TOWARDS AN AUTOMATED APPROACH FOR MONITORING VEGETATION GREEN-UP DYNAMICS USING VEHICLE DASHCAMS IN URBAN ENVIRONMENTS

Sally Jane Crudge

Nottingham Geospatial Institute, University of Nottingham, UK

Dissertation submitted to The University of Nottingham in partial fulfilment of the degree of Master of Research in Geospatial Systems

August 2020

ABSTRACT:

Phenological events are highly sensitive to climatic variation, and temporal phenological shifts have significant impact on ecosystem function. Vegetation in urban environments holds significant value in providing ecosystem services, of which will become increasingly important as urban populations grow. Insights into vegetation phenological transitions have typically long been monitored through satellite imaging analysis and ground-based field measurements, but these methods are limited by financial costs and coarse resolutions, both spatially and temporally. Despite an increase in the growth of fixed digital camera networks for monitoring vegetation phenology, there still exists a data gap in urban settings. Findings of this study showcased that time series imagery of street level trees in urban environments is obtainable from vehicle dashcams. The YOLOv3 deep learning algorithm demonstrated suitability for automating stages of processing towards deriving a greenness metric. However, further work is required to determine an optimum sized detector training dataset, which also proportionally represents trees across the phenological cycle. Questions remain as to how error caused by scene illuminance variation can be mitigated and as to how full automation from raw data to the final green-up metric can be reached.

ACKNOWLEDGMENTS

I would like to express my gratefulness to EPSRC and the Centre for Doctoral Training in Geospatial Systems for the opportunity to purse postgraduate research under a fully funded studentship.

I would like to thank my supervisors, Professor Doreen Boyd and Professor Giles Foody, for their direction and valuable insight throughout this dissertation.

Finally, I would like to thank my parents Jane and Stephen Crudge, for their continued support and for letting me turn the dining room table into a desk during this period of home working. I would also like to thank them, along with my youngest brother Robert, for collecting all the dashcam imagery on 'The Loop', which has been used in this study.

TABLE OF CONTENTS

ABSTRACT: 2
ACKNOWLEDGMENTS 3
TABLE OF CONTENTS 4
ABBREVIATIONS 6
LIST OF FIGURES 7
LIST OF TABLES 8
1. Introduction 9
2. Literature Review 10
2.1 Green Infrastructure and Ecosystem Services 10
2.2 Remote Sensing Vegetation Phenology 11
2.2.1 Pixel Based Analysis: 11
2.2.2 Object Based Analysis: 12
2.3 Deep Learning for Vegetation Detection 13
2.4 The Urban Environment: Citizen Science 14
3. Aims and Objectives 16
Aim: 16
Objectives 16
4. Methodology 17
4.1 Dashcam Imagery 17
4.2 Data Capture 17
4.3 Survey Location 17
4.4 Video Pre-Processing 18
4.5 Deep Learning 19
4.5.1 Data Annotations 19
4.5.2 YOLOv3 Implementation 20
4.5.3 Training 20
4.5.4 Inference 20
4.6 Training Data Reduction 21
4.7 Image Processing 21
4.8 Greenness Plotting 21
4.9 Evaluation Metrics 22
4.10 Field Survey 22

4.11 ArcGIS Collector App 22
4.11.1 Data Collection 23
4.11.2 Data Download 23
4.12 Nature's Calendar 23
4.13 Data Comparisons: Google Street View 23
5. Results 24
5.1 Object Detection Evaluation Metrics 24
5.2 Image Processing and Deriving Greenness 27
5.3 ESRI Collector App – Field Mapping 33
5.4 Nature's Calendar 34
5.5 Data Comparisons: Google Street View 35
6. Discussion 38
6.1 Vehicle Dashcams: capturing roadside vegetation imagery 38
6.2 YOLOv3: Detecting roadside trees at different phenological stages 39
6.3 Cloud Computing: Google Colab 40
6.4 Deriving Greenness 40
6.5 Citizen Science: volunteered dashcam footage for vegetation monitoring 41
6.6 Applications 42
6.7 Further Research Scope 43
7. Conclusion 44
8. REFERENCES 45
9. APPENDIX 51
9.1 Project Management Plan 51
9.2 Data Management Plan 52

ABBREVIATIONS

NOAA = National Oceanic and Atmospheric Administration

AVHRR = Atmospheric Administration's Advanced Very High-Resolution Radiometer

NASA = National Aeronautics and Space Administration

NDVI = Normalised Difference Vegetation Index

HLS = Harmonised Landsat Sentinel

LiDAR = Light Detection and Ranging

UN = United Nations

SDG = Sustainable Development Goals

CNN = Convolutional Neural Network

UAV = Unmanned Aerial Vehicle

YOLO = You Only Look Once

VOTT = Visual Object Tagging Tool

RGB = Red Green Blue

HSV = Hue Saturation Value

2D = Two Dimension

3D = Third Dimension

GPU = Graphics Processing Unit

CSV = Comma Separated Values

AOI = Area of Interest

ROI = Region of Interest

DOY = Day of Year

ExG = Excess Green

Gcc = Green Chromatic Coordinate

IoU = Intersection Over Union

mAP = Mean Average Precision

TP = True Positive

FP = False Positive

FN = False Negative

OSGB = Ordnance Survey Great Britain

API = Application Programming Interface

LIST OF FIGURES

Figure 1. Ecosystem Services from Urban Green Infrastructure 10
Figure 2. Map of Survey Transect Route 18
Figure 3. Example of the manually drawn bounding box in VOTT. 20
Figure 4. Example comparison of detector outputs between leaf off and leaf on. 24
Figure 5. Example of the original detector results (left) and the detector results with a reduced training size detector (right) at AOI_4 on DOY 121. 24
Figure 6. Average precision for the original YOLOv3 object detector, where IoU is 50%. 25
Figure 7. Average precision for the original YOLOv3 object detector, where IoU is 75%. 25
Figure 8. Average precision for the reduced training YOLOv3 object detector, where IoU is 50%. 26
Figure 9. Average precision for the reduced trainingYOLOv3 object detector, where IoU is 75%. 26
Figure 10. AOI examples of clipped RGB images, the converted HSV image, and the ROI. 27
Figure 11. The Green Chromatic Coordinates plotted for AOI_2, calculated using the RGB. 28
Figure 12. The Green Chromatic Coordinates plotted for AOI_2, calculated using the HSV. 28
Figure 13. The Green Chromatic Coordinates plotted for AOI_2, calculated using the ROI HSV. 28
Figure 14. The Green Chromatic Coordinates plotted for AOI_3, calculated using the RGB. 29
Figure 15. The Green Chromatic Coordinates plotted for AOI_3, calculated using the HSV. 29
Figure 16. The Green Chromatic Coordinates plotted for AOI_3, calculated using the ROI HSV. 29
Figure 17. The Green Chromatic Coordinates plotted for AOI_4, calculated using the RGB. 30
Figure 18. The Green Chromatic Coordinates plotted for AOI_4, calculated using the HSV. 30
Figure 19. The Green Chromatic Coordinates plotted for AOI_4, calculated using the ROI HSV. 30
Figure 20. The Green Chromatic Coordinates plotted for AOI_5a, calculated using the RGB. 31
Figure 21. The Green Chromatic Coordinates plotted for AOI_5a, calculated using the HSV. 31
Figure 22. The Green Chromatic Coordinates plotted for AOI_5a, calculated using the ROI HSV. 31
Figure 23. The Green Chromatic Coordinates plotted for AOI_5b, calculated using the RGB. 32
Figure 24. The Green Chromatic Coordinates plotted for AOI_5b, calculated using the HSV. 32
Figure 25. The Green Chromatic Coordinates plotted for AOI_5b, calculated using the ROI HSV. 32
Figure 26. A map to locate each AOI in relation to the survey transect. 34
Figure 27. AOI_2 in Google Street View. 36
Figure 28. AOI_3 in Google Street View. 36
Figure 29. AOI_4 in Google Street View. 37
Figure 30. AOI_5 in Google Street View. 37

LIST OF TABLES

Table 1. Descriptions of Area of Interests 18
Table 2. A sample of training images, taken periodically across the capture period from the test images of the six areas of interest, with the aim of representing the vegetation at various growth stages. 19
Table 3. Start of green-up date for each AOI. 33
Table 4. First leaf records from the 2020 spring green-up, obtained from the Woodland Trust's Nature's Calendar Citizen Science Project. 35

1. INTRODUCTION

Phenology is defined as the study of the timing of recurring biological events, the causes of their timing regarding biotic and abiotic forces, and the interrelation among phases of the same or different species (Leith 1974). Phenological events in nature are highly sensitive to climatic variations, and subsequently the shifts in phenological timings have a significant impact on ecosystem functions (Richardson et al., 2013). In the United Kingdom, evidence suggests that just a 1-degree change in temperature can change the onset of vegetation budburst by 4 to 5 days, although variable between species (Abernethy et al. 2016). Phenological transitions have been recognised in influencing ecosystem productivity (Richardson et al., 2010), and thus the provision of ecosystem services. Nature in urban areas has long been accepted as providing ecosystem services which are consumed directly by humans, and as a necessity for sustaining urban ecology (Bolund and Hunhammer, 1999, Mullaney et al., 2015).

Phenological patterns have long been monitored though field observations, however towards the latter half of the last century, there was a shift to satellite remote sensing methods for monitoring. The relative coarse spatial resolution of many available satellite remote sensing datasets results in pixel scale processing, eliminating information on induvial trees because of the mixed pixel problem. This approach is also disadvantaged by the general coarse temporal resolution of many sensors, a factor of significant importance when considering the rate of phenological transitions in nature. However, there are number of high spatial and temporal commercial satellites on the market that remove such issues. A common, cost effective alternative to satellite approaches is the use of fixed digital cameras for phenological monitoring, which has allowed for the transition to object-based (individual tree) analysis. The novelty of dashcam imagery enlists a new opportunity for analysis, that of a different perspective to many existing object-based analyses such as aerial photography, by offering a ground (road) level view. There is potential of obtaining dashcam data as a crowd-sourced sensor-network, deriving Big Data to explore the phenological patterns of urban vegetation on a national scale.

In this study, a time series of digital imagery was collected using a commercially available dashcam mounted to the interior of a car. Imagery has been used to interpret the timing of phenological greening in urban trees. The YOLOv3 (You Only Look Once Version Three) object detector algorithm has been incorporated into the methodology and evaluated as a possible technique for automating the process in attaining the Green Chromatic Coordinate, as a metric for greenness. Challenges of dashcam imagery for phenology monitoring have been taken into account, and image processing techniques have been experimented with in an attempt to reduce the effects of variation in scene illuminance. On reflection of results obtained, the potential of sourcing volunteered dashcam imagery from Citizen Scientists has been considered, to work in the direction of Big Data to expand geospatial coverage.

2. LITERATURE REVIEW

2.1 Green Infrastructure and Ecosystem Services

Ecosystem services are defined as the direct and indirect benefits human populations derive from ecosystem function (Costanza et al., 1997). Urban ecosystem services contribute to the quality of life for urban citizens; although, these services are site specific, with their value significantly varying across urban zones around the globe (Bolund and Hunhammer, 1999). Bolund and Hunhammer (1999) identified seven different urban ecosystems which are natural by definition, however they have been manipulated and manged by man in such an urban setting. These ecosystems are lakes, streams, lawns/parks, cultivated land, wetlands, street trees, and urban forests. These specifically urban ecosystems are often referred to as blue and green areas. Commonly researched examples of the ecosystem services provided by trees and urban forests include (Figure 1), lowering air temperature subsequently reducing smog and improving air quality (Akbari et al., 2001), carbon sequestration (Nowak and Crane, 2002), rainfall interception and runoff reduction, and water treatment (Xiao and McPherson, 2002). Urban trees can also be a disservice to urban settings, producing allergens, leaf litter and obstructing views (Escobedo et al., 2011), nonetheless, perhaps menial when compared to the many hazards produced by the urban setting itself.



Figure 1. Ecosystem Services from Urban Green Infrastructure

With 68% of the global population projected to live in urban areas by 2050 (UN DESA, 2019), there will be an increased dependence on ecosystem services and strengthening value of trees as urban infrastructure playing a significant role in the United Nation's Sustainable Development Goals (SDGs). A network analysis with the aim of distilling the role of ecosystem services in the SDGs (Wood et al., 2018), showed that SDGs targets rarely are dependent on a single ecosystem service, and that most services contribute to two or more goals. This study drew upon the opportunity of synergies for policymakers to act on, between the poverty-agriculture-water-nature nexus, highlighting the potential of ecosystem services contributing to food provision, water provision, and habitat and biodiversity maintenance. These specific factors were linked to the targets in within four SDGs: SDG1, SDG2, SDG6, and SDG15. Additionally, another key result highlighted the role of ecosystem services in carbon storage and sequestration contributing to targets across seven or more of the SDGs. This research

concluded the importance of managing and enhancing key ecosystem services for achieving multiple SDG targets.

Phenological transitions have been thought as an influencing factor of productivity and gross photosynthesis in ecosystems (Richardson et al., 2013), and subsequently, the provision of ecosystem services. This study's findings suggested that spring productivity anomalies have effects which carry over to the following autumn, subsequently causing productivity anomalies. There is suggestion that these results could be from phenological anomalies, however this is not clear. Another study has produced results challenging existing experiments investigating plant phenological responses to climate change. For instance, observational data suggests species that flower earliest in spring as having the highest temperature sensitivities, which is not what was shown in the study's experimental data (Wolkovich et al., 2012). Research by Dallimer et al., (2016) suggested that the end of season for vegetation in cities was consistently later compared to surrounding rural zones, which could have several benefits for urban areas through ecosystem services. However, the reasons for this later end of season are complex, and differing species between urban zones are surrounding rural areas could play a significant factor to this. Nonetheless, a greater comprehension of phenological variations in urban environments is required. Such knowledge could improve the understanding of climate and vegetation interactions, subsequently aiding with the identification of shortfalls or surfeits in ecosystem services.

2.2 Remote Sensing Vegetation Phenology

2.2.1 Pixel Based Analysis:

In advance of traditional field-based monitoring, remote sensing techniques have been applied for monitoring phenological events. Research which now spans decades has successfully monitored phenological changes in Earth vegetation dynamics using satellite derived imagery. One of the preliminary studies by Justice et al., (1986) utilised the National Oceanic and Atmospheric Administration's (NOAA) Atmospheric Administration's Advanced Very High-Resolution Radiometer (AVHRR) sensor to examine phenology, developing a Normalised Difference Vegetation Index (NDVI) thresholding technique. Reed et al., (1994) staged a similar study across the United States, using the same sensor, for a 24-month period, 1990 to 1991. Conclusions cited the potential of high-resolution sensors such as the Landsat Thematic Mapper or the French Satellite Pour l'Observation de la Terre to give better insight into regional phenological variability. A comparison study by Schwartz and Reed (1999) examined the start of season data as inferred on the ground, to of that suggested by NDVI from NOAA's AVHRR. Their results suggested that the satellite sensors detected the start of season in understorey vegetation rather than dominant tree species. The biweekly, temporal resolution of the satellite also limited comparisons with ground data. In addition to the optical portion, the microwave portion of the electromagnetic spectrum has also been used in phenological studies. For example, the radar backscatter measurements from a scatterometer on board NASA's Quick Scatterometer which has responded to canopy dynamics throughout the growing season, was similar to trends observed by the Moderate Resolution Imaging Spectroradiometer (Frolking et al., 2006).

Developments in algorithms for the multispectral sensors with improved spatial resolution and increased revisit frequency has seen substantial usage for analysis across large areas of land. With a spatial resolution of 30 metres, the Landsat Thematic Mapper is an improvement on NOAA sensors, despite this, its revisit period of 16 days is insufficient for operationally monitoring land surface phenology. The Sentinel-2 constellation has reduced this problem,

with NASAs Harmonised Landsat Sentinel (HLS) product, which produces imagery sub-weekly at 30 metres. This product has proven effective in land surface phenological studies (Cai et al., 2018, Pastick et al., 2020); although refinements in the HLS are required to address issues with cloud cover and correcting the bidirectional reflectance distribution function (Bolton et al., 2020). Nonetheless, the mixed-pixel problem due to sensor resolutions limit the scope of analysis when using many satellite remote sensing sensors.

