IEEE Sensors Council

# GreenScan: Toward Large-Scale Terrestrial Monitoring the Health of Urban Trees Using Mobile Sensing

Akshit Gupta, *Member, IEEE*, Simone Mora, Fan Zhang, Martine Rutten, R. Venkatesha Prasad, and Carlo Ratti

Abstract—Healthy urban greenery is a fundamental asset to mitigate climate change phenomena such as extreme heat and air pollution. However, urban trees are often affected by abiotic and biotic stressors that hamper their functionality, and whenever not timely managed, even their survival. While the current greenery inspection techniques can help in taking effective measures, they often require a high amount of human labor, making frequent assessments infeasible at city-wide scales. In this article, we present GreenScan, a ground-based sensing system designed to provide health assessments of urban trees at high spatio-temporal resolutions, with low costs. The system uses thermal and multispectral imaging sensors fused using a custom computer vision model to estimate two tree health indexes. The evaluation of the system was performed through data collection experiments in Cambridge, USA. Overall, this work illustrates a novel approach for autonomous mobile ground-based tree health monitoring on city-wide scales at high temporal resolutions with low costs.

Index Terms— Drive-by sensing, greenery health, mobile sensing, sensors.
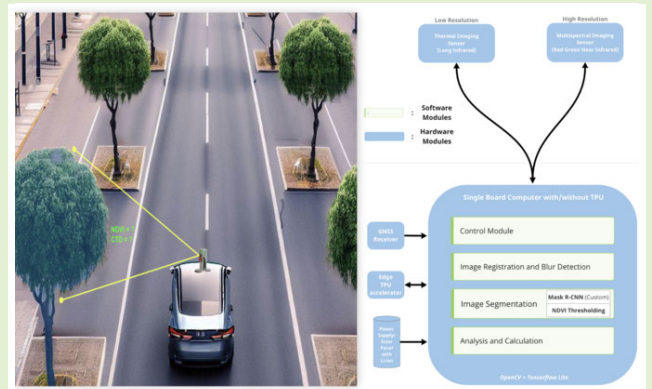
Akshit Gupta was with the MIT Senseable City Laboratory, Cambridge, MA 02139 USA, USA. He is now with the Faculty of Civil Engineering and Geosciences, TU Delft, 2628 CD Delft, The Netherlands, and also with the Faculty of Electrical Engineering Mathematics, and Computer Science, TU Delft, 2628 CD Delft, The Netherlands (e-mail: a.gupta-5@tudelft.nl).

Simone Mora is with the Senseable City Laboratory, Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA 02139 USA, and also with the Department of Computer Science, Norwegian University of Science and Technology, 7491 Trondheim, Norway (e-mail: moras@mit.edu).

Fan Zhang is with the Institute of Remote Sensing and Geographical Information System, School of Earth and Space Sciences, Peking University, Beijing 100871, China (e-mail: fanzhanggis@pku.edu.cn).

Martine Rutten is with the Faculty of Civil Engineering and Geosciences, TU Delft, 2628 CD Delft, The Netherlands (e-mail: m.m.rutten@tudelft.nl).

R. Venkatesha Prasad is with the Faculty of Electrical Engineering Mathematics, and Computer Science, TU Delft, 2628 CD Delft, The Netherlands (e-mail: r.r.venkateshaprasad@tudelft.nl).

Carlo Ratti is with the Senseable City Laboratory, Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA 02139 USA, and also with ABC Department, Politecnico di Milano, 20133 Milan, Italy (e-mail: ratti@mit.edu).

Digital Object Identifier 10.1109/JSEN.2024.3397490

I. INTRODUCTION

URBAN greenery improves the resilience of cities to climate change. Nowadays, protecting, managing, and restoring greenery ecosystems is fundamental for climate-resilient development, given the multiple risks posed to humanity and nature by global warming and climate change as per the latest UN-IPCC report [1]. In cities, tree canopies and vegetation provide a wide range of ecosystem services such as air filtering, carbon sequestration, reduced energy consumption, increased biodiversity, and decreased local temperatures [2], [3]. However, urban trees are experiencing an ample amount of *abiotic* stressors (e.g., soil salinity, heat waves) and *biotic* stressors (caused by living agents such as insects and bacteria) that are exacerbated due to climate change [4], [5], [6]. As a result, their functionality, productivity, and survival are of increasing concern [7]. Trees with poor health cannot provide most of their beneficial ecosystem services [8], [9]. For instance, trees with low transpiration rates do not cool the environment sufficiently and trees with low growth rates have a reduced shading effect. By 2050, it is expected that about two-thirds of urban tree species worldwide will fail to provide the desired climate-positive benefits [10].

The practice of measuring and monitoring urban trees began over a century ago [11]. Today, the health of trees can be monitored using manual inspection by arborists with good

TABLE I
COMPARISON OF SENSING APPROACHES ALONG WITH THE WORKING MECHANISM,
COST, AND QUALITY OF ASSESSMENT TO ANALYZE TREE HEALTH

Approach		Principle	Quality of Assessment	Cost*
Manual inspection		Depends on method	Generally high, varying based on method	$$$$
Embedded sensing		Depends on method	High but lower than manual inspection	$$$
Handheld imaging	Multi/Hyper-spectral imaging	Properties of chlorophyll (photosynthesis) and cell structure	High-quality quantitative value	$$
	Thermal	Cavities, temperature gradient and water stress	Cavities: qualitative value, water stress and temperature gradient: quantitative value	$
	LiDAR	Geometrical parameters such as Leaf Area Index (LAI) and leaf density	Low-quality quantitative value	$ to $$
Street-view based (visible spectrum)		Uses image processing to quantify amount of greenery	No health assessment, only quantity of greenery	$
Remote sensing (multi/hyper-spectral, thermal imaging, LiDAR)		Depends on the type of imaging	Top-level view only	$

(* refers to relative cost where $ is the lowest cost and $$$$ is the highest cost for large-scale evaluation of multiple trees based on the scale in [12])

quality results [12]; yet the high labor cost leads to assessments performed infrequently at very low temporal resolutions such as once every 3–5 years. Technology-assisted monitoring methods can complement manual inspections [13]. However, these methods are impeded by variable data quality, low spatial granularity (remote sensing), or high operational costs (airborne sensing) [14]. Furthermore, most of these methods are unable to quantify the vegetation elements below the tree canopy such as green walls, short trees, or shrubs [15], [16]. All these challenges lead to the lack of urban tree health data in cities and appropriate urban forest management. For instance, adverse health conditions in trees being discovered only after severe damage are already inflicted. Furthermore, from an urban planning perspective, intricate relationships of urban trees with other micro-scale ecosystem services such as air quality improvements and benefits to public health are difficult to quantify. For instance, inappropriate placement of trees in outdoor environments can be detrimental as they can serve to trap air pollutants [17].

Recently, several projects have investigated developing novel alternatives for environmental sensing. For instance, applying artificial intelligence (AI)-based methods on Google Street View (GSV) images to detect the presence of trees [15], [18], or using drive-by strategies to measure air pollution [19], [20] in a cost-efficient way. For instance, it was demonstrated that just ten random taxis could capture data over one-third of streets in Manhattan (New York City) in a single day using drive-by sensing [21]. In addition, citizen-science-based approaches [22] have also been successful in measuring urban environmental parameters [23]. All these methods are set within the domain of opportunistic sensing and are aimed at developing platforms that can be deployed and operated without the need of an expensive or a dedicated infrastructure. Thus, allowing democratic access to even underresourced cities that are affected by climate change in a disproportionate manner [24].

Following on this trend and the critical need for protecting and managing urban forestry, in this work, we develop a novel system, named GreenScan, which measures the health of urban trees on city-wide scales from ground level (terrestrially). The system fuses high-quality data from low-cost thermal and multispectral imaging sensors using custom computer vision

models to generate two complementary tree health indexes, namely, normalized difference vegetation index (NDVI) and canopy temperature depression (CTD), which, respectively, indicate the photosynthetic capacity and water stress levels of a tree. GreenScan was designed both for deployment in citizen-science paradigms by being carried by pedestrians, or in drive-by sensing approaches by being mounted on urban vehicles such as taxis and garbage trucks. Thus, enabling terrestrial urban tree health measurements with high spatial and temporal resolutions at low costs for cities and municipalities around the world.

In this article, we first give a brief overview of the state-of-the-art technology tools and methods to monitor the health of urban greenery with focus on low-cost solutions. We present the design of GreenScan and describe the implementation of the hardware and software components. We evaluate the system with 40 urban trees in uncontrolled outdoor environments and analyze the performance of the system. Finally, we conclude by identifying the immediate future research that can be enabled through large-scale deployment of GreenScan while discerning the limitations of this work.

II. RELATED WORK

Currently, the health of trees is monitored through manual inspection by human experts, remote and airborne sensing through satellites or UAVs, direct installation of embedded sensors on/near the tree, handheld imaging-based sensing, or opportunistic sensing using street view imaging [13]. A comparison of these methods in terms of working mechanism, cost, and quality of assessment is shown concisely in Table I.

Manual inspection involves the work of arborists (human experts) inspecting trees visually, often with the aid of tools such as borers (to extract a wooden core sample from the tree for laboratory analysis) or resistographs (to measure the electrical resistance of the trunk). These methods usually provide a high-quality assessment, but they are time-consuming due to the amount of human labor involved to perform a tree-by-tree assessment. Furthermore, although effective, methods that require drilling and penetration in the living wood may create an entry path for pathogens or may alter the structural integrity of a tree. For a review on these methods, see [12].

