DEVELOPMENT OF MONITORING OPEN-PIT MINE REHABILITATION BASED ON UAV PHOTOGRAMMETRY AND DEEP LEARNING

TEJENDRA KUMAR YADAV

A Thesis Submitted to the Graduate School of Naresuan University
in Partial Fulfillment of the Requirements
for the Master of Science in (Geographic Information Science)
2020

DEVELOPMENT OF MONITORING OPEN-PIT MINE REHABILITATION BASED ON UAV PHOTOGRAMMETRY AND DEEP LEARNING

TEJENDRA KUMAR YADAV

A Thesis Submitted to the Graduate School of Naresuan University
in Partial Fulfillment of the Requirements
for the Master of Science in (Geographic Information Science)
2020

Thesis entitled " Development of monitoring open-pit mine rehabilitation based on UAV photogrammetry and deep learning"

By TEJENDRA KUMAR YADAV

has been approved by the Graduate School as partial fulfillment of the requirements for the Master of Science in Geographic Information Science of Naresuan University

Oral Defense Committee

..... Chair
(Associate Professor Chatchai Waisurasingha, Ph.D.)

..... Advisor
(Assistant Professor Nattapon Mahavik, D.Sc.)

..... Co Advisor
(Polpreecha Chidburee, Ph.D.)

..... Internal Examiner
(Assistant Professor Kampanart Piyathamrongchai, Ph.D.)

Approved

.....
(Professor Paisarn Muneesawang, Ph.D.)

Dean of the Graduate School

Title DEVELOPMENT OF MONITORING OPEN-PIT MINE REHABILITATION BASED ON UAV PHOTOGRAMMETRY AND DEEP LEARNING

Author TEJENDRA KUMAR YADAV

Advisor Assistant Professor Nattapon Mahavik, D.Sc.

Co-Advisor Polpreecha Chidburee, Ph.D.

Academic Paper Thesis M.S. in Geographic Information Science, Naresuan University, 2020

Keywords UAV photogrammetry Land cover classification Deep learning (DL) Convolution neural network (CNN)

ABSTRACT

Mine reclamation is essential as the extractive organizations are bounded by-laws that have been established by stakeholders to ensure the restoration of mined areas. This study focuses on the application of Unmanned Aerial Vehicle (UAV) photogrammetry and deep learning. The use of a UAV improves safety compared to other surveying systems in mines. In addition, monitoring utilizing drones delivers fast, real-time results, and minimizes human exposure in unsafe ground conditions at mines area. UAVs have revitalized the mining industry through Artificial Intelligence. Deep learning reignited the pursuit of artificial intelligence towards machine to perform related tasks in an automated way. The recent advances of Deep learning (DL) for computer vision tasks, especially for Convolution Neural Network (CNN) models, the potentials the automatically classification of land cover using UAV Photogrammetry. Orthophoto and Digital Surface Model (DSM) are the photogrammetric results used for land cover classification.

This research aims to employ the classification of land cover for monitoring mine reclamation using DL from the UAV photogrammetric results (orthophoto and DSM) at Mae Moh mine in Lampang, Thailand. Two vegetation areas were selected (Pattern and complex) to perform the classification using DL with CNN also, the height of the trees was calculated using results from UAV photogrammetry. The land cover was classified into five classes, comprising: 1) trees, 2) shadow, 3) grassland, 4) barren

land, and 5) others. The effectiveness of both datasets was examined to verify whether orthophoto or combination of orthophoto with DSM for land cover classification.

Land cover classes were, thus, classified with accuracy. The experimental findings revealed that effective results for land cover classification over test area were obtained by DL through the combination of orthophoto and DSM with an Overall Accuracy (OA) of 0.904, Average Accuracy (AA) of 0.681, and Kappa (K) of 0.937 for study area 1 and OA of 0.751, AA of 0.636, K of 0.684 for study area 2. Our experiments presented that land cover classification from combination orthophoto with DSM was more precise than using orthophoto only. This research provides framework for conducting an analytical process, a UAV approach with DL based evaluation of mine reclamation with safety, also providing a time series information for future efforts to evaluate reclamation.

ACKNOWLEDGEMENTS

I would like to express my great appreciation to the Ministry of Industry Nepal, Department of Mines and Geology, Government of Nepal for providing the opportunity and granting me the study leave to continue my Master's degree. I would like to honor Thailand international cooperation agency (TICA) for providing me a scholarship under the Thailand international postgraduate program (TIPP). I also want to specially mention Naresuan University Faculty of Agriculture, Natural Resource and Environment, Department of Natural Resource and environment, Phitsanulok, Thailand for providing me the necessary environment and accomplishing the degree.

I would like to express my deep sense of gratitude to my advisor team Dr. Polpreecha Chidburee and Assistant Professor Dr. Nattapon Mahavik for their consultant, guidance, valuable comments, recommendations, suggestions, and their critical view to shaping the thesis. I would like to thank they for their friendly attitude, sharing knowledge, frankly sharing the ideas which lead to the completion of the research.

I am immensely grateful to Associate Professor Dr. Chatchai Waisurasingha and Assistant Professor Dr. Kampanart Piyathamrongchai Members of the advisory committee for their suggestions. I wish to express my sincere thanks and gratitude to department of Natural Resource and Environment, Naresuan University for providing the necessary environments and equipment to conduct the research.

I would also like to appreciate the helping hands of all staffs at the faculty of Agriculture, Natural Resource and Environment for helping in the academic works and to run the research works especially Ms. Sutthisa Sanhan and Mr. Ittithep onpan Special mention also goes to my friends and juniors who have been helping me during the study as well as an entire period at the University.

I would like to express my gratitude and love to my family whose support, Sacrifice, and love have always been an energetic force for my effort in life.

TEJENDRA KUMAR YADAV

TABLE OF CONTENTS

Page

List of tables

Page

List of figures

Page

Abbreviations

Artificial Intelligence (AI)

Convolutional Neural Networks (CNN)

Deep Learning (DL)

Digital surface model (DSM)

Global navigation satellite system (GNSS)

Ground control point (GCP)

Ground Sampling Distance (GSD)

Land cover (LC)

Object-based image analysis (OBIA)

Real-time kinematic (RTK)

Root Mean Square Error (RMSE)

Unmanned aerial vehicle (UAV)

World Geodetic System 1984 (WGS84)

CHAPTER 1
INTRODUCTION

1.1 Background

There is a need to quantify land cover and its changes over time in a precise and timely way for monitoring human and physical environments at the reclaimed mines area. Mining of the minerals for construction works is human activity, especially open-pit mines represent a physical trace of land modification by human activities (Tarolli & Sofia, 2016). Moreover, open-pit mines influence vegetation, land cover soil, and hydrology condition (Mossa & James, 2013; Osterkamp & Joseph, 2000; Tarolli, 2014). The modern mining process involves prospecting of ore bodies, analysis of profit of a proposed mine, monitoring the mine conditions during mining, extraction of desired minerals, and finally, rehabilitation of the land after the mining activity is closed. However, the demand for mineral resources has significantly increased with the growth of the industry and urbanization, leading to a greater need for mining activities. In recent years, urbanization and industrialization have the increasing demands for the materials for building, base materials, industrial materials (Kobayashi, Watando, & Kakimoto, 2014), and the trend will not cease according to the prediction (Vidal, Goffé, & Arndt, 2013). Mining activities contribute to development through several channels, ranging from employment and development activities.

The reclamation in mines has become the focus of the mining works. Decent engineering design is essential for reducing the effects of the hazards in the mining, which can help to realize the higher level of the mine reclamation works. For example, after finishing mine activity, plantation of trees is vital for the rehabilitation at the mines because the miners are degrading the land and the vegetation by the mining activities and they are supposed to make the land to their natural state and also the mine area is bound by legislation to conclude the rehabilitation works. Therefore, it is also a requirement of monitoring for the post-mining activity (especially land cover, ecological and geological hazard monitoring) for land rehabilitation in mining, which is also the key to provide reasonable planning.

The application of Unmanned Aerial Vehicles (UAV) is essential in mining during recent years, as UAV explore ways of improving the monitoring methods. UAV

helps mining companies to survey large fields at a lower price, less time, and with higher safety margins. The potential usages for the application of drones in mining are seemingly endless. Drone aerial surveys, field mapping, and monitoring can be done in real-time, depending on the system (Rathore & Kumar, 2015). Companies are no longer required to operate land vehicles in rugged environments, frequently with unstable and unsafe terrain. To date, the application of UAV has been limited to operations such as stock control, transport management, dumping dams, safety, emergency response, and monitoring. UAVs play an important role in the mining industry. The following are the benefits of the use of drones in the mining industry:

1. Monitoring – The first and foremost priority of the mining industry is safety, and various kinds of monitoring play a major role in increasing mine safety. Monitoring by drones is an important new type of monitoring and the use of point cloud and image processing to regulate the monitoring. Monitoring with drones is important in mining as traditional monitoring has access problems and safety issues.
2. Cost-Effective - In the past, a team of surveyors was necessary for the surveying equipped with the device at the field. This method was labor-intensive and requires a lot of effort including a large number of errors. The application of UAV as a survey tool takes good planning but improves the use of staff and findings can be obtained in real-time with great accuracy, saving the cost.
3. Development and Resource Replacement – With the application of UAV and sensors, the productivity of the mines can be increased also it optimizes the use of equipment and manpower. Employing UAV can potentially be more cost-effective compared to tripod-based new technologies such as GPS, total station, and LiDAR surveys.
4. Safety and Security- The use of drones helps avoid unexpected accidents by identifying unsafe hazards and identify points of threat. With specialized remote sensors including monitoring techniques, drones help to ensure continuous monitoring of extremely complex mining environments. A UAV can perform continuous monitoring of processing plants and other

structures for safety risks, and inspect structures for unsafe rock movements. It can also help in emergency escape and disaster monitoring.

UAV photogrammetry has revolutionized our ability to observe the surface of the earth. Technological developments over the last decade have allowed more and more accurate location and time monitoring of land and water supply by improved sensors. The Ordnance Survey, an extractive corporation, needs square kilometers of aerial photography per year to upgrade its imagery for the monitoring of reclaimed mines and topographical items. The number of sensors being flown is increasingly growing, both in terms of spatial resolution and temporal frequency, due to better instrumentation and demand from the authorities and end-users concerned to maximize the data currency. The vast majority of the ground features were collected manually by on-site observations and aerial photo interpretations, which are highly labor-intensive and time-consuming. UAVs, more generally referred to as drones, are widely being used in research studies to resolve issues related to the fast and high-resolution processing of spatial datasets (Gago et al., 2015; Haala, Cramer, Weimer, & Trittler, 2011; Webster, Westoby, Rutter, & Jonas, 2018). UAVs can be operated at low altitude, allowing high-resolution data to be obtained (Flener et al., 2013); The low cost of UAVs enables time-and spatially versatile observation schemes, due to ease of implementation and use, to promote the acquisition of data (Feng, Liu, & Gong, 2015). Arguably, this large archive of aerial imagery is highly under-utilized and could be mined for much more information efficiently and effectively through modern techniques UAV photogrammetry, and deep learning (DL). Accurate and up-to-date information on land cover is also required to keep pace with evolving conditions and to improve appropriate policy and decision-making.

Recent trends in Artificial Intelligence (AI) and machine learning, particularly in the emerging field of deep learning, have transformed the way we store, interpret and manipulate geospatial data. These are largely motivated by the surge of interest in deep machine learning, as the new frontier of AI, where the most representative and hierarchical attributes are trained from end to end, spatially (Arel, Rose, & Karnowski, 2010). Deep learning approaches have gained considerable success not only in classical computer vision activities, such as target identification, image recognition, and

robotics, but also in many other practical applications (F. Hu, Xia, Hu, & Zhang, 2015; Nogueira, Penatti, & Dos Santos, 2017). They have made significant advances from state-of-the-art documentation in several areas, and have drawn considerable interest in both academics and industrialized cultures. Deep learning is basically about representation learning or function learning, where the most representative and unequal attributes are taught at the end to end, hierarchically (C. Chen, Zhang, Su, Li, & Wang, 2016). Unlike their shallow equivalents, such as the support vector machine (SVM) and multi-layer perceptron, DL methods do not focus on prior features extraction or human feature architecture, but rather learn higher-level features representation through the models themselves to improve generalization capabilities (Arel et al., 2010). Also, deep levels of interpretation have a strong ability to describe robust features of complicated patterns and classifications such as ground cover, functional locations, etc. Deep learning approaches are good match for capturing this form of feature representation with high-level semantics (Nogueira et al., 2017).

In past few years, deep learning with convolution neural networks (CNNs) have gained substantial interest in the community's image processing (Krizhevsky, Sutskever, & Hinton, 2012). They were originally designed for image classification, where the image is attributed to a particular real-world category depending on its nature, such as natural scenes used in computer vision applications (Maggiori, Tarabalka, Charpiat, & Alliez, 2017). However, the land classifications (LC) classifications follow the actual criteria of image classification, which expects all pixels in the whole image to be classified and labeled in land cover categories.

In summary, DL methods have significant advantages for the learning of dynamic feature representations, they are still feasible. LC classification has not been discussed until now. In comparison, LC are simply abstractions or generalizations of the terrain in the reclaimed mining area. The classification systems are presented at different levels, nested within each other hierarchically over the same geographical space. Different land features exist at different scales, and their monitoring goals are highly dependent on important implementations (Heydari & Mountrakis, 2018). There is also an open issue as to how to better implement or refine DL-based approaches to solve complex LC classification challenges using UAV imagery.

1.2 Statement of problems

The reclamation of vegetation on mining sites is one of the principal rehabilitation strategies used to address the environment (Hartemink & Minasny, 2014). Reclamation is a complex and long process that can be unsuccessful due to a range of living and nonliving factors (Ramoelo et al., 2015). Open-pit mining causes the comprehensive destruction of the solid environment, water bodies, and environment. The reclamation of open-pit mines is regulated in all the developing countries. Reclamation of the degraded land at the mines area involves priority for the government as well as public sector (Ockendon et al., 2018). Introducing through planting showing encroaching the vegetation to the rehabilitation of the mining landscape conditions (Korjus, Laarmann, Sims, Paluots, & Kangur, 2014). For instance, in Thailand, we can find different levels of legislations and regional government all of which require monitoring periodically of mine rehabilitation (Ministry of Industry Thailand 2009). Rehabilitation of the open-pit mine and mining integration techniques is an effective way to solve the environmental problems of the open-pit mine. Mine rehabilitation is an important part of mining technology.

These problems are related to the employment of affective monitoring techniques aimed at accurate observation of mining-related effects. Besides, there are many advantages to employ Geomatics approaches for reclamation monitoring, using UAV Photogrammetry and deep learning. The increasing availability of UAV recent advances in sensor technologies and analytical capabilities are rapidly expanding the number of potential UAV applications. UAV-based environmental monitoring applications, work has focused on land cover mapping (Akar, 2017; Bryson, Reid, Ramos, & Sukkarieh, 2010).

In comparison to the other monitoring methods, UAVs based on different kinds of sensors acquire multi-source data for continuous monitoring, and the centimetre-scale images can be obtained rapidly when the UAV flew at a low altitude. As a dynamic, continuous, and economical data acquisition method, UAVs have now various advantages in monitoring the mines compare to the traditional methods. The comparison of the different equipment used for monitoring is shown in table1.

Table 1 Comparison between UAV Photogrammetry and other monitoring methods in mining

Monitoring methods	Total station	Satellite remote sensing	InSAR	UAV photogrammetry
Data collecting time	Long	Short	Short	Shortest
Data processing time	Depend on computer	Depend on computer	Depend on computer	Depend on computer
Price	Low	High	High	Low
Working condition	Weather dependent	All-weather	All-weather	Weather dependent
Monitoring Area	Small	Large	Large	Medium

Source: Arango & Morales, 2015; Cigna et al., 2017; Padró, Muñoz, Planas, & Pons, 2019

Mine site rehabilitation typically needs to be addressed at fine spatial scales due to the heterogeneous nature of the substrate and topography that influence rehabilitation processes. Characterizing the spatial distribution of land-cover at high spatial resolution can provide key information to assist in landscape function. Object-based approaches are used to classify landscapes, whose features can be observed at multiple spatial scales. These approaches are also employed to analyze ecological patterns that are suitable for extraction through object-based segmentation and classification methods with the use of high spatial resolution imagery or image data in complex environments. Thus, the application of UAV photogrammetry can help to provide information related to land cover at high spatial resolution.

In conclusion, the study focuses on the monitoring of the mine rehabilitation using the imagery from UAV. The research focused on UAV in the application of 3D modeling, slope stability, mapping but there is insufficient information related to mine rehabilitation monitoring. Mining areas are being impacted by the mine works and need to be rehabilitated (Xiang, Chen, Sofia, Tian, & Tarolli, 2018; Yucel & Turan, 2016). The method of UAV monitoring for mine rehabilitation used in this study will be useful for government and private agencies to evaluate the mining areas, especially for decision-making purposes. Therefore, UAV photogrammetry for monitoring rehabilitation in the open-pit mine can be for legislation purposes.

1.3 Aim and objectives of the study

The study aims to monitor the mine rehabilitation works at the open-pit mine site after the completion of the mine work using UAV photogrammetry. The following are the objectives of the study:

1) To employ UAV photogrammetry for supporting the rehabilitation monitoring in open-pit mines.
2) To develop the algorithm based on Deep learning for the land cover classification of the rehabilitation monitoring in the mine works.

1.4 Research questions

The research question for the above objectives are as follow:

- How is the effectiveness of the UAV photogrammetric approach compared to other monitoring methods for monitoring rehabilitation in the mines?
- How did the algorithm develop from deep learning benefits for the classification of land cover?

