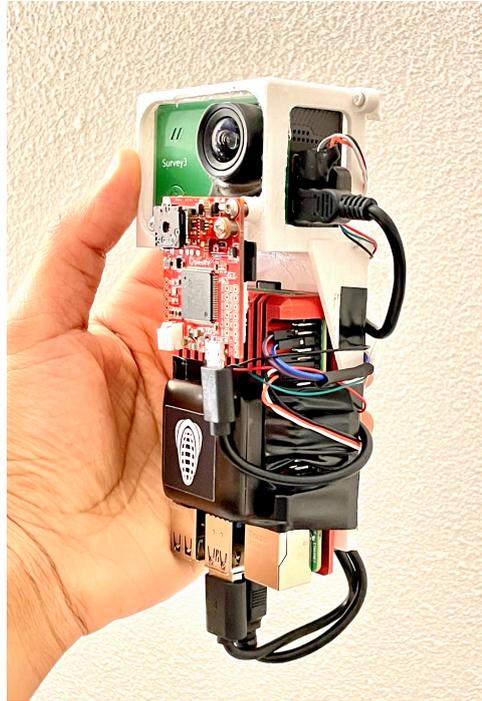


Figure 3.2: All hardware components encased with the 3D printed case

USB with OpenMV Cam H7 as the mediator. The MAPIR Survey 3W only has support for PWM signals for control commands. Hence, scripts were coded to pulse modulate GPIO pins on Raspberry Pi which were in turn connected to the micro HDMI port of MAPIR Survey 3W. These scripts mounted and unmounted the memory card installed in MAPIR Survey 3W with the raspberry pi and triggered image capture.

All the hardware components were encased in a 3D printed case as shown in Figure 3.2. This made it easy to carry during data collection. Further, the case was designed such that it is suitable to be attachable to moving vehicles using magnets currently used in the CityScanner project as shown in Figure 3.3. *(This casing was designed by a member of the CityScanner team in consultation with the main author of this work.)*

## 3.2 System Architecture

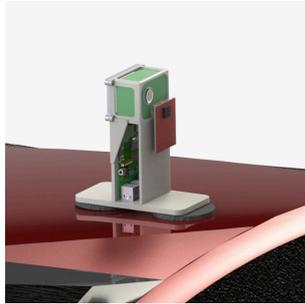
The block diagram of the entire system along with all the major modules is shown in Figure 3.4. Recall from previous Section 3.1.1 that a raspberry pi is employed as the central brain of our system and all the other modules are interfaced with it. Herewith, we will give a brief overview of all the other major modules as shown in the architecture diagram. We begin by elaborating on the hardware modules, then move towards embedded control module and finally towards image processing and calculation modules. The sequence of steps applied during image processing are also visually shown in Figure 3.5.



(a) Concept casing for the system with magnets



(b) The system attached to the top of a car



(c) A closeup view of the system attached to the roof of a car

Figure 3.3: Concept casing with magnets attachable to moving vehicles (Developed by a member of CityScanner [6] Team)

1. **FLIR Lepton 3.5 and OpenMV Cam H7:** The thermal imaging sensor is attached to openMV cam H7 using a FLIR Lepton adapter module. This module communicates with Raspberry Pi via remote procedure call over USB. The custom python code running on OpenMV is configured to measure temperature in range of  $-10^{\circ}$  to  $40^{\circ}$  C and returns grayscale image data normalised to this temperature range. This means that a measured temperature of  $-10^{\circ}$  C or lower will be scaled to 0 (black) as pixel value and a measured temperature of  $40^{\circ}$  C or higher will be scaled to 255 (white) as pixel value. This temperature range was chosen by considering the temperature of tree canopies observed during the data collection experiments (refer Section 3.6). A narrow temperature range is usually preferred to decrease the effect of non-uniform noise across the sensor. The micropython code running on OpenMV also handles the serving of the callbacks initiated by the **Control** module.
2. **MAPIR Survey 3W with GPS Receiver:** The multispectral imaging sensor is attached to the raspberry pi over USB and the mounting/unmounting of the memory card along with the capturing of images is handled using PWM signals over the micro HDMI port of MAPIR Survey. Further, a GPS receiver is attached to MAPIR Survey in order to geo-tag

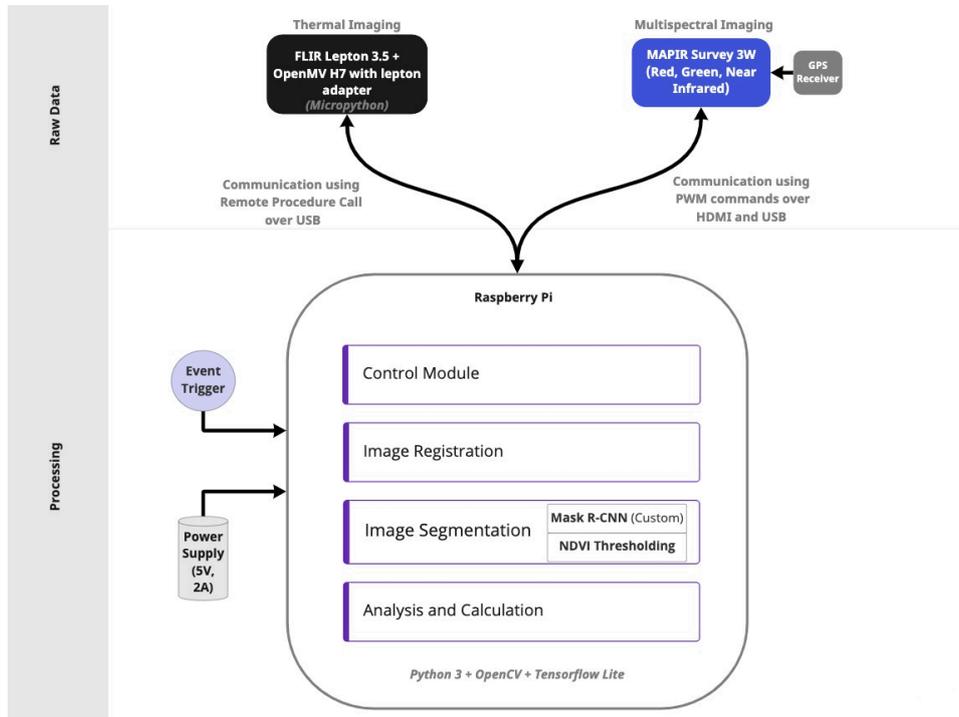


Figure 3.4: **Architecture Diagram of the system**

all the images captured.

3. **Event Trigger:** This indicates any event which signals the beginning of processing on raspberry pi from capturing the images to generating NDVI and CTD values. At the current state of the prototype, a press of a push button connected to GPIO 2 of raspberry pi was used as an event trigger for the data collection experiments. However, this trigger can also be an event signalling the co-location of the system with particular GPS coordinates. For instance, when the the system is deployed on moving vehicles, the co-location of the system with GPS coordinates fetched from a tree inventory database will be used to trigger this processing.
4. **Control Module:** This software module handles the event trigger, signals the sensors to capture the images and transfers the captured images to the raspberry pi for further analysis. For the OpenMV Cam, this involves initiation of callbacks requesting the transfer of current image frame from the thermal imaging sensor. For the multispectral imaging sensor, this involves generating PWM signals to capture an image, mounting the memory card installed in the multispectral imaging sensor with the raspberry pi, transferring the captured image to the raspberry pi and finally, unmounting the memory card from the raspberry pi.
5. **Image Registration Module:** Image registration is the task of matching or aligning images taken by two different sensors or different viewpoints into a single coordinate system for further analysis [62]. It involves map-

ping points from one image to corresponding points in another image. In our system, both the multispectral and thermal imaging sensors have different FOV (field of view) and are not co-aligned. This software module registers the images taken by the thermal and multispectral imaging sensors. In order to align both multispectral (RGN) and thermal images, this module handles the linear translation of RGN images in both horizontal and vertical directions. Further, to compensate for wider FOV of multispectral sensor, this module also handles the zooming in of the RGN images. For our prototype, in case of RGN images, the values of translation in X (horizontal) and Y (vertical) directions were found to be +50 and +150 pixels respectively and zoom scale was found to be 0.57 (1 indicates no magnification and 0 indicates pixel level magnification) to perfectly overlay thermal and RGN images. These configuration parameters were manually found by taking multiple RGN and thermal images and overlaying them. An instance of inputs and outputs utilizing this module are shown in Figure 3.5. Further, in order to scale this module for large scale production of the system in the future, we also tried automatic image registration using three traditional image registration algorithms namely SIFT, SURF and ORB [5]. However, none of the algorithms were able to find useful keypoints or features in the thermal images. We believe that this was due to the low resolution of thermal images.

6. **Image Segmentation Module: This is the most computationally intensive software module of our system.** Recall that the aim of our system is to calculate NDVI and CTD values of trees in the images. However, these values should only be calculated for the leaves in the tree canopy and not the wooden parts which include branches and trunk. This is solved using a fusion of custom Mask R-CNN model and pixel-wise NDVI analysis. Hence, given a RGN image, this task can be broken into two sub-problems as follows:

- **Detecting the canopy part of the trees even in cases where the image contains multiple trees:** This is solved using a custom Mask R-CNN model. The Mask R-CNN model is discussed in more detail in Section 3.4 and it outputs the instances of the tree canopies in the image by generating a mask (segmentation) over their canopies as shown in Figure 3.5.
- **Once the canopy of the tree is detected, the segmentation of only the the leaves of the tree without the wooden parts and sky:** Recall from Section 2.3 that non-vegetation elements such as trunks, branches and sky have very low NDVI values compared to vegetation elements which have significantly higher NDVI. Thus, we employ a thresholding method which first calculates the individual NDVI of each pixel in the segmentation mask given by Mask R-CNN and then, eliminates pixels with NDVI values below a certain threshold as shown in Figure 3.5. The calculation of NDVI for each pixel is simply computed by deriving the raw values in the red and near infrared channels of the pixel and plugging them in (2.1). In order to eliminate noise along the edges of tree canopy, median filtering is also employed.

The end result using the above two-stage approach gives segmentation of leaves present on the canopy of a tree while eliminating the sky, wooden branches and trunk of the tree. Since our multispectral imaging sensor is uncalibrated, the raw NDVI values generated by it are relative. Hence, a value of 0.02 was used threshold cutoff value to eliminate non-vegetation elements in the image. This value was derived using the analysis of the images captured during data collection and comparing the output quality with different threshold values. An instance of inputs and outputs using this module is also shown in Figure 3.5

7. **Analysis and Calculation Module:** With the availability of segmentation i.e. tree canopy mask consisting of only leaves for RGN images, the same mask can also be used for thermal images due to image registration. This module handles the calculation of final NDVI and CTD for the tree.

The CTD value is computed by calculating the raw temperature value for each pixel by converting its color intensity value in the grayscale thermal image (using (3.1)), computing the mean temperature over all pixels in the canopy and subtracting the ambient air temperature from the mean canopy temperature (refer (2.2)).

The temperature of each pixel is calculated as:

$$T_{pixel} = \frac{P_{value}}{255} * (T_{max} - T_{min}) + T_{min} \quad (3.1)$$

where  $P_{value}$  is the pixel value in normalised thermal image,  $T_{min}$  and  $T_{max}$  are configured temperature range for the FLIR Lepton 3.5 respectively ( $-10^\circ$  and  $40^\circ$  C in our case).

Then, as per (2.2), CTD is calculated as:

$$CTD = \overline{T_{pixel}} - T_{air} \quad (3.2)$$

where  $\overline{T_{pixel}}$  is the average canopy temperature for all segmented pixels in the image and  $T_{air}$  is the air temperature respectively.

To calculate the mean NDVI, each RGN pixel in the image is split into its 3 constituting channels (red, green and near infrared). The raw NDVI value for each pixel is calculated from red and near infrared channels as per (2.1). To compensate for the aperture adjustment, focal adjustment and other mechanical adjustments performed by the multispectral imaging sensor, this raw NDVI is normalised by dividing with a correction factor similar to the dynamic range of a camera [11].

Mathematically, our corrected NDVI is calculated as:

$$NDVI_{corrected} = \frac{NDVI_{raw}}{|NDVI_{max}|} * |NDVI_{min}| \quad (3.3)$$

where  $NDVI_{raw}$  is the raw NDVI of a pixel,  $NDVI_{max}$  and  $NDVI_{min}$  are maximum and minimum NDVI values among all pixels in the segmented image.

Finally, the corrected NDVI for the entire canopy is computed by taking the mean over all the corrected NDVI for all pixels in the segmented

image. While this approach of adding a correction factor does not calibrate the corrected NDVI to the absolute ground truth NDVI, it can be used for relative comparison between the calculated values (thus meeting our system requirements).

The final calculated CTD and NDVI values along with other intermediate calculations such as raw NDVI, Canopy temperature and related indexes like GNDVI (Green NDVI: same as NDVI but with replacement of red band by green in (2.1)), SR (Simple Ratio: NIR/Red), PercentHealthy0.1 (% canopy with raw NDVI > 0.1) are stored on the raspberry pi in a .csv file.

### 3.3 Visual Image Processing Pipeline

A visualization showcasing the processing of both thermal and RGN images in each module is shown in Figure 3.5. As discussed above, the images from the thermal and multispectral imaging sensors are fetched using the control module based on the event trigger which begins this processing of images in a sequential manner.

### 3.4 Development of Custom Mask R-CNN

As discussed briefly in the previous section, after the image registration of the thermal and RGN images, we need to segment the canopy of a tree even if the image contains multiple trees from the RGN image. For our system to operate completely autonomously, the images captured on the basis of an event trigger will be unsupervised and may contain other objects in the image such as cars, buildings, grass, snow in addition to multiple trees. Hence, it becomes imperative to individually identify all the tree canopies in an image and feed them to the calculation and analysis module. The custom mask R-CNN aims to solve this problem by providing instance segmentation of the tree canopies in the image.

#### 3.4.1 Motivation for a custom instance segmentation model

In our knowledge, there is no pre-existing model available for **instance segmentation** of tree canopies or even trees for the standard RGB (red, green, blue) images. Our problem is further complicated as our input is RGN images from the multispectral imaging sensor instead of standard RGB images. For instance, pre-trained models like Deeplabv3 [40] which are able to perform **semantic segmentation** of trees and vegetation on standard RGB images perform poorly on RGN images as shown in Figure 3.6. These models are trained either using CityScapes [7] or ADE20k [105] datasets. Further, these models only allow semantic segmentation of trees so detecting the health of individual tree canopies in the image is not feasible using them. Hence, it becomes imperative to develop a custom model from scratch which works well on RGN images and can detect individual instances of tree canopies.

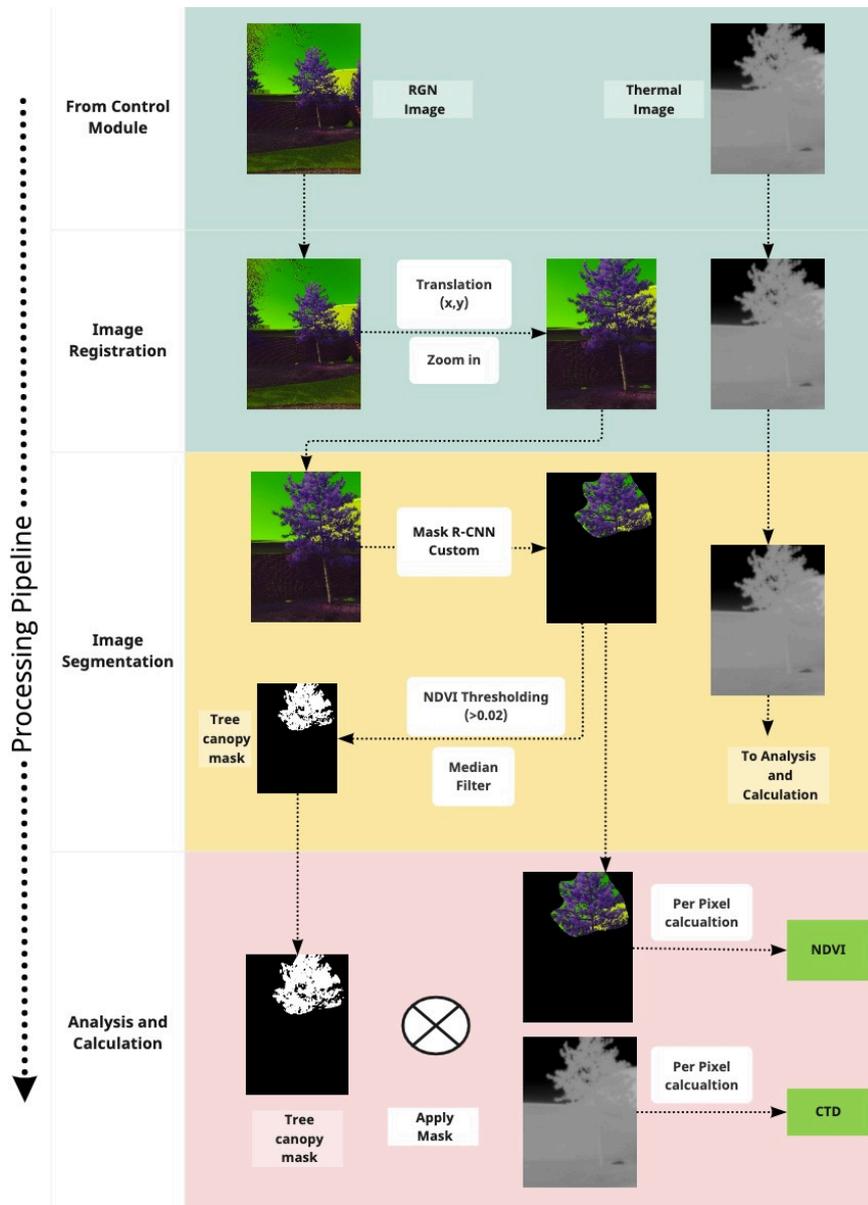
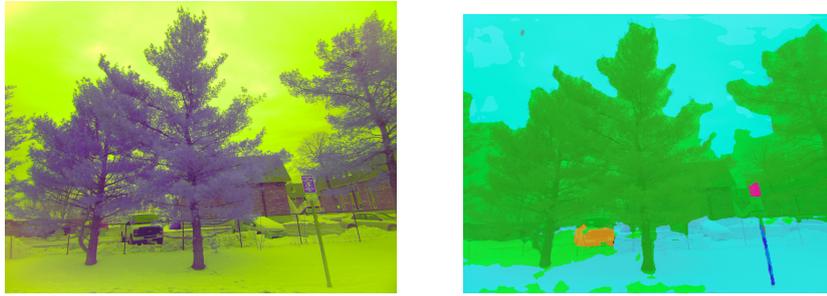


Figure 3.5: A visualisation of how the images are processed at each step

### 3.4.2 Dataset Creation

Any deep learning model requires training data in order to optimise the weights and activations of the layers. However, there does not exist any dataset with labels for instances of trees or tree canopies in RGN images. Hence, we manually created the dataset using the RGN images collected during the data collection experiments (See Section 3.6). Here, each tree canopy in the image was manually annotated using the popular image annotation tool called LabelMe [99]. During



(a) Input RGN Image captured from MAPIR Survey 3W (b) Segmentation output from DeepLabv3 (semantic segmentation model) trained on ADE20k

Figure 3.6: Performance of pre-trained DeepLabv3 model (Semantic Segmentation). Notice how the model overestimates the segmentation of trees by including buildings, sky and snow as vegetation

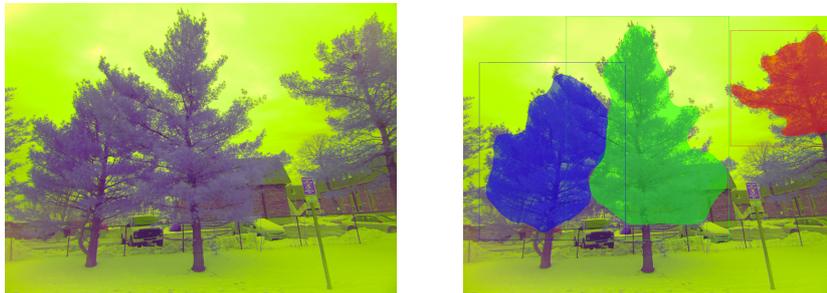
annotation, only tree canopies which were completely present in the image were labelled. After this process, our dataset consisted of 51 annotated RGN images with 2 classes namely tree canopies and background.

### 3.4.3 Training

Our dataset consists of a relatively small number of images to train a deep learning model like Mask R-CNN from scratch. Recall from Section 2.4.3 transfer learning combined with data augmentation can be theoretically employed in order to develop a custom model by re-training an existing model pre-trained on a different dataset. For our purposes, we used a Mask R-CNN model pre-trained [29] on COCO [8] (a dataset with 330K images) with ResNet101 as the backbone. Tensorflow was used as the framework of choice. We retrained only the head layers (the top layers without the backbone) on our dataset. We also generated synthetic data by augmenting the original dataset with flips in horizontal and vertical directions and further applying Gaussian blur. This increased our training dataset size by 50% and also acts as a regularizer. The configuration parameters used for training are shown in Section A.1. The visual output results from our model are shown in Figure 3.7.

### 3.4.4 Model Quantization

Mask R-CNN is a relatively heavy model from both training and inference point of view. While this model can be deployed on the cloud servers like Amazon Web Services or Google cloud platform during inference for fast processing, raw image data would need to be sent to the server. However, our system requirements includes data privacy constraints. Thus, the developed Mask R-CNN was optimized to run on the edge at the cost of possible minute performance reductions. To achieve this, the model built on Tensorflow was converted to Tensorflow-lite with dynamic range quantization [23]. Dynamic range quantization means that only the weights of the layers consisting of 32 bit Floats are stored as 8 bit ints in the converted model whereas the activations of the layers



(a) Input RGN Image captured from MAPIR Survey 3W (b) Segmentation output from our Custom Mask R-CNN (instance segmentation model) trained using transfer learning

Figure 3.7: Performance of our Custom Mask R-CNN. Notice how the model detects each instance of the tree canopy in the image and considers all the other objects as background

are quantized during runtime. Our custom Mask R-CNN built over Tensorflow took around 15 seconds per inference of an image on a raspberry pi 4 while the tensorflow-lite model reduced the inference time to 7 seconds with one-fourth of the CPU usage as the original tensorflow model.

### 3.5 (Ground truth) Tree Health Dataset

Municipalities in cities obtain ground truth tree health data through city wide surveys every few years. For instance, in the city of Cambridge, USA, this survey is performed every 5 years whereas in the city of Delft, The Netherlands, this survey is performed every 2 years. To help with the evaluation of our system, we obtained ground truth tree health dataset for the city of Cambridge, USA through Cambridge Urban Forest Master Plan. This dataset is from the summer of 2018 and was created through a combination of manual in-person arborists, satellite based remote sensing and aerial LiDAR incorporation [1]. The dataset classifies the health conditions of trees in three categories namely good, poor and fair. It contains information about 47,063 trees out of which 35,821 are in good health, 5,176 are in fair health and 6,066 are in poor health. Hence, most of the trees (> 75%) are rated as having good health condition. In addition to this, the dataset contains information about the tree species, the satellite based NDVI, the latitude and the longitude, whether located on a street or not, the shape length and shape area of the canopy, flood tolerance and drought tolerance. This dataset was provided as Shapefiles (.shp, a dataformat used by Geographical Information Systems (GIS)) and was loaded to the online platform CARTO [37] (a GIS and spatial analysis tool) as shown in Figure 3.8.

While this dataset is from 2018 and the tree health conditions are subject to variability in four years, this is the most accessible ground truth dataset that was obtained during this work. Further, the staleness of data also necessitates the need for advancements in this field of tree health monitoring to ensure faster remedial actions to save trees in poor health.