2.2.2 Object Based Analysis:

A number of high spatial and high temporal, commercial earth observation satellites have emerged as part of a new market in recent years. Planet Labs Inc commercially operate a constellation of more than 150 satellites, capturing the entire Earth's landmass every day at sub 5 metre resolution, offering scope for object-based analysis in an aerial view. Similarly, DigitalGlobe also have a number of commercial satellites in orbit, including their WorldView and QuickBird satellites. WorldView-3 has been used for tree classification; senescence phenology variables were particularly valuable in classifying at the species and genus level (Fang et al., 2020). Yan et al., (2018) integrated vegetation phenology with few high-resolution multi-temporal WorldView-2 images, finding it to be more effective than using more high-frequency lower-resolution multi-temporal images, for improving urban vegetation classification.

An alternative perspective is offered through the use of fixed digital cameras, which enable object-based analysis at a ground level. Richardson et al., (2007) first demonstrated the suitability of a fixed digital for identifying spring green up in vegetation using repeat photography, as an inexpensive and accessible method. This primary study derived a "greenness" index from the mean values of the Red Green Blue (RGB) bands in each image; expressed through two metrics: excess green ($ExG = 2G - [R + B]$) and the green chromatic coordinate ($Gcc = G/[R + G + B]$). Crimmins and Crimmins (2008) used the same approach but with the integration of local metrological and edaphic variables, also extracting induvial flowers from some species. Their findings suggested that this site-specific approach could greatly enhance the understanding of environmental triggers of phenological events. They also highlighted the potential scope of engaging Citizen Scientists to aid with data collection efforts. Kennan et al., (2014) critically assed the relationship between colour indices, leaf physiology, and canopy structure using repeat digital photography. Findings from their mixing model analysis suggested the development of key physiological leaf traits and canopy leaf area lag behind camera-derived green-up in spring, indicating the seasonal cycle of canopy greenness is driven by a combination of plant changes.

A common issue of this ground based remote sensing method, is the effects of changes in scene illumination, influenced by solar incidence angle and variable weather conditions. Sonnentag et al., (2012) explored the effectiveness of different indices, concluding that the *Gcc* is generally more effective than ExG in supressing the effects of changes in scene illumination. To further smooth the effects of changes in scene illumination, Sonnentag et al., (2012) also adopted a moving-window approach to calculate a three-day *Gcc*. Following the increase in repeat digital photography in research, Brown et al., (2016) drew on its potential for developing a global phenocam network. Although it was suggested that such network would require a number of standardisations, across the metadata, data collection, and the creation of software and bespoke scripts for consistent analytical methodology.

Some studies have broadened coverage to a network of cameras, such as through motorway traffic cameras (Morris et al., 2013), and cameras in public spaces (Graham et al., 2010). Several methodologies exist for extracting greenness metrics from fixed digital camera data: data driven approaches (Bothmann et al., 2017), statistical methods (Graham et al., 2010), and deep learning (Yalcin, 2017). There is evidence to suggest that the potential of spatial coverage from ground level sensors has yet to be fully unlocked. Google Street View imagery has been used an input dataset for urban tree mapping and species classification (Wegner et al., 2016), similarly to a dashcam, it provides a street level view of a tree. Advantageously, this dataset frequently has multiple images as taken along the road, providing a view of the tree from different angles and directions. However, it does not provide a frequent temporal dataset, which is required for deriving phenological insights, in addition to a multitude of other properties and metrics such as disease onset, and biomass.

2.3 Deep Learning for Vegetation Detection

Object detection in computer vision has been dramatically improved by deep learning computational models composed of multiple processing layers which learn representations of data, with multiple levels of abstraction (LeCun et al., 2015). With its origins stemming from a sought after understanding of the human brain, it was not until the early 2000s that the idea of neural networks for useful feature extraction became feasible. One of the most notable milestones came in 2006, when Hinton et al., (2006) pioneered the Deep Belief Network algorithm which could learn deep, directed networks, a single layer at a time, unsupervised. Neural networks have since become an integral component of research in the computer vision field, frequently outperforming the once state-of-the-art machine learning methods (Voulodimos et al., 2018).

Convolutional Neural Networks (CNNs) are deep learning algorithms with convolutional layers, designed for processing gridded data patterns, such as images. The first layers of a CNN mathematical model typically are a 'filter bank layer' and a 'non-linearity layer' which extract characteristics, with a third fully connected 'feature pooling layer' mapping characteristics, a classification for example, into a final product. Inputs and outputs are described as feature maps, and these are comprised as sets of arrays. In the example of a colour image, this would be a 2D array containing the colour channel of the input image, with the output feature map representing a particular feature at all locations on the input (LeCun et al., 2010). CNNs have the power to make frequently correct and strong assumptions about the nature of images, such as pixel statistics and locality, they are characterised by fewer connections compared to standard neural networks making them easier to train (Krizhebsky et al., 2012).

Vegetation detection has long been a subject of remote sensing, but now deep learning approaches are increasingly being proven as an advantageous technique when implemented in such analytical image processing applications. Vegetation detection in imagery has been used is used in phenological studies (Zhang et al., 2009), vegetation mapping (Minh et al., 2018) crop monitoring (Torres-Sánchez et al., 2015), and vegetation management (Mills et al., 2010). Deep learning algorithms have been demonstrated with success in detecting individual trees from satellite imagery (Li et al., 2017), and now there is an increasing trend of applying such algorithms to imagery collected from unmanned aerial vehicles (UAVs). Studies using RGB images (Safonova et al., 2019) in analysis are however detecting solely tree crowns (e.g. Weinstein et al., 2019, Santos et al., 2019), due to the aerial position of the platform. Individual tree detection has been achieved using pointclouds acquired from

UAV-borne LiDAR (Windrim and Bryson, 2018), which advantageously gives a view on an individual tree in the 3D but removes the potential of analysis in an RGB colour space.

There are fewer studies which have demonstrated neural networks for detecting trees in ground level imagery. Wegner et al., (2016) exemplified using a Faster R-CNN algorithm to detect trees and classify species in an urban environment, from dense aerial imagery and street-level imagery, sourced from Google Street View. They also generated an estimated geographic coordinate for detected objects based on platform geometry. Their model combining multiple views of street imagery obtained a higher mean average precision compared to a pure aerial view detector, when using the same Faster R-CNN.

Exponential growth in the computer vision and deep learning fields have advanced to single neural network application for real time object detection. The You Only Look Once (YOLO) algorithm frames object detection as a regression problem to spatially separated bounding boxes with associated class probabilities (Redmon et al., 2016). The architecture is extremely fast as a single neural network works in one evaluation on full images nonetheless, further research is required to fully assess its level of accuracy compared to that of traditional CNNs. YOLOv2 has been used for tree detection, comparing 3D images sourced from a 360◦ camera and LiDAR pointclouds, with high detection accuracy (Itakura and Hosoi, 2020). The standard YOLO deep learning algorithm operates by dividing an image into a 13x13 grid, and the minimum classifier bounding box can therefore have a maximum size of a single cell. This results in some problems when detecting small objects. However, methods of scaling up the grid has been used to improve the model and has produced high accuracies in in detecting single fruits on apple trees for instance (Bresilla et al., 2019). Further improvements have been applied to the YOLOv3 algorithm; it uses multi-scale prediction to detect the final target. Similarly, the YOLOv3 detector has successfully been used for detecting apples for tree fruiting phenology (Tian et al., 2019). There is little to no evidence which suggests the YOLO deep learning algorithm has been used to detect trees at different phenological stages, when focusing on the leaf development and green-up.

2.4 The Urban Environment: Citizen Science

Frontiers in ecological research and public engagement are continuously expanded by the coupling of new technologies and Citizen Science, with substantial scope to accelerate the collation of useful information through innovative uses of existing technology (Newman et al. 2012). Crowd-generated data with sensor-derived content presents great potential in leading a better understanding of the manner in which cities function as systems (Crooks et al. 2016), a system vegetation is encompassed within. It has been recognised that dashcam video sharing will greatly extend urban surveillance (Park et al. 2016) as such, acquiring dashcam imagery opens scope to create a complete and consistent dataset, at a road level scale that provides intel on the green infrastructure of cities and towns.

Local authorities are routinely required to allocate expenditure to collect vital data and information, which is implemented into urban monitoring and management plans. For many social and environmental datasets, geographic location is a powerful, yet intuitive, cataloguing structure (Connors et al., 2012). Citizen Science is the process of involving members of the public in scientific projects as researchers (Kruger and Shannon, 2000), its value has long been recognised and now a plethora of Citizen Science projects and initiatives exist, designed to assist with various aspects of urban monitoring, which in turn contribute to planning and management efforts.

The sources of pollution and how to reduce it is a key challenge for urban areas around the globe, and a number of participatory sensor models have been used to monitor air quality (Dutta et al., 2009, Leonardi et al., 2014) and sampling models for water quality (Jollymore et al., 2017), with other models taking bottom-up approaches (Jiang et al., 2016). Geo-localised and citizen user-generated data has been used to map noise pollution in cities through mobile applications (Maisonneuve et al., 2009). Citizen Scientists are also frequently called upon for urban ecological studies, such as bird species surveys (McCaffrey, 2005), bumblebee population monitoring (Lye et al., 2012), assessing small mammal behaviour (Williams et al., 2015, Anton et al., 2018) and reporting invasive plant species (Gallo and Waitt, 2011).

Access to green space has increasingly become an integral part of wellbeing and quality of life for residents of urban environments, and in tandem with the growth in understanding ecosystem services, there has been increasing effort in monitoring urban green infrastructure to ensure survival and longevity. Tree inventories hold value for a range of purposes including managing and maintaining risks and benefits associated with trees (McPherson et al., 2005), while also frequently used for more generalised monitoring in terms of condition and distribution (Nowak and Greenfield, 2012). Such inventories and databases have been developed with Citizen Science contributions ranging from new urban planting (Vogt et al., 2017) to developing inventories specifically for routine street pruning (Crown et al., 2018). Despite many inventories produced in the field, one study explored virtual surveying using street level imagery for tree locating and species identification (Berland et al., 2019). An overall trend in Citizen Science initiatives concerning urban vegetation suggests a positive and cost-effective method for data collection. However, it has been noted that with only minimal training, Citizen Science tree observations are not always consistent with that of the expert, although if used appropriately data is still valuable depending on the purpose and required level of accuracy (Roman et al., 2017).

Existing literature suggests that the integration of using citizen volunteered data with dashcams as the sensor has yet to be exploited to its full potential. There is evidence of the incorporation of dashcams for urban monitoring, such as community policing (Park et al., 2017) and pedestrian flow estimation (Hara et al., 2018), with one study taking a crowd-sourcing approach to collect information on road conditions (Aihara et al., 2017). There is little to no evidence however which suggests the active engagement of volunteered dashcam data from citizens specifically for ecological monitoring in urban environments.

Although dependent on the data usage, a repeatedly cited concern for both local authorities and citizens involved in Citizen Science initiatives for urban monitoring, is the data quality. Seemingly, there is an inherent requirement for standardisation for such initiatives, with sensor calibration and clear guidance a paramount consideration. Conrad and Hilchey (2010) suggest that in order to develop an improved understanding of the ecological, social and economic benefits of Citizen Science initiatives, there are two major gaps which require addressing. They identified the need of sound evidence presented from successful initiatives, whereby projects are compared and contrasted, and secondly, the need for more examples of Citizen Science derived data used by decision makers, to aid with the identification of potential linkages or barriers across projects.

3. AIMS AND OBJECTIVES

Aim: *To develop automated steps for extracting vegetation green dynamics in an urban environment from a road level dataset.*

Objectives

1. Assess the suitability of vehicle dashcams for capturing roadside vegetation imagery
2. Evaluate the suitability of the YOLOv3 deep learning algorithm automatically detecting roadside trees at different phenological stages
3. Consider the appropriate image processing techniques for deriving a greenness metric
4. Consider the potential of Citizen Science volunteered dashcam footage for vegetation monitoring

4. METHODOLOGY

4.1 Dashcam Imagery

A commercially available dashcam, the Nextbase 322GW (NBDVR322GW), was used for data capture. The sensor resolution is 2.12M Pixel (1936 x 1097) with a recording resolution of 1920 x 1080 (HD) at 60 frames per second (fps), with a 140° lens angle and a f1.6 focal ratio. The recording format is MP4 and saved to 60 second clips. Files save onto a removeable 128GB MicroSD storage card. Pixel dimensions of a single frame are 1366 x 768, with a resolution of 96dpi and a 24-bit depth. The dashcam was mounted to the interior of a car, connected by a 12-24Vdc car power cable.

4.2 Data Capture

Data was acquired daily across a 90-day capture period from 03/03/2020 to 31/05/2020. This was selected based on findings of a similar study in the same geographic region. Across 18 sites in England, Morris et al., (2013) recording the start and end of green-up between 01/03 to 01/05 and 22/04 to 01/06 during the 2012 season.

As a result of connection failure (power cable dislodging when crossing speed bumps) the survey was incomplete for the following dates: 21/03/20, 14/05/20, 25/05/20. Daily data was typically captured between 0745 and 0830, although the capture time did vary outside this throughout daylight hours during the 90-day capture period, due to factors including traffic and driving conditions. This characterises a common use of dashcams during weekday work commutes, and the leisurely drives more typical of a weekend. The variation in the timing of the data capture would better mirror what would be expected in the context of sourcing volunteered geographic information and Citizen Science approaches for data.

4.3 Survey Location

Melksham (51.3704, -2.1376), a small market town in South-West England. The climate, based on information sourced from the nearest weather station in the City of Bath (51.3754, -2.3585 at 50m altitude above msl) is subject to an Atlantic Maritime air mass. The oceanic climate is characterised by warmer summers (July average of 21.7c) and cooler winters (January average of 7.6c) with rainfall all year round with the average annual total of 814.1mm between 1981 – 2010.

(For additional information see: https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-climate-averages/gcnk62de6). The survey transect (Figure 2) is characterised by various deciduous tree species, including: Silver Birch, Lime, Chestnut, Hazel, Rowan and London Plane.



Figure 2. Map of Survey Transect Route

4.4 Video Pre-Processing

Areas of interest (AOIs) containing a tree(s) were manually extracted, in the open-source VLC media player. The video clips were panned through, and then viewed frame by frame, to derive the 'optimum view'. This was objectively selected as the frame which contained the entire tree in a single scene, which was then extracted and saved (aoi#_YY/MM/DD.png). There was slight variation to this 'optimum view' across the 90-day capture period. This was in part due to some minor movement of the dashcam position, and vehicles obscuring the tree of interest at the most optimum view. A total of seven AOIs equated to 523 images as the test image dataset (see Table 1).

Table 1. Descriptions of Area of Interests

Area of Interest	Descriptive/Locative Name	No. of Test Images	Single Tree or Multiple Trees?	Species (')
AOI_1	LimeRow	88	Multiple	Lime
AOI_2	ChestnutTree	88	Single	Horse Chestnut
AOI_3	LibraryTree	87	Single	Rowan
AOI_4	RoundaboutTree	87	Single	Plane
AOI_5	FireStation	87	Multiple	Acer
AOI_6	SemingtonRd	86	Multiple	Birch, Long Leaved Lime

4.5 Deep Learning

The You Only Look Once (YOLO) deep learning algorithm frames object detection as a regression problem to spatially separated bounding boxes with associated class probabilities (Redmon et al., 2016). The YOLO architecture applies a single forward pass neural network to the whole image and predicts the boxes and probabilities. It works by dividing the image into a grid, and each one of these grid cells will predict N number of possible bounding boxes in the image, in addition to a confidence value of it actually enclosing an object and the probability of such object being in a particular class. Most bounding boxes are then eliminated because of low confidence or they are enclosed by another bounding box with a very high confidence score. And this is a technique called non-maximum suppression. As such is extremely fast in comparison to other neural network algorithms. YOLO version 3 (YOLOv3), an update containing minor design changes which improved accuracy on the previous two versions (Redmon and Farhadi, 2018), is the algorithm applied to this methodology.

The following key steps were taken as part of the YOLOv3 implementation:

1. Data Annotating
2. Data Training
3. Inference

4.5.1 Data Annotations

A subset of the test images was extracted systematically from the test images (Table 2), which equated to 42 images. Two methods of image augmentation were applied to the images (twice duplicated), to have a median blur filter applied and horizontally flipped, facilitated in Python using the Pillow and OpenCV libraries. The total of number of training images for annotation was 126. Microsoft's Visual Object Tagging Tool (VOTT) was used to annotate images. Within VOTT, the directory containing the 126 training images was set and the images was added as a connection to the software. Firstly, a class 'Tree' tag was created. Then for each of the 126 training images, a bounding box was manually drawn around each tree in every image and tagged with the 'Tree' class (Figure 3). The project was then exported under the name 'Annotations'. During this, an export CSV file was created containing each of the training image filenames and the coordinates of each bounding box drawn.

Table 2. A sample of training images, taken periodically across the capture period from the test images of the six areas of interest, with the aim of representing the vegetation at various growth stages.

DAY OF CAPTURE PERIOD	DAY OF YEAR	DATE (DD/MM/YYYY)
DAY 1	63	03/03/2020
DAY 15	77	17/03/2020
DAY 30	92	01/04/2020
DAY 45	107	16/05/2020
DAY 60	122	01/05/2020
DAY 75	137	16/05/2020
DAY 90	152	31/05/2020



Figure 3. Example of the manually drawn bounding box in VOTT.

4.5.2 YOLOv3 Implementation

The source code for running the YOLOv3 object detection algorithm was downloaded from a GitHub Repo, available at: https://github.com/AntonMu/TrainYourOwnYOLO.git. The object detector training and inference was implemented in Python, facilitated through Google Colabratory (Colab). All data was uploaded for training and processing was uploaded to the Google Drive.

Google Colab is a free cloud computing platform, available as an online environment in the layout of the Jupyter Notebook graphical user interface, for writing and executing code in Python. The platform enables users to mount their Google Drive to directly access files for processing and facilitates the direct cloning of GitHub repositories. It also allows users to upload existing notebook and share notebooks created. Google Colab supports free a free, cloud-based graphics processing unit (GPU) for hardware acceleration through virtual machines, which is particularly advantageous for deep learning applications using libraries such as Keras, TensorFlow, and PyTorch. However, the maximum runtime lifetime for connecting to virtual machines on this online, cloud computing platform is 12 hours.

4.5.3 Training

Firstly, a script was run to convert the exported VOTT CSV to the required YOLOv3 format. This generated two text files which would later be used in training, data_train.txt and data_classes.txt. Following this, a script was run to download pre-trained YOLOv3 weights and convert them to Keras format. These pre-trained weights have been pre-trained on the ImageNet 1000 dataset. The training script was then run, taking approximately 8 hours to complete.

4.5.4 Inference

Once training had completed, a detector script implemented the YOLOv3 algorithm on the test images. This process of object detection took a total of 26 minutes. Output products included the original test images with bounding boxes and confidence scores, and a Detection Results CSV file containing the file paths and the bounding box coordinates.

4.6 Training Data Reduction

The data annotation, training and inference steps were repeated using the same training dataset but reduced in size by 50%. Half of the original training images were randomly removed, reducing the training dataset to 63 images. This stage has been carried out in order to consider the effects of a smaller training set on the YOLO performance and the subsequent accuracy of the bounding boxes, with the intention of ascertaining the most efficient method of developing an appropriate training dataset.

4.7 Image Processing

Following a visual review of the object detector results, the Detection Results CSV was used to filter the results and remove outliers. A total of 64 bounding boxes had a Y minimum coordinate of 0, suggesting the bounding box had reached the edges of the image. Based on optimum view test image selection, it was assumed that any coordinates with a Y minimum coordinate had not produced a good bounding box enclosing any trees. A further 62 bounding boxes were removed where the tree class confidence score was <0.30. This was decided following a visual analysis, where accurate tree bounding boxes were observed despite a low confidence score of between 0.30 and 0.50. The remaining bounding box coordinates were then used to crop all test images to coordinates, reducing the image size for analysis to a new area of interest, a single tree. These cropped images were exported as new files.

Images were manually re-organised into AOI groups within the file folder structure. Another manual check was used to remove some duplicate dates, where the YOLOv3 object detector had created multiple bounding boxes on a single image. The image, which was determined as the most accurate bounding box, i.e. containing the most tree, was retained.

For the next stages of processing, the following Python libraries and modules were used: Pillow, OpenCV, Pandas, Matplotlib, NumPy, OS, and Glob. Using the Detection Results CSV file, all images were cropped to the corresponding bounding box coordinates.

Images were then converted from Red, Green, Blue (RGB) values to the Hue, Saturation, Value (HSV) colour space to mitigate potential error due to solar variance and brightness in the image, effecting scene illuminance (Yang et al., 2015). In the HSV colour space, distribution of colour for a given coloured object is invariant with respect to brightness variation (Chaves-Gonzalez et al., 2010). As such, analysis on the images can take place in a colour space where colour is not correlated with brightness.

Cropped images still contained background features that were not desired as part of image when deriving the Gcc as the greenness index specifically from the vegetation. As such, a user defined region of interest (ROI) for each of the images was generated, using the Python resizeimage library. A specified rectangle was used to resize and crop all images, which defined the ROI, from each image's centroid. For aoi1, aoi2, and aoi3, the rectangle specification measured 200x100 pixels. For the two trees at aoi4, the ROI rectangle was specified as 150x100. These images were exported as a new file.

4.8 Greenness Plotting

The Green Chromatic Coordinate (Gcc) (1) was then derived for each AOI, across all ~90 images. The NumPy library was used to determine the mean value for each of the RGB channels, in the RGB image. This was repeated on the red, green, and blue segments of the HSV images, and for a third time on the cropped HSV ROI.

$$Gcc = \frac{green}{red + greeb + blue} \quad (1)$$

All Gcc values were transposed into a CSV file, in date order (03/03/2020 – 31/05/2020), against a day of year (DOY) column. The CSV files for all AOIs were read to plot the data on a scatter graph, with DOY as the x axis and the Gcc on the y axis. To further smooth the effects of variable scene illuminance in the data, a line graph was plot as a three-day rolling mean (Sonnentag et al., 2012). This step was repeated on the original cropped RGB images, the converted HSV images, and the ROI HSV images, in order to view and visually compare the effectiveness of these processing techniques on the data.

4.9 Evaluation Metrics

The performance of the object detection model was evaluated with the Mean Average Precision (mAP) metric on the AOIs used in images analysis, which equated to 32 bounding boxes. Code to perform evaluation was sourced from the following GitHub Repo: https://github.com/rafaelpadilla/Object-Detection-Metrics (Padilla et al., 2020). The ground truth bounding boxes manually drawn for training were compared to the predicted bounding box outputs from inference, to derive the Intersection over Union (IoU) (2). This was determined by finding the area of overlap and area of union between the two sets of images. For each image used in evaluation, a ground truth text file contained the class name and the bounding box coordinates, with a corresponding detection text file which contained the class name, confidence score, and bounding box coordinates.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (2)$$

Two different IoU thresholds, 50% and 75%, were used to determine the accuracy of the detector. A IoU score of >0.5 is considered a 'good' prediction, and a score of >0.75 as 'great', in each instance classified as a True Positive (TP). In each instance, a score below each of the IoU threshold is considered a 'poor' detection and is classified as a False Positive (FP). The absence of a bounding box detection where a ground truth is positive is classified a False Negative (FN). Precision (3) and Recall (4) were then derived to further evaluate the detectors performance.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (3)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (4)$$

The Precision and Recall results were then used to determine the mAP accuracy of the detector across the dataset, which was plotted on a graph, for both the 50% IoU threshold and the 75% IoU threshold. The evaluation steps were also repeated on the second trained YOLOv3 detector trained with a reduced training set.

4.10 Field Survey

A field survey was completed along the dashcam survey transect, to collect information on the species of trees and the location of the trees. To reduce manual data input and transferral, ESRI's ArcGIS Collector App, a mobile mapping application, was the selected method of data collection.

4.11 ArcGIS Collector App

Prior to data collection, the base map was prepared in desktop ArcMap. A point feature class 'Tree_Survey' was created with the following attributes: Number, Species, Comments, and Location. Binary large objects were enabled to allow photo attachments. The feature class was projected in Ordnance Survey Great Britain (OSGB) 1936. The base map and feature class was then published as a service ad 'Tree_Survey_Map', with editable permissions granted. The map was then available as content on ArcGIS Online. The Collector App was

installed onto a 4G iOS mobile device. After logging into through ArcGIS Online credentials on the Collector App, the 'Tree_Survey_Map' could be view and used.

4.11.1 Data Collection

The field survey was carried out on foot. When trees, which featured in an area of interest frame used in the image analysis, were identified it was recorded as a point feature with the map on the Collector App. The feature attribute fields were populated, and a photograph of the tree was taken and attach. Each feature was automatically uploaded onto ArcGIS Online once submitted, using cellular (4G) data. Data was stored on the cloud until later download. The location accuracy information was based on the GPS location as recorded by the iOS device.

4.11.2 Data Download

In ArcGIS Online, data collected in the field was visible as a cloud hosted map. The feature class containing tree species and locative information was downloaded as a zipped File Geodatabase and saved locally. Data was then imported to ArcMap, plotted against the survey transect line, before being exported as a map.

4.12 Nature's Calendar

The Woodland Trust's Citizen Science platform, Nature's Calendar, allows members of the public to upload their observations of phenological events onto an interactive Species and Events Map. This was used to obtain observations of first tree leafing, in Wiltshire and surrounding counties, to situate the phenological outputs produced in this study. The records of six common tree species were input to a table with the dates recorded by Citizen Scientists of first leaf.

4.13 Data Comparisons: Google Street View

A screenshot of each of the six AOIs was taken from the closest position available to match that of the frame used in analysis, in Google Street View via Google Maps. This was used as a comparative tool to demonstrate some existing sources of street level imagery.

5. RESULTS

5.1 Object Detection Evaluation Metrics

The results of the YOLOv3 object detector model produced good results overall with the evaluation dataset producing values >0.83, obtaining a mAP accuracy of 87.32% with the IoU threshold set at 50% (see Figure 6) and obtaining a mAP accuracy of 70.73% with the IoU threshold set at 75% (see Figure 7). The performance of the detector consistently identified trees in the images, however there was some variation in the accuracy of the bounding boxes produces. There were instances when the detector output multiple bounding boxes for a single tree, 23% of test images in AOI_2 had multiple bounding boxes for a single tree, 18% in AOI_3, 22% in AOI_4, and 12% in AOI_5. Although these were straightforward to eliminate from further analysis, as typically where a tree had two or more bounding boxes, one box had a much higher confidence score hence the remaining could be removed. A trend observed across all AOIs was that the detector performed poorer on trees with leaf off, and better with leaf on (see Figure 4 for an example).
AOI_1 and AOI_6 were omitted from further analysis following the detector inference, due to the added complexity of obtaining and organising the bounding boxes for image processing where a single image contained >2 trees, particularly due to overlapping. Additional time would be required to develop suitable Python code for bounding box extraction.

The results of the YOLOv3 detector with a reduced training dataset produced a comparably better accuracy results, obtaining a mAP of 94.82% with the IoU threshold set at 50% (see Figure 8) and obtaining a mAP accuracy of 74.99% with the IoU threshold set at 75% (see Figure 9). This detector produced better results but with lower confidence scores overall (see Figure 5 for an example).



Figure 4. Example comparison of detector outputs between leaf off and leaf on.



Figure 5. Example of the original detector results (left) and the detector results with a reduced training size detector (right) at AOI_4 on DOY 121.

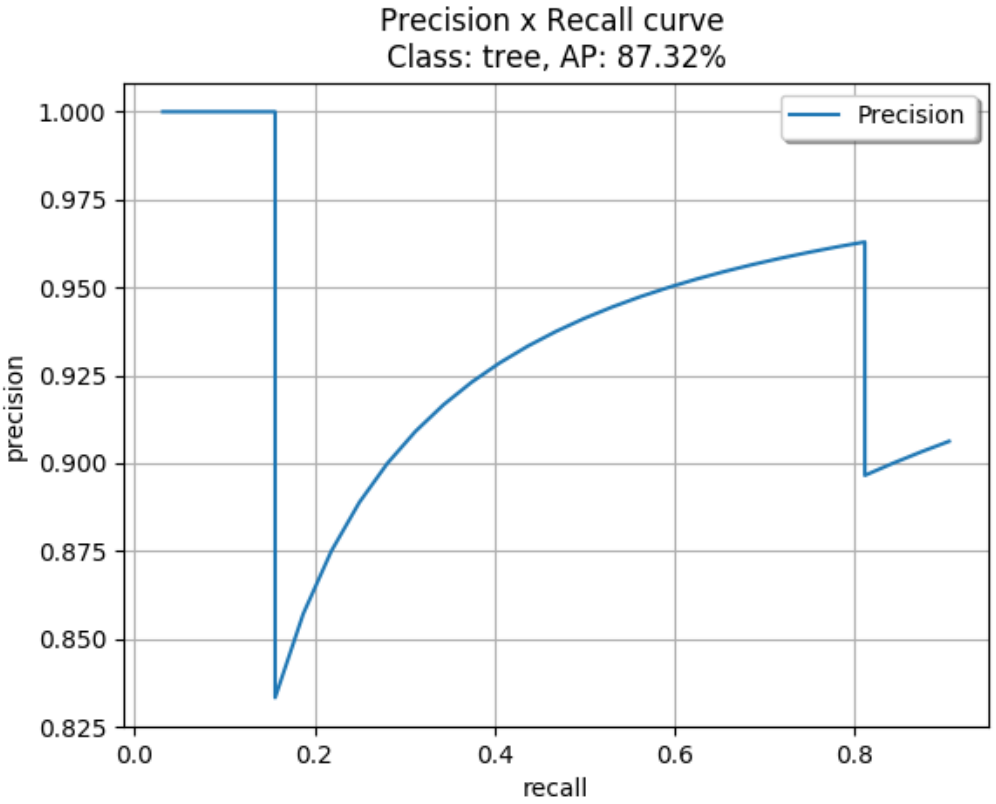


Figure 6. Average precision for the original YOLOv3 object detector, where IoU is 50%.

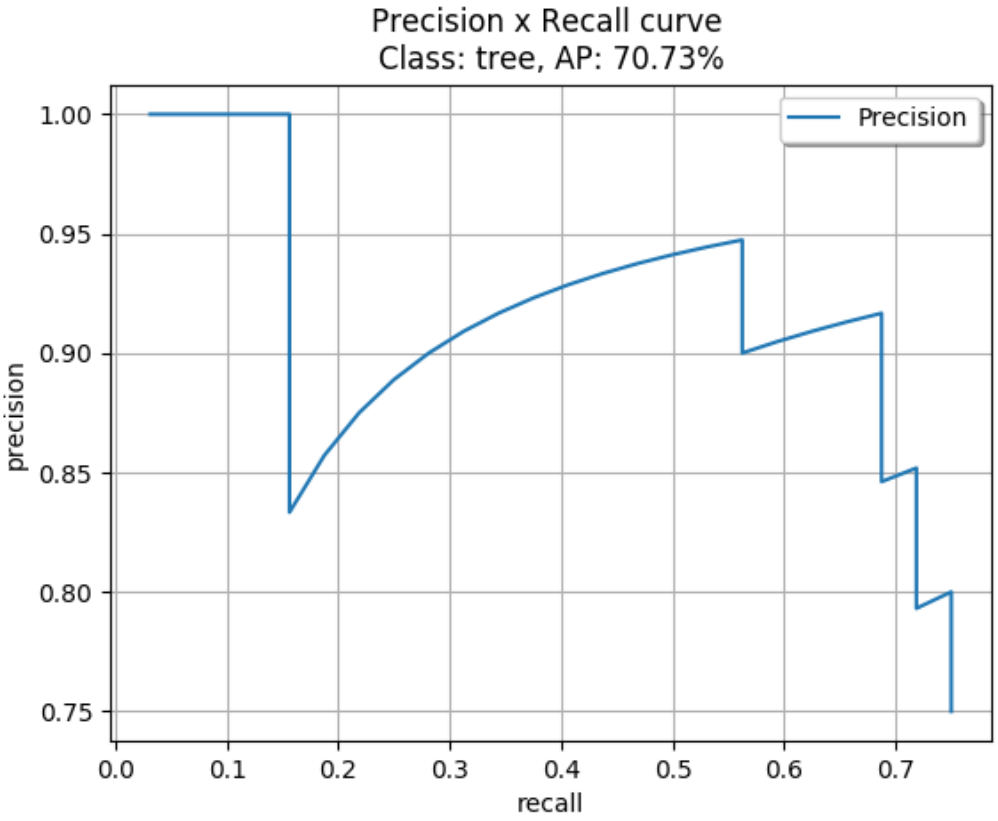


Figure 7. Average precision for the original YOLOv3 object detector, where IoU is 75%.

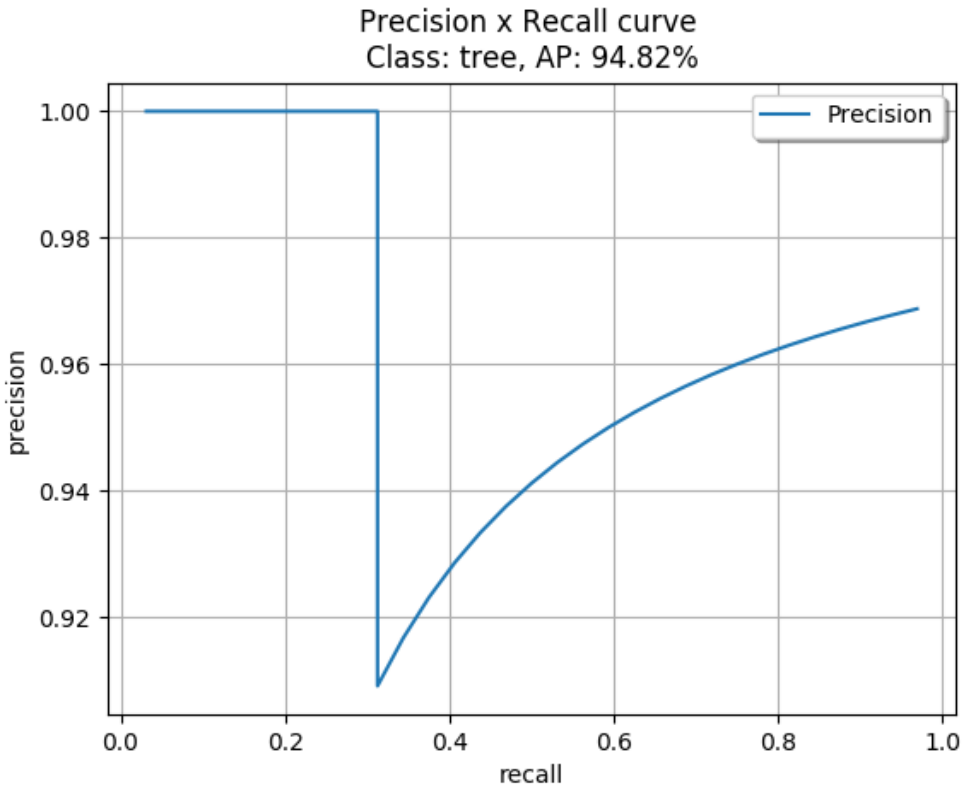


Figure 8. Average precision for the reduced training YOLOv3 object detector, where IoU is 50%.

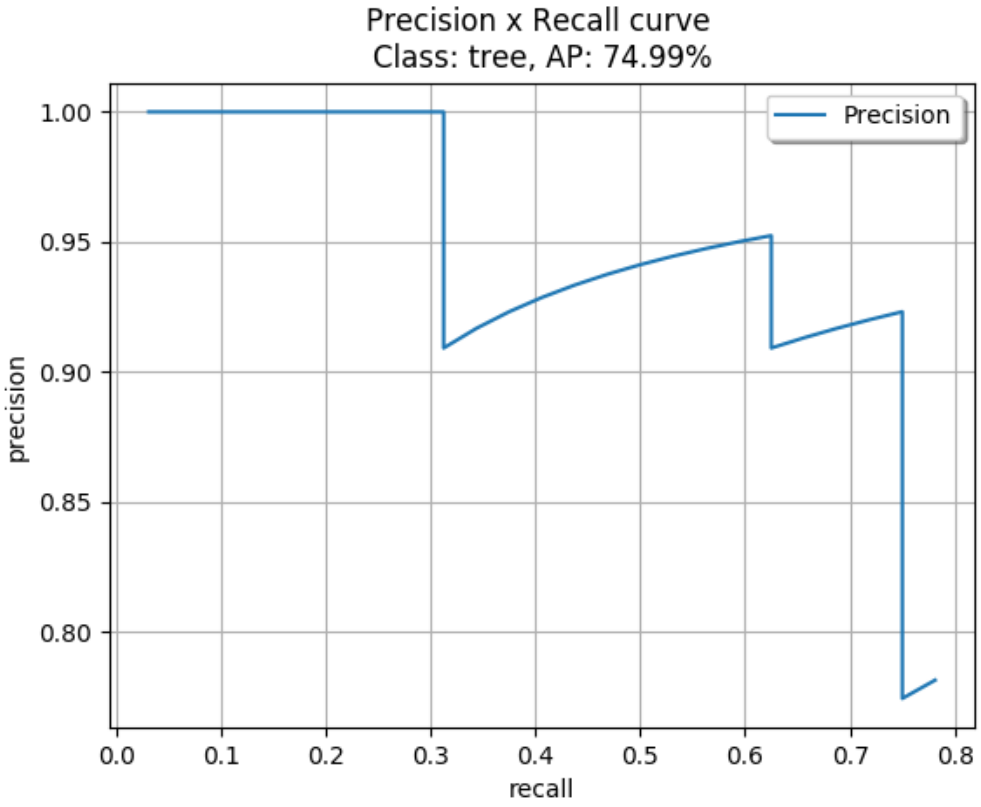


Figure 9. Average precision for the reduced trainingYOLOv3 object detector, where IoU is 75%.

5.2 Image Processing and Deriving Greenness

All clipped bounding boxes of the original RGB test images were converted to the HSV colour space, which was then also clipped to new coordinates from each images' centroid (see Figure 10).

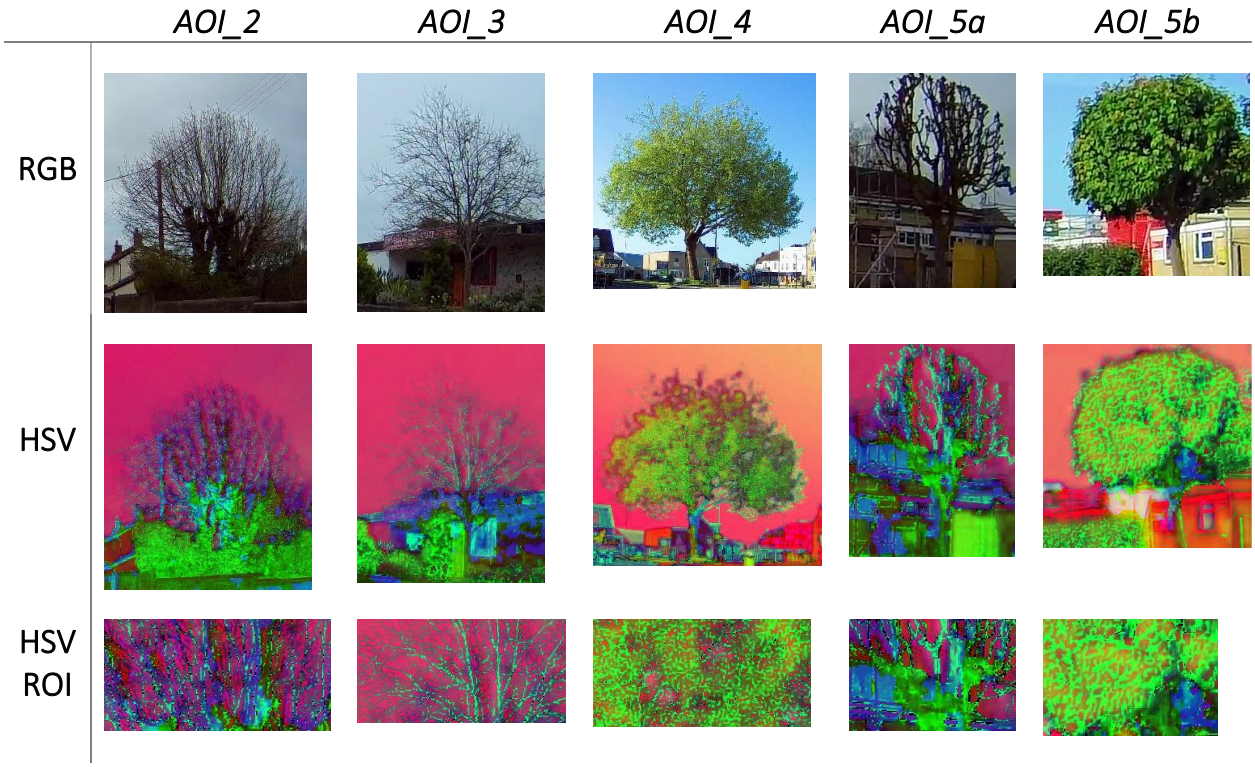


Figure 10. AOI examples of clipped RGB images, the converted HSV image, and the ROI.

The Green Chromatic Coordinate was derived from the original clipped RGB image, the RGB-HSV converted, and the region of interest from the HSV image, creating three graphs for the following area of interests: AOI_2 (Figures 11-13), AOI_3 (Figures 14-16), AOI_4 (Figures 17-19), and AOI_5 (Figures 20-22). A positive correlation was observed between increasing greenness overtime for all AOIs using all three image processing techniques with the exception of the HSV and ROI graphs for AOI_5b. It appeared that the HSV had a stretching effect on the Gcc values up until the start of green-up, after which it relatively smoothed the values.

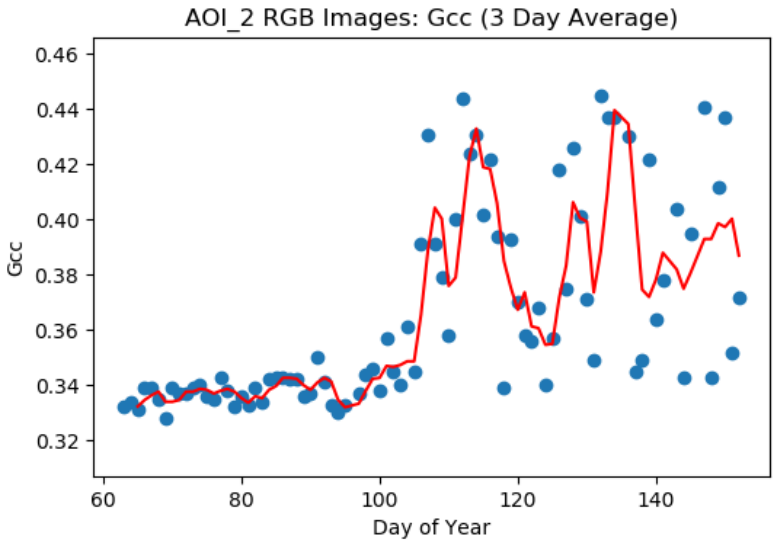


Figure 11. The Green Chromatic Coordinates plotted for AOI_2, calculated using the RGB.

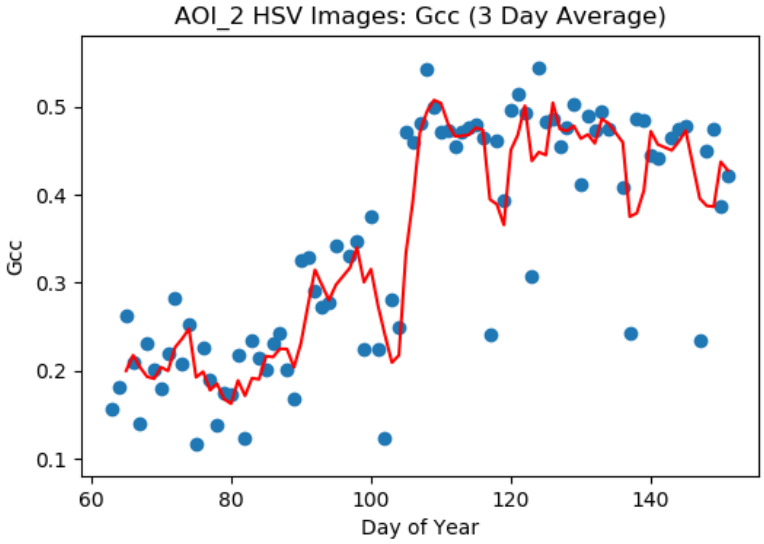


Figure 12. The Green Chromatic Coordinates plotted for AOI_2, calculated using the HSV.

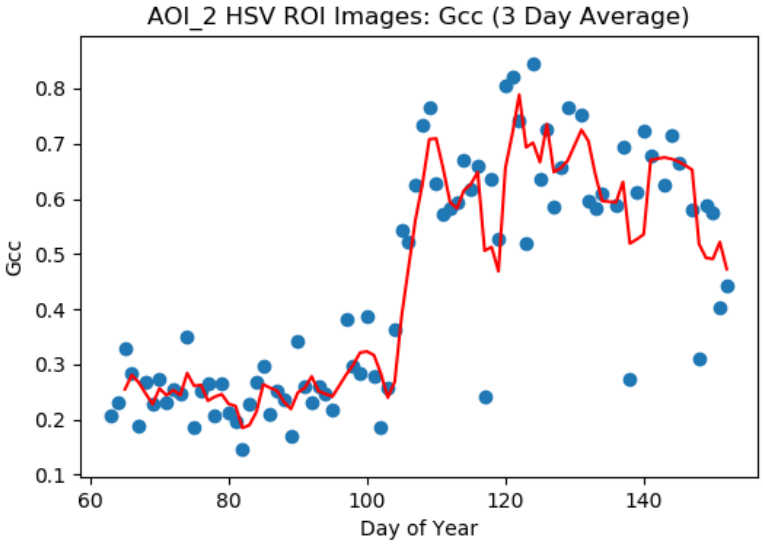


Figure 13. The Green Chromatic Coordinates plotted for AOI_2, calculated using the ROI HSV.

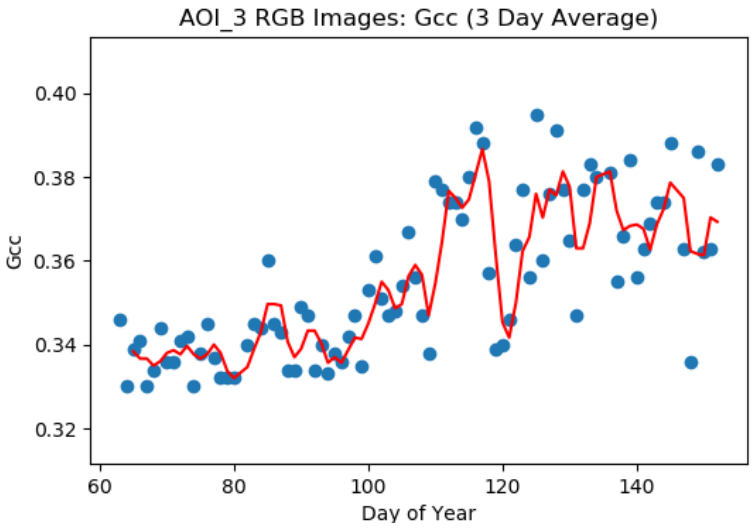


Figure 14. The Green Chromatic Coordinates plotted for AOI_3, calculated using the RGB.

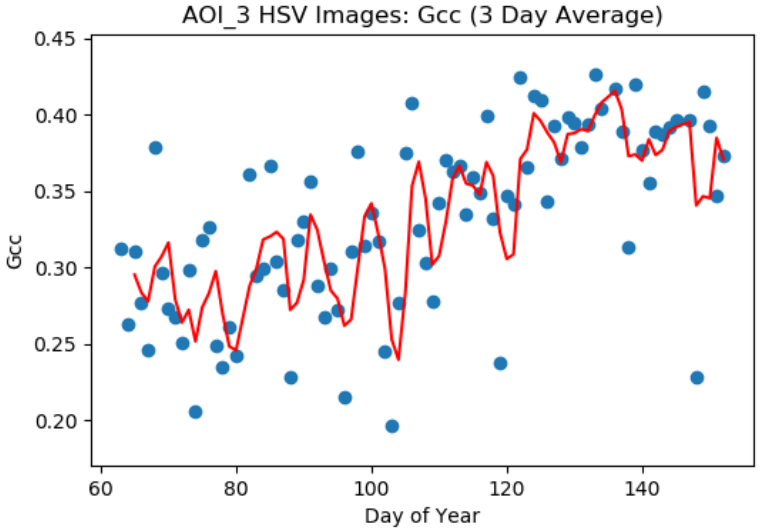


Figure 15. The Green Chromatic Coordinates plotted for AOI_3, calculated using the HSV.

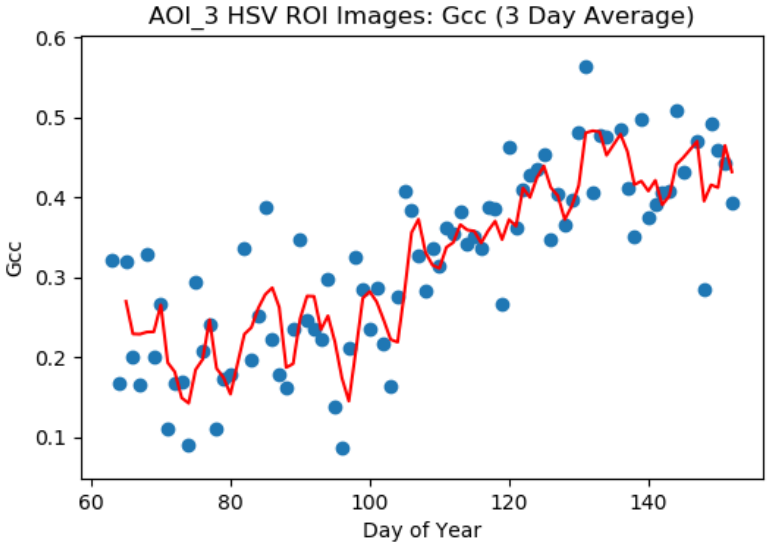


Figure 16. The Green Chromatic Coordinates plotted for AOI_3, calculated using the ROI HSV.

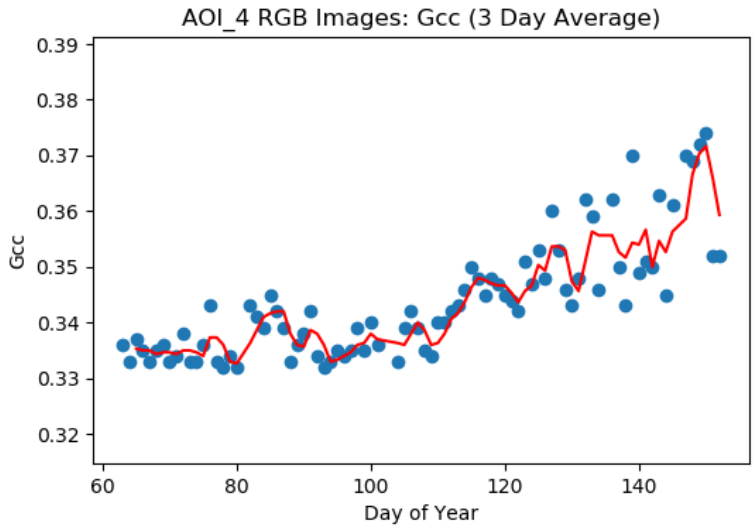


Figure 17. The Green Chromatic Coordinates plotted for AOI_4, calculated using the RGB.

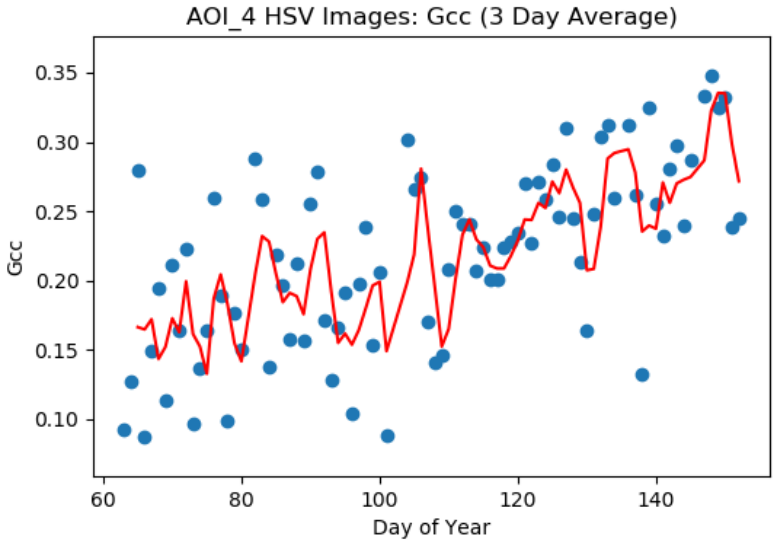


Figure 18. The Green Chromatic Coordinates plotted for AOI_4, calculated using the HSV.

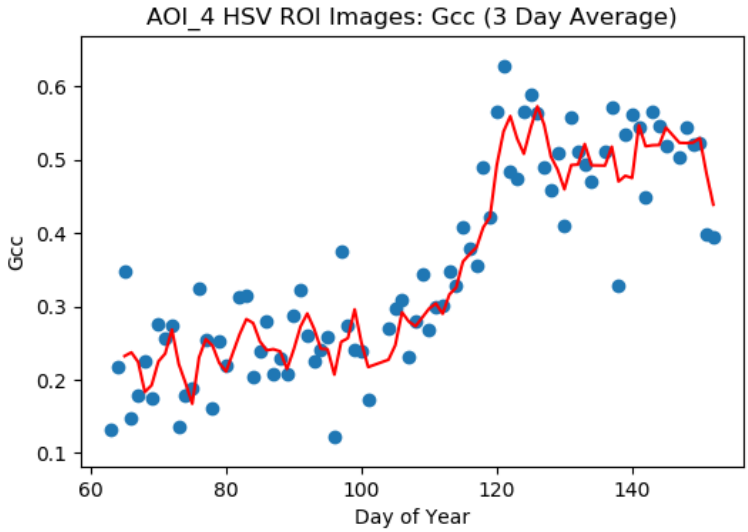


Figure 19. The Green Chromatic Coordinates plotted for AOI_4, calculated using the ROI HSV.

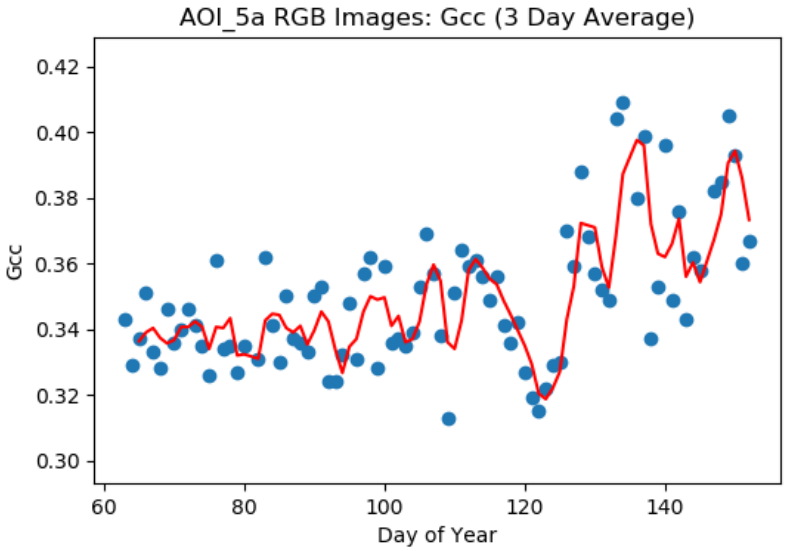


Figure 20. The Green Chromatic Coordinates plotted for AOI_5a, calculated using the RGB.

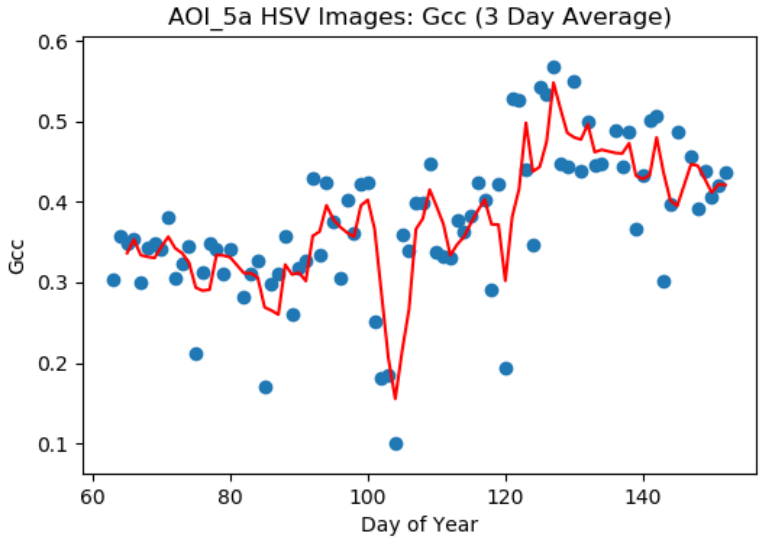


Figure 21. The Green Chromatic Coordinates plotted for AOI_5a, calculated using the HSV.

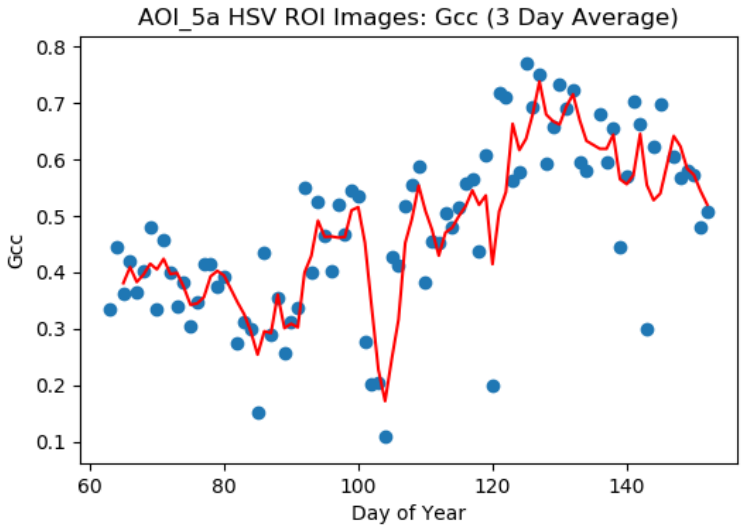


Figure 22. The Green Chromatic Coordinates plotted for AOI_5a, calculated using the ROI HSV.

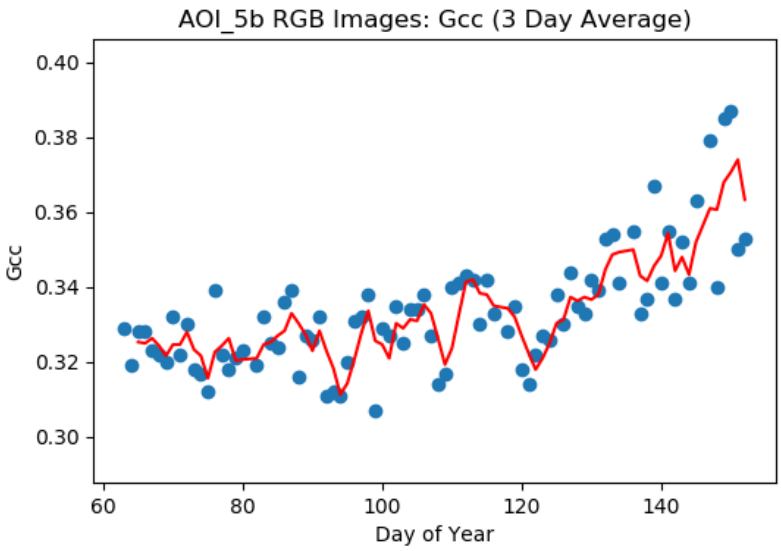


Figure 23. The Green Chromatic Coordinates plotted for AOI_5b, calculated using the RGB.

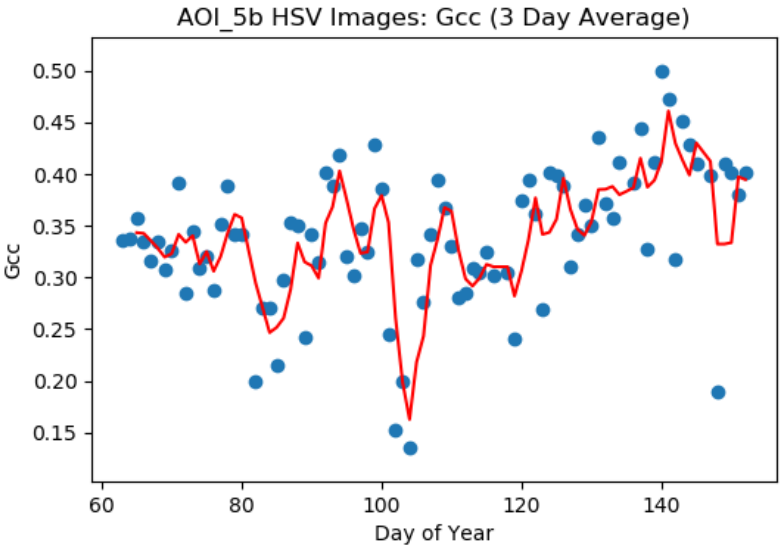


Figure 24. The Green Chromatic Coordinates plotted for AOI_5b, calculated using the HSV.

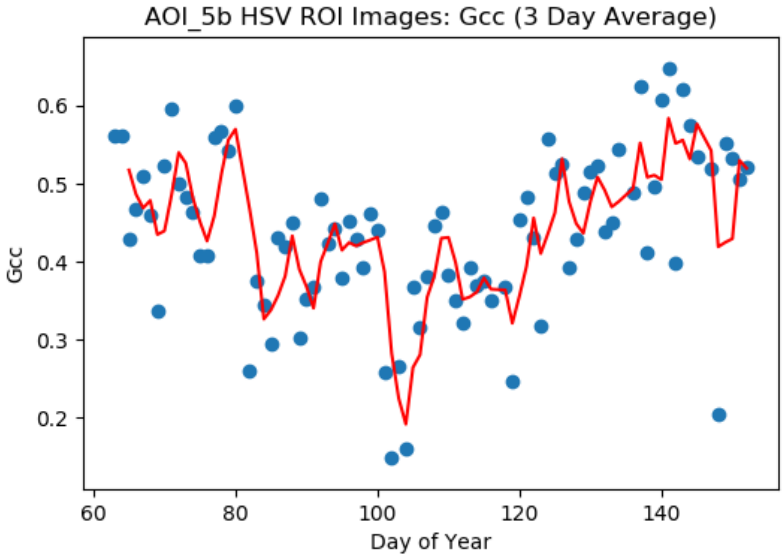


Figure 25. The Green Chromatic Coordinates plotted for AOI_5b, calculated using the ROI HSV.

The start of green-up was determined based on the 3-day rolling mean Gcc coordinate of each AOI, from the ROI graphs. Both the Horse Chestnut tree in AOI_2 and the Rowan tree in AOI_3 observed the onset of greenness on the 14th April, which can be attributed to budburst. AOI_2 plateaued in greening just four days later on the 18th April, with AOI_3 greener slower, reaching a plateau of greenness on approximately the 9th May. The Plane tree of AOI_4 began green-up 10 days later than AOI_2 and AOI_3, on the 24th April, reaching maximum greenness on approximately 29th April.

Table 3. Start of green-up date for each AOI.

AOI	Species	Start of Green-Up (Date)
AOI_2	Horse Chestnut	14th April 2020
AOI_3	Rowan	14th April 2020
AOI_4	Plane	24th April 2020
AOI_5a	Acer	23rd April 2020
AOI_5b	Acer	4th May 2020

It was difficult to distinguish the onset on greenness in the trees in AOI_5. For tree AOI_5a there was an evident positive relationship between an increase in greenness over time, although there were significant fluctuations, however green-up can be pin-pointed to 23rd April. AOI_5a began to plateau in greenness on approximately the 4th May. The green-up was easiest identifiable in the pure RGB images for AOI_5b, and despite the same species as the other tree in AOI_5, green-up occurred 11 days later on the 4th May. The fluctuations in the data make it impossible to confidently select an approximate greenness plateau date for AOI_5b, assuming it occurred within the 90 day data capture period.

5.3 ESRI Collector App – Field Mapping

The ESRI Collector App was an effective method for mapping the AOIs. Utilising the global positioning system embedded into the mobile device, which had downloaded on it the Collector App, was quick and straightforward. Although, the range of the positioning accuracy did vary but not exceeding 20metres, which subsequently resulted in the points mapped not being very precise. However, this was sufficient for the purpose of mapping the AOIs featuring trees in relation to other trees and the survey transect itself. Once the data had been collected, it was immediately hosted in the cloud, and later downloaded to be opened in ArcMap to create a map (see Figure 26).

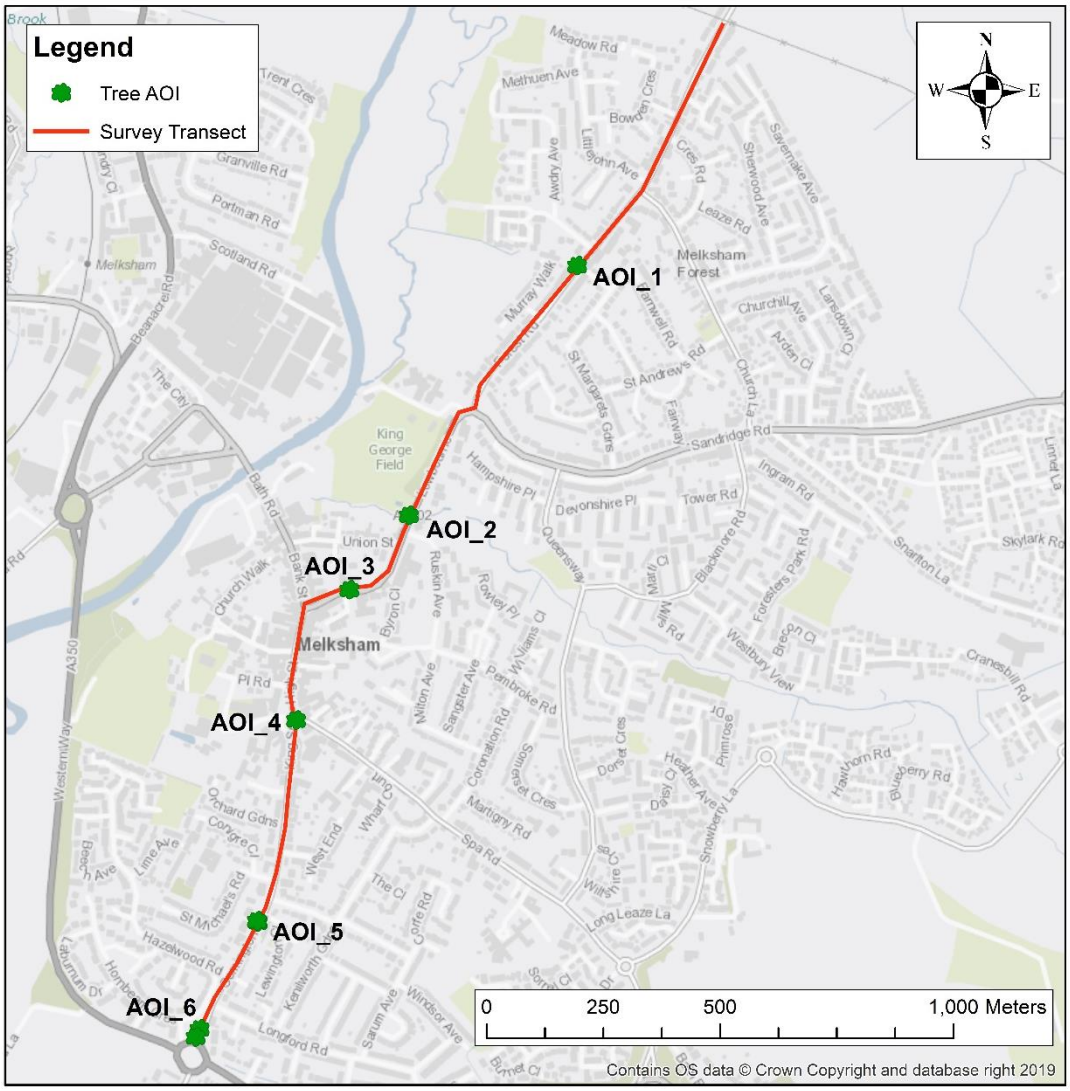


Figure 26. A map to locate each AOI in relation to the survey transect.

5.4 Nature's Calendar

The earliest first leaf recorded by Citizen Scientists in the county and surrounding counties of the study was on the 13th March 2020 with the latest first leaf recording made on the 24th April 2020 (Table 4), for the species given. Of the information obtained from these six species, there were only 12 data entries, with the maximum number of entries for a single species not exceeding three, all of which were different dates.

Table 4. First leaf records from the 2020 spring green-up, obtained from the Woodland Trust's Nature's Calendar Citizen Science Project.

Species	Horse Chestnut	Rowan	Birch	Field Maple	Ash	Beech
Citizen Science Recorded First Leaf Date	22/03	13/03	26/03	03/04	09/04	13/04
	25/03			04/04	10/04	15/04
					12/04	24/04

5.5 Data Comparisons: Google Street View

When comparing Google Street View imagery to the images captured by the dashcam, it provides a very similar view. However, the camera used for Google Street View data capture is roof mounted, coupled with the 360-degree camera used, it provides a wider field of view compared to the dashcam which is 140 degrees and forward facing only. Google Street View imagery is only available as a single day of capture. It is free to view, and the google_streetview API available for installation in Python does enable panorama download, however a longitude and latitude is required and there is no option for automatic batch download. See Figures 27 to 30 for the corresponding Google Street View image to the closest optimum view capture by the dashcam.



Figure 27. AOI_2 in Google Street View.



Figure 28. AOI_3 in Google Street View.



Figure 29. AOI_4 in Google Street View.



Figure 30. AOI_5 in Google Street View.

6. DISCUSSION

6.1 Vehicle Dashcams: capturing roadside vegetation imagery

The results demonstrate that vehicle dashcams, when used as a sensor, can facilitate the data collection of repeat roadside vegetation imagery. A dashcam enlists the opportunity to capture high resolution, repeat imagery, at a road level view subsequently enabling specific object-based analysis on urban trees. Alternative, ground level datasets have been used in urban vegetation studies, for example Google Street View imagery (Wegner et al., 2016, Branson et al., 2018), however, the frequency of this dataset is insufficient for phenological studies. Dashcams however offer a relatively inexpensive method of remote sensing, offering a dynamic and continuous view of the urban environment, with high temporal frequency possible.

This method of data capture does however entail some limitations. Primarily, the digital camera of the dashcam constrains analysis to the RGB channels of the images, in absence of the near infrared or any other bands. The NDVI has been the primary metric used in satellite remote sensing studies investigating plant phenology (Stockli and Vidale, 2004), with the Gcc thought to provide to produce the best quality results when using digital RGB cameras (Sonnentag et al., 2012). When comparing the two indices, vegetation measures of phenology derived from digital cameras have been shown to highly correlate with the NDVI derived a multispectral satellite sensor (St. Peter et al., 2018). Nonetheless, use of the Gcc in this study based on images derived from the dashcam, have observed an increasing greenness trend overall.

Unlike the use of fixed digital cameras for repeat photography (Richardson et al., 2007, Crimmins and Crimmins, 2008), the variability in the dashcam's platform (vehicle) position and the dynamic nature of an urban environment presents a number of challenges. In some instances, the AOI i.e. a tree was obscured, in part or entirely, by passing or parked vehicles. This either resulted in omitting that particular AOI on the given day, in the instance of a parked vehicle. When a vehicle was passing, where possible, an AOI was still extracted, however not at what was determined its most optimum view. Similarly, this occurred when the platform was required to seek a different road position, for instance, passing cyclists, passing buses at bus stops, and avoiding road works. This has little to no effect on the later analysis due to the benefits of the object detection methods. The dashcam was also subject to some other physical movements. For example, on three occasions during the data capture period of this study, the dashcam became disconnected due to the power cable dislodging when crossing a speed bump.

When comparing Google Street View imagery to the images captured by the dashcam, it provides a very similar view. The field of view is also comparably wider due to the 360-camera used, in addition to the camera being roof mounted. Subsequently, the images capture more imagery compared to the dashcam, potentially including trees which could not necessarily feature in the same field of view as the dashcam at a given location. Google Street View also offers a multi view of a given tree, which a dashcam does not provide as the camera is forward facing only. The image resolution is also higher than that of the dashcam, providing greater vegetation detail enabling easier species identification, and the positioning on Google Maps can also provide precise longitude and latitude coordinate when viewing the tree in the aerial imagery. Despite a useful source of data for developing a street tree inventory (Wegner et al., 2016), it is limited by the lack of temporal range and repeatability. However, it could prove a valuable secondary dataset for confirming species, obtaining the trees' position in space, and using Google Street View as part of the deep learning detector training dataset.

6.2 YOLOv3: Detecting roadside trees at different phenological stages

Overall, the detector obtained a mAP accuracy of 87.32% with a 50% IoU threshold and a 70.73% mAP accuracy with a 75% IoU threshold. However, the detector with a reduced training dataset produced better accuracy results in comparison, obtaining a mAP of 94.82% with the IoU threshold set at 50% and a 74.99% mAP accuracy the 75% IoU threshold. However, this does mean the detector trained with a reduced dataset did observe a greater difference when increasing the IoU threshold, with mAP decrease of 19.83% whereas the original detector only observed a decrease of 16.59% when increasing the IoU threshold. These findings suggest that further research is required in order to establish an optimum training dataset as to improve efficiency in the methodology, from the dashcam imagery obtained through to the end greenness metric obtained. It is possible that by default, the random reduction in the training dataset resulted in a better proportional representation of the trees at different stages of phenology, which had been recognised as a fault in the visual analysis of the first detector, thus improving bounding box results. Nonetheless, it is worth considering that deep learning algorithms such as YOLO are black-box models, in addition to the fact the detector outputs compared to the ground truths was just a sample.

Ultimately, the application of the YOLOv3 detection algorithm for this project is to repeatedly detect the same tree from multiple images over time. Despite relatively good results observed overall, an improved mAP accuracy for the detector is desirable to produce more consistent results. Precise intersection of the ground truth bounding boxes and the detection bounding boxes is desirable in this instance as it is a determining factor in data consistency when deriving the greenness metric, as a key element to this study is the temporal analysis in comparing specific trees. Inconsistency in the detector precision, results in error when deriving the greenness metric as this analysis subsequently not be made to the same portion of the tree. However, when used on dashcam imagery, this deep learning approach does enable object-based analysis, which is not a viable option from traditional satellite remote sensing techniques, either due to resolution limitations or the financial viability. Additionally, the deep learning approach largely reduces positioning issues related to the sensor itself. The nature of the dashcam imagery and the urban environment means data is subject to changing views, for example dashcam movement and changing road position of the platform (vehicle) due to overtaking cyclists, avoiding roadworks etc, results in a specific tree is not repeatedly captured from the same position. However, by obtaining the bounding box coordinates of a given tree, it can be extracted, and any further analysis made solely on the tree, thus the 'optimum view' of a tree is not a necessity.

Further research should be made into comparing the YOLOv3 detector to other deep learning algorithms, such as CNNs, and make consideration to the trade-offs between detector inference speed and the accuracy for the purpose of this application.

A limitation of using deep learning algorithms is the requirement of image annotations to produce a training dataset for the detector. The literature suggests that in object detection-based research, pre-existing labelled image catalogues are frequently used either for training or testing, for example the Pasadena Urban Trees dataset (Wegner at al., 2016). However, unlike work demonstrated by Wegner's et al., (2016) study into deep learning for tree detection, similarly at street level, this project aimed to detect trees at across a section of the phenological cycle, capturing trees during late winter through to early summer as part of the methodology. Hence there was a requirement that the training dataset used to train the

YOLOv3 detector included annotated images across the phenological time frame of interest, both trees with leaf off and leaf on, and the growth stages between this.

However, the performance of the detector on trees off was comparably weaker than when trees had greened up. It is possible that this is due to the added complexity of visible bare branches, which requires more learning for the detector to work effectively. There was an intention to incorporate an improved training dataset in this project's methodology, due to computational limitations and time constraints of this project, this was not possible. The greenness results established in this study could be used to inform the dates of green-up in different species, and thus a better representation of phenological variation could be obtained and incorporated into an improved detector. As per the findings discussed above, developing an optimum model will require balance in the training dataset size in addition to proportional representation of tree phenology. Additionally, when considering the potential reproducibility of the detector in other geographical locations, increased tree species heterogeneity is an advised area of improvement for the training data.

6.3 Cloud Computing: Google Colab

Google Colab, the free cloud computing platform, was utilised for processing. This web-based platform facilitates coding in the Python programming language. It has a user friendly, graphical user interface in the style of Jupyter Notebook, making it simplistic to use. A number of Python libraries are pre-installed making set-up of many projects quick and straightforward, although specific libraries that are not Python standards require installation. Advantageously, this platform can directly mount to the Google Drive, enabling entirely cloud based processing and connection to files, although the free storage on Google Drive is limited to 15GB but more can be purchased. Google Colab also enables users to access Graphical Processing Units (GPU) on Google's remote computers, which can be used to speed up processing for more computationally demanding work, such as those using deep learning methods. However, Google Colab is limited to a 12-hour maximum runtime, which if processing exceeds this, the runtime disconnects, and variables are lost. This became problematic for this project, when an attempt was made to advance and improve the detector by altering and increasing the training dataset. However, the computational time required to train the detector with this increased training dataset exceeded the 12- hour runtime and thus could not be achieved. The platform also has a user max usage limit, which also constrained the speed at which processing could be completed.

6.4 Deriving Greenness

Repeat images from dashcam imagery can be used to successfully capture the green-up of urban trees when applying the Green Chromatic Coordinate metric. Unlike visual observations of phenological events throughout seasons, the collection of repeat digital imagery allows for a quantitative analysis in capturing a change in greenness from the leaf-off to leaf-on change to trees.

However, the accuracy results of the YOLOv3 object detector algorithm had secondary impacts on the image analysis. Due to the primary variable of phenology studies being a change in time, the area or region of interest should be kept consistent. A complexity of using dashcam imagery for vegetation-based phenology studies, which is not observed in existing repeat photography research, is the challenge of non-static imagery. The approach applied aimed to derive a consistently precise bounding box of a given tree, so the greenness metric

could be applied unfailing without a change in the object of analysis. A region of interest approach was used to remove redundant data which featured within a bounding box of a tree for example, buildings, bins, and vehicles. However, the lack of consistency in bounding boxes, particularly for the trees in AOI_5 which had smaller crowns and a longer trunk in relation to the depth of the crown, resulted in inconsistencies when plotting greenness over time. Subsequently, the green-up trend was best visible in the pure RGB images for AOI_5b, with the conversion to the HSV colour space, and cropping of the image to a smaller ROI skewing the data.

As discussed, unlike existing repeat photography research for vegetation phenology, using a dashcam as a sensor for capturing imagery is subject to several challenges. Exploiting the data captured by a dashcam, typically installed by drivers to provide evidence in the instance of an accident, will generally only capture imagery in a single instance when passing an AOI while driving by. Whereas the fixed digital cameras can be set up to take photographs at specified intervals across a day. Subsequently, having only a single window of capture removes potential of obtaining an image during the day which is least impacted by the changes in scene illuminance. Additionally, the nature of the dashcam data means it is subject to interruptions such as traffic, roadworks, road crossings etc, hence even if the start time of capture is maintained, the moment at which imagery of a given AOI is captured, may differ. With the research methodology designed in consideration of the potential use of Citizen Science volunteered dashcam imagery, it is also worth considering the probable commuting hour of many drivers. The data collected for this project was largely captured during a morning commute, and subsequently there was significant change in the scene illuminance as a result of changing daylight, particularly up until the transition from daylight saving hours to British Summer Time. As such, the HSV colour space was used for analysis (Yang et al., 2015), which was effective but more so in smoothing the Gcc coordinates after the onset of greenness. It is possible the it was less effective at smoothing the data prior to this point as the images were affected by background features visible through tree branches, that were not visible once in leaf. Applying a three-day rolling average to the Gcc coordinates (Sonnentag et al., 2012) to reduce the effects of variance in scene illuminance was broadly effective in smoothing the data and enabled a clearer visual on the trends observed.

The Gcc results indicated green-up occurring between the 14th April and 5th May 2020 across the four species examined. This is broadly in keeping with the dates recorded by Citizen Scientists as part of the Woodland Trust's Nature's Calendar project where 12 date entries of first leaf on trees within the region occurred between 13th March and 24th April. However, the volume of this volunteered data is insufficient to draw absolute conclusions from, in addition to a lack of species representation. It is recommended that if this methodology is repeated, then field observations should be made for validation purposes.

6.5 Citizen Science: volunteered dashcam footage for vegetation monitoring

The AA, a British motoring association, found in a 2019 driver poll of their members, that 24% of respondents already own a dashcam, with an additional 18% stating they were seriously considering buying one. With 38.4 million licensed vehicles in Great Britain (Department for Transport, 2019), the dashcam user figures could be used to estimate a potential 16.1 million dashcams in vehicles capturing roadside imagery across the 396,700 km of paved roads over England, Scotland, and Wales (ONS, 2020). This volume of Big Data

is unmatched at such a spatial resolution and frequency, offering huge scope to attain geospatial coverage on a national level. The data collected by a subset of these dashcam users who make repeated commuter journeys from home to the workplace, would be of particular interest to phenology studies due to the high repeatability in terms of passing a specific location on a given route.

Crimmins and Crimmins (2008) have highlighted the opportunities for engaging the public, by encouraging sharing of repeat photography for capturing phenological events. They suggest that a website would allow for immediate remote access for interpretation, through automatic uploads of images. Also highlighting that in turn, the website would serve as a platform for a long-term archive of time-series photographs. OpenStreetCam (openstreetcam.org) is a platform for collecting and providing free and open street level imagery, with mappers able to share dashcam footage and other street level photography. The platform is available as an app (both on Android and iOS devices) enabling mappers to capture and upload imagery directly. This also enables linkage between internet of things devices, particularly with dashcams which are Bluetooth enabled, hence this platform would be a great asset in facilitating Citizen Science contributions to aid with the evolution this project. At present, however, OpenStreetMap does not facilitate batch download of data which would be a limitation if deriving such imagery for research applications.

Treezilla (treezilla.org) is a Citizen Science project which takes a collaborative approach to mapping, measuring, and monitoring trees across the United Kingdom and the Republic of Ireland. The project aims to encourages a number of organisations, businesses, local authorities, and members of the public to contribute to the dataset of trees, with a focus on urban trees, to gather more information and insights into ecosystem services. A challenge recognised by the project relates to being able to co-ordinate Citizen Scientists at scale (Goodenough et al., 2017). Encouraging the dashcam users to volunteer their footage, as Citizen Scientists, would aid with the scale challenges presented by field-based projects such as Treezilla. Crooks et al., (2016) have recognised that sensor-derived content from crowd-sourced data has great potential in bettering the understanding of city systems, with Part et al., (2016) acknowledging that sharing dashcam footage will greatly extend urban surveillance. The methodology developed in this project opens doors towards an automated process for deriving green-up metrics for urban trees and indicates scope to ascertain more metrics relating to other aspects of vegetation such as species, health, and size.

6.6 Applications

The methodology developed in this project has showcased a dynamic sensor for repeat photography, applied to capturing images of vegetation, in an urban setting, for phenological analysis. Primarily, insight into phenological patterns is a key indicator to measure the effects of climate change on vegetation and the growing season (Chielewski and Rotzer, 2001). It is recognised that phenological transitions influence ecosystem productivity (Richardson et al., 2010), and subsequently impacting the provision of ecosystem services in urban environments which are of huge value to urban populations (Bolund and Hunhammer, 1999). Significantly, research suggests that the United Nation's Sustainable Development Goals are both dependent on, and targets contributed to by, ecosystem services (Wood et al., 2018). There is limited understanding between the interactions of climate and plant phenology specifically in urban settings (see Luo et al., 2007). The synergies between phenology, plant productivity, the provision of ecosystem services, and sustainable urban environments, highlights the importance of a reliable and consistent dataset which subsequently could be used to inform urban planning decision makers.

Existing methods of repeat photography are lacking geospatial coverage. The methodology demonstrated here shows scope to expand geospatial coverage when obtaining repeat photography of vegetation in urban settings. Stages of automation in processing also show the potential capability of analysing big data and thus offering insights of vegetation greening across multiple locations. Results suggest that this method could be an input dataset required by urban planners for decision making, in addition to acting as a suitable measure for key actions as part of the SDG targets. On a local scale, near-to-real time vegetation insights could also be of benefit for local councils and authorities in general vegetation management. Further methodological development is required to offer a fully automated process, from capture, to processing, through to obtaining the phenological metric.

6.7 Further Research Scope

This research project has demonstrated that dashcams are a suitable sensor for capturing roadside vegetation imagery. There are several immediate methodological improvements that would be of benefit to explore, in addition to some new research avenues to investigate. Initially, an improved detector based on a more systematically selected training dataset should be tested. This should be designed with the intention of better representing vegetation throughout the phenological transition between leaf-off and green-up. This is based on the detector's poorer performance in outputting bounding boxes on trees during leaf-off. Other recommendations based on findings and areas of potential:

- The comparison of the efficiency and effectiveness of YOLOv3 against other deep learning algorithms, such as the traditional Convolutional Neural Network.
- A continuation of the capture period into the autumn season would enable an investigation into capturing the leaf off, senescence.
- Further exploration into automating big datasets of dashcam imagery, with regards to a crowdsourced direction of data collection.
- Investigation into the feasibility of advancing computer vision techniques for analysing dashcam footage for identifying individual tree species.
- Application of computer vision techniques for identifying roadside invasive species to aid with the management of non-native plants such as Himalayan Balsam and Japanese Knotweed.
- Experimenting with dual angle dashcam usage could also be explored as a method of 3D modelling urban vegetation infrastructure.
- Exploring the capacity of dashcam footage as a method for identifying stresses in vegetation. For example, the effects of disease, fire, drought etc.
- Investigation into using dashcams for monitoring the success of roadside reforestation efforts in rural environments.

7. CONCLUSION

The focal aspect of this dissertation has been the use of vehicle dashcam imagery for obtaining phenological insights of trees, specifically the greening up stage of the cycle, in an urban environment. The data source used in this project shows scope to expand geospatial coverage when obtaining repeat photography of vegetation in urban settings. The results suggest that dashcam imagery could offer a solution to the data problem for phenological studies, associated with spatial and temporal frequency, when using satellite remote sensing techniques, and the geospatial coverage limitation faced by using fixed digital cameras for repeat photography.

The image processing has progressed in the form of an object-based analysis, experimenting with the deep learning YOLOv3 object detection algorithm, with outputs indicating significant suitability for automating processing. The methodology shows promise in developing towards fully automated steps for analysis, and as such, could be used to process Big Data. Increasing numbers of dashcam users on roads, coupled with the increasing connectivity of internet of things devices highlights the potential of sourcing volunteered dashcam imagery from Citizen Scientists. In the long term, increased geospatial coverage of urban tree data capture for monitoring the timing of phenological events would aid with gaining a better understanding of ecosystem services and the interactions between plant productivity and climate change. Additionally, this object-based approach offers specific information of individual trees, which could be a valuable source of intel for urban planners and decision makers, particularly in the context of measuring the UN SDG targets.

A major challenge faced by using the a dashcam as the sensor and source of imagery is related to the effects of scene illuminance variation. The results of the methods used to supress this issue do suggest success. However, it is thought that the accuracy of the object detector bounding boxes did affect these results. The YOLOv3 detector results performed better on trees with leaf on, and further research should explore developing a training dataset with a better representation of trees at different stages of phenology. The unexpected accuracy results produced when comparing the original training dataset for the detector to the reduce training dataset brings the question, of what is the optimum detector training dataset for this application? This question is of importance when considering the efficiency of reproducibility and generalisability, if expanding analysis in a geospatial context. Therefore, it is recommended that further research is used to find an optimum training dataset size.

8. REFERENCES

Aihara, K., Imura, H., Piao, B., Takasu, A. and Tanaka, Y., 2017. Mobile crowdsensing to collect road conditions and events. In Smart Sensors at the IoT Frontier (pp. 271-297). Springer, Cham.

Akbari, H., Pomerantz, M. and Taha, H., 2001. Cool surfaces and shade trees to reduce energy use and improve air quality in urban areas. *Solar energy*, *70*(3), pp.295-310.

Anton, V., Hartley, S., Geldenhuis, A. and Wittmer, H.U., 2018. Monitoring the mammalian fauna of urban areas using remote cameras and citizen science. Journal of Urban Ecology, 4(1), p.juy002.

Berland, A., Roman, L.A. and Vogt, J., 2019. Can Field Crews Telecommute? Varied Data Quality from Citizen Science Tree Inventories Conducted Using Street-Level Imagery. Forests, 10(4), p.349.

Bolton, D.K., Gray, J.M., Melaas, E.K., Moon, M., Eklundh, L. and Friedl, M.A., 2020. Conti nental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sens. Environ*, *240*, p.111685.

Bolund, P. and Hunhammar, S., 1999. Ecosystem services in urban areas. *Ecological economics*, *29*(2), pp.293-301.

Bothmann, L., Menzel, A., Menze, B.H., Schunk, C. and Kauermann, G., 2017. Automated processing of webcam images for phenological classification. *PloS one*, *12*(2).

Branson, S., Wegner, J.D., Hall, D., Lang, N., Schindler, K. and Perona, P., 2018. From Google Maps to a fine-grained catalog of street trees. *ISPRS Journal of Photogrammetry and Remote Sensing*, *135*, pp.13-30.

Bresilla, K., Perulli, G.D., Boini, A., Morandi, B., Corelli Grappadelli, L. and Manfrini, L., 2019. Single-shot convolution neural networks for real-time fruit detection within the tree. Frontiers in plant science, 10, p.611.

Cai, Z., Melaas, E., Friedl, M.A., Eklundh, L. and Jönsson, P., 2018. A Method for Robust Estimation of Vegetation Seasonality from Landsat and Sentinel-2 Time Series Data. *Remote Sensing; 4*, *10*.

Chaves-González, J.M., Vega-Rodríguez, M.A., Gómez-Pulido, J.A. and Sánchez-Pérez, J.M., 2010. Detecting skin in face recognition systems: A colour spaces study. *Digital Signal Processing*, *20*(3), spp.806-823.

Chmielewski, F.M. and Rötzer, T., 2001. Response of tree phenology to climate change across Europe. *Agricultural and Forest Meteorology*, *108*(2), pp.101-112.

Connors, J.P., Lei, S. and Kelly, M., 2012. Citizen science in the age of neogeography: Utilizing volunteered geographic information for environmental monitoring. Annals of the Association of American Geographers, 102(6), pp.1267-1289.

Conrad, C.C. and Hilchey, K.G., 2011. A review of citizen science and community-based environmental monitoring: issues and opportunities. Environmental monitoring and assessment, 176(1-4), pp.273-291.

Costanza, R., d'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'neill, R.V., Paruelo, J. and Raskin, R.G., 1997. The value of the world's ecosystem services and natural capital. *nature*, *387*(6630), pp.253-260.

Crimmins, M.A. and Crimmins, T.M., 2008. Monitoring plant phenology using digital repeat photography. *Environmental management*, *41*(6), pp.949-958.

Crooks, A.T., Croitoru, A., Jenkins, A., Mahabir, R., Agouris, P. and Stefanidis, A., 2016. User-generated big data and urban morphology. *Built Environ*, *42*(3), pp.396-414.

Crown, C.A., Greer, B.Z., Gift, D.M. and Watt, F.S., 2018. Every Tree Counts: Reflections on NYC's Third Volunteer Street Tree Inventory. Arboriculture & Urban Forestry, 44(2).

Dallimer, M., Tang, Z., Gaston, K.J. and Davies, Z.G., 2016. The extent of shifts in vegetation phenology between rural and urban areas within a human-dominated region. *Ecology and Evolution*, *6*(7), pp.1942-1953.

Department for Transport, 2019. Vehicle Licensing Statistics: 2019 Quarter 1 (Jan – Mar). Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/812253/vehicle-licensing-statistics-january-to-march-2019.pdf (Accessed 3rd June 2020).

Department for Transport, 2020. Road Lengths in Great Britain 2019. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/860685/road-lengths-in-great-britain-2019.pdf (Accessed 3rd June 2020).

Dutta, P., Aoki, P.M., Kumar, N., Mainwaring, A., Myers, C., Willett, W. and Woodruff, A., 2009, November. Common sense: participatory urban sensing using a network of handheld air quality monitors. In Proceedings of the 7th ACM conference on embedded networked sensor systems (pp. 349-350).

Escobedo, F.J., Kroeger, T. and Wagner, J.E., 2011. Urban forests and pollution mitigation: Analyzing ecosystem services and disservices. *Environ Pollut*, *159*(8-9), pp.2078-2087.

Fang, F., McNeil, B.E., Warner, T.A., Maxwell, A.E., Dahle, G.A., Eutsler, E. and Li, J., 2020. Discriminating tree species at different taxonomic levels using multi-temporal WorldView-3 imagery in Washington DC, USA. *Remote Sens. Environ*, *246*, p.111811.

Frolking, S., Milliman, T., McDonald, K., Kimball, J., Zhao, M. and Fahnestock, M., 2006. Evaluation of the SeaWinds scatterometer for regional monitoring of vegetation phenology. *Journal of Geophysical Research: Atmospheres*, *111*(D17).

Gallo, T. and Waitt, D., 2011. Creating a successful citizen science model to detect and report invasive species. BioScience, 61(6), pp.459-465.

Goodenough, J., Ansine, J. and Wheeler, P., 2017. Using a citizen science tool to model the health benefits of roadside trees.

Graham, E.A., Riordan, E.C., Yuen, E.M., Estrin, D. and Rundel, P.W., 2010. Public Internet-connected cameras used as a cross-continental ground-based plant phenology monitoring system. *Glob Chang Biol, 16*(11), pp.3014-3023.

Hara, Y., Uchiyama, A., Umedu, T. and Higashino, T., 2018, October. Sidewalk-level People Flow Estimation Using Dashboard Cameras Based on Deep Learning. In *2018*

Eleventh International Conference on Mobile Computing and Ubiquitous Network (ICMU) (pp. 1-6). IEEE.

Hinton, G.E., Osindero, S. and Teh, Y.W., 2006. A fast learning algorithm for deep belief nets. Neural computation, 18(7), pp.1527-1554.

Itakura, K. and Hosoi, F., 2020. Automatic Tree Detection from Three-Dimensional Images Reconstructed from 360° Spherical Camera Using YOLO v2. Remote Sensing, 12(6), p.988.

Jiang, Q., Kresin, F., Bregt, A.K., Kooistra, L., Pareschi, E., Van Putten, E., Volten, H. and Wesseling, J., 2016. Citizen sensing for improved urban environmental monitoring. Journal of Sensors, 2016.

Jollymore, A., Haines, M.J., Satterfield, T. and Johnson, M.S., 2017. Citizen science for water quality monitoring: Data implications of citizen perspectives. Journal of Environmental Management, 200, pp.456-467.

Keenan, T.F., Darby, B., Felts, E., Sonnentag, O., Friedl, M.A., Hufkens, K., O'Keefe, J., Klosterman, S., Munger, J.W., Toomey, M. and Richardson, A.D., 2014. Tracking forest phenology and seasonal physiology using digital repeat photography: a critical assessment. *Ecological applications*, *24*(6), pp.1478-1489.

Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

Kruger, L.E. and Shannon, M.A., 2000. Getting to know ourselves and our places through participation in civic social assessment. Society & Natural Resources, 13(5), pp.461-478.

LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. nature, 521(7553), pp.436-444.

LeCun, Y., Kavukcuoglu, K. and Farabet, C., 2010, May. Convolutional networks and applications in vision. In Proceedings of 2010 IEEE international symposium on circuits and systems (pp. 253-256). IEEE.

Leonardi, C., Cappellotto, A., Caraviello, M., Lepri, B. and Antonelli, F., 2014, October. SecondNose: an air quality mobile crowdsensing system. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (pp. 1051-1054).

Li, W., Fu, H., Yu, L. and Cracknell, A., 2017. Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. Remote Sensing, 9(1), p.22.

Luo, Z., Sun, O.J., Ge, Q., Xu, W. and Zheng, J., 2007. Phenological responses of plants to climate change in an urban environment. *Ecological Research*, *22*(3), pp.507-514.

Lye, G.C., Osborne, J.L., Park, K.J. and Goulson, D., 2012. Using citizen science to monitor Bombus populations in the UK: nesting ecology and relative abundance in the urban environment. Journal of Insect Conservation, 16(5), pp.697-707.

Maisonneuve, N., Stevens, M., Niessen, M.E., Hanappe, P. and Steels, L., 2009. Citizen noise pollution monitoring.

McCaffrey, R.E., 2005. Using citizen science in urban bird studies. Urban habitats, 3(1), pp.70-86.

McPherson, G., Simpson, J.R., Peper, P.J., Maco, S.E. and Xiao, Q., 2005. Municipal forest benefits and costs in five US cities. Journal of forestry, 103(8), pp.411-416.

Mills, S.J., Castro, M.P.G., Li, Z., Cai, J., Hayward, R., Mejias, L. and Walker, R.A., 2010. Evaluation of aerial remote sensing techniques for vegetation management in power-line corridors. IEEE Transactions on Geoscience and Remote Sensing, 48(9), pp.3379-3390.

Minh, D.H.T., Ienco, D., Gaetano, R., Lalande, N., Ndikumana, E., Osman, F. and Maurel, P., 2018. Deep recurrent neural networks for winter vegetation quality mapping via multitemporal SAR Sentinel-1. IEEE Geoscience and Remote Sensing Letters, 15(3), pp.464-468.

Mohan, A., Gauen, K., Lu, Y.H., Li, W.W. and Chen, X., 2017, May. Internet of video things in 2030: A world with many cameras. In *2017 IEEE International Symposium on Circuits and Systems (ISCAS)* (pp.1-4). IEEE.

Morris, D.E., Boyd, D.S., Crowe, J.A., Johnson, C.S. and Smith, K.L., 2013. Exploring the potential for automatic extraction of vegetation phenological metrics from traffic webcams. *Remote Sensing*, *5*(5), pp.2200-2218.

Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S. and Crowston, K., 2012. The future of citizen science: emerging technologies and shifting paradigms. *Front Ecol Environ*, *10*(6), pp.298-304.

Nowak, D.J. and Crane, D.E., 2002. Carbon storage and sequestration by urban trees in the USA. *Environ Pollut*, *116*(3), pp.381-389.

Nowak, D.J. and Greenfield, E.J., 2012. Tree and impervious cover change in US cities. Urban Forestry & Urban Greening, 11(1), pp.21-30.

Padilla, R., Netto, S.L. and da Silva, E.A., 2020, July. A Survey on Performance Metrics for Object-Detection Algorithms. In *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)* (pp. 237-242). IEEE.

Pande-Chhetri, R., Abd-Elrahman, A., Liu, T., Morton, J. and Wilhelm, V.L., 2017. Object-based classification of wetland vegetation using very high-resolution unmanned air system imagery. *Euro. J. Remote Sens*, *50*(1), pp.564-576.

Park, S., Ilincai, E.S., Oh, J., Kwon, S., Mizouni, R. and Lee, U., 2017, May. Facilitating Pervasive Community Policing on the Road with Mobile Roadwatch. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 3538-3550).

Park, S., Kim, J., Mizouni, R. and Lee, U., 2016, May. Motives and concerns of dashcam video sharing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 4758-4769).

Pastick, N.J., Dahal, D., Wylie, B.K., Parajuli, S., Boyte, S.P. and Wu, Z., 2020. Characterizing Land Surface Phenology and Exotic Annual Grasses in Dryland Ecosystems Using Landsat and Sentinel-2 Data in Harmony. *Remote Sensing*, *12*(4), p.725.

Redmon, J. and Farhadi, A., 2018. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

Richardson, A.D., Andy Black, T., Ciais, P., Delbart, N., Friedl, M.A., Gobron, N., Hollinger, D.Y., Kutsch,W.L., Longdoz, B., Luyssaert, S. and Migliavacca, M., 2010. Influence of spring and autumn phenological transitions on forest ecosystem productivity. *Philos T R Soc B*, *365*(1555), pp.3227-3246.

Richardson, A.D., Jenkins, J.P., Braswell, B.H., Hollinger, D.Y., Ollinger, S.V. and Smith, M.L., 2007. Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. *Oecologia*, *152*(2), pp.323-334.

Richardson, A.D., Keenan, T.F., Migliavacca, M., Ryu, Y., Sonnentag, O. and Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agr Forest Meteorol*, *169*, pp.156-173.

Roman, L.A., Scharenbroch, B.C., Östberg, J.P., Mueller, L.S., Henning, J.G., Koeser, A.K., Sanders, J.R., Betz, D.R. and Jordan, R.C., 2017. Data quality in citizen science urban tree inventories. Urban Forestry & Urban Greening, 22, pp.124-135.

Safonova, A., Tabik, S., Alcaraz-Segura, D., Rubtsov, A., Maglinets, Y. and Herrera, F., 2019. Detection of fir trees (Abies sibirica) damaged by the bark beetle in unmanned aerial vehicle images with deep learning. Remote Sensing, 11(6), p.643.

Santos, A.A.D., Marcato Junior, J., Araújo, M.S., Di Martini, D.R., Tetila, E.C., Siqueira, H.L., Aoki, C., Eltner, A., Matsubara, E.T., Pistori, H. and Feitosa, R.Q., 2019. Assessment of CNN -Based Methods for Individual Tree Detection on Images Captured by RGB Cameras Attached to UAVs. *Sensors*, *19*(16),p.3595.

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A.M., Friedl, M., Braswell, B.H., Milliman, T., O'Keefe, J. and Richardson, A.D., 2012. Digital repeat photography for phenological research in forest ecosystems. *Agricultural and Forest Meteorology*, *152*, pp.159-177.

St Peter, J., Hogland, J., Hebblewhite, M., Hurley, M.A., Hupp, N. and Proffitt, K., 2018. Linking phenological indices from digital cameras in Idaho and Montana to MODIS NDVI. *Remote Sensing*, *10*(10), p.1612.

Stöckli, R. and Vidale, P.L., 2004. European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset. *International Journal of Remote Sensing*, *25*(17), pp.3303-3330.

The AA, 2019. Highlights from our 2019 Driver Poll surveys. Available at: https://www.theaa.com/about-us/public-affairs/aa-populus-driver-poll-summaries-2019#june2019 (Accessed 3rd June 2020).

Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E. and Liang, Z., 2019. Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Comput Electron Agric*, *157*, pp.417-426.

Torres-Sánchez, J., López-Granados, F. and Peña, J.M., 2015. An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. Computers and Electronics in Agriculture, 114, pp.43-52.

United Nations Department of Economic and Social Affairs (2019) World Urbanisation Prospects: The 2018 Provision. United Nations, New York. Available at: https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf (Accessed 26th May 2020).

Vogt, J., Gillner, S., Hofmann, M., Tharang, A., Dettmann, S., Gerstenberg, T., Schmidt, C., Gebauer, H., Van de Riet, K., Berger, U. and Roloff, A., 2017. Citree: A database supporting tree selection for urban areas in temperate climate. Landscape and Urban Planning, 157, pp.14-25.

Voulodimos, A., Doulamis, N., Doulamis, A. and Protopapadakis, E., 2018. Deep learning for computer vision: A brief review. Computational intelligence and neuroscience, 2018.

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

Williams, R.L., Stafford, R. and Goodenough, A.E., 2015. Biodiversity in urban gardens: Assessing the accuracy of citizen science data on garden hedgehogs. Urban ecosystems, 18(3), pp.819-833.

Windrim, L. and Bryson, M., 2018. Forest Tree Detection and Segmentation using High Resolution Airborne LiDAR. arXiv preprint arXiv:1810.12536.

Wolkovich, E.M., Cook, B.I., Allen, J.M., Crimmins, T.M., Betancourt, J.L., Travers, S.E., Pau, S., Regetz, J., Davies, T.J., Kraft, N.J. and Ault, T.R., 2012. Warming experiments underpredict plant phenological responses to climate change. *Nature*, *485*(7399), pp.494-497.

Wood, S.L., Jones, S.K., Johnson, J.A., Brauman, K.A., Chaplin-Kramer, R., Fremier, A., Girvetz, E., Gordon, L.J., Kappel, C.V., Mandle, L. and Mulligan, M., 2018. Distilling the role of ecosystem services in the Sustainable Development Goals. *Ecosystem services*, *29*, pp.70-82.

Xiao, Q. and McPherson, E.G., 2002. Rainfall interception by Santa Monica's municipal urban forest. *Urban ecosystems*, *6*(4), pp.291-302.

Yalcin, H., 2017, August. Plant phenology recognition using deep learning: Deep-Pheno. In *2017 6th International Conference on Agro-Geoinformatics* (pp. 1-5). IEEE.

Yan, J., Zhou, W., Han, L. and Qian, Y., 2018. Mapping vegetation functional types in urban areas with WorldView-2 imagery: Integrating object-based classification with phenology. *Urban Forestry & Urban Greening*, *31*, pp.230-240.

Yang, W., Wang, S., Zhao, X., Zhang, J. and Feng, J., 2015. Greenness identification based on HSV decision tree. *Information Processing in Agriculture*, *2*(3-4), pp.149-160.

Zhang, X., Friedl, M.A. and Schaaf, C.B., 2009. Sensitivity of vegetation phenology detection to the temporal resolution of satellite data. International Journal of Remote Sensing, 30(8), pp.2061-2074.

9. APPENDIX

9.1 Project Management Plan

Appendix Table 1. Project Management Plan: Gantt Chart

	March				April				May				June				July				August			
Task/ Time	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4
Background Reading																								
Data Collection																								
Introduction/ Aim & Objectives Chapters																								
Literature Review Chapter																								
Imagery Pre-processing																								
Metric Extraction: Code Development																								
Metric Extraction: Code Implementation																								
Plot Metric Results																								
Methodology Chapter																								
Results Chapter																								
Discussion Chapter																								
Concluding Chapters																								
Bibliography, Figures & Tables																								
Complete Draft																								
Final Edits & Review																								
Submission																								

9.2 Data Management Plan

1) Provide the title and briefly describe the aim and objectives of your MRes project.

Title:
TOWARDS AN AUTOMATED APPROACH FOR MONITORING VEGETATION GREEN-UP DYNAMICS USING VEHICLE DASHCAMS IN URBAN ENVIRONMENTS

Research Aim:
To develop automated steps for extracting vegetation green dynamics in an urban environment from high spatial resolution data.

Objectives:

1. Assess the suitability of vehicle dashcams for capturing roadside vegetation imagery
2. Evaluate the suitability of the YOLOv3 deep learning algorithm automatically detecting roadside trees at different phenological stages
3. Consider the appropriate image processing techniques for deriving a greenness metric
4. Consider the potential of Citizen Science volunteered dashcam footage for vegetation monitoring

2) What data will be produced? (Data types, format, standards, scale and method)

The data produced will be a 90 day series of dashcam imagery.

Although the road distance travelled on the survey transect for the data capture will be consistent, the journey length will vary subject to external factors such as traffic, hence footage may vary in length day to day. The default format of the dashcam imagery is 60 second clips in MP4 format.

Video files vary in size from 20MB to 160MB.

Processing will be facilitated in Python 3. A working Jupyter Notebook file and Google Colabratory notebook will be used to write the code.

Processing will generate the follows products:

- A manually annotated dataset of trees (126 training images)
- A CSV file containing annotation bounding box coordinates and class
- A set of 523 test images (manually extracted frames)
- Test images clipped to YOLOv3 output bounding box coordinates
- A greenness % graph of each AOI analysed

3) What metadata standards will you use? (Metadata content and format)

The DCC ISO 19115 metadata standard will be used as a template for recording metadata for collected data.
Recorded metadata will include:

- Filename
- Item type (MP4)

- Size
- Clip Length
- Frame Dimensions
- Data rate (kbps)
- Total Bitrate (kbps)
- Frame rate (fps)
- Creation Timestamp (DD/MM/YYYY 00:00:00)
- Location (long lat)
- Spatial Reference

The default file name for each clip includes the: date, time, and clip number for each created file. See below for an example:

200302_071152_015_FH.MP4

The dashcam records the longitude and latitude of each video which although this is not required for the proposed methodology, this data shall be recorded from the raw data files

4) How will your data be structured and stored? (Project storage)

All data will be download on to a laptop's hard drive from the micro SD memory card. MP4 files will be divided into week folders, i.e. VIDEO_W1 through to VIDEO_W13. These are saved on a personal laptop (unable to backup due to data size and upload speed issues)

Default file products from the dashcam stores the footage as 60 second clips. On the hard drive, within each week folder, day folders will be created one for each day across the capture period i.e. DAY_1 through to DAY_90. Here, all 60 second clips from each transect will be moved. This data is saved to a personal laptop and on an external hard drive.

The manually extracted AOI frames have been saved as png files, in an allocated folder AOI#_YY/MM/DD.png. These have been uploaded onto the OneDrive and downloaded onto the office desktop.

All processing outputs facilitated in the Google Colab cloud computing platform have been saved to a personal Google Drive account and downloaded onto the office desktop.

5) How will the data be shared during and after the project? (Access, data sharing and reuse)

During the project, data will be shared between the principal researcher and supervisors, in addition to CDT Directors if requested.

Data will initially be retained for the duration of the PhD (until October 2023) which will be a evolution of this MRes research project, towards the end of which data sharing will be reviewed.

6) Outline the approach to data selection and long-term preservation?

The code developed in the automation stage will be done so in Python in order to allow reproducibility in an open sourced software. Additionally, files will be converted to an open archival format for the same reason.

7) Who has responsibility for implementing the DMP and are resources required?

The lead researcher on this MRes project has responsibility for implementing the Data Management Plan.

Beyond the dashcam which facilitates data capture, a hard drive is required for data storage. Access to the University OneDrive for data back up is also required.

All developed code will be copied and saved to the OneDrive.

This equipment and storage is sufficient to undertake the study. If the project materialises differently, and the data exceeds the standard storage limit on the OneDrive, the relevant contact is known, and more storage can be requested.

END