1.5 Purpose of study

UAV technology is developing and progressed rapidly from year to year for mapping applications. Moreover, UAV is a low budget with time constraints and less manpower rather than satellites of aircraft with expensive flight costs, time-consuming and weather dependent data collection, restricted workability, limited flying time, and low ground resolution mapping process (Aasen, Burkart, Bolten, & Bareth, 2015; Bendig et al., 2014; Dash, Watt, Pearse, Heaphy, & Dungey, 2017; Tian et al., 2017). The UAV photogrammetry can be used for monitoring the rehabilitation of the mine sites as it provides rapid data collection, faster, low cost and utilizes less manpower. Moreover, the 3D model can be generated using UAV images performing digital image processing (DIP). Data processing rapid and accurate results can be obtained. Hence this study attempts to improve the methodology for monitoring the mine rehabilitation assessment using the UAV technology. This technology can be used by government organizations, private agencies, and local authorities for planning and making decision purposes. This study accounts for the land cover mapping of the rehabilitated mining

areas. The study focus on the digital elevation model (DEM), and UAV images for land cover classification and monitoring for the rehabilitated open-pit mine areas.

1.6 Significance of the study

Monitoring reclamation for open-pit mining with UAV photogrammetry is useful at it provides expedient and accurate spatial and thematic information. Monitoring the restoration after the extraction is the legal requirement for the mine companies and public administration. It is resulting in the systematic and sampling of the restoration area of interest, which improves the airborne spatial resolution. Allows the monitoring in inaccessible restricted and dangerous areas, also the imagery for visual inspections, photogrammetric processing, and GIS software provides spatial information about the vegetation development with relatively low-cost equipment used in the study. i.e. low weight platform sensor, low-cost reference panel. Also, UAV photogrammetry is possible to generate useful results such as orthophoto, DEM, digital surface model (DSM), vegetation indices, and thematic land cover maps. The use of UAV photogrammetry gives the new idea for monitoring the open-pit mine rehabilitation. Moreover, UAV is a low budget project with time constraints and less manpower rather than satellites imagery with expensive flight costs, time-consuming and weather dependent data collection, restricted workability, limited flying time, and low ground resolution mapping process (Ren, Zhao, Xiao, & Hu, 2019).

The monitoring of the rehabilitation of the mine works will guide the government, organizations, and private individuals for the developmental activities. The area with the mine works is prone, land cannot be kept barren it needs to be with vegetation after the completion of the mine works. The study will suggest the appropriate UAV photogrammetry techniques in an open-pit mining environment and represent a valid support susceptible area for the mine area to monitor with high spatial and temporal resolution. This approach can be considered as a valid tool to monitor the impacts of mining.

1.7 Benefits of study

UAV photogrammetry has shown the possibility for the future in mining applications. UAV can minimize the time consuming and labour-intensive rapidly in comparison to the traditional surveying methods(Park & Choi, 2020). Periodic data is

obtained from the UAV ensures much-improved resolution than satellite images (Hastaoğlu, Gül, Poyraz, & Kara, 2019). At the same time, the payload and the time consumed for the study are less and are accessible to environmental factors such as terrain. Furthermore, hyperspectral cameras, thermal cameras, infrared cameras, and LIDAR suitable for UAV platforms are still expensive, limiting their use (Colomina & Molina, 2014; Lucieer, Jong, & Turner, 2014; Turner, Lucieer, & De Jong, 2015). At present, the bulk of sensors used in the mining applications are still focused on 3D reconstruction and terrain Surveying. It can be seen that UAVs have a great possibility in mine monitoring at a small or large scale. Low-cost sensor development is necessary to expand the service range of UAV photogrammetry in mining applications and achieve better research results. The use of UAVs helps to obtain information for complex mining environments.

Monitoring of the rehabilitation of open-pit mines represents a challenge for the upgraded supporting the planning. The analysis of open-pit mines through rehabilitation improves the mechanism for responsible environmental effects with the classification helps to find the appropriate strategies for the rehabilitation (Al-Najjar et al., 2019). UAVs have been attractive in many research fields to obtain the latest information of the target areas, owing to their high mobility, high resolution, and low cost (Remondino, Barazzetti, Nex, Scaioni, & Sarazzi, 2011). Nowadays, UAV-photogrammetry has reached a level of practical reliability and become a useful platform for spatial data acquisition. Hence it is expected that UAV-photogrammetry can acquire topographical data in a short time and generate a high-resolution digital model of the environment with the required accuracy.

Studies have been done for using the approaches and tasks for the classification of land cover. Al-Najjar et al. (2019) used fused DSM and CNN for land cover classification. The studies also, therefore, vary according to the technique used. Feng et al. (2015) developed model for the classification of forest and urban area, model contained 200 trees trained on handcrafted texture feature.

1.8 Thesis Structure

Chapter 1 gives the general introduction of this thesis. It comes with commercial and academic needs and towards deep learning as the methods, focusing on the challenges and opportunities for land cover and land use classification using DL with UAV photogrammetry.

Chapter 2 provides a concise literature review of the traditional and deep learning-based methods in land cover classification using UAV photogrammetry and discusses the research directions.

Chapter 3 presents the methodology for UAV photogrammetry and land cover classification with deep learning using CNN in which land cover at reclaimed mines area is classified. Also, the evaluation methods for determining the accuracy of the research is described

Chapter 4 summarizes the results obtained from this research and answers the research questions. Also, the results from the classification methods are evaluated.

Chapter 5 concludes the remarks of the thesis followed by the recommendations and future works.

CHAPTER 2
LITERATURE REVIEW

2.1 Research overview

land cover classification contributes to the up-to-date evidence on both the current state of and changes on the surface of the land. Mine surface information includes relates to the land cover at the mines area. Also, it is necessary to develop an automatic classification methodology. Monitoring plant development and their conditions on rehabilitated mining areas are one of the most important factors in assessing their adaptation in the restored ecosystem. Plant succession strongly depends on soil parameters and soil formation processes in the restored or rehabilitated area soil is a critical component that interacts with vegetation, climate, and animals (Bradshaw & Hüttl, 2001). Method for land cover classification mapping utilizes modern machine learning techniques. The basic concept of image classification can be shown in two steps: first, learning the classification model from labeled reference data, and second, its prediction to all pixels of the imagery. The excellence of the image classification depends on the machine learning algorithm (Maxwell, Warner, & Fang, 2018). Moreover, the accuracy is influenced by the class, extent, spatial and semantic distribution, and positional correctness of the reference data. Some related studies focus on vegetation land cover to describe the vegetation cover maps (Tassopoulou et al., 2019), agriculture and monitoring the vegetation (Khaliq, Musci, & Chiaberge, 2018; Kussul, Mykola, Shelestov, & Skakun, 2018).

Land cover classification creates a common understanding of land cover nomenclatures to produce global, regional, and national data sets able to be designed at different scales, levels of detail, and geographic location. The land cover classification provides a general framework of rules from which more exclusive conditions can be derived to create specific legends. The system may be used to specify any land cover feature anywhere in the world, using a set of independent diagnostic criteria that allow correlation with existing classifications and legends. Researchers have developed and studied methods for land cover classification, such as the Maximum Likelihood Classifier (Otukei & Blaschke, 2010), Random Forests (Breiman, 2001), and Support

Vector Machine (SVM) (Suykens & Vandewalle, 1999) methods. There is a high probability for mixed classes and other problems

The advancement in the deep learning techniques provided an alternative to traditional land cover classifiers. Deep learning brought about around 2006 (G. E. Hinton, Osindero, & Teh, 2006), became well known in the computer vision community around 2012, since one supervised version of deep learning networks CNN made a breakthrough for scene classification tasks (Krizhevsky et al., 2012; Yann LeCun, Yoshua Bengio, & Geoffrey Hinton, 2015), and has reached out to many applications and other academic areas in recent years as it continues to advance technologies in areas, like speech recognition (G. Hinton et al., 2012), medical diagnosis (Suk, Lee, Shen, & Initiative, 2014), autonomous driving (Huval et al., 2015), or even the gaming world (Silver et al., 2016). When compared with other traditional classifiers, deep learning does not require feature engineering, which attracted many researchers from the remote sensing community to test its usability for landcover mapping (Alshehhi, Marpu, Woon, & Dalla Mura, 2017; X. Ma, Wang, & Wang, 2016; Vetrivel, Gerke, Kerle, Nex, & Vosselman, 2018). The review papers on object-based image analysis (OBIA) (G. Chen, Weng, Hay, & He, 2018; L. Ma et al., 2017) emphasize the need for testing deep learning techniques under the OBIA framework (G. Chen, Weng, Hay, & He, 2018; L. Ma et al., 2017.

2.2 Applications of UAV photogrammetry in mining

The method of collecting useful data with UAV is the advancement in technology. Recently UAV also known as drones are used as platforms for data acquisition in natural and complex landscapes also, the UAV (Colomina & Molina, 2014; Salvini, Mastrorocco, Seddaiu, Rossi, & Vanneschi, 2017), play an important role for monitoring. In digital photogrammetry, UAVs equipped with compact or reflex cameras allow high-resolution images of inaccessible areas to be obtained with relatively low costs, avoiding shadow zones compared to terrestrial surveys, acquired from the ground. The technological advances have recently enabled an alternative to an inflexible fixed network of sensors or the labor-intensive and potentially slow deployment of personnel (Coifman, McCord, Mishalani, Iswalt, & Ji, 2006). Initially, satellites and manned aircraft were used for monitoring purposes, but several qualities,

cost, and safety issues have proven these methods to be inefficient. Recently, UAV in the monitoring, management, and control are starting to take center stage (Kanistras, Martins, Rutherford, & Valavanis, 2013; Puri, 2005). The UAVs commonly is considered to be one of the most dynamic and multi-dimensional technologies of the modern era (Park & Choi, 2020). This technology is rapidly strengthening its presence in multiple fields of the human life, varying from commercial tasks from exploration to monitoring (Park & Choi, 2020)the mines area. UAVs are predicted to be the most dynamic growth sector of the world aerospace market this decade , as mentioned by (Kanistras et al., 2013). The UAVs cover a large area in short times with an extreme low cost. The lower cost can also be achieved, since all the equipment is reusable to a different point of interest (Barmpounakis, Vlahogianni, & Golias, 2016). Nevertheless, this technology is progressing rapidly and can be safely termed as a future-proof technology with the widespread commercial availability and decreasing costs. It may also serve as a foundation for more advanced studies related to UAVs but there has been no such detailed framework based on the existing literature.

2.3 Land cover classification

Land cover (LC) information at reclaimed mines is essential for a range of geospatial purposes, such as urban planning, regional management, and environmental management (X. Liu et al., 2017). It also acts as a framework for understanding the dynamic relationships between human behaviors and environmental change (Patino & Duque, 2013). Many predictive models (e.g. ecosystem, hydrological, and transport models) include LC as input variables to simulate natural and anthropogenic processes and the workings of the Earth's surface (Verburg, Neumann, & Nol, 2011). Earth observations from a range of sites, including satellite, airborne, along with and Public Observatories, offer excellent opportunities to classify LC characteristics at mines and improvements across various scales (Katherine Anderson, Ryan, Sonntag, Kavvada, & Friedl, 2017). With the exponential advancement of sensors and computers, vast volumes of remotely sensed spatial resolution (VFSR) images are now commercially available with submeter resolution, enabling the processing of detailed LC details in fine spatial detail (Pesaresi et al., 2013; Zhao, Zhong, & Zhang, 2016).

LC classification using remotely sensed data can be a very challenging task due to the spatial and temporal sensitivity of the imagery. However, LC is much more complicated due to the indirect interaction between trends of land cover and spectral responses reported in photographs. These land uses are usually defined in terms of functions or socio-economic practices rather than tangible types of land cover, which can only be inferred explicitly by interpreting the color, texture, or shape of the picture features (M. Li, Stein, & Bijker, 2016). Sometimes the same land-use types are distinguished by distinctive physical properties as a result, such spectral and spatial ambiguity and heterogeneity make the automated LC classification using VFSR images an incredibly challenging task.

During the last few decades, significant effort has been made to establish automated LC classification methods using remotely sensed data. These techniques, in particular in terms of land cover environments, are constructed mainly based on the spectral characteristics reflected by the physical properties. Pixel Classification Methods, Support Vector Machine (SVM), and Random Forest (RF) have been developed to learn nonlinear spectral feature space at pixel level independently of its statistical properties (Zhong, Zhu, & Zhang, 2015). However, these pixel-based approaches cannot guarantee high classification precision, especially at the fine spatial resolution, because single fine pixels will lose their thematic sense and the discriminatory efficiency of various types of LC (Xia et al., 2017). Object-based approaches, as part of the OBIA, have dominated ground cover detection using VFSR imagery over the last decade (Blaschke et al., 2014). A variety of experiments used OBIA methods to extract land cover information from VFSR images, using spectral, textural, and geometric information from image objects made up of relatively homogeneous neighboring pixels (Myint, Gober, Brazel, Grossman-Clarke, & Weng, 2011).

The major challenges of these object-based approaches are the choice of segmentation scales to obtain objects that conform to particular types of land cover in which over-segmentation and under-segmentation typically occur within the same picture (Ming, Li, Wang, & Zhang, 2015). To date, current methods remain insufficient to properly analyze the data, and no successful solution has been suggested for

LC classification using remotely sensed VFSR imagery. Classification of the LC is less studied due to the ambiguity of the spatial composition and configuration. This difference has resulted in a semantic divide between details from the data and information relevant to users and applications (Bratasanu, Nedelcu, & Datcu, 2010). To bridge such a semantic gap, a variety of studies have attempted to integrate specialist information or ancillary data as a geographical framework for the extraction of land-use features. Additional regional details for defining the LC might not be appropriate for certain regions (M. Li et al., 2016), and contextual environments are also challenging to define and classify the land cover even if complicated constructs or patterns may be identifiable and distinguishable for human experts (Oliva-Santos, Maciá-Pérez, & Garea-Llano, 2014).

2.4 Traditional land cover classification approaches

During the last decade, a significant attempt has been done to establish automated LC classification methods using remotely sensed imagery. For LC, conventional classification techniques may be categorized into pixel-based and object-based methods based on the specific processing, either per pixel or per object (Salehi, Ming Zhong, & Dey, 2012). Pixel-based approaches are commonly used to identify individual pixels in specific land-cover categories based on spectral reflection, without considering the effect of neighboring pixels (Verburg et al., 2011). These methods are also constrained in classification performance due to speckle and increased inter-class variation relative to the coarse or medium spatial resolution of remotely sensed data. post-classification methods have been developed to address the limitation of pixel-based approaches e.g. (Hester, Cakir, Nelson, & Khorram, 2008; McRoberts, 2014). However, such approaches exclude small classes with few pixels or single-family classes, such as small fields. Object-based techniques, as part of the OBIA, have regulated land cover identification using imagery throughout the last decade (Blaschke et al., 2014). These OBIA approaches are focused on similar characteristics objects, which are formed of pixel values all around the image, for the classification of land cover components. The key concerns in implementing these object-based approaches are the collection of segmentation scales to obtain objects corresponding to particular types of land cover in which over-segmentation and under-segmentation typically occur

within the same image (Ming et al., 2015). As a result, the task of classifying LCs is very difficult, particularly for mine restoration, which exhibits high intra-class variations with a wide variety of land cover classes.

In the meantime, such objects also interact with each other by occlusions and shadows, that present significant difficulty to distinguish them accurately and consistently. Till present, no practical approach has been suggested for the automatic classification of land cover using remotely sensed imagery. A variety of studies used OBIA for land use classification including object details with a collection of low-level characteristics (such as spectra, texture, shape) of landforms e.g. (Blaschke, 2010; Blaschke et al., 2014; S. Hu & Wang, 2013).

However, these OBIA methods ignore semantic functions or spatial structures limiting the ability to use low-level features in the representation of semantic features. In this sense, researchers have indicated a two-step workflow, where object-based land covers were initially collected, accompanied by accumulation of objects using spatial contextual descriptive indicators on well-defined LC units. Descriptive measures are usually derived from spatial data to determine their properties or graph-based approaches that structure spatial relations (Walde, Hese, Berger, & Schmullius, 2014). However, additional geographic data for defining the LC may not be appropriate for certain areas, and geographical contexts are also difficult to define and classify complicated patterns may be identifiable and differentiated for human experts (Oliva-Santos et al., 2014).

2.5 Problems in traditional land cover classification approaches

The traditional method for land cover classification is hand-drawn in application combined with structure and architecture. They usually require two distinct but complementary steps for the extraction and classification of features (Volpi & Tuia, 2016). Feature extraction is performed by unique operators on specific parts of the image (e.g. image patches, pixels or areas, objects, etc.) to translate the original feature into compact and/or abstract representations that can be conveniently separated by a classifier (Sun, Zhao, Huang, Yan, & Dissanayake, 2014). These modified spatial features are then used, along with the actual spectra, to train supervised classifiers (e.g. Vector Machine Support), to understand the semantic content of the imagery (C. Chen

et al., 2016). The efficiency of any classifier used is greatly influenced by the transformations used and the resulting spatial characteristics. Popular examples of such operators involve a variety of parameters (Reis & Taşdemir, 2011), classification techniques (Pingel, Clarke, & McBride, 2013), and gradient orientation (Cheng et al., 2013).

However, the hand-drawn approach also requires a time-consuming trial-and-error technique for the extraction and selection of features (Volpi & Tuia, 2016). These hand-drawn features are also task-specific and can be useful for a particular area and/or problem. In comparison, the low-level features used, accompanied by coarse classifier architectures, are inadequate to undermine the underlying semantics or structures due to the absence of larger feature representations (Liu, Hang, Song, & Li, 2017). Thus, minimal classification efficiency has been accomplished for utilizing images that are spectral and structurally challenging.

2.6 Overview of Deep Learning

Deep learning provides a new perspective on classification, where stable, abstract, and observable features are taught end-to-end, spatially, from raw data (e.g. image pixels) to functional labels, which is a key benefit relative to previous approaches (Nogueira et al., 2017). Several deep learning approaches have been suggested, namely deep networks (DN) (J. Chen, Li, Chang, Sofia, & Tarolli, 2015), deep Boltzmann systems (DBMS) (Qin, Guo, & Sun, 2017), stacked self-encoders (SDE) (Yao, Han, Cheng, Qian, & Guo, 2016) and CNNs (Maggiori et al., 2017). Among them, the CNN model is the most well-established process, with remarkable efficiency and considerable popularity in the area of computer vision and pattern recognition (Schmidhuber, 2015), likely visual recognition (Farabet, Couprie, Najman, & LeCun, 2012; Krizhevsky et al., 2012), image processing (Yang, Yin, & Xia, 2015) and scene annotation (Othman, Bazi, Alajlan, Alhichri, & Melgani, 2016). Deep learning convenient throughout the field of image processing (X. X. Zhu et al., 2017) .