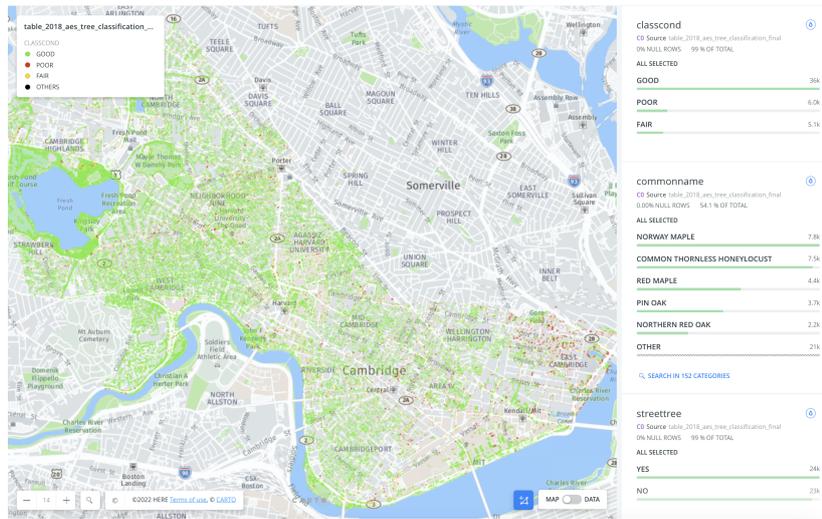


Figure 3.8: Tree health dataset after importing to the CARTO platform

### 3.6 Data Collection Experiments

We collected RGN and thermal images through our developed system on three separate days in Cambridge, USA during the month of February, 2022. A push button was used as the event trigger for the system. Hence, we used the developed system in a citizen science paradigms with the 3D printed casing. In total, we collected data for 49 trees spread over two species namely Red pine and Eastern White Pine trees. The multispectral imaging sensor was configured with shutter speed of 1/60s and ISO of 50. The thermal imaging sensor was configured to measure temperature in range of (-10, 40). On the first day of data collection, the raspberry pi had to be restarted due to a loose power connection. Since the thermal cameras requires at least 10 minutes before stable readings can be taken again, 9 data points concerning CTD had to be removed from the dataset. This is elaborated more in Chapter 4.

**Species Constraints:** There are two types of trees namely evergreen and deciduous trees. During winters, deciduous trees loose their leaves, thus hampering NDVI calculation. Hence, our analysis was constrained only to evergreen trees due to data collection in the winters. The species namely Red pine and Eastern White Pine were selected because they still have leaves in the winter (evergreen) and they are the most widespread and easily accessible evergreen trees found from CARTO in the city of Cambridge around MIT. The sites of data collection experiments are shown in Figure 3.9.

A distribution of the data collected on each of the three separate days is shown in Table 3.1.

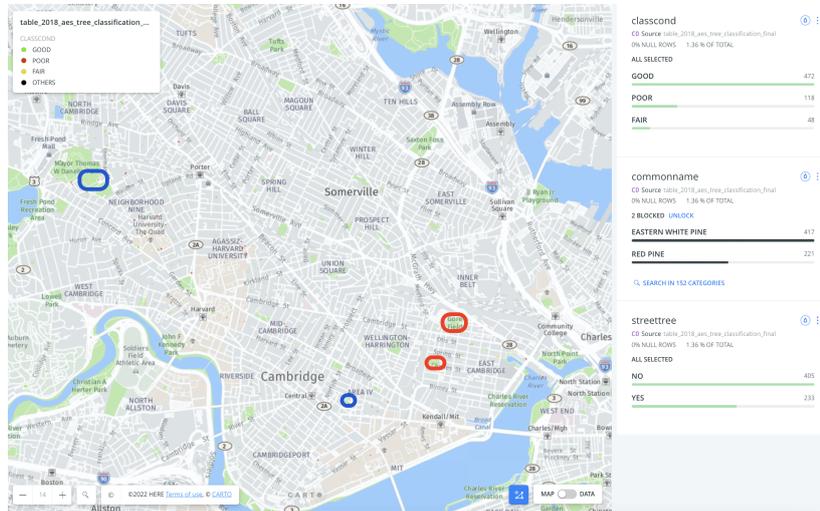


Figure 3.9: The trees were analysed in these locations. The red boxes indicate the Red Pine trees and the blue boxes indicate the Eastern White Pine trees.

Date	Species	Area	Number of Trees	Health Distribution	Weather
23/2/22	Red Pine	Gore Field, John Ahern Field	33	Good:20 Fair:7 Poor:5	8°C, Sunny
24/2/22	Eastern White Pine	Columbia Park, Cherry Street	4	Good:1 Fair:0 Poor:3	-3°C, Cloudy
1/3/22	Eastern White Pine	Danehy park, Garden Street	13	Good:6 Fair:1 Poor:6	0°C, Cloudy

Table 3.1: Data collection experiments undertaken in Cambridge, USA. The health refers to the health condition from ground truth dataset.



Figure 3.10: Using the system with push button trigger

## Chapter 4

# Analysis and Results

In this chapter, we will evaluate the custom Mask R-CNN model and discuss the main takeaways from this evaluation in Section 4.1. This is followed by the analysis of the tree health parameters namely NDVI and CTD measured by our system and their comparison with the ground truth dataset in Section 4.2 along with the main takeaways from this analysis.

### 4.1 Mask R-CNN analysis

#### 4.1.1 Training Curve

As discussed in Section 3.4.3, the model was trained on Google Cloud Platform on a N1 Instance with 2 vCPUs and 13 GB memory. Our manually annotated dataset (refer Section 3.4.2) consisting of 51 images was split in the ratio of 70:30 for training : testing. During retraining, each epoch took approximately 3 hours on the N1 instance. The configuration settings during training the model are elaborated in Appendix A.1. The model weights were saved at end of every epoch.

The training curve of the model is shown in Figure 4.1. It is seen from the training curve that only a small number of epochs (3 in our case) are sufficient to reach the optimal validation loss on the test set owing to retraining of only the head layers of Mask R-CNN.

#### 4.1.2 Performance

##### Full Model

To measure the performance of our model, we calculated the standard evaluation metrics [9] as recommended by COCO. Specifically, we measured mean Average Precision (mAP) / Average Precision (as per [9]) at different IoU thresholds. The performance of our custom Mask R-CNN without quantization is shown in the first entry of Table 4.1. Recall from Section 2.4.1 that increasing the IoU means that the segmentation mask generated by the model should intersect more strictly with the ground truth mask, thus requiring the model to have better object localization. The same behaviour is seen in Figure 4.2 where increasing

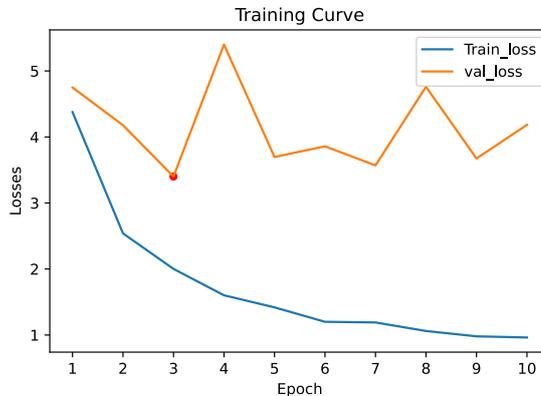


Figure 4.1: **The Training curve of Mask R-CNN with epochs=10 and batch size=4, the red point indicates point of minimum loss**

<b>Model</b>	$AP(IoU=0.5:0.95:0.05)$	$AP(IoU=0.5)$	$AP(IoU=0.75)$
Custom Mask R-CNN TF	0.489	0.938	0.500
Custom Mask R-CNN TF-lite (Dynamic Quantization)	0.491	0.938	0.500

Table 4.1: **Performance of custom R-CNN model (Full and Quantized model)**

the IoU from 0.5 to 0.95 leads to a decrease in Average Precision of the model as expected.

For context, the originally published Mask R-CNN [55] achieved an  $AP(IoU=0.5:0.95:0.05)$  of 33.1 on COCO where the problem is more complex and involves segmenting 81 different objects. Our  $AP(IoU=0.5:0.95:0.05)$  is higher than the originally published Mask R-CNN [55]. However, our problem is much more simple with the segmentation of only 2 objects (tree canopy and background). Nevertheless, Mask R-CNN being state-of-the-art model for instance segmentation and our high AP values inspire confidence in usage of this model for real world deployments in uncontrolled environments.

In order to measure the stability of our results, a k-fold cross validation was also performed with k=3, in order to evaluate the performance of the model on different training and test splits. The results of the k-fold cross validation are shown in Table 4.2 further showcasing the stability of our results.

<b>Cross Validation Fold</b>	<b>1</b>	<b>2</b>	<b>3</b>
$AP(IoU=0.5)$	0.82	0.87	0.75

Table 4.2: **Results of 3-Fold cross validation of custom Mask R-CNN model**

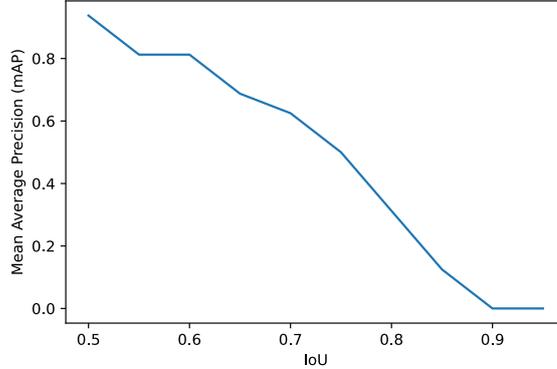


Figure 4.2: The AP scores with increasing IoU thresholds as per COCO metrics [8] for the full model

### Quantized Model

As discussed in Section 3.4, the full model was optimized to run on the edge (raspberry pi 4 in our case) using dynamic quantization as data privacy is an important requirement of our system. A comparison of inference time and model size comparing both the full model and the quantized model are shown in Table 4.3. Most importantly, as expected from [23], the inference time was reduced from 15 seconds to 7 seconds on the raspberry pi 4.

An example of segmentation outputs generated by the full model and quantized model on the same image is shown in Figure 4.3. Since our system does not need to be real time, batch processing can be applied once all the images are collected at the end of the day in a real world large scale deployment.

Model	Inference Time	Model Size
Custom Mask R-CNN TF	15s	255.9 MB
Custom Mask R-CNN TF-lite (Dynamic Quantization)	7s	65 MB

Table 4.3: Comparison between full and quantized model

The performance of the quantized model in terms of Average Precision is shown in the second entry of Table 4.1. From Table 4.1, it is seen that there is no significant reduction in performance using quantization.

In fact, it may appear that the  $AP^{(IoU=0.5:0.95:0.05)}$  for quantized model is slightly increased compared to the full model. On further exploring this anomaly, it was found that this behaviour is exhibited due to our annotated dataset, where most images contain only one full tree canopy as ground truth. Thus, a model (non-quantized model) generalising better to find partially visible tree canopies in addition to the full tree canopy is penalised in terms of precision (False Positive). Further, it is seen from Figures 4.2 and 4.4 that the perform-



(a) Segmentation output from Mask R-CNN using full Tensorflow model  
 (b) Segmentation output from Mask R-CNN using Tensorflow-lite (quantized)

Figure 4.3: **Outputs from custom Mask R-CNNs in Tensorflow and the quantized Tensorflow-lite model showcasing no major visual performance reductions**

ance of the quantized model decreases more than the full model at higher IoUs (IoU= 0.85 for the quantized model compared to IoU=0.90 for the full model) signifying it is marginally poorer at object localisation compared to the full model.

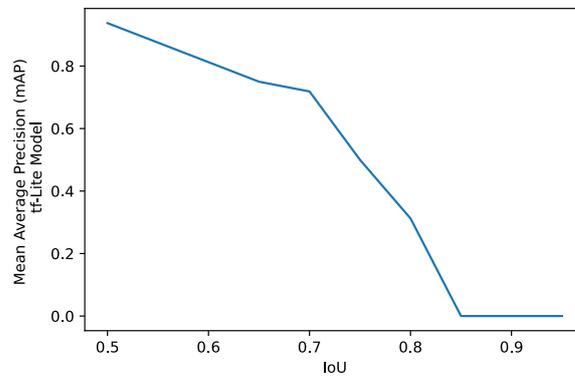


Figure 4.4: **The AP scores with increasing IoU thresholds as per COCO metrics [8] for the quantized tf-lite model**

### 4.1.3 Takeaways

Our custom Mask R-CNN was found to be suitable for segmenting tree canopies from multispectral (RGN) images with high AP values despite the small dataset size. A larger dataset for training will further allow for even improved generalisations. Further, no performance reduction was observed on quantiz-

ing the model. The inference time of the quantized model is half compared to the non-quantized model. Given our use case and system requirements (Section 1.2), batch processing of the captured data is sufficient to maintain data privacy as well as system performance. For instance, in case the system is deployed on moving vehicles, the captured data can be processed while the vehicle is waiting at traffic lights.

## 4.2 Analysis of Tree Health

In this section, we will evaluate the data measured by our system during the data collection experiments (Section 3.6) with the ground truth dataset obtained from the municipality of Cambridge. We used three parameters from the ground truth dataset namely Ground Truth Condition (Health), Remote NDVI and Area of tree (measured using aerial LiDAR) out of all the parameters present in the dataset. The choice of comparison with these three parameters is justified and explained here:

- Ground Truth Condition (health): The ground truth dataset classifies tree health condition in three categories namely good, fair and poor. For evaluation purposes, these parameters were converted to 3, 2, 1 respectively. This parameter will function as the ground truth health condition that our system serves to identify.
- Remote NDVI: NDVI using satellites was measured in the summer of 2018. Since we are measuring NDVI terrestrially instead of from overhead like satellites as well in winters than summers, a comparison and correlation with this parameter served to give interesting results. Unless otherwise stated, our system measured NDVI will always refer to corrected NDVI as discussed in Section 3.2.
- Area of tree (measured using aerial LiDAR): The tree canopy area from overhead was measured using aerial LiDAR. Recall from the Chapter 2, in earlier works [44], the usefulness of aerial LiDAR measured attributes was questioned and a correlation with this parameter will help future researchers in this field regarding the incorporation of LiDAR.

**Data Cleaning:** During the first day of data collection experiment, the raspberry pi hung up due to unknown reasons leading to a forced restart. On the third day of the experiments, owing to cold temperatures, the power supply had to be changed during data collection. These interruptions and restarts resulted in unstable readings of the canopy temperature by the thermal imaging sensor for 9 trees. As a result, these 9 data points were removed from our dataset generated using data collection experiments. A scatter plot of CTD and NDVI before and after data cleaning is shown in Appendix Figure A.1. The distribution of trees after data cleaning is shown in Appendix Table A.1.

### 4.2.1 High-level analysis

A comparison of our system measured NDVI and Remote NDVI is shown in Figure 4.5. As seen in the Figure, our measured NDVI varies in a similar way to the Remote NDVI.

Further, a comparison showing the variation of raw NDVI (NDVI without correction), measured NDVI and Remote NDVI is shown in Appendix Figure [A.3](#).

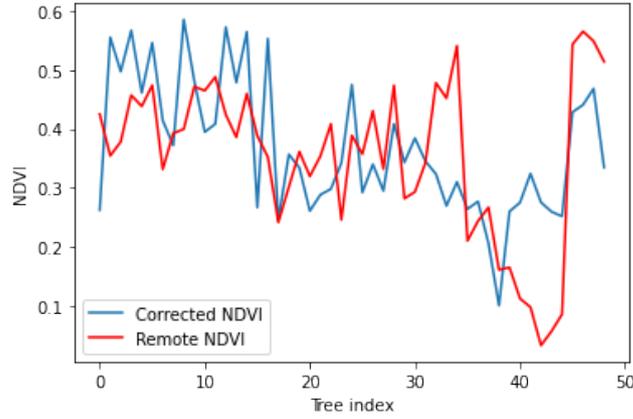


Figure 4.5: **Variation of measured NDVI vs Remote NDVI for trees observed during data collection experiments. Tree index refers to tree ID in our dataset.**

Pearson’s correlation coefficient ( $r$ ) was measured to calculate the strength of linear relationship between our measured parameters and ground truth data.

The correlation matrix comprising of all of our measured parameters with the three ground truth parameters namely Ground Truth Condition, Remote NDVI and Area is shown in Figure [4.6](#). Further, the correlation results between our measured NDVI and CTD with the ground truth parameters is shown in Table [4.4](#).

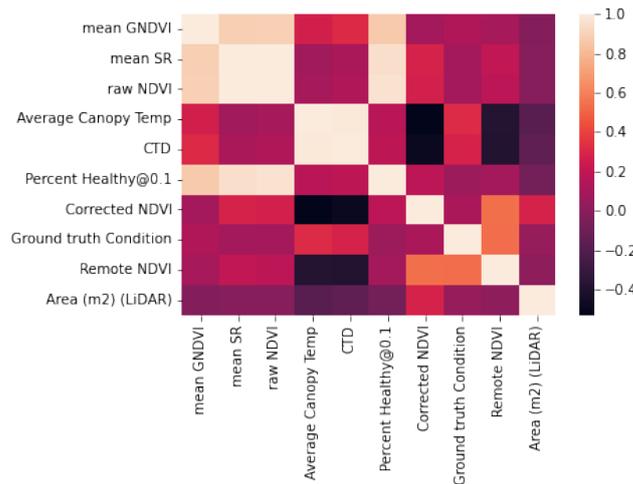


Figure 4.6: **Correlation matrix between our measured values and ground truth parameters**

Variables		Pearson Correlation (r)	Significant at (p < 0.05)
NDVI	Remote NDVI	0.54	Yes
CTD	Remote NDVI	-0.38	No
NDVI	Ground Truth Health	0.11	No
CTD	Ground Truth Health	0.28	Yes
NDVI	Area (m2) (LiDAR)	0.28	Yes
CTD	Area (m2) (LiDAR)	-0.15	No

Table 4.4: **The correlation between our measured parameters and ground truth parameters**

From the correlation matrix, it is clear that there is almost no correlation between NDVI and CTD. Thus, they are independently measuring two different attributes related to tree health and useful to incorporate in the system. From Table 4.4 it is seen that there is a moderately strong and significant correlation ( $r=0.54$  with  $p < 0.05$ ) between our measured NDVI and remote NDVI. For context, in recent works [42], the correlation between NDVI measured using two different satellites was found to be 0.74. Even though our NDVI and the ground truth Remote NDVI were measured four years apart and the former was measured from the ground while later was measured from overhead, this strong correlation shows the validity of our approach as well as motivates further research in ground based NDVI measurements. Further, a Bland-Altman plot widely used in clinical diagnosis to showcase the agreement between two methods (instead of strength of relationship like Pearson’s coefficient) is shown in Appendix Figure A.4. From the Bland-Altman plot, it is seen that there is strong agreement between the two methods (Remote NDVI and Terrestrial NDVI) with 98% (more than 95%) points lying between the difference  $\pm 1.96SD_{difference}$  range.

Further, it is seen that the CTD has weak-moderate correlation ( $r=0.28$  with  $p < 0.05$ ) with ground truth tree health condition. However, we infer that this correlation can also be attributed to the skewed distribution of the dataset where more trees are rated as having good condition compared to poor and fair conditions with more trees being measured with high canopy temperatures on a hot day. This is explored in depth in the next subsection.

Interestingly, due to unknown reasons, NDVI is found to have significant and weak correlation ( $r=0.28$  with  $p < 0.05$ ) with the Area of tree canopy obtained from the ground truth dataset.

## 4.2.2 Low-level health condition analysis

The distribution of CTD and NDVI with respect to health conditions from the ground truth dataset is shown in Figure 4.7. Further visualisation of these distributions with respect to health conditions for both the species observed is also shown in Figures 4.8 and 4.9. The mean NDVI and CTD for each species with respect to their ground truth health conditions is also shown in Tables 4.5

and 4.6

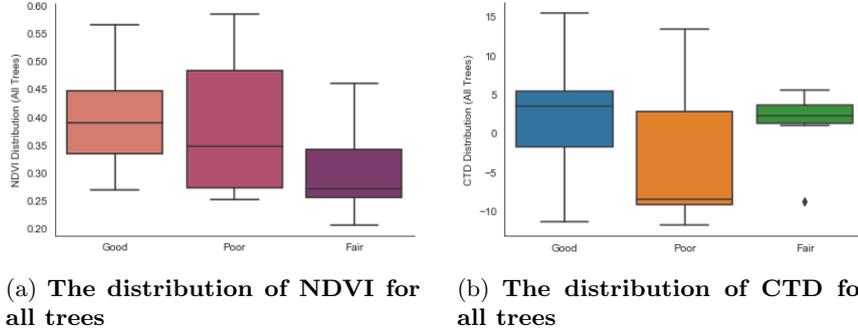


Figure 4.7: The distribution of NDVI and CTD for all trees with respect to health

Species / Health	Good	Fair	Poor
Red Pine	$0.37 \pm 0.07$	$0.28 \pm 0.05$	$0.28 \pm 0.03$
Eastern White Pine	$0.49 \pm 0.08$	0.46	$0.43 \pm 0.12$

Table 4.5: The mean measured NDVI for each species with respect to the ground truth health condition

Species / Health	Good	Fair	Poor
Red Pine	$4.63 \pm 3.64$	$2.89 \pm 1.78$	$6.99 \pm 4.85$
Eastern White Pine	$-9.1 \pm 1.88$	-8.59	$-9.1 \pm 1.88$

Table 4.6: The mean measured CTD for each species with respect to the ground truth health condition

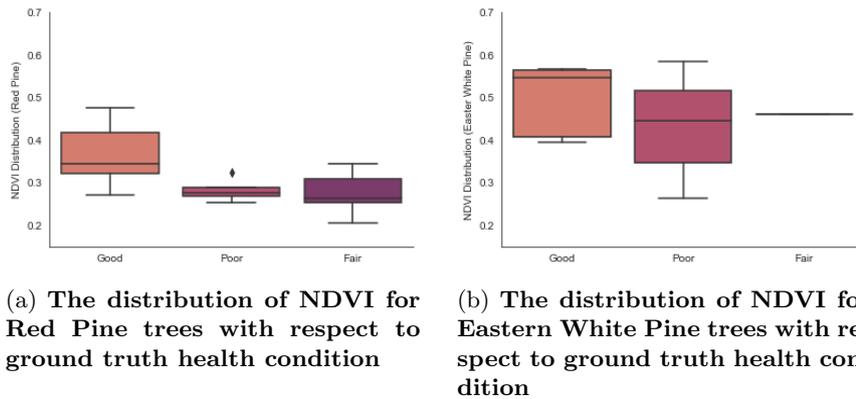


Figure 4.8: The species-wise distribution of NDVI with respect to ground truth health condition

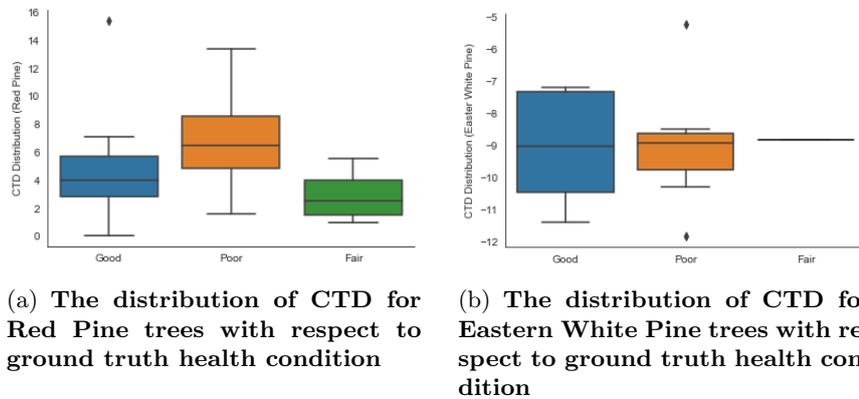


Figure 4.9: The species-wise distribution of CTD with respect to ground truth health condition

From the NDVI distribution in Figure 4.7, it is seen that the distribution of NDVI is mostly in agreement with the ground truth health conditions. This implies that the trees in good health have higher measured NDVI values than trees in poor and fair condition as expected (refer Section 2.3). From the CTD distribution in Figure 4.7, it is seen that CTD is higher for trees in good condition than trees in poor and fair condition but a species wise analysis following up contradicts this result.

### Species Wise Analysis

Moving on to species-wise analysis, it is inferred from Table 4.5 and Figure 4.8 that the distribution of NDVI values is in agreement with health condition of trees for both Red Pine and Rastern White Pine species. In fact, mean NDVI values for Red pine as shown in Table 4.5 are significant ( $p < 0.05$ ) for good and poor condition trees. The mean NDVI values for eastern white pine are insignificant ( $p > 0.05$ ). However, from the NDVI distributions for Eastern White pine in Figure 4.8, it is seen that majority of distribution for good conditions trees still have higher NDVI values than poor and fair condition trees.

From the species wise CTD distributions as shown in Figure 4.9 and Table 4.6, no significant relations could be determined except that the CTD values for poor condition red pine trees tended to be higher than red trees in fair and good condition. However, the relations were insignificant ( $p > 0.05$ ).

### 4.2.3 Takeaways

The NDVI measured by our system was found to have significant and moderately strong correlation ( $r=0.54$  with  $p < 0.05$ ) with remote NDVI from ground truth dataset. Further, the Bland-Altman plot in Appendix Figure A.4 shows strong agreement between the remote NDVI approach and our terrestrial NDVI approach. This illustrates the validity of our method to measure NDVI terrestrially even though the ground truth dataset is four years old. A comparison with a more recent dataset in the future will be very interesting. Further, for

both Red pine and Eastern White Pine species, the measured NDVI distributions were also found to be in theoretical agreement with their ground truth health conditions. While a higher CTD was found for Red Pine trees in poor condition than good and fair health condition trees, the same patterns were not applicable for Eastern White Pine trees. This inference about CTD is similar to earlier work such as [60, 32], where it was concluded that the tree species under observation has an important influence in determining correlation with ground truth water stress. However, it is also possible that the data collection for Red Pine trees being performed on a hot day compared to sub-zero temperatures during data collection of Eastern White Pine trees; the applicability of CTD might be questionable for extremely cold temperatures.

From the correlation matrix, it was also observed that both NDVI and CTD have no correlation with each other. Hence, they are independently measuring two different attributes related to tree health and are useful to incorporate in system. Using the NDVI and CTD values, it is straightforward to assign a continuous comparative ranking for the health of trees under observation.

## Chapter 5

# Discussion and Future Work

In this chapter, we expand on the discussions of the obtained results from Chapter 4. Further, we discuss our methodology including the data collection experiments and point out the limitations of our work. From these discussions, possible research directions for future work are derived in Section 5.2.

### 5.1 Discussion and limitations of our approach

#### 5.1.1 Generalisation of Mask R-CNN performance

From Section 4.1, our custom Mask R-CNN developed using transfer learning was found to be suitable for segmenting tree canopies from multispectral (RGN) images despite the small dataset used for training along with its quantized version. While its difficult to say with certainty about its generalisability in different lighting conditions other than those encountered in our dataset (cloudy and sunny), theoretically, we believe that an increase in number of images in the dataset will further help to improve its performance. Further, the inference time of the quantized model is still seven seconds. However, our use case allows for batch processing of images such as when the taxi is not moving or when the garbage truck is at a halt to collect garbage at a site. Also, in order to significantly improve the inference time of Mask R-CNN if needed, the raspberry pi can be replaced with single board computers containing a TPU like NVIDIA Jetson.

#### 5.1.2 Generalisation of tree health results

As discussed in Section 4.2, the NDVI measured by our system has significant correlation with remote NDVI from ground truth dataset along with the Bland-Altman plot showcasing the validity of our system. Our analysis constrained to evergreen trees belonging to two different species gave similar agreement between tree health and the NDVI measured by our system. However, due time constraints, we collected data for 40 trees (data collected for 49 trees). Hence, it is difficult to generalise this performance with certainty to different tree species

in different seasons spread across different parts of the world. Further, our ground truth dataset is four years old and tree health is expected to change at a higher temporal resolution. Therefore, a comparison with more recent ground truth data set along with a more extensive study will be very interesting. *(While we do advocate an extensive study, in the field of tree health monitoring, an evaluation on limited number of trees is widespread as displayed in Tables 2.2 and 2.3 of Chapter 2. Hence, we recommend a deeper collaboration between environmental and computer scientists to find the optimal variety and number of trees for evaluation for large-scale generalisation of a technology)*

### 5.1.3 Comparison with comparable earlier works

Our work is unique in the sense that it is suitable to be deployed for large-scale monitoring of trees without any human intervention. Further, our main results as discussed extensively in Chapter 4 showcase the validity of our approach. Earlier works in the field which measured tree health terrestrially and used low cost sensors are shown concisely in Table 5.1 (they are elaborated in Chapter 2 in conjunction with the sensors used by them) along with the research gaps in them and how our work fills this research gap. The two utmost comparable works [64] lacks in requiring manual human based analysis of images and [47] lacks in outputting only raw values without ground truth validation combined with controlled upward facing deployment respectively. Our work fulfills this research gap by being completely autonomous and suitable to be deployed for opportunistic sensing along with validated ground truth directly provided by a municipality.

Works	Autonomous (No human intervention needed)	Approach	Ground Truth Comparison	Evaluation
[47]	Yes	Mobile (Cars)	No	172 trees, only system output
[32] and [60]	Yes	Handheld	Yes	44 Images (trees not mentioned)
[65]	Yes	Handheld	No	8 trees
[33] and [93]	Yes	Mobile (Ro- bot)	Yes	2 trees in controlled lab environ- ment
[64]	No	Mobile (Cars)	Yes	20 trees
<b>Our work</b>	<b>Yes</b>	<b>Mobile (Cars) and Citizen Science</b>	<b>Yes</b>	40 trees

Table 5.1: **A concise comparison of our work with comparable earlier works in the field. A comprehensive summary of earlier related works is also displayed in Tables 2.2 and 2.3 of Chapter 2**

#### 5.1.4 CTD analysis

While the measured CTD was found to be higher for poor health condition Red Pine trees than good and fair health condition Red Pine trees, the same pattern was not seen for Eastern White Pine trees. Due to time and seasonal constraints, we could not collect data for Eastern White Pine trees when the ambient air temperature was above zero. A further research is required in this aspect to find if the CTD applicability for water stress is species dependent as inferred in [60] and [32] or if the applicability of CTD is questionable for extremely cold temperatures. *(We question the applicability of CTD for extremely cold temperatures due to: in case the behaviour of tree canopy follows the behaviour of human body. Then in extreme cold temperatures, the human body tries to raise the body temperature. A similar behaviour pattern in trees implies that a good health condition tree will have a higher CTD in cold temperatures and vice-versa in hot temperatures. We consulted a tree expert to peruse this insight and are awaiting for confirmation.)*

#### 5.1.5 Deployment: low speed citizen science data collection vs high speed vehicles

Our data collection was performed in a science based paradigm where a person moving at low speed was carrying the system. For opportunistic moving vehicles like taxis and garbage trucks, this can be equated to instances when the vehicle is stationary at a traffic light or at very low speeds after resuming from a halt. If we deploy the same system on moving vehicles, we expect the multispectral imaging sensor to still capture **blur-free** images at urban city level speeds as it was configured to capture images at 60 fps (frames per second) in our experiments. Further, a lot of algorithms are already available for blur reduction in high resolution images [90]. A speed threshold can also be incorporated so that the system will capture the images only when the speed of the vehicle is below a certain threshold speed. However, the thermal imaging sensor used in the system was frame limited to 8.7 fps (frames per second) as per U.S. government export regulations. Hence, a different thermal sensor with higher frame rate might be needed to be incorporated into the system for high speed moving vehicles. We plan a follow up study in the summer of 2022 where the system is deployed on moving vehicles for large scale data collection.

#### 5.1.6 Thermal imaging for detecting internal decay in trunks

Our analysis of tree health was limited to evaluating the canopy part of the trees only. However, during the course of this thesis, we also employed the thermal cameras to try and identify internal cavities and decay in the trunk of the tree as per earlier works discussed in Section 2.2.3. The earlier works employing this technology to detect internal cavities are contradictory where some works [65] [38] [39] advocate using thermal imaging to detect internal decays in trunks while others [36] reject the usefulness of thermal imaging for the same. During our thesis, we found thermal imaging to be useful in identifying superficial surface occlusions generated using broken branches. However, we could not detect any temperature gradients on the trunk of trees indicating presence of cavities or internal decay with varying health conditions. Thus, we align with the findings

of [36] where it is inferred that only very large cavities with cross-sectional area of around 76% of the trunk are visible using thermal imaging and usually, thermal images are not useful to detect internal decays in the trunk of trees. To help the arboriculture community, we have open-sourced our dataset comprising of 140 thermal images of tree trunks collected during this thesis.

## 5.2 Future Work

### 5.2.1 Feasibility of modelling based classification

From the correlation matrix in Figure 4.6, it is seen that there is no correlation between CTD and NDVI values. Hence, to develop an autonomous model to classify tree health, both these measured parameters are useful. A scatter plot between NDVI and CTD values for Red Pine trees is shown in Appendix A.5. From the scatter plot, it is seen that most of the fair and poor condition trees are concentrated around a cluster between NDVI (0.20-0.35) and CTD (0-7). Hence, a simple white-box machine learning algorithms like SVMs with kernel or unsupervised clustering based algorithms can be used to distinguish between good, poor and fair health condition trees. While we do not advocate this approach on small size datasets such as the one collected in this work, future research building on this work with large datasets can incorporate such models to screen the health of trees based on these measured parameters.

### 5.2.2 Mapping relation between UHI effect with Tree health

Urban surfaces like road pavements, building facades and vegetation play an important role in the urban heat island effect. For instance, studies such as [43] and [34] found that increased greenery is associated with reduction in land surface temperatures (LST). Similarly, Venter et al. [91] use satellite derived land surface temperature and NDVI to find a link between them in the city of Oslo. It is found that increasing vegetation is correlated with reducing LST, thus reducing health risk from heat exposure. However, all these studies use satellite derived parameters for both vegetation quantification and land surface temperature measurements. Since our system is developed to be deployed terrestrially, a study extending our system with sensors measuring ambient air temperature and mean radiant temperature along with tree health can be interesting to find an answer to the research question: how much the health of trees (instead of the quantity) influences the urban heat island effect and affects human thermal comfort.

### 5.2.3 Follow up study with large-scale evaluation

Following up from our discussions in the previous section, a future study employing this system is planned in the summer of 2022. The main aim of this study will be to evaluate the system on a large scale with deployment on uncontrolled moving vehicles and automated collection of large datasets. Also, the system would be evaluated on a varied species of trees and we are in talks for collaboration with a leading arboriculture company for ground truth data provision.

#### **5.2.4 Selective and Smart Irrigation**

Currently, trees in urban cities are irrigated at regular fixed intervals which results in some trees being over irrigated and some trees being under irrigated. A high temporal and spatial resolution data related to urban tree health derived from our system can help in precise selective watering of trees and optimize irrigation as a function of their water use. This will help in optimal water resource management and contribute to the existing efforts in the field of selective irrigation.

# Chapter 6

## Conclusions

In this chapter, we will summarise the problem statement, the research questions and the research gap that this work identified and built upon. We will briefly describe the methodology used to tackle the problem and summarise the results obtained. We also enlist the main research contributions of this work and finally, conclude by presenting a brief discussion of limitations of this work and future research paths originating naturally based on this thesis.

Urban greenery is essential for developing resilient cities in the face of climate change. Nowadays, urban trees are experiencing atypical amount of natural and human-induced stresses which affects their functionality, productivity and survival. The current major methods of monitoring urban tree health relies on inspection by arborists (a tree expert), satellite based imagery and airborne sensing by UAVs or aeroplanes. The assessment by arborists while providing highest quality of inspection involves high costs leading to a low temporal and spatial resolution at a city wide scale. Satellite based imagery has high spatial resolution with low spatial granularity and data quality being dependent on availability of clear skies. Airborne sensing involves high operational costs and is unsuitable for highly urbanised environments due to aviation regulations. Further, both airborne sensing and satellite imagery can only capture an overhead view of urban tree canopies, thus missing or misinterpreting vegetation elements present below the canopies. Given the recent investigations in opportunistic sensing at low costs, the goal of this thesis was to answer the following research question: *How can we terrestrially measure the health of urban trees at both high spatial and temporal resolutions autonomously in a low cost fashion?*. The main measurable attributes of the developed method included suitability for mobile sensing, assurance of data privacy, low cost and assessment quality capable of distinguishing between trees with varying health conditions and validated using ground truth.

### 6.1 Methodology and Results

To tackle the research question, we performed an in-depth survey of the current state-of-the-art technologies employed for tree health monitoring and identified a research gap in them. A comparative analysis articulated thermal and hyperpectral/multispectral imaging sensors as a promising technologies to employ

in our system due to low costs, quality of assessment and the flexibility of tree parameters measured by them.

A prototype system consisting of thermal imaging sensor, multispectral imaging sensor and raspberry pi was developed from ground up. Complementing software consisted mainly of four modules namely control module for controlling the sensors, image registration module for data fusion of these sensors, image segmentation module for identifying the tree canopies from raw sensor data in uncontrolled environments and analysis & calculation module to generate the two independent tree health parameters namely NDVI and CTD. A custom Mask R-CNN model was developed using transfer learning for image segmentation on multispectral images and the model was quantized to process images on the edge ensuring data privacy.

The system was evaluated by collecting data for 49 trees in Cambridge, USA and comparing the measured parameters of the developed system with ground truth dataset for the same city. The developed Mask R-CNN performed admirably with an  $AP^{IoU=0.50} = 0.938$  despite the small dataset used for training. The tree health analysis revealed significant and moderately-strong correlation ( $r=0.54$  with  $p - value < 0.05$ ) between our measured NDVI and the remote NDVI from the ground truth dataset. Also, our measured NDVI and remote NDVI were in agreement with each other as per the Bland-Altman test. Further, for both the species of trees analysed, our measured NDVI distributions were found to be in theoretical agreements with ground truth tree health conditions. For CTD, a pattern with a theoretical agreement was applicable for one of the species observed, but the same pattern was not seen in the other species examined. Using the measured NDVI and CTD values, it is straightforward to assign a continuous ranking of trees under observation. The significant results illustrated the validity of our approach which involves measuring the health of trees terrestrially through a system developed for opportunistic low cost sensing capable of collecting data at high temporal and spatial resolutions on a city-wide scale.

## 6.2 Research Contributions

The main research contributions originating from our work are briefly summarised below.

Our work led to the development of a novel system capable of measuring the health of urban trees terrestrially at high temporal and spatial resolutions in a low cost opportunistic fashion. In the process of development of this system, we developed a custom Mask R-CNN which can perform instance segmentation of tree canopies in multispectral images with processing on the edge. This mask R-CNN based approach allowed us to extract useful information from low resolution thermal images.

We also reviewed and performed a comparative analysis of the state of the art methods and technologies used to measure the health of trees. Further, we elaborate and discuss on our experience of using two of these technologies and address the research gap of large-scale opportunistic sensing based urban tree health monitoring.

### 6.3 Limitations

We elaborated and discussed some assumptions and choices made during this work, the main ones of which are briefly described here.

The main limitation of our work is the generalisability of tree health results obtained. Due to winter season, our analysis was only limited to two species of evergreen trees. Further, our ground truth dataset is from the summer of 2018 and a comparison with a more recent dataset along with arborist interaction will be interesting. It is possible and expected that some trees in the dataset might have different health conditions than in the ground truth dataset. Thus, the quality of our results specifically regarding CTD values could have been improved by incorporating the new health labels in the ground truth dataset.

Further, we collected data using our system in a citizen science based paradigm where a person moving at low speed was carrying the system. While our system and method is suitable to be deployed on moving vehicles, the thermal imaging sensor used in the system might need to be replaced with another thermal imaging sensor with a higher frame rate.

### 6.4 Future work

We discussed and identified several future research paths originating from this work, the most interesting ones of which are briefly summarised as follows.

Naturally originating from our main limitation, a follow up study with deployment on moving vehicles with large scale data collection is scheduled in the summer of 2022. A large scale data collection will also allow the employment of white box model-based classification techniques.

Further, the developed system can be extended by adding ambient air temperature and mean radiant temperature sensors in order to quantify a relationship between tree health (rather than quantity of trees) with the Urban Heat Island effect and human Thermal Comfort in cities.

In essence, a comprehensive urban platform measuring the quantity of trees, their health, quantifying the urban heat island effect and air pollution at high spatio-temporal resolutions can be substantially useful in improving the resilience of urban cities in the face of climate change.

## Chapter 7

# Scientific Papers

The work presented in this thesis has led to two papers, the drafts of which are attached herewith. These will be submitted for publication in the coming weeks.

# Tools and methods for Monitoring health and attributes of the urban greenery: A review of tools and methods

Akshita Gupta <sup>\*1,3</sup>, Simone Mora <sup>†1</sup>, and Yakir Preisler <sup>‡2</sup>

<sup>1</sup>Massachusetts Institute of Technology

<sup>2</sup>Harvard University

<sup>3</sup>Technische Universiteit Delft

April 18, 2022

## 1 Introduction

Urban forests, parks and greenery increase cities' resilience due to their ability to mitigate the effects of climate change [53] such as the Urban Heat Island (UHI), a phenomenon that leads to an air temperature difference as high as 10° C [29] in highly urbanized cities compared to the rural or sub-urban surrounding.

In cities, urban canopies and vegetation provide a wide range of ecosystem services such as air filtering, carbon sequestration, reduced energy consumption (via direct cooling and shading) and lower local temperatures [24, 32]. Trees mainly reduce the local heat index by both shading, decreasing the direct incoming net radiation; and transpiration, transferring latent heat energy from the leaf to the atmosphere which in turn reduces the amount of hot air trapped in its surroundings [29]. Both these factors compound their effects to enhance the cooling benefits of trees [29], improving the perceived tolerance of the pedestrians with the environment (Human Thermal Comfort) [29, 15].

However, urban trees are experiencing an ample amount of *abiotic* (external, non biological factors e.g. drought, salinity) and *biotic* (Biological factors, e.g. Insects, bacteria etc.) stressors that are being exacerbated due to climate change [25, 50, 5]. As a result, their functionality, productivity, and survival are of increasing concern [77]. Indeed, trees with poor health cannot provide most of the aforementioned ecosystem services [28, 30]. For instance, trees with low transpiration rates do not cool the environment sufficiently and trees with low growth rate have reduced shading effect.

Although frequent inspections can identify and correct stresses, keeping trees healthy; the greenery is a high-maintenance asset because it is constantly changing due to natural and human-induced causes. Globally, the total cost of inspection, maintenance and settlement of tree damages is estimated to be more than \$2 trillion USD [79]. Maintaining sarge trees is particularly costly [28]; yet large trees can provide up to 8 times more ecosystem benefits compared to smaller ones [47]. In addition, large trees also have higher economic value than young tree, due to their historical-cultural values and due to their size that improves their cooling efficiency.

Nowadays, a tree's health can be inspected by arborists (a tree expert) with good quality results, but usually at high costs, leading to an assessment that has a low spatial and temporal resolution. However, in recent years, a few technology-assisted screening methods have been developed to complement manual methods. Satellite-based imaging can cover large areas although at a low space granularity [21], with data quality depending on the availability of clear skies [21] and low-time resolution because satellites revisit the same spot only every few days. Airborne sensing using Unmanned Aerial Vehicles (UAVs) or aeroplanes leads to an increased spatial granularity [21]; yet it involves high operational costs and may be not suitable in highly urbanized environments due to aviation authority regulations. Further, both airborne sensing and satellite imagery can only capture the overhead view

---

\*aksgupta@mit.edu

†moras@mit.edu

‡ypreisler@seas.harvard.edu

of the urban canopy. As a result, vegetation elements such as green walls, short trees or shrubs present under the tree canopy are often missed [43].

Recently, a number of projects have investigated the use of low-cost technological alternatives to survey the amount of urban greenery present. For instance, by using Google street view (GSV) images to detect the presence of trees [42, 64], or to calculate the species diversity and changes in urban cities [11]. These projects are set within the field of opportunistic and low-cost sensing aimed at developing environmental platforms that can be deployed and operated without the need of an established, expensive infrastructure.

In this paper, we review the state of the art of low-cost and opportunistic technologies to detect and map the health of urban greenery. We aim at understanding what type of health information can be captured, and how those methods compare with traditional ones. We discuss how tree health conditions impact on the expected ecosystem services and we discuss how different stakeholders can utilize greenery health data for the design of cities. We conclude the paper highlighting existing research gaps and discussing design choices for the development of new approaches.

## 2 Characterizing greenery attributes and health

Trees consist of (i) below-ground root system, serving as nutrient, water allocator and transporter, and as a mechanic stabiliser, (ii) a woody trunk and branches, that transfer water (via the xylem conduits) from the soil to the leaves and allocate soluble sugars (products of photosynthesis) to the required locations in the tree; and (iii) leaves, the main site of the plant gas exchange with the atmosphere - the photosynthesis machinery sequestering carbon dioxide and transpiring water vapor.

Early detection of conditions such as cavities, wounds, diseases etc. in trees can help in taking preemptive action to save the tree and prevent damage to human lives as well as property and to maintain the tree functionality [38]. Unlike the simple visual detection of tree damage, tree physiological stress and damage is often undetectable to human eyes and the 'point of no return' can be long after the visual evidence to a human eye [57, 26, 27]. Further, trees under stress reduce their transpiration rate to prevent excessive water loss, and as a result store less CO<sub>2</sub> and decrease their growth rate. Such trees have a weak defensive mechanism and their general health state is damaged, making them vulnerable to insect attack and diseases as well as increasing the chances of mortality. If a tree health state is poor, its effective contribution to the urban environment is limited, and unlike rural environments, fast measures are needed to improve its health. Urban trees experience ongoing stress due to a unfavourable habitat, mainly originated by; soil limitations, energy limitations and management constrains. Thus, the ability to continuously monitor the tree for its health, water use, growth rate and greenness together with understanding the soil quality trends, is crucial for identifying early signs of stress and allowing early intervention and management in order to save trees, money, and reduce environmental damage.

**Soil limitations** Urban soils can be highly variable within small spatial scales and have typical characteristics such as different bulk densities due to soil compaction, the contents of organic matter, or the patchy distribution of coarse natural or human-made materials such as large gravel or construction waste, as often even the parent material is anthropogenic in origin. Urban-forest soil health is the primary determinant of urban-tree health. Soil limitations and quality lead to restricted space for roots to develop, preventing trees to have proper growth and eventually reduce the tree lifespan significantly [31]. The limited soil volume usually available to the urban tree (with respect to forest trees) also suffer from polluted runoff - containing hazardous elements, and increasing the soil salinity thus reducing the root ability to extract water from the soil. Occasional flooding event, that are projected to increase with climate change, will also leads to root oxygen limitation (Anoxia).

**Energy limitation** The main energetic sources required for a tree to function and grow, are radiation, heat, and water. The urban environment has limited sunlight availability due the buildings around it (urban canyon), allowing only short times spans of sufficient sunlight. Heat energy is crucial for most developmental stages and for all biochemical process in the photosynthesis process, according to each species thermal optimum [20]. The main driving force for the water movement in the tree stem, from the roots to the top of the canopy, is the vapor pressure gradient between the air (unsaturated) to the leaves (usually in near saturation of water vapor) [51], also called as the atmospheric demand for water

(VPD - Vapor Pressure Deficit). Above-average air temperatures and heatwave events, which both often occurs in urban environments [63], can lead to an excessive water loss due to high VPD. Trees under drought stress, mainly resulted from high temperatures or low soil water availability, close their stomata (the small pores in the leaf that allow gas exchange CO<sub>2</sub> uptake and H<sub>2</sub>O evaporation) to regulate water loss via transpiration and as a result reduce their local cooling effect. Another outcome of this 'life saving' stomatal regulation, is a reduced photosynthesis that leads to reduced growth. Water availability for urban trees can be unlimited, mainly thanks to irrigation solutions, but the demand for water, is very high and therefore, it is important to match the supply to the increasing demand - to preserve tree health.

**Management constrains** Trees in urban environment, experience intensive pruning, physical damages and poor maintenance due to high competition on land by the other urban residents' constraining also the canopy crown volume. Urban trees also suffer from pollution of vehicles, industry, runoff etc.

## 2.1 Inspection Techniques

In general, arborists measure tree health firstly by visual inspection and non-invasive techniques for screening, diagnostic or evaluation purposes [41] [71]. Soil quality, and soil water quality can be measured by soil sampling and laboratory physical and chemical analysis, by soil sensors that measures electrical conductivity, temperature and water content. Water limitation can be detect by the same soil moisture sensors, by sensing the air dryness (RH and temperature), and by measuring the tree water consumption and status (using sap flow sensors or various leaf scale temporal measurements). Insect and physical damage to leaves and other organs can be visually manually detected, and also by using image analysis deep learning [60, 49, 52].

However, these techniques may not provide sufficient information about the presence of damages and their extent. For instance, external symptoms of decay may be absent even in the presence of internal decay [41]. This in turn may lead to delayed action, when the tree has already reached its terminal stages [10]. In order to verify the health and provide complete information about the extent of damage in trees, visual inspection is sometimes combined with more invasive methods like electrical resistance (involves calculating voltage difference by attaching electrodes and passing electric current through the trunk) or destructive instruments like increment borers (a wooden core from the tree is extracted and analysed). Methods that require drilling and penetration in the sapwood (living wood), thus creating an entry for pathogens or altering the structural integrity of. However, both invasive and non-invasive methods require intensive manual human labour with low spatial scalability as the analysis is done one tree at a time. Overall, the general rule is to start with the least damage causing method for screening the health of trees such as stress wave velocity [71] and then apply a more aggressive technique such as boroscope to get more information about the damage or decay [54]. Currently, various manual inspection methods exist and they are summarised based in Table 2.1. More details can be found in References [71], [41] and [23].

## 3 Inspection strategies for Greenery health

In this section, we review new techniques for low cost opportunistic sensing of urban tree health. The aim in these methods is to generate an accurate representation of single or multiple parameters of a tree such such as chlorophyll, water transpiration etc. Herewith, each approach was further analysed to critically to determine its advantages and disadvantages.

### 3.1 Material and Methods

The reviewed works were scouted by performing a query search on the Google Scholar and Web of Science portals between June 2021 to February 2022. The keywords used in the search were: tree, tree health, vegetation health, vegetation index, mapping, monitoring, inspection, methods, techniques, sensing, sensor, mobile, thermal camera, thermal imaging, infrared, thermography, drive-by, vehicle, vehicular sensor network. The keywords were combined with "AND" and "OR" operations as well as the use of brackets ("()"). Amongst the results generated, the abstracts of the works were read to select relevant papers which were further analysed. The reference list of relevant works was also taken into account and repeated references were also examined. Works older than the year 2000 were excluded from the study; although only two works have been found in the period 2000-2010. After

Method	Property Measured	Output Detail/Resolution/Quality of Assessment	Cost*	Level of invasiveness
<b>Increment borer</b>	Visual inspection	Low	\$	Invasive
<b>Boroscope</b>	Internal visual inspection from inside	Moderate	\$\$	Invasive
<b>Resistograph</b>	Penetration resistance	Moderate	\$\$	Invasive
<b>Shigometer</b>	Single probe electrical resistivity	Moderate	\$	Invasive
<b>Fractometer</b>	Strength and stiffness	Moderate	\$	Invasive
<b>Stress wave velocity</b>	Single path acoustic wave velocity	Low	\$\$	Non-Invasive
<b>Electrical resistance</b>	Multi-probe electrical resistivity	Moderate	\$\$	Non-Invasive
<b>Stress wave tomography</b>	Multi-path acoustic wave velocity	High	\$\$\$	Non-Invasive
<b>Electromagnetic tomography</b>	Electromagnetic wave permittivity	High	\$\$\$	Non-Invasive
<b>Nuclear magnetic resonance (NMR)</b>	Magnetic property	High	\$\$\$\$	Non-Invasive
<b>Electronic nose</b>	Odour	Low	\$\$	Non-Invasive
<b>Gamma-ray computed Tomography</b>	Gamma ray transmissivity	High	\$\$\$\$	Non-Invasive

Table 1: Summary of manual techniques for detecting tree damage (*\* refers to relative cost where \$ is lowest cost and \$\$\$\$ is highest* )

this process, we retained fifty-one papers that were read in depth. After this process, nine papers were found to be irrelevant in the bigger context of the problem and were excluded. Each paper was annotated as well as relevant interesting parts were digitally highlighted for later analysis.

Each of the methods reviewed is classified based on the primary and secondary sensors used and is summarized briefly in Table 2 and 3.

### 3.2 Embedded sensing methods

Embedded methods involves the use of static sensors which are directly attached to the trees or are placed near the trees to measure a property at regular intervals. Usually, they generate data at high temporal resolution with little or no human supervision required. The physical property measured by these methods can vary from detection of sudden vibrations to measure of water uptake and transpiration. Potamitis et al. [56] explores the use of an accelerometer-based sensor which is attached to the tree to monitor the presence of movement of insects or larvae in the internal part of the tree by transmitting the internal vibrations of the trunk. This approach is deemed to be cost efficient as it reduces repeated visits by a human to examine the tree for presence of insects at regular intervals. Shabandri et al. [65] uses various sensors to detect a multiple physical phenomenon such as sudden tree movements, availability of sufficient sunlight and soil moisture; and it sends the gathered

Works	Primary Sensors Used	Auxillary Sensors	Inspection Level	Metric Computed	Deployment Type	Real World Evaluation
[13] and [14]	Thermal Camera	NA	Manual	Cool zones	Handheld	over 2000 trees selected over 10 year
[54]	Thermal Camera	NA	Manual	Temperature gradient, cool zones	Handheld	2 sample trees
[67]	Thermal Camera	NIR modified Camera, LiDAR	Manual	tree temperature rise and disease progress	UAV	40 sample trees
[45]	Thermal Camera	NA	Manual	tree temperature rise and disease progress	Handheld	15 sample trees
[12]	Thermal Camera	NA	Manual	Temperature gradient (Internal Defects)	Handheld	48 trees
[68]	Thermal Camera	NA	Manual	Temperature gradient (Internal Defects)	Handheld	20 sample trees
[21]	Thermal Camera	Visible Camera	Automatic	TWSI from thermal data, LAIe from RGB data	Car	172 trees on 2.52km trees
[48]	Thermal Camera	NA	Manual	CWSI	Handheld	5 sample trees
[34] and [7]	Thermal Camera	Visible Camera	Automatic	Canopy Temperature (Water Stress)	Handheld	44 images (trees not mentioned)
[37]	NDVI sensor, IR thermometer,	Thermal Camera, visible camera, multispectral camera	Manual	NDVI, canopy temperature correlation	Car	20 sample trees
[38]	Thermal Camera	NA	Automatic	Detection of structural defect in trunk	Handheld	8 sample trees
[8] and [72]	NDVI Sensors	LiDAR	Automatic	NDVI value	Mobile Robot	Lab environment
[19]	Hyperspectral camera	LiDAR	Automatic	NDVI value	Stationary	175 trees, setup ambiguous
[74]	RGB Camera	NA	Automatic	Estimated NDVI value after modifying camera	Handheld	controlled indoor setup
[4]	Custom Hyperspectral camera	NA	Automatic	NDVI, NDRE, LAI	Handheld	NA
[36]	Satellite based NDVI	Satellite based LST	Manual	Relation NDVI and LST	Satellite	Entire Mongolia: Coarse Resolution
[81]	Multispectral Remote Sensing Data	NA	Manual	NDVI	NA	8962 Trees
[17]	Satellite based NDVI	Satellite based LST	Manual	Relation NDVI and LST	Satellite	Entire Mongolia: Coarse Resolution
	Airborne Hyperspectral Camera	Airborne LiDAR	Automatic	LAI and NDVI	Airplane	118 trees

Table 2: A summary of emerging approaches analysed for detecting tree health

Works	Primary Sensors Used	Auxillary Sensors	Inspection Level	Metric Computed	Deployment Type	Real World Evaluation
[80]	LiDAR: Vehicle borne and airborne	Infrared Camera	Automatic	Custom health classification in 4 levels	Helicopter and Car	220 trees
[62]	LiDAR	NA	Automatic	LAI	Handheld	3 sample trees
[58]	Array of fixed light sensors	Smartphone	Automatic	LAI	Stationary, Handheld	Exact number not mentioned
[1]	GSV Images	NA	Automatic	Greenery View Index (No health assessment)	Mobile	NA
[75]	Baidu Street View Images	NA	Automatic	Panoramic Greenery View Index (No health assessment)	Mobile	Sanya, China
[44]	GSV Images	NA	Automatic	Shading effect of trees (No health assessment)	Mobile	Downtown area of Boston
[56]	Accelerometer based sensor detecting vibrations	NA	Manual	Presence of insects	Stationary, Attached to tree	11 sample trees
[10]	electrical impedance spectroscopy	NA	Manual	Presence of disease and hydration stress	Stationary, Attached to tree	24 sample trees
[65]	Temperature, moisture, carbon sensor	Air quality sensor, LDR, vibration sensor	Manual	Real time readings for alerts	Stationary, Attached to tree	NA
[76, 3]	Sensors to measure surrounding ambient features	NA	Automatic	Binary health classification	Stationary, Attached around the tree	100 and 14 sample trees respectively
[3, 73]	Sensors to measure surrounding ambient features	NA	Automatic	Binary health classification	Stationary, Attached around the tree	14 and 15 sample trees respectively
[79]	Sensors to measure surrounding ambient features	Information like location, age, site condition etc.	Automatic	Five point health classification	Stationary, Attached around the tree	1418 trees

Table 3: Contd: A summary of emerging approaches analysed for detecting tree health

data to a centralised server to show real time alerts in an app. In [76] and [3], sensors are installed in the surrounding of the tree which are used as input features for developing a machine learning based algorithm to classify between healthy and unhealthy trees. In these works, various features such as air temperature, humidity, soil humidity and soil acidity are fed into a neural network based model in case of [76] and regression model in case of [3]. In a follow up work [73], the same approach is extended with real-time transmission of the sensor data to a server using low-energy wireless networks. In a similar approach, Wu et al. [79] use the bifurcation of features into static and dynamic categories and feed them to a neural network to output a new index namely Urban Tree Health Index (UTHI). The static features are composed of parameters such as tree species, age, rooting area etc. whereas dynamic features consist of parameters such as air temperature, air humidity, soil moisture, tilt angle etc. The dynamic features are fed into a recurrent based neural network (RNN) layer and the static features are fused with the output of the RNN into fully connected neural network layers.

In all these works [76] [3] [73] [79], the ground truth tree health data is obtained using manual inspection. In other works, Borges et al. [10] propose an Electrical Impedance Spectroscopy (EIS system) to assess the physiological stress in the trees. In this work, a pair of electrodes are placed in the trunk of the trees at diametric positions and an alternating current or a voltage with varying frequencies is passed and the resulting impedances are measured. It is found that there is a strong correlation with the ratio of impedances at predefined frequencies with the physiological stress in the trees. In [35] dendrometers (an analog or digital tool to measure trunk growth and shrinkage) are used to detect long term seasonal growth patterns, daily cycles of water uptake and shorter patterns like swelling after rainfall and subsequent drying [18, 61]. Further, dendrometers can be combined with sap flow sensors, which measure water flux through the stem, to estimate both transpiration as well as hydration at tree level and to detect early signs of tree mortality [57]. A throughout review of additional static sensors is provided in [69]. In general, all these methods require installation of a sensor either on or near the trees. Thus, leading to high cost of deployment on a per tree basis with the benefits of reduced human labour thereafter. Further, with the development of new types of sensors, new tree properties can be measured in future.

### 3.3 Imaging Based Methods

In this section, we discuss imaging based techniques such as thermal imaging, hyper spectral imaging and LiDAR based methods to measure the health of trees. The works are classified into each section based on the primary sensor used. A subset of the works reviewed were found to use a combination of multiple approaches and they are classified as multi-sensory methods. The methods in this section allow analysis of multiple trees using a single sensing equipment. This makes them suitable for low cost sensing at the expense of varying quality of sensing. However, it is seen that the analysis of data collected for trees is usually done manually and deterministically by a human.

**3.3.1 Thermal Imaging** Thermal imaging based on IR (InfraRed) radiation emitted from biological materials is one of the most emerging technological approaches for tree health monitoring. It has been mainly used for either detecting cavities in tree trunks or measuring the water stress in trees.

Early works in this field [13, 14] relied on experts manually reviewing thermal images to identify conditions such as cavities and zones of decay in the bark or branches of a tree (wooden part). The areas with cavities appear cooler in thermal footprint than the rest of the bark surface. However, the extent as well as the cause of the decay can hardly be estimated using thermal images alone. Moreover, to provide reliable results, the tree surface has to be shielded from direct sunlight since the sun-exposed areas may show higher temperature in the thermal images; hiding potential damages. Further, since water absorbs the IR radiation, the part which is examined cannot be wet. Also, the surface of the tree being examined has to be free of moss or other vegetation.

Leong et al. [41] argued that while thermal imaging is a good technique to screen trees for possible damage (binary classification), more advanced techniques are needed for an accurate health inspection. A recent review on thermal imaging techniques for tree health assessment by Vidal and Pitarma [71] found that there is no generalised temperature gradient pattern in terms of shape and size, along the bark to detect damage across various species of trees. Although the area near a deteriorated tissue might have a lower temperature, different conditions can generate different temperature gradients

along the trunk of two different trees even if they belong to the same species. The main works analysed by the authors are also summarised concisely in Figure 1 in appendix (As shown from [71]).

Pitarma et al. [54] give an overview of the intricacies as well as the complexities of thermal imaging to detect the health of branches and the trunk taking into account the atmospheric temperature, the exposure to sun and the thermal environment of observations. The method used in this work again relies on the temperature differences between various parts of the tree. Specifically, it is illustrated that even if a part of the tree is exposed to sun, inferences can still be made about the health of a branch by comparing its temperature to atmospheric temperature. However, these inferences require expert knowledge.

On a separate note, while most of previous works focus on identifying colder parts in the tree's trunk as an indicator of poor health, others [67][45] argue that the early onset of diseases in trees is in-turn associated with temperature increase. For instance, Smigaj et al. [67] used a thermal camera mounted on a UAV to detect sub-degree temperature rise in the leaf and canopy temperature of trees affected by *red band needle blight*; a common disease. The infection level of each tree was manually assessed at ground level for validation. Similarly, Majdak et al. [45] found that infected trees have higher trunk temperature than un-infected trees and the differences are more noticeable on warm and sunny days than on cold and cloudy days. In a study that leveraged a drive-by sensing approach, Fuentes et al. [21] mounted both thermal and visible imaging cameras on top of moving vehicles to monitor the green infrastructure of Melbourne, Australia, at a tree-by-tree scale. The tree growth was estimated in terms of Leaf Area Index (LAI) and Tree Water Stress Index (TWSI). The LAI was measured by applying computer vision algorithms to the RGB images of tree canopy to segment leaves from the rest of the environment. The TWSI was measured using the canopy temperatures obtained by the thermal camera. The approach was also deemed privacy preserving as the cameras were mounted in an upward looking fashion facing towards the sky and the images are captured above the pedestrian level. Other work employing thermal cameras for water stress detection include [48], [68] and [37].

Finally, Kwok et al. [38] proposed a machine learning based method to automatically extract abnormal tree parts potentially containing cavities from thermal images installed at static positions. The method used k-mean clustering and Sobel gradient filter to identify potential cavities and the evaluation was performed on a dataset specifically created for this study and consisted of four tree species along with manual verification by an arborist. Similarly, Jiménez-Bello et al. [34] developed an automated method to calculate plant water stress by using a fusion of RGB (visible) and thermal imaging. It was found that the type of tree under examination had an important influence in determining the correlation with ground truth water stress results. Interestingly the process was fully automated using unsupervised classification in RGB images. In the follow study [7], further intricacies involved in usage of thermal imaging to detect water stress are enlightened. It was seen that the size of leaves may have a significant effect on the performance of thermal imaging. Further, the absolute values of CWSI and canopy temperature do not yield accurate correlation with actual water stress and thus, relative comparison with control trees is required. On a separate note, Burcham et al. [12] suggest that thermal imaging does not provide accurate results about the internal condition of trees containing decay and cavities and can only be used to detect superficial bark surface damage like detached bark or mechanical damage.

**3.3.2 Hyperspectral/Multispectral Imaging** In Hyperspectral and multispectral imaging, various bands in the electromagnetic spectrum are captured which may or maynot be in the visible range. This captured data is then used to calculate various vegetation indexes, the most popular of which is NDVI. NDVI stands for Normalized Difference Vegetation Index (NDVI). NDVI relies on the ratio between visible red band and the near infrared (NIR) signals reflected by the vegetation and is widely used for vegetation health assessment. The NDVI index relies on the property of the chlorophyll present in the leaves which absorb red light and the cell structure of the leaves which reflects NIR. Higher NDVI values symbolise healthy photosynthetic capacity while lower values symbolise poor health or presence of stress in trees or absence of vegetation. This technique, while already being measured using satellites and drones can also be measured terrestrially to give an estimate of tree health. This terrestrial measurement can be done using either active NDVI sensors which have their

own energy source or passive NDVI sensors such as hyperspectral or multispectral cameras. Both of these types of sensors are readily available in the market from various manufacturers [78] [2]. Huang et al. [33] give a comprehensive review on the effectiveness of NDVI as a measure of the health of vegetation, but they also argue for the need of calibration of NDVI sensors. Further, Bahe et al. [6] suggest evidence that NDVI values can give accurate stress detection results when comparing data within one species and not across diverse species.

In recent works, Bietresato et al. [8] uses a mobile robot with NDVI and LiDAR sensors to detect health of five plants in a controlled environment. LiDAR is used to model the plant volume whereas NVDI sensors by OptRx [78] (popular in literature) are used to measure plant health and discriminate between vegetation, flower pots and background. The preceding work is continued in [72] and the vegetation thickness and NDVI index are combined to give a diagnostic matrix to give region wise vegetation index map. The approach of combining NVDI and LiDAR is also used in [19] where individual tree are identified from point clouds using both hyperspectral and depth information. Then, for health analysis, geometric parameters of the trees such as height, inclination and crown diameter are determined using LiDAR data and physiological parameters to be determined from hyper-spectral data are left to be explored in a further study. In [74], in order to reduce the cost associated with NDVI sensors, the NDVI index is calculated by modifying a regular RGB camera to remove the NIR rejection filters. In [4], the authors provide the design and implementation of a handheld generic sensor that can configured to give various characteristics of a plant such as NDVI, LAI and nutrient requirements. The maximum distance of the target under measurement from the sensor in this study is 4m.

As discussed earlier, NDVI is also measured using satellites or other airborne approaches, e.g. employing aeroplanes or UAVs. In [36], the NDVI index calculated using satellite data is correlated with Land Surface Temperature (LST). Degerickx et al. [17] utilised LiDAR data for individual tree segmentation and hyperspectral imaging from an aeroplane to detect tree health. The authors in [81] used multispectral remote sensing data to measure health of trees at both pixel level and whole tree level. Lausch et al. [39] gave an overview of various remote sensing strategies available for sensing forest health and advocate the fusion of terrestrial data along with remote sensing based approach to achieve a better health assessment.

**3.3.3 LiDAR** LiDAR stands for Light detection and ranging. It is used to determine variable distances to an object under consideration or surface by targeting it with a pulsed laser and measuring the time for the reflected light to return to the receiver. In [46], the authors used a mobile LiDAR system in a two stage process. In the first stage, the species of the tree is identified by first determining if the tree is coniferous or deciduous and then examining the branching behaviour. In the second stage, the health of the tree is calculated by analysing the point density of the tree which involves an estimation of the leaves surrounding each branch in the tree. Wu et al. [80] compares LiDAR based airborne laser scanning (ALS) and LiDAR ground based mobile laser scanning (MLS) for tree detection, tree species classification and vitality classification. It is found that while ALS in general, gives better performance to MLS, the combination of both ALS and MLS surpasses the performance achieved by only either of them. For tree health alone, more useful features were extracted from MLS compared to ALS. Degerickx et al. [17] derive Leaf Area Index (LAI) for tree health quantification using laser penetration metric of LiDAR. It is found that specifically for LAI, LiDAR performed better than hyperspectral data. Similarly, [62] also uses low cost LiDAR sensor (\$129) for ground based LAI measurement to quantify health of oil palm trees using intensity data.

**3.3.4 Multi-Sensory Approaches** The works discussed in this section use a combination of previously discussed methods. This enables them to measure more than one parameter related to tree health than is possible with a single sensor or overcome the drawbacks of a single sensor.

Kim and Glenn [37] uses a multi modal system comprising of thermal camera, IR thermometers, multi spectral camera and NDVI sensors to detect plant water stress. The use of multiple sensors is based on the hypothesis that water stress causes: 1. leaf temperature changes which are identified using thermal camera and IR thermometers 2. leaf color changes which are identified using multispectral camera and NDVI sensors. The sensors were mounted on a mobile vehicle and evaluation was performed in a controlled environment consisting of irrigated and non irrigated trees. It was found

that the array IR thermometer can serve as low cost alternative to thermal camera, at the expense of bulkiness. Further, the NDVI sensor was found to perform better than the multispectral camera to distinguish between dry and irrigated trees.

Qu [58] proposed two methods to detect tree health based on the leaf area index (LAI). The first approach uses fixed light sensors installed above and below the canopy to measure received and transmitted solar radiation and quantify the LAI. The second approach uses handheld device containing two smartphones which determine the LAI by applying object segmentation on the captured image. Both the methods were evaluated in a controlled environment and the measured LAI values were compared against a commercially available handheld device (LAI-2000) which is used to measure LAI.

Further, Google and Baidu panoramic street view images have also been used to quantify the extent of urban greenery. Li et al. [44] quantified the shading effect of urban trees using panoramic google street view (GSV) images. In [82], the authors used custom hemispherical view panoramic photos and employed image processing on the resulting photos to give a quantification of urban greenery at pedestrian level. The resulting quantification of urban greenery called panoramic green cover index (PGCI) was also correlated with land surface temperatures and NDVI of the urban areas. A similar approach is used in [1] and [75] where the authors used panoramic images from Google and Baidu respectively. However, these street view based approaches have only been used to quantify the spread of urban greenery rather than it's health.

## 4 Discussion

### 4.1 Preliminary Comparison

A comparison of the surveyed methods in terms of working mechanism, cost and quality of tree health assessment and level of invasiveness is shown in Table 4.

While manual methods like nuclear wave resonance or stress wave tomography have high quality of assessment, they are time consuming and infeasible in terms of cost for analysing each and every tree in the urban area. Further, the amount of labour involved is also huge. Terrestrial static methods again lead to high cost for analysing each and every tree due to sensors costs with the benefit of reduced human labour. The quality of analysis is also lower than manual methods. On the other hand, although infrared thermography has been used to compute TWSI [21] [48], most of the perused literature [71] [14] uses it for detection of presence or absence of decay only (binary classification). NDVI sensors give more quantitative information about the health of trees, however, commercially available NDVI sensors (OptRx) are more expensive (cost > \$3000 per sensor) than thermal cameras (\$300-\$3000). Sometimes, the NDVI may reach the maximum value due to the chlorophyll content causing peak greenness. Hence, any health issue may become difficult to detect until the problem progresses enough to reduce the NDVI value from the peak maximum value [22]. This may delay the detection of the problem and prohibit gauging the health accurately. LiDAR based detection methods as discussed in [17] [62] are mostly used to calculate LAI. However, the LAI value varied with respect to commercial handheld LAI measurement devices. Street view based methods [44, 82, 1, 75] while cost effective are only able to quantify the extent of urban greenery at pedestrian level rather than its health.

### 4.2 Privacy Aspects

Thermal cameras are susceptible to privacy invasion, however, a number of techniques are available to tackle this. A subset of these include: [55] where the authors enforce privacy at sensor level. The main idea relied on the assumption that the human facial skin temperature lies in a particular temperature range. Based on this assumption, all the parts of the thermal image lying in this temperature range are blurred out by appropriate modifications to the sensor hardware and firmware. Compared to the traditional and straightforward approach of capturing facial features first and then running post processing on the image to blur the face, this idea is interesting in the sense that no facial features are actually captured in the first place. Similarly, another approach to preserve privacy that was used in [21] to measure tree health was to mount the cameras so that they are mostly upward looking and the frame level is always above the pedestrian height as shown in Figure 5 [21].

Approach	Working Mechanism	Quality of Assessment	Cost	Level of invasiveness
<b>Manual Methods</b>	Depends on Method	Varying on method, generally high	\$\$\$\$	Depends on the method
<b>Terrestrial static methods</b>	Depends on Method	Lower than manual methods	\$\$\$	Depends on method
<b>NDVI</b>	Properties of Chlorophyll	High quality Quantitative Value	\$\$	low
<b>Thermal Imaging</b>	Cavities, temperature gradient and water stress	Cavities, temp. gradient: Mostly Binary Classification, Water Stress: Quantitative Value	\$	low
<b>LiDAR</b>	Uses laser penetration metrics for parameters like LAI and leaf density	Low quality quantitative Value	\$ to \$\$\$	low
<b>Street view based methods</b>	Image processing to detect greenery	No health assessment, only extent of greenery	\$	low
<b>Remote Sensing</b>	Satellite Imagery	Top level view only	\$	low

Table 4: A comparison of approaches to analyse tree health

### 4.3 Use cases

**4.3.1 Healthy greenery and Urban Heat Island Effect** Urban surfaces like road pavements, building facades and green vegetation play an important role in UHI effect. Some of the existing works [40][9] in this domain use an infrared thermal camera to measure the temperature of surface at the pedestrian height. However, these works are based on the assumption that, there is a high correlation between land surface temperature (LST) and ambient air temperature [16]. The City Scanner nodes already contain the ambient temperature & humidity sensors and the addition of a sensor to detect vegetation health can provide a quantitative relation of these parameters with the UHI effect at hyper localised level. The remainder part of this section will give a brief overview of the works perused related to the UHI effect.

Lee et al. [40] analysed three streets in South Korea using a manually held infrared camera and concluded the effect of various urban elements on UHI. Interestingly, while trees provide a shading effect during daytime to reduce temperatures, they also retain more heat during night time. Further, they found that surface temperature were higher near restaurant buildings.

In [16], the authors used satellite based NVDI to suggest increased greenery being associated with reduction in LST as also found in [9]. Binarti et al. [9] used a fusion of satellite imagery and handheld thermal infrared cameras to get an accurate mapping of both horizontal surface temperatures (like roofs of buildings) and vertical surface temperatures(walls, facades). The authors found that while different thermal properties of surface materials effect the LST, the effect on ambient air temperature was insignificant.

Linking UHI and urban tree canopy, Venter et al. [70] use satellite derived land surface temperature

and NDVI to find a link between them in the city of Oslo. It is found that increasing vegetation is correlated with reducing LST, thus reducing health risk from heat exposure. [59] evaluates different characteristics of tree species and studies the cooling effect due to various tree variables such as canopy density, tree height, leaf thickness etc. The work relies on external reviewed studies, all of which follow different study protocols and methodology. Nevertheless, the authors are able to suggest some context specific planting design guidelines.

**4.3.2 Selective Watering** Currently, trees are irrigated at regular fixed intervals which results in some trees being over irrigated and some trees being under irrigated, due to high heterogeneity in water consumption by individual trees and soil hydraulic properties. A high temporal and spatial resolution data related to urban greenery can help in precised selective watering of trees and optimize irrigation as a function of their water use. Using precised urban irrigation (a practice that is becoming more common in agriculture[66]), can improve cooling by transpiration, save water expenses and improve tree health and soil fertility.

## 5 Conclusion

The manual methods of tree health inspection like nuclear wave resonance or stress wave tomography have high quality of assessment, they are time consuming and infeasible in terms of cost for analysing each and every tree in the urban area. Further, the amount of labour involved is also huge. While, satellite and airborne based remote sensing approaches can cover large areas but with low resolution per pixel [21], they are only able to analyse the vegetation from overhead view. Further, the presence of background materials such as grass or shrubs within the same pixel affects the reliability of observations [76]. In contrast, ground based sensing can look at vegetation elements in a more holistic manner. Based on the emerging technologies perused, the use of thermal imaging to analyse attributes of tree such as water stress index and cavities in trees seems promising and suitable for future research. Hyperpectral/multispectral imaging can be used to generate a number of vegetation indexes, the most popular of which is NDVI. Contradictory studies were seen on the usage of LiDAR with some works such as [62] claiming no increase in performance with its addition. It is seen that most of the works perused which measured tree health from ground level except two [21] [37] use manual judgment and processing by humans on the data collected. This makes them inefficacious to be deployed on large scale with data collection at regular time intervals. In essence, although there are ample works which have tried to measure tree health from ground level, measuring this health autonomously without human intervention and scaling it to be suitable for large scale deployment is an open field of research. Given the current advances in the field of artificial intelligence, low cost sensing and edge computing and the strong need to develop climate adaptive cities, new methods for monitoring urban greenery need to be explored at the intersection of environmental and computational science.

## References

- [1] URL [http://papers.cumincad.org/data/works/att/caadria2020\\_028.pdf](http://papers.cumincad.org/data/works/att/caadria2020_028.pdf).
- [2] Ndvi sensor support. URL <https://www.apogeeinstruments.com/ndvi-sensor-support/>.
- [3] URL [https://crec.ifas.ufl.edu/extension/trade\\_journals/2017/2017\\_July\\_monitoring.pdf](https://crec.ifas.ufl.edu/extension/trade_journals/2017/2017_July_monitoring.pdf).
- [4] 2021. URL <https://www.ispag.org/proceedings/?action=download&item=1682>.
- [5] Richard P Allan, Ed Hawkins, Nicolas Bellouin, and Bill Collins. Ipecc, 2021: Summary for policymakers. 2021.
- [6] Michael Bahe, Ryan Murphy, Matthew Russell, Joseph Knight, and Gary Johnson. Suitability of a single imager multispectral sensor for tree health analysis. *Urban Forestry & Urban Greening*, 63:127187, 05 2021. doi: 10.1016/j.ufug.2021.127187.
- [7] C. Ballester, M. A. Jiménez-Bello, Juan R. Castel, and Diego S. Intrigliolo. Usefulness of thermography for plant water stress detection in citrus and persimmon trees. *Agricultural and Forest Meteorology*, 168:120–129, 2013.

# Large-scale monitoring the Health of Urban Trees using Mobile Sensing

Akshit Gupta\*  
MIT Senseable City Lab  
Cambridge, MA, USA  
aksgupta@mit.edu

Simone Mora  
MIT Senseable City Lab  
Cambridge, MA, USA  
moras@mit.edu

Fan Zhang  
MIT Senseable City Lab  
Cambridge, MA, USA  
zhangfan@mit.edu

## ABSTRACT

To be written

### ACM Reference Format:

Akshit Gupta, Simone Mora, and Fan Zhang. 2022. Large-scale monitoring the Health of Urban Trees using Mobile Sensing. In *Proceedings of conference Title (Conference name)*. ACM, New York, NY, USA, 11 pages. <https://doi.org/XXXXXXXX.XXXXXX>

## 1 INTRODUCTION

Urban forests, parks and greenery increase cities' resilience due to their ability to mitigate the effects of climate change. As per the latest IPCC Report [8], global warming and climate change presents multiple risks to humans as well as nature and protecting, managing and restoring ecosystems is fundamental for climate resilient development. In cities, urban canopies and vegetation provide a wide range of ecosystem services such as air filtering, carbon sequestration, reduced energy consumption and decreasing local temperatures [20, 25].

However, urban trees are experiencing an ample amount of *abiotic* and *biotic* stressors that are exacerbated due to climate change [11, 21, 34]. As a result, their functionality, productivity, and survival are of increasing concern [40]. Trees with poor health cannot provide most of the aforementioned ecosystem services [22, 24]. For instance, trees with low transpiration rates do not cool the environment sufficiently and trees with low growth rate have reduced shading effect.

Greenery has been a high-maintenance asset because it is constantly changing due to natural and human-induced stresses. Nowadays, tree's health can be inspected by arborists (a tree expert) with good quality results, but usually at high costs, leading to an assessment that has a low spatial and temporal resolution. However, in recent years, a few technology-assisted screening methods have been developed to complement inspection by arborists. Satellite-based imaging can cover large areas although at a low spatial granularity [18], with data quality depending on the availability of clear skies [18] and low-time resolution because satellites revisit the same spot

only every few days. Airborne sensing using Unmanned Aerial Vehicles (UAVs) or aeroplanes leads to an increased spatial granularity [18]; yet it involves high operational costs and is not suitable in highly urbanized environments due to aviation authority regulations. Further, both airborne sensing and satellite imagery can only capture the overhead view of the urban canopy. As a result, vegetation elements such as green walls, short trees or shrubs present under the tree canopy are missed [31] or even misinterpreted as tree canopy [1].

Recently, a number of studies have investigated the use of low-cost technological alternatives to monitor the environment in urban cities; for instance, using Google Street View (GSV) images to detect the presence of trees, such as [30, 36], or using drive-by sensing to measure air pollution in an area [4]. These projects are set within the field of opportunistic sensing and are aimed at developing platforms that can be deployed and operated without the need of an established, expensive infrastructure and human operator.

Following on this trend and the critical need for managing and protecting urban ecosystems, in this work, we aim to answer the following research question: how can we measure terrestrially the health of urban trees in a low cost opportunistic fashion at a high spatial and temporal resolution? Considering the advantages of drive-by sensing and the low variability of tree health within a day, the system developed in this work will either be suitable to be deployed independently in a drive-by sensing approach like CityScanner[4] or be suitable to be carried by humans in a citizen science based approach. This system serves as the main research contribution of our work. Deployment of this method on large scale will contribute towards protecting and managing nature effectively for cities and municipalities around the world.

## 2 RELATED WORK

A comparison of the surveyed methods in terms of working mechanism, cost and quality of tree health assessment and level of invasiveness is shown in Table 1 from our previous work.

While manual methods like nuclear wave resonance or stress wave tomography have high quality of assessment, they are time consuming and infeasible in terms of cost for analysing each and every tree in the urban area. Further, the amount of labour involved is also huge. Terrestrial static methods again lead to high cost for analysing each and every tree due to sensors costs with the benefit of reduced human labour. The quality of analysis is also lower than manual methods. On the other hand, although infrared thermography has been used to compute TWSI [18] [33], most of the perused literature [37] [14] uses it for detection of presence or absence of decay only (binary classification). NDVI sensors give more quantitative information about the health of trees, however, commercially

\*Akshit Gupta is also a student at TU Delft, The Netherlands

Unpublished working draft. Not for distribution.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted by ACM, provided that the fee of \$15.00 is paid directly to ACM. This permission is granted without fee for non-profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

Conference name, Date, Location

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/XXXXXXXX.XXXXXX>

2022-04-20 22:24. Page 1 of 1-11.

needs  
to  
be  
up-  
dated

105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116

Approach	Working Mechanism	Quality of Assessment	Cost*	Level of invasiveness
Manual Methods	Depends on Method	Varying on method, generally high	\$\$\$\$	Depends on the method
Terrestrial static methods	Depends on Method	Lower than manual methods	\$\$\$	Depends on method
NDVI	Properties of Chlorophyll	High quality Quantitative Value	\$\$	low
Thermal Imaging	Cavities, temperature gradient and water stress	Cavities, temp. gradient: Mostly Binary Classification, Water Stress: Quantitative Value	\$	low
LiDAR	Uses laser penetration metrics for parameters like LAI and leaf density	Low quality quantitative Value	\$ to \$\$\$	low
Street view based methods	Image processing to detect greenery	No health assesment, only extent of greenery	\$	low
Remote Sensing	Satellite Imagery	Top level view only	\$	low

**Table 1: A comparison of approaches to analyse tree health**  
 (\* refers to relative cost where \$ is the lowest cost and \$\$\$\$ is the highest cost for large-scale evaluation of multiple trees )

available NDVI sensors (OptRx) are more expensive (cost > \$3000 per sensor) than thermal cameras (\$300-\$3000). Sometimes, the NDVI may reach the maximum value due to the chlorophyll content causing peak greenness. Hence, any health issue may become difficult to detect until the problem progresses enough to reduce the NDVI value from the peak maximum value [19]. This may delay the detection of the problem and prohibit gauging the health accurately. LiDAR based detection methods as discussed in [17] [35] are mostly used to calculate LAI. However, the LAI value varies with respect to commercial handheld LAI measurement devices. Street view based methods [2, 32, 38, 42] while cost effective are only able to quantify the extent of urban greenery at pedestrian level rather than its health.

### 2.1 Research gaps and influence on design

From the above suvey, it is clear that the manual inspection methods are infeasible due to both time and cost constraints. Further, some of them are invasive as discussed in Section ???. While, satellite and airborne based remote sensing approaches can cover large areas but with low resolution per pixel [18], they are only able to analyse the vegetation from overhead view. Further, the presence of background materials such as grass or shrubs within the same pixel will affect the reliability of observations [39]. In contrast, ground based sensing can look at vegetation elements in a more holistic manner. Based on the emerging technologies perused, the use of thermal imaging to analyse attributes of tree such as water stress index and cavities in trees seems promising and suitable for our problem statement. Hyperpectral/multispectral imaging can also generate a number of vegetation indexes, the most popular of which is NDVI. Contradictory studies were seen on the usage of LiDAR with some works such as [35] claiming no increase in performance with its addition. It is seen that most of the works perused which measured tree health from ground level except two [29] [28] used manual judgment and processing by humans on the

data collected and employed handheld data collection. This makes them inefficacious to be deployed on large scale with data collection at regular time intervals.

In essence, although there are ample works which have tried to measure tree health from ground level, measuring this health autonomously without human intervention and scaling it to be suitable for drive-by sensing is an open field of research. Given the current advances in the field of computer vision and availability of low cost multispectral & thermal cameras which can be integrated with microcontrollers and single board computers, we employed both these technologies in the developed system and employ traditional image processing and custom computer vision models to generate the tree health attributes automatically.

## 3 METHODOLOGY

### 3.1 System Architecture

The block diagram of the entire system architecture along with all the modules is shown in Figure 1. A Raspberry Pi is employed as the central brain of our system and all the other modules are interfaced with it. Herewith, we will give a brief overview of all the other major modules as shown in the architecture diagram. The first four modules are related to hardware while the remaining modules are related to image processing and the applied steps are visualised in the image processing pipeline as shown in Figure.

All the hardware components were encased in a 3D printed case as shown in Figure 2. This made it easy to carry during data collection. Further, the case was designed such that it is suitable to be attachable to moving vehicles using magnets as shown in Figure 3.

- (1) **FLIR Lepton 3.5 and OpenMV Cam H7:** The thermal imaging sensor is attached to openMV cam H7 using a FLIR Lepton adapter module. This module communicates with Raspberry Pi via remote procedure call over USB. The custom

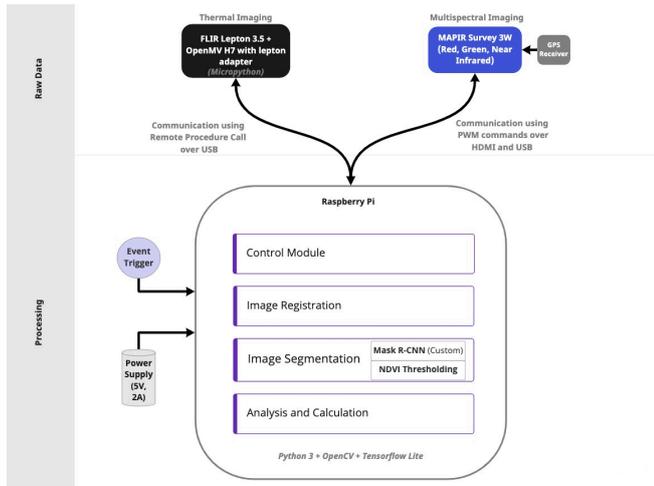


Figure 1: Architecture Diagram of the system

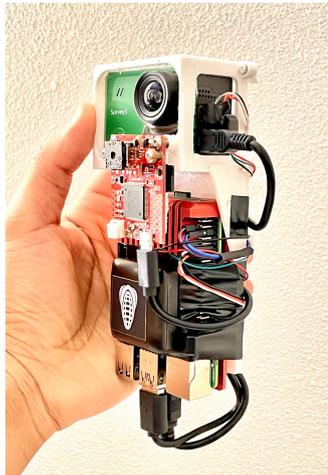


Figure 2: All hardware components encased with the 3D printed case

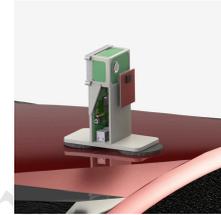
python code running on OpenMV is configured to measure temperature in range of  $-10^{\circ}$  to  $40^{\circ}$  C and return grayscale image data normalised to this temperature range. This means that a measured temperature of  $-10^{\circ}$  C or lower will be scaled to 0 (black) as pixel value and a measured temperature of  $40^{\circ}$  C or higher will be scaled to 255 (white) as pixel value. This temperature range was chosen as per the  $\pm 10$  of lowest and maximum temperature of trees found during the data collection experiments (Section ??). A narrow temperature range is usually preferred to decrease the effect of non uniform noise across the sensor. The micropython code running on OpenMV also handles the serving of the callbacks initiated by the **Control** module.

- (2) **MAPIR Survey 3W with GPS Receiver:** The multispectral imaging sensor is attached to the raspberry pi over USB and the mounting /unmounting of the memory card along with capturing images is handled using PWM signals over the



(a) Concept casing for the system with magnets

(b) The system attached to the top of a car



(c) A closeup view of the system attached to the roof of a car

Figure 3: Concept casing with magnets attachable to moving vehicles

micro HDMI port of MAPIR Survey. Also, a GPS receiver is attached to MAPIR Survey in order to geo-tag all the images captured. Further, this sensor automatically turns on as whenever the power is supplied to the raspberry pi.

- (3) **Event Trigger:** This indicates any event which signals the beginning of processing on raspberry pi from capturing the images to generating NDVI and CTD values. At the current state of the prototype, a press of a push button connected to GPIO 2 of raspberry pi was used as an event trigger for the data collection experiments. Thus, as of now, it can be considered a citizen science based device. However, this trigger can also be an event signalling the co-location of the system with particular GPS coordinates. For instance, when the the system is deployed on moving vehicles, GPS coordinates fetched from a tree inventory database can be used to trigger this processing.
- (4) **Control Module:** This software module handles the event trigger, signals the sensors to captures the images and transfers the captured images to the raspberry pi. For the OpenMV Cam, this involves initiation of callbacks requesting the transfer of current image frame from the thermal imaging sensor. For the multispectral imaging sensor, this involves generating PWM signals to capture an image, mounting the memory installed in the multispectral imaging sensor with the raspberry pi, transferring the captured image to raspberry pi and finally, unmounting the memory from the raspberry pi.
- (5) **Image Registration Module:** Image registration is the task of matching or aligning images taken by two different sensors or different viewpoints into a single coordinate system for further analysis [27]. It involves mapping points from one

image to corresponding points in another image. In our system, both multispectral and thermal imaging sensors have different FOV (field of view) and are not co-aligned. This software module registers the images taken by the thermal and multispectral imaging sensors. In order to align both multispectral (RGN) and thermal images, this module handles the linear translation of RGN images in both horizontal and vertical directions. Further, to compensate for wider FOV of multispectral sensor, this module also handles the zooming in of the RGN images. For our prototype, in case of RGN images, the values of translation in X and Y direction were found to be +50 and +150 pixels respectively and zoom scale was found to be 0.57 (1 indicates no magnification and 0 indicates magnification) to perfectly overlay thermal and RGN images. These parameters were manually found by taking multiple RGN and thermal images and overlaying them. An instance of inputs and outputs utilizing this module is shown in 4. Further, in order to scale this module for large scale production of the system in the future, we also tried automatic image registration using three traditional image registration algorithms namely SIFT, SURF and ORB [3]. However, none of the algorithms were able to find useful keypoints or features in the thermal images. We believe that this was due to the low resolution of thermal images (160x120).

(6) **Image Segmentation Module: The is the most computationally intensive software module of our system.**

Recall that the aim of our system is to calculate NDVI and CTD values of trees in the images. However, these values should only be calculated for the leaves in the tree canopy and not the wooden parts which include branches and trunk. This is solved using a fusion of custom Mask R-CNN and pixel-wise NDVI analysis. Hence, given a RGN image, this task can be broken into two sub problems:

- **Detecting the canopy part of the tree even if the image contains multiple trees:** This is solved using a custom made Mask R-CNN model. The Mask R-CNN model is discussed in more detail in Section ?? and it outputs the instances of the tree canopies in the image by generating a mask (segmentation) over their canopies.
- **Once the canopy of the tree is detected, the segmentation of only the the leaves of the tree without the wooden parts and sky:** Non-vegetation elements such as trunks, branches and sky have very low NDVI values compared to vegetation elements which have significantly higher NDVI. Thus, we employ a thresholding method which first calculates NDVI of each pixel over the segmentation mask given by Mask R-CNN and then, eliminates pixels with NDVI values below a certain threshold. The calculation of NDVI for each pixel is simply computed by deriving the raw values in the red and near infrared channels of the pixel and plugging them in (4). In order to eliminate noise, median filtering is also employed.

The end result using two stage approach gives segmentation of leaves present on the canopy of a tree while eliminating the sky, wooden branches and trunk of the tree. Since our multispectral imaging sensor is uncalibrated, the raw NDVI values generated by it are relative. Hence, a value of 0.02 was

used as cutoff value to eliminate non-vegetation elements in the image. This value was derived using the analysis of the images captured during data collection. An instance of inputs and outputs utilizing this module is also shown in 4.

- (7) **Analysis and Calculation Module:** With the availability of segmentation i.e. tree canopy mask consisting of only leaves for RGN images, the same mask can also be used for thermal images due to image registration. This module handles the calculation of final NDVI and CTD for the tree. The CTD value is computed by calculating the raw temperature value for each pixel by converting its color intensity value in the grayscale thermal image (See (2)), computing the mean temperature over all pixels in the canopy and subtracting the ambient air temperature from the mean canopy temperature. Mathematically, CTD is calculated as:

$$CTD = T_{canopy} - T_{air}, \quad (1)$$

where  $T_{canopy}$  and  $T_{air}$  are canopy temperature and air temperature respectively. The temperature of each pixel is calculated as:

$$T_{pixel} = \frac{P_{value}}{255} * (T_{max} - T_{min}) + T_{min} \quad (2)$$

where  $P_{value}$  is the pixel value in normalised thermal image,  $T_{min}$  and  $T_{max}$  are configured temperature range for the FLIR Lepton 3.5 respectively ( $-10^{\circ}$  and  $40^{\circ}$  C in our case. Then, as per (1), CTD is calculated as:

$$CTD = \overline{T_{pixel}} - T_{air} \quad (3)$$

where  $\overline{T_{pixel}}$  is the average canopy temperature for all segmented pixels in the image and  $T_{air}$  is the air temperature respectively.

To calculate the mean NDVI, each RGN pixel in the image is split into its 3 constituting channels (red, green and near infrared). The raw NDVI value for each pixel is calculated from red and near infrared channels as per (4). To compensate for the aperture adjustment, focal adjustment and other mechanical adjustments performed by the multispectral imaging sensor, the raw NDVI is normalised by dividing with a correction factor similar to the dynamic range of a camera [7]. Mathematically, NDVI is mathematically calculated as:

$$NDVI = \frac{NIR - Red}{NIR + Red}, \quad (4)$$

where  $NIR$  and  $Red$  are near infrared reflectance and visible red reflectance from the leaves of the tree respectively. Mathematically, our corrected NDVI is calculated as:

$$NDVI_{corrected} = \frac{NDVI_{raw}}{|NDVI_{max}|} * |NDVI_{min}| \quad (5)$$

where  $NDVI_{raw}$  is the raw NDVI of a pixel,  $NDVI_{max}$  and  $NDVI_{min}$  are maximum and minimum NDVI values among all pixels in the segmented image.

Finally, the corrected NDVI for the entire canopy is computed by taking the mean over all the corrected NDVI for all pixels in the segmented image. While this approach of adding a correction factor does not calibrate the corrected NDVI to the absolute truth NDVI, it can be used for relative

comparison between the calculated values (which meets our system requirements).

The final calculated CTD and NDVI along with other intermediate calculations such as raw NDVI, Canopy temperature and related indexes like GNDVI (Green NDVI: same as NDVI but with replacement of red band by green in (4), SR (Simple Ratio: NIR/Red), PercentHealthy0.1 (% canopy with raw NDVI > 0.1) were stored on the raspberry pi in a .csv file.

### 3.2 Visual Image Processing Pipeline

A visualization of processing of images after each module is shown in 4. As discussed above, the images from the thermal and multispectral imaging sensors are fetched using the control module based on the event trigger which begins the processing of images in a sequential manner.

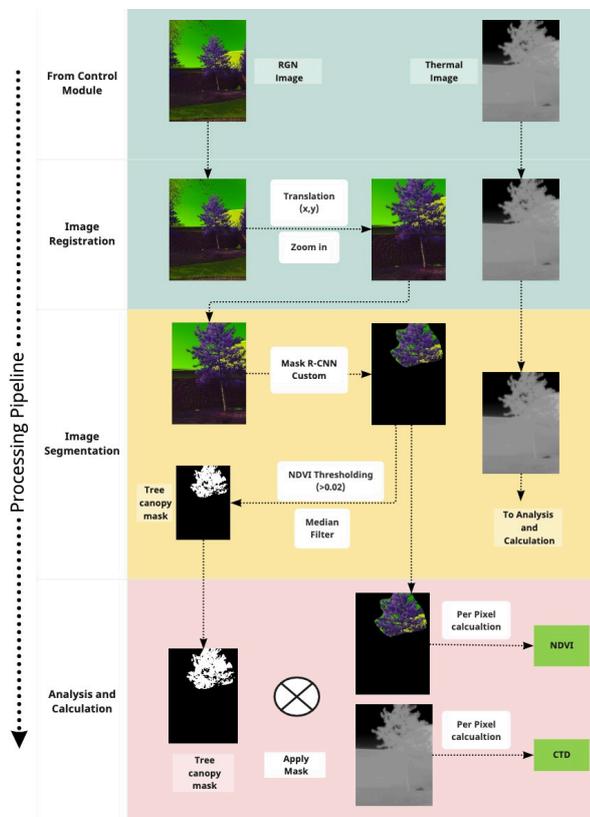


Figure 4: A visualisation of how the images are processed at each step

### 3.3 Development of Custom Mask R-CNN

For our system to operate completely autonomously, the images captured on the basis of an event trigger will be unsupervised and may contain other objects in the image such as cars, buildings, grass, snow in addition to multiple trees. Hence, it becomes imperative to individually identify all the tree canopies in an image and feed them to the calculation and analysis module. The custom mask

R-CNN aims to solve this problem by providing instance segmentation of the tree canopies in the image. In our knowledge, there is no pre-existing model available for instance segmentation tree canopies or even trees for the standard RGB images. Our problem is further complicated as our input is RGN images from the multispectral imaging sensor instead of standard RGB images. For instance, pre-trained models like Deeplabv3 [15] which are able to perform semantic segmentation of trees and vegetation on standard RGB images perform poorly on RGN images.

### 3.4 Training Data

Any deep learning model requires training data in order to optimise the weights and activations of the layers. However, there does not exist any dataset with labels for instances of trees or tree canopies for RGN Images. Hence, we manually created the dataset using the RGN images collected during the data collection experiments (See Section ??). Here, each tree canopy in the image was manually annotated using the popular image annotation tool called LabelMe [41]. During annotation, only tree canopies which were completely present in the image were labelled. After this process, our dataset consisted of 51 annotated RGN images with two classes namely tree canopies and background.

### 3.5 Training Process and Training Curve

Our dataset consists of a relatively small number of images to train a deep learning model like Mask R-CNN from scratch. transfer learning combined with data augmentation can be theoretically employed in order to develop a custom model by using an existing model pre-trained on a different dataset. For our purposes, we used a Mask R-CNN pre-trained [10] on COCO[5] (a dataset with 330K images) with ResNet101 as the backbone. We retrained only head layers (the top layers without the backbone) on our dataset. The batch size was configured as 4 and no. of epochs as 10. The training was performed on Google Cloud platform with N1 instance with 13GB memory and 2vCPUS. We also generated synthetic data by augmenting the original dataset with flips in horizontal, vertical direction and applying Gaussian blur. This increased our training dataset size by 50% and also acts as a regularizer. Our manually annotated dataset (refer Section 3.4) consisting of 51 images was split in the ratio of 70 : 30 for training : testing. During retraining, each epoch took approximately 3 hours on the N1 instance. The training curve of the model is shown in 5. It is seen from the training curve that only a small number of epochs (3 in our case) are sufficient to reach the optimal validation loss on the test set owing to retraining of only head layers. The visual output results from our model are shown in Figure 6.

### 3.6 Model Quantization

Mask R-CNN is a relatively heavy model from both training and inference point of view. Hence, the developed Mask R-CNN was optimized to run on the edge at the cost of possible minute performance reduction. For this, the model built on tensorflow was converted to tensorflow-lite with dynamic range quantization [9]. Dynamic range quantization means that that only the weights of the layers in 32 bit Floats in the full model are stored as to 8 bit INTs while the activations of the layers are quantized during runtime.

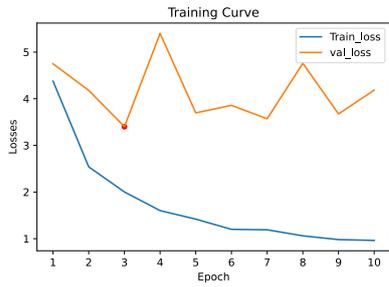
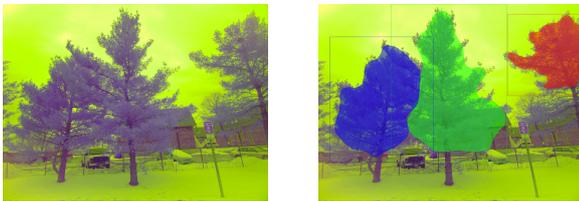


Figure 5: The Training curve of Mask R-CNN with epochs=10 and batch size=4, the red point indicates point of minimum loss



(a) Input RGN Image captured from MAPIR Survey 3W (b) Segmentation output from our Custom Mask R-CNN (instance segmentation model) trained using transfer learning

Figure 6: Performance of our Custom Mask R-CNN. Notice how the model detects each instance of the tree canopy in the image and considers all the other objects as background

Our custom made Mask R-CNN built over tensorflow took around 15 seconds per inference of an image on a raspberry pi 4 while the tensorflow lite model reduced the inference time to 7 seconds with one-fourth of the CPU usage as the original tensorflow model.

## 4 EVALUATION

We evaluated our system using a ground truth dataset from the municipality of Cambridge, USA. We also conducted three data collection experiments to collect data about urban trees. In this section, we elaborate on these datasets followed by the preliminary results obtained.

### 4.1 (Ground truth) Tree Health Dataset

Municipalities in cities obtain ground truth tree health data through city wide surveys in years. For instance, in the city of Cambridge, USA, this survey is performed every 5 years whereas for the city of Delft, The Netherlands, this survey is performed every 2 years. To help with the evaluation of our system, we obtained ground truth tree health dataset for the city of Cambridge, USA through Cambridge Urban Forest Master Plan. This dataset is from 2018 and was obtained through a combination of manual in person arborists, satellite based remote sensing and aerial LiDAR [1]. The dataset classifies the health conditions of trees in three categories namely good, poor and fair. The dataset contains information about 47,063

trees out of which 35,821 are in good health, 5176 are in fair health and 6066 are in poor health. Hence, most of the trees (> 75%) are rated as having good health condition. In addition to this, the dataset contains information about the tree species, common name, the satellite based NDVI, the latitude and the longitude, whether located on a street or not, the shape length and shape area of the canopy, flood tolerance and drought tolerance. This dataset was provided as Shapefiles (.shp, a dataformat used by Geographical Information Systems (GIS)) and was loaded to the online platform CARTO [13] (a GIS and spatial analysis tool).

### 4.2 Data Collection Experiments

We collected RGN and thermal images through our system on three separate days in Cambridge, USA during the month of February, 2022. A push button was used as the event trigger for the system. Hence, we used the developed system as a citizen science project with the 3d printed casing. In total, we collected data for 49 trees spread over two species namely Red pine and Eastern White Pine trees. The multispectral imaging sensor was configured with shutter speed of 1/60s and ISO at 50. The thermal imaging sensor was configured to measure temperature in range (-10, 40). On the first day of data collection, the raspberry pi had to be restarted due to a loose power connection.

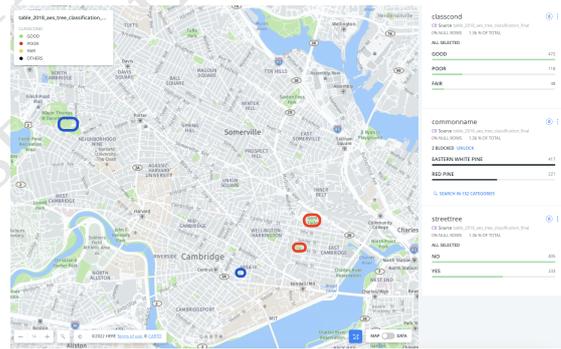


Figure 7: The trees were analysed in these locations. The red boxes indicate the Red Pine trees and the blue boxes indicate the Eastern White Pine trees.

**Species Constraints:** There are two types of trees namely evergreen and deciduous trees. During winters, deciduous trees loose their leaves, thus hampering NDVI calculation. Hence, our analysis was constrained to evergreen trees due to data collection in the winters. The species namely Red pine and Eastern White Pine were selected because they are evergreen and they are the most widespread and easily accessible evergreen trees found from CARTO in the city of Cambridge. The sites of data collection experiments are shown as Figure 7.

**Data Cleaning:** During the first day of data collection experiment, the raspberry pi hung up due to unknown reasons leading to a forced restart. On the third day of the experiments, owing to cold temperatures, the power supply had to be changed during data collections. These interruptions and restarts resulted in unstable readings in a sequence of readings for the canopy temperature by the thermal imaging sensor. As a result, these 9 data points were

removed from our dataset generated using data collection experiments. In the end, our dataset was reduced to contain 40 trees. A distribution of the data collected from each of the tree species after data cleaning is shown in Table 2.

Species	Number of Trees	Health Distribution
Red Pine	26	Good: 15 Fair: 7 Poor: 4
Eastern White Pine	14	Good: 5 Fair: 1 Poor: 8

Table 2: Distribution of trees after data cleaning

### 4.3 Performance of custom R-CNN

To measure the performance of our custom Mask R-CNN model, we calculated the standard evaluation metrics [6] as used by COCO. Specifically, we measured mean Average Precision (mAP) / Average Precision (as per [6]) at different IoU thresholds. The performance of our custom Mask R-CNN without quantization is shown in Table 3. In order to measure the stability of our results, a k-Fold cross validation was also performed with k=3, in order to evaluate performance of the model on different training and test splits as shown in Table 4.

Cross Validation Fold	1	2	3
$AP(IoU=0.5)$	0.82	0.87	0.75

Table 4: Results of 3-Fold Cross Validation of custom R-CNN model

**4.3.1 Quantized Mask R-CNN model.** As discussed earlier, the full model was optimized to run on the edge (Raspberry Pi in our case) using dynamic quantization. A comparison of inference time and model size comparing both the full model and the quantized model are shown in Table 5. Most importantly, as expected [9], the inference time was reduced from 15 seconds to 7 seconds on the raspberry pi.

An example of segmentation outputs generated by the full model and quantized model on the same image is shown in Figure 10. The performance of the quantized model is in terms of Average Precision is shown in Table 3.

Model	$AP(IoU=0.5:0.95:0.05)$	$AP(IoU=0.5)$	$AP(IoU=0.75)$
Custom Mask R-CNN TF	0.489	0.938	0.500
Custom Mask R-CNN TF-lite (Dynamic Quantization)	0.491	0.938	0.500

Table 3: Performance of custom R-CNN model (Full and Quantized model)

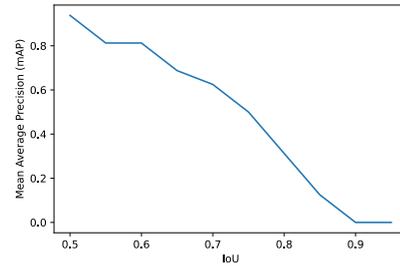


Figure 8: The AP scores with increasing IoU thresholds as per COCO metrics [5] for the full Model

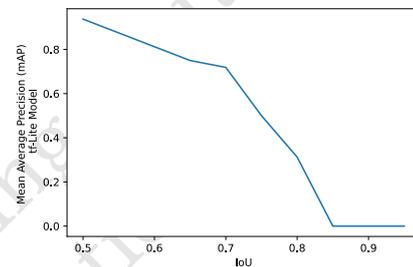


Figure 9: The AP scores with increasing IoU thresholds as per COCO metrics [5] for the quantized tf-lite model

Model	Inference Time	Model Size
Custom Mask R-CNN TF	15s	255.9 MB
Custom Mask R-CNN TF-lite (Dynamic Quantization)	7s	65 MB

Table 5: Comparison between full and quantized model



(a) Segmentation output from Mask R-CNN using full Tensorflow model (b) Segmentation output from Mask R-CNN using Tensorflow-lite (quantized)

Figure 10: Outputs from custom Mask R-CNNs in Tensorflow and the quantized Tensorflow-lite model

## 5 RESULTS FOR THE HEALTH OF TREES

We extracted three parameters from the ground truth dataset namely Ground Truth Condition (Health), Remote NDVI and Area of tree (measured using aerial LiDAR) from all the parameters present in the dataset as these were deemed to be the most appropriate parameters for our evaluation.

A comparison of our system measured NDVI and Remote NDVI is shown in Figure 11. As seen in the figure, our measured NDVI varies in a similar way to the Remote NDVI.

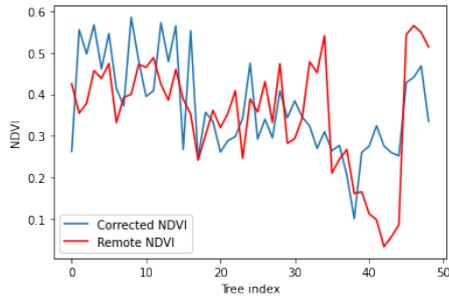


Figure 11: Variation of measured NDVI vs Remote NDVI for trees observed during data collection experiments

Pearson’s correlation coefficient ( $r$ ) was measured to calculate the strength of linear relationship between our measured values and ground truth data.

The correlation matrix comprising of all of our measured values with the three ground truth parameters namely Ground Truth Condition, Remote NDVI and Area is shown in Figure 6.3.1. Further, the correlation results between the measured NDVI and CTD with ground truth parameters is shown in 6.

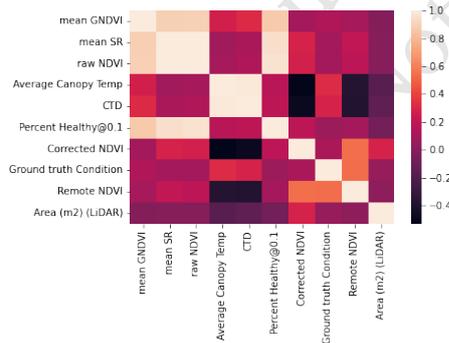
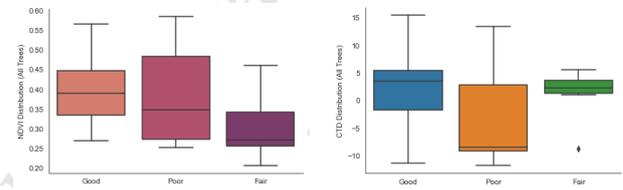


Figure 12: Correlation matrix between our measured values and ground truth parameters

The distribution of CTD and NDVI with respect to health conditions from the ground truth dataset is shown in Figure 13. Further elaboration of these distributions for each of the species is shown in Figures 14 and 15. The mean NDVI and CTD for each species is also shown in Tables 7 and 8.

Variables		Pearson Correlation ( $r$ )	Significant at ( $p < 0.05$ )
NDVI	Remote NDVI	0.54	Yes
CTD	Remote NDVI	-0.38	No
NDVI	Ground Truth Health	0.11	No
CTD	Ground Truth Health	0.28	Yes
NDVI	Area (m2) (LiDAR)	0.28	Yes
CTD	Area (m2) (LiDAR)	-0.15	No

Table 6: The correlation between our measured values and ground truth parameters



(a) The distribution of NDVI for all trees (b) The distribution of CTD for all trees

Figure 13: The distribution of NDVI and CTD for all trees with respect to health

Species / Health	Good	Fair	Poor
Red Pine	$0.37 \pm 0.07$	$0.28 \pm 0.05$	$0.28 \pm 0.03$
Eastern White Pine	$0.49 \pm 0.08$	0.46	$0.43 \pm 0.12$

Table 7: The means of measured NDVI across species and health

Species / Health	Good	Fair	Poor
Red Pine	$4.63 \pm 3.64$	$2.89 \pm 1.78$	$6.99 \pm 4.85$
Eastern White Pine	$-9.1 \pm 1.88$	-8.59	$-9.1 \pm 1.88$

Table 8: The means of measured CTD across species and health

## 6 DISCUSSION

### 6.1 Mask R-CNN performance analysis

The results of the k-fold cross validation are shown in 4 show-case the reliability of our results. Increasing the IoU means that the segmentation mask generated by model should intersect more strictly with the ground truth mask, thus requiring the model to have better object localization. The same behaviour is seen in Figure 8 where increasing the IoU from 0.5 to 0.95 leads to decrease in Average Precision of the model as expected. The originally published Mask R-CNN [23] achieved an  $AP^{(IoU=0.5:0.95:0.05)}$  of 33.1 on COCO where the problem is more complex and involves segmenting 81

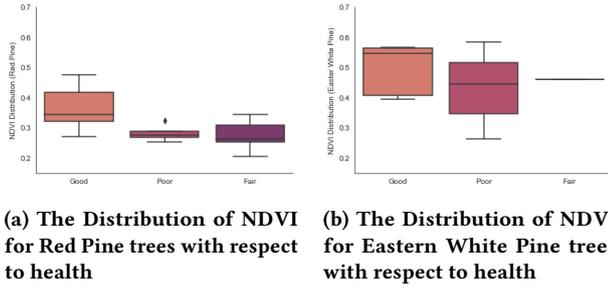


Figure 14: The Distribution of NDVI for Red pine and Eastern White Pine trees with respect to health

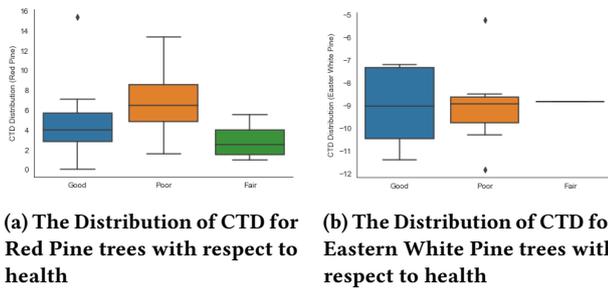


Figure 15: The Distribution of CTD for Red pine and Eastern White Pine trees with respect to health

different objects. Our AP is higher than the originally published Mask R-CNN [23]. However, our problem is much simpler with the detection of only 1 object (tree canopy). Nevertheless, since Mask R-CNN being state of the art model for instance segmentation, our AP values inspire confidence in usage of our model for real world scenarios.

From the table 3, it is seen that there is no significant reduction in performance using quantization. The inference time of the quantized model is half compared to the non-quantized model. Since our system does not need to be real time, batch processing can be applied once all the images are collected at the end of a day in a real world scenario. It may appear that the  $AP(IoU=0.5:0.95:0.05)$  for quantized model is increased slightly compared to the full model. On further exploring this anomaly, it was found that this behaviour is exhibited due to our annotated dataset where most images contain only one full tree canopy as ground truth. Thus, a model (non-quantized model) generalising better to find partially visible tree canopies in addition the the full tree canopy is penalised in terms of Precision (False Positive). Further, it is seen from Figures 9 and 9, that the performance of the quantized model decreases more than the full model at higher IoUs (IoU= 0.85 for quantized model compared to 0.90 for full model) signifying it is slightly poorer at object localisation compared to the full model.

Nevertheless, a larger dataset for training will further allow for even improved generalisations. Given our use case and system requirements (Section ??), batch processing of the captured data is sufficient to maintain data privacy as well as system performance.

For instance, the system can process the captured data while the moving vehicle is waiting at traffic lights.

## 6.2 Tree health analysis

6.2.1 High-level analysis. From the figure 6.3.1 comprising the correlation matrix, it is clear that there is almost no correlation between NDVI and CTD. Thus, they are independently measuring two different attributes related to tree health and useful to incorporate in the system. From 6, it is seen that there is a moderately strong correlation ( $r=0.54$  with  $p < 0.05$ ) between our measured NDVI and remote NDVI. For context, in recent works [16], the correlation between NDVI measured using two different satellites was found to be 0.74. Even though our NDVI and the ground truth Remote NDVI were measured four years apart and the former was measured from the ground while later was measured from overhead, this strong correlation shows the validity of our approach as well as motivates further research in ground based NDVI measurements.

Further, it is seen that the CTD has weak-moderate correlation ( $r=0.28$  with  $p < 0.05$ ) with ground truth tree health condition. However, we infer that this correlation can be attributed to the skewed distribution of the dataset where more trees are rated as having good condition compared to poor and fair conditions. Interestingly, due to unknown reasons, NDVI is found to have significant and weak correlation ( $r=0.28$  with  $p < 0.05$ ) with the area of tree canopy obtained from the ground truth dataset.

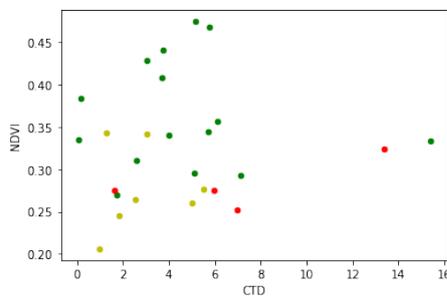
From the NDVI distribution in Figure 13, it is seen that the distribution of NDVI is mostly in agreement with the ground truth health conditions. This implies that the trees in good health have higher measured NDVI values than trees in poor and fair condition. From the CTD distribution Figure 13, it is seen that CTD is higher for trees in good condition than trees in poor and fair condition.

6.2.2 Species Wise Analysis. Moving on to species wise analysis, it is inferred from table 7 and Figure 14, that the distribution of NDVI values is in agreement with health condition of trees for both red pine and eastern white pine species. In fact, mean NDVI values for Red pine as shown in 7 are significant ( $p < 0.05$ ) for good and poor condition trees. The mean NDVI values for eastern white pine are insignificant ( $p > 0.05$ ). However, from the NDVI distributions for Eastern White pine in Figure 14, it is seen that majority of distribution for good conditions trees still have higher NDVI values than poor and fair condition trees.

Overall, the NDVI measured by our system was found to have significant ( $r=0.54$  with  $p < 0.05$ ) and moderately strong correlation with remote NDVI from ground truth dataset. This illustrates the validity of our approach to measure NDVI terrestrially although the ground truth dataset is four years old. Further, for both red pines and eastern white pines, the measured NDVI distributions were found to be in agreement with their ground truth health conditions. While a higher CTD was found for red pine trees in poor condition than good and fair health condition trees, the same pattern conclusion was not applicable for eastern white pine trees. This inference about CTD is similar to earlier work such as [26] [12], where it was concluded that the tree species under observation has an important influence in determining correlation with ground truth water stress.

### 6.3 Future Work

6.3.1 *Feasibility of modelling based classification.* From the correlation matrix in Figure , it is seen that is no correlation between CTD and NDVI values. Hence, to develop an autonomous model to classify tree health, both these measured parameters are useful. A scatter plot between NDVI and CTD values for red pine trees is shown in Figure 16. From the scatter plot, it is seen that most of the fair and poor condition trees are concentrated around a cluster between NDVI (0.20-0.35) and CTD (0-7). Hence, simple white-box machine learning algorithms like SVMs with kernel or logistic regression classifier can be used to distinguish between good, poor and fair condition trees. While we do not advocate this approach on small size datasets such as the one collected in this work, future research building on this work with large dataset can incorporate such models to screen the health of trees based on these measured parameters.



**Figure 16: Scatter plot between NDVI and CTD for red pine trees. The color of the points indicate the ground truth health with red denoting poor, yellow denoting fair and green denoting good condition trees.**

## 7 CONCLUSION

Nowadays, urban trees are experiencing atypical amount of natural and human-induced stresses which affects their functionality, productivity and survival. The current methods for monitoring the health of urban trees mainly comprises of manual inspection by arborists and remote sensing. However, all these methods are riddled with various challenges involving scalability, spatio-temporal resolutions and quality of assessment. In this work, we developed a novel system to measure tree health autonomously from ground level in urban cities. The system can be deployed both in a drive-by sensing paradigm on moving vehicles such as taxis and garbage trucks or be carried by humans in a citizen science paradigm. A custom Mask R-CNN model developed using transfer learning was employed to fuse the data collected by low cost thermal and multispectral imaging sensors on the edge device. The approach was evaluated through data collection experiments performed in Cambridge, USA. The developed Mask R-CNN performed admirably with an  $AP^{IoU=0.50} = 0.938$  despite the small dataset used for training. The tree health analysis revealed moderately-strong correlation ( $r=0.54$  with  $p-value < 0.05$ ) between our measured NDVI and the remote NDVI from the ground truth dataset. Further, for both the species of trees analysed, our measured NDVI distributions were

found to be in theoretical agreements with ground truth tree health conditions. For CTD, a pattern with a theoretical agreement was applicable for one of the species observed, but the same pattern was not seen in the other species examined. Our work illustrates the potential of terrestrial level tree health monitoring and motivates further research in this field.

## REFERENCES

- [1] URL <https://www.cambridgema.gov/-/media/Files/publicworksdepartment/urbanforestmaterplan/techncialreportappendix.pdf>.
- [2] URL [http://papers.cumincad.org/data/works/att/caadria2020\\_028.pdf](http://papers.cumincad.org/data/works/att/caadria2020_028.pdf).
- [3] Opencv: Feature detection and description. URL [https://docs.opencv.org/3.4/db/d27/tutorial\\_py\\_table\\_of\\_contents\\_feature2d.html](https://docs.opencv.org/3.4/db/d27/tutorial_py_table_of_contents_feature2d.html).
- [4] URL <http://senseable.mit.edu/cityscanner/>.
- [5] Common objects in context, . URL <https://cocodataset.org/#home>.
- [6] Common objects in context evaluation metrics, . URL <https://cocodataset.org/#detection-eval>.
- [7] Dynamic range in digital photography. URL <https://www.cambridgeincolour.com/tutorials/dynamic-range.htm>.
- [8] Climate change 2022: Impacts, adaptation and vulnerability. URL <https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/>.
- [9] Post-training quantization nbps: nbps; tensorflow lite. URL [https://www.tensorflow.org/lite/performance/post\\_training\\_quantization](https://www.tensorflow.org/lite/performance/post_training_quantization).
- [10] Waleed Abdulla. Mask r-cnn for object detection and instance segmentation on keras and tensorflow. [https://github.com/matterport/Mask\\_RCNN](https://github.com/matterport/Mask_RCNN), 2017.
- [11] Richard P. Allan, Ed Hawkins, Nicolas Bellouin, and Bill Collins. *Ipcc*, 2021: Summary for policymakers. 2021.
- [12] C. Ballester, M. A. Jiménez-Bello, Juan R. Castel, and Diego S. Intrigliolo. Usefulness of thermography for plant water stress detection in citrus and persimmon trees. *Agricultural and Forest Meteorology*, 168:120–129, 2013.
- [13] Carto. Location intelligence amp; gis for cloud natives. URL <https://carto.com/>.
- [14] Alessandra Catena and Giorgio Catena. Overview of thermal imaging for tree assessmentf. *Arboricultural Journal*, 30, 03 2008. doi: 10.1080/03071375.2008.9747505.
- [15] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation, Dec 2017. URL <https://arxiv.org/abs/1706.05587>.
- [16] Alessia Cogato, Vinay Pagay, Francesco Marinello, Franco Meggio, Peter Grace, Massimiliano De, and Massimiliano Migliorati. Assessing the feasibility of using sentinel-2 imagery to quantify the impact of heatwaves on irrigated vineyards. *Remote Sensing*, 11:1–19, 12 2019. doi: 10.3390/rs11232869.
- [17] J. Degerickx, D.A. Roberts, J.P. McFadden, M. Hermy, and B. Somers. Urban tree health assessment using airborne hyperspectral and lidar imagery. *International Journal of Applied Earth Observation and Geoinformation*, 73:26–38, 2018. ISSN 0303-2434. doi: <https://doi.org/10.1016/j.jag.2018.05.021>. URL <https://www.sciencedirect.com/science/article/pii/S0303243418302095>.
- [18] Sigfredo Fuentes, Eden Tongson, and Claudia Gonzalez Viejo. Urban green infrastructure monitoring using remote sensing from integrated visible and thermal infrared cameras mounted on a moving vehicle. *Sensors*, 21(1), 2021. ISSN 1424-8220. doi: 10.3390/s21010295. URL <https://www.mdpi.com/1424-8220/21/1/295>.
- [19] Nadina J. Galle, William Brinton, Robin Vos, Bidroha Basu, Fábio Duarte, Marcus Collier, Carlo Ratti, and Francesco Pilla. Correlation of worldview-3 spectral vegetation indices and soil health indicators of individual urban trees with ex-captions to topsoil disturbance. *City and Environment Interactions*, 11:100068, 2021. ISSN 2590-2520. doi: <https://doi.org/10.1016/j.cacint.2021.100068>. URL <https://www.sciencedirect.com/science/article/pii/S2590252021000131>.
- [20] E. Gregory McPherson. Accounting for benefits and costs of urban greenspace. *Landscape and Urban Planning*, 22(1):41–51, 1992. ISSN 0169-2046. doi: [https://doi.org/10.1016/0169-2046\(92\)90006-L](https://doi.org/10.1016/0169-2046(92)90006-L). URL <https://www.sciencedirect.com/science/article/pii/016920469290006L>.
- [21] Nancy Grimm, Stanley Faeth, Nancy Golubiewski, Charles Redman, Jianguo Wu, Xuemei Bai, and John Briggs. Global change and the ecology of cities. *Science (New York, N.Y.)*, 319:756–60, 03 2008. doi: 10.1126/science.1150195.
- [22] Kathryn L. Hand and Kieron J. Doick. Understanding the role of urban tree management on ecosystem services. URL <https://www.forestresearch.gov.uk/research/understanding-role-urban-tree-management-ecosystem-services/>.
- [23] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn, Jan 2018. URL <https://arxiv.org/abs/1703.06870>.
- [24] Deborah Hilbert, Lara Roman, Andrew Koester, Jess Vogt, and Natalie Van Doorn. Urban tree mortality: A literature review, 09 2018.
- [25] Sarah E. Hobbie and Nancy B. Grimm. Nature-based approaches to managing climate change impacts in cities. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1794):20190124, jan 2020. doi: 10.1098/rstb.2019.0124. URL <https://doi.org/10.1098/rstb.2019.0124>.

1045  
not  
sure  
if  
needed  
to  
men-  
tion,  
this  
will  
be  
our  
fu-  
ture  
work  
orig-  
inat-  
ing  
from  
here  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
needs  
up-  
dat-  
ing  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102

1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160

- 1161 [26] Miguel Jiménez-Bello, Carlos Ballester, Juan Castel, and Diego Intrigliolo. Development and validation of an automatic thermal imaging process for assessing plant water status. *Agricultural Water Management*, 98:1497–1504, 08 2011. doi: 10.1016/j.agwat.2011.05.002. 1219
- 1162 1220
- 1163 [27] Emna Kamoun. Image registration: From sift to deep learning, Mar 2021. URL <https://www.sicara.ai/blog/2019-07-16-image-registration-deep-learning>. 1221
- 1164 [28] James Y. Kim and David M. Glenn. Multi-modal sensor system for plant water stress assessment. *Computers and Electronics in Agriculture*, 141:27–34, 2017. ISSN 0168-1699. doi: <https://doi.org/10.1016/j.compag.2017.07.009>. URL <https://www.sciencedirect.com/science/article/pii/S0168169916311073>. 1222
- 1165 [29] Coco Yin Tung Kwok, Man Sing Wong, Hon Li, Karena Ka Wai Hui, Florence Wan Yee Ko, Herman Yiu Kay Shiu, and Zihan Kan. Detection of structural tree defects using thermal infrared imaging. 2020. 40th Asian Conference on Remote Sensing: Progress of Remote Sensing Technology for Smart Future, ACRS 2019 ; Conference date: 14-10-2019 Through 18-10-2019. 1223
- 1166 [30] Xiaojiang Li, Chuanrong Zhang, Wei Li, Robert M. Ricard, Qingyan Meng, and Weixing Zhang. Assessing street-level urban greenery using google street view and a modified green view index. *Urban Forestry & Urban Greening*, 14:675–685, 2015. 1224
- 1167 [31] Xiaojiang Li, Chuanrong Zhang, Weidong Li, Robert Ricard, Qingyan Meng, and Weixing Zhang. Assessing street-level urban greenery using google street view and a modified green view index. *Urban Forestry & Urban Greening*, 14(3): 675–685, 2015. ISSN 1618-8667. doi: <https://doi.org/10.1016/j.ufug.2015.06.006>. URL <https://www.sciencedirect.com/science/article/pii/S1618866715000874>. 1225
- 1168 [32] Xiaojiang Li, Carlo Ratti, and Ian Seiferling. Quantifying the shade provision of street trees in urban landscape: A case study in boston, usa, using google street view. *Landscape and Urban Planning*, 169:81–91, 2018. ISSN 0169-2046. doi: <https://doi.org/10.1016/j.landurbplan.2017.08.011>. URL <https://www.sciencedirect.com/science/article/pii/S0169204617301950>. 1226
- 1169 [33] Attila Nagy. Thermographic evaluation of water stress in an apple orchard. *Journal of Multidisciplinary Engineering Science and Technology*, 2:2210–2215, 01 2015. 1227
- 1170 [34] Sophie A. Nitoslawski, Nadine J. Galle, Cecil C. Konijnendijk van den Bosch, and James W. N. Steenberg. Smarter ecosystems for smarter cities? a review of trends, technologies, and turning points for smart urban forestry. *Sustainable Cities and Society*, 2019. 1228
- 1171 [35] Siti Rosli, Fazida Hashim, Thinal Raj, W Mimi Diyana W Zaki, and Aini Hussain. A rapid technique in evaluating tree health using lidar sensors. *International Journal of Engineering and Technology(UAE)*, 7:118–122, 08 2018. doi: 10.14419/ijet.v7i3.17.16634. 1229
- 1172 [36] Ian Seiferling, Nikhil Naik, Carlo Ratti, and Raphael Proulx. Green streets quantifying and mapping urban trees with street-level imagery and computer vision. *Landscape and Urban Planning*, 165:93–101, 09 2017. doi: 10.1016/j.landurbplan.2017.05.010. 1230
- 1173 [37] Daniele Vidal and Rui Pitarma. Infrared thermography applied to tree health assessment: A review. *Agriculture*, 9(7), 2019. ISSN 2077-0472. doi: 10.3390/agriculture9070156. URL <https://www.mdpi.com/2077-0472/9/7/156>. 1231
- 1174 [38] Chen Wei, Meng, Wu Zhiwei, Zhang Yong, and Yang Jaemo. Evaluating greenery around streets using baidu panoramic street view images and the panoramic green view index. *Forests*, 10:1109, 12 2019. doi: 10.3390/f10121109. 1232
- 1175 [39] Yang Wei, Hao Wang, Kim Fung Tsang, Yucheng Liu, Chung Kit Wu, Hongxu Zhu, Yuk-Tak Chow, and Faan Hei Hung. Proximity environmental feature based tree health assessment scheme using internet of things and machine learning algorithm. *Sensors*, 19(14), 2019. ISSN 1424-8220. doi: 10.3390/s19143115. URL <https://www.mdpi.com/1424-8220/19/14/3115>. 1233
- 1176 [40] Zoey Werbin, Leila Heidari, Sarabeth Buckley, Paige Brochu, Lindsey Butler, Catherine Connolly, Lucila Houttuin Bloemendaal, Tempest McCabe, Tara Miller, and Lucy Hutyra. A tree-planting decision support tool for urban heat island mitigation. 10 2019. doi: 10.1101/821785. 1234
- 1177 [41] Wkentar. Wkentar/labelme: Image polygonal annotation with python (polygon, rectangle, circle, line, point and image-level flag annotation). URL <https://github.com/wkentar/labelme>. 1235
- 1178 [42] Yonglin Zhang, Shanlin Li, Xiao Fu, and Rencai Dong. Quantification of urban greenery using hemisphere-view panoramas with a green cover index. *Ecosystem Health and Sustainability*, 7(1):1929502, 2021. doi: 10.1080/20964129.2021.1929502. URL <https://doi.org/10.1080/20964129.2021.1929502>. 1236
- 1179 1237
- 1180 1238
- 1181 1239
- 1182 1240
- 1183 1241
- 1184 1242
- 1185 1243
- 1186 1244
- 1187 1245
- 1188 1246
- 1189 1247
- 1190 1248
- 1191 1249
- 1192 1250
- 1193 1251
- 1194 1252
- 1195 1253
- 1196 1254
- 1197 1255
- 1198 1256
- 1199 1257
- 1200 1258
- 1201 1259
- 1202 1260
- 1203 1261
- 1204 1262
- 1205 1263
- 1206 1264
- 1207 1265
- 1208 1266
- 1209 1267
- 1210 1268
- 1211 1269
- 1212 1270
- 1213 1271
- 1214 1272
- 1215 1273
- 1216 1274
- 1217 1275
- 1218 1276

# Bibliography

- [1] Cambridge urban forest master plan. URL <https://www.cambridgema.gov/-/media/Files/publicworksdepartment/urbanforestmasterplan/technicalreportappendix.pdf>.
- [2] Embedded and networked systems tu delft. URL <https://www.tudelft.nl/ewi/over-de-faculteit/afdelingen/software-technology/embedded-and-networked-systems>.
- [3] Development of an urban greenery evaluation system based on deep learning and google street view. URL [http://papers.cumincad.org/data/works/att/caadria2020\\_028.pdf](http://papers.cumincad.org/data/works/att/caadria2020_028.pdf).
- [4] Ndvi sensor support. URL <https://www.apogeeinstruments.com/ndvi-sensor-support/>.
- [5] Opencv: Feature detection and description. URL [https://docs.opencv.org/3.4/db/d27/tutorial\\_py\\_table\\_of\\_contents\\_feature2d.html](https://docs.opencv.org/3.4/db/d27/tutorial_py_table_of_contents_feature2d.html).
- [6] City scanner, . URL <http://senseable.mit.edu/cityscanner/>.
- [7] The cityscapes dataset, . URL <https://www.cityscapes-dataset.com/>.
- [8] Common objects in context, . URL <https://cocodataset.org/#home>.
- [9] Common objects in context evaluation metrics, . URL <https://cocodataset.org/#detection-eval>.
- [10] Coral usb accelerator. URL <https://coral.ai/products/accelerator>.
- [11] Dynamic range in digital photography. URL <https://www.cambridgeincolour.com/tutorials/dynamic-range.htm>.
- [12] 5 factors influencing radiometric temperature measurements. URL <https://www.flir.com/discover/oem/cores/5-factors-influencing-radiometric-temperature-measurements/>.
- [13] Climate change 2022: Impacts, adaptation and vulnerability. URL <https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/>.
- [14] Kernel2 cameras: Modular array cameras. URL <https://www.mapir.camera/pages/kernel2-array-cameras>.

- [15] Flir lepton 3.5. URL <https://www.flir.eu/products/lepton/?model=3.5+Lepton&vertical=lwir&segment=oem>.
- [16] Far infrared thermal sensor array (32x24 res). URL <https://www.melexis.com/en/product/MLX90640/Far-Infrared-Thermal-Sensor-Array>.
- [17] Openmv cam h7 r2. URL <https://openmv.io/collections/products/products/openmv-cam-h7-r2>.
- [18] Senseable stockholm lab, . URL <https://www.senseablestockholm.org>.
- [19] Mit senseable city lab, . URL <https://senseable.mit.edu/>.
- [20] Monitoring tree health using water-uptake rate. URL [https://crec.ifas.ufl.edu/extension/trade\\_journals/2017/2017\\_July\\_monitoring.pdf](https://crec.ifas.ufl.edu/extension/trade_journals/2017/2017_July_monitoring.pdf).
- [21] Stm32 32-bit arm cortex mcus. URL <https://www.st.com/en/microcontrollers-microprocessors/stm32-32-bit-arm-cortex-mcus.html>.
- [22] Survey3w camera - red+green+nir (rgn, ndvi). URL <https://www.mapir.camera/collections/survey3/products/survey3w-camera-red-green-nir-rgn-ndvi>.
- [23] Post-training quantization: Tensorflow lite, . URL [https://www.tensorflow.org/lite/performance/post\\_training\\_quantization](https://www.tensorflow.org/lite/performance/post_training_quantization).
- [24] Tensorflow, . URL <https://www.tensorflow.org/>.
- [25] Tensorflow lite. URL <https://www.tensorflow.org/lite>.
- [26] Vegetation indices ndvi, evi, gndvi, cvi, true color. URL <https://www.soft.farm/en/blog/vegetation-indices-ndvi-evi-gndvi-cvi-true-color-140>.
- [27] 2021. URL <https://www.ispag.org/proceedings/?action=download&item=1682>.
- [28] Advanced techniques specialization, Jul 2021. URL <https://www.deeplearning.ai/program/tensorflow-advanced-techniques-specialization/>.
- [29] Waleed Abdulla. Mask r-cnn for object detection and instance segmentation on keras and tensorflow. [https://github.com/matterport/Mask\\_RCNN](https://github.com/matterport/Mask_RCNN), 2017.
- [30] Richard P Allan, Ed Hawkins, Nicolas Bellouin, and Bill Collins. Ipc, 2021: Summary for policymakers. 2021.
- [31] Michael Bahe, Ryan Murphy, Matthew Russell, Joseph Knight, and Gary Johnson. Suitability of a single imager multispectral sensor for tree health analysis. *Urban Forestry & Urban Greening*, 63:127187, 05 2021. doi: 10.1016/j.ufug.2021.127187.

- [32] C. Ballester, M. A. Jiménez-Bello, Juan R. Castel, and Diego S. Intrigliolo. Usefulness of thermography for plant water stress detection in citrus and persimmon trees. *Agricultural and Forest Meteorology*, 168:120–129, 2013.
- [33] M Bietresato, G Carabin, D D’Auria, R Gallo, G Ristorto, F Mazzetto, R Vidoni, A Gasparetto, and L Scalera. A tracked mobile robotic lab for monitoring the plants volume and health. In *2016 12th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, pages 1–6, 2016. doi: 10.1109/ mesa.2016.7587134.
- [34] Floriberta Binarti, Pranowo Pranowo, and Soesilo Leksono. Thermal infrared images to identify the contribution of surface materials to the canopy layer heat island in hot-humid urban areas. *Environmental and Climate Technologies*, 24:604–623, 10 2020. doi: 10.2478/rtuuct-2020-0037.
- [35] Elisabeth Borges, Mariana Sequeira, A. Cortez, H. Catarina Pereira, Tania Pereira, Vânia Almeida, João Cardoso, Teresa Maria Vasconcelos, Isabel Duarte, and Neusa Nazaré. Bioimpedance parameters as indicators of the physiological states of plants in situ a novel usage of the electrical impedance spectroscopy technique. *International Journal on Advances in Life Sciences*, 6:74 – 6, 01 2014.
- [36] Daniel Burcham, E. Leong, Yok Fong, and Puay-Yok Tan. An evaluation of internal defects and their effect on trunk surface temperature in casuarina equisetifolia l. (casuarinaceae). *Arboriculture & Urban Forestry*, 38: 277–286, 11 2012. doi: 10.48044/jauf.2012.037.
- [37] Carto. Location intelligence and gis for cloud natives. URL <https://carto.com/>.
- [38] Alessandra Catena. Thermography reveals hidden tree decay. *Arboricultural Journal*, 27(1):27–42, 2003. doi: 10.1080/03071375.2003.9747360. URL <https://doi.org/10.1080/03071375.2003.9747360>.
- [39] Alessandra Catena and Giorgio Catena. Overview of thermal imaging for tree assessment. *Arboricultural Journal*, 30, 03 2008. doi: 10.1080/03071375.2008.9747505.
- [40] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation, Dec 2017. URL <https://arxiv.org/abs/1706.05587>.
- [41] Taihan Chen, Haonan Pan, Mengrong Lu, Jian Hang, Cho Kwong Charlie Lam, Chao Yuan, and David Pearlmutter. Effects of tree plantings and aspect ratios on pedestrian visual and thermal comfort using scaled outdoor experiments. *Science of The Total Environment*, 801:149527, 2021. ISSN 0048-9697. doi: <https://doi.org/10.1016/j.scitotenv.2021.149527>. URL <https://www.sciencedirect.com/science/article/pii/S0048969721046015>.
- [42] Alessia Cogato, Vinay Pagay, Francesco Marinello, Franco Meggio, Peter Grace, Massimiliano De, and Massimiliano Migliorati. Assessing the

- feasibility of using sentinel-2 imagery to quantify the impact of heat-waves on irrigated vineyards. *Remote Sensing*, 11:1–19, 12 2019. doi: 10.3390/rs11232869.
- [43] Andrew M. Coutts, Richard J. Harris, Thu Phan, Stephen J. Livesley, Nicholas S.G. Williams, and Nigel J. Tapper. Thermal infrared remote sensing of urban heat: Hotspots, vegetation, and an assessment of techniques for use in urban planning. *Remote Sensing of Environment*, 186: 637–651, 2016. ISSN 0034-4257. doi: <https://doi.org/10.1016/j.rse.2016.09.007>. URL <https://www.sciencedirect.com/science/article/pii/S0034425716303509>.
- [44] J. Degerickx, D.A. Roberts, J.P. McFadden, M. Hermy, and B. Somers. Urban tree health assessment using airborne hyperspectral and lidar imagery. *International Journal of Applied Earth Observation and Geoinformation*, 73:26–38, 2018. ISSN 0303-2434. doi: <https://doi.org/10.1016/j.jag.2018.05.021>. URL <https://www.sciencedirect.com/science/article/pii/S0303243418302095>.
- [45] David M Drew and Geoffrey M Downes. The use of precision dendrometers in research on daily stem size and wood property variation: a review. *Dendrochronologia*, 27(2):159–172, 2009.
- [46] Yunhe Feng, Chenglu Wen, Pengdi Huang, Cheng Wang, and Jonathan Li. Detection and health analysis of individual tree in urban environment with multi-sensor platform. In *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 7548–7551, 2018. doi: 10.1109/igarss.2018.8518369.
- [47] Sigfredo Fuentes, Eden Tongson, and Claudia Gonzalez Viejo. Urban green infrastructure monitoring using remote sensing from integrated visible and thermal infrared cameras mounted on a moving vehicle. *Sensors*, 21(1), 2021. ISSN 1424-8220. doi: 10.3390/s21010295. URL <https://www.mdpi.com/1424-8220/21/1/295>.
- [48] Nadina J. Galle, William Brinton, Robin Vos, Bidroha Basu, Fábio Duarte, Marcus Collier, Carlo Ratti, and Francesco Pilla. Correlation of worldview-3 spectral vegetation indices and soil health indicators of individual urban trees with exceptions to topsoil disturbance. *City and Environment Interactions*, 11:100068, 2021. ISSN 2590-2520. doi: <https://doi.org/10.1016/j.cacint.2021.100068>. URL <https://www.sciencedirect.com/science/article/pii/S2590252021000131>.
- [49] Ross Girshick. Fast r-cnn, Sep 2015. URL <https://arxiv.org/abs/1504.08083>.
- [50] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation, Oct 2014. URL <https://arxiv.org/abs/1311.2524>.
- [51] Chiew Loon Goh, Ruzairi Abdul Rahim, Mohd Hafiz Fazalul Rahiman, Masturah Tunnur Mohamad Talib, and Zhen Cong Tee. Sensing wood decay in standing trees: A review. *Sensors and Actuators A: Physical*, 269:

- 276–282, 2018. ISSN 0924-4247. doi: <https://doi.org/10.1016/j.sna.2017.11.038>. URL <https://www.sciencedirect.com/science/article/pii/S0924424717300237>.
- [52] E. Gregory McPherson. Accounting for benefits and costs of urban green-space. *Landscape and Urban Planning*, 22(1):41–51, 1992. ISSN 0169-2046. doi: [https://doi.org/10.1016/0169-2046\(92\)90006-L](https://doi.org/10.1016/0169-2046(92)90006-L). URL <https://www.sciencedirect.com/science/article/pii/016920469290006L>.
- [53] Nancy Grimm, Stanley Faeth, Nancy Golubiewski, Charles Redman, Jianguo Wu, Xuemei Bai, and John Briggs. Global change and the ecology of cities. *Science (New York, N.Y.)*, 319:756–60, 03 2008. doi: [10.1126/science.1150195](https://doi.org/10.1126/science.1150195).
- [54] Kathryn L. Hand and Kieron J. Doick. Understanding the role of urban tree management on ecosystem services. URL <https://www.forestresearch.gov.uk/research/understanding-role-urban-tree-management-ecosystem-services/>.
- [55] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn, Jan 2018. URL <https://arxiv.org/abs/1703.06870>.
- [56] Jelle A. "Hiemstra, Hadas Saaroni, and Jorge H." Amorim. *The Urban Heat Island: Thermal Comfort and the Role of Urban Greening*, pages 7–19. Springer International Publishing, Cham, 2017. ISBN 978-3-319-50280-9. doi: [10.1007/978-3-319-50280-9\\_2](https://doi.org/10.1007/978-3-319-50280-9_2). URL [https://doi.org/10.1007/978-3-319-50280-9\\_2](https://doi.org/10.1007/978-3-319-50280-9_2).
- [57] Deborah Hilbert, Lara Roman, Andrew Koeser, Jess Vogt, and Natalie Van Doorn. Urban tree mortality: A literature review, 09 2018.
- [58] Sarah E. Hobbie and Nancy B. Grimm. Nature-based approaches to managing climate change impacts in cities. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1794):20190124, jan 2020. doi: [10.1098/rstb.2019.0124](https://doi.org/10.1098/rstb.2019.0124). URL <https://doi.org/10.1098/rstb.2019.0124>.
- [59] Sha Huang, Lina Tang, Joseph Hupy, Yang Wang, and Guofan Shao. A commentary review on the use of normalized difference vegetation index (ndvi) in the era of popular remote sensing. *Journal of Forestry Research*, 32, 05 2020. doi: [10.1007/s11676-020-01155-1](https://doi.org/10.1007/s11676-020-01155-1).
- [60] Miguel Jiménez-Bello, Carlos Ballester, Juan Castel, and Diego Intrigliolo. Development and validation of an automatic thermal imaging process for assessing plant water status. *Agricultural Water Management*, 98:1497–1504, 08 2011. doi: [10.1016/j.agwat.2011.05.002](https://doi.org/10.1016/j.agwat.2011.05.002).
- [61] Michael G Just and Steven D Frank. Evaluation of an easy-to-install, low-cost dendrometer band for citizen-science tree research. *Journal of Forestry*, 117(4):317–322, 2019.
- [62] Emna Kamoun. Image registration: From sift to deep learning, Mar 2021. URL <https://www.sicara.ai/blog/2019-07-16-image-registration-deep-learning>.

- [63] Arnon Karnieli, M. Bayasgalan, Y. Bayarjargal, Nurit Agam, S. Khudulmur, and C. Tucker. Comments on the use of the vegetation health index over mongolia. *International Journal of Remote Sensing - INT J REMOTE SENS*, 27:2017–2024, 05 2006. doi: 10.1080/01431160500121727.
- [64] James Y. Kim and David M. Glenn. Multi-modal sensor system for plant water stress assessment. *Computers and Electronics in Agriculture*, 141:27–34, 2017. ISSN 0168-1699. doi: <https://doi.org/10.1016/j.compag.2017.07.009>. URL <https://www.sciencedirect.com/science/article/pii/S0168169916311073>.
- [65] Coco Yin Tung Kwok, Man Sing Wong, Hon Li, Karena Ka Wai Hui, Florence Wan Yee Ko, Herman Yiu Kay Shiu, and Zihan Kan. Detection of structural tree defects using thermal infrared imaging. 2020. 40th Asian Conference on Remote Sensing: Progress of Remote Sensing Technology for Smart Future, ACRS 2019 ; Conference date: 14-10-2019 Through 18-10-2019.
- [66] Angela Lausch, Stefan Erasmi, Douglas J. King, Paul Magdon, and Marco Heurich. Understanding forest health with remote sensing -part i—a review of spectral traits, processes and remote-sensing characteristics. *Remote Sensing*, 8(12), 2016. ISSN 2072-4292. doi: 10.3390/rs8121029. URL <https://www.mdpi.com/2072-4292/8/12/1029>.
- [67] E. Leong, Daniel Burcham, and Yok Fong. A purposeful classification of tree decay detection tools. *Arboricultural Journal*, 34, 06 2012. doi: 10.1080/03071375.2012.701430.
- [68] Xiaojiang Li, Chuanrong Zhang, Wei Li, Robert M. Ricard, Qingyan Meng, and Weixing Zhang. Assessing street-level urban greenery using google street view and a modified green view index. *Urban Forestry & Urban Greening*, 14:675–685, 2015.
- [69] Xiaojiang Li, Chuanrong Zhang, Weidong Li, Robert Ricard, Qingyan Meng, and Weixing Zhang. Assessing street-level urban greenery using google street view and a modified green view index. *Urban Forestry & Urban Greening*, 14(3):675–685, 2015. ISSN 1618-8667. doi: <https://doi.org/10.1016/j.ufug.2015.06.006>. URL <https://www.sciencedirect.com/science/article/pii/S1618866715000874>.
- [70] Xiaojiang Li, Carlo Ratti, and Ian Seiferling. Quantifying the shade provision of street trees in urban landscape: A case study in boston, usa, using google street view. *Landscape and Urban Planning*, 169:81–91, 2018. ISSN 0169-2046. doi: <https://doi.org/10.1016/j.landurbplan.2017.08.011>. URL <https://www.sciencedirect.com/science/article/pii/S0169204617301950>.
- [71] A Majdak, R Jakus, and M Blazenec. Determination of differences in temperature regimes on healthy and bark-beetle colonised spruce trees using a handheld thermal camera. *iForest - Biogeosciences and Forestry*, (3):203–211, 2021. doi: 10.3832/ifor3531-014. URL <https://iforest.sisef.org/contents/?id=ifor3531-014>.

- [72] Henry Mak and Baoxin Hu. Tree species identification and subsequent health determination from mobile lidar data. pages 1365–1368, 07 2014. doi: 10.1109/IGARSS.2014.6946688.
- [73] E. Gregory McPherson. Center for Urban Forest Research, Pacific Southwest Research Station, USDA Forest Service, 2004.
- [74] Attila Nagy. Thermographic evaluation of water stress in an apple orchard. *Journal of Multidisciplinary Engineering Science and Technology*, 2:2210–2215, 01 2015.
- [75] Sophie A. Nitoslawski, Nadine J. Galle, Cecil C. Konijnendijk van den Bosch, and James W. N. Steenberg. Smarter ecosystems for smarter cities? a review of trends, technologies, and turning points for smart urban forestry. *Sustainable Cities and Society*, 2019.
- [76] Kevin P. O’Keeffe, Amin Anjomshoaa, Steven H. Strogatz, Paolo Santi, and Carlo Ratti. Quantifying the sensing power of vehicle fleets. *Proceedings of the National Academy of Sciences*, 116(26):12752–12757, 2019. doi: 10.1073/pnas.1821667116. URL <https://www.pnas.org/doi/abs/10.1073/pnas.1821667116>.
- [77] Raspberry Pi. raspberry pi 4 model b. URL <https://www.raspberrypi.com/products/raspberry-pi-4-model-b/>.
- [78] Rui Pitarma, João Crisóstomo, and Maria Eduarda Ferreira. Contribution to trees health assessment using infrared thermography. *Agriculture*, 9(8), 2019. doi: 10.3390/agriculture9080171. URL <https://www.mdpi.com/2077-0472/9/8/171>.
- [79] Ilyas Potamitis, Iraklis Rigakis, Nicolaos-Alexandros Tatlas, and Stelios Potirakis. In-vivo vibroacoustic surveillance of trees in the context of the iot. *Sensors*, 19(6), 2019. ISSN 1424-8220. doi: 10.3390/s19061366. URL <https://www.mdpi.com/1424-8220/19/6/1366>.
- [80] Yakir Preisler, Fedor Tatarinov, José M Grünzweig, and Dan Yakir. Seeking the “point of no return” in the sequence of events leading to mortality of mature trees. *Plant, Cell & Environment*, 44(5):1315–1328, 2021.
- [81] Yonghua Qu. *Leaf Area Index: Advances in Ground-Based Measurement*, pages 359–378. Springer Berlin Heidelberg, Berlin, Heidelberg, 2019. ISBN 978-3-662-48297-1. doi: 10.1007/978-3-662-48297-1\_11. URL [https://doi.org/10.1007/978-3-662-48297-1\\_11](https://doi.org/10.1007/978-3-662-48297-1_11).
- [82] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks, Jan 2016. URL <https://arxiv.org/abs/1506.01497>.
- [83] Eva Rocha and Steffen Holzkämper. Assessing urban climate effects on pinus sylvestris with point dendrometers: a case study from stockholm, sweden. *Trees*, pages 1–10, 2021.

- [84] Siti Rosli, Fazida Hashim, Thinal Raj, W Mimi Diyana W Zaki, and Aini Hussain. A rapid technique in evaluating tree health using lidar sensors. *International Journal of Engineering and Technology(UAE)*, 7:118–122, 08 2018. doi: 10.14419/ijet.v7i3.17.16634.
- [85] Ian Seiferling, Nikhil Naik, Carlo Ratti, and Raphael Proulx. Green streets: Quantifying and mapping urban trees with street-level imagery and computer vision. *Landscape and Urban Planning*, 165:93–101, 09 2017. doi: 10.1016/j.landurbplan.2017.05.010.
- [86] Bilal Shabandri, Sahith Reddy Madara, and Piyush Maheshwari. *IoT-Based Smart Tree Management Solution for Green Cities*, pages 181–199. 01 2020. ISBN 978-981-15-0662-8. doi: 10.1007/978-981-15-0663-5\9.
- [87] Magdalena Smigaj, Rachel Gaulton, S. Barr, and Juan Suarez Minguez. Uav-borne thermal imaging for forest health monitoring: Detection of disease-induced canopy temperature increase. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XI-3/w3:349–354, 08 2015. doi: 10.5194/isprsarchives-XL-3-W3-349-2015.
- [88] Magdalena Smigaj, Rachel Gaulton, Juan C. Suarez, and Stuart L. Barr. Use of miniature thermal cameras for detection of physiological stress in conifers. *Remote Sensing*, 9(9), 2017.
- [89] Chiara Torresan, Marta Benito Garzón, Michael O’Grady, Thomas Matthew Robson, Gianni Picchi, Pietro Panzacchi, Enrico Tomelleri, Melanie Smith, John Marshall, Lisa Wingate, Roberto Tognetti, Lindsey E. Rustad, and Dan Kneeshaw. A new generation of sensors and monitoring tools to support climate-smart forestry practices. *Canadian Journal of Forest Research*, 0(0):1–15, 0. doi: 10.1139/cjfr-2020-0295. URL <https://doi.org/10.1139/cjfr-2020-0295>.
- [90] Fagun Vankawala, Amit Ganatra, and Amit Patel. A survey on different image deblurring techniques. *International Journal of Computer Applications*, 116(13):15–18, 2015.
- [91] Zander S. Venter, Norun Hjertager Krog, and David N. Barton. Linking green infrastructure to urban heat and human health risk mitigation in oslo, norway. *Science of The Total Environment*, 709:136193, 2020. ISSN 0048-9697. doi: <https://doi.org/10.1016/j.scitotenv.2019.136193>. URL <https://www.sciencedirect.com/science/article/pii/S0048969719361893>.
- [92] Daniele Vidal and Rui Pitarma. Infrared thermography applied to tree health assessment: A review. *Agriculture*, 9(7), 2019. ISSN 2077-0472. doi: 10.3390/agriculture9070156. URL <https://www.mdpi.com/2077-0472/9/7/156>.
- [93] Renato Vidoni, Raimondo Gallo, Gianluca Ristorto, Giovanni Carabin, Fabrizio Mazzetto, Lorenzo Scalera, and Alessandro Gasparetto. Byelab: An agricultural mobile robot prototype for proximal sensing and precision farming. page V04at05a057, 11 2017. doi: 10.1115/imece2017-71216.

- [94] Hao Wang, Yang Wei, Hongxu Zhu, Yucheng Liu, Chung Kit Wu, and Kim Fung Tsang. Nb-iot based tree health monitoring system. In *2019 IEEE International Conference on Industrial Technology (ICIT)*, pages 1796–1799, 2019. doi: 10.1109/ICIT.2019.8755153.
- [95] Liangju Wang, Yunhong Duan, Libo Zhang, Tanzeel U. Rehman, Dongdong Ma, and Jian Jin. Precise estimation of ndvi with a simple nir sensitive rgb camera and machine learning methods for corn plants. *Sensors*, 20(11), 2020. ISSN 1424-8220. doi: 10.3390/s20113208. URL <https://www.mdpi.com/1424-8220/20/11/3208>.
- [96] Chen Wei, Meng, Wu Zhiwei, Zhang Yong, and Yang Jaemo. Evaluating greenery around streets using baidu panoramic street view images and the panoramic green view index. *Forests*, 10:1109, 12 2019. doi: 10.3390/f10121109.
- [97] Yang Wei, Hao Wang, Kim Fung Tsang, Yucheng Liu, Chung Kit Wu, Hongxu Zhu, Yuk-Tak Chow, and Faan Hei Hung. Proximity environmental feature based tree health assessment scheme using internet of things and machine learning algorithm. *Sensors*, 19(14), 2019. ISSN 1424-8220. doi: 10.3390/s19143115. URL <https://www.mdpi.com/1424-8220/19/14/3115>.
- [98] Zoey Werbin, Leila Heidari, Sarabeth Buckley, Paige Brochu, Lindsey Butler, Catherine Connolly, Lucila Houttuijn Bloemendaal, Tempest McCabe, Tara Miller, and Lucy Hutyra. A tree-planting decision support tool for urban heat island mitigation. 10 2019. doi: 10.1101/821785.
- [99] Wkentar. Wkentar/labelme: Image polygonal annotation with python (polygon, rectangle, circle, line, point and image-level flag annotation). URL <https://github.com/wkentar/labelme>
- [100] Sam Worley. What is optrx? URL <https://www.agleader.com/blog/what-is-optrx/>.
- [101] Chung Kit Wu, Kim Fung Tsang, Yucheng Liu, Hao Wang, Hongxu Zhu, Cheon Hoi Koo, Wai Hin Wan, and Yang Wei. An iot tree health indexing method using heterogeneous neural network. *IEEE Access*, 7:66176–66184, 2019. doi: 10.1109/ACCESS.2019.2918060.
- [102] Jianwei Wu, Wei Yao, and Przemyslaw Polewski. Mapping individual tree species and vitality along urban road corridors with lidar and imaging sensors: Point density versus view perspective. *Remote Sensing*, 10(9), 2018. ISSN 2072-4292. doi: 10.3390/rs10091403. URL <https://www.mdpi.com/2072-4292/10/9/1403>.
- [103] Qingfu Xiao and E. Gregory Mcpherson. Tree health mapping with multispectral remote sensing data at uc davis, california. *Urban Ecosystems*, 8: 349–361, 2005.
- [104] Yonglin Zhang, Shanlin Li, Xiao Fu, and Rencai Dong. Quantification of urban greenery using hemisphere-view panoramas with a green cover index. *Ecosystem Health and Sustainability*, 7(1):1929502, 2021.

doi: 10.1080/20964129.2021.1929502. URL <https://doi.org/10.1080/20964129.2021.1929502>.

- [105] Bolei Zhou, Hang Zhao, Xavier Puig, Tete Xiao, Sanja Fidler, Adela Barriuso, and Antonio Torralba. Semantic understanding of scenes through the ade20k dataset, Oct 2018. URL <https://arxiv.org/abs/1608.05442>.

# Appendix A

## APPENDIX

### A.1 Settings during Mask R-CNN training

The following settings were used while training the custom Mask R-CNN model "head" layers.

- **Batch Size:** 4
- **Number of Epochs:** 10
- **Network backbone:** ResNet101
- **Data Augmentation:** Probability (P)= 0.5. Apply Sequentially:
  - **Flip Left to right:** P = 0.5
  - **Flip Up to Down:** P = 0.5
  - **Guassian Blur:** sigma = 5.0
- **Train-test split:** 36 : 15 images

### A.2 Data Cleaning

Species	Number of Trees	Health Distribution
Red Pine	26	Good: 15 Fair: 7 Poor: 4
Eastern White Pine	14	Good: 5 Fair: 1 Poor: 8

Table A.1: **Distribution of trees after data cleaning**

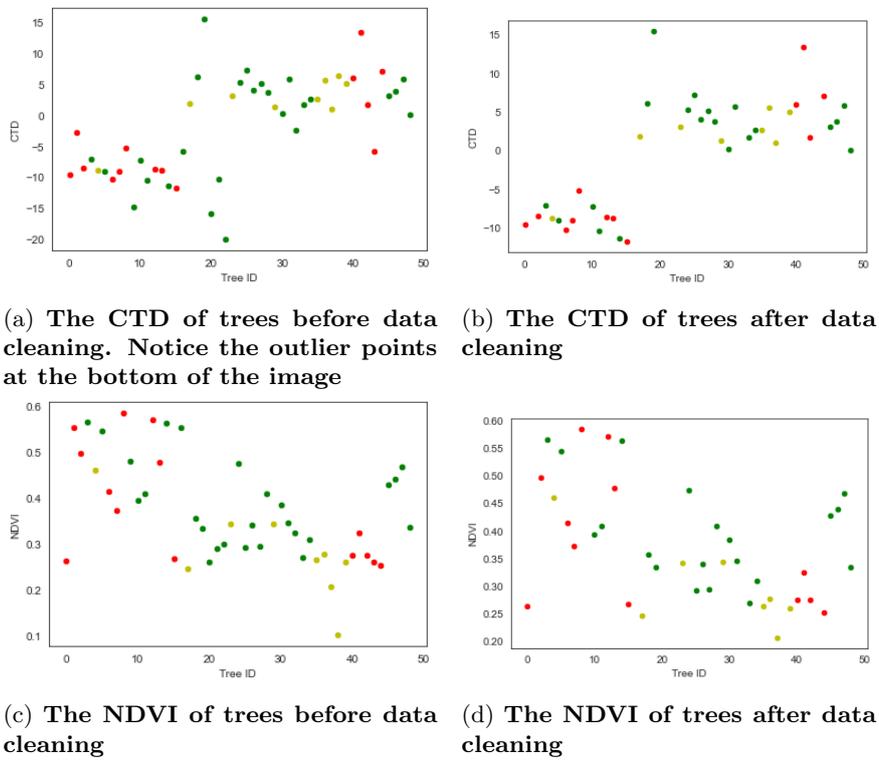


Figure A.1: The CTD and NDVI of trees before and after data cleaning. The color of the points indicate the ground truth health with red denoting poor, yellow denoting fair and green denoting good condition trees.

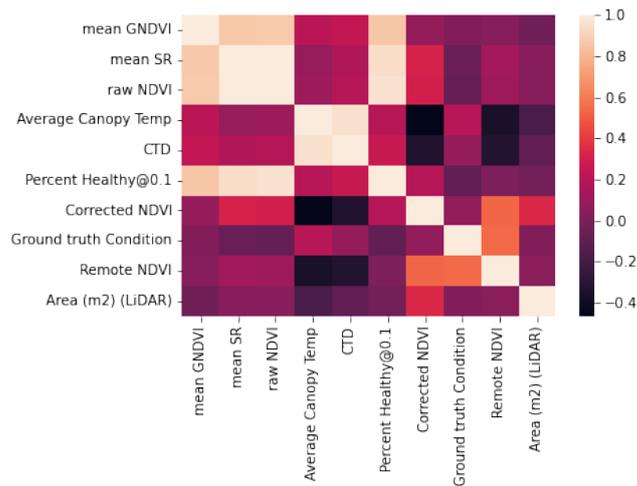


Figure A.2: Correlation between our measured values and ground truth parameters before data clean

### A.3 Additional Results

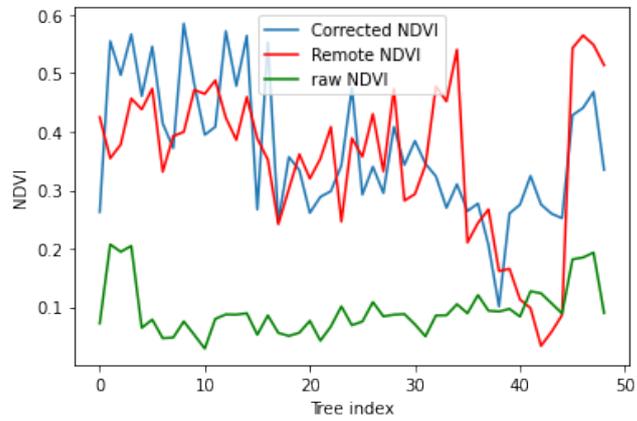


Figure A.3: Variation of raw NDVI measured NDVI and Remote NDVI for trees observed during data collection experiments

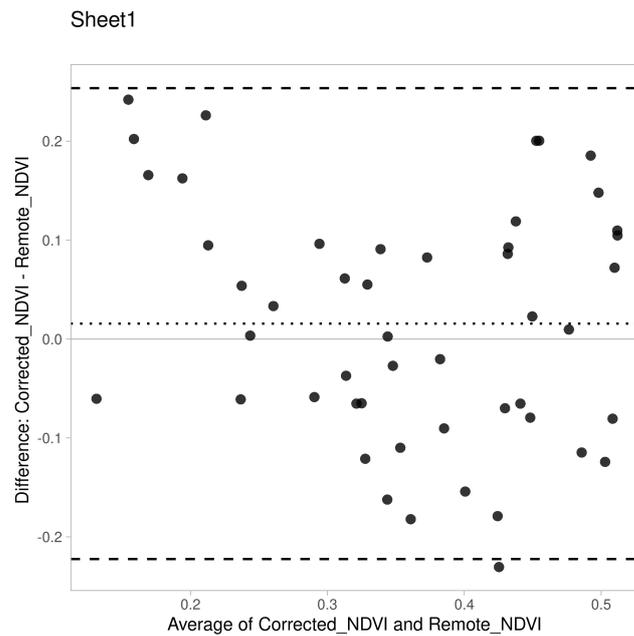


Figure A.4: The Bland-Altman plot showcasing the agreement between Corrected NDVI and Remote NDVI. The dashed-middle line shows the mean difference. The top most and bottom most lines indicate the upper and lower lines of agreement respectively.

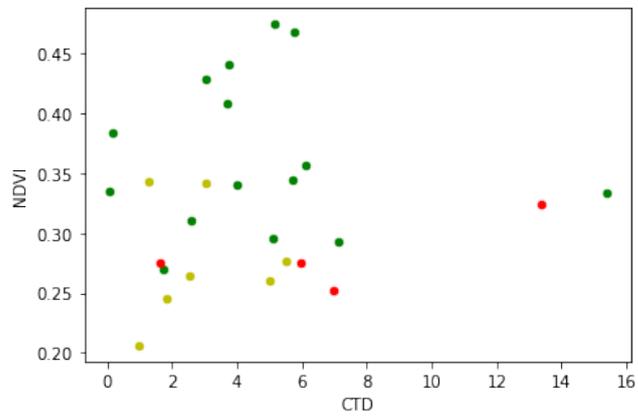


Figure A.5: Scatter plot between NDVI and CTD for red pine trees. The color of the points indicate the ground truth health with red denoting poor, yellow denoting fair and green denoting good condition trees.

**Table 3.** A short summary of infrared thermography applied to tree health analysis.

Study	Main Focus	Findings
Al-doski et al. [72]	Pest detection.	Pest infestation detected.
Ballester et al. [74]	Water stress detection on citrus and persimmon trees.	Water stress detected.
Bellett-Travers & Morris [61]	Relationship between surface temperature and radial wood thickness.	No apparent relationship in most of the trees; strong relationship when there was a gradual change in radial wood thickness caused by a cavity.
Burcham et al. [59]	Effect of mechanically induced internal voids on <i>Dracaena fragrans</i> L. stem temperature.	Only able to identify reductions temperature in internal defects with at least 76% of the stem cross-sectional area.
Burcham et al. [70]	Evaluate the relationship between the internal defects and trunk surface temperature in <i>Casuarina equisetifolia</i> L.	Does not provide accurate results about the internal condition of trees.
Catena & Catena [60]	Review in order to assess the accuracy, reliability, and costs.	Does not automatically distinguish between different kinds of alteration; does not accurately provide the extent of the damages found; provides enough information to decide regarding the need for remedial action or a more detailed kind of assessment; non-invasive, fast, reasonable prices, in real time.
Catena, G. [68,71]	Internal cavities in trees detection.	Enables the detection of cavities.
Catena, A. [79] and Catena et al. [69]	Damages in the roots.	Enables that damages in the roots can be deduced in real time.
Crisóstomo et al. [32]	Considerations over IRT as applied to the state of the tree healthiness.	Enables healthiness evaluation.
Crisóstomo et al. [52]	Quercus pygmentosis tree Analysis.	Evaluating its healthiness status.
García-Tejero et al [76]	Water stress detection in almond trees.	Water stress detected.
Giuliani & Flore [77]	Water stress detection in apple trees.	Water stress detected.
Goh et al. [29]	Review of the current sensing methods used for decay detection in trees.	Comparing methods concerning the fundamental of measurements, hardware implementation, damage caused to the tree and the ease of use.
Hoffmann et al. [73]	Detecting the larval stage of goat moth's larvae in young tree species.	Was not able to detect.
Jones et al. [75]	Water stress detection on grapevine.	Water stress detected.
Leong et al. [31]	Evaluating the current tree decay detection tools.	Classifying the tree decay detection tools in terms of measurement speed, resolution and accuracy.

Figure A.6: Thermal Imaging related works analysed in [92]