Embedded sensing involves the deployment of sensors in the bark of a tree (the outer wooden part of a tree) or in the soil. These sensors can rely on physical, chemical, or electrical phenomenon to detect the presence of parasites, e.g., detecting sudden minimal bark vibrations produced by parasites' locomotion and feeding [25] as well as water uptake and transpiration, e.g., measuring electrical impedance using a pair of electrodes placed in the trunk at opposite positions [26]. These methods generate data at high temporal resolutions with little or no human supervision required; yet at the cost of installing and maintaining one or more sensors per tree. For a review on these methods, see [27].

Imaging-based methods involve the use of optical sensors such as thermal imaging sensors, HMI (hyperspectral or multispectral imaging) sensors, or light detection and ranging (LiDAR). Thermal imaging is based on infrared (IR) radiation emitted from materials and it is mainly used to: 1) measure cavities and physical damages in the living wood [28], [29]; 2) detect infections caused by insects and bacteria [30], [31]; and 3) calculate water stress levels by measuring the temperature of the leaves in the canopy [14]. On the other hand, HMI sensors capture various bands in the electromagnetic spectrum, usually near-IR and parts of the visible spectrum. These captured data are used to calculate various vegetation indexes, the most popular being NDVI. HMI sensors are often used for remote sensing applications [32], although static sensors also exist [33]. Calibration methods are critical to achieve quality results [34]. LiDAR sensors can be used to measure geometrical parameters such as the leaves surrounding a branch and trunk diameter and to estimate the leaf area index (LAI) [35]. However, contradictory studies have been observed on the usage of LiDAR with some works such as [36] claiming no increase in health classification performance with its addition. Usually, LiDAR and HMI sensor approaches are often deployed in tandem in both airborne [35] and ground-based [37] approaches.

Recently, street-view-based methods based on red, green, and blue (RGB) images usually involving the use of Google Street View images have become popular. These methods are used to quantify the presence of urban greenery [15], [38], catalog species [39], and shading effects [40]. While these approaches are cost-effective and scalable, they are only able to quantify the extent of urban greenery at a terrestrial level rather than its health.

When imaging-based sensors are deployed on satellites, airplanes, or unmanned aerial vehicles (UAVs), high spatial coverage can be achieved. However, satellites have a low temporal resolution due to infrequent revisit time and data quality being dependent on the availability of clear skies [14]. Data collection using UAVs and airplanes involves high operational costs and is unsuitable for highly urbanized environments due to aviation regulations. Most importantly, both airborne sensing and satellite imagery can only capture an overhead view of urban tree canopies. As a result, lower vegetation elements such as green walls, short trees, or shrubs are often missed or misinterpreted [15].

For a systematic review of the technological methods and tools for greenery health monitoring, see [13].

TABLE II
Concise Comparison of Our Work With Earlier Works in the Field Measuring Tree Health Terrestrially

Works	Autonomous (No human intervention needed)	Approach	Ground Truth Comparison	Evaluation
[14]	Yes	Mobile (Cars)	No	172 trees (only raw system output with no comparison)
[46] and [43]	Yes	Handheld	Yes	44 images (trees not mentioned)
[42]	Yes	Handheld	No	8 trees
[47] and [44]	Yes	Mobile (Robot)	Yes	2 trees (in controlled lab environment)
[41]	No	Mobile (Cars)	Yes	20 trees
This work	**Yes**	**Mobile (Cars) and Citizen Science**	**Yes**	40 trees

A. Research Gaps and Influence on Design

We seek to provide a scalable system that provides high-quality data with low costs. Comparing the different approaches in Table I, it emerges that ground-based (terrestrial) sensing approaches combined with imaging-based methods can look at vegetation elements in a holistic manner with high-quality data gathered either through drive-by sensing or citizen science paradigms. In addition, the advances in deep-learning-based computer vision models for imaging data in the past decade enable the development of a system that is broad in scope. Narrowing this down, the past studies measuring tree health from ground level and using low-cost imaging sensors are also shown concisely in Table II. These studies are limited by requiring manual analysis of images by humans [41], outputting only raw data without ground-truth validation [14], [42] or requiring controlled system deployment and operation [14], [43], [44].

Our work builds upon all these insights. Hence, in Green-Scan, we use HMI imaging and thermal imaging sensors to autonomously measure two health indexes, namely, NDVI and CTD from ground level. GreenScan is designed to be completely autonomous (by using computer vision model based on deep learning) and is suitable to be deployed in noncontrolled environments, and the early results are compared with a ground-truth dataset provided by a municipality. While data from HMI and thermal imaging can be used to generate a number of health indices, we carefully chose NDVI and CTD indices after explicit considerations. Our choices are driven by: 1) NDVI remains one of the most important and popular indices used in the domain [34] and the ground-truth dataset provided by the municipality contains remote NDVI for a fair scientific comparison and 2) CTD is one of the relatively simpler metrics for assessing properties of a tree such as
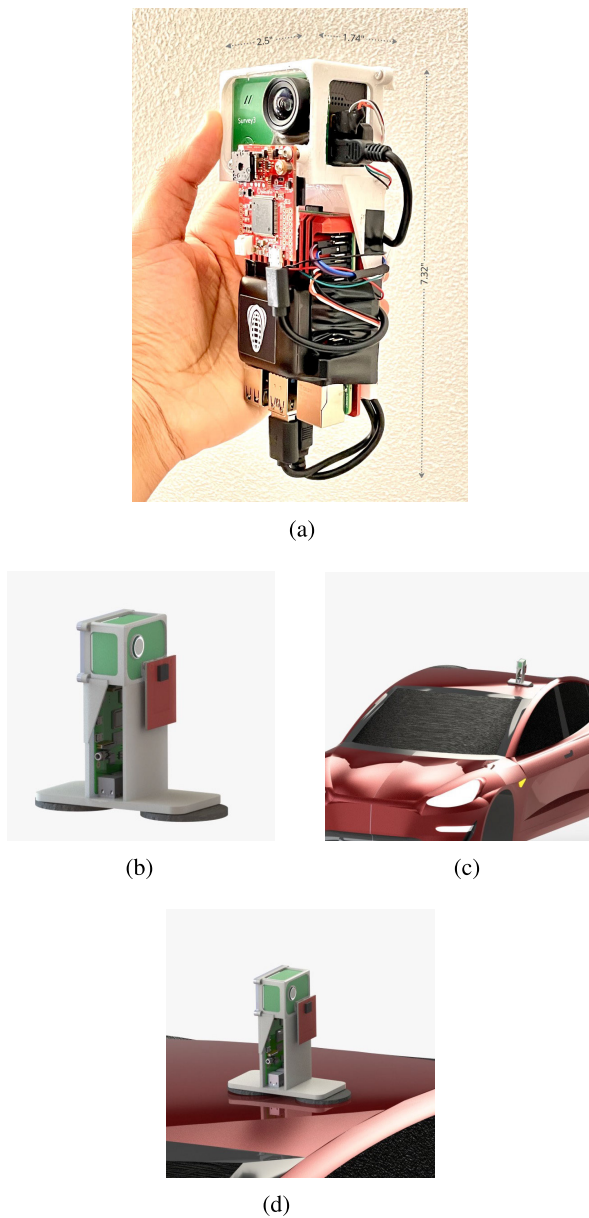
(a)

(b)

(c)

(d)

Fig. 1. Visualization of the current GreenScan system and the concept casing. (a) All the hardware components encased with the 3-D-printed case. Dimensions in inches: 7.32″ × 2.50″ × 1.74″. (b) Concept casing with magnets. (c) System attached on the top of a car. (d) Close-up view of the system on the roof of a car.

water consumption and its resilience to drought and heat stress events [45], and it uses different wavelengths than NDVI (thermal imaging sensors instead of HMI sensors), in turn, generating two complementary health parameters for urban trees.

III. METHODOLOGY

The GreenScan system integrates low-cost thermal and multispectral imaging sensors which are attached to a single-board computer. The system processes the imaging data generated by these sensors using a custom computer vision model to generate the two tree health indexes, namely, CTD and NDVI. All these components were encased in a 3-D-printed case as shown in Fig. 1(a). The case was designed such that it is suitable to be attachable to moving vehicles without any

alterations using magnets, as shown in Fig. 1(b)–(d). In this section, we aim to explain all the major modules of GreenScan.

A. System Architecture

The block diagram of the entire GreenScan system architecture is shown in Fig. 2. The first five modules are related to hardware, while the remaining four modules are related to software.

1) Hardware Modules: In the following, we first provide the generic description of each hardware module followed by the concrete implementation of the same in the GreenScan system.

1) *Thermal Imaging Sensor:* A thermal imaging sensor with radiometric calibration (to measure true temperature) is attached to the central single-board computer and captures long-wave IR images normalized to a suitable temperature range with low pixel resolution. A narrow temperature range is preferred to decrease the effect of nonlinear noise across the sensor as the low-cost thermal imaging sensors are constrained in terms of resolution. These long-wave IR imaging data are used for the generation of CTD which indicates the water stress levels of a tree.

   For concrete implementation, we used FLIR Lepton 3.5 (spectrum: long-wave-IR at 8–14 μm) attached to an OpenMV cam H7 using an FLIR Lepton adapter module. These captured thermal images with a pixel resolution of 160 × 120 which are normalized to a suitable temperature range ($-10°$–$40°$ C). This temperature range was chosen based on the lowest and highest temperature (± 10 °C) of trees found during the data collection experiments (in Section IV). The OpenMV cam H7 communicates with Raspberry Pi via remote procedure call (RPC) over USB.