These previous works reflect the research emphasis and the topics of deep learning in the field of remote sensing. Note, while this section discusses the most major contributions in the research, it will not include a complete description of deep learning in remote sensing (Zhang, Chen, Wang, Wang, & Dai, 2017; X. X. Zhu et al., 2017).

Instead, the goal is to provide a succinct description of the deep learning approaches used to identify LC using remotely sensed images. The research emphasis on CNNs because they're the most popular and well-established deep learning approach implemented in the field of remote sensing. Deep learning with CNNs are a variant of multilayer neural networks primarily designed to process large-scale representations of sensory data in the form of several arrays, considering local and global static properties (Y LeCun, Y Bengio, & G Hinton, 2015). The fundamental aspect of CNNs is their translational variability by a patch-based process in which a higher-level entity within an image patch can be identified even though the pixels containing the object are modified or blurred.

Deep learning with CNNs were developed to solve an image classification problem, i.e. to assign an entire image to a semantic class, such as a digit (LeCun, Bottou, Bengio, & Haffner, 1998) or an object type (Krizhevsky et al., 2012). The equivalent challenge in the remote sensing domain is to solve the remotely sensed classification role where an image patch is assigned to a particular category, such as vegetation, grassland, barren land, etc. These types of land cover classification activities are directly connected to object identification (Zhang et al., 2017) and position (Long, Gong, Xiao, & Liu, 2017) where translational invariance is the main advantage of CNN to represent data with higher-order characteristics, such as land cover locations. However, this attribute is a major limitation in the LC classification for pixel-level, from which distorted boundaries are created between ground surface objects. Here, we study the classification of LC using CNNs to expand on these challenges in depth and to define research gaps.

2.7 Deep learning for land cover classification

Land cover (LC) classification via CNN can be divided into two types based on processing, namely patch-based and pixel-based techniques. The patch-based methods for LC classification include an image patch going over the given dataset pixel-by-pixel, with strongly overlapping patches used for land cover calculations (Fu, Liu, Zhou, Sun, & Zhang, 2017). Deep learning with CNN models has explored some advances with patch-based. For example, Mnih (2013) suggested a patch-based CNN model develop large-scale visual factors for aerial object labeling. The model created a

complex classification patch, rather than separating a single value image type, where its spatial contextual features were trained to further differentiate the land cover classes. Längkvist, Kiselev, Alirezaie, and Loutfi (2016) used the pixel-wise CNN with patches for the classification of land cover, outperforming the current classification approaches. Sharma, Liu, Yang, and Shi (2017) obtained image patches for all potential positions throughout medium-resolution satellite data and categorized them into land cover classes respectively. However, a patch-wise approach has the drawback of adding objects at the boundary of the defined patches, and the use of the patches adds too many redundant computations, therefore, significantly reducing the actual usefulness of the system for land cover classification (Fu et al., 2017; Maggiori et al., 2017). New studies have changed the emphasis on patch-based CNN for land cover classification towards developing pixel-level frameworks for pixel labeling with remotely sensed imagery (Volpi & Tuia, 2016). Additionally, classification of land cover (Wang, Wang, Zhang, Xiang, & Pan, 2017; Zhao et al., 2016) proposed for the task of classification such as building, grassland, and trees (X. Liu et al., 2017). These FCN methods include convolution and down-sampling along with concurrent up-sampling to sustain the resolution of the outcome map to be like the input image, where the class likelihoods for a feature map were generated for pixel-wise image classification (C. Chen et al., 2016). However, the convolution uses the local features as background, and there is an exchange between down-sampling, which enables the network to see a broad context but removes fine spatial information for specific boundary delineation (Marmanis et al., 2018).

Besides, the up-sampling layers are done in a sense of interpolation at the pixel level that appears to over-smooth the target with inadequate spatial information during the inference stage (Q. Liu et al., 2017). As a result, the FCN models face difficulty in pixel-wise dense classification. Some other research has aimed to limit the disparity due to down sampling and up-sampling, by using the convolution to improve the strength of the expected class labels or by introducing skip links inside the network architectures, such that the fine resolution information was initiated after up-sampling (Marmanis et al., 2018). These extension approaches resulted in distorted boundary demarcations when extended to remotely sensed information of several objects enclosed within each other.

Others took the CNN as a coarse classifier for image classification and further resolved the boundaries during the post-classification process by using the initial image as the guideline for detailed classification Maggiori et al. (2017). For example, Längkvist et al. (2016) combined the regular pixel-wise CNN with features extracted to smooth the classification results by average post-processing. Zhao et al. (2016) presented a contour-preserving CNN approach for image classification and smoothed the classification results by post-processing using a conditional random field (CRF). Likewise, Fu et al. (2017) used FCN-based methods for classification and then conducted the CRF procedure as post-processing to optimize the boundaries. (Marmanis et al., 2018) added a special framework of FCN (SegNet) and filtered the results using CRF for image classification. However, such post-processing techniques (either by averaging over classified regions or using a CRF approach) will only partly resolve the boundary problems created by CNN models by mapping the outputs at the expense of sacrificing fine spatial information. Sometimes, some small features of linearly formed objects were easily removed by post-processing processes, which is inappropriate in the case of remotely sensed image classification.

2.8 Accuracy assessment

This section describes the evaluation metrics used for UAV photogrammetry and image classification. The UAV photogrammetry approach should be validated for the precision result. The Root Mean Square Error (RMSE) is calculated for the evaluation of the photogrammetric works. For the image classification validation compares the model and evaluates using the ground truth. Overall performance of a deep learning architecture in image classification task is described in terms of overall accuracy, average accuracy and per class accuracy of pixel-wise labelling. The performance of the classified algorithm was evaluated using confusion matrix which shows a classification model's efficiency on a series of test data for which the true values are identified. The confusion matrix is used for the validation of the classified results where the row entries are the reference data, and the column entries include the number of pixels expected by the classifier belonging to the column class obtained using MATLAB. From the confusion matrix, indices are calculated for each class, and the average of all the values across classes serves for the multiclass purposes. The

Overall accuracy (OA) Average accuracy (AA) and Kappa (K) is calculated for the validation of results. OA is the simplest and one of the most popular accuracy measures and is computed by dividing the total correct (the sum of the major diagonal) by the total number of pixels in the confusion matrix (Congalton, 1991). The accuracy were averaged respectively by (Fung & LeDrew, 1988) to get the average accuracy. The kappa coefficient of agreement was introduced to the remote sensing community in the early 1980s as an index to express the accuracy of an image classification used to produce map (Congalton, Oderwald, & Mead, 1983; Rosenfield & Fitzpatrick-Lins, 1986).

CHAPTER 3
METHODOLOGY

3.1 Study area in Mae Moh mine

The research area Mae Moh mine is located at Mae Moh District, Lampang Province, Thailand. The Mae Moh mine is the largest Lignite mine with over 630 million tons throughout Thailand, it is one of the places where lignite has been providing fuel for the power plants operated by the Electricity Generation Authority of Thailand (EGAT). The region has a vegetated area of 135 km^2. The reclamation area of the mine was grown with a number of trees for five, ten, fifteen and twenty years. Fig 1 shows the mines area at Mae Moh mine Lampang. The diverse landforms at reclaimed mines area render the study area suitable as an experimental area for this study.

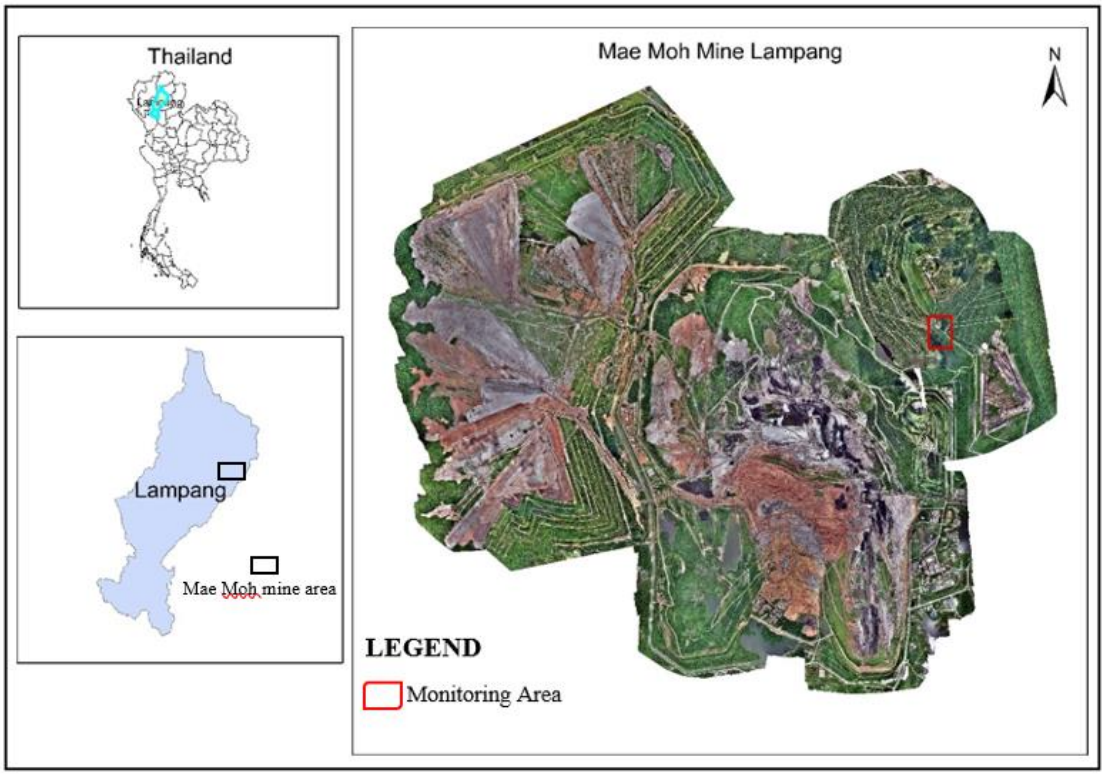


Figure 1 Study Area in red box at Mae Moh mine Lampang

3.1.1 Study area 1

Study area 1 shows the area with the pattern of vegetation (trees at the separate interval). The southwest and northeast UTM coordinates (Zone 47N, WGS84) of this area are (579550, 579630) and (2027500, 2027620), respectively, and covers an area of

2,702.85 m². The selection of the study area was based on its characteristics of vegetation at the reclaimed mines area, which includes wide range of slope values. The height of the tress was between 1 - 3 meters Figure 2 shows the location of the study area 1.

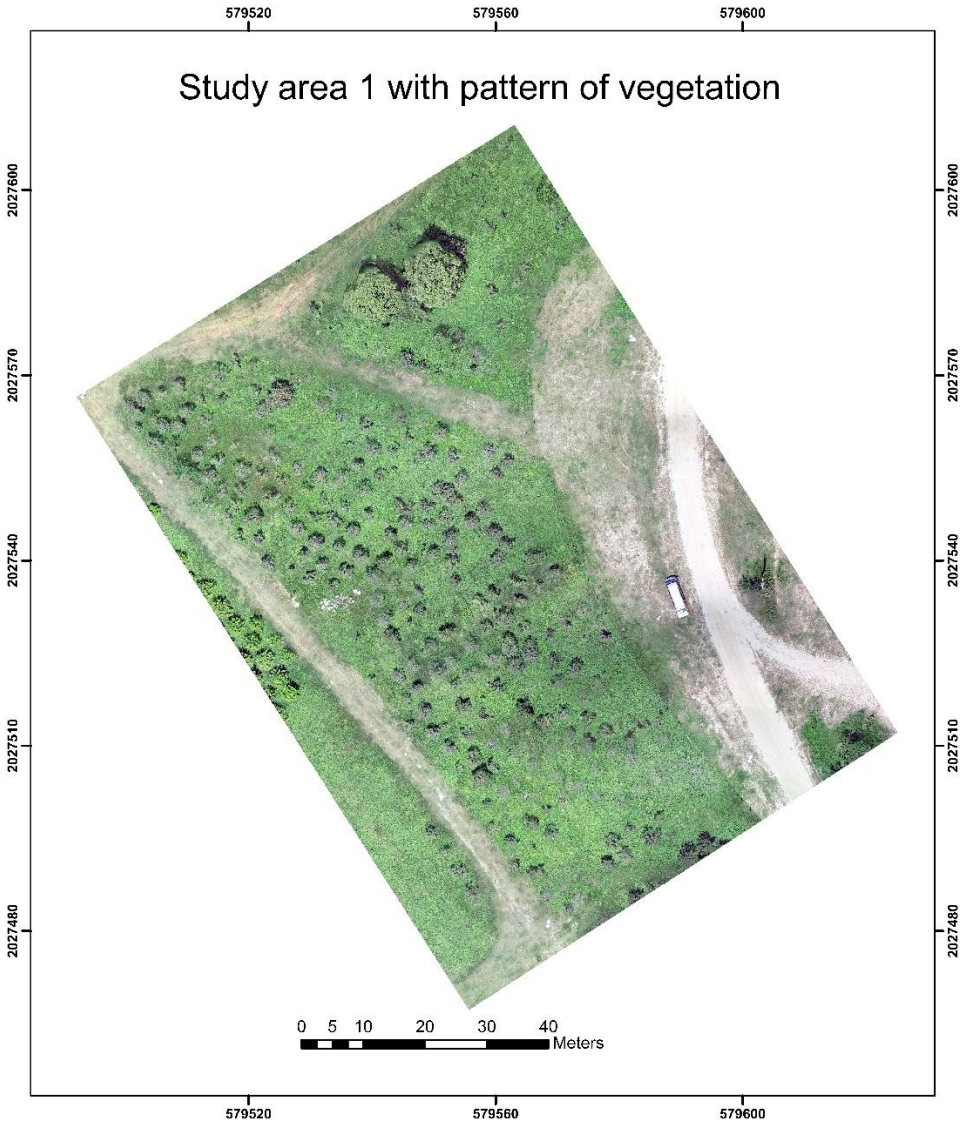


Figure 2 Study area 1 with the pattern of the vegetation at the Mae Moh mine Lampang

3.1.2 Study Area 2

Study area 2 shows the area with complex vegetation types (trees not separated) The southwest and northeast UTM coordinates (Zone 47N, WGS84) of this boundary are (579700, 579800) and (2027140, 2027270), respectively, and covers an area of 2,519.98 m². The selection of the study area was based on its characteristics of

vegetation at the reclaimed mines area, which includes flat area. The height of the tress was between 4 -10 meters Figure 3 shows the location of the study area 2.

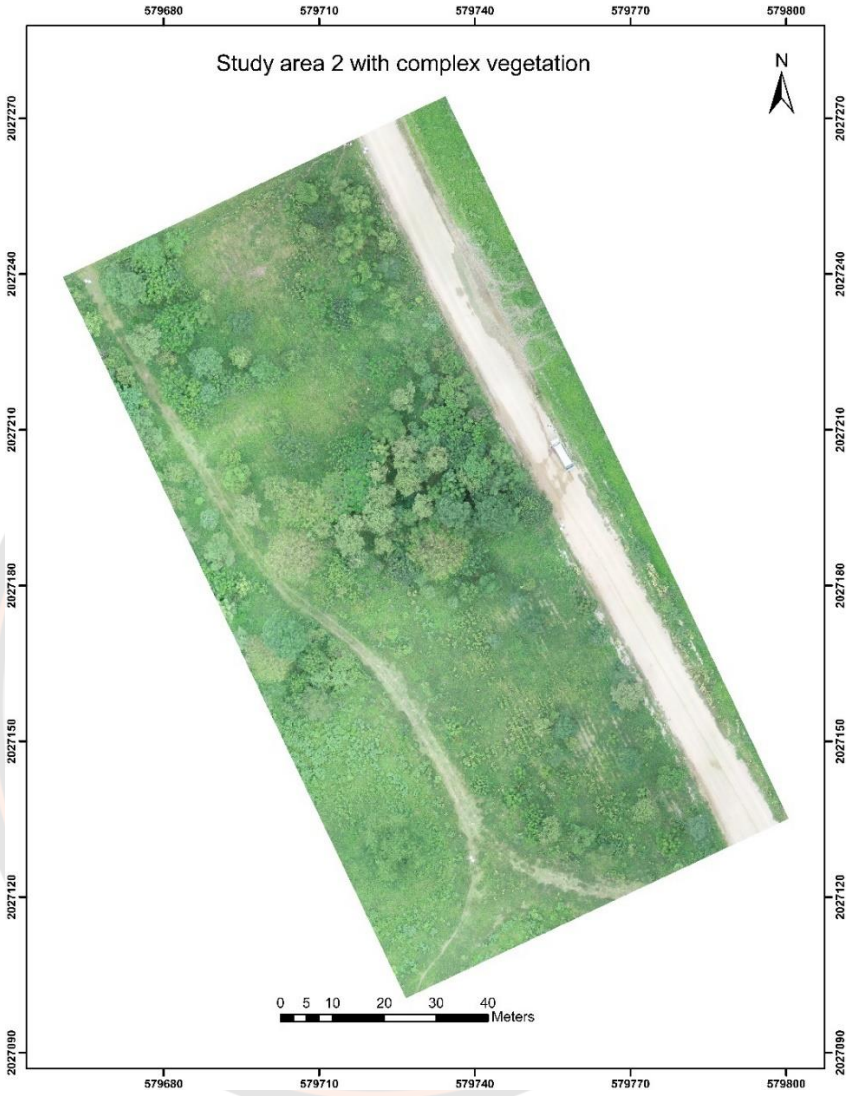


Figure 3 Study area 2 with complex vegetation at the Mae Moh mine Lampang

There were two areas were selected for the research as the areas show the different vegetation types. The two-study areas were selected to classify the land cover at the reclaimed mine and predicting the results of the pattern and complex vegetation types. Also, the slope variation was different in the two areas. Study area 1 has a variation in the slope whereas the second study area is flat.

3.1.3 Comparison between study area 1 and 2

The study area 1 and 2 comprises of the pattern and complex vegetation types. Study area 1 has the training and test area which were captured using DJI phantom 4 pro on 2nd October 2020 whereas study area 2 has training area and test area captured at different time using different drones. Due to the time constrains for research, data for the test area was requested to the EGAT team. The training images were captured on 2nd October by the research team using DJI phantom 4 pro and the test image were captured on 21st July 2020 by the EGAT team using Wingtra one.

3.2 Methodology in Research

3.2.1 Overall methods for monitoring using UAV photogrammetry and deep learning

The overall methodology describes the workflow procedure of the methodology used in this research for the UAV photogrammetry and deep learning. Fig 4 describes the methodology used for the research. There are three steps assigned for the methodology to conduct the research i.e. (1) Data preparation, (2) Data processing, and (3) Data analysis.

1) Data preparation: - Data preparation is also known as data collection describes the device used in the research and the ways of collecting the data for the research. It also describes the ground control points, flight plan, Imagery from UAV, and photogrammetric processing. The details of the data collection are provided in section 3.3.
2) Data processing: Data processing is another important factor for conducting it includes the photogrammetric processing and the generation of orthophoto and DSM. Also, it describes the deep learning approach for land cover classification using a convolution neural network (CNN) in MATLAB. The details of data processing are provided in section 3.4
3) Data Analysis: Data analysis is the concluding factor of the research. It describes the accuracy assessment and the land cover information using the UAV imagery and the deep learning approach using CNN. The details of the data analysis are provided in section 3.5.

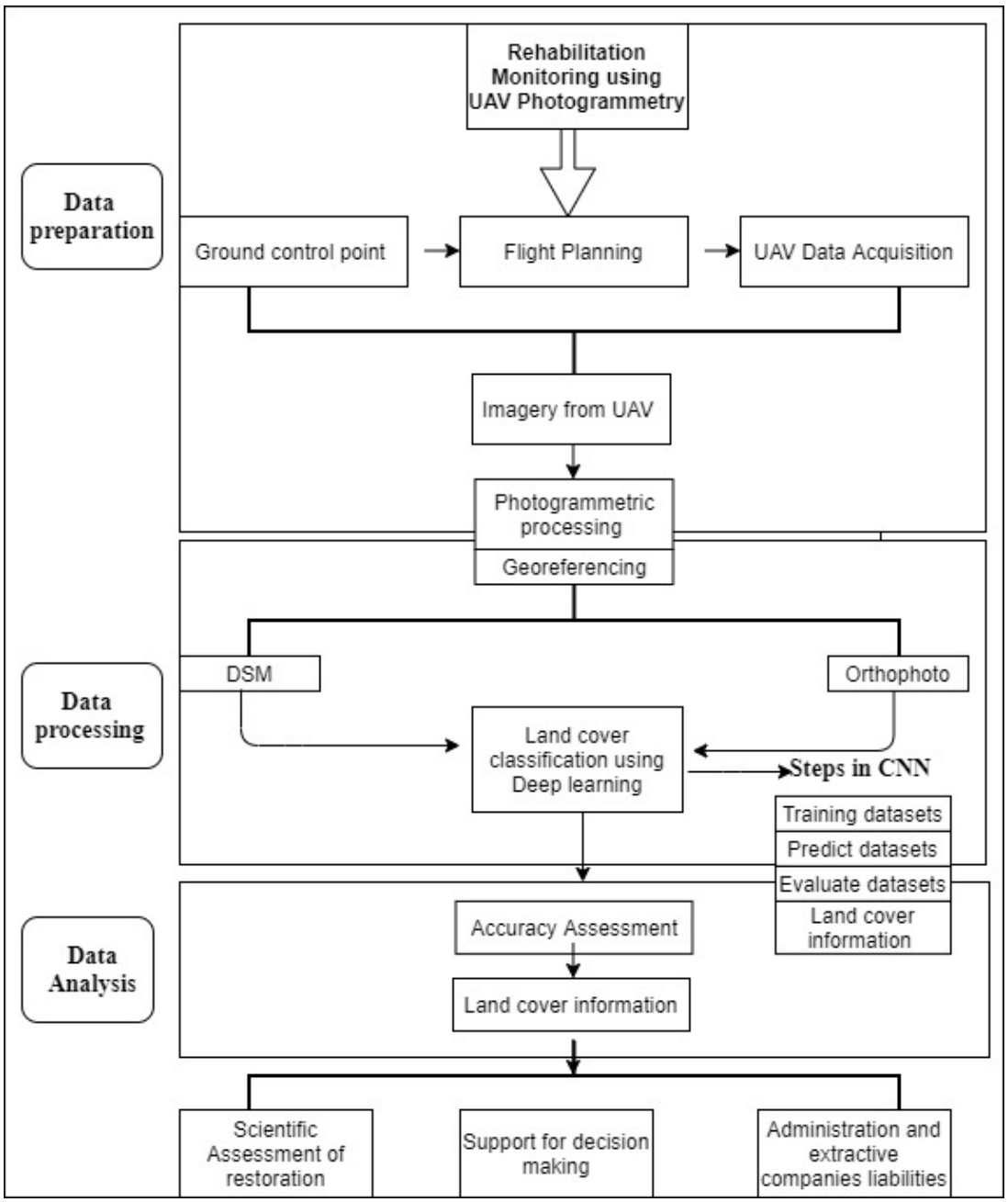


Figure 4 Overall workflow of the research using UAV photogrammetry and the Deep learning approach

3.2.2 UAV photogrammetric approach

The methodology is described for the UAV photogrammetry approach. Fig 5 describes the methodology for the UAV photogrammetric approach and the manual survey to meet the requirements of objective 1 of the research. The UAV photogrammetry comprises the UAV flight plan, collecting imagery from the UAV and

the photogrammetric results (orthophoto and DSM). The details of the flight plan, imagery from UAV and photogrammetric processing is described in section 3. For the UAV photogrammetry approach using the DSM and DTM the height of the tree is calculated to ensure the evaluation of the photogrammetric work. Also, from the manual survey in which the height of the trees was measured using Haga (the device used for measuring the height of the trees) to compare between traditional and UAV approach. The location of the trees was also recorded manually using the GPS. And finally, the evaluation of the height of the trees was intended from the UAV photogrammetry and the manual survey method.

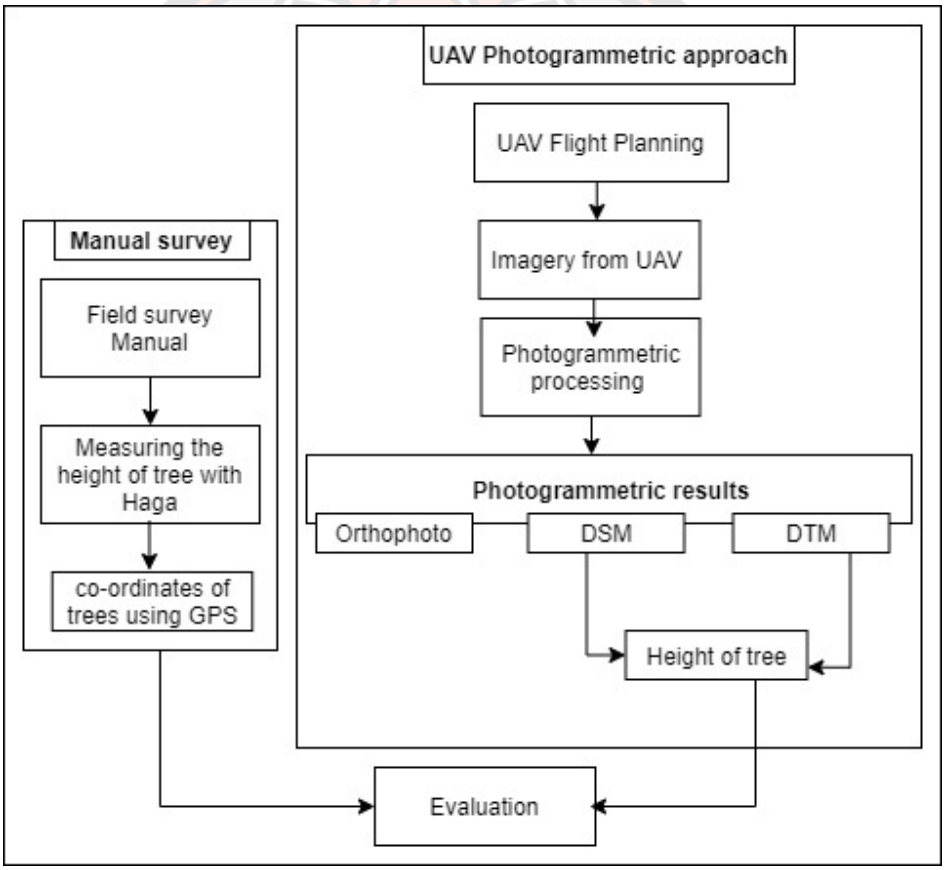


Figure 5 Workflow of measuring the height of trees using Manual and automated method

3.2.3 Land cover classification using Deep learning

The research aims to classify land cover at a selected region of interest by using deep learning with the CNN classification method. Deep learning is adaptable to modern techniques for feature extraction and classification. This technique has shown

capable results and potential in the field of image classification (Al-Najjar et al., 2019). The purpose of selecting deep learning CNN is to work on unstructured data effectively. It can automatically extract features that are required for the detection or classification of the land cover. This patch-level analysis is used with deep learning methods, especially CNN to classify with pixel-level and object-level feature extraction (Sameen, Pradhan, & Aziz, 2018). Also, the images are divided into a grid of tiles of m*m and then each patch is separately analyzed. The size of the image patch used to train the CNN is determined based on the spatial resolution of the RGB image and the expected size of the objects.

Convolution Neural network (CNN): CNN is a class of deep neural networks that are used to analyze visual imagery. They have applications in image and video recording image classification, medical image analysis, and natural language processing. CNN is a multilayer perceptron. Multilayer perceptron can be said as the connected networks that in each neuron in one layer is connected to all neurons in the next layer. These networks classify the image by using the number of convolutional layers that improve the labeling process (Y. Zhu & Newsam, 2015). The deep learning method can automatically learn hierarchically contextual features from the input image and the architecture provide a suitable platform to analyze high-resolution images (Bergado, Persello, & Gevaert, 2016). The information can be extracted at the pixel level, object level, and patch level. Fig 6 shows the flowchart of the deep learning with CNN.

The CNN is based on the three layers

1) Convolution layer
2) Pooling layer
3) Fully connected layer

1. Convolutional Layer: It includes the values of the input image or raw pixel values. It is a layer of access to all the other layers. The activations of perceptron related to the receptive fields of the previous layer are calculated by this layer. Each perceptron is bound to a spatially local area of the input volume, as discussed above. The convolutional parameters of a layer include:

a. The input for the next layer is the number of outputs.

b. The kernel size regulates the local spatial region of the input value.

c. The pixel skips the sliding pane.

d. Padding refers to the scale of the layer.

2. Pooling Layer: This layer is used specifically to resize the spatial representations and accumulate them. for example, is called max pooling. It is quite common to a pooling layer between convolutional layers periodically.

3. Fully-connected Layer: These are generally the last couple of layers for CNN. Perceptron are fully connected to all activations of the previous layer in a fully connected layer. The distinction between a fully connected layer and a convolution layer is that only a local area in the input is connected to the perceptron in the convolution layer, while all the perceptron in the fully connected layer is connected to all the input perceptron (input to the fully connected layer).

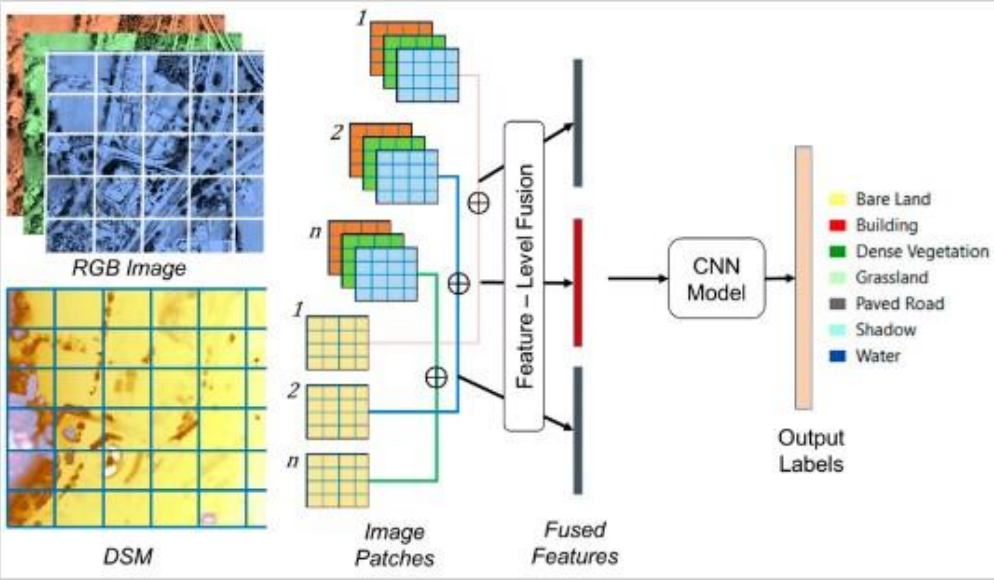


Figure 6 Flowchart of CNN with deep learning model
Source: Al-Najjar et al., 2019

3.3 Data collection

Geographical data can be collected in many ways and different systems are used today. All systems have both advantages and disadvantages and the choice of which method to use usually depends on the project, time acquired for the project, and the budget. This section briefly presents the collection of the images at the Mae Moh

reclaimed mine using the UAV approach. The study sites were surveyed in October 2020 when the vegetation was properly developed with the leaves. Images were captured using the UAV Phantom 4 pro with the normal camera. The resolution of the camera was 20 megapixels. The images were captured with the pix4d software in the single grid and double grid mission of the flight. The UAV was programmed for automatic flight mode at 150 m above ground level with a horizontal speed of 4 m/s from the waypoint. The images were collected continuously at an interval of 2.4 s. There was 80% overlap was along and across the flight plan. The flight plan was uploaded to the autopilot, the system had all the information needed to complete the survey and return it. Each flight block consisted of 200-300 images. The duration of the flight was around 30-40 min.

3.3.1 UAV system for survey

UAV consists of the aircraft and other support equipment. The aircraft, referred is a UAV is operated without an aircrew aboard and is instead controlled by remote control or autonomously. The features of the UAV camera (DJI phantom 4) as shown in table 2 and Fig 7 shows the UAV used for the photogrammetric surveying at Mae Moh mine Lampang.

Table 2 Features of the camera used for monitoring

Parameters	Value
Sensor	1-inch CMOS
Effective pixels	20 Megapixel
Lens	FOV 84° 8.8 mm/24 mm f/2.8-f/11 auto focus at 1m-∞
Mechanical Shutter Speed	8-1/2000 s
Electronic Shutter Speed	8-1/8000 s
Image Size	3:2 Aspect Ratio: 5472×3648
Photo type	JPEG

Figure 7 UAV DJI phantom 4 for the data collection at Mae moh mine Lampang

3.3.2 Flight planning for UAV

The Flight plan is a component designed to facilitate the exploitation of data obtained from the UAV. It provides a hardware-independent interface that isolates payload components from the autopilot specificities. We used the pix 4D software to plan the flight for the UAV. The flight plan was performed using DJI phantom Pro UAV with pix4D mobile application for autonomous flight planning. The description of the flight parameters while capturing the imagery at Mae Moh mine is as shown in Table 3.

Table 3 Parameters for flight plan at reclaimed area in Mae Moh mine Lampang

Parameters	Value
Speed	5.8 m/s
Flying height	150 meters
Front overlap	80%
Side overlap	80%
Camera angle	75^0

The drone was equipped with autonomous flying modes for image capturing mode. Image capture was done in a pre-defined flying path. Coupled with the excellent

obstacle avoidance systems, the DJI Phantom 4 was used for surveying and monitoring purposes.

3.3.3 Image acquisition of UAV survey

The image was collected using a normal RGB camera. To obtain the best images, the camera was calibrated by setting its parameters on the ground. Usually, it is required to set the parameters before the flight and leave the shutter speed and diaphragm opening to function automatically. To examine and evaluate camera calibration one has to pick the drone and move around 180^0 to calibrate. Several factors are influencing a flight plan, including the desired resolution, the area to be covered, and the height variations over the terrain. The expected resolution has a direct impact on the height of flight. The area to be covered determines the number of flight lines. Moreover, the variation of the height of flight influences the overlapping values between images. The resolution, pixel size, or ground sampling distance (GSD) is the size of the projected pixel on the ground. It is directly dependent on the sensor's size and height of flight. As far as the terrain is not entirely flat in the study area, thus this value is an average of the different pixel sizes in the dataset. Indeed, it dictates the accuracy which is possible to vectorize objects on an ortho-photo.

In this study, the vectorization was realized on the point cloud. In the end, the global set of pictures presented a good level of overlapping and was of good quality. The Study area 1 and 2 consists of 272 and 376 UAV imagery respectively.

3.3.4 GNSS survey for georeferencing

To provide an independent network of GNSS, a total of 5 targets were deployed at both the study area A Trimble receiver was used to collect coordinates of the site amenities and ground control points (GCP) for the UAV imagery. To obtain the best resolution possible it is necessary to use ground control points. Placement of target and measuring location of targets by GNSS is in many cases fast and efficient. Five targets were placed as homogeneously as possible on the site to obtain the best referencing on the whole model. Fig 8 shows the GCP used for the georeferencing indicating A for study area 1 and B for study area 2 and Fig 9 shows GNSS survey for Target as GCP in georeferencing at the study area.

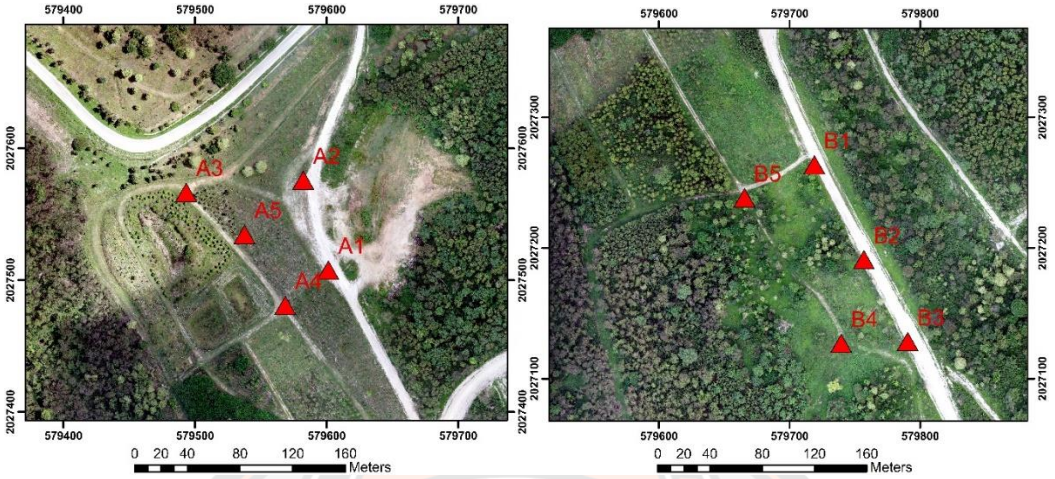


Figure 8 GCP used for the georeferencing using the GNSS indicating A for study area 1 and B for study area 2

Figure 9 GNSS survey for Target as GCP in georeferencing at the study area

3.3.5 Ground truth survey for validation

Ground truth data for this study is collected with the GPS for measuring the height of the trees and determining the locations of the trees on the orthophoto for the validation. The height of the trees was recorded manually using the device Haga (a device used for measuring the height of trees). Fig 10 Measuring the height of the trees using Haga. Fig 11 represents the location of the trees at the study area.

Figure 10 Measuring the height of the trees using Haga (a device used for measuring height)

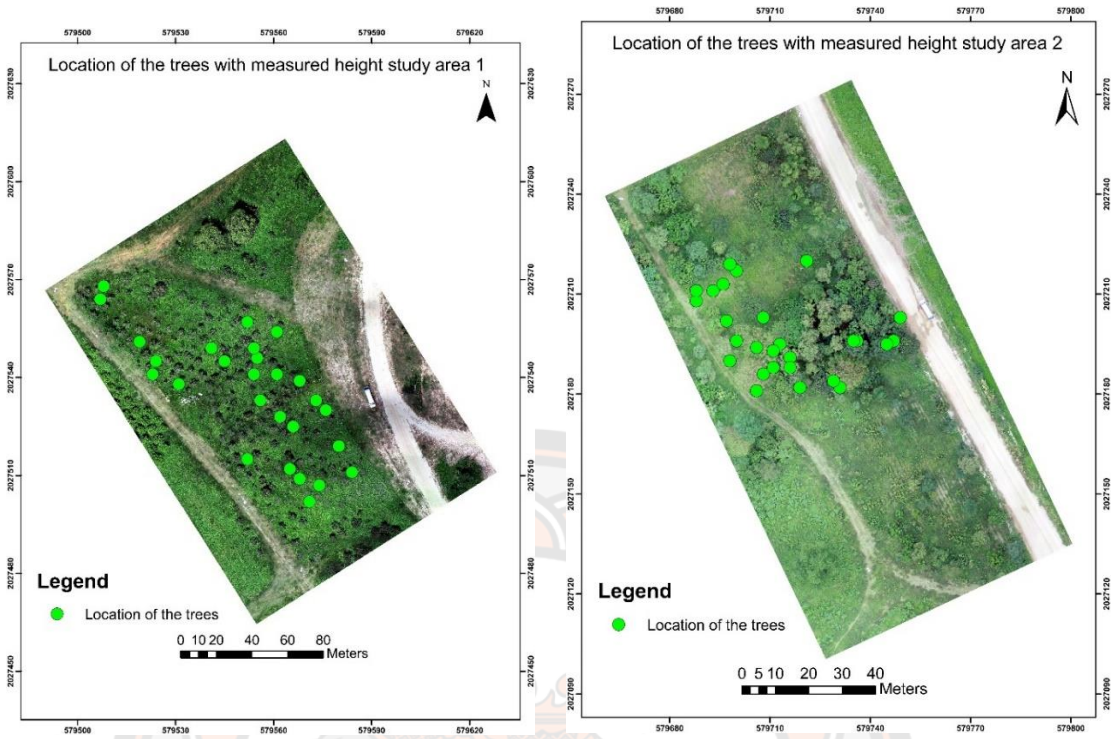


Figure 11 Location of the trees at the study area 1 and 2

3.4 Data processing and quality assessment

3.4.1 Photogrammetric processing

In the area of study, we will be able to see three-dimensional earth surface image on a smaller scale, commonly called a model. That is why data mapping results by using this drone, besides, produce images with a very high resolution it also produces a three-dimensional view of the surface from the recorded object. The photogrammetric processing is done using the software Agisoft Photoscan as shown in Fig 12. Align photo is done to identify the points in each photo and make the matching process of the same point in two or more photos. The align photos process generates the initial 3D model, camera position, and photo in every recording, and sparse point clouds to be used in the next stage.

After aligning the photos, the process is data rectification. Rectification is a re-exposure of a photograph so that the tilt of the image is lost and at the same time adjusts the average photo scale to one another. Rectification of aerial photographs using Ground Control Point (GCP). With rectification, we make photos completely upright / without tilt and average scales are following the desired scale of each photo. while the

depth parameter filtering shows how the treatment of a suspected high point. Its characteristics are usually the value of its altitude is much larger or much smaller than the dots around it.

The next step is a 3D model or mesh. This is one of the main outputs of air photo processing in Agisoft Photoscan. The 3D model will be used as the basis for making DEM both Digital Surface Model (DSM) and Digital Terrain Model (DTM) and also orthophoto. At the time of build, a mesh appeared Mesh Parameter option. For Surface Type, there are two options, Height Field and Arbitrary. The next step in the Aerial Photogrammetric process is to build a digital elevation model (DEM). Digital Elevation Model is a digital field model in raster/grid format that is usually used in spatial analysis.

The last step in an aerial Photogrammetric process is to build an orthophoto and DEM. Orthophoto is generated after making stage Dense Point Clouds, Mesh, and DEM. The Orthophoto is an aerial photograph that has corrected geometric error using DEM data and GCP data so that it can be utilized for the benefit of mapping without any scale inconsistency along with photo coverage.

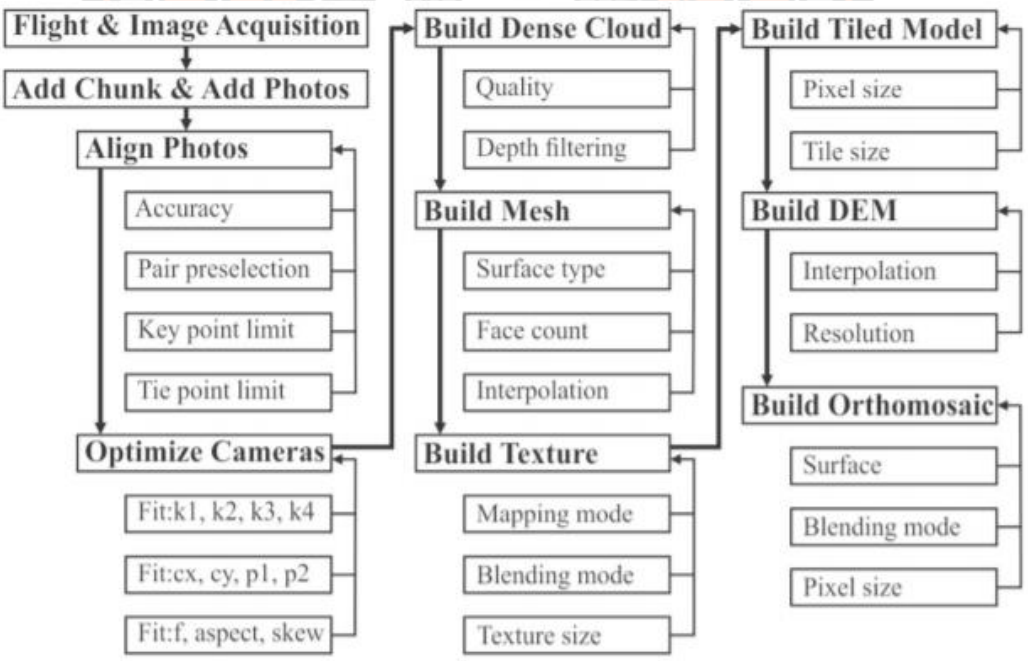


Figure 12 Processing steps in the Agisoft Photoscan

Source: Yucel & Turan, 2016

3.4.2 Accuracy assessment of photogrammetric results

The calculation of spatial error can be implemented in numeral ways (Z. Li, 1993), but positional accurateness is assessed by RMSE (Greenfeld, 2001). This statistic is calculated as the square root of the average of a set of squared variances among calculated values and independent control measurements (of the same points) of superior accuracy. Since it shows all error influences, the RMSE is represented as an absolute accuracy measure (Kraus, 2011). The RMSE is calculated for the difference between values predicted by DSM and the value observed manually. The RMSE represents the square root of the DSM and the differences between predicted and observed values for calculating the height of the trees. The root mean square error (RMSE) is a method for calculating a model's error in predicting quantitative results. RMSE is suitable for the prediction task, because it measures inaccuracies on all ratings, positive and negative. RMSE is obtained using equation (1) as shown below.

$$RMSE = \sqrt{\frac{\sum(x_1 - x_2)^2}{n}} \qquad (1)$$

The accuracy of the photogrammetric results is done using ArcGIS and Microsoft excel. The height and location of the tree are located manually and from the photogrammetric results DEM. Also, using the DEM the average height is calculated from the maximum and minimum height at the same location with trees. The average height is obtained from the mean of the values from the DEM. From the average values of DEM, the average height is calculated by subtracting the height from the DSM and the average height of DEM. Similarly, the standard deviation is calculated, and hence to perform the accuracy assessment, the set of 28 known ground points was calculated for two study areas as they were located within the area covered by the aerial images. The GCPs were used for georeferencing the reconstructed model in Agisoft Photoscan, the accuracy evaluation was yielded by treating the remaining 28 points as an independent data set GCPs to validate the positional accuracy of the orthophotograph and DSM.

3.5 Data analysis and validation

The collected raw data may be noisy, incomplete, or inconsistent, and before using the data as input in the model, it is required to pre-process it by removing errors and outliers and fill missing values. Other tasks include the integration of datasets and transform the data into an appropriate format, so it is readable depending on the tool deployed to perform the deep learning process. The training dataset is applied to show the model how to estimate the feature which will be able to predict the efficiency of observation for the analysis (James, Witten, Hastie, & Tibshirani, 2013). the validation dataset is applied to choose a suitable architecture for the model. Finally, the testing dataset allows the model to iterate and tune the different parameters until the model is ready to be deployed. The analysis focuses on the effectiveness of the method and on the computational criteria (training and testing time) for its application (Sammut & Webb, 2011). A high analysis refers to a simplistic DL model that poorly maps the relations among characteristics and results (under-fitting) whereas a variance means a DL framework that matches the training data but does not generalize well to predict new data (Boutaba et al., 2018).

3.5.1 Land cover classification with MATLAB

Accurate land cover classification is based on remotely sensed data is essential for relevant authorities to evaluate the time or event-based shift in reclaimed mining areas. Every approach that has user flexibility in region selection offers a great deal of convenience during research, since the analyzer may need to focus on a single area of interest instead of working with all remotely sensed data. Image classification refers to the task of extracting information classes from a raster image. Pixels are the smallest units represented in an image, and image classification uses the reluctance statistics for individual pixels to group them. The classification of land cover is based on the labeled data defined in the training set.

Land cover classification in MATLAB has 3 steps described:

1) Create a training set of pixels belonging to a certain group from orthophoto and DSM.

For the process of image classification, we need the orthophoto and DSM to be labeled according to their features in MATLAB using the image labeler application available in MATLAB. The data labeling process is done using the pixel labeling using the polygon feature. There was no prelabelled dataset available for the study area thus, the manual labels were generated for the specific period. However, extra time was allocated in this study to properly label the images. Precise labeling is highly encouraged for this research to obtain a precise result. The correct method of labeling features for deep learning is by marking the pixels of the mentioned topographies in detail. All visible areas should be marked and those must be within its boundaries. Extra attention should be given to annotate marks, shaded areas, or anything on the ground level. Images were finally extracted in the PNG format after completion of the labeling. MATLAB was employed for this step since it offered the most straightforward process. It provided a user-friendly interface and an exportable setting. The labeled result from this setting can be easily applied for training the data set. Another important item for deep learning images is to make sure that both the original image and the resulting image are of the same resolution. The later section will dive into how the model is trained and the final model after testing.

2) Training the Model.

The next step in the development of a deep learning model is to train it using the labeled data and orthophoto in a dataset. Model training was the easiest initial option where it has a built-in function in MATLAB. A line of code can represent a whole architecture of UNet layers. It reduced the overall coding time. However, it was slightly problematic when an error was found since most functions cannot be examined directly. The images were first loaded into the workspace by storing them in Datastore. This step was done for image files by directing them to the corresponding folder location in the computer without separately loading them into the MATLAB workspace. Original images were stored with the image Datastore function, and the mask images were stored with pixel Label Datastore function. An image Datastore is commonly used for applications related to image classification tasks necessarily; however, pixel Label Datastore uses the read function in

MATLAB but more specific to reading pixel label data for the classification task. It combined the mask image, classes given, and ID of each class. The class for each pixel is defined with the code in MATLAB from the marked image. The pixel value ranges from 0 or 225 respectively. The original image was transformed into grayscale which uses an 8-bit pixel range from 0 to 255, so the labels generated were using the same range. Both datastores were then combined in one datastore that connects original images to the corresponding mask. UNet network architecture creation in MATLAB was straightforward. The function used was UNet layers, and the parameters feed were the sizes of images and the number of classes assigned to the mask. This input will let the network know that images stored in datastores are of such sizes and pixels that can be labeled as either of the five classes.

Fig 13 displays the accuracy and loss during the training and validation data during the model training. It shows that the training accuracy increases and loss gradually decrease toward zero the validation loss was initially very high then it decreases and follows a similar trend after 15 epochs as the model generalizes. The performance of this model served as the model was performing well. If the validation curve does not follow the trend of the training curve asymptotic toward zero, then the model is overfitting. Overfitting means the model is simply memorizing the training data by heart, thus unable to predict new data properly (Goodfellow, Bengio, Courville, & Bengio, 2016).On the contrary, the model is considered underfitting when the training loss is not asymptotic toward zero meaning it is too simple to learn the complexity of the current data structure (Géron, 2019). It means that the best result was obtained by the model as there is consistency for accuracy and loss after the 10^{th} epoch.

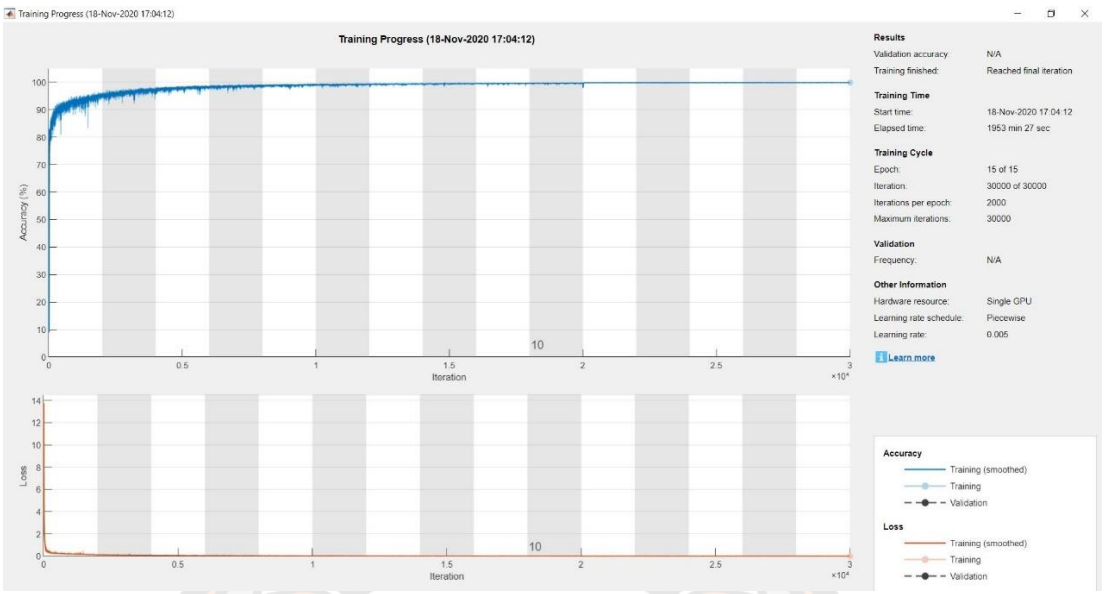


Figure 13 Accuracy and loss of the model during training the dataset

3) Prediction of the results

It ran through parameters values for training in the MATLAB during training and validation parameters. It was first tested qualitatively by overlaying the predicted mask with the original image. A common model evaluation would start by observing the accuracy, and the confusion matrix to determine how well the model performed. However, the deep learning model is a new approach for monitoring the reclaimed mine in the mining industry.

This initial test result showed that the model was able to predict all pixels on the image used in the training data. The resulting prediction quantitatively proved effective vegetation detection. Areas selected were meant to test the model's capability of monitoring the vegetation, the model predicted the vegetation with the test image.

3.5.2 Preparing ground truth data

The ground truth object contains information about the data source, label definitions, and marked label annotations for a set of ground truth labels. These are the label which is performed in the image labeler application in the MATLAB using point, polyline, polygon, bounding box, and segmenting tools available as shown in Fig 14. The Ground truth data comprises the information about the data source, defined label, and the labels marked for the annotations. Image annotations are designed to increase

the productivity of labeling and also efficiency to reduce the efforts. Also, the labeling interface allows adapting the specific needs of the area for the classification. All tasks for image classification on multiclass, drop-down on a long list, and systematic classification is done to handle complex ontologies. The point, polyline, polygon, bounding box, and segmenting tools are available.

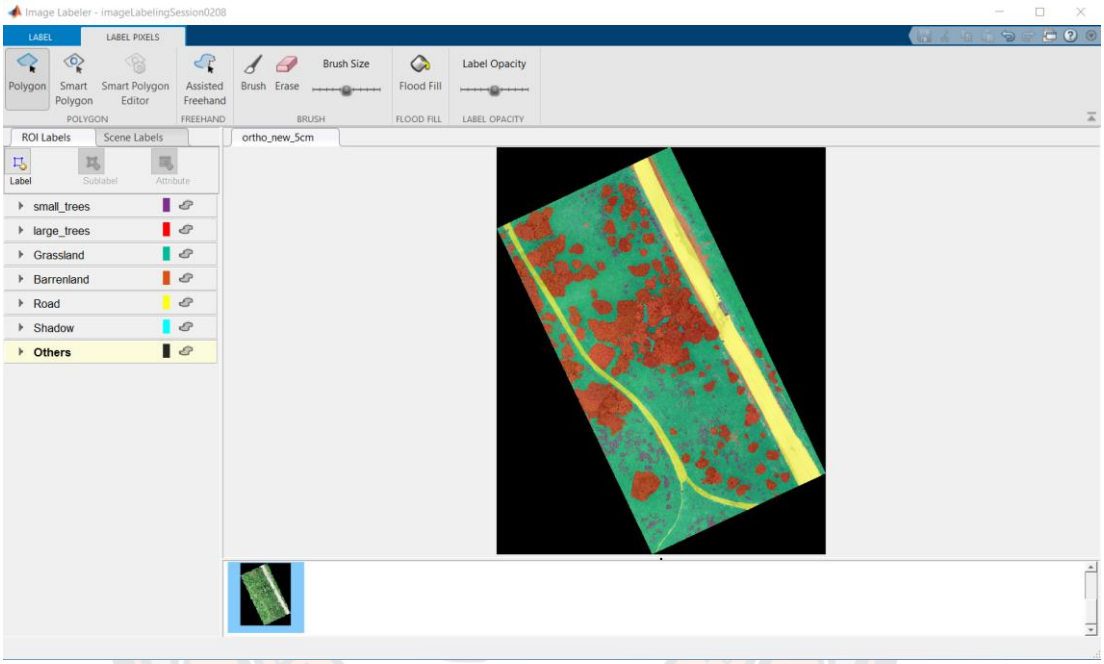


Figure 14 Labeling the image in MATLAB using image labeler application

3.5.3 Validation of classification results

For the evaluation of the deep learning, the statistical methods, major misclassified classes will undergo a visual explanation for identifying the inconsistency in the classifier. This test considered how many pixels were correctly labeled. Three metrics, namely overall accuracy (OA), Per class accuracy (PA), Average accuracy (AA) and Kappa(K) coefficient were used to evaluate the quantitative performance of different classifiers. The parameters needed to calculate these metrics acquired from the confusion matrix.

Overall accuracy represents the correctly classified areas for the whole image and is calculated by dividing the number of correctly classified pixels to the total number of pixels in the confusion matrix. (T_{ab} or the sum of major diagonal) by the

entire sum of pixels (N) in the confusion matrix, as shown in equation (2), and PA was calculated using equation (3);

$$\mathrm{OA} = \frac{\sum \mathrm{T}_{ab}}{\mathrm{N}} \quad (2)$$

$$\mathrm{PA} = \frac{\sum \mathrm{T}_{ab}}{p_{\mathrm{i}}} \quad (3)$$

Where T_{ab} is the total number of the correctly classified pixels in row m and column n and p_i is the total number of pixels in the row. Average accuracy (AA) was computed as in equation (4) defining C as the number of the classes

$$AA = \frac{\sum_1^m PA}{C} C \quad (4)$$

The Kappa coefficient is the proportion of agreement after the possibility agreement is excluded (Landis & Koch, 1977). Kappa provides data on comparing the effectiveness of a classifier, Several researchers have clarified the suitability of the kappa coefficient in accuracy evaluation of image classification to include more clear measurements such as averages or per-class accuracy and confusion matrices (Foody, 2020; Pontius Jr & Millones, 2011). Alternatively, the Kappa coefficient is the standard calculation of the definite and predicted values reliability of a confusion matrix that takes into consideration of the non-diagonal elements (Foody & Mathur, 2004). Kappa analysis is a powerful way to validate a single mistake matrix and compare the difference between error matrices (Ben-David, 2008). 0 to 1 is the value range for kappa. The analysis of Kappa produces a K metric, a quantifiable estimate of the accurately labeled pixels and the degree of consensus (Landis & Koch, 1977). The kappa was calculated as in equation (5):

$$k = \frac{N \sum_{p,q=1}^m T_{ab} - \sum_{p,q=1}^m o_i, p_j}{N^2 - \sum_{p,q=1}^m R_i, S_j} \quad (5)$$

Where m is the number of classes, T_{ab} is the number of correctly classified pixels in row p and column q, o_i is the total number of pixels in row p, p_j is the total number of pixels in column q and N is the total number of pixels.

Model validation is used to predict how effective an estimator is on data which has been trained on and applicable to new input. The training and validation time on the parallel axis and a reliability metric is plotted by validation curves. The accuracy curve is a standard measure of reliability derived from the matrix of uncertainty and displays both accurate and inaccurate classifications (Abd & Alnajjar, 2013). The accuracy is calculated as shown in equation (6).

$$Accuracy = \frac{\sum_{p=1} C_{ab}}{\sum_{p=1}^{n} \sum_{q=1}^{n} C_{pq}} \quad (6)$$

Where C_{ab} is the correct classification on the diagonal, n is the number of classes, C_{pq} is the number of times items of class p were classified as class q (an incorrect classification) and $\sum_{p=1}^{N} \sum_{q=1}^{m} C_{pq}$ is the total number of samples that were analyzed.

Another metric for understanding the potential of DL is the loss curve which proves whether the method of optimization and relative development in learning improves in several epochs through the training (Foody, 2002). The loss function is obtained using equation (7).

$$\text{Loss} = 1 - \text{accuracy} \quad (7)$$

CHAPTER 4
RESULTS AND DISCUSSION

In this section, the results from the UAV photogrammetry and the deep learning are described. For deep learning using the results from UAV photogrammetry, the best parameters are applied to the training data, and later the trained model is used to classify the test data enabling the classification of land cover at reclaimed mine. The final map is discussed in the section and analyzed based on the metrics presented in the section. Finally, a visual evaluation of the map is shown to understand the correct classification results.

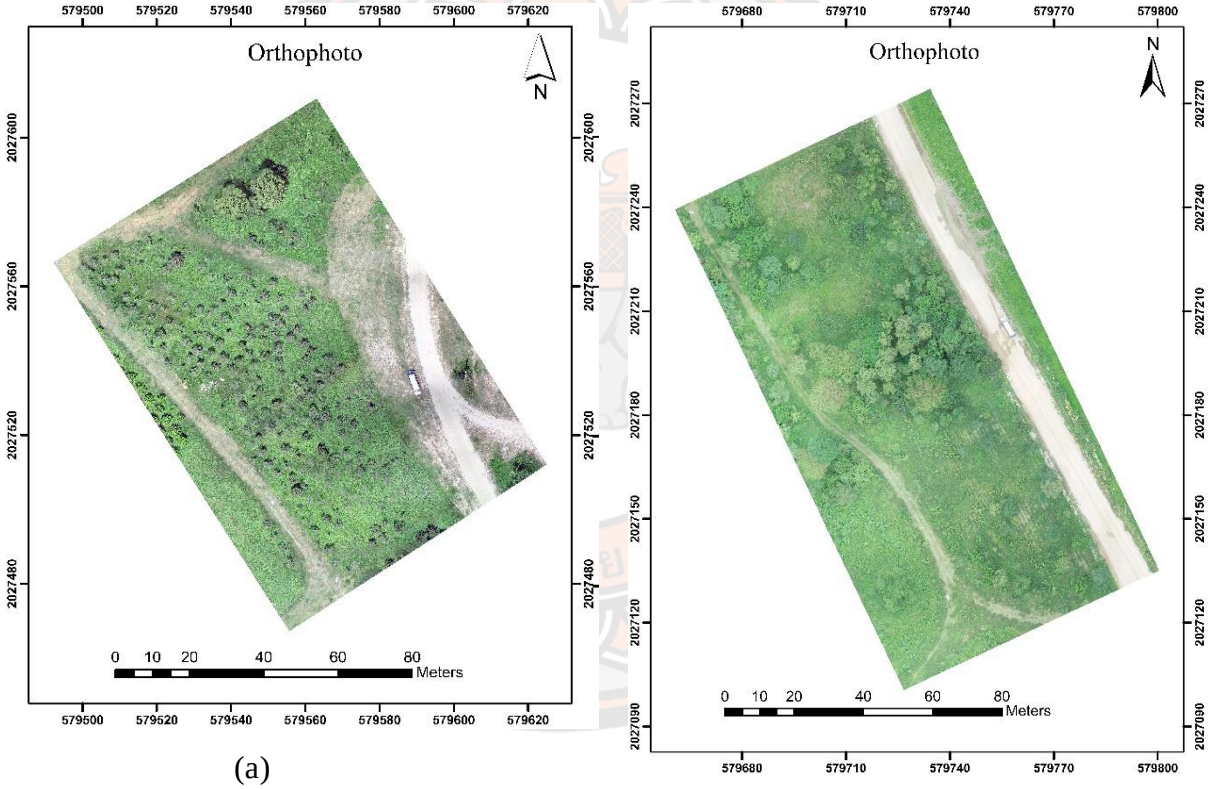
4.1 Photogrammetric results of Study area 1 and 2

Our results show that the UAV images used in our studies were able to capture the vegetation at the reclaimed, mines with high resolution. In the two study areas presented here,1) area with the pattern of vegetation and 2) area with complex vegetation, orthophoto and DSM generated was in an automated way. The UAV used in this thesis was easy to use and highly flexible in time and space, as it was transported and operated by a single person and without applying for flight permissions (autonomous flight). Ground control points were also needed for validation. UAV-technology has, however, the potential to reduce fieldwork in comparison to the amount that would have been necessary to assess the entire surveyed area by fieldwork. Therefore, UAVs are especially suited for areas that are difficult to access (mines area). The potential of UAVs in remote areas is, at the moment, hampered by flight regulations that restrict their use to the range of vision. Nevertheless, it is easier to transport and a backpack with ground-control-station equipment to reclaimed mines area. Collecting the UAV imagery is cost-effective as described by (Karen Anderson & Gaston, 2013). Finally, the image data were analyzed to extract the desired information automatically.

4.1.1 Orthophoto

A set of aerial images for study areas 1 and 2 respectively were taken from automated flight paths was used for the generation of orthophotos. In areas where the terrain is flat, corrections for tilt are enough. The correction process is called rectification and requires four Ground Control Points for each image. When the terrain

is not flat, relief distortions have to be eliminated which requires elevation data that usually is presented in the form of a DEM. The orientation parameters were obtained from photogrammetric measurement. The image is then orthorectified, which means that the image is cleared from scale distortions due to terrain height and camera tilt, resulting in the same scale everywhere in the image. An orthophoto was transformed from a perspective projection. Finally, generated orthophotos for study areas 1 and 2 in this dataset are tiled together in Figure 15 (a), (b). In that figure, occluded cells in each of the tiles are filled using those in overlapping orthophotos. Two orthophotos produced from UAV image enables GIS-based studies of the reclaimed mine.



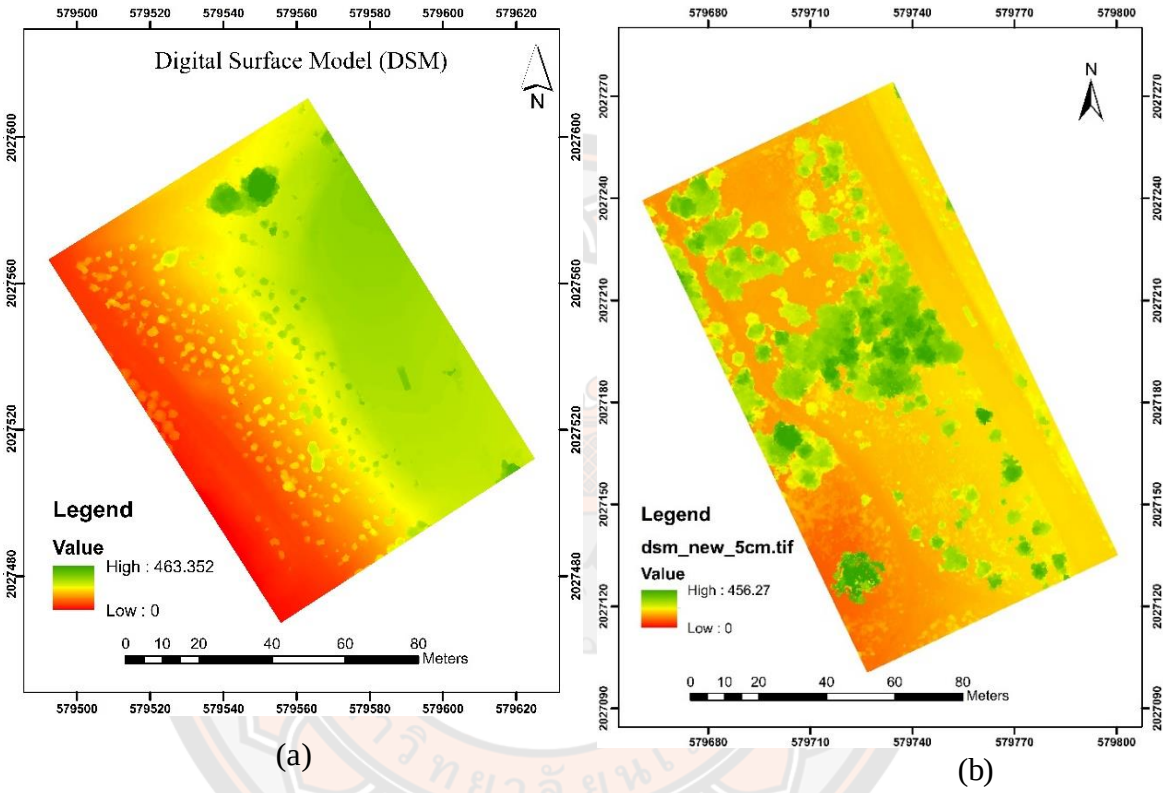
(a) (b)

Figure 15 Orthophoto of (a) study area 1 and (b) study area 2

4.1.2 DSM

A set for study areas 1 and 2 respectively were taken from automated flight paths was used for the generation of the DSM. A DEM is usually provided by photogrammetry for monitoring the mine reclamation analysis, the UAV image data is advantageous since the image data is used for the automated generation of elevation models and manual measurements. The orientation parameters were obtained from photogrammetric

measurement. DSM is used for interpretation of the height due to the high point density. DSM was produced covering the mines study area 1 and 2 because the manual measurements were very time-consuming. The DSM model covers the reclaimed mines area as shown in Fig 16 (a), (b) which is covered by vegetation. Automatically generated DSM has a resolution of 5 cm for both study areas 1 and 2.



(a) (b)

Figure 16 DSM of (a) study area 1 and (b) study area 2

4.2 Accuracy results of photogrammetric accuracy for study area 1 and 2

In this study, height was extracted to assess the accuracies of the photogrammetric works. The accuracy of all photogrammetric projects was evaluated using the surveyed points for trees and the calculation of the height from DSM and DTM using ArcGIS. Also, the average DTM and from the maximum height was calculated as shown in the table for study area 1. Also, to compare quantitatively the difference between UAV and the traditional approach, the mean difference (MD), the RMSE were used in this study. For the study are 1 table 4 shows the height of the trees and error calculated using DSM and DTM results from the UAV photogrammetry.

Table 4 Study area 1 with the height of DSM, DTM showing the difference between DSM and DTM and error

No.	Height of tree	DSM	DTM	DSM-DTM	Error
1	2.9	456.636	453.749	2.9	0.0
2	3.7	455.062	451.908	3.2	-0.5
3	2.7	455.391	452.875	2.5	-0.2
4	2.8	453.991	451.936	2.1	-0.7
5	3.2	453.840	451.773	2.1	-1.1
6	2.8	455.930	453.453	2.5	-0.3
7	3.3	457.490	453.855	3.6	0.3
8	2.9	456.924	454.260	2.7	-0.2
9	3.3	454.925	452.795	2.1	-1.2
10	2.5	455.193	453.093	2.1	-0.4
11	2.8	453.382	450.660	2.7	-0.1
12	2.7	450.195	448.220	2.0	-0.7
13	2.9	451.965	449.397	2.6	-0.3
14	3.7	453.376	448.894	4.5	0.8
15	3.8	453.980	451.026	3.0	-0.8
16	2.5	453.264	450.475	2.8	0.3
17	2.8	452.495	450.285	2.2	-0.6
18	3.1	453.730	450.901	2.8	-0.3
19	3.0	452.889	449.727	3.2	0.2
20	3.8	452.570	448.749	3.8	0.0
21	3.2	449.760	446.465	3.3	0.1
22	3.7	449.515	446.083	3.4	-0.3
23	3.5	449.449	446.294	3.2	-0.3
24	3.8	449.131	445.436	3.7	-0.1
25	3.3	448.218	445.102	3.1	-0.2
26	3.7	449.818	446.283	3.5	-0.2
27	3.7	448.915	445.425	3.5	-0.2
28	3.3	451.012	447.414	3.6	0.3

The height was extracted accordingly as the method used in study area 1 from DSM and DTM. Also, the results from the traditional approach were used for the calculation of the error at study area 2. Table 5 shows the height of the trees and the error calculated using DSM and DTM results from the UAV photogrammetry.

Table 5 Study area 2 with the height of DSM, DTM showing the difference between DSM and DTM and error

No.	Height of tree	DSM	DTM	DSM- DTM	Error
1	6.0	456.633	451.052	5.6	-0.4
2	5.5	453.993	449.417	4.6	-0.9
3	4.0	452.412	449.090	3.3	-0.7
4	9.0	458.590	449.772	8.8	-0.2
5	6.0	448.896	443.638	5.3	-0.7
6	2.0	447.752	445.665	2.1	0.1
7	7.0	452.297	444.571	7.7	0.7
8	5.5	448.980	443.631	5.3	-0.2
9	5.0	448.987	444.765	4.2	-0.8
10	7.0	450.777	443.881	6.9	-0.1
11	7.0	450.556	444.464	6.1	-0.9
12	10	453.004	443.251	9.8	-0.2
13	6.0	450.225	444.357	5.9	-0.1
14	5.0	448.927	443.150	5.8	0.8
15	3.0	446.293	443.133	3.2	0.2
16	6.5	449.124	443.012	6.1	-0.4
17	7.0	449.400	443.065	6.3	-0.7
18	5.0	448.418	443.363	5.1	0.1
19	5.0	448.115	443.162	5.0	0.0
20	5.5	448.580	443.276	5.3	-0.2
21	7.0	449.805	443.203	6.6	-0.4
22	7.0	450.475	443.047	7.4	0.4
23	7.0	450.114	443.252	6.9	-0.1
24	7.0	449.902	443.463	6.4	-0.6
25	6.5	451.461	444.143	7.3	0.8
26	8.5	451.986	444.331	7.7	-0.8
27	9.0	452.548	443.908	8.6	-0.4
28	7.0	449.988	443.821	6.2	-0.8

For the accuracy of the photogrammetric results, RMSE was calculated for study areas 1 and 2. In both study areas, the based on the UAV photogrammetric approach achieved better results in terms of RMSE compared to the two study areas (Table 6). The height from the DSM and DTM were calculated using ArcGIS and the mathematical calculation using excel. The mean for study area 1 was -0.2 and standard deviation (SD) was 0.4 likewise for study area 2 mean of -0.2 and SD of 0.5 were

calculated respectively for both the study area to compute the values for RMSE. The height of 28 trees was calculated manually and from the UAV photogrammetric works which resulted in an RMSE for the study area 1 and 2 was similar. Moreover, DSM and DTM revealed limited differences in prediction accuracy in terms of RMSE at the reclaimed mine area Mae Moh. Furthermore, the Min and Max values were significant. The visual analysis of the field reference against predicted values confirmed the accurate results. It was possible to observe that the effect of the values from the traditional survey and UAV photogrammetry.

Table 6 Evaluation of photogrammetric results with traditional survey on the study area 1 and 2

Study Area	Mean (m)	SD (m)	RMSE (m)	Min(m)	Max(m)
1.	-0.2	0.4	0.5	-1.2	0.8
2.	-0.2	0.5	0.5	-0.9	0.8

4.3 Validation results of land cover classification for study area 1 and 2

This section describes about the validation results from the study area 1 and 2 It ran through all training and test dataset in MATLAB. It was first tested by overlaying the predicted mask with the original image. A common model evaluation would start by looking at the accuracy, precision, Overall accuracy, Average accuracy, and per class accuracy to determine how good the model performs. However, image classification for monitoring the reclamation of the mines area using a deep learning model is a new approach in the mining industry

4.3.1 Study area 1

As displayed in Figures 17 (c) and (d) the model was effective when tested on the training data using orthophoto. This initial test result showed that the model was able to predict all pixels on the same image used in the test data. Figure 18 (c) and (d) reveals that the model responds qualitatively well for orthophoto using the model for the same time period. The test area included the height information from the DSM to predict the better result. The model was effective in predicting the vegetation accurately when orthophoto was combined with DSM. The testing procedure of the deep learning model is done using the test dataset. This test shows the capability of the model for the

land cover classification at a reclaimed mine. The other area was used for test data besides the training data. Each predicted patch was 256*256 in size showing the land cover classification. The resulting predicted model was effective for the test area. There are shadows in the study area, classifying the shadows at the reclaimed area shows the accurate classification of the vegetation. Next, the test was performed for the accuracy of the correctly classified pixel. This test considered how many pixels were correctly labeled. The model is evaluated by calculating per class accuracy. The parameters needed to calculate these metrics were acquired through the confusion matrix.

Prediction results on the training area study area 1 based on the training dataset

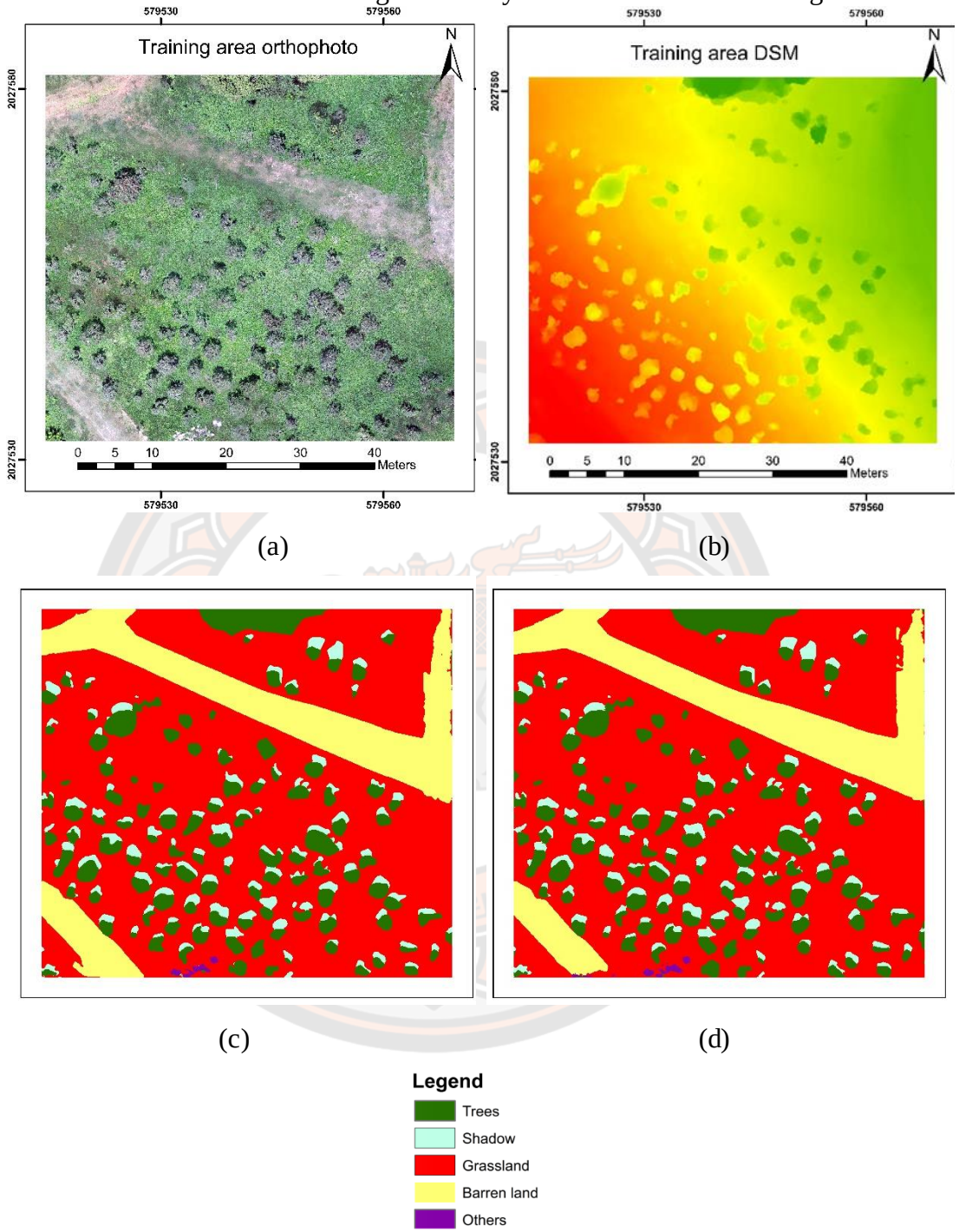


Figure 17 Training datasets and results after classification using DL (a) Orthophoto, (b) DSM, (c) land cover map from orthophoto, and (d) land cover map from orthophoto and DSM with the legend

Based on the training results for each class, Table 7 shows the per-class accuracy (PA) achieved by the proposed model. The results from the training data suggest that the deep learning with a combination of orthophoto and DSM was able to classify almost all classes excluding others with relatively high accuracy. The maximum per-class accuracy of the orthophoto and a combination of orthophoto with DSM for the grassland were 0.992 and 0.994, respectively. Focusing on the classification of vegetation area, the PA of combining orthophoto with DSM for both trees and grassland was relatively higher than the orthophoto.

Table 7 Training Dataset with per-class accuracy obtained by different class

Class	Orthophoto	Orthophoto and DSM
Trees	0.978	0.983
Shadow	0.967	0.969
Grassland	0.992	0.994
Barren land	0.982	0.970
others	0.772	0.824

For validation of prediction results using DL, the evaluation criteria for Overall accuracy (OA), Average accuracy (AA), and kappa(K) are presented in Table 8. The OA for the orthophoto and orthophoto with DSM were 0.987, AA was higher for orthophoto with DSM 0.948 and the K (0.991) was similar for both the models.

Table 8 Accuracy metrics of the classification methods for the training dataset

	Model	OA	AA	K
Training	Orthophoto	0.987	0.938	0.991
	Orthophoto and DSM	0.987	0.948	0.991

Prediction results on the test area study area 1 based on the training dataset

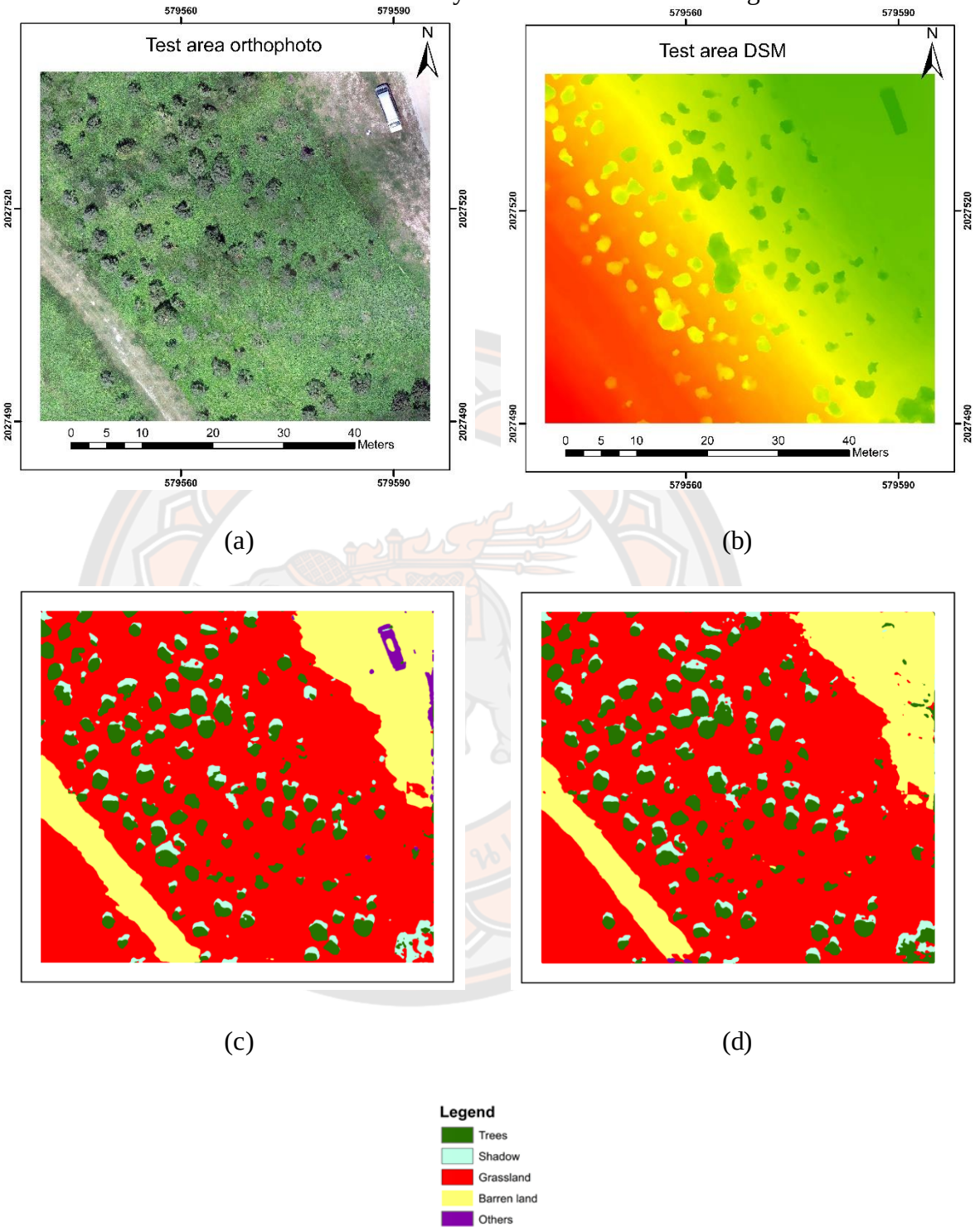


Figure 18 Test datasets and results after classification using DL (a) Orthophoto, (b) DSM, (c) land cover map from orthophoto, and (d) land cover map from orthophoto and DSM with the legend

Also, from Table 9, in terms of the classification results of the test dataset, the PA of combination orthophoto with DSM was more accurate than orthophoto excluding barren land and other class. The accuracy of trees was 3.9% greater when orthophoto was combined with DSM and the accuracy of grassland increased by 1.3% when orthophoto was combined with DSM. Likely, the classification results for the vegetation obtained from combination of orthophoto with DSM were significantly potential for the DL approach. However, the other class was misclassified when orthophoto was combined with DSM because of the inequalities in the dataset of other classes for training with DL, dataset of the labeled pixel for GT was the least of all. It is noticed that a contributing factor may also be the lack of adequate sampling (Marcos, Volpi, Kellenberger, & Tuia, 2018). On the other hand, most of the sampled pixels in the training dataset were for grassland the highest precision of classification was predicted by grassland for both approaches.

Table 9 Test Dataset with per-class accuracy obtained by different

Class	Orthophoto	Orthophoto and DSM
Trees	0.731	0.770
Shadow	0.770	0.788
Grassland	0.920	0.933
Barren land	0.963	0.913
others	0.800	0.000

For validation of prediction results using DL, the evaluation criteria for OA, AA, and K are presented in Table 10 When the orthophoto was combined with DSM, the overall accuracy (OA) for the test dataset was 0.900 and 0.904, respectively. The OA and K for the test dataset using the orthophoto and a combination of orthophoto were 0.837, 0.681, and 0.935, 0.937 respectively. Moreover, in the case of the testing for a combination of orthophoto with DSM, AA is lower accuracy as compared to the orthophoto due to misclassification error in other classes, as there was the least number of classified pixels for the others class which is shown in Table11.

Table 10 Accuracy metrics of the classification methods for the test dataset

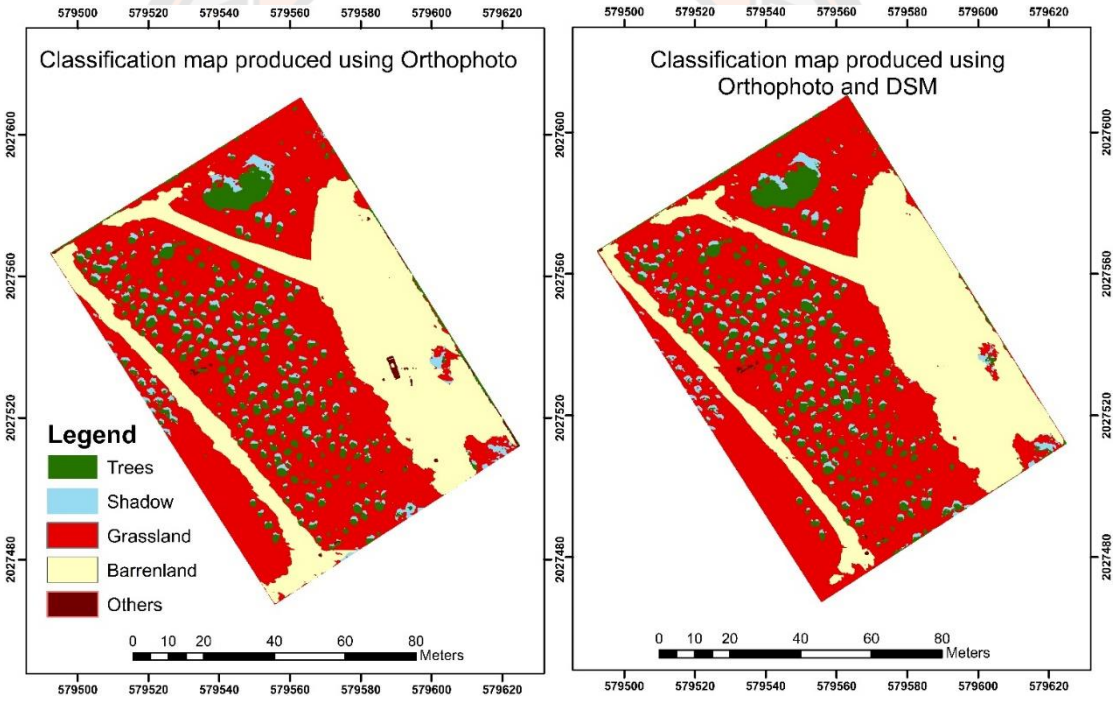
	Model	OA	AA	K
Testing	Orthophoto	0.900	0.837	0.935
	Orthophoto and DSM	0.904	0.681	0.937

Table 11 Number of pixels per land cover class in the ground truth

Land cover class	No. of pixel	Square meter
Trees	138,452	346.130
Shadow	519,76	129.940
Grassland	725,547	1813.868
Barren land	163,008	407.520
others	2,156	5.390

Five land cover groups considered in this research are (i) trees, (ii) shadow, (iii) grassland, (iv) barren land, and (v) others. The database provides ground truth labels that associate each pixel with one of five semantic classes. Total 1,011,839 pixels of the test area dataset were labeled in the image labeler, as shown in table 11. Each class was labeled according to its respective class using the orthophoto and DSM. The overall area was classified using the training area which shows that all the area of the overall image was classified especially for the vegetation area.

In conclusion, developing a machine learning model is a challenging process yet produces an effective way for monitoring reclaimed mines. This study had a challenging data acquisition process, but it allowed the researcher to explore another acquisition method. The vegetation classes were categorized well-using orthophoto and combining orthophoto with DSM. The study area has the test and train area which were taken from the same area and same time period. Images captured UAV platform showed the potential for land cover classification. Fig 19 shows the classification of the overall area for the test and train dataset. In addition, the UNet architecture provided favorable results in the vegetation detection algorithm. Some misclassifications were present for the other class due to the least number of labels for the other class.



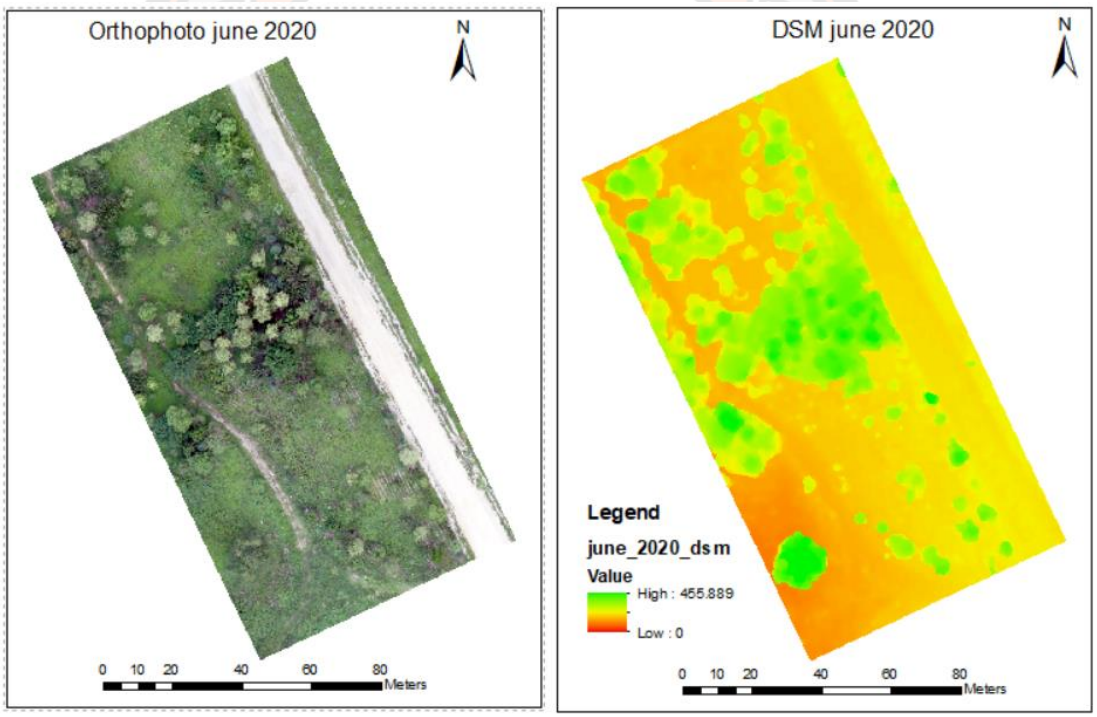
(a) (b)

Figure 19 Classification map using orthophoto and combination of orthophoto with DSM

4.3.2 Study area 2

The study area 2 consists of two datasets of the same area at a different time consisting of a) July 2020 and b) October 2020. The dataset for training was collected on 2nd October 2020 and the test data was collected on 22nd July 2020 by officials of Mae Moh mine using different UAV. Previously collected data was requested for

conducting the research which was used for predicting the result. After training of DL with CNN, land cover classification for the train and test dataset is predicted. Validation curves (accuracy and loss) of the two models evaluated the results using orthophoto and a combination of orthophoto with DSM. The accuracy and loss curves demonstrate the consistency of the training over the iterations for the DL with CNN. The accuracy curves in both the methods saturated after the 10th epoch. After the iterations were completed, the accuracy curves of both the model were identical. The findings revealed that the learning rate was positively linked to model efficiency. After 15 iterations, the loss curves reached a smooth stage, with the highest precision. The accuracy of the 0.1 curves was the highest when the iteration was stopped. Fig 20 shows the orthophoto and DSM used for the prediction of land cover using DL.

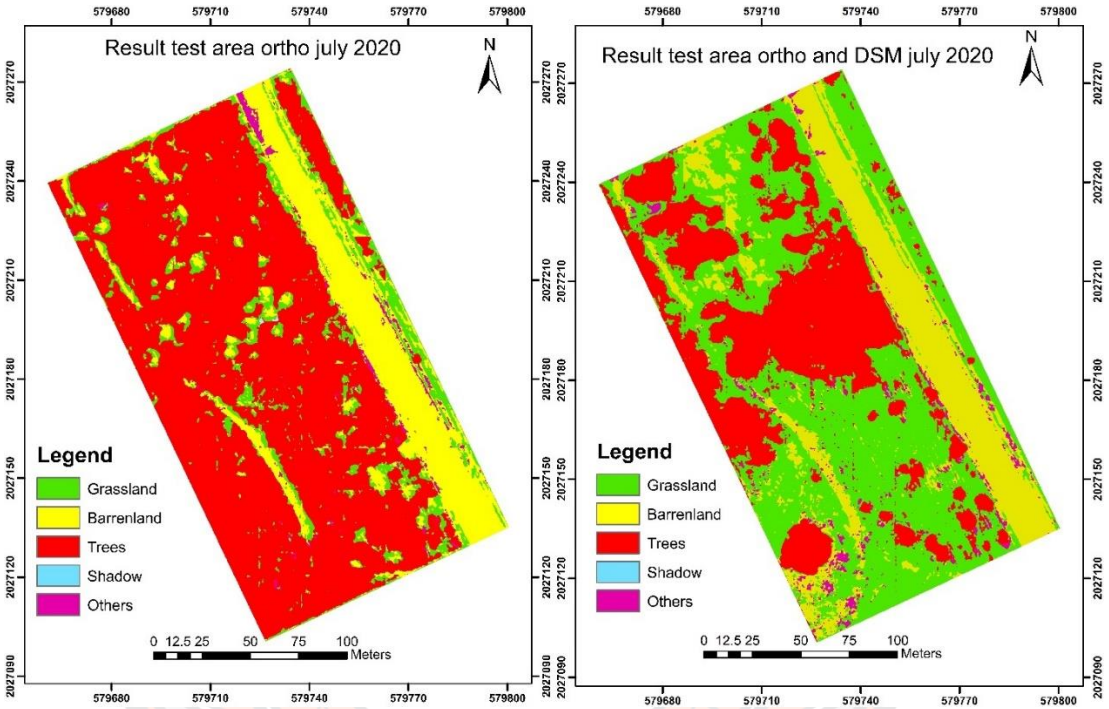


(a) (b)

Figure 20 Test area (a) orthophoto and (b) DSM (July 2020)

According to a visual interpretation between Figure 21 (a) and (b) the results were very different for the prediction. While using the only orthophoto the classes could not be identified and the results were not precise where when orthophoto was combined with the DSM the land cover classes were distinguished. Likely, there is insignificantly

identified for the type of land cover when orthophoto was combined with the DSM. In addition, the tree canopy was appropriately classified while using orthophoto with DSM. In conclusion, the DL for land cover classification demonstrates an efficient approach for monitoring the vegetation at the reclaimed mines when orthophoto was combined with DSM.



(a) (b)

Figure 21 Land cover classification results of test area using (a) orthophoto (b) orthophoto and DSM of July 2020

For validation of prediction results using DL, the evaluation criteria for OA, AA, and K were presented in Table 12. The overall accuracy (OA) for testing datasets using orthophoto and combining orthophoto and DSM were 0.395 and 0.751, respectively. Moreover, in the case of the testing orthophoto have lower accuracy as compared to the combination of orthophoto with DSM which shows that while combining orthophoto with DSM shows the better result of prediction as in Fig 22 orthophoto only cannot predict precise results using the data at different times. However, the OA and K of training using orthophoto and a combination of orthophoto with DSM are similar.

Table 12 validation results for the train and test area study area 2

Dataset	Model	OA	AA	K
Training	Orthophoto	0.990	0.964	0.987
	Orthophoto and DSM	0.994	0.979	0.992
Testing	Orthophoto	0.395	0.482	0.417
	Orthophoto and DSM	0.751	0.636	0.684

Based on the training and test results for each class, Table 13 shows the per-class accuracy (PA) achieved by the proposed model. The results from the test data suggest that deep learning with a combination of orthophoto and DSM was able to classify almost all classes with relatively high accuracy than orthophoto. The maximum per-class accuracy for test data using orthophoto and a combination of orthophoto with DSM for the grassland were 0.913 and 0.947, respectively. Focusing on the classification of vegetation area for the test area at different time, the PA of combining orthophoto with DSM for both trees and grassland were relatively higher than the orthophoto. Also, from Table 13, in terms of the classification results of the test dataset, the PA of combination orthophoto with DSM was more accurate than orthophoto excluding shadow and other classes.

The accuracy of trees was increased by 0.088 when orthophoto was combined with DSM and the accuracy of grassland increased by 0.550 when orthophoto was combined with DSM. Likely, the classification results for the vegetation obtained from orthophoto and DSM were significantly potential for the DL approach for change in time series data. However, the other class was misclassified when orthophoto was combined with DSM. Because of the inequalities in the dataset of other classes for training with DL, the dataset of the labeled pixel for GT was the least of all. It is noticed that a contributing factor may also be the lack of adequate sampling (Marcos et al., 2018). On the other hand, most of the sampled pixels in the training dataset were for grassland and trees with the highest precision of classification using orthophoto and DSM.

Table 13 Per class accuracy for the train and test area using orthophoto and DSM

Classes	Per Class accuracy			
	Training area		Test area	
	Orthophoto	Orthophoto and DSM	Orthophoto	Orthophoto and DSM
Grassland	0.993	0.994	0.117	0.665
Barren land	0.988	0.993	0.913	0.947
Trees	0.986	0.993	0.844	0.931
Shadow	0.892	0.940	0.521	0.000
Others	0.960	0.972	0.000	0.000

In conclusion, developing a machine learning model is a challenging process for the time series data yet produces an effective way for monitoring reclaimed mines. The vegetation classes were categorized well-using orthophoto and combining orthophoto with DSM. The study area has the test and train area which were taken at the same area and different periods. Images captured using the UAV platform showed the potential for land cover classification. Fig 21 shows the classification of the overall area for the test dataset from orthophoto and combination of orthophoto with DSM using the trained model at different time and same area. In addition, the UNet architecture provided favorable results in the vegetation detection algorithm when orthophoto was combined with DSM. The shadow and other classes were not classified while combining orthophoto and DSM due to the least number of labels for the classes.

4.4 Discussion

UAV data, as presented in this study, could potentially contribute to mine reclamation studies by providing land cover information meeting research needs. Land cover information provides a comprehensive insight into how the environment is useful for many conservation applications in reclaimed mines, such as investigation of spatial effects. Furthermore, it is dynamic to observe a large number of individuals within a forest community in order to adequately capture the expected variation in timing events (Donnelly, Grant, & O'Reilly, 2017). A UAV dataset, preferably acquired using the same camera at the different times, could provide vision into whether these issues have significant influences Furthermore, the ability to monitor mine reclamation with UAVs (although an accurate species identification still requires field observation) could help in understanding the effects of the underlying spatial complexity of the objects which

might be present within the resolved satellite imagery (White et al. 2009). The findings of this research support the use of UAV data for this application, as canopy of vegetation (based on the pixels).

The findings of the study show that the UAV photogrammetry approach for mine reclamation provides photogrammetric results, consisting of orthophoto and DSM, which is possibly more useful information for monitoring land cover. The deep learning technique can employ the automatic extraction of land clover using the trained datasets from orthophoto and/or DSM at the pixel level. The effectiveness of land cover was determined using the DL technique with training and test datasets from between orthophoto and a combination of orthophoto with DSM. The results comparison of the land clover classification based on DL revealed that the use of both orthophoto and DSM provides more accuracy than orthophoto, especially the vegetation area. As the DSM can offer useful information of height above ground surface especially for vegetation (such as trees), adding this information helps to improve accuracy (Al-Najjar et al., 2019). On the other hand, the use of orthophoto and DSM for DL revealed that the misclassification for barren land and other class due to adequate labeled data mentioned by Marcos et al. (2018). However, for mine reclamation, the vegetation area was more important due to the regulation of plantation from the concerned authorities and natural recovery.

In our land-cover classification task, it is clear that the reliability of deep-learning approaches (Al-Najjar et al., 2019; S. Li, Hao, Gao, & Kang, 2018). The spatial resolution of the orthophoto and DSM available in this analysis was 5 cm, which was adequate for vegetation classification. In the terms of the shadow, it is necessary to separate the shadow area from the tree crown measurement because calculating the vegetation area would be more precise when excluding shadow (Adeline, Chen, Briottet, Pang, & Paparoditis, 2013; Movia, Beinat, & Crosilla, 2016). The results for the image classification as in Figures 17 and 18 analyzed shows precise results based on Tables 8 and 10 for the land cover especially vegetation. Relevantly, the timeline for estimating land cover from digital images are based on models is also substantially improved relative to conventional approaches. Applying the DL methods to trained U-Net models, land cover maps are generated. Few computationally effective CNN for detection, semantic labeling has been added (Krizhevsky, Sutskever, & Hinton, 2017;

S. Li et al., 2018; Tang & Yuan, 2015). Land cover classification is essential for mine reclamation, as it can help reduce mine recovery to assess the mine field. Several studies have been developed to create effective architectures that can be performed in land cover based on DL (Alshehhi et al., 2017; Wu, Yu, Huang, & Yu, 2015) . In addition, regarding the state of art technologies, introducing DL and UAV into the periodic monitoring of mine reclamation could be considered as an effective result for reducing discrepancies among public administration and private companies, and at the same time, contributing to monitoring the sustainable development of extractive mining activities.

DL with a UAV photogrammetric approach platform could be more suitable for much more extensive restorations and perhaps further improvements could be achieved using multispectral sensors with more bands or hyperspectral sensors to enhance the spectral information. This study was limited by UAV imagery with a normal (RGB) camera at the mines area with the pattern of vegetation. These results indicated that the combination of orthophoto with DSM provides a comprehensive database of UAV imagery at an affordable cost, which can be exploited for land cover mapping and detecting the changes of vegetation at reclaimed mines area. The accuracy of image classification is as improved by Y. Li, Zhang, and Shen (2017), the accuracy of OA, AA, and K was higher for the training data using the combination of orthophoto and DSM with the spatial resolution of 5 cm/pixel.

The test and the train study area 1 were taken at the same time and using the same camera whereas for the study area 2 the images were of different time between 2 October 2020 (train dataset) and 21 July 2020 (test dataset). The results for the study area 1 show precise results for both the test and train data as shown in figure 17 (c, d) and 18 (c, d). the objects in the train and test areas are classified easily using orthophoto and combination of orthophoto with DSM. Also, the misclassification for barren land and other class due to adequate labeled data mentioned by (Marcos et al., 2018). Similarly, for the study area 2 as in Fig 21 (a), (b) using the test imagery at different time the result is more precise when orthophoto is combined with DSM (Fig 21 b).

In our experiment, it was observed that the vegetation area specially trees were classified well for both the study pattern and complex vegetation while combining

orthophoto with DSM from table 9 and 13 Deep learning with CNN provided effective techniques to deal with these problems and to improve accuracy. Moreover, a useful method is to use the pixels and guide the maps used by (Wang et al., 2017) for classification. Xu, Xie, Feng, and Chen (2018) designed attention unit to fuse different scale feature maps and designed a global attention unit to eliminate the effects on object covering, which effectively improved the accuracy. The cases of vegetation (trees) being often misclassified as vegetation may diminish but if we add DSM to the model it enhances the vegetation features. Trees, are one of the most prominent classes in the reference map for the testing area at different time. Also, (Al-Najjar et al., 2019) proposed the method for land cover classification using fused DSM which improved the accuracy. These methods contribute to improve the performance of the model.

This experimental study demonstrated two successful applications using the different time of dataset for the classification method. This technique attempted to simulate the pattern of vegetation and classification for interpretation. The proposed convolution neural network and its associated training techniques can be extended to monitor reclaimed mines areas. In fact, the deep learning with CNN technique should be able to be trained and to detect almost all vegetation patterns perceivable for monitoring the reclaimed mines. CNN extracts deep spatial-related features for fixed input patch size. The DL with CNN offered land cover classification results in their mapping accuracies using between Orthophoto and combination of orthophoto with DSM. Therefore, the classification using CNN solve the problem of monitoring scale across the mines area.

CHAPTER 5
CONCLUSION AND RECOMMENDATION

5.1 Conclusion

A comprehensive review of UAVs photogrammetry concerning traditional photogrammetry is presented at the reclaimed mines as is efficient and low cost. This described the applications of UAVs in photogrammetry for monitoring the reclaimed mines. Based on the literature review and results of reclaimed mines area, classification methods using UAV were presented. Using the review and the formulated classification UAV system is low-cost, open-source, manually, and assisted controlled and autonomously flying UAVs. The coordinates in the flight planning are frequently defined relative to the start point, which allows flight planning independently from existing maps and coordinate systems. Finally, the capability of using UAVs in inaccessible and dangerous areas, and the improvements in the data processing open up new applications in UAV photogrammetry at the reclaimed mines area. Images captured through the UAV photogrammetry showed the potential to serve as suitable input data to monitor the vegetation at the reclaimed mine.

The deep learning with CNN presented in this thesis is feasible to automate the detection and monitoring of reclaimed mines which does not require a long waiting period between data acquisition sessions at different intervals of time. However, this thesis took that concept and utilize automation in the monitoring reclaimed mine from UAV imagery. Developing a deep learning model is a challenging process yet will produce an effective way for monitoring the reclaimed mine. Deep learning has had a revolutionary effect on computer vision. The deep learning with CNN provided favorable results in land cover classification at reclaimed mines. Deep learning with UAV has become a widespread resource for aerial imagery. The proposed approach in this thesis has shown that it can be applied to photogrammetric results for automatic land cover classification from UAV images at reclaimed mine areas. A large survey area can be mapped through the use of a UAV system without interrupting ongoing production operