

**MONITORING CROP HEALTH, GROWTH AND ITS STAND
COUNT ATTRIBUTES USING UAV BASED PRECISION
AGRICULTURE : A STUDY IN TROPICAL FARMLAND OF
THAILAND**

BY

SUMAN GHIMIRE

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE
(ENGINEERING AND TECHNOLOGY)
SIRINDHORN INTERNATIONAL INSTITUTE OF TECHNOLOGY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2017**

**MONITORING CROP HEALTH, GROWTH AND ITS STAND
COUNT ATTRIBUTES USING UAV BASED PRECISION
AGRICULTURE: A STUDY IN TROPICAL FARMLAND OF
THAILAND**

BY

SUMAN GHIMIRE

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE
(ENGINEERING AND TECHNOLOGY)
SIRINDHORN INTERNATIONAL INSTITUTE OF TECHNOLOGY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2017**

MONITORING CROP HEALTH, GROWTH AND ITS STAND COUNT ATTRIBUTES
USING UAV BASED PRECISION AGRICULTURE: A STUDY IN TROPICAL FARMLAND
OF THAILAND

A Thesis Presented

By
SUMAN GHIMIRE

Submitted to
Sirindhorn International Institute of Technology
Thammasat University
In partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE (ENGINEERING AND TECHNOLOGY)

Approved as to style and content by

Advisor


(Asst. Prof. Dr. Teerayut Horanont)

Committee Member


(Assoc. Prof. Dr. Bunyarit Uyyanonvara)

Committee Member and
Chairperson of Examination Committee


(Asst. Prof. Dr. Sarawut Ninsawat)

MAY 2018

Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisor Asst. Prof. Dr. Teerayut Horanont (School of Information, Computer, and Communication Technology (ICT), SIIT, Thammasat University), whose constant motivation and academic advice tremendously furnished my research work and as a result helped me accomplish my study. It is a great honor to work under his supervision.

I would also like to express my sincere appreciation to my committee members Assoc. Prof. Dr. Bunyarit Uyyanonvara (School of Information, Computer, and Communication Technology (ICT), SIIT, Thammasat University), and Asst. Prof. Dr. Sarawut Ninsawat (Remote Sensing and Geographic Information System FoS, Asian Institute of Technology, Thailand) for their endless support and constructive comments throughout the research work.

I am deeply grateful to Asst. Prof. Dr. Paiboon Sreearunothai (School of Bio-Chemical Engineering and Technology (BCET), SIIT, Thammasat University), who expertise has greatly benefitted me in the measurement of sensitivity and radiometric calibration of modified infrared camera using Black Comet C-200 TEC Stellar Net Spectrometer.

I would also like to thank RS and GIS Department at Asian Institute of Technology, Thailand for their assistance in field validation, using LI-COR LAI 2000 Plant Canopy Analyzer.

I owe my deepest gratitude to my parents, Mr. Rajendra Sharma Ghimire and Mrs. Sunita Ghimire, for their continuous encouragement and moral support to complete this report as well as my study at SIIT, Thammasat University, Thailand.

Lastly, I would like to thank Thammasat University: TU Basic and Applied Research Grant 2017 for their generous financial support, without which this research would not have been possible.

Abstract

MONITORING CROP HEALTH, GROWTH AND ITS STAND COUNT ATTRIBUTES USING UAV BASED PRECISION AGRICULTURE: A STUDY IN TROPICAL FARMLAND OF THAILAND

by

SUMAN GHIMIRE

Bachelor of Engineering in Geomatics Engineering, Kathmandu University, 2015

Master of Science in Engineering and Technology, SIIT, Thammasat University, 2018

Unmanned aerial vehicle (UAV) equipped with multispectral sensor has become an active research topic for crop health monitoring and has been widely used across many regions. However, the high cost associated with multispectral sensors suggests us to shift to a cheaper alternative in order to be implemented by average farmers from developing countries. We evaluated the feasibility of a lightweight (38 gm) mobius action camera with wide FOV in monitoring crop health through the removal of IR filter, and replacing it with Wratten 25A red filter. The research implements Structure from motion (SfM) for creating orthomosaic, and computes Normalized Difference Vegetation Index (NDVI). Finally, the results of NDVI from the modified camera was validated with the ground measurement of LAI carried using LI-COR LAI 2000, with a linear correlation that resulted in coefficient of determination of (R^2) 0.843. The result demonstrates that the modified camera is potential in agricultural health monitoring.

Likewise, this research also implemented SfM algorithm using UAV imagery supported with global positioning system (GPS) to generate multitemporal crop surface

models, which assessed crop growth throughout the growth season. Banana plantation which normally takes 9-12 months from sowing to harvesting the fruit, had been sowed on the first week of November, 2016 and harvested on the mid of September, 2017. The field data acquisition was performed three times, particularly on 25th January, 26th April & 16th September 2017 in an area of 0.186 sq.km, with the area's centroid coordinates at N $14^{\circ} 15.133'$ E $100^{\circ} 53.393'$ and Z 10 meters (WGS84). This study followed a methodology based on SfM to create a multitemporal surface models (DTM & DSM), followed by the difference method to generate canopy height model (CHM) which was used to assess crop growth. The growth ranged between 2.31-4.89 m for the period between Jan-Apr while the period between Apr-Sept demonstrated negative growth as a result of harvesting carried on September 10. The methodological framework adopted in this study will enable the spatial analysis of crop growth within banana plantation, enabling wide range of applications in the improvement of crop management.

Furthermore, several studies has been performed for object detection from ground view perspective and has been the key topic of interest for computer vision communities, however very less has been explored in detecting objects in an aerial imagery. The convolutional neural network implemented in this study was based open source tensorflow implementation of the darknet framework named, Darkflow, which has been modified to a near real-time multi-scale detector implementing YOLOv2 object detection model to improve the performance on aerial imagery. To detect the palm trees, the YOLO v2. Neural Net was modified and fine-tuned on our dataset consisting 255 images of palm trees, each of 4000 x 3000 resolution taken at 70m above the ground. The images were manually annotated which consisted of 595 annotations representing validation dataset, whereas 1000 annotations representing training dataset required for training and accuracy assessment of the ConvNet. The annotations were created in xml format, using a python script which takes the manual input from the user regarding the bounding box of the object to be detected. Likewise, we applied pre-trained weights and configuration files for the PASCAL VOC datasets, which was modified by

changing the no. of classes to 1 and the no. of filters in the last convolutional layer was modified to 30, which fits our purpose of detecting palm trees.

Likewise, the batch size was set to 64, subdivisions to 8, learning rate to 0.0001 and the datasets were initially trained until 6500 iterations on GPU server consisting of 32 GB of NVidia Tesla P100-SXM2. Finally, the precision and recall for our object detection model was observed to be 45.59% and 65.87%, making our total accuracy to 55.23%, which requires further improvements before it could be directly applied for palm tree counting purpose. These initial results demonstrate that provided a large training dataset (approx. 5 times more than current) with good quality labeled images and intensive training time, YOLO v2 net can accurately detect palm trees in our project area. As for now, the training spends more time per epoch resizing than training due to large resolution images, therefore, to increase performance the future prospects of our work would focus on tiling the images into multiple sections of 666 x 500 pixel blocks using OpenCV, before actually feeding the images for training which is expected to optimize the network.

Keywords: NDVI, Modified Infrared CMOS sensor, UAV Photogrammetry, Multi-temporal CSM, Crop Growth, Deep-Learning, Object Detection, Crop Counting.

Table of Contents

Chapter	Title	Page
	Signature Page	i
	Acknowledgements	ii
	Abstract	iii
	List of Figures	ix
	List of Tables	xi
1	Introduction	1
	1.1 Background	1
	1.2 Problem Identification	5
	1.3 Objectives	6
2	Literature review	7
3	Materials and Methods	19
	3.1 Monitoring crop health using modified infrared action camera	19
	3.1.1 Study Site	19

3.1.2	Modified Camera System	21
3.1.3	Methodological Framework	23
3.1.3.1	Geometric calibration of wide angle lens	24
3.1.3.2	Radiometric Calibration	25
3.1.3.3	Feature Detection, Matching and Alignment	26
3.1.3.4	3D Reconstruction, Textured Mesh, and Orthomosaic	27
3.1.3.5	NDVI	29
3.2	Monitoring crop growth using multitemporal crop surface model	30
3.2.1	Data Acquisition	30
3.2.2	Methodological Framework	31
3.3	Palm tree counting in an aerial imagery using deep learning	33
3.3.1	Training Dataset	33
3.3.2	Methodological Framework	34
3.3.2.1	Network Architecture	36
3.3.2.2	YOLO Object Detection Algorithm	37
4	Results and Discussions	42
4.1	Monitoring crop health using modified infrared action camera	42
4.1.1	Alignment Accuracy	42
4.1.2	Spectral calibration of the modified camera	43
4.1.3	NDVI	44

4.2 Monitoring crop growth using multitemporal crop surface model	50
4.2.1 SfM Accuracy	50
4.2.2 Surface Modeling	51
4.2.3 Crop Growth Statistics	52
4.3 Palm tree counting in an aerial imagery using deep learning	61
4.3.1 Accuracy of classification	61
4.3.2 Accuracy of Object Detection	62
5 Conclusion and Recommendations	66
References	68

List of Figures

Figure	Page
3.1 The study area in Pathumthani province, Thailand. The red polygon represents the distribution of banana farm while the yellow polygon represents the area for the validation of the NDVI values	19
3.2 (a) Unmanned aerial system design for field data collection, (b) Flight overview	19
3.3 Spectrometer (Black Comet C-200 TEC ranging from 200-1100nm)	21
3.4 Spectral response of CMOS based camera sensor ((LoopTechnology, 2009))	21
3.5 Transmission properties of the Wratten 25A Filter (red) ((Fastie, 2013))	22
3.6 Spectral sensitivity of mobius action camera illuminated with monochromatic light of 850 nm	22
3.7 Workflow for modified infrared CMOS sensor	23
3.8 Image Alignment where, the blue polygons on the top represents each camera positions.	26
3.9 The study area in Pathumthani province, Thailand represented by yellow polygons. The GCPs are represented by 7 small red markers, and the plots represented by six small 3x 3 m square plots.	29
3.10 Framework for crop growth monitoring	31
3.11 The study area consisting of palm plantations in Nakhon Pathom province, Thailand	33
3.12 Methodological workflow for palm tree counting using deep learning framework	35
3.13 Network Architecture	36

3.14	Object localization example, where the green polygon represents the ground labelled truth, while the red polygon represents the algorithm detected bounding box	39
4.1	Setup of the calibration targets for the spectral calibration of the lens	44
4.2	NDVI computation from modified infrared camera after calibration	45
4.3	Field measurements with LI-COR LAI 2000 Plant Canopy Analyzer	46
4.4	Scatterplot depicting the linear relationship between field estimate of LAI and NDVI result from modified infrared camera	47
4.5	Digital Terrain Model (DTM) Raster	52
4.6	Digital Surface Model (DSM) Raster	52
4.7	Orthomosaic, captured on January 25, 2017	53
4.8	Orthomosaic, captured on April 26, 2017	54
4.9	Orthomosaic, captured on September 16, 2017	54
4.10	Pictorial representation of black spots observed in Plot 2	57
4.11	Spatial representation of crop growth between the periods Jan-Apr, 2017	57
4.12	Spatial representation of crop growth between the periods Apr-Sept, 2017	58
4.13	Digital Elevation Model (Raster) of Bangkadi Campus Area	58
4.14	Assessing the height of car parking shed using SfM point clouds	59
4.15	Field measurements with a measuring tape for validation of the result	59
4.16	Scatterplot representing the loss function of the CNN model	62
4.17	Initial results of palm tree detection using Yolov2 detector	63

List of Tables

Tables	Page
3.1 Flight plan description	20
3.2 Camera Specification	21
3.3 Details of ground control points in decimal degree (WGS 84)	28
3.4 Description of Flight Plan	31
3.5 Accuracy comparison between different object detection models	38
4.1 Camera orientation & lens distortion coefficients	43
4.2 NDVI measurements from modified camera and field measurements of LAI for validation of the result	47
4.3 Uncertainties in camera's positional & orientation parameters	50
4.4 Crop growth statistics (Jan-Apr, 2017)	55
4.5 Crop growth statistics (Apr-Sept, 2017)	55

Chapter 1

Introduction

1.1 Background

Remote sensing technologies has been providing timely and accurate information related to crop health and productivity; and has been widely used in agricultural health monitoring (Atzberger (2013) at both regional and global scale since decades. One of the important characteristics for any agricultural health monitoring system is to timely disseminate information related to plant health, growth and yield to the farmers. With recent advancements in technology, light-weight airborne remote sensing (C. Zhang & Kovacs, 2012) equipped with multispectral sensors presents a unique advantage over traditional satellite borne images (Lamb & Brown, 2001) in terms of high spatial & temporal resolution with reduced effect of cloud cover during data acquisition. However, the associated costs with such multispectral sensors are generally high which suggests researchers to focus on cheaper alternatives that could be financially implemented by farmers from developing countries.

Numerous studies has been performed in the application of UAV based imagery in precision agriculture (Alexandridis et al., 2017; Allahyari, Mohammadzadeh, & Nastis, 2016; Shaw, Lark, Williams, Chadwick, & Jones, 2016). All these studies suggest in improving georeferencing, image mosaicking algorithms and more concrete and automated workflows, as well as directly involve farmers in different design phases for better data interpretation and to provide good services to the farmers. The remote sensing platforms including air and space borne suffers atmospheric scattering in the blue & green region, consequently it is recommended using larger wavelength such as red and NIR for the purpose of agricultural applications (Nijland et al., 2014). Therefore, an inexpensive digital camera setup, with internal infrared filter removed and replaced with a blue blocking filter allowed blue channel to record the NIR light

with red recorded in its original channel (Hunt et al., 2010; Zigadlo, Holden, Schrader, & Vogel, 2001), resulting in a promising technique for agricultural monitoring. In recent years, a lot of research have implemented such modification of standard RGB digital camera to near infrared for assessment of crop health (Hunt, Cavigelli, Daughtry, McMurtrey, & Walthall, 2005; Hunt et al., 2010; Rabatel, Gorretta, & Labbe, 2014), but very less has been explored in CMOS based action camera model (Ghazal, Khalil, & Hajjdiab, 2015; J. Wijitdechakul, S. Sasaki, Y. Kiyoki, & C. Koopipat, 2016). For instance, (Hunt et al., 2010) studied the ability of CCD based digital color infrared photograph in crop monitoring using UAV. Their results in terms of green normalized vegetation index (GNDVI) was found to have a good correlation with leaf area index (LAI) suggesting their approach to be potential in providing accurate information related to crop health. Similarly, (J. Wijitdechakul et al., 2016) demonstrates a dual action camera model: one normal RGB camera while the other IR filter removed modified camera in UAV platform for real-time agricultural area management using SPA process. Their system was able to detect the healthy vs non-healthy plantations, and notify farmers about the unhealthy plantation area for improved decision making in agricultural practices.

The action cameras offers a unique advantage in UAV photogrammetry due to its lightweight and low associated costs (Hastedt, Ekkel, & Luhmann, 2016). Significant number of research has been performed for testing the applicability of action cameras for UAV photogrammetry (Balletti, Guerra, Tsioukas, & Vernier, 2014; D'Agostino, Antuono, & Pepe; Hastedt et al., 2016). For instance, (Hastedt et al., 2016) highlights different camera calibration approaches on the GoPro Hero4 and evaluated the potential of a pre-correction in rectifying the initial distortion generated from the wide angle lens for the application in photogrammetric works. They also discussed on the challenges introduced by wide angle lens in terms of their short principal distance and high radiometric distortion, which causes diversion from the central projective model. An efficient method for image matching was applied (S. Agarwal et al., 2011), where they identified a small patches of features for each image using tree detection techniques

instead of matching all the image together which still preserved the features for structure from motion and significantly reduced the processing time. Similarly, linear time structure from motion was implemented for large scale image reconstruction (Wu, 2013b) by introducing preemptive feature matching, which reduced the image matching process by 95%, and also recovered good feature description match for the reconstruction process. Likewise, (D. Turner, Lucieer, & Watson, 2012) presented an approach for robust radiometric and geometric calibration to implement UAV photogrammetry using SfM algorithm for the generation of image orthomosaic. The images were processed to generate a 3D point clouds in arbitrary space coordinates, which was later transformed into global coordinate system using; a. direct camera EXIF file, b. ground control points. Thus, the point cloud was used to create DTM which was followed by the creation of orthomosaic. Their results demonstrated an absolute spatial accuracy of 65-120 cm with direct georeferencing method while 10-15 cm using GCP technique.

The banana plantation, originally from South-east Asian region is able to grow in variety of environments, and mostly favorable for hot & humid temperatures similar to Thailand; exhibits an important role as a staple diet and has been helping millions of farmers generate income. The conventional remote sensing technologies i.e. space borne systems offers nondestructive and remote estimation for crop monitoring; however, these systems are often limited to low spatial and temporal resolution making it less reliable for precision agriculture tool such as biomass estimation, early detection of disease and yield prediction (Chang, Jung, Maeda, & Landivar, 2017). Recent advancements in UAV and sensor platform presents a wide possibilities for overcoming the limitations of traditional remote sensing through its potential in generating high spatial & temporal resolution data (Wang Li et al., 2016; Malambo et al., 2018). The 3D point cloud of the scene generated using structure from motion (SfM) techniques presents an opportunity for reliable crop height determination, which aided through automated preprocessing & robust image matching approach, enables the use of

inexpensive consumer cameras to be embedded on UAV platform for the applications in mapping and precision agriculture.

An emerging remote sensing tool namely, light detection and ranging (LiDAR) enables the accurate representation of 3D properties of forest canopy and vegetation structure (Ghimire, Xystrakis, & Koutsias, 2017), and overcomes saturation problem due to its high penetration capabilities (Wang Li et al., 2016). Although the terrestrial laser scanner offers more detailed 3D point cloud, they are often limited to their scalability in large vegetation areas, while airborne laser scanning are most suitable for large scale crop monitoring but tend to overestimate canopy area and underestimate canopy height (Wang Li et al., 2016; Wang Li et al., 2015). Crop monitoring requires frequent data acquisition throughout the growing season which is not feasible using laser scanners due to its high cost; therefore, it is evident that the rapidly decreasing UAV systems offers cheap alternative to expensive laser scanners and carries huge possibilities in studying trends in precision agriculture (Leberl et al., 2010; Malambo et al., 2018).

Significant number of research has been performed on the application of crop monitoring (Bendig, Bolten, & Bareth, 2013; J Bendig et al., 2013; Chang et al., 2017; Hunt et al., 2010) using the application of UAV photogrammetry. Numerous research has also been performed on the application of deep learning for robust object detection in a ground based imagery (LeCun, Bengio, & Hinton, 2015; Simonyan & Zisserman, 2014; Tian, Luo, Wang, & Tang, 2015). However, very less has been explored in object detection based on satellite and aerial imagery (Carlet & Abayowa, 2017; Weijia Li, Fu, Yu, & Cracknell, 2016; Puttemans, Van Beeck, & Goedemé, 2018), and is slowing grabbing huge research interests. For the deep neural network detectors to detect an object of interest, it is necessary to supply huge amount of data during the training phase, therefore it is very important to integrate deep learning with big data (Carlet & Abayowa, 2017). We aim to count the stand attributes in palm tree plantations using deep learning based on drone based imagery, which is highly advantageous to the farmers in providing valuable insights to their farmland including, yield prediction,

irrigation management, growth monitoring; and is considered as an important indicator in precision agriculture to assess valuable information related to biomass.

1.2 Problem Identification

Banana plantation exhibits an important part in ensuring food safety and financial assistance for millions of farmers in Thailand. It's also a popular staple food which is rich in vitamins and minerals, and contributes approximately 2.3 million USD each year in the economy (Anupunt, 2002). However, the plantation is affected by several biotic and abiotic stress factors. Biotic stress includes fungi, bacteria, viruses, weeds and pests. Abiotic stress are caused by surrounding environment for example water, temperature. Specifically, the disease Black Sigatoka also termed as black leaf streak (Mobambo et al., 1993) affects banana plantation by blackening & decreasing the leaf area responsible for photosynthetic process which significantly reduces the yield. Therefore, predicting plantation performance to crop stress conditions is vital for an enriched decision making tool to be used by farmers for developing better response mechanism and resilient agricultural systems.

This study aims to enable the spatial analysis within banana plantation, enabling wide range of application in the improvement of plantation's management & scheduling using modified infrared CMOS camera model. This study on monitoring the photosynthetic activity using NDVI indices in banana plantation, throughout the growth season, will provide farmers with valuable information regarding the crop health/disease of individual plantations to help prevent the loss in yield. Similarly, monitoring crop growth using SfM based multitemporal crop surface model (DTM, DSM); and stand count attributes using deep learning neural network framework implemented in tensor flow (Dark flow), and multi-scale detector (YOLOv2) will help farmers access important indicators in precision agriculture such as crop yield, biomass and an estimate of plant health. Finally, the methodological workflow adapted in this

study presents an inexpensive crop health monitoring system for average farmers belonging from developing countries.

1.3 Objectives

In order to fill the existing gap in agricultural crop monitoring in developing countries, the objectives of this research are as follows:

1. Monitoring crop health using modified infrared camera attached to UAV.
2. Assessing plantation growth throughout the growth season, using SfM based multitemporal crop surface model (DSM and DTM), acquired using UAV.
3. Assessing stand count attributes from aerial imagery using deep learning object detection models.

Chapter 2

Literature Review

The application of geospatial technologies coupled with sensor platform has given us tremendous opportunities to explore in the field of precision agriculture (C. Zhang & Kovacs, 2012). (Moran, Inoue, & Barnes, 1997) discussed the limitation of satellite imagery in terms of poor revisiting times, cloud coverage and coarse spatial resolution. However, the recent advancements in the field of UAV gives it the unique capabilities over satellite remote sensing in terms of high temporal resolution (Grenzdörffer, Engel, & Teichert, 2008). The use of UAV in agricultural monitoring can provide farmers with quick, low cost, real time and accurate data (Yue, Lei, Li, & Zhu, 2012). Their research adapted an improved SIFT algorithm for feature extraction and matching which yielded good results when creating mosaic. Afterwards they applied object oriented information extraction method using Feature Analyst 4.2 providing a terrain classification. Their results were significant and provided a quick and low cost methodology for monitoring crop pests. SIFT algorithm is capable of generating huge number of features that could be implemented as a tie-point, which in turn supplies tons of observations for entering bundle block adjustment process resulting in improved accuracy (Wu, 2011; Wu, Agarwal, Curlless, & Seitz, 2011; Y. Zhang, Xiong, & Hao, 2011). (Wu, 2013a) introduced a preemptive feature matching based on SIFT algorithm that was able to increase accuracy as well as decrease processing time of the feature matches.

Sustainable farm management strategies depends on accurate prediction of crop yield, which are normally generated from the computation over remotely-sensed data, i.e. vegetation indices. These indices combined with crop surface model, forms a more concrete methodology which has proven to improve the accuracy in for crop yield prediction (Geipel, Link, & Claupein, 2014). They gathered multitemporal RGB images using unmanned aerial vehicle from early to mid-growth phase, which was processed to simple vegetation index for crop classification, whereas crop height information was extracted from multitemporal crop surface models at multiple resolutions. Their results

demonstrated coefficient of determination (R^2) value of 0.74, and best results were obtained at a spatial resolution of 0.04 m/pixel. Their results proved that combining vegetation indices with crop height information could form a more concrete methodology for predicting corn yield.

Significant number of research has been performed for testing the applicability of action cameras for UAV photogrammetry (Balletti et al., 2014; Hastedt et al., 2016). Calibration of action camera namely GoPro Hero 3 for photogrammetric purpose has been evaluated (Balletti et al., 2014), with the recommendation of use of high spatial resolution during data acquisition. They applied chessboard pattern and OpenCV algorithms to generate distortion free images which significantly improved the accuracy in image registration. The accuracy of initial distorted images in alignment was 0.035 m, which later halved to 0.015 m using undistorted images. An investigation on vegetation cover estimation using modified color infrared GoPro Hero 4 action camera and its NDVI imagery (Ghazal et al., 2015). However, their research doesn't speak much about the validation of the obtained NDVI images. (Hastedt et al., 2016) highlights the promising advantage of light weight cameras in UAV photogrammetry overcoming the restrictions in payload. They applied different camera calibration approaches on their wide angle GoPro Hero4 and evaluated the potential of a pre-correction in rectifying the initial distortion for the applicability in photogrammetric works. They further discussed on the challenges introduced by wide angle lens characterizing their short principal distance and high radiometric distortion, diverting from the central projective model. On a similar study made by Balletti et al 2014 on GoPro Hero3 recommended the use of highest possible resolution during image acquisition.

A camera setup, with infrared filter removed and replaced with a blue blocking filter allowed blue channel to record the NIR light (Zigadlo et al., 2001), forming a promising methodology in monitoring vegetation. Although other standard RGB digital camera has been modified to near infrared for assessment of crop health (Hunt et al., 2005;

Hunt et al., 2010; Rabatel et al., 2014), but very less has been explored in action camera series (Velasquez, Argueta, & Mazariegos, 2016; J. Wijitdechakul et al., 2016) for their potential in monitoring crop health. (Hunt et al., 2010) studied the ability of digital color infrared camera modified through the replacement of infrared filter with a red light blocking filter in crop health monitoring by embedding it on UAV platform. The acquisition was carried over two fields of winter wheat having different fertilization rate situated on Queen Anne's County, Maryland, USA. Their results in terms of GNDVI taken at 210 m altitude over two different fertilization rate of wheat was found to have good correlation with LAI, which suggested that their approach was potential in providing accurate information on crop health. Moreover, their camera setup (Hunt et al., 2010) had the advantage of being low cost, lightweight and compact which made it ideal for UAV, enabling high resolution images at low cost. The other advantage was the channels NIR-Green-Blue are inbuilt recognized by the camera, therefore it doesn't require separate camera to record each channels which causes more trouble during registration which decreases the spatial and radiometric resolution when combined together. Similarly, this camera setup doesn't require any post-processing and can be inspected immediately. The tarpaulins targets were measured using Field Spec Pro FR Spectroradiometer. A black dyed paper is in most cases reflective in the NIR region, which could be applied to test whether the cameras channels are sensitive to NIR light in the absence of Spectroradiometer (Hunt et al., 2010)

In recent years, there have been an increasing amount of research studies in the application of UAV based imagery for precision agriculture (Alexandridis et al., 2017; Allahyari et al., 2016; Khanal, Fulton, & Shearer, 2017; Shaw et al., 2016). The applications of different types of unmanned aerial vehicle(UAV) with different remote sensors was reviewed by (Gago et al., 2015) to access water stress for precision agriculture. (J. Berni, Zarco-Tejada, Sepulcre-Cantó, Fereres, & Villalobos, 2009) mapped the canopy conductance and crop water stress index (CWSI) in olive applying high resolution thermal imaging. Their model took into account the parameters affecting the temperature difference between air and tree canopy like vapor pressure

deficit and wind speed while calculating CWSI. Their methodology enabled the analysis of water use spatially within orchards plantations, ultimately leading to improvement in irrigation management. Traditionally, four methods are used by growers to schedule irrigation: fixed irrigation intervals, judging plant condition by eye, determining soil water content, and estimating crop water use from meteorological data. Irrigation management can be drastically enhanced through constantly observing the crop water content properties, instead of completely relying upon soil content properties or modeled evapotranspiration (Clarke, 1997). A major difficulty in applying infrared thermometry to irrigation management is the nadir viewing instruments cannot be used if exposed soil is within the sensors field of view. A dry bare soil can have a midday temperature that can be $>20^{\circ}\text{C}$ above air temperature and 30°C higher than a non-stressed canopy temperature (Herbert & Jackson, 1985). For reducing such abnormalities resulting from the difference between the soil and canopies, the undesirable soil background effect is eliminated using the Soil adjusted vegetation index (SAVI) as mentioned in (Clarke, 1997). The application of Multispectral Airborne Sensors in detecting crop water stress was studied by (Clarke, 1997) using an empirical approach on drip-irrigated muskmelon. The CWSI is based on the fact that the plants receiving sufficient water has a functional transpiration through their leaves making the leaves cooler and as the water is deficit, the stomata in the leaves closes stopping the transpiration and leading to increased temperature at the leaves. Their methodology applied red and near infrared images together which eliminated undesirable soil background effect using the Soil adjusted vegetation index (SAVI). The combination of SAVI with paired radiant surface temperature from thermal imagery was aimed to reduce the error occurred during the data acquisition from a thermal sensor. The combination was also used to derive trapezoidal two dimensional index empirically which was able to detect water stress even with low percentage of canopy cover.

(Bellvert, Zarco-Tejada, Girona, & Fereres, 2014) compared the accuracy in spatial characterization of the water status across their field plots between field measurements

and thermal remote sensing from UAV across an 11 hectare plantation. The canopy water stress was assessed from the temperature of the canopy, which had a good correlation with leaf water potential with R^2 value of 0.83. The data acquisition time for the thermal camera were 07:30, 9:30 and 12:30 h. The data assessed on 07:30 was useless as it was impossible to distinguish soil with the vegetation canopy temperature. The results were highly correlated at 12:30 h and is suggested as most preferred period for the data acquisition of the thermal imagery. Their results from the thermal images was accurate in assessing the variable water status across their project area.

Similarly, (Möller et al., 2007) applied the fusion of thermal & RGB images for the determination of crop water status in grapevine. Their study demonstrated that combination of thermal and visible imaging can significantly improve their accuracy. Likewise, (Shah) worked on image processing pipeline using thermal imaging to determine the water stress in Walnut trees. The backbone of their algorithm lies in the correction of images for lens distortion and brightness, image mosaicking algorithm following SIFT and RANSAC together with a nonlinear equation which converts radiometric data to surface temperature. (Labbé, Lebourgeois, Jolivot, & Marti) concluded that thermal images could be used to determine the water content of the vegetation for their application in irrigation management strategies. The images acquired from thermal sensors could help in irrigation monitoring if they can be combined with images in the visible and near infrared bands. Furthermore, the image needs to be geometrically and radiometrically calibrated to consider the drift and atmospheric conditions to yield better results.

The identification of water stress condition in an orchard plantation using leaf measurement of chlorophyll fluorescence's and PRI was investigated by (Zarco-Tejada, González-Dugo, & Berni, 2012) using a hyperspectral imagery and a thermal sensor. Their results demonstrated positive correlation between PRI and crown temperature from airborne platform with the ground measurement of stomatal conductance and water potential. They applied atmospheric correction methodology depending on

MODTRAN radiative modeling for obtaining surface temperature. They used a portable weather station to measure atmospheric condition like temperature, humidity and pressure during the flight to feed the data into the model. Their results showed that the atmospheric correction methods conducted with thermal camera was successful in estimating the surface temperature of the vegetation. (Ochoa et al., 2016) presented a hyperspectral imaging system with a highly sensitive visible and NIR camera, and additionally an optical spectrograph for detecting the disease named Black Sigatoka (BS), which as the name suggests creates a dark black dots in the leaves of banana impeding the photosynthesis process.

Similarly, (J. A. Berni, Zarco-Tejada, Suárez, & Fereres, 2009), applied thermal & multispectral remote sensing for the purpose of monitoring vegetation. They applied helicopter based system which was embedded with an inexpensive thermal in the region 7.5–13- μm with 40 cm resolution, and narrow band sensors in the region 400-800 nm with 20 cm resolution; and flight was taken over agricultural plots on summer of 2007. They implemented atmospheric correction with MODTRAN to obtain surface reflectance and temperature which helped in the estimation and validation of biophysical parameters such as leaf area index, and water stress using vegetation indices obtained using their hyperspectral and narrowband sensors such as NDVI, PRI and soil adjusted vegetation indices. The study on sensor web-enabled infrastructure was performed by (Geipel, Jackenkroll, Weis, & Claupein, 2015) for the application in precision farming. Their sensor infrastructure was based on (Bröring et al., 2011) and constitutes of 4 layers mainly: 1.Sensor, 2.Integration, 3.Web Interface and 4. Application Interface. Their infrastructure enabled the users and systems to read the sensor data in a robust way giving the functionality of build applications over the web services.

There are various challenges for obtaining optimum processing of multispectral images, for instance, (Laliberte, Goforth, Steele, & Rango, 2011) described some of the challenges and solutions associated with obtaining radiometrically and geometrically

calibrated orthomosaic for accurately classifying rangeland vegetation. They applied object-based image classification approach, and automated batch processing to generate species-wise classification with an accuracy of 87 %. Their results obtained good correlation between ground spectral reflectance and spectral reflectance from airborne/satellite data for their selected vegetation/soil targets, with an accuracy of $R^2 = 0.92$. (Honkavaara et al., 2012) investigated on the applications of UAV embedded with a hyperspectral camera and high resolution RGB based camera for precision agriculture. They developed an image processing pipeline for robust production of high density point clouds and integrated the orthomosaic with the hyperspectral reflectance, which was applied in the process of biomass estimation. They also discussed on the factors such as image quality, processing framework, surface models which highly affects the accuracy of biomass estimation. Their results confirmed that it is possible to apply light weight, low cost imaging for UAV remote sensing and proved to be powerful and cost efficient technology for possible remote sensing applications.

(N. Agarwal, 2009) applied an efficient method for image matching where he identified small amount of feature match for each images, instead of matching all the image to each other which still preserved enough feature matches for structure from motion increasing efficiency in terms of time. (Wu, 2013a) studied on the large scale image reconstruction from linear time structure from Motion. They introduced a preemptive feature matching that was able to reduce the image pair match by 95%, and still recovered to detect good feature match for the reconstruction process. They were able to examine the complexities in time for the gradient bundle adjustment methods. Their results show that many sub steps in the process of image reconstruction like feature detection, matching, filtering and stitching required $O(n)$, where n being the number of images in the function of time using their novel bundle adjustment strategy which previously required $O(n^4)$ in the function of time. Their method also maintained high level of precision through regularizing the triangulation phase across the feature matches, until every feature map are triangulated.

(D. Turner et al., 2012) presented an approach for robust radiometric and geometric calibration to enable accurate UAV photogrammetry using SfM algorithm. The flight images were processed to create a 3D point clouds in an arbitrary coordinate system, which was later transferred into real world coordinate system using two techniques, either with direct georectification approach that utilized the estimated camera coordinates through camera EXIF file, or through a Ground Control Point (GCP). The point cloud was used to produce DTM which was required for the correction of the images, and subsequently producing an orthomosaic of the project area. An absolute spatial accuracy of 65-120 cm was achieved using direct georectification, whereas a more accurate results were obtained with GCP technique of 10-15 cm. Similarly, the comparison between two approach for determining the crop height determination namely, 1. difference method and 2. statistical method has been compared by (Grenzdörffer, 2014), with recommendation on difference method due to its simple and accurate results when supplied a high resolution reference DTM whereas the statistical approach doesn't necessarily requires a reference DTM for computation of CHM.

Likewise, (Juan et al., 2016) studied the application of UAV for plant phenotyping analysis using two different UAV platforms, mainly an octacopter and a quadcopter for the purpose of monitoring plant growth, cover and yield forecasting in tomato & potato plantation. The octacopter was embedded with two sensors, namely, 12 Megapixel RGB camera and Tetracam multispectral camera recording RGNIR wavelengths. The flight was taken on March 17, 2016 at an altitude of 30 m, maintaining an effective front and side overlap of 80% and 70% respectively. These data were processed using SfM algorithm to generate an orthomosaic and surface models which resulted in the extraction of crop height, cover and plant health using vegetation indices acquired using Tetracam multispectral camera for their entire field plots. Similarly, the application of cheap modified infrared cameras with the IR mirror filters disassembled from a normal RGB camera and replaced with two band pass filters will be particularly tested for its application in monitoring plant health in banana plantations (Hunt et al., 2010).

(Krizhevsky, Sutskever, & Hinton, 2012) trained a deep ConvNet, which comprised of 60 M parameters having five convolution layers & subsequently, maxpool layers together with two fully connected layers, on the ImageNet datasets consisting of millions of dataset to classify 1000 classes. They implemented an efficient GPU framework and a regularization technique for the rapid performance and prevent overfitting during the training phase. Another popular dataset, namely, MS COCO (Lin et al., 2014) is comprised of 2.5 M labelled dataset captured from 328 K ground based images that is able to accurately classify 91 different object types. Likewise, there is another publicly accessible dataset namely, PASCAL VOC (Everingham, Van Gool, Williams, Winn, & Zisserman, 2010) has been trained from 20k ground based training images, and is able to classify 20 different object types.

(Carlet & Abayowa, 2017) presented an improved the performance of YOLOv2 detector for the purpose of fast vehicle detection in an aerial imagery which performs cutting edge detection at 4x speed. Their dataset included several aerial imagery which are publicly available, some of them are: Vehicle Detection in Aerial Imagery (VEDAI), AFVID and DLR3k etc. They made their neural net shallower to increase its output resolution and changed the net shape to match the aspect ratio of data, which increased the speed making it near real time object detector for aerial imagery. However, their precision and recall (Carlet & Abayowa, 2017) when compared to faster RCNN (Ren, He, Girshick, & Sun, 2015) is still slower. Unlike Yolov2, many literatures has also implemented region proposal based neural networks (Girshick, 2015; Ren, He, Girshick, & Sun, 2017) for the purpose of real time object detection applications (Gavrila & Philomin, 1999; Redmon, Divvala, Girshick, & Farhadi, 2016; Ren et al., 2017). For instance, (Ren et al., 2015) implemented region proposal based approach, where they applied 300 regions proposal/image, and achieved an accuracy in terms of detection of 73.2% mean average pixel (mAP) for VOC 2007 dataset and 70.4% mAP for 2012 dataset. Likewise, (He, Zhang, Ren, & Sun, 2016) implemented the residual learning approach for training the deeper neural net, on ImageNet dataset (Krizhevsky et al., 2012) with their network consisting of 152 layers, which is nearly 8

times larger than some popular networks like VGG net (Simonyan & Zisserman, 2014), yet manage to attain minimal complexities. Their result achieved minimal error percentage of 3.57%, when trained on ImageNet dataset, likewise, achieved 28% improved results on COCO dataset; as a result of which it managed to score first position in ILSVRC 2015 classification competition (He et al., 2016). On other hand, (Simonyan & Zisserman, 2014) applied convolutional approach to train a deep neural network for the purpose of large scale detection, where their network comprised of 19 layers implemented with 3 x 3 kernels which demonstrated good results on both localization and classification.

Unlike region proposal based object detection algorithms (Ren et al., 2015), Single Shot Multibox Detector (Liu et al., 2016a) as the name suggests follows a unified detection approach implementing a sole deep neural network. Their approach estimates a predefined set of bounding boxes which could be scaled over different aspect ratio depending on the objects to detect. This method also terminates the creation of region based proposals, together with other resampling techniques and integrates all workflow in a sole neural net, making the overall procedure simple and uncomplicated. Their results (Liu et al., 2016a) were tested on VOC 2007 dataset (Everingham et al., 2010), and attained 74.3% mAP for an input image dimension of 300x300, whereas, an accuracy of 76.9% mAP for input dimension of 512x512, both at 58 frames/second. The results from the conventional SSD was impressive, however (Jeong, Park, & Kwak, 2017) further enhanced the performance of the detector by swapping the VGGNet (Sujana, Abisheck, Ahmed, & Chandran, 2017) in original detector to ResNet (Targ, Almeida, & Lyman, 2016). This modification further improved the performance, which when trained with VOC dataset (Everingham et al., 2010) resulting in 78.5% mAP with input image dimension 300x300 at 35 frames/second, whereas the accuracy further improved with 512x512 input dimension at 80.8% mAP but with 16.6 frames/second. Similarly, (Girshick, Donahue, Darrell, & Malik, 2014) integrated the region proposal technique with the convolutional implementation, referred as RCNN, which enhanced the performance by boosting the precision by 30% i.e. attained mAP 53.3% when

trained on VOC dataset (Everingham et al., 2010). They also suggested in applying transfer learning in cases where there are insufficient training dataset which was followed by fine tuning to significantly improve the overall performance.

Furthermore, (Puttemans et al., 2018) performed research in finding an optimal algorithm for automated and robust detection on aerial images consisting palm tree plantation, where they applied object detection algorithm based on (Viola & Jones, 2001) which followed the principle of boosted cascade of simple features. With significant reduction in computation costs and the addition of robust deep learning architectures which performs well on both classification and detection; it has become feasible for solving the complex task of quick and real time object detection and classification in an aerial imagery. Furthermore, they (Puttemans et al., 2018) mentioned that many other pretrained network such as Caffe model zoo (Jia et al., 2014) might serve as an alternative to their approach.

YOLO object detection algorithm (Redmon & Farhadi, 2017), uses a single neural network that predicts the class probabilities and bounding box directly for the input full images in one evaluation; enabling this architecture to work extremely fast (Redmon et al., 2016). This algorithm helps neural network output precise bounding box. Likewise, with the implication of a finer grids like (19 x 19) (Redmon & Farhadi, 2017) decreases the chance of assigning same grid to multiple objects, and also allows neural network to output the bounding box in any aspect ratio which aren't dictated by striding size as in sliding window classifiers (Girshick, 2015) as well as outputs much precise coordinates. Furthermore, the algorithm is a convolutional implementation ie. we are supplying all the grids in a full size image as a single convolutional network, which makes the network pretty fast and enables real time object detection (Redmon et al., 2016; Redmon & Farhadi, 2017).

One of the main concerns in deep learning is subjected to the training data and the hardware for computation, and without large amount of data, our research focused

on transfer learning, where we utilized existing deep learned models which was trained on large dataset such as Pascal VOC (Everingham et al., 2010) for adapting to our task of object detection through fine tuning the weights of the convolutional neural network onto our single new object class (i.e. Palm trees) . All the above literatures are very significant and outlines proposed methodologies together with the limitations which will be vital in adapting a methodological workflow for our research. These literatures suggests that it is possible to apply the modified near infrared remote sensing, multitemporal crop surface models and deep learning for monitoring crop health, growth and stand count attributes. The purpose of this study is to highlight and discuss related literatures, which can be highly potential in their ability of producing accurate, faster and efficient results to be adapted in the workflow of our project.

Chapter 3

Materials and Methods

3.1 Monitoring crop health using modified infrared action camera.

3.1.1 Study Site

The study area consists of *Musa acuminata* plantation, a species native to Southeast Asian region; which covers an area of 0.186 sq.km as represented by the red polygons in Figure 3.1. The area's center is located at N 14° 15.133' E 100° 53.393' and Z 10 meters (WGS84) in Pathumthani province, Thailand. The area in yellow polygon covering an area of 0.0445 sq.km indicates the dedicated area for the comparison and validation of NDVI results with the field estimates of LAI using LI-COR LAI 2000 Plant Canopy Analyzer.

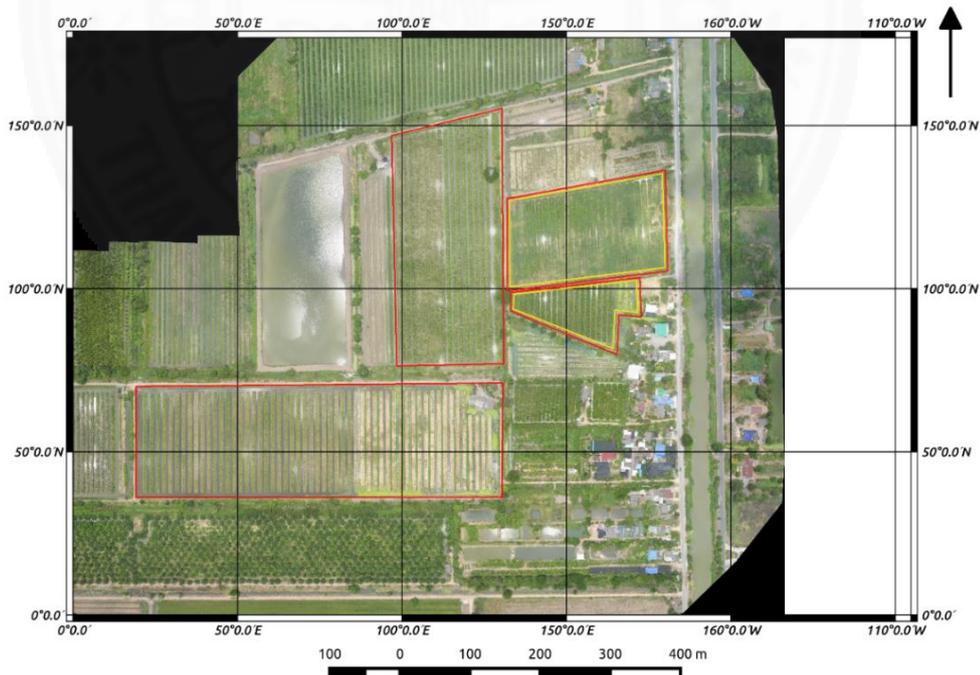


Figure 3.1: The study area in Pathumthani province, Thailand. The red polygon represents the distribution of banana farm while the yellow polygon represents the area for the validation of the NDVI values.

The field data collection was carried on 7th July, 2017 at 12:30 hr. in a bright sunny day. The flight plan was performed in DJI Ground Station Pro with flight parameters description as shown in the Table 3.1 along with the flight plan and unmanned aerial system presented in Figure 3.2

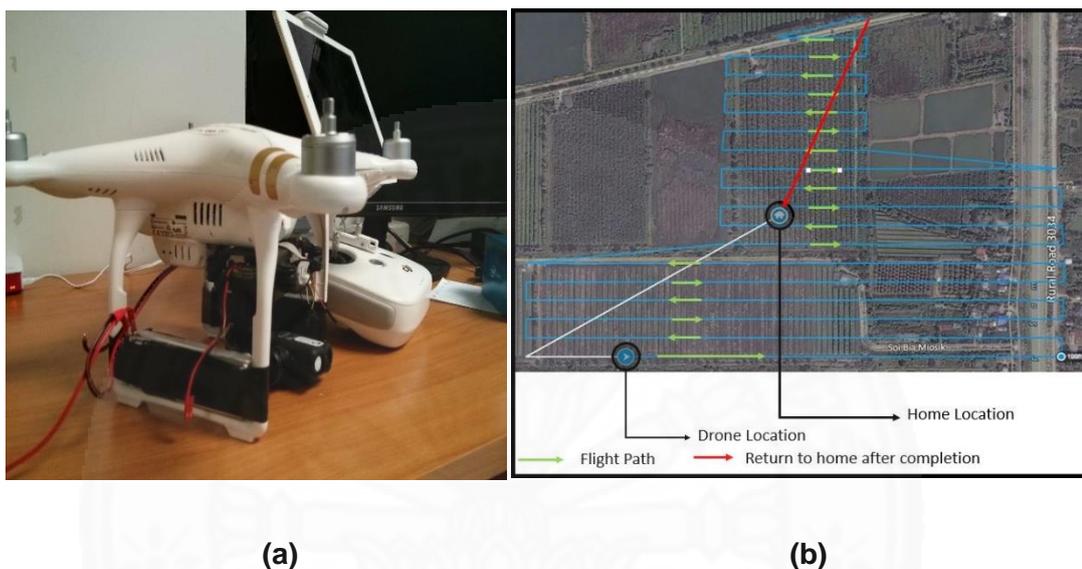


Figure 3.2: (a) Unmanned aerial system design for field data collection, (b) Flight overview

Table 3.1: Flight plan description

Parameters	Value
Shooting angle	Parallel to main path
Capture mode	Hover and capture
Flight course mode	Inside
Speed	5.8 m/s
Altitude	200 meters
Ground sampling distance	8.7 cm/pix
Front overlap	80%
Side overlap	80%

3.1.2 Modified Camera System

The camera system used for the research was a lightweight (39 gm.) mobius action camera with the dimension of (5 cm * 2.5 cm * 2.5 cm) which is a CMOS based camera sensor with spectral response as shown in Figure 3.4. Further details on the camera sensor are given in Table 3.2.

Table 3.2: Camera Specification

Lens	A lens (87 ⁰ FOV)
Resolution	2304 * 1536 pixels
Sensor Type	CMOS Mobius
Focal Length	2.1mm
Sensor Width	5.07mm
Sensor Height	3.38 mm
Sensor Size	1/2.7 inch

The spectral sensitivity of the mobius action camera was observed in a completely dark room using Black Comet C-200 TEC Stellar Net Spectrometer as shown in Figure 3.3. After the removal of infrared blocking filter, the images were illuminated with 850 nm monochromatic led light and the results demonstrates that the mobius action camera was more sensitive to near infrared spectrum in its blue channel and least in red channel as shown in Figure 3.6. Likewise, the spectral response of the CMOS based camera is demonstrated in Figure 3.4.



Figure 3.3: Spectrometer (Black Comet C-200 TEC ranging from 200-1100nm)

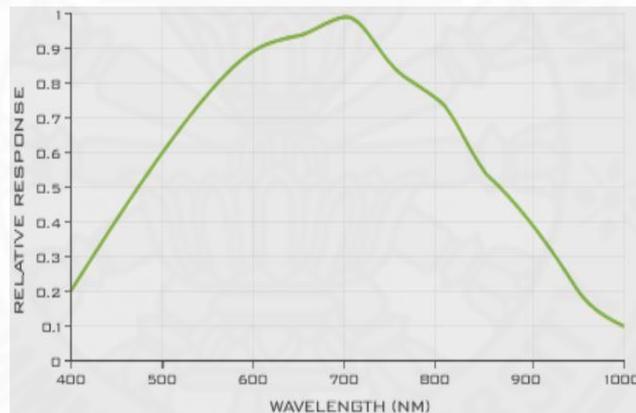


Figure 3.4: Spectral response of CMOS based camera sensor ((LoopTechnology, 2009))

Since, the sensitivity of NIR was higher at camera's blue channel, the camera was modified by replacing the IR filter with Wratten 25 A red filter whose transmission properties are presented in Figure 3.5, which is a gelatin filter that restricts the blue and green bands ranging (440- 600 nm) while allows the red (550-850nm) and near infrared (800-1000 nm) (Velasquez et al., 2016). Thus, the camera sensor records Red-Green-NIR configuration with NIR recorded in the camera's blue channel. The configuration

further requires calibration to correct the NIR wavelength leakage into the red channel which is discussed under preprocessing section below.

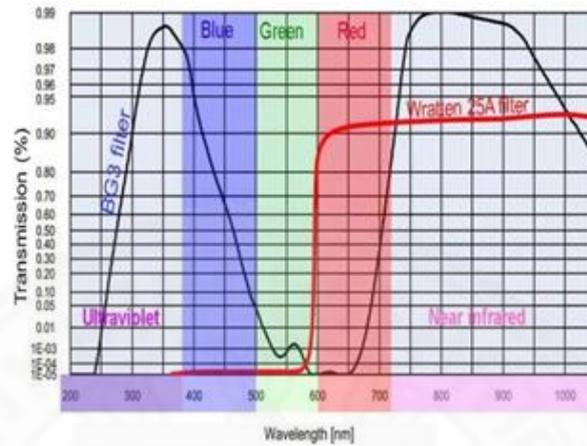


Figure 3.5: Transmission properties of the Wratten 25A Filter (red) ((Fastie, 2013)

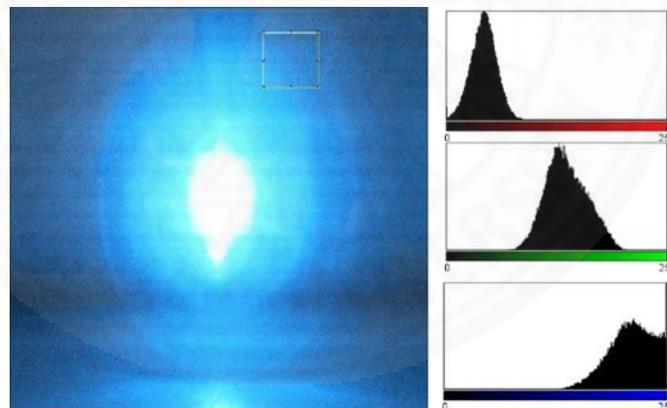


Figure 3.6: Spectral sensitivity of mobius action camera illuminated with monochromatic light of 850 nm.

3.1.3 Methodological Framework

The methodological workflow has been presented in Figure 3.7. First, the multiple images are captured using modified mobius action camera embedded in unmanned aerial system, which are then preprocessed and calibrated before running Sfm algorithm based on (LLC, 2017). The output of the Sfm algorithm yields an orthomosaic, using which we compute vegetation index, NDVI. Finally the field validation of NDVI was carried out by computing LAI.

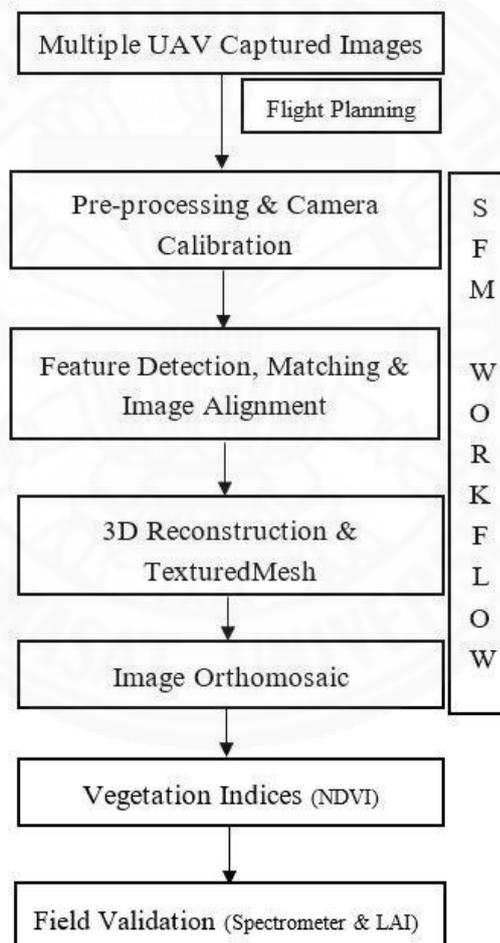


Figure 3.7: Workflow for modified infrared CMOS sensor

3.1.3.1 Geometric calibration of wide angle lens

The imaging system, in most of the action camera with wide angle lens deviates from central projective model. Additionally, the CMOS sensors in action camera doesn't follow global shutter camera models which forms the basis for structure from motion (SFM) approach (Hedborg, Forssén, Felsberg, & Ringaby, 2012). Therefore, it is desired to apply pre-correction in the image for producing images close to central projective model (Hastedt et al., 2016). An initial correction step was implemented, where each images collected in the field were carefully examined manually and initial correction was applied. The correction for the remaining distortion was estimated using standard checkerboard, implementing the pre calibration approach similar to (Harwin, Lucieer, & Osborn, 2015). For the correction of lens distortion, Brown's distortion model (Duane, 1971; Hastedt et al., 2016; Kelcey & Lucieer, 2012) has been implemented which computes the tangential as well as the radial sections of the lens distortion. A set of 8 images of the checkerboard kept in fixed position was taken from various angles and the images were loaded into the Agisoft Lens. The freely available Lens software then generates camera orientation parameters and lens distortion coefficients using brown's model (Duane, 1971) based on bundle adjustment of the matched corners of the checkerboard pattern. The generated parameters were exported to xml file format which was later supplied into Agisoft Photoscan to calibrate the field images.

3.1.3.2 Radiometric Calibration

The goal of calibration is to convert the pixel value of each channel of photo to reflectance. For the calibration of the images, we implemented the plugin (Horning, 2013) which works with Fiji image processing software (Schindelin et al., 2012). First, the plugin is used to calculate the camera calibration variables and thus uses those variables on other multiple images. The image which requires correction should have same camera properties such as ISO, shutter speed as the image which was applied to

determine the camera calibration parameters. The plugin allows flexibility in subtracting required percentage of NIR channel's pixel values from the visible ones. That is because after modification, some percentage of NIR light gets mixed in all the visible channel. The value to subtract will be determined by taking an image over 850 nm LED illumination and carefully studying the histogram of each channels. The plugin also supports removing the effect of the gamma correction (Lebourgeois et al., 2008) that is normally applied when the image is converted to a JPEG inside the camera to make the camera sensor mimic the response of a human eye. The camera sensor records light intensity linearly but our eyes are more sensitive to low-light conditions than they are to brighter lighting so a gamma correction was applied. The linear regression was correlated between average pixel value for red band VS the reference reflectance recorded by standard reflectance target at 600 nm and similar for NIR but at 850 nm. Finally, the slope and aspect of the linear equation was applied using gain offset method on each bands to produce the reflectance image.

3.1.3.3 Feature Detection, Matching and Alignment

Image matching is one of the most time taking process in structure from motion algorithms (SfM) (Geert Verhoeven, 2011). To reduce processing time, the implemented Sfm algorithm (Javernick, Brasington, & Caruso, 2014; LLC, 2017) firstly detects the pair of image which share the same view and creates a set of descriptors for each points and finally detects its equivalence points across the images based on a similar approach to SIFT (Lowe, 2004) as mentioned in (Sona, Pinto, Pagliari, Passoni, & Gini, 2014). Furthermore, the value of 50,000 and 5,000 were supplied as the number of key points and tie points to be extracted from each image, i.e. the algorithm (LLC, 2017) extracts 50,000 points out of which it selects 5000 best points from each image for alignment process which helps reduce the associated processing time. These descriptors and its correspondence across the image allows the creation of 3D sparse point cloud reconstruction (G. Verhoeven, Doneus, Briese, &

Vermeulen, 2012) and camera positions (Hartley & Zisserman, 2004; Ullman, 1979) as shown in the Figure 3.8.

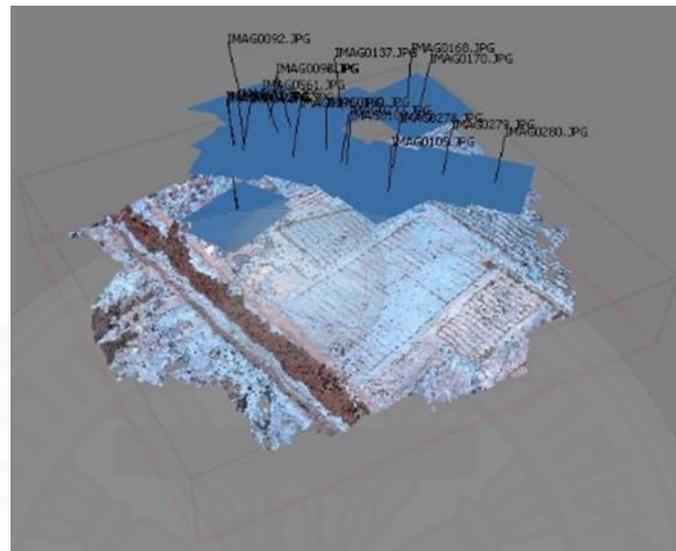


Figure 3.8: Image Alignment where, the blue polygons on the top represents each camera positions.

3.1.3.4 3D Reconstruction, Textured Mesh and Orthomosaic

This step applies the camera calibration parameters previously obtained from Agisoft Lens (LLC, 2017) using brown's distortion model (Duane, 1971) and mentioned in (Ridolfi, Buffi, Venturi, & Manciola, 2017) to remove the lens distortion before further processing of 3D reconstruction (Geert Verhoeven, 2011). Given a set of aligned sparse point clouds and calibration coefficients, the algorithm applies classic bundle adjustment (Bendig et al., 2015; Sona et al., 2014) to generate densely populated point clouds. The dense reconstruction utilizes all pixel value (Scharstein & Szeliski, 2002) which allows handling of even small details of the scene represented as a mesh (Geert Verhoeven, 2011). Furthermore, the mesh were textured using the multiple images. Since our modified camera model didn't have GPS embedded within the system, the dense reconstructed textured mesh is still in an arbitrary space. To transform to absolute

coordinate space, we collected 8 ground control points (GCP) using a GPS whose coordinates has been presented in Table 3.3.

Table 3.3: Details of ground control points in decimal degree (WGS 84)

Label	N	E	Z
B	14.251748	100.891241	17.502
C	14.253361	100.889043	8.996
D	14.251846	100.889074	8.245
G	14.25157	100.891255	15.216
H	14.251544	100.890993	12.860
I	14.252313	100.891231	19.706
J	14.252908	100.889062	9.938
K	14.252137	100.889055	8.133

After the set of GCP were supplied and identified manually in the interface, the algorithm (LLC, 2017) computes seven parameters Helmert Transformation (Javernick et al., 2014; G. Verhoeven et al., 2012) which can be represented by equation 3.1-3.3.

$$X_A = C_X + (1 + s \times 10^{-6}) \cdot (X_L - R_Z \cdot Y_L + R_Y \cdot Z_L) \quad 3.1$$

$$Y_A = C_Y + (1 + s \times 10^{-6}) \cdot (R_Z \cdot X_L + Y_L - R_X \cdot Z_L) \quad 3.2$$

$$Z_A = C_Z + (1 + s \times 10^{-6}) \cdot (-R_Z \cdot X_L + R_X \cdot Y_L + Z_L) \quad 3.3$$

Where,

S= scale factor

C_X, C_Y, C_Z = translation matrix

R_X, R_Y, R_Z = rotation matrix

X_A, Y_A, Z_A = absolute coordinate system of the matched feature points

X_L, Y_L, Z_L = image coordinate of the matched feature points

The Helmert transformation parameters like scale factor, translation and rotation matrix were prior computed using matched feature points and GCP (D. Turner et al., 2012). The concluding phase for the process was to merge the images into a single orthomosaic covering whole project location. As all the images were georectified, it was a simplified step to generate the orthomosaic applying a simple orthomosaic algorithm implemented in (LLC, 2017).

3.1.3.5 NDVI

NDVI (Tucker et al., 2001) is an indicator that provides an estimate of plant health using the visible and near-infrared channels from a multispectral camera platform. NDVI, has found a wide application (Meng, Du, & Wu, 2013) in vegetation performance studies, as it enables farmers to predict their crop yield, access plant health, estimate biophysical characteristics like LAI and many others. As suggested in many literatures (Bravo, Moshou, West, McCartney, & Ramon, 2003; L.-y. Fan, Y.-z. Gao, H. Brück, & C. Bernhofer, 2009; Marti, Bort, Slafer, & Araus, 2007), NDVI has positive correlation with biophysical vegetation parameters, like, plant's photosynthetic activity, leaf area and the biomass. NDVI is given by the equation 3.4.

$$\text{NDVI} = (\text{NIR}-\text{RED}) / (\text{NIR}+\text{RED}) \quad 3.4$$

These indices takes on values between -1 and 1 with -1 being no vegetation (rock, water, bare soil) and 1 being the healthy and highly photosynthetic vegetation.

3.2 Monitoring crop growth using multitemporal crop surface model.

3.2.1 Data Acquisition

The project area in the tropical farmland of Thailand (N 14.25 E100.89, decimal degree-WGS84) consists of *Musa acuminata* plantation; covering an area of 0.186 sq.km as represented in Figure 3.9. The data acquisition was carried on January 15, April 26 and September 16 at 12:30 hr. in a bright sunny day. The flight plan was performed with DJI Ground Station Pro with flight parameters description as shown in the Table 3.4.

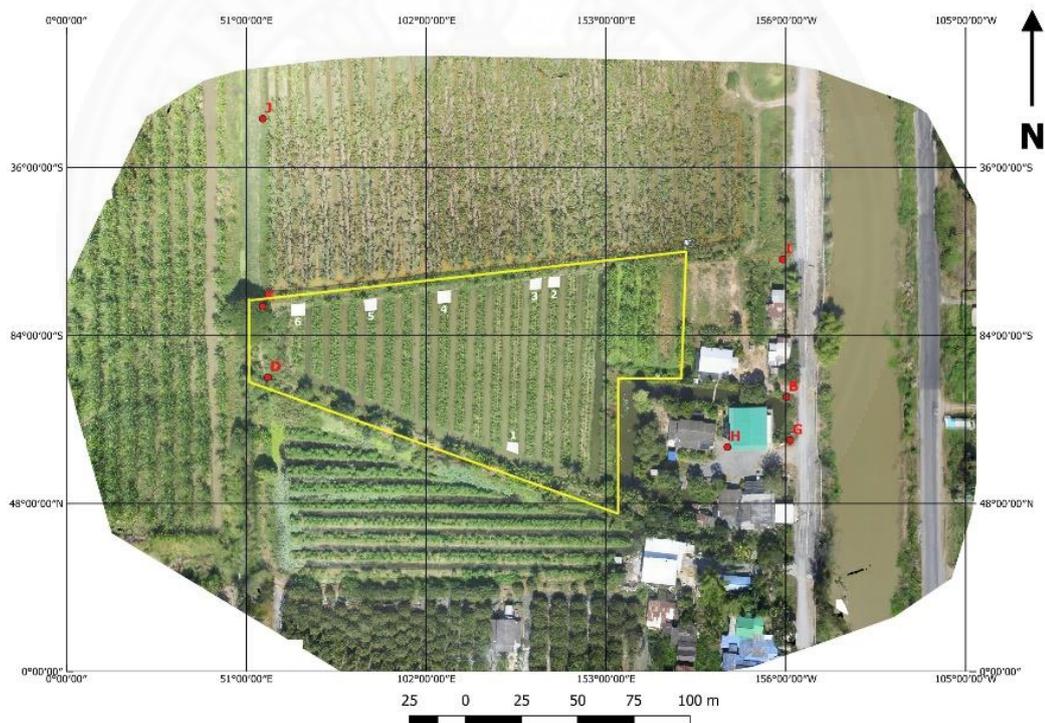


Figure 3.9: The study area in Pathumthani province, Thailand represented by yellow polygons. The GCPs are represented by 7 small red markers, and the plots represented by six small 3x 3 m square plots.

Table 3.4: Description of Flight Plan

Parameters	Value
Shooting angle	Parallel to main path
Capture mode	Hover and capture
Flight course	Inside
Speed	5.8 m/s
Altitude	100 meters
Resolution	4.78 cm/pix
Front overlap	75%
Side Overlap	70%

3.2.2 Methodological Framework

The methodological framework adapted in this study has been demonstrated in Figure 3.10. Firstly, the multiple images captured from UAV were processed for its automatic interior and exterior orientation calibration, before extracting features for the matching process. After the images were aligned, the algorithm (Malambo et al., 2018) computes bundle adjustment to generate a sparse point cloud of the scene.

Prior to data acquisition phase, we setup seven uniformly distributed ground control points (GCP's) in the project area as demonstrated in Figure 3.9, which allows to generate accurate geocorrection and referencing of the UAV data. This is because the approximate locations stored onboard GPS wasn't accurate enough for direct georeferencing (Chang et al., 2017). The GCP's were manually detected in the raw images and their respective coordinates were supplied. The processing of the 89 raw images for each dates were carried on CPU: Intel(R) Core(TM) i7-3770 CPU @ 3.40GHz with RAM: 7GB and GPU: Intel(R) HD Graphics 400, taking 49 minutes for entire processing.

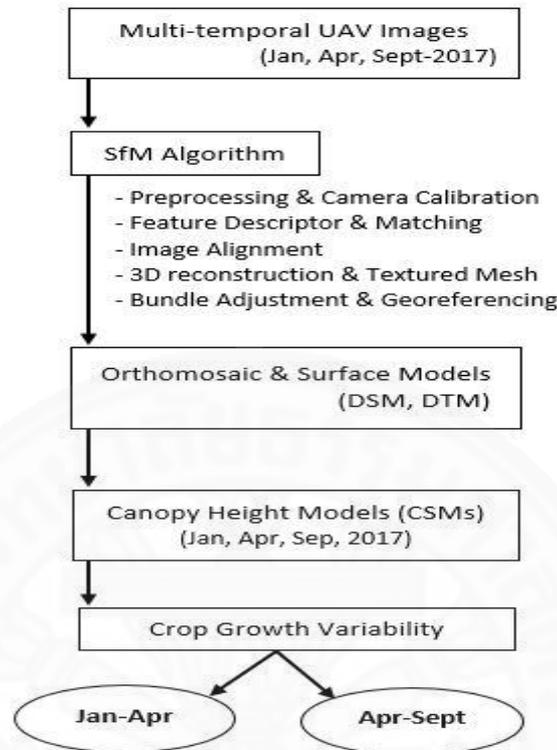


Figure 3.10: Framework for crop growth monitoring.

The algorithm applied in this study follows an iterative bundle adjustment with better image matching approach (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012) such as scale invariant feature matching (Lowe, 1999) for accurate dense terrain reconstruction. Following the dense reconstruction and orthomosaic, digital surface model (DSM) raster were generated using inversed distance weights method together with sharp noise filtering and surface smoothing (Tuominen et al., 2015). Although the flight plan were consistent accross different acquisition dates, the generated orthomosaic and surface models may result in slight difference in spatial resolution depending on weather conditions and accuracy of GPS (Chang et al., 2017). The resolution of the surface models ranged between 3.78-4.78 cm/pixel but for the purpose of maintaining consistent resolution accross dates, the data were resampled to 5 cm/pixel.

The representation of ground surface using DTM was achieved by normalizing the point clouds from discrete ground points. The discrete ground points were interpolated using natural neighbour algorithm (Chang et al., 2017) based on Voronoi tessellation of the data resulting in a smoother surface reconstruction (Ledoux & Gold, 2005). Finally, the canopy height model (CHM) was generated by subtracting the DSM from DTM using a difference method. The major advantage using this method is that with accurate DTM, the crop height is accurately reliable throughout the growing season (Grenzdörffer, 2014). Now that we have the multitemporal CHM data across different acquisition dates, it is straightforward to monitor crop growth by comparing CHM. In this study, we assessed the crop growth between the dates January- April, 2017 and April-September, 2017, to provide valuable insights to crop growth throughout the growth season.

3.3 Palm tree counting in an aerial imagery using deep learning.

3.3.1 Training Dataset

The project area consists of palm tree plantation, which covers an area of 0.062 sq.km. The area's center is located at N $14^{\circ} 3' 40.6''$ E $100^{\circ} 8' 10.3''$ (WGS84) in Nakhon Pathom province, Thailand. A total of 255 images were collected from the plantation area using UAV on 8th February, 2018 at an altitude of 70 meters from the ground. These images were used to generate the training dataset, together with the validation dataset for the purpose of training our deep neural network. A total of 1595 annotations were created manually, out of which 1000 annotations were provided as training dataset and 595 were provided as validation and test datasets. The project area has been demonstrated in Figure 3.11.

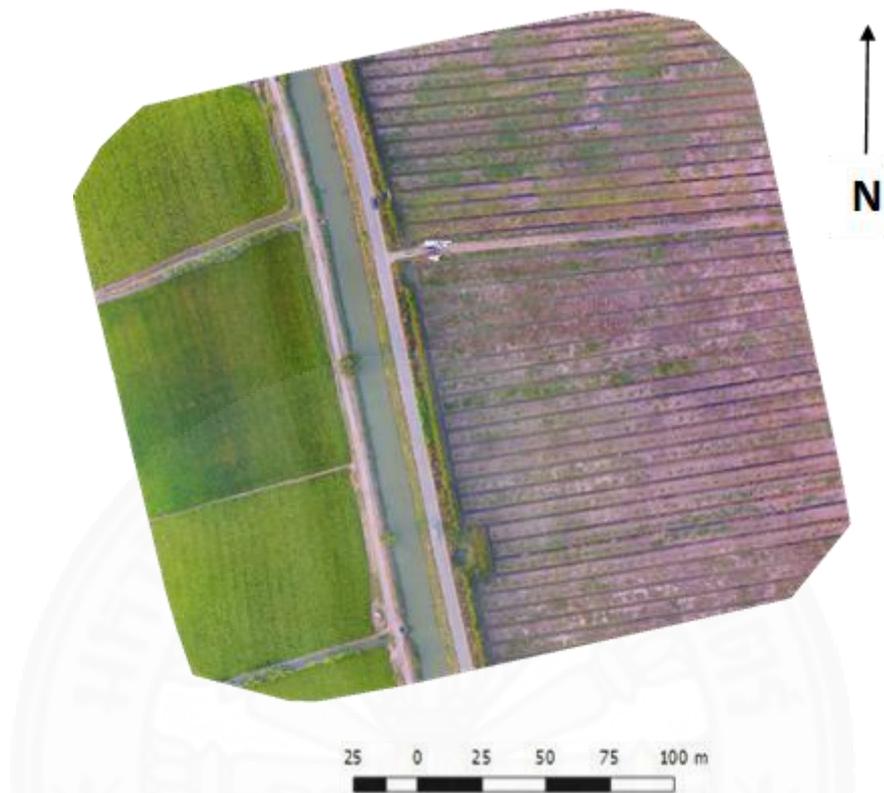


Figure 3.11: The study area consisting of palm plantations in Nakhon Pathom province, Thailand.

3.3.2 Methodological Framework

The methodological workflow has been described in Figure 3.12. The high resolution (4000 x 3000) aerial images of palm trees collected using UAV were manually annotated which consisted of 595 annotations representing validation dataset, whereas 1000 annotations representing training dataset which are required for the purpose of training the ConvNet, and subsequently for accuracy assessment of the network. The convolutional neural network implemented in this study was based open source tensorflow (Abadi et al., 2016) implementation of the darknet framework (Redmon, 2013) named, darkflow (Trieu, 2017). The variables of CNN are continuously adjusted accross many layers until the best overall accuracy is achieved when compared to the 595 annotations of validation dataset . The modified version of YOLO (Redmon

& Farhadi, 2017) comes with pretrained weights and models which were initially trained on Pascal VOC dataset (Everingham et al., 2010) and MS COCO (Lin et al., 2014), however for detecting smaller objects, such as Palm trees in aerial imagery with a single number of classes, the network must be modified and finely tuned (Carlet & Abayowa, 2017), which has been described in sections below.

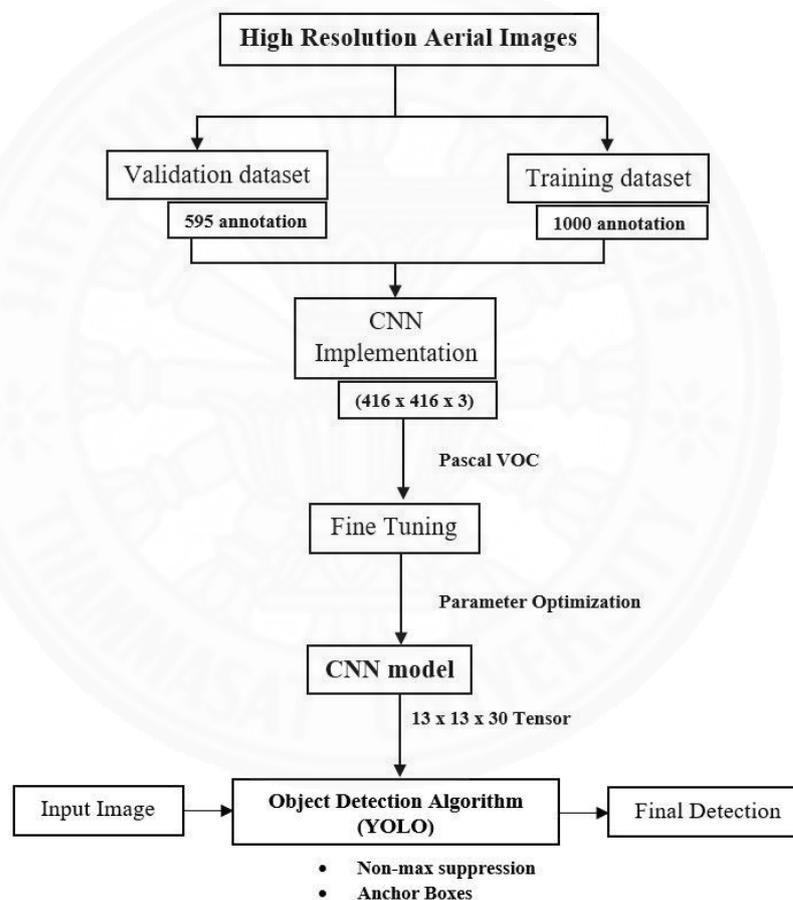


Figure 3.12: Methodological workflow for palm tree counting using deep learning framework.

3.3.2.1 Network Architecture

The network architecture follows an approach similar to (Carlet & Abayowa, 2017) and has been demonstrated in the Figure 3.13. The architecture consists of 23 convolutional networks, with the kernel size ranging between 3 x 3 and 1 x 1. The striding (s) and padding (p) were made consistent in each of the 23 ConvNet, with the value of 1. The activation function for the (n-1) layers were leaky ReLU (Redmon et al., 2016) while the last ConvNet had a linear activation function. A total of 6 maxpool layers with kernel size 2 x 2 and s = 2 has also been introduced in the architecture, which follows with increase in number of channels and decrease in net resolution.

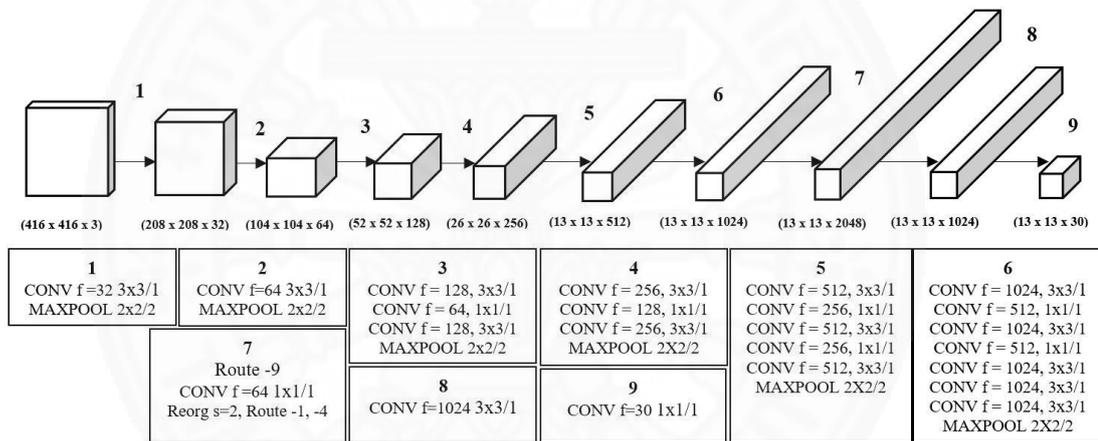


Figure 3.13: Network Architecture

The output of the network is a 13 x 13 x 30 tensor, which is a modified version of YOLO v2 object detection which supports localization of smaller objects through its fine grained feature map (Redmon & Farhadi, 2017) compared to the earlier version which supports detection in 7 x 7 grid as in (Redmon et al., 2016). Likewise, we implemented Route and Reorg in 7th operation of the network, which brings feature from previous layer at 26 x 26 x 512 mentioned in (Carlet & Abayowa, 2017). This added layer binds and stacks the higher resolution features to the lower ones, demonstrating similar architecture to identity mapping in ResNet (He et al., 2016). This

adder layer modifies the $26 \times 26 \times 512$ feature map to $13 \times 13 \times 2048$ feature map giving 1% rise in performance (Redmon & Farhadi, 2017). Similarly, the improved version of YOLOv2 (Redmon & Farhadi, 2017) added batch normalization (Ioffe & Szegedy, 2015; Zhuang, Wang, Yamins, & Hu, 2017) after each convolutional layer which improved the convergence of the model and terminated the need of any other regularization including dropout layer (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014); increasing mean average precision (mAP) by 2%. Furthermore, the model classifier was pretrained at 224×224 on Pascal VOC (Everingham et al., 2010) for classification, while later increments its resolution to 416×416 for detection, which makes the network robust and can be adjusted for higher resolution. This modification of higher resolution classifiers increases the mAP by 4 % (Redmon & Farhadi, 2017).

3.3.2.2 YOLO Object Detection Algorithm

You Look Only Once, abbreviated as YOLO is a near real time object detector whose improved version YOLOv2 launched in late 2016 surpassed all other detection methods for instance, Faster R-CNN (Ren et al., 2015) and SSD (Liu et al., 2016b) in terms of both speed and detection (Redmon & Farhadi, 2017). The mAP of 76.8 was achieved on VOC dataset at 67 FPS, while the fine version at 544×544 achieved even more accuracy of 78.6 mAP at 40 FPS. The detailed statistics on detection accuracy of state-of-the-art YOLOv2 compared to other object detection models has been presented in Table 3.5.

Table 3.5: Accuracy comparison between different object detection models.

Detection Frameworks	Train	mAP	FPS
Fast R-CNN	2007+2012	70.0	0.5
Fast R-CNN VGG-16	2007+2012	73.2	7
Faster R-CNN ResNet	2007+2012	76.4	5
YOLO	2007+2012	63.4	45
SSD300	2007+2012	74.3	46
SSD500	2007+2012	76.8	19
YOLOv2 288 x 288	2007+2012	69.0	91
YOLOv2 352 x 352	2007+2012	73.7	81
YOLOv2 416 x 416	2007+2012	76.8	67
YOLOv2 480 x 480	2007+2012	77.8	59
YOLOv2 544 x 544	2007+2012	78.6	40

Source: (Redmon & Farhadi, 2017)

This algorithm (Redmon et al., 2016) combines methodologies using a single neural network that predicts the class probabilities & bounding box directly for full size images in one look; which makes this architecture work extremely fast. For each of the 13 x 13 grid cells, we defined the output vector y given by equation 3.5.

$$y = \begin{pmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \end{pmatrix} \quad 3.5$$

Where, p_c = probability for object's occurrence i.e.> if $p_c = 0$ (no object (background)) and if $p_c = 1$ (object detected which in our case is Palm tree), and b_x , b_y , b_h and b_w are the parameters to describe the bounding box of the predicted objects which are assigned relative to their grid cell. The bounding boxes are then normalized (Redmon et al., 2016) respect to the grid which makes the values of b_x and b_y falls between 0 and 1, whereas b_h and b_w can be greater than 1.

The Yolo algorithm then takes the midpoint of the object, and assigns the object to the grid cell containing the midpoint, and if the object extends to other grid cells, it

pretends as if the other grids has no part of that object (Redmon et al., 2016). Furthermore, we applied pre-trained weights and configuration files from the PASCAL VOC datasets, which was fine-tuned by changing the number of classes to 1, fitting our purpose of detecting palm trees. Likewise, the number of filters in the last convolutional layer was changed to 30 based on the equation 3.6 where, number of anchors was set to 5 and coords represent the 4 parameters namely, b_x , b_y , b_h and b_w used to define the bounding boxes; an approach similar to (Carlet & Abayowa, 2017).

$$\text{No. of filters (f)} = \text{number of anchors (classes + coords + 1)} \quad 3.6$$

While training, we optimized the loss function as the sum of squared error between our ground truth validation (y) dataset and neural network prediction (\hat{y}). We train the network until the loss function is minimized. Let's suppose we have n parameters for our feature vector y , then the loss function (Redmon et al., 2016) is given by equation 3.7

$$\mathcal{L}(\hat{y} - y) \left\{ \begin{array}{l} = (\hat{y}_1 - y_1)^2 + (\hat{y}_2 - y_2)^2 + \dots + (\hat{y}_n - y_n)^2, \quad \text{if } y_1 = 1 \\ = (\hat{y}_1 - y_1)^2, \quad \text{if } y_1 = 0 \end{array} \right\} \quad 3.7$$

(1) Intersection over Union (IoU)

This IoU function (Redmon et al., 2016; Redmon & Farhadi, 2017) defines the measure for determining the accuracy of our object localization algorithm. This function compares the predicted bounding box from deep learned architecture with the ground truth. In our case, we judge the bounding box to be accurate if IoU threshold ≥ 0.5 similar to (Carlet & Abayowa, 2017). Finally, the most accurate localization results are obtained if both the bounding boxes overlaps perfectly i.e. $\text{IoU} = 1$. An example of object localization is demonstrated in Figure 3.14, where IoU is given by equation 3.8.



Figure 3.14: Object localization example, where the green polygon represents the ground labelled truth, while the red polygon represents the algorithm detected bounding box.

$$\text{IoU} = \frac{\text{(Size of intersection between two bounding boxes)}}{\text{(Size of union between the two bounding boxes)}} \quad 3.8$$

(2) Non-max suppression

One of the problem of object detection is that our algorithm may find multiple detection of the same object (i.e. Detecting same object twice) which is the reason we implement non-max suppression (Huang, Yang, Deng, & Yu, 2015; Redmon et al., 2016) to make sure our algorithm detects each object only once. The algorithm is described below:

1. Discard all boxes with $p_c \leq 0.6$

For each grid in 13×13 i.e. 169 positions, we output bounding boxes together with the probability of that bounding box being a good one. Finally, we discard all the low probability bounding boxes.

2. WHILE, any remaining boxes:
 - Pick the box with largest p_c , and output that as a prediction.

- Discard any remaining box with IoU ≥ 0.5 with the box output in the previous step.

The while loop continues until there are still any remaining bounding boxes to be fed into this loop. In short, non-max suppression cleans up multiple detections based on probability p_c associated with each of the detections. This technique further adds 2-3 % in mAP (Redmon et al., 2016).

(3) Anchor Boxes

Object detection sometimes encounters two or more objects appearing in the same grid cell, hence we use anchor boxes to solve such complexities which also allows the learning algorithm to specialize better. This problem is especially encountered if we use coarse grid cells (7 x 7) rather than fine resolution grid cells (19 x 19) as two objects having the same midpoint is comparatively rare in fine grained resolution (Redmon & Farhadi, 2017). However, for aerial imagery as objects are comparatively smaller therefore, we applied 5 anchor boxes, which were computed from k-means approach on bounding boxes in VOC dataset, similar to (Carlet & Abayowa, 2017; Luo, Peng, Zhu, & Li, 2018).

Chapter 4

Results and Discussions

4.1 Monitoring crop health using modified infrared action camera.

4.1.1 Alignment Accuracy

An approach similar to (Balletti et al., 2014) has been implemented for the alignment of the 30 images using algorithm defined in (LLC, 2017). First of all, the images were corrected for the initial distortion using Agisoft Lens (LLC, 2017), the results of which are presented in Table 4.1, where k_1 , k_2 , k_3 represents radial distortion coefficients and p_1 , p_2 being tangential coefficients, generated in Lens. Before applying the calibration parameters, a pre-correction step (Balletti et al., 2014; Hastedt et al., 2016) was introduced to increase the accuracy, reliability, and provide valid parameters to correct the wide angle images into central projective mathematical model applied in most of the Sfm algorithms using GIMP 2.8 (Distortion, 2013).

After the pre-correction step, the images were aligned using Agisoft Photoscan 1.2.4 (LLC, 2017) using generic image pair selection supplying the key point and tie point limit of 50,000 and 5,000 respectively. The medium accuracy was chosen for the alignment process which subsequently reduced the matching and alignment time to 49 seconds and 25 seconds respectively. The accuracy (Root mean square reprojection error) of the alignment process improved significantly from 0.0975 m to 0.0365 m with the mean key point size and effective overlap of 6.89094 pixel and 2.7732 respectively, after the introduction of the precorrection step.

Table 4.1: Camera orientation & lens distortion coefficients

	Results	Std. error
f (mm)	2.5	0.1650
c_x (mm)	1161.14	0.4224
c_y (mm)	800.322	0.3503
k1	-0.0614605	0.0015
k2	0.0010953	0.00388
k3	-0.0131311	0.00319
p1	-0.00049972	6.912E ⁻⁰⁵
p2	0.00107982	6.0737E ⁻⁰⁵

4.1.2 Spectral calibration of the modified camera

The spectral calibration of the modified mobius camera was performed with 4 calibration targets using Black Comet C-200 TEC Stellar Net Spectrometer as shown in Figure 4.1. The calibration process was carried on July 7, 2017 simultaneous to the drone flight data acquisition, to maintain same atmospheric conditions. The spectrometer readings of % reflectance were then linearly correlated with the normalized image pixel values (DN) collected using modified camera during same time of the day, in a plugin (Horning, 2015) where % reflectance the reflectance values were supplied using a csv file. The reference spectrum for visible (red) had a good correlation with image pixel values at 600 nm with R^2 value of 0.9059, and near-infrared spectrum had a better correlation particularly at 850 nm with R^2 value of 0.8652. The plugin (Horning, 2013) then applies the calibration parameters of regression model to the directory of images for the calibration of modified infrared camera. Finally, the calibrated image is used for the calculation and validation of NDVI results, which has been described in the sections below.



Figure 4.1: Setup of the calibration targets for the spectral calibration of the lens.

4.1.3 NDVI

NDVI (Tucker et al., 2001) is an indicator that uses the visible and near-infrared channels to predict plant performance. NDVI has found a wide application (Meng et al., 2013) in vegetative studies as it has been widely used to estimate crop yields, assess plant health, estimate biophysical characteristics like LAI and many others. As mentioned in many literatures (Bravo et al., 2003; L.-y. Fan et al., 2009; Marti et al., 2007), NDVI has positive correlation with field estimates of leaf area index and biomass. These indices take on values between -1 and 1; with -1 being no vegetation (rock, water, bare soil) and 1 being the healthy and highly photosynthetic vegetation. The result of NDVI has been demonstrated in Figure 4.2.

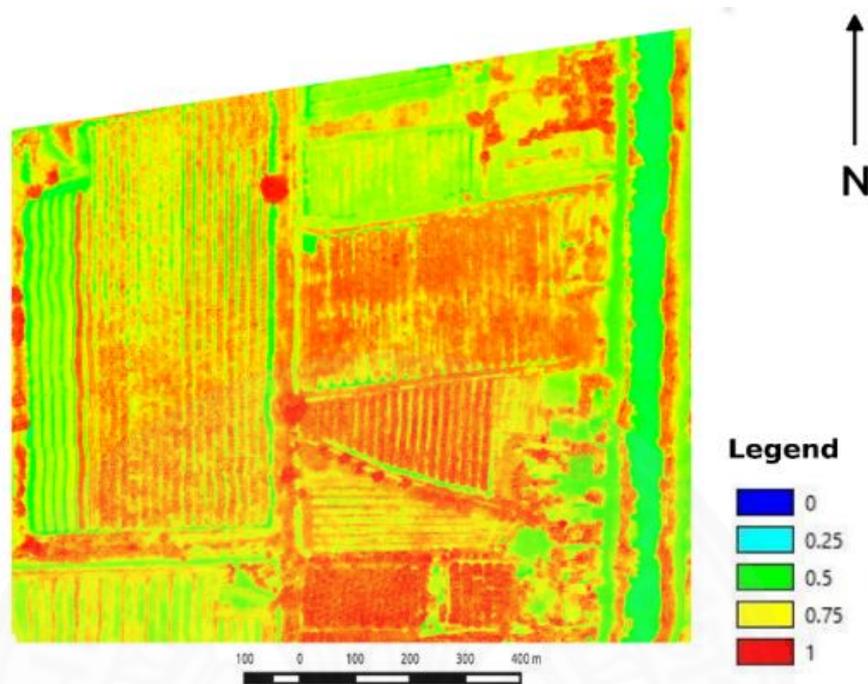


Figure 4.2: NDVI computation from modified infrared camera after calibration.

Finally, we designed 6 square plots, each measuring a length of 5m distributed in an area of 0.0455 sq.km for the validation of NDVI. The field measurement of LAI followed a nondestructive approach by using LI-COR LAI 2000 Canopy Analyzer as shown in Figure 4.3. Simultaneous to the ground measurements, aerial image with modified near-infrared camera was recorded at an altitude 60 m, directly above the plot center in each individual plots.

At an altitude of 60 m, there occurs banana plantation with smaller pixel sizes, which means there will be more pure banana pixels, pure bare-soil pixels, and less mixed pixels. Therefore, NDVI can be estimated using only the pure banana pixels captured through the modified camera system over the field plot. Particularly 5 different images captured during the validation process over each field plot were radiometrically calibrated using (Horning, 2013) plugin, and the calibrated images were used to compute NDVI, which were averaged before correlating with the field measurements

of LAI. The linear relationship between NDVI and field measured LAI was investigated to validate the results from our modified infrared camera; following a similar approach as mentioned in (Hunt et al., 2010). This is based on the literature (L. Fan, Y. Gao, H. Brück, & C. Bernhofer, 2009) which demonstrated NDVI as a good estimator of LAI. Some of the plots encountered saturation in NDVI values above the LAI of 4.5 (D. W. Turner, Fortescue, & Thomas, 2007), which was based on the fact that most of the vegetation indices suffer saturation at some point of LAI which has been mentioned in (Gitelson, Kaufman, & Merzlyak, 1996; Hunt et al., 2010). The plots suffering from saturation were removed as suggested in (Hunt et al., 2010), resulting in an improved correlation with coefficient of determination (R^2) of 0.843 shown in Figure 4.4, where the relationship between the two variables was expressed linearly as: $LAI = 4.0374 NDVI + 0.8066$, with the data demonstrated in Table 4.2. Finally, the result indicates that our modified CMOS camera could potentially be applied for agricultural health monitoring.



Figure 4.3: Field measurements with LI-COR LAI 2000 Plant Canopy Analyzer.

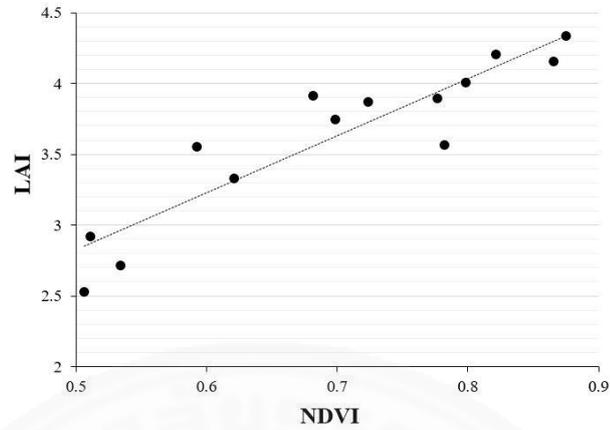


Figure 4.4: Scatterplot depicting the linear relationship between field estimate of LAI and NDVI result from modified infrared camera.

Table 4.2: NDVI measurements from modified camera and field measurements of LAI for validation of the result

Plot	Coordinates (WGS84, Decimal degree)			NDVI	LAI (90)	Std. error
	N	E	Z			
8	14.25139	100.8903	8	0.621	3.33	0.19
				0.681	3.92	0.25
				0.534	2.72	0.07
9	14.25224	100.8899	9.7	0.798	4.01	0.3
				0.821	4.21	0.29
				0.776	3.9	0.12
10	14.25221	100.8899	8.9	0.782	3.57	0.2
				0.865	4.16	0.18
				0.511	2.92	0.25
11	14.25218	100.8896	8.4	0.723	3.876	0.08
				0.875	4.34	0.12
				0.506	2.53	0.05
13	14.25215	100.8893	6.4	0.922	5.99	0.02
				0.918	6.24	0.15
15	14.25214	100.8891	7.5	0.698	3.75	0.14
				0.592	3.554	0.2

Many different approaches has been applied for the correction introduced by wide angle lens in CMOS sensors and their applications for UAV photogrammetry. For instance, calibration of action camera namely, GoProHero3 for photogrammetric purpose has been evaluated (Balletti et al., 2014), with the recommendation of use of high spatial resolution during data acquisition. They applied chessboard pattern and OpenCV algorithms to generate distortion free images which significantly improved the accuracy in image registration. The accuracy of initial distorted images in alignment was 0.035 m, which later halved to 0.015 m using undistorted images. An algorithm for robust and automatic camera calibration using chessboard pattern has been presented by (Douterloigne, Gautama, & Philips, 2009). They further discussed on the fluctuation on the camera parameters, especially on the second order un-distortion parameters, which was explained by different random camera positions giving more weight to one area of image than another. As a result, they concluded on using as much views as possible for correction and generating accurate results. One must pay great attention towards image capture of the checkerboard pattern to get rid of bad lighting, low luminance and gloss effect when imaging from different viewing angles which might create confusion & block the algorithm from creating proper match between images (Balletti et al., 2014). The advantages of light weight action camera in UAV photogrammetry has been highlighted by (Hastedt et al., 2016), particularly through evaluation of GoPro Hero4 using different acquisition modes. Their results suggests on using the pre-corrected images together with pre-calibrated interior orientation parameters which provided valid parameters and most reliable and accurate results. They have recommended the use of GoPro Studio for applying pre correction of image and Agisoft Lens for generating statistical information with chessboard pattern related to the orientation parameters. Our methodological framework followed similar approach as mentioned by (Balletti et al., 2014; Hastedt et al., 2016). However, our results were less accurate but in acceptable region compared to their results because their study area didn't include dense vegetation, where the automated feature detection and matching algorithm introduces some blunders (Remondino, Barazzetti, Nex, Scaioni, & Sarazzi, 2011).

Furthermore, it is highly recommended to have sufficient overlap (~80-90%) between each consecutive image for feature matching and 3D reconstruction. Similarly, the sufficient ground control points must be introduced and distributed throughout the project area, particularly around the periphery to ensure accurate and reliable calibration results for UAV photogrammetry. The data from our modified action camera was free from Jello Effect (vibration, damping) due to the use of Image Stabilizing Gimbal especially designed for Mobius Action Camera (Quantum 3-axis brushless Gimbal). Further, the acquisition altitude was 200 m covering large area which supplied enough feature matches for SfM algorithm.

Various methods and protocols have been developed to test the feasibility of a modified camera system in agricultural health monitoring. For instance, (Hunt et al., 2010) studied the ability of digital color infrared camera modified through replacement of infrared filter with red blocking filter in crop monitoring using UAV. The acquisition was carried out over winter wheat with two different fertilization rate in Queen Anne's County, Maryland, USA. Their results in terms of green normalized vegetation index (GNDVI) was found to have good correlation ($R^2= 0.85$) with LAI suggesting their approach to be potential in providing accurate information on crop health. They collected imagery at 105 m and 210 m altitude and applied Tarpaulins of various color to check the spectral and radiometric properties of the modified camera system. The application of a dual camera system for agricultural health monitoring has been presented by (Jinmika Wijitdechakul, Shiori Sasaki, Yasushi Kiyoki, & Chawan Koopipat, 2016), however this dual system decreases the spatial and radiometric accuracy while registering the images due to intrinsic difference of grey level distribution between NIR and visible images(Hunt et al., 2010; Rabatel et al., 2014). A study on the potential of digital camera has been performed by (Lebourgeois et al., 2008) for allowing its potential use as multispectral sensors to monitor crop by examining a sequence of radiometric corrections to minimize the distortion subjected

to camera optics and environmental phenomena. Their results suggested the use of preprocessing of the raw images and correction of vignetting effect for the modified infrared camera. Similarly, (Lelong et al., 2008) applied the combination of digital camera and spectral filters to design multispectral sensor for the application in precision farming. The images were computed for NDVI, which was compared with field measurement of LAI. Their results demonstrated a moderate correlation between NDVI results and LAI with R^2 of 0.82 and root squared error of 0.57; concluding that their results from standard cameras is potential for precision agriculture; and further suggests researchers on simplifying and improving preprocessing step to establish more accurate results. Our calibration approach followed an approach similar to (Bourgeon, Paoli, Jones, Villette, & Gée, 2016) with direct measurements using spectroradiometer in the field under sunlight conditions to convert DN to reflectance measurements for the calculation of NDVI. Similar to their results, we observed the variations of light intensity with some saturation in the white calibration targets (DN >250) and some calibration targets having strong sensitivity to light along with some effects of environmental conditions and noise. It is therefore suggested to optimize the radiometric calibration; and one must carefully select the number of required calibration targets with excluding targets with high saturation (in our case white target has been excluded).

4.2 Monitoring crop growth using multitemporal crop surface model

4.2.1 SfM Accuracy

A similar approach (Malambo et al., 2018) has been implemented for multiview 3D reconstruction from multiple UAV images. A total of 83 images were processed for each date with 56988 key point extraction per image. Furthermore, the difference of 7.01% was observed between initial and optimized internal camera parameters with 16621 matches per calibrated image. The uncertainties in camera position and orientation parameters has been presented in Table 4.3.

Table 4.3: Uncertainties in camera's positional & orientation parameters.

	X	Y	Z	Omega	Phi	Kappa
Absolute	0.145	0.145	0.353	0.370	0.178	0.070
Relative	0.007	0.007	0.008	0.008	0.0012	0.002

The results of bundle block adjustment demonstrates root mean projection error of 0.278 pixels where, 1623558 and 552470, number of 2D and 3D point's observations were supplied for the adjustment process.

4.2.2 Surface Modeling

The raster representation of the terrain elevation data including both surface and ground elevation for the data acquisition on January, 2017 has been demonstrated in Figure 4.5 and Figure 4.6. The surface models were generated using inverse distance weighting with spacing of 100 cm and ground sampling distance of 4.78 cm/pixel. The total processing time for DSM and DTM were 9 minutes and 2 minutes respectively. Finally, the CHM were generated by applying the difference between the generated surface models --namely, DSM and DTM.

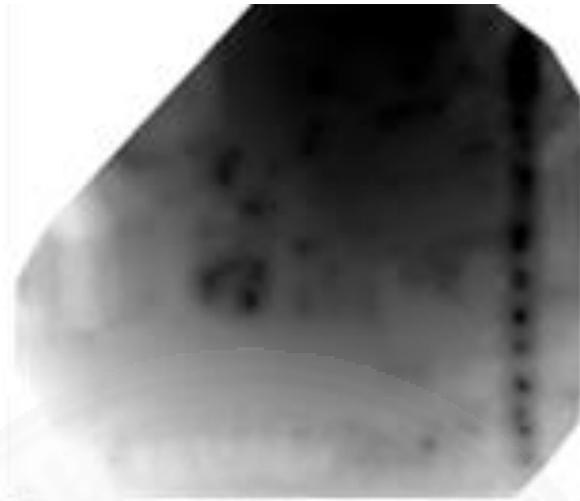


Figure 4.5: Digital Terrain Model (DTM) Raster

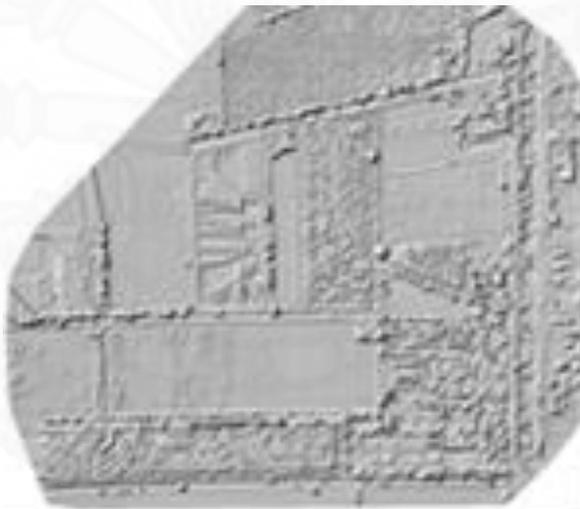


Figure 4.6: Digital Surface Model (DSM) Raster

4.2.3 Crop Growth Statistics

The plot wise investigation of crop growth was carried out by subtracting the generated multitemporal CHM across different acquisition dates. The multitemporal orthomosaic across three different acquisition dates has been presented in Figure 4.7- Figure 4.9. The statistical analysis within each plots were carried out using zonal

statistics function within QGIS 2.18.12 Las Palmas ((2017), 2017), where CHM between different acquisition dates i.e. (Apr-Jan) and (Sept-Apr), were subtracted to access the crop growth within our 6 plots as demonstrated in

Table 4.4 and Table 4.5.



Figure 4.7: Orthomosaic, captured on January 25, 2017



Figure 4.8: Orthomosaic, captured on April 26, 2017



Figure 4.9: Orthomosaic, captured on September 16, 2017

Table 4.4: Crop growth statistics (Jan-Apr, 2017)

Plots	Mean	Med.	Std.	Min	Max	Range
1	2.423	2.657	1.882	-0.526	4.83	6.157
2	0.803	0.807	0.421	-0.338	2.494	2.832
3	3.296	3.753	1.207	-2.287	4.285	7.423
4	2.106	1.966	1.173	0.046	4.644	4.597
5	3.219	3.534	0.968	0.209	4.479	4.27
6	1.778	2.132	2.105	-2.264	4.419	6.683

Table 4.5: Crop growth statistics (Apr-Sept, 2017)

Plots	Mean	Med.	Std.	Min	Max	Range
1	-3.149	-3.235	2.131	-5.641	-3.561	6.406
2	-0.373	-0.646	0.804	-2.448	-1.906	4.268
3	-1.531	-1.792	1.788	-5.215	-3.780	6.184
4	-2.917	-2.817	1.308	-4.895	-3.871	4.143
5	-0.766	-0.374	1.881	-4.895	-3.256	6.610
6	-1.701	-1.633	2.918	-4.967	-2.916	9.741

The plotwise monitoring for the crop growth was based on the assumption of subtracting the maximum values of CHM in each plots as represented in Figure 3.9 across various acquisition dates i.e. (Apr-Jan) and (Sept-Apr). The maximum value corresponding to each plot represents the actual growth within banana plantation between the periods. The maximum crop growth between the growth periods (Jan-Apr, 2017) was observed in plot 1, with the value corresponding to 4.83 m whereas the minimum growth was observed in plot 2 with the value corresponding to 2.492 m.

Furthermore, the plots in our project area were so designed that it contains a single banana tree per plot, with additional grasses growing in the periphery. The period in April 26, represents matured plantation where the plots are entirely covered with the banana's canopy as represented in Figure 4.8. The minimum values in the Table 4.4 & Table 4.5 appears to have abnormalities due to the canopy covering entire plot area during the April period, and the minimum Z value in CHM was confused to be one of the lowest canopy, which is the result behind unpredictable results in the minimum section and also due to presence of other grasses around the periphery of the plot. Our assumption was only to take into consideration the maximum values of CHM within each plot across different acquisition dates, therefore, we simply ignore other variables for instance, minimum, from the Table 4.4 & Table 4.5, which is also supported by the fact that minimum Z values makes no sense while monitoring crop growth.

The observations on the phase (Apr-Sept, 2017) demonstrated the negative banana growth as a result of harvesting carried out on September 10. The orthomosaic in Figure 4.9, captured on September 16 visually demonstrates the post-harvest situation, where most of the banana plantations from the field plot were cut down for the plantation of pineapple plantation. Likewise, the post-harvest yield of the plot 2 was subsequently lower compared to other plots which could be effect of several biotic and abiotic stress factors such as fungi, bacteria and water deficiency (Wairegi, van Asten, Tenywa, & Bekunda, 2010). On careful field investigation of plot 2, we observed black spots on the leaves of banana plantation, which suggests the presence of black sigatoka disease (Stover, 1980). The black spots as represented by the red circles in Figure 4.10, reduces the photosynthesis contributing leaf area and affects the yield within the plantation (Chillet, Abadie, Hubert, Chilin-Charles, & de Lapeyre de Bellaire, 2009). The image was taken with a modified infrared camera during the field data acquisition on July 15, 2017. Furthermore, the crop growth across different growth phases has been spatially represented in Figure 4.11 and Figure 4.12.



Figure 4.10: Pictorial representation of black spots observed in Plot 2.

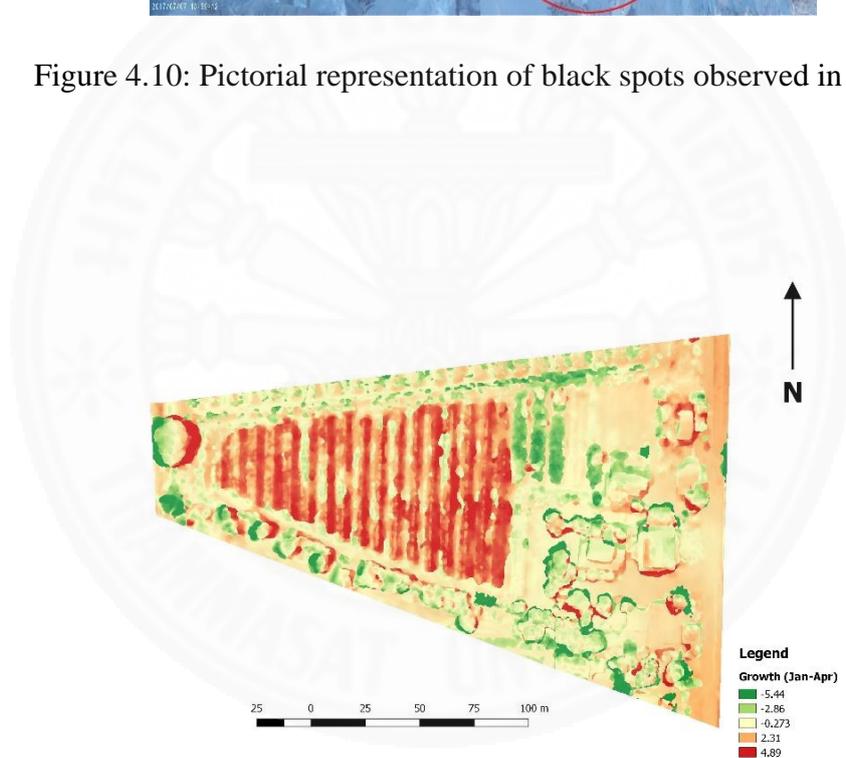


Figure 4.11: Spatial representation of crop growth between the periods Jan-Apr, 2017.

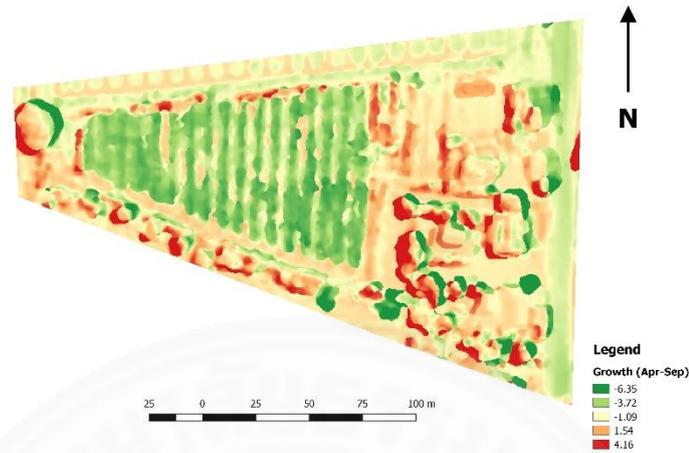


Figure 4.12: Spatial representation of crop growth between the periods Apr-Sept, 2017.

The validation of SfM based elevation model was carried out in SIIT, Bangkadi Campus with the area's centroid at $13^{\circ} 58.836' N$, $100^{\circ} 33.280' E$ as it was not possible to carry out field validation in banana plantation due to the harvesting carried out on 1st week of September. The digital elevation model of the area is represented by Figure 4.13

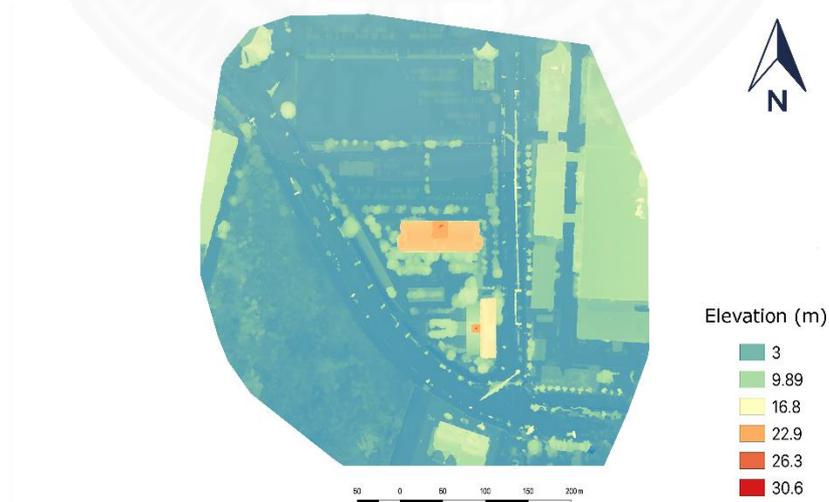


Figure 4.13: Digital Elevation Model (Raster) of Bangkadi Campus Area.

We accessed the car park shed height using SfM dense point clouds as demonstrated in the Figure 4.14, where the obtained result had the height of 3.34 m. This result was validated with the field measurements using tape which resulted in a height of 3.29 m as demonstrated in Figure 4.15.

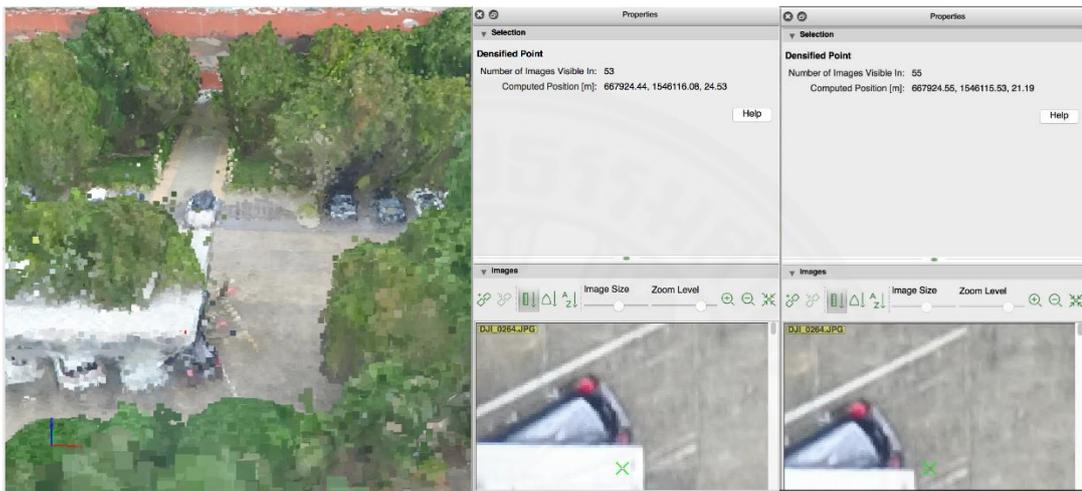


Figure 4.14: Assessing the height of car parking shed using SfM point clouds.



Figure 4.15: Field measurements with a measuring tape for validation of the result.

Various photogrammetric techniques has been developed and validated to extract vital information related to crop growth and biomass using stereo UAV images. For instance, (Malambo et al., 2018) validated a methodological framework for extracting plant height over time using SfM based surface models generated from high resolution UAV images with an accuracy ($R^2 = 0.88-0.97$) when compared with TLS measurements. Likewise, an estimation of above-ground biomass through UAV generated CSM on barley has demonstrated an accuracy of $R^2 = 0.82$, while comparing with their ground measurement of dry biomass (Bendig et al., 2014). Furthermore, our methodology followed a very similar difference method approach for generating CHM from UAV imagery as in (Chang et al., 2017). Their results demonstrated root mean squared error (RMSE) of 0.33 m when compared with field measurements of crop height.

Many literatures suggests the need of a quick, robust methodology for the estimation of biomass which is the key indicator for assessing the food security, enhanced decision making, and food threats management (Becker-Reshef, Vermote, Lindeman, & Justice, 2010; Wang Li et al., 2016). The field based direct estimation of biomass comprises a costly, time consuming and tedious workflow consisting destructive sampling of plants and oven drying to constant weight which is challenging to scale over large areas (Wang Li et al., 2016). Therefore, the UAV derived estimate of biomass seems promising methodology and has huge potential for its application by farmers in rapid assessment for carbon financing and predicting yield. Although the UAV photogrammetry demonstrates promising results , there are various factors associated to uncertainties in generation of DTM, georeferencing and unpredictable data acquisition conditions (Chang et al., 2017). Wind (Mesas-Carrascosa et al., 2015) acts as a major source of uncertainty in acquisition of high quality images as the crop movement during data acquisition introduces discrepancy in feature matching process within SfM algorithm resulting in positional error; ultimately generating underestimation of actual crop height. Finally, it is highly recommended to take into account the wind speed and other

weather conditions; prior to field data acquisition. Likewise, the distribution of GCP was carried throughout the study area with the available accuracy of 1m, however, it is preferred to use more accurate DGPS for accurate CHM measurements. All the above literatures suggests SfM as cheap and robust alternative for assessing plant growth which if assessed over time with other agro-ecological attributes like biomass is potential in accessing the physiological traits & environmental influence on crop performance (Araus & Cairns, 2014; Ghanem, Marrou, & Sinclair, 2015).

4.3 Palm tree counting in an aerial imagery using deep learning

4.3.1 Accuracy of Classification

The accuracy of our model was determined by the 595 validation datasets which were manually annotated using our high resolution aerial imagery of palm tree plantations, while the training was carried with 1000 samples. Likewise, our model consisted of 23 convolution layers, with kernel size of 3x3 and 1x1, whereas consistent maxpooling of kernel size 2x2 has been maintained across the model. The model had the batch size of 64 and subdivisions of size 8 i.e. the model loaded 64 images for each iterations and splits the image into mini-batches consisting 8 images per mini-batch. Furthermore, we modified the parameters such as, number of classes to 1 and number of filters in the last convolution to 30 for the purpose of detecting the one class of palm plantation in aerial imagery. The average loss function of our CNN model was significantly lower at 6500 iterations, and the classification was stopped resulting in R^2 0.9582 as demonstrated in Figure 4.16.

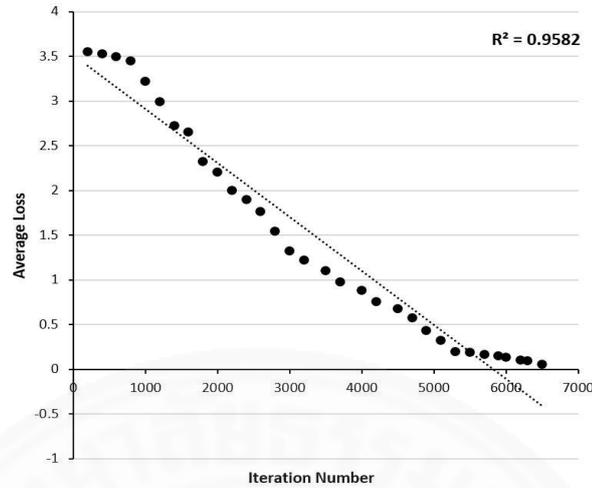


Figure 4.16: Scatterplot representing the loss function of the CNN model.

4.3.2 Accuracy of Object Detection

To assess the accuracy of our object detection model, we calculated two metrics, namely, Precision and Recall, as mentioned in (Carlet & Abayowa, 2017; Weijia Li et al., 2016; Puttemans et al., 2018). Precision, can be referred to as probability that the detected tree is valid, whose relationship is given by equation 4.1, whereas recall, also referred as detection rate is probability that tree is detected in ground truth, whose relationship is given by equation 4.2. The total accuracy is defined by the average of both the precision and recall.

$$\text{Precision} = \frac{\text{No. of correct detection of Palm trees } (N_C)}{\text{Total Palm trees detected } (N_T)} \quad 4.1$$

$$\text{Recall} = \frac{\text{No. of correct detection of Palm trees } (N_C)}{\text{Total Palm trees in Field Measurement } (N_{GT})} \quad 4.2$$

Furthermore, each of our 13 x 13 grid cells is responsible for predicting the output vector (y) which contains the bounding box (b_x , b_y , b_h and b_w) and confidence scores

(Redmon et al., 2016), which is given by the relation described in equation 4.3 as follows:

$$\text{Confidence Score} = P_c (\text{Object}) \times \text{IoU}_{\text{based truth}} \quad 4.3$$

If the grid cell detects no object i.e. $P_c = 0$, then the confidence score must be zero, whereas, if the object is detected i.e. $P_c=1$, then confidence score must be equal to the IoU between the model predicted and ground truth. Finally, the precision and recall for our object detection model was observed to be 45.59% and 65.87%, making our total accuracy to 55.23%, which requires further improvements before it could be applied for palm tree counting purpose, and is discussed in section below. The result of object localization and detection using our architecture has been demonstrated in Figure 4.17, where, red polygon represents machine detected, green polygon represents human labeled and machine detected whereas, yellow represents the training dataset.



Figure 4.17: Initial results of palm tree detection using YOLOv2 detector.

A very few research studies have applied deep learning to count palm plantations on satellite and aerial imagery (Weijia Li et al., 2016; Puttemans et al., 2018). A study on robust detection of coconut plantation in aerial images (Puttemans et al., 2018) has been performed based on the principle of boosted cascade of weak

classifiers based on approach recommended by (Viola & Jones, 2001), and integral channels by (Dollár, Tu, Perona, & Belongie, 2009); which yield accurate tree detection. Their cascade classifiers had the average precision of 94.56 %, whereas their deep learned model achieved an accuracy of 97.4%. Their future prospects of the research discussed on combining their deep learning network with region proposal networks such as (Ren et al., 2015) which will make the entire pipeline work even faster. Similarly, the application of deep learning for robust palm tree counting in a high resolution aerial imagery has been studied by (Weijia Li et al., 2016) in Malaysia. They trained and optimized their ConvNet using manually annotated samples of palm trees in remote sensed images, and later applied sliding window approach for object detection. Their object detector was able to localize palm trees with a remarkable accuracy of 96 % when compared to ground truth labels.

The aerial imagery used for the object detection purpose had the resolution of 4000 x 3000 pixels, which causes the training phase to spend more time per epoch resizing our input images than the actual training. Our architecture during the training phase resizes the images to 416 x 416, which means that the palm trees are resized down by almost a factor of 10. Therefore, the future prospect of this research would focus on tiling the images into multiple sections of 666 x 500 pixel blocks using script in OpenCV, before actually feeding the images for training which is expected to optimize the network. Likewise, the number of training dataset used in our neural network was significantly lower, compared to other similar studies involving palm tree counting such as (Weijia Li et al., 2016; Puttemans et al., 2018).

Hence, the future work would also focus on:

- Increasing the training dataset 5x times with proper quality labeled images.
- Applying finer grid size of 19x19 which would support better localization of smaller objects in aerial imagery as mentioned in (Redmon & Farhadi, 2017).

- Implementing the boosted cascade approach for accurate object localization and detection, based on (Viola & Jones, 2001), and integral channels by (Dollár et al., 2009) which demonstrated promising results for palm tree counting in a research by (Puttemans et al., 2018).
- Combining our architecture with region proposal based on (Ren et al., 2015) and sliding window approach based on (Weijia Li et al., 2016) to further improve the performance.



Chapter 5

Conclusions and Recommendations

This study assessed the feasibility of inexpensive modified infrared CMOS based action camera mounted on DJI Phantom 3 Professional in monitoring the crop health in banana plantations. The action camera modified by replacing the infrared filter with blue blocking filter had the camera configuration of R-G-NIR, out of which R and NIR channels were selected for further processing of NDVI; which was tested and validated against the field measurement of LAI. There was a moderate correlation between the NDVI results from the modified camera and the field estimates of LAI with an R² value of 0.843. Moreover, this system could potentially serve as an effective monitoring tool; however, further research needs to focus on effective band separation for the modified camera by choosing a more suited band pass filters to achieve accurate results. This study is expected to assist farmers in agricultural decision making by constantly monitoring crop health throughout the growth season. Finally, the further studies need to focus on automation of SfM algorithm, reducing the complexities in preprocessing, distortion and radiometric calibration and simplified workflow to provide fast and accurate output to the end users.

The proposed methodology for crop growth monitoring with UAV across different dates has great potential to enable the assessment of valuable information related to crop yield, biomass and estimate of crop health. This study is expected to assist farmers in monitoring crop growth throughout their farm to help increase management practices and site specific agricultural decision making. Finally, the future prospects of our work involves:

- Correlating the crop height from SfM based surface modeling with field measurement of biomass to help better understand crop vitality and predict yield.

- Comparing & validating the SfM based measurement of crop height with more accurate estimates from terrestrial and airborne based laser scanners.

There is a huge advantage of counting plantations, which can provide farmers valuable insights to their farmland including, yield prediction, irrigation management, growth monitoring; ultimately helping improve farm productivity and maximizes profit. The implemented deep learning architecture has demonstrated some potential in palm tree counting, however for direct implementation the accuracy of the classifier must be improved. There is a huge trade-off in choosing an object detection algorithms in terms of speed and accuracy. The implemented object detector namely, YOLOv2 detector outperforms all other approaches such as sliding window and boosted cascade techniques in terms of speed and achieves near-real time speed; while underperforms in terms of accuracy i.e. precision and recall, when compared to sliding window based region proposal techniques. Our primary purpose was to detect palm trees on the fly for fertilizer spraying purpose, which requires detector to perform at near real time, therefore speed was one of the major concern. Hence, our future work intends to increase the training dataset (5x times), and use fine grids (19x19) for better object localization; which is expected to increase the performance without having to compromise with the speed. Moreover, the future prospects of this research would focus on exploring different classification and object detection approach, and integrate together to maximize the performance in terms of both accuracy and speed.

References

- (2017), Q. G. D. T. (2017). Quantum GIS Geographic Information System (2017) Retrieved September 15, 2017, from <http://qgis.osgeo.org>
- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., . . . Isard, M. (2016). *TensorFlow: A System for Large-Scale Machine Learning*. Paper presented at the OSDI.
- Agarwal, N. (2009). S. and Snavely, I. Simon, S. Seitz, and R. Szeliski. Building rome in a day. *ICCV, Kyoto, Japan, 1*.
- Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S. M., & Szeliski, R. (2011). Building rome in a day. *Communications of the ACM, 54*(10), 105-112.
- Alexandridis, T. K., Andrianopoulos, A., Galanis, G., Kalopesa, E., Dimitrakos, A., Katsogiannos, F., & Zalidis, G. (2017). An Integrated Approach to Promote Precision Farming as a Measure Toward Reduced-Input Agriculture in Northern Greece Using a Spatial Decision Support System *Reference Module in Earth Systems and Environmental Sciences*: Elsevier.
- Allahyari, M. S., Mohammadzadeh, M., & Nastis, S. A. (2016). Agricultural experts' attitude towards precision agriculture: Evidence from Guilan Agricultural Organization, Northern Iran. *Information Processing in Agriculture, 3*(3), 183-189. doi: <https://doi.org/10.1016/j.inpa.2016.07.001>
- Anupunt, P. (2002). Banana in Thailand. *Advancing banana and plantain R&D in Asia and the Pacific-Vol. 11*, 149.
- Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends in plant science, 19*(1), 52-61.
- Atzberger, C. (2013). Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sensing, 5*(2), 949-981.
- Balletti, C., Guerra, F., Tsioukas, V., & Vernier, P. (2014). Calibration of action cameras for photogrammetric purposes. *Sensors, 14*(9), 17471-17490.
- Becker-Reshef, I., Vermote, E., Lindeman, M., & Justice, C. (2010). A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment, 114*(6), 1312-1323.

- Bellvert, J., Zarco-Tejada, P. J., Girona, J., & Fereres, E. (2014). Mapping crop water stress index in a 'Pinot-noir' vineyard: comparing ground measurements with thermal remote sensing imagery from an unmanned aerial vehicle. *Precision agriculture*, 15(4), 361-376. doi: 10.1007/s11119-013-9334-5
- Bendig, J., Bolten, A., & Bareth, G. (2013). UAV-based Imaging for Multi-Temporal, very high Resolution Crop Surface Models to monitor Crop Growth Variability Monitoring des Pflanzenwachstums mit Hilfe multitemporaler und hoch auflösender Oberflächenmodelle von Getreidebeständen auf Basis von Bildern aus UAV-Befliegungen. *Photogrammetrie-Fernerkundung-Geoinformation*, 2013(6), 551-562.
- Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., & Bareth, G. (2014). Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. *Remote Sensing*, 6(11), 10395-10412.
- Bendig, J., Willkomm, M., Tilly, N., Gnyp, M., Bennertz, S., Qiang, C., . . . Bareth, G. (2013). Very high resolution crop surface models (CSMs) from UAV-based stereo images for rice growth monitoring in Northeast China. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, 40, 45-50.
- Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., . . . Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation*, 39, 79-87. doi: <http://dx.doi.org/10.1016/j.jag.2015.02.012>
- Berni, J., Zarco-Tejada, P., Sepulcre-Cantó, G., Fereres, E., & Villalobos, F. (2009). Mapping canopy conductance and CWSI in olive orchards using high resolution thermal remote sensing imagery. *Remote sensing of Environment*, 113(11), 2380-2388.
- Berni, J. A., Zarco-Tejada, P. J., Suárez, L., & Fereres, E. (2009). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Transactions on Geoscience and Remote Sensing*, 47(3), 722-738.
- Bourgeon, M.-A., Paoli, J.-N., Jones, G., Villette, S., & Gée, C. (2016). Field radiometric calibration of a multispectral on-the-go sensor dedicated to the characterization of vineyard foliage. *Computers and Electronics in Agriculture*, 123, 184-194. doi: <http://dx.doi.org/10.1016/j.compag.2016.02.019>
- Bravo, C., Moshou, D., West, J., McCartney, A., & Ramon, H. (2003). Early disease detection in wheat fields using spectral reflectance. *Biosystems Engineering*, 84(2), 137-145.

- Bröring, A., Echterhoff, J., Jirka, S., Simonis, I., Everding, T., Stasch, C., . . . Lemmens, R. (2011). New generation sensor web enablement. *Sensors*, *11*(3), 2652-2699.
- Carlet, J., & Abayowa, B. (2017). Fast Vehicle Detection in Aerial Imagery. *arXiv preprint arXiv:1709.08666*.
- Chang, A., Jung, J., Maeda, M. M., & Landivar, J. (2017). Crop height monitoring with digital imagery from Unmanned Aerial System (UAS). *Computers and Electronics in Agriculture*, *141*, 232-237.
- Chillet, M., Abadie, C., Hubert, O., Chilin-Charles, Y., & de Lapeyre de Bellaire, L. (2009). Sigatoka disease reduces the greenlife of bananas. *Crop Protection*, *28*(1), 41-45. doi: <https://doi.org/10.1016/j.cropro.2008.08.008>
- Clarke, T. R. (1997). An empirical approach for detecting crop water stress using multispectral airborne sensors. *HortTechnology*, *7*(1), 9-16.
- D'Agostino, P., Antuono, G., & Pepe, F. ACTION CAMERAS IN DIGITAL PHOTOGRAMMETRY. EXPERIENCES OF COMPARISON OF DENSE POINT CLOUDS.
- Distortion, G. L. (2013). Lens Distortion. Retrieved July 22, 2017, from <https://docs.gimp.org/en/plugin-in-lens-distortion.html>
- Dollár, P., Tu, Z., Perona, P., & Belongie, S. (2009). Integral channel features.
- Douterloigne, K., Gautama, S., & Philips, W. (2009, 12-17 July 2009). *Fully automatic and robust UAV camera calibration using chessboard patterns*. Paper presented at the 2009 IEEE International Geoscience and Remote Sensing Symposium.
- Duane, C. B. (1971). Close-range camera calibration. *Photogramm. Eng.*, *37*(8), 855-866.
- Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, *88*(2), 303-338.
- Fan, L.-y., Gao, Y.-z., Brück, H., & Bernhofer, C. (2009). Investigating the relationship between NDVI and LAI in semi-arid grassland in Inner Mongolia using in-situ measurements. *Theoretical and applied climatology*, *95*(1-2), 151-156.
- Fan, L., Gao, Y., Brück, H., & Bernhofer, C. (2009). Investigating the relationship between NDVI and LAI in semi-arid grassland in Inner Mongolia using in-situ measurements. *Theoretical and Applied Climatology*, *95*(1), 151-156. doi: 10.1007/s00704-007-0369-2

- Fastie, C. (2013). Transmission properties of KODAK Wratten 25 A Filter. In W. A. Filter (Ed.). <https://publiclab.org/notes/cfastie/11-12-2013/red-filter-rising>.
- Gago, J., Douthe, C., Coopman, R., Gallego, P., Ribas-Carbo, M., Flexas, J., . . . Medrano, H. (2015). UAVs challenge to assess water stress for sustainable agriculture. *Agricultural water management*, 153, 9-19.
- Gavrila, D. M., & Philomin, V. (1999). *Real-time object detection for" smart" vehicles*. Paper presented at the Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on.
- Geipel, J., Jackenkroll, M., Weis, M., & Claupein, W. (2015). A Sensor Web-Enabled Infrastructure for Precision Farming. *ISPRS International Journal of Geo-Information*, 4(1), 385-399.
- Geipel, J., Link, J., & Claupein, W. (2014). Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. *Remote Sensing*, 6(11), 10335-10355.
- Ghanem, M. E., Marrou, H., & Sinclair, T. R. (2015). Physiological phenotyping of plants for crop improvement. *Trends in plant science*, 20(3), 139-144.
- Ghazal, M., Khalil, Y. A., & Hajjdiab, H. (2015, 7-10 Dec. 2015). *UAV-based remote sensing for vegetation cover estimation using NDVI imagery and level sets method*. Paper presented at the 2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT).
- Ghimire, S., Xystrakis, F., & Koutsias, N. (2017). Using Terrestrial Laser Scanning to Measure Forest Inventory Parameters in a Mediterranean Coniferous Stand of Western Greece. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*. doi: 10.1007/s41064-017-0024-1
- Girshick, R. (2015). Fast r-cnn. *arXiv preprint arXiv:1504.08083*.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). *Rich feature hierarchies for accurate object detection and semantic segmentation*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote sensing of Environment*, 58(3), 289-298.
- Grenzdörffer, G. (2014). Crop height determination with UAS point clouds. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(1), 135.

- Grenzdörffer, G., Engel, A., & Teichert, B. (2008). The photogrammetric potential of low-cost UAVs in forestry and agriculture. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 31(B3), 1207-1214.
- Hartley, R., & Zisserman, A. (2004). *Multiple View Geometry in Computer Vision* (2 ed.). Cambridge: Cambridge University Press.
- Harwin, S., Lucieer, A., & Osborn, J. (2015). The Impact of the Calibration Method on the Accuracy of Point Clouds Derived Using Unmanned Aerial Vehicle Multi-View Stereopsis. *Remote Sensing*, 7(9), 11933.
- Hastedt, H., Ekkel, T., & Luhmann, T. (2016). EVALUATION OF THE QUALITY OF ACTION CAMERAS WITH WIDE-ANGLE LENSES IN UAV PHOTOGRAMMETRY. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 41.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Hedborg, J., Forssén, P. E., Felsberg, M., & Ringaby, E. (2012, 16-21 June 2012). *Rolling shutter bundle adjustment*. Paper presented at the 2012 IEEE Conference on Computer Vision and Pattern Recognition.
- Herbert, C. V., & Jackson, D. C. (1985). Temperature effects on the responses to prolonged submergence in the turtle *Chrysemys picta bellii*. II. Metabolic rate, blood acid-base and ionic changes, and cardiovascular function in aerated and anoxic water. *Physiological Zoology*, 670-681.
- Honkavaara, E., Kaivosoja, J., Mäkynen, J., Pellikka, I., Pesonen, L., Saari, H., . . . Rosnell, T. (2012). Hyperspectral reflectance signatures and point clouds for precision agriculture by light weight UAV imaging system. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. I-7*, 353-358.
- Horning, N. (2013). Photo Monitoring Plugin. Retrieved July 22, 2017, from <https://github.com/nedhorning/PhotoMonitoringPlugin/graphs/contributors>
- Horning, N. (2015). Introducing the calibration plugin for ImageJ/Fiji. Retrieved July 22, 2017, from <https://publiclab.org/notes/nedhorning/07-22-2015/introducing-the-calibration-plugin-for-imagej-fiji>
- Huang, L., Yang, Y., Deng, Y., & Yu, Y. (2015). Densebox: Unifying landmark localization with end to end object detection. *arXiv preprint arXiv:1509.04874*.

- Hunt, E. R., Cavigelli, M., Daughtry, C. S., McMurtrey, J. E., & Walthall, C. L. (2005). Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precision agriculture*, 6(4), 359-378.
- Hunt, E. R., Hively, W. D., Fujikawa, S. J., Linden, D. S., Daughtry, C. S., & McCarty, G. W. (2010). Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring. *Remote Sensing*, 2(1), 290-305.
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.
- Javernick, L., Brasington, J., & Caruso, B. (2014). Modeling the topography of shallow braided rivers using Structure-from-Motion photogrammetry. *Geomorphology*, 213, 166-182. doi: <http://dx.doi.org/10.1016/j.geomorph.2014.01.006>
- Jeong, J., Park, H., & Kwak, N. (2017). Enhancement of SSD by concatenating feature maps for object detection. *arXiv preprint arXiv:1705.09587*.
- Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., . . . Darrell, T. (2014). *Caffe: Convolutional architecture for fast feature embedding*. Paper presented at the Proceedings of the 22nd ACM international conference on Multimedia.
- Juan, E., Murilo, M., Juan, L., Carlos, A., Jinha, J., & Anjin, C. (2016). Unmanned Aerial System (UAS) for Precision Agriculture and Management Decisions. *2016 ASABE Annual International Meeting*, 1. doi: 10.13031/aim.20162428013
- Kelcey, J., & Lucieer, A. (2012). Sensor correction of a 6-band multispectral imaging sensor for UAV remote sensing. *Remote Sensing*, 4(5), 1462-1493.
- Khanal, S., Fulton, J., & Shearer, S. (2017). An overview of current and potential applications of thermal remote sensing in precision agriculture. *Computers and Electronics in Agriculture*, 139, 22-32. doi: <https://doi.org/10.1016/j.compag.2017.05.001>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *Imagenet classification with deep convolutional neural networks*. Paper presented at the Advances in neural information processing systems.
- Labbé, S., Lebourgeois, V., Jolivot, A., & Marti, R. Thermal infra-red remote sensing for water stress estimation in agriculture.
- Laliberte, A. S., Goforth, M. A., Steele, C. M., & Rango, A. (2011). Multispectral remote sensing from unmanned aircraft: Image processing workflows and applications for rangeland environments. *Remote Sensing*, 3(11), 2529-2551.

- Lamb, D. W., & Brown, R. B. (2001). PA—Precision Agriculture. *Journal of Agricultural Engineering Research*, 78(2), 117-125. doi: <http://dx.doi.org/10.1006/jaer.2000.0630>
- Leberl, F., Irschara, A., Pock, T., Meixner, P., Gruber, M., Scholz, S., & Wiechert, A. (2010). Point clouds. *Photogrammetric Engineering & Remote Sensing*, 76(10), 1123-1134.
- Lebourgeois, V., Bégué, A., Labbé, S., Mallavan, B., Prévot, L., & Roux, B. (2008). Can commercial digital cameras be used as multispectral sensors? A crop monitoring test. *Sensors*, 8(11), 7300-7322.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436.
- Ledoux, H., & Gold, C. (2005). An efficient natural neighbour interpolation algorithm for geoscientific modelling. *Developments in spatial data handling*, 97-108.
- Lelong, C. C., Burger, P., Jubelin, G., Roux, B., Labbé, S., & Baret, F. (2008). Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. *Sensors*, 8(5), 3557-3585.
- Li, W., Fu, H., Yu, L., & Cracknell, A. (2016). Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. *Remote Sensing*, 9(1), 22.
- Li, W., Niu, Z., Chen, H., Li, D., Wu, M., & Zhao, W. (2016). Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. *Ecological Indicators*, 67, 637-648.
- Li, W., Niu, Z., Huang, N., Wang, C., Gao, S., & Wu, C. (2015). Airborne LiDAR technique for estimating biomass components of maize: A case study in Zhangye City, Northwest China. *Ecological Indicators*, 57, 486-496.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., . . . Zitnick, C. L. (2014). *Microsoft coco: Common objects in context*. Paper presented at the European conference on computer vision.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016a). *SSD: Single Shot MultiBox Detector*, Cham.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016b). *Ssd: Single shot multibox detector*. Paper presented at the European conference on computer vision.

- LLC, A. (2017). Agisoft Photoscan User Manual: Professional Edition. *Agisoft Photoscan*. Version 1.2.4. Retrieved June 22, 2017, from http://www.agisoft.com/pdf/photoscan-pro_1_3_en.pdf
- Loop Technology. (2009). Spectral response of CMOS sensor model. In S. R. o. C. b. c. sensor (Ed.). <http://www.looptechnology.com/cmso-sensor-cameras.asp#.WW42E4jyuM9>: Loop Technology.
- Lowe, D. G. (1999). *Object recognition from local scale-invariant features*. Paper presented at the Computer vision, 1999. The proceedings of the seventh IEEE international conference on.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91-110.
- Luo, C., Peng, Y., Zhu, T., & Li, L. (2018). An optimization framework of video advertising: using deep learning algorithm based on global image information. *Cluster Computing*, 1-13.
- Malambo, L., Popescu, S., Murray, S., Putman, E., Pugh, N., Horne, D., . . . Avant, R. (2018). Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. *International Journal of Applied Earth Observation and Geoinformation*, 64, 31-42.
- Marti, J., Bort, J., Slafer, G., & Araus, J. (2007). Can wheat yield be assessed by early measurements of Normalized Difference Vegetation Index? *Annals of Applied biology*, 150(2), 253-257.
- Meng, J., Du, X., & Wu, B. (2013). Generation of high spatial and temporal resolution NDVI and its application in crop biomass estimation. *International Journal of Digital Earth*, 6(3), 203-218.
- Mesas-Carrascosa, F.-J., Torres-Sánchez, J., Clavero-Rumbao, I., García-Ferrer, A., Peña, J.-M., Borra-Serrano, I., & López-Granados, F. (2015). Assessing optimal flight parameters for generating accurate multispectral orthomosaics by UAV to support site-specific crop management. *Remote Sensing*, 7(10), 12793-12814.
- Mobambo, K. N., Gauhl, F., Vuylsteke, D., Ortiz, R., Pasberg-Gauhl, C., & Swennen, R. (1993). Yield loss in plantain from black sigatoka leaf spot and field performance of resistant hybrids. *Field Crops Research*, 35(1), 35-42. doi: [https://doi.org/10.1016/0378-4290\(93\)90134-9](https://doi.org/10.1016/0378-4290(93)90134-9)
- Möller, M., Alchanatis, V., Cohen, Y., Meron, M., Tsipris, J., Naor, A., . . . Cohen, S. (2007). Use of thermal and visible imagery for estimating crop water status of irrigated grapevine. *Journal of experimental botany*, 58(4), 827-838.

- Moran, M. S., Inoue, Y., & Barnes, E. (1997). Opportunities and limitations for image-based remote sensing in precision crop management. *Remote sensing of Environment*, 61(3), 319-346.
- Nijland, W., De Jong, R., De Jong, S. M., Wulder, M. A., Bater, C. W., & Coops, N. C. (2014). Monitoring plant condition and phenology using infrared sensitive consumer grade digital cameras. *Agricultural and Forest Meteorology*, 184, 98-106.
- Ochoa, D., Cevallos, J., Vargas, G., Criollo, R., Romero, D., Castro, R., & Bayona, O. (2016). *Hyperspectral imaging system for disease scanning on banana plants*. Paper presented at the SPIE Commercial+ Scientific Sensing and Imaging.
- Puttemans, S., Van Beeck, K., & Goedemé, T. (2018). *Comparing Boosted Cascades to Deep Learning Architectures for Fast and Robust Coconut Tree Detection in Aerial Images*. Paper presented at the Proceedings of the international conference on computer vision theory and applications.
- Rabatel, G., Gorretta, N., & Labbe, S. (2014). Getting simultaneous red and near-infrared band data from a single digital camera for plant monitoring applications: Theoretical and practical study. *Biosystems Engineering*, 117, 2-14.
- Redmon, J. (2013). 2016. Darknet: Open Source Neural Networks in C.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). *You only look once: Unified, real-time object detection*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. *arXiv preprint*.
- Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., & Sarazzi, D. (2011). UAV photogrammetry for mapping and 3d modeling—current status and future perspectives. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38(1), C22.
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). *Faster r-cnn: Towards real-time object detection with region proposal networks*. Paper presented at the Advances in neural information processing systems.
- Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6), 1137-1149.
- Ridolfi, E., Buffi, G., Venturi, S., & Manciola, P. (2017). Accuracy Analysis of a Dam Model from Drone Surveys. *Sensors*, 17(8), 1777.

- Scharstein, D., & Szeliski, R. (2002). A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision*, 47(1), 7-42. doi: 10.1023/a:1014573219977
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., . . . Schmid, B. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7), 676-682.
- Shah, P. Image Processing Aerial Thermal Images to Determine Water Stress on Crops.
- Shaw, R., Lark, R. M., Williams, A. P., Chadwick, D. R., & Jones, D. L. (2016). Characterising the within-field scale spatial variation of nitrogen in a grassland soil to inform the efficient design of in-situ nitrogen sensor networks for precision agriculture. *Agriculture, Ecosystems & Environment*, 230, 294-306. doi: <https://doi.org/10.1016/j.agee.2016.06.004>
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sona, G., Pinto, L., Pagliari, D., Passoni, D., & Gini, R. (2014). Experimental analysis of different software packages for orientation and digital surface modelling from UAV images. *Earth Science Informatics*, 7(2), 97-107. doi: 10.1007/s12145-013-0142-2
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929-1958.
- Stover, R. (1980). Sigatoka Leaf Spots of. *Plant disease*, 64(8), 751.
- Sujana, S. R., Abisheck, S. S., Ahmed, A. T., & Chandran, K. S. (2017). *Real time object identification using deep convolutional neural networks*. Paper presented at the Communication and Signal Processing (ICCSP), 2017 International Conference on.
- Targ, S., Almeida, D., & Lyman, K. (2016). Resnet in Resnet: generalizing residual architectures. *arXiv preprint arXiv:1603.08029*.
- Tian, Y., Luo, P., Wang, X., & Tang, X. (2015). *Pedestrian detection aided by deep learning semantic tasks*. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- Trieu, T. H. (2017). Darkflow. Retrieved April 25, 2017, from <https://github.com/thtrieu/darkflow>
- Tucker, C. J., Slayback, D. A., Pinzon, J. E., Los, S. O., Myneni, R. B., & Taylor, M. G. (2001). Higher northern latitude normalized difference vegetation index and

- growing season trends from 1982 to 1999. *International journal of biometeorology*, 45(4), 184-190.
- Tuominen, S., Balazs, A., Saari, H., Pölönen, I., Sarkeala, J., & Viitala, R. (2015). Unmanned aerial system imagery and photogrammetric canopy height data in area-based estimation of forest variables. *Silva fennica*, 49.
- Turner, D., Lucieer, A., & Watson, C. (2012). An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (SfM) Point Clouds. *Remote Sensing*, 4(5), 1392.
- Turner, D. W., Fortescue, J. A., & Thomas, D. S. (2007). Environmental physiology of the bananas (*Musa* spp.). *Brazilian Journal of Plant Physiology*, 19, 463-484.
- Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London B: Biological Sciences*, 203(1153), 405-426.
- Velasquez, L. C., Argueta, J., & Mazariegos, K. (2016, 13-16 Oct. 2016). *Implementation of a low cost aerial vehicle for crop analysis in emerging countries*. Paper presented at the 2016 IEEE Global Humanitarian Technology Conference (GHTC).
- Verhoeven, G. (2011). Taking computer vision aloft – archaeological three-dimensional reconstructions from aerial photographs with photscan. *Archaeological Prospection*, 18(1), 67-73. doi: 10.1002/arp.399
- Verhoeven, G., Doneus, M., Briese, C., & Vermeulen, F. (2012). Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *Journal of Archaeological Science*, 39(7), 2060-2070. doi: <http://dx.doi.org/10.1016/j.jas.2012.02.022>
- Viola, P., & Jones, M. (2001). *Rapid object detection using a boosted cascade of simple features*. Paper presented at the Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on.
- Wairegi, L. W., van Asten, P. J., Tenywa, M. M., & Bekunda, M. A. (2010). Abiotic constraints override biotic constraints in East African highland banana systems. *Field Crops Research*, 117(1), 146-153.
- Westoby, M., Brasington, J., Glasser, N., Hambrey, M., & Reynolds, J. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 179, 300-314.
- Wijitdechakul, J., Sasaki, S., Kiyoki, Y., & Koopipat, C. (2016, 29-30 Sept. 2016). *UAV-based multispectral image analysis system with semantic computing for*

- agricultural health conditions monitoring and real-time management*. Paper presented at the 2016 International Electronics Symposium (IES).
- Wijitdechakul, J., Sasaki, S., Kiyoki, Y., & Koopipat, C. (2016). *UAV-based multispectral image analysis system with semantic computing for agricultural health conditions monitoring and real-time management*. Paper presented at the Electronics Symposium (IES), 2016 International.
- Wu, C. (2011). VisualSFM: A visual structure from motion system.
- Wu, C. (2013a). *Towards linear-time incremental structure from motion*. Paper presented at the 2013 International Conference on 3D Vision-3DV 2013.
- Wu, C. (2013b). *Towards linear-time incremental structure from motion*. Paper presented at the 3DTV-Conference, 2013 International Conference on.
- Wu, C., Agarwal, S., Curless, B., & Seitz, S. M. (2011). *Multicore bundle adjustment*. Paper presented at the Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on.
- Yue, J., Lei, T., Li, C., & Zhu, J. (2012). The application of unmanned aerial vehicle remote sensing in quickly monitoring crop pests. *Intelligent Automation & Soft Computing*, 18(8), 1043-1052.
- Zarco-Tejada, P. J., González-Dugo, V., & Berni, J. A. (2012). Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote sensing of Environment*, 117, 322-337.
- Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision agriculture*, 13(6), 693-712.
- Zhang, Y., Xiong, J., & Hao, L. (2011). Photogrammetric processing of low-altitude images acquired by unpiloted aerial vehicles. *The Photogrammetric Record*, 26(134), 190-211.
- Zhuang, C., Wang, Y., Yamins, D., & Hu, X. (2017). Deep Learning Predicts Correlation between a Functional Signature of Higher Visual Areas and Sparse Firing of Neurons. *Frontiers in Computational Neuroscience*, 11(100). doi: 10.3389/fncom.2017.00100
- Zigadlo, J. P., Holden, C. L., Schrader, M. E., & Vogel, R. M. (2001). Electronic color infrared camera: Google Patents.