2) *Multispectral Imaging Sensor:* A multispectral imaging sensor is attached to the central single-board computer and captures red, green, and near-IR (RGN) imaging data with high pixel resolution. The near-IR and red imaging data are used for the generation of the NDVI. Furthermore, these high-resolution imaging data are used for segmentation of the tree canopy from the images using the custom deep learning model as described in the *image segmentation module.*

   For implementation, MAPIR Survey 3W (spectrum: red at 660 nm, green at 550 nm, near-IR at 850 nm) was attached to the Raspberry Pi over USB and captured RGN imaging data with a pixel resolution of 4000 × 3000. To control the MAPIR Survey 3W for triggering the capture and transfer of images, pulsewidth modulation (PWM) signals over the micro-HDMI port of MAPIR Survey 3W are used.

3) *GNSS Receiver:* A global navigation satellite system (GNSS) receiver with support for GPS, GLONASS, and Galileo is used to find the current location of the system and geo-tag all the images of the trees captured.

   For this prototype, the RGN images captured were geo-tagged using the standard GPS adapter available for MAPIR Survey 3W.
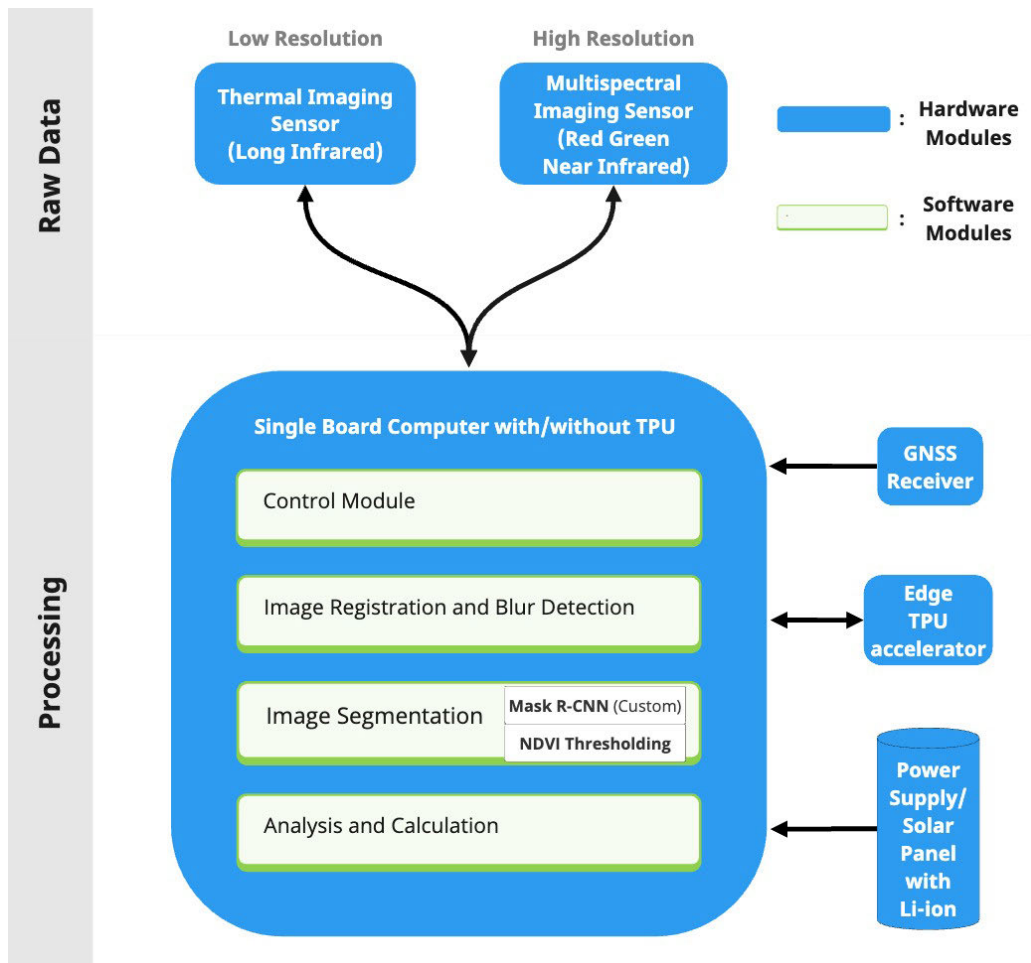
Fig. 2. Architecture diagram of the GreenScan system.

4) *Single-Board Computer With/Without Edge TPU:* A single-board computer without/with onboard edge tensor processing unit (TPU) or USB edge TPU accelerator (such as USB TPU accelerator from coral.ai) acts as the central brain of the system integrating all the hardware and software components. The edge TPU allows to speed up deep learning operations, improving the images processed per second without sending any data to the cloud.

   Raspberry Pi 3 was used as the single-board computer in the GreenScan system running the software modules (see Section III-A2).

5) *Power Supply/Solar Panel:* A lithium-ion battery (10 000 mAh) is used to ensure uninterrupted power supply to the system along with support for charging over a solar panel or a standard power adapter (5 V/2 A).

2) Software Modules: Herewith, we provide the description of each software module followed by the concrete implementation of the same in the GreenScan system. A visualization of processing the images after each software module is also shown in Fig. 3.

1) *Control Module:* This module handles all the embedded communication with the hardware. It includes detecting the event trigger, signaling the sensors to capture the images, and transferring the captured images to the central single-board computer.

   The event trigger signals the beginning of processing on the Raspberry Pi, and in the current prototype, a press of a push button is used as an event trigger for the data collection experiments. This event trigger can also be the co-location of the system with particular GPS coordinates fetched from a tree inventory database. For the thermal imaging sensor, this involves the initiation of callbacks requesting the transfer of the current image frame from FLIR Lepton 3.5. For the multispectral imaging sensor, this involves generating PWM signals to capture an image, mounting the memory card installed in the MAPIR Survey 3W with the Raspberry Pi, transferring the captured image to Raspberry Pi, and finally, unmounting the memory card from the Raspberry Pi.

2) *Image Registration:* Image registration involves matching or aligning images taken by two different sensors into a single coordinate system for further analysis [48]. It includes detecting key points from one image and mapping them to another image. Since both the multispectral and thermal imaging sensors have different field of views (FoVs) and are unaligned, this modules aligns the multispectral images to the thermal images through linear translation in both the horizontal and vertical directions. Besides, to compensate for wider FOV of the
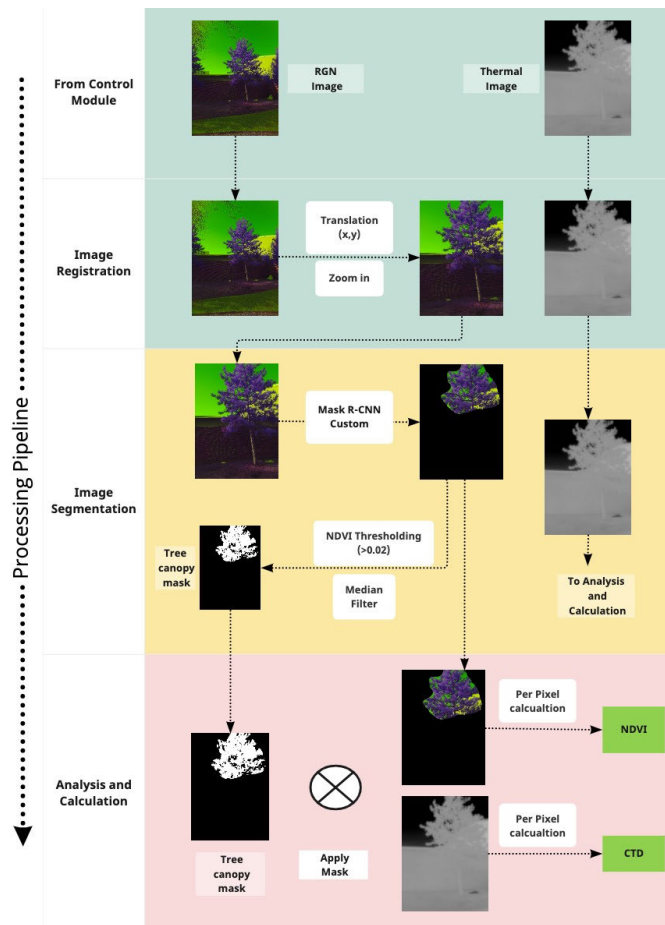
Fig. 3. Visualization of processing the images at each software module on the Raspberry Pi.

multispectral sensor, this module also handles zooming in on the multispectral images.

For the current prototype, the values of translation in th X- and Y-directions were found to be +50 (right) and +150 (upward) pixels, respectively, and the zoom scale was found to be 0.57 (where 1 indicates no magnification and 0 indicates 100 % magnification) to perfectly overlay the thermal and RGN images. These parameters were found by manually taking multiple RGN and thermal images and overlaying them. An instance of inputs and outputs using this module is shown in Fig. 3.

Furthermore, automatic image registration using three image registration algorithms, namely, SIFT, SURF, and ORB [49] was also tested. However, these algorithms were not able to detect useful keypoints or features in the thermal images possibly due to the low resolution (160 × 120).

3) *Image Segmentation:* This is the most computationally intensive software module of the system. Recall that the aim of our system is to calculate the NDVI and CTD values for each tree in the images. However, these values should be calculated only for the leaves in tree canopy excluding the wooden parts, such as the trunk and branches. This is solved using a fusion of custom-developed mask regional convolutional neural network (Mask R-CNN) and pixelwise NDVI analysis. Mask R-CNN [50] is an object detection and instance segmentation model that identifies and then

draws a precise mask around the detected object. Given a multispectral RGN image captured using the multispectral imaging sensor, this task can be broken into two subproblems as follows.

a) *Detect the Canopy Part of the Trees Even in Cases Where the Image Contains Multiple Trees:* This is solved using a custom-developed Mask R-CNN model. The Mask R-CNN model is trained using transfer learning and discussed in more detail in Section III-B. It segments the instances of the tree canopies in the RGN image by generating a mask (segmentation) over them as shown in Fig. 4.

b) *Remove Noise:* Once the canopy of the tree is detected, there is segmentation of only the leaves of the tree without the wooden branches and sky. The nonvegetation elements such as trunks, branches, and sky have very low NDVI values compared with vegetation elements which have significantly higher NDVI. Thus, we use a thresholding-based method which first calculates the individual NDVI of each pixel in the segmentation mask generated by Mask R-CNN and then eliminates pixels with NDVI values below a certain threshold. The calculation of NDVI for each pixel is computed by plugging the raw values of the red and near-IR channels of the pixel in (4). To eliminate noise along the edges of tree canopy, median filtering is also used.

The end result using the above two-stage approach gives segmentation of only leaves present in the tree canopy while eliminating the sky, wooden branches, trunk, and other street objects such as buildings and cars in the multispectral image. Since both the thermal and multispectral images are registered, the same mask of leaves in tree canopy can also be used for thermal images.

With the MAPIR Survey 3W used in the GreenScan system, a value of 0.02 was used as the cutoff to eliminate nonvegetation elements in the image. This value was derived using the analysis of the images captured during data collection experiments. An instance of inputs and outputs using this module is also shown in Fig. 3.

4) *Analysis and Calculation Module:* This module handles the calculation of final NDVI and CTD for each tree in the FoV of the imaging sensors.

The CTD value is computed by calculating the raw temperature value for each pixel in the grayscale thermal image as per (2), computing the mean temperature over all the pixels in the canopy and subtracting the ambient air temperature from the mean canopy temperature as per (3). CTD is calculated as

$$\text{CTD} = T_{\text{canopy}} - T_{\text{air}} \qquad (1)$$

where T_{canopy} and T_{air} are the canopy temperature and air temperature, respectively, in °C.

The temperature of each pixel is calculated as

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

where P_{value} is the pixel value in normalized thermal image, and T_{min} and T_{max} are the configured temperature range for the thermal imaging sensor, respectively ($-10\,°C$ and $40\,°C$ in our case). Then, as per (1), CTD is calculated as

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

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

To calculate the NDVI, each pixel in the RGN image is split into its three constituting channels RGN. The raw NDVI value for each pixel is calculated from red and near-IR channels as per (4). To compensate for the aperture adjustment, the focal adjustment, and other mechanical adjustments performed by the multispectral imaging sensor, the raw NDVI is normalized by applying a correction factor similar to the dynamic range of a camera [51] as shown in (5). NDVI is calculated as

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

where NIR and Red are the values of near-IR channel and visible red channel for each pixel, respectively. The corrected NDVI is calculated as

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

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

Finally, the NDVI for the entire canopy is computed by taking the mean over the corrected NDVI values for all the pixels consisting of leaves in the segmented tree canopy.

B. Development of Custom Mask R-CNN

For the system to operate autonomously, the images will be captured in an unsupervised fashion. Thus, in addition to multiple trees in a single image, they may contain other objects such as cars, buildings, grass, and snow. Hence, it is imperative to individually identify all the tree canopies in an image and feed them to the analysis and calculation module. The custom Mask R-CNN part of the image segmentation module solves this by providing instance segmentation of the tree canopies in the image. To our knowledge, there is no preexisting model available for instance segmentation of tree canopies or even trees in standard RGB images. The problem is further complicated as our input are RGN images from the multispectral imaging sensor instead of standard RGB images. For instance, we found that pretrained models such as Deeplabv3 [52], which can perform semantic segmentation of trees and vegetation on standard RGB images, perform poorly on RGN images.
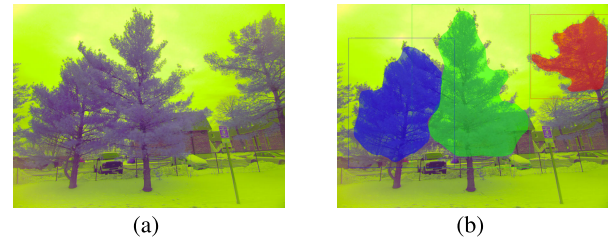

Fig. 4. Performance of our custom Mask R-CNN. Note how the model detects each instance of the tree canopy in the image and considers all the other objects as background. (a) Input RGN image captured from MAPIR Survey 3W. (b) Segmentation output from our custom Mask R-CNN instance segmentation model trained using transfer learning.

1) Training Data: Any deep learning model requires training data to optimize the weights and activations of the layers. However, there does not exist a 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 IV-B). Each tree canopy in the image was manually annotated using a popular image annotation tool called LabelMe [53]. During annotation, only tree canopies that were completely present in the image were labeled. After this process, our dataset consisted of 51 annotated RGN images with two classes, namely, tree canopies and background.

2) Training Process and Training Curve: Our dataset consists of a relatively small number of images to train a deep learning model such as Mask R-CNN from scratch. Transfer learning combined with data augmentation was used to develop a custom model using an existing model pretrained on a different dataset. For this, we used a Mask R-CNN pretrained [54] on COCO [55] (a dataset with 330K images) with ResNet101 as the backbone. We retrained only the head layers (the top layers without the backbone) on our dataset. The batch size was configured as four and the number of epochs was ten. The training was performed on the Google Cloud Platform with an N1 instance with 13-GB memory and 2vCPUs. We also generated synthetic data by augmenting the original dataset with flips in the horizontal and vertical directions and applying Gaussian blur. This increased the training dataset size by 50% and acted as a regularizer. The manually annotated dataset (refer Section III-B1) consisting of 51 images was split in the ratio of 70: 30 for training: testing. During retraining, each epoch took approximately 3 h on the $N1$ instance. The training curve of the model is shown in Fig. 5. It is seen from the training curve that only a small number of epochs are sufficient to reach the optimal validation loss on the test set owing to the retraining of only the head layers. The visual output results from our model are shown in Fig. 4.

3) Model Quantization: Mask R-CNN is a relatively heavy model both from training and inference points 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 into TensorFlow-lite with dynamic range quantization [56]. Dynamic range quantization means that only the weights of the layers in 32-bit FLOAT of the full model are stored as
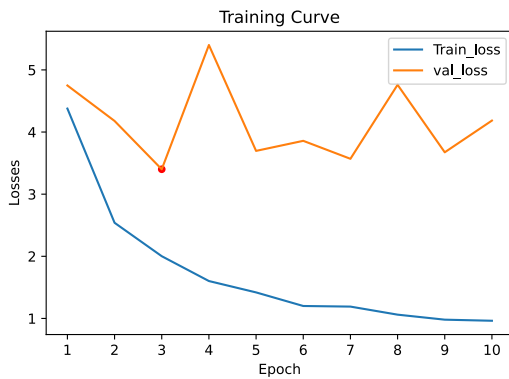
Fig. 5. Training curve of Mask R-CNN with epochs = 10 and batch size = 4. The red point indicates point of minimum loss and training losses are as defined in [50].

8-bit INTs while the activations of the layers are quantized during runtime. Our custom-made Mask R-CNN built over TensorFlow took around 15 s per inference of an image on a Raspberry Pi 3 while the TensorFlow-lite model reduced the inference time to 7 s with one-fourth of the CPU usage as the original TensorFlow model.

IV. EVALUATION

The system was evaluated using a dataset obtained from the municipality of Cambridge, USA, as the ground-truth reference. We also conducted three data collection experiments to collect data of urban trees using the GreenScan system. In this section, we elaborate on this dataset and the data collection experiments followed by the obtained results.

A. (Ground Truth) Tree Health Dataset

Municipalities in cities obtain tree health data through city-wide surveys over years. For instance, in the city of Cambridge, USA, a survey is performed every five years whereas, for the city of Delft, The Netherlands, a survey is performed every one–two years depending on the previously rated condition of the tree. For the evaluation of our work, we obtained the tree health dataset for the city of Cambridge, USA through the Cambridge Urban Forest Master Plan [16] to be used as the ground-truth reference. A 2018 dataset was obtained from the municipality. This dataset was created through a combination of manual in-person arborist visits, satellite-based remote sensing, and aerial LiDAR [16]. The dataset classifies the health conditions of trees into 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 a good health condition. In addition, the dataset contains information about the tree species, common name, the satellite-based NDVI, the latitude and the longitude, location, the shape length and shape area of the canopy, and other parameters. This dataset was provided as Shapefiles [.shp, a data format used by geographical information systems (GISs)] and was loaded to the online platform CARTO [57] (a GIS and spatial analysis tool). On a side note, the staleness of data in terms of time also necessitates the advancement in this field of tree health monitoring.
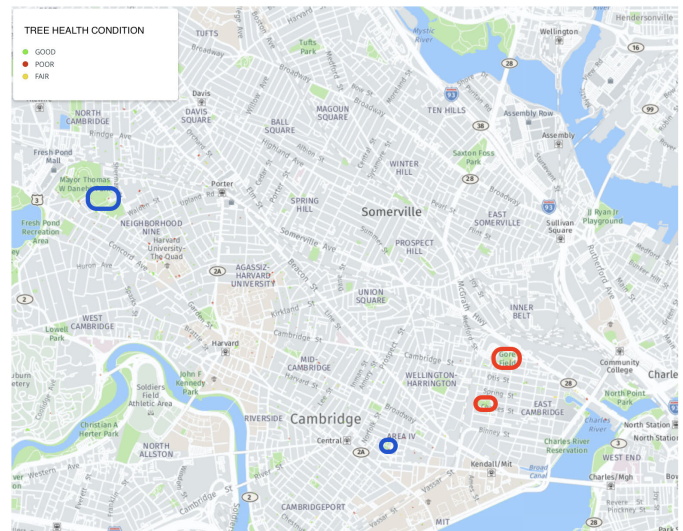

Fig. 6. Trees were analyzed in these locations. The red boxes indicate the red pine trees and the blue boxes indicate the eastern white pine trees.

B. Data Collection Experiments

We collected multispectral (RGN) and thermal images through the 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 GreenScan system as a citizen science project with the 3-D-printed casing (pedestrian moving at walking speed in a straight line at a distance of 8–20 m from the tree). 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 a shutter speed of 1/60 s and ISO of 50. The thermal imaging sensor was configured to measure temperature in range of $(-10, 40)$ °C. The sites of data collection experiments are shown in Fig. 6, chosen based on the species and accessibility.

1) Species Constraints: There are two types of trees, namely, evergreen and deciduous trees. During winters, deciduous trees lose their leaves, thus hampering NDVI calculation. Hence, our analysis was constrained to evergreen trees due to data collection in the winter. 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.

2) Data Cleaning: During the first day of the data collection experiments, 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 values for a sequence of readings related to the canopy temperature by the thermal imaging sensor. As a result, these 11 data points were removed from our dataset generated using data collection experiments. In the end, our dataset was reduced to contain 40 trees. Distribution of the data collected from each of the tree species after data cleaning is shown in Table III.

TABLE III
DISTRIBUTION OF TREES AFTER DATA CLEANING

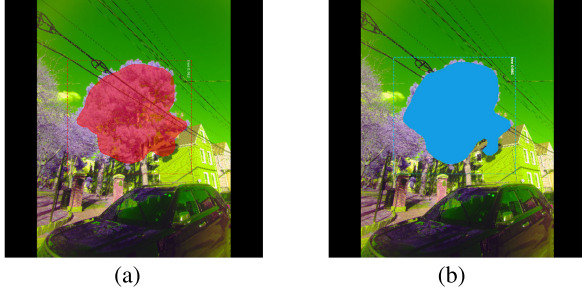| Species	Number of Trees	Health Distribution
Red Pine	26	Good: 15 Fair: 7 Poor: 4
Eastern White Pine	14	Good: 5 Fair: 1 Poor: 8

(a)   (b)

Fig. 7. Outputs from the full and quantized custom Mask R-CNNs. (a) Segmentation output from Mask R-CNN using full TensorFlow model. (b) Segmentation output from Mask R-CNN using TensorFlow-lite (quantized).

C. Performance of Custom Mask R-CNN

To measure the performance of our custom Mask R-CNN model, we calculated the standard evaluation metrics as used by COCO [58]. Specifically, we measured mean average precision (mAP)/average precision (AP) at different intersection over union (IoU) thresholds (as per [58]). The performance of our custom Mask R-CNN with and without quantization is shown in Table IV. A comparison of inference time and model size comparing both the full model and the quantized model is also shown in Table IV. To measure the stability of our results, k-fold cross-validation [59] was also performed with $k = 3$, to evaluate the performance of the model on different training and test splits as shown in Table V. These results for different train–test splits as shown in Table V showcase the reliability of our model.

From Table IV, it is seen that there is no significant reduction in performance using quantization. The inference time of the quantized model is half compared with the nonquantized model along with the reduction of the model size. An example of segmentation outputs generated by the full model and quantized model on the same image is also shown in Fig. 7. This also showcases the similar performance for both the full and optimized models in a visual form.

From Table IV, it may appear that $\text{AP}^{(\text{IoU}=0.5:0.95:0.05)}$ [58] for the quantized model is increased slightly compared with the full model. On further exploring this anomaly, it was found that this behavior is exhibited due to our annotated dataset where most of the images contain only one full tree canopy as ground truth. Thus, a model (nonquantized model) generalizing better to find partially visible tree canopies in addition to the full tree canopy is penalized in terms of precision (false positive). Furthermore, it is seen from Fig. 8 that the performance of the quantized model decreases more than the full model at higher IoUs (IoU = 0.85 for quantized model compared with 0.90 for full model) signifying that it is slightly poorer at object localization compared with the full model.
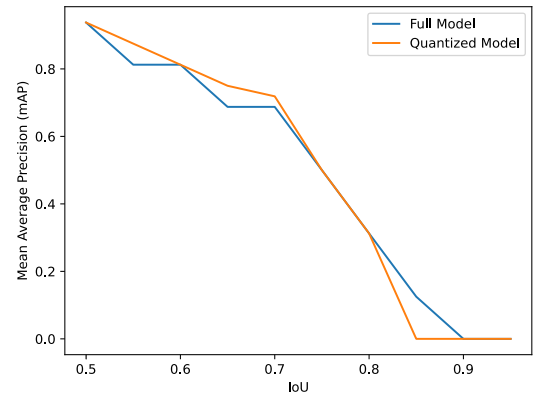

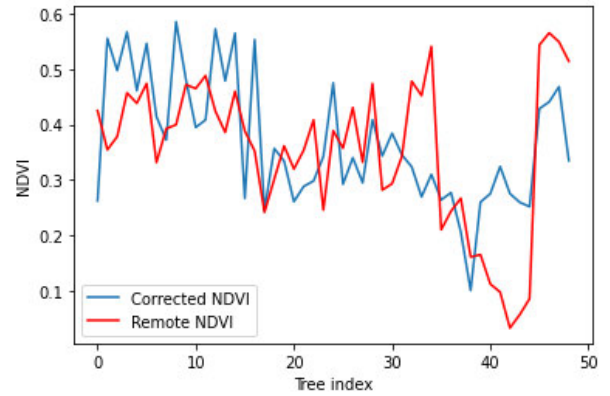Fig. 8. AP scores with increasing IoU thresholds as per COCO metrics [55] for the full and quantized models.

Fig. 9. Variation of measured NDVI versus remote NDVI for trees observed during data collection experiments. The tree index refers to an individual tree ID.

D. 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 the tree (measured using aerial LiDAR) from all the parameters present in the dataset.

A comparison of our system-measured NDVI and remote NDVI is shown in Fig. 9. As seen in this figure, our measured NDVI is distributed similar to the remote NDVI for an individual tree (denoted by tree index in Fig. 9).

Two datasets can be highly correlated but strongly disagree. Hence, a Bland–Altman plot [60] widely used to showcase the agreement between two monitoring methods measuring the same attribute was used. It plots the difference between the corresponding measurement values against the average of those values. From the Bland–Altman plot shown in Fig. 10, it is seen that there is a strong agreement between the two methods (remote NDVI and our measured NDVI) with all the points (representing data for each tree) except one lying within the 95% limits of agreement (average difference within ± 1.96 standard deviation of the difference, as the difference between values follows a normal distribution).

Pearson's correlation coefficient (r) was also measured to calculate the strength of the linear relationship between our measured values and ground-truth data. The correlation matrix comprising all our measured values with the three ground-truth parameters, namely, ground-truth condition, remote NDVI, and

TABLE IV
PERFORMANCE OF CUSTOM R-CNN MODEL (FULL AND QUANTIZED MODELS)

Model	$AP^{(IoU=0.5:0.95:0.05)}$	$AP^{(IoU=0.5)}$	$AP^{(IoU=0.75)}$	$InferenceTime$	$ModelSize$
Custom Mask R-CNN TF	0.489	0.938	0.500	15s	255.9 MB
Custom Mask R-CNN TF-lite (Dynamic Quantization)	0.491	0.938	0.500	7s	65 MB

TABLE V
RESULTS OF THREEFOLD CROSS-VALIDATION
OF CUSTOM MASK R-CNN MODEL

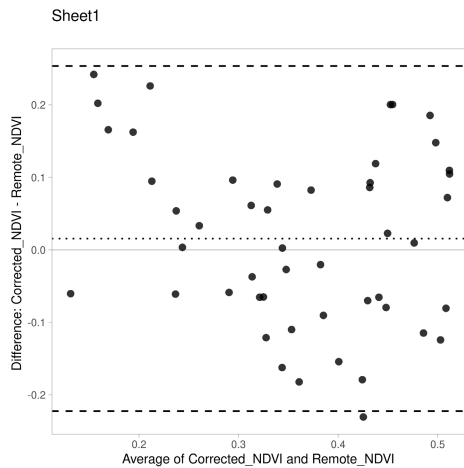| Cross Validation Fold	1	2	3
$AP^{(IoU=0.5)}$	0.82	0.87	0.75

Fig. 10. Bland–Altman plot showcasing the agreement between our measured NDVI and remote NDVI. The dashed-middle line shows the mean difference. The top most and bottom most lines refer to the 95% limits of agreement, respectively.
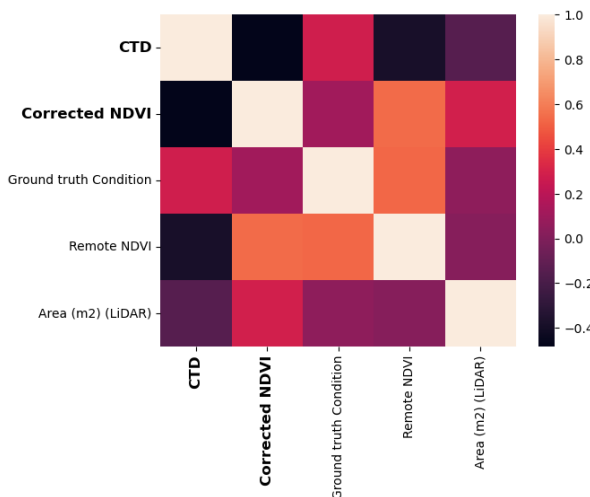

Fig. 11. Correlation matrix between our measured values (in bold) and parameters from the ground-truth dataset.

TABLE VI
CORRELATION BETWEEN OUR MEASURED VALUES
AND GROUND-TRUTH PARAMETERS

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

TABLE VII
MEAN OF MEASURED NDVI AND CTD ACROSS SPECIES AND HEALTH

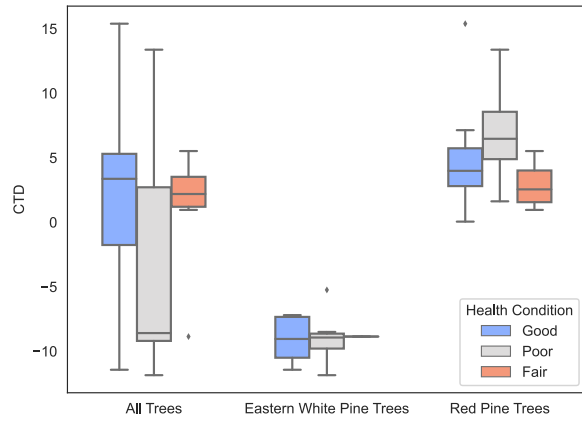| Species / Health	Good	Fair	Poor
Red Pine (NDVI)	0.37 ± 0.07	0.28 ± 0.05	0.28 ± 0.03
Eastern White Pine (NDVI)	0.49 ± 0.08	0.46	0.43 ± 0.12
Red Pine (CTD)	4.63 ± 3.64	2.89 ± 1.78	6.99 ± 4.85
Eastern White Pine (CTD)	-9.1 ± 1.88	-8.59	-9.1 ± 1.88

NDVI distribution in Fig. 12(a), it is seen that the extent of agreement of NDVI with respect to the ground-truth health conditions varies with respect to the species. For red pine trees, the trees in good health condition have higher measured NDVI values than trees in poor and fair condition. A similar conclusion is drawn from the CTD distribution in Fig. 12(b). The mean NDVI and CTD for each species are also shown in Table VII.

1) High-Level Tree Health Analysis: From Fig. 11, 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. In recent works such as [61], the correlation between remote NDVI measured using two different satellites was found to be 0.74 (moderately strong). In this work, Table VI shows moderately strong correlation ($r = 0.54$ with $p < 0.05$) between our measured NDVI (ground-based) and remote NDVI. This moderately strong correlation serves to showcase the validity of our approach for ground-based NDVI measurement using multispectral imaging sensors.

Since there is no ground-truth reference attribute for CTD which indicates water stress of trees, we checked the correlation of CTD with ground-truth health condition as shown in Table VI. The weak–moderate correlation ($r = 0.28$ with $p < 0.05$) between CTD and ground-truth tree health condition can be attributed to the skewed distribution of the dataset where more trees are rated as having good conditions compared with poor and fair conditions. Further analysis of

area is shown in Fig. 11. Furthermore, the correlation results between the measured NDVI and CTD with the ground-truth parameters are shown in Table VI.

The distribution of CTD and NDVI with respect to ground-truth health conditions is shown in Fig. 12. From the
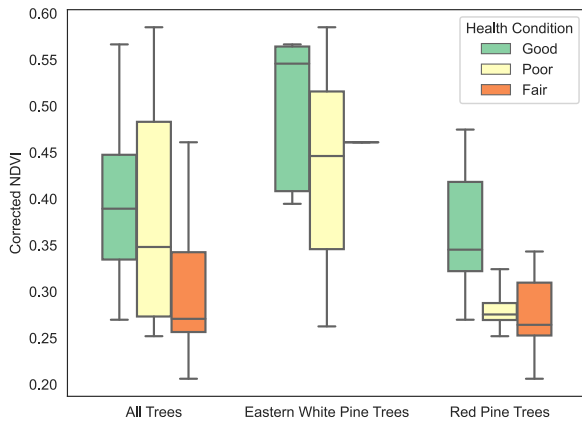
(a)

(b)

Fig. 12. Distribution of NDVI and CTD for the trees with respect to health. (a) Distribution of NDVI for the trees. (b) Distribution of CTD for the trees.

CTD distribution for all the trees in Fig. 12(b) shows the high variability of CTD for trees in poor condition leading to this overall weak–moderate correlation.

2) Specieswise Tree Health Analysis: From the NDVI distributions for eastern white pine in Fig. 12(a), it is seen that the good condition trees are generally distributed to have higher NDVI values than poor and fair condition trees. Thus, a simple threshold-based classification algorithm can easily flag trees which might not specifically be in good health conditions. At a scale of tens of thousands of trees in a city, this can lead to a significant amount of cost savings.

From Fig. 12(b), while a higher CTD is found for red pine trees in poor condition than good and fair health condition trees, the same pattern is not applicable for eastern white pine trees. This inference about CTD is similar to earlier works such as [43] and [46], where the tree species under observation has a significant influence on the results obtained from thermal imaging. Hence, further studies with varied species are required to measure the stability of CTD with respect to ground-truth health conditions.

V. LIMITATIONS AND FUTURE WORK

Subsequent investigations stemming from this work and using our approach as a foundational framework are expected
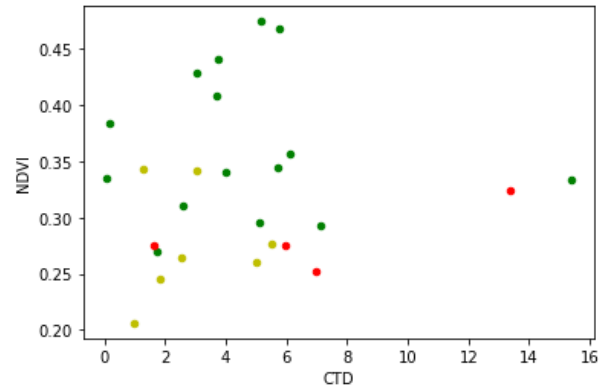

Fig. 13. Scatter plot between NDVI and CTD (in °C) for red pine trees. The color of the points indicates the ground-truth health with red denoting poor, yellow denoting fair, and green denoting good condition trees.

to illuminate and solve several novel challenges, subset of which are delineated as follows.

1) Feasibility of Modeling-Based Classification: From the correlation matrix in Fig. 11, it is seen that there is no correlation between CTD and NDVI values. Hence, in autonomous models to classify tree health, both these measured parameters are useful features. A scatter plot between NDVI and CTD values for red pine trees is shown in Fig. 13. 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 such as support vector machines (SVMs) with kernel [62] or logistic regression classifiers [63] can autonomously distinguish between good, and poor or fair condition trees. Furthermore, the methodology can be expanded by adding human-in-the-loop validation at intermediate steps to enhance the performance of the system.

2) Direction of Movement and Robust Positioning System: At present, our methodology is contingent upon aligning the system consisting the imaging sensors with the trees' orientation. However, given the intended practical application in real-world scenarios, the angle between the tree canopy and the camera's direction could potentially impact the segmentation of tree canopies. This can be improved by adding a simple image selection algorithm such that the majority of image frame is occupied by pixels belonging to a tree canopy. Furthermore, our current positioning method relies on GPS coordinates sourced from the tree survey dataset and the GNSS module within the system, which inherently faces uncertainties in positioning. Thus, an alternative positioning approach using real-time kinematics (RTKs) can enhance the positioning robustness.

3) Scalability in Different Weather Conditions and Geographical Boundaries: First, the effect of different weather conditions with reduction of visibility and sunlight directly facing the imaging sensor lenses needs further exploration. Second, the deployment and validation of the system in cities with different topographies and with geographical domain shifts can help in enhancing the generalization of the approach using large-scale training and validation datasets.

VI. CONCLUSION

Urban greenery provides various environmental services such as carbon sequestration and cooling making them essential for building climate-adaptive cities. Currently, urban trees are experiencing atypical amounts of natural and human-induced stresses leading to their volatile health. Yet, high costs make it infeasible for cities to perform frequent inspections on a large scale, leading to adverse health conditions being discovered only after severe damage. The current popular methods for monitoring the health of urban trees rely on an in-person inspection performed by arborists and remote sensing based on satellites or airborne imagery. 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 called GreenScan to measure tree health autonomously from the ground level in urban cities. The GreenScan system fuses data from low-cost thermal and multispectral imaging sensors using custom computer vision models optimized for efficiency to generate the tree health indexes, namely, NDVI and CTD. The custom Mask R-CNN model fine-tuned using transfer learning was used to fuse the data collected by the imaging sensors on the edge device. Deployment can be performed both in a drive-by sensing paradigm on moving vehicles such as taxis and garbage trucks and in a citizen-science sensing paradigm by humans. Initial evaluation of the system was performed through data collection experiments in Cambridge, USA. The custom Mask R-CNN developed 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 between our measured NDVI and the remote NDVI obtained from the ground-truth dataset. Furthermore, our measured NDVI distributions can be used to flag trees that are specifically not in good health conditions. For the measured CTD, a pattern with a theoretical agreement was applicable for one of the species observed. However, further large-scale evaluation studies over multiple species would help in improving the generalizability of the system. In essence, this work illustrates the potential of autonomous ground-based urban tree health monitoring on city-wide scales at high temporal resolutions and motivates future research at the intersection of environmental science and computer science.

REFERENCES

[1] (2022). *Climate Change 2022: Impacts, Adaptation and Vulnerability.* [Online]. Available: https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/

[2] E. Gregory McPherson, "Accounting for benefits and costs of urban greenspace," *Landscape Urban Planning*, vol. 22, no. 1, pp. 41–51, Sep. 1992. [Online]. Available: https://www.sciencedirect.com/science/article/pii/016920469290006L

[3] S. E. Hobbie and N. B. Grimm, "Nature-based approaches to managing climate change impacts in cities," *Philos. Trans. Roy. Soc. B*, vol. 375, Jan. 2020, Art. no. 20190124, doi: 10.1098/rstb.2019.0124.

[4] N. Grimm et al., "Global change and the ecology of cities," *Science*, vol. 319, no. 5864, pp. 756–760, 2008.

[5] S. A. Nitoslawski, N. J. Galle, C. K. Van Den Bosch, and J. W. N. Steenberg, "Smarter ecosystems for smarter cities? A review of trends, technologies, and turning points for smart urban forestry," *Sustain. Cities Soc.*, vol. 51, Nov. 2019, Art. no. 101770.

[6] IPCC, *Summary for Policymakers*. New York, NY, USA: Cambridge Univ. Press, 2021, pp. 3–32.

[7] Z. R. Werbin et al., "A tree-planting decision support tool for urban heat mitigation," *PLoS ONE*, vol. 15, no. 10, Oct. 2020, Art. no. e0224959, doi: 10.1371/journal.pone.0224959.

[8] K. L. Hand and K. J. Doick. *Understanding the Role of Urban Tree Management on Ecosystem Services*. Accessed: Dec. 21, 2022. [Online]. Available: https://www.forestresearch.gov.uk/research/understanding-role-urban-tree-management-ecosystem-services/

[9] D. Hilbert, L. Roman, A. Koeser, J. Vogt, and N. van Doorn, "Urban tree mortality: A literature review," *Arboriculture Urban Forestry*, vol. 45, no. 5, pp. 167–200, Sep. 2019. [Online]. Available: https://auf.isa-arbor.com/content/45/5/167

[10] K. Huang, "Urban forests facing climate risks," *Nature Climate Change*, vol. 12, no. 10, pp. 893–894, Oct. 2022.

[11] W. Solotaroff, *Shade-Trees in Towns and Cities: Their Selection, Planting, and Care as Applied to the Art of Street Decoration; Their Diseases and Remedies; Their Municipal Control and Supervision*. Hoboken, NJ, USA: Wiley, 1912.

[12] E.-C. Leong, D. C. Burcham, and Y.-K. Fong, "A purposeful classification of tree decay detection tools," *Arboricultural J.*, vol. 34, no. 2, pp. 91–115, Jun. 2012.

[13] A. Gupta, S. Mora, Y. Preisler, F. Duarte, V. Prasad, and C. Ratti, "Tools and methods for monitoring the health of the urban greenery," *Nature Sustainability*, pp. 1–9, Mar. 2024, doi: 10.1038/s41893-024-01295-w.

[14] S. Fuentes, E. Tongson, and C. Gonzalez Viejo, "Urban green infrastructure monitoring using remote sensing from integrated visible and thermal infrared cameras mounted on a moving vehicle," *Sensors*, vol. 21, no. 1, p. 295, Jan. 2021, doi: 10.3390/s21010295.

[15] X. Li, C. Zhang, W. Li, R. Ricard, Q. Meng, and W. Zhang, "Assessing street-level urban greenery using Google street view and a modified green view index," *Urban Forestry Urban Greening*, vol. 14, no. 3, pp. 675–685, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1618866715000874

[16] *Cambridge Urban Forest Master Plan Preliminary Report*. Accessed: Dec. 21, 2022. [Online]. Available: https://www.cambridgema.gov/-/media/Files/publicworksdepartment/urbanforestmasterplan/techncialreportappendix.pdf

[17] N. H. Wong, C. L. Tan, D. D. Kolokotsa, and H. Takebayashi, "Greenery as a mitigation and adaptation strategy to urban heat," *Nature Rev. Earth Environ.*, vol. 2, no. 3, pp. 166–181, Jan. 2021.

[18] I. Seiferling, N. Naik, C. Ratti, and R. Proulx, "Green streets—Quantifying and mapping urban trees with street-level imagery and computer vision," *Landscape Urban Planning*, vol. 165, pp. 93–101, Sep. 2017.

[19] A. Anjomshoaa, F. Duarte, D. Rennings, T. J. Matarazzo, P. de Souza, and C. Ratti, "City scanner: Building and scheduling a mobile sensing platform for smart city services," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4567–4579, Dec. 2018.

[20] S. Mora, A. Anjomshoaa, T. Benson, F. Duarte, and C. Ratti, "Towards large-scale drive-by sensing with multi-purpose city scanner nodes," in *Proc. IEEE 5th World Forum Internet Things (WF-IoT)*, Apr. 2019, pp. 743–748.

[21] K. P. O'Keeffe, A. Anjomshoaa, S. H. Strogatz, P. Santi, and C. Ratti, "Quantifying the sensing power of vehicle fleets," *Proc. Nat. Acad. Sci. USA*, vol. 116, no. 26, pp. 12752–12757, Jun. 2019, doi: 10.1073/pnas.1821667116.

[22] J. Silvertown, "A new dawn for citizen science," *Trends Ecology Evol.*, vol. 24, no. 9, pp. 467–471, Sep. 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016953470900175X

[23] *Assessing Air Quality Through Citizen Science*, Eur. Environ. Agency, Copenhagen, Denmark, Eur. Commission, 2020.

[24] C. Ordóñez and P. N. Duinker, "Assessing the vulnerability of urban forests to climate change," *Environ. Rev.*, vol. 22, no. 3, pp. 311–321, Sep. 2014. [Online]. Available: http://www.jstor.org/stable/envirevi.22.3.311

[25] I. Potamitis, I. Rigakis, N.-A. Tatlas, and S. Potirakis, "In-vivo vibroacoustic surveillance of trees in the context of the IoT," *Sensors*, vol. 19, no. 6, p. 1366, Mar. 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/6/1366

[26] E. Borges et al., "Bioimpedance parameters as indicators of the physiological states of plants in situ," *Int. J. Adv. Life Sci.*, vol. 6, pp. 1–2, Jan. 2014.

[27] C. Torresan et al., "A new generation of sensors and monitoring tools to support climate-smart forestry practices," *Can. J. Forest Res.*, vol. 51, no. 12, pp. 1751–1765, Dec. 2021, doi: 10.1139/cjfr-2020-0295.

[28] A. Catena and G. Catena, "Overview of thermal imaging for tree assessment," *Arboricultural J.*, vol. 30, no. 4, pp. 259–270, Mar. 2008, doi: 10.1080/03071375.2008.9747505.

[29] R. Pitarma, J. Crisóstomo, and M. E. Ferreira, "Contribution to trees health assessment using infrared thermography," *Agriculture*, vol. 9, no. 8, p. 171, Aug. 2019. [Online]. Available: https://www.mdpi.com/2077-0472/9/8/171

[30] M. Smigaj, R. Gaulton, S. L. Barr, and J. C. Suárez, "Uav-borne thermal imaging for forest health monitoring: Detection of disease-induced canopy temperature increase," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 3, pp. 349–354, Aug. 2015, doi: 10.5194/isprsarchives-xl-3-w3-349-2015.

[31] A. Majdák, R. Jakus, and M. Blaženec, "Determination of differences in temperature regimes on healthy and bark-beetle colonised spruce trees using a handheld thermal camera," *iForest-Biogeosciences Forestry*, vol. 14, no. 3, pp. 203–211, Jun. 2021. [Online]. Available: https://iforest.sisef.org/contents/?id=ifor3531-014

[32] A. Lausch, S. Erasmi, D. King, P. Magdon, and M. Heurich, "Understanding forest health with remote sensing—Part I—A review of spectral traits, processes and remote-sensing characteristics," *Remote Sens.*, vol. 8, no. 12, p. 1029, Dec. 2016, doi: 10.3390/rs8121029.

[33] L. Wang, Y. Duan, L. Zhang, T. U. Rehman, D. Ma, and J. Jin, "Precise estimation of NDVI with a simple NIR sensitive RGB camera and machine learning methods for corn plants," *Sensors*, vol. 20, no. 11, p. 3208, Jun. 2020, doi: 10.3390/s20113208.

[34] S. Huang, L. Tang, J. P. Hupy, Y. Wang, and G. Shao, "A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing," *J. Forestry Res.*, vol. 32, no. 1, pp. 1–6, Feb. 2021.

[35] J. Degerickx, D. A. Roberts, J. P. McFadden, M. Hermy, and B. Somers, "Urban tree health assessment using airborne hyperspectral and LiDAR imagery," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 73, pp. 26–38, Dec. 2018, doi: 10.1016/j.jag.2018.05.021.

[36] S. F. Rosli, F. H. Hashim, T. Raj, W. M. D. W. Zaki, and A. Hussain, "A rapid technique in evaluating tree health using LiDAR sensors," *Int. J. Eng. Technol.*, vol. 7, no. 3, pp. 118–122, Aug. 2018.

[37] J. Wu, W. Yao, and P. Polewski, "Mapping individual tree species and vitality along urban road corridors with LiDAR and imaging sensors: Point density versus view perspective," *Remote Sens.*, vol. 10, no. 9, p. 1403, Sep. 2018. [Online]. Available: https://www.mdpi.com/2072-4292/10/9/1403

[38] X. Chen, Q. Meng, D. Hu, L. Zhang, and J. Yang, "Evaluating greenery around streets using Baidu panoramic street view images and the panoramic green view index," *Forests*, vol. 10, no. 12, p. 1109, Dec. 2019.

[39] S. Branson, J. D. Wegner, D. Hall, N. Lang, K. Schindler, and P. Perona, "From Google maps to a fine-grained catalog of street trees," 2019, *arXiv:1910.02675*.

[40] X. Li, C. Ratti, and I. Seiferling, "Quantifying the shade provision of street trees in urban landscape: A case study in Boston, USA, using Google street view," *Landscape Urban Planning*, vol. 169, pp. 81–91, Jan. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169204617301950

[41] J. Y. Kim and D. M. Glenn, "Multi-modal sensor system for plant water stress assessment," *Comput. Electron. Agricult.*, vol. 141, pp. 27–34, Sep. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168169916311073

[42] C. Kwok et al., "Detection of structural tree defects using thermal infrared imaging," in *Proc. 40th Asian Conf. Remote Sens., Prog. Remote Sens. Technol. Smart Future*, 2020.

[43] M. A. Jiménez-Bello, C. Ballester, J. R. Castel, and D. S. Intrigliolo, "Development and validation of an automatic thermal imaging process for assessing plant water status," *Agricult. Water Manag.*, vol. 98, no. 10, pp. 1497–1504, Aug. 2011.

[44] R. Vidoni, "ByeLab: An agricultural mobile robot prototype for proximal sensing and precision farming," in *Proc. ASME Int. Mech. Eng. Congr. Expo.*, Nov. 2017.

[45] S. B. Lepekhov, "Canopy temperature depression for droughtand heat stress tolerance in wheat breeding," *Vavilov J. Genet. Breeding*, vol. 26, no. 2, pp. 196–201, Apr. 2022.

[46] C. Ballester, M. A. Jiménez-Bello, J. R. Castel, and D. S. Intrigliolo, "Usefulness of thermography for plant water stress detection in citrus and persimmon trees," *Agricult. Forest Meteorol.*, vol. 168, pp. 120–129, Jan. 2013.

[47] M. Bietresato et al., "A tracked mobile robotic lab for monitoring the plants volume and health," in *Proc. 12th IEEE/ASME Int. Conf. Mech. Embedded Syst. Appl. (MESA)*, Aug. 2016, pp. 1–6.

[48] E. Kamoun. (Mar. 2021). *Image Registration: From Sift to Deep Learning*. [Online]. Available: https://www.sicara.ai/blog/2019-07-16-image-registration-deep-learning

[49] *Opencv: Feature Detection and Description*. Accessed: Dec. 21, 2023. [Online]. Available: https://docs.opencv.org/3.4/db/d27/tutorial_py_table_of_contents_feature2d.html

[50] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," 2017, *arXiv:1703.06870*.

[51] *Dynamic Range in Digital Photography*. Accessed: Dec. 21, 2023. [Online]. Available: https://www.cambridgeincolour.com/tutorials/dynamic-range.htm

[52] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, *arXiv:1706.05587*.

[53] Wkentaro. *Wkentaro/Labelme: Image Polygonal Annotation With Python (Polygon, Rectangle, Circle, Line, Point and Image-Level Flag Annotation)*. [Online]. Available: https://github.com/wkentaro/labelme

[54] W. Abdulla. (2017). *Mask R-CNN for Object Detection and Instance Segmentation on Keras and TensorFlow*. [Online]. Available: https://github.com/matterport/Mask_RCNN

[55] *Common Objects in Context*. [Online]. Available: https://cocodataset.org/#home

[56] *Post-Training Quantization Tensorflow Lite*. [Online]. Available: https://www.tensorflow.org/lite/performance/post_training_quantization

[57] Carto. *Location Intelligence & GIS For Cloud Natives*. [Online]. Available: https://carto.com/

[58] *Common Objects in Context Evaluation Metrics*. [Online]. Available: https://cocodataset.org/#detection-eval

[59] J. Brownlee. *A Gentle Introduction to K-Fold Cross-Validation*. [Online]. Available: https://machinelearningmastery.com/k-fold-cross-validation/

[60] J. Martin Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 327, no. 8476, pp. 307–310, Feb. 1986. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140673686908378

[61] A. Cogato, V. Pagay, F. Marinello, F. Meggio, P. Grace, and M. D. A. Migliorati, "Assessing the feasibility of using Sentinel-2 imagery to quantify the impact of heatwaves on irrigated vineyards," *Remote Sens.*, vol. 11, no. 23, p. 2869, Dec. 2019.

[62] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, Sep. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231220307153

[63] C.-Y.-J. Peng, K. L. Lee, and G. M. Ingersoll, "An introduction to logistic regression analysis and reporting," *J. Educ. Res.*, vol. 96, no. 1, pp. 3–14, Sep. 2002.

Akshit Gupta (Member, IEEE) received the B.E. degree in electronics and communication engineering from the University of Delhi, New Delhi, India, in 2017, and the M.Sc. degree in embedded systems and the M.Sc. degree in computer science from TU Delft, Delft, The Netherlands, in 2022, where he is currently pursuing the Ph.D. degree.

From 2021 to 2022, he was a Visiting Student Researcher with the Senseable City Laboratory, MIT, Cambridge, MA, USA. His specialization in the domain of optimization and networking aspect of embedded systems and the artificial intelligence domain for computer science enables him to modularize and ideate, based on raw hardware and software from the ground up. His current research interests include the Internet of Things, network science, embedded systems, edge AI, and likes to solve problems in domains of hardware-based sensing and applied machine learning.

Mr. Gupta received annotation entrepreneurship '22 by TU Delft, DJ100 '21 (the 100 sustainable young frontrunners in the Netherlands), and the merit scholarship for academic excellence '14–'17 by the University of Delhi and is the winner of multiple hackathons and tech competitions listed at www.iamakshitgupta.com.

Simone Mora received the M.Sc. degree in computer engineering from UniBg, Bergamo, Italy, in 2009, and the Ph.D. degree in computer science from NTNU, Trondheim, Norway, in 2015.

He is a Research Scientist with the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. During his graduated studies, he has been a Visiting Scholar with City London University, London, U.K., and MIT. In 2018, he co-founded a company that developed an ideation toolkit to tackle the UN's Sustainable Development Goals. He coauthored more than 50 publications in peer-reviewed international journals. He develops methods and tools for rapid design and rapid prototyping of Internet-connected sensors and human–computer interfaces. He does research on novel sensing technologies and their applications for future sustainable cities.

Dr. Mora serves as a Guest Editor for *MDPI Electronics* journal and a Reviewer for the Association of Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE).

Fan Zhang received the B.E. degree in electronic engineering from Beijing Normal University, Beijing, China, in 2012, and the M.Sc. and Ph.D. degrees in geoinformation science from the Chinese University of Hong Kong, Hong Kong, in 2013 and 2017, respectively.

From 2017 to 2022, he was a Senior Research Associate with the Massachusetts Institute of Technology, Cambridge, MA, USA. Since 2022, he worked as an Assistant Professor with the Peking University, Beijing. His research interests include intersection of urban informatics, data-driven approaches for urban studies, and geographic artificial intelligence.

Dr. Zhang received the Geospatial World 50 Rising Stars Award in 2022 and the Global Young Scientist Award in Frontier Science and Technology from WGDC in 2020. He is currently an Associate Editor of *Transactions in Urban Data, Science, and Technology*, and a Guest Editor of *ISPRS Journal of Photogrammetry and Remote Sensing*. He has served as a reviewer for more than 50 SCI journals in GIS and urban studies. He was included in Stanford's list of the world's top 2% scientists in 2022.

Martine Rutten received the Ph.D. degree in remote sensing for water resources management from the Delft University of Technology, Delft, The Netherlands, in 2015.

She is an Associate Professor of Water Management and Climate Adaptation. She has a background in civil engineering with specializations hydraulic engineering with ETH Zürich, Zürich, Switzerland, and water resources management with TU Delft, Delft. She has a broad integrated perspective that is reflected in her research portfolio which spans a wide set of topics, from nexus modeling, to optimal control to citizen science. She has a strong background in transdisciplinary learning for water resources management and vast experience with learning communities and living labs.

R. Venkatesha Prasad received the Ph.D. degree from Indian Institute of Science, Bengaluru, India, in 2004.

He is an Associate Professor with the Embedded and Networked Systems Group, Delft University of Technology, Delft, The Netherlands. He has participated in several European and Dutch projects in the area of the Internet of Things, 60-GHz communications, personal networks, and cognitive radios. He has around 250 publications in the peer-reviewed international journals and conferences. His research interests include tactile Internet, the Internet of Things, 60-GHz MMW networks, and 5G.

Dr. Prasad has served on the editorial board of many IEEE Transactions. He has contributed to several standards, including IEEE P1918.1 and P1906.1. He is a Senior Member of ACM.

Carlo Ratti received the joint master's degree from the Politecnico di Turin, Turin, Italy, and the École Nationale des Ponts et Chaussées, Champs-sur-Marne, France, and the M.Phil. and Ph.D. degrees from the University of Cambridge, Cambridge, U.K., in 1996 and 2001, respectively.

He is currently the Founder and the Director of the MIT Senseable City Laboratory, Cambridge, MA, USA. An architect, and an Engineer by training, he practices in Italy, and teaches with the Massachusetts Institute of Technology, Cambridge, MA, USA. He has coauthored more than 200 publications and holds several patents. His work has been exhibited worldwide at venues, such as the Venice Biennale, Venice, Italy; the Design Museum Barcelona, Barcelona, Spain; the Science Museum, London, U.K.; GAFTA, San Francisco, CA, USA; and The Museum of Modern Art, New York, NY, USA. His Digital Water Pavilion at the 2008 World Expo was hailed by Time Magazine as one of the "Best Inventions of the Year." He has been included in Esquire Magazine's *Best and Brightest* list, Blueprint Magazine's *25 People Who Will Change the World of Design*, and Forbes Magazine's *Names You Need To Know* in 2011. He was a Presenter at TED 2011.

Dr. Ratti is serving as a member for the World Economic Forum Global Agenda Council for Urban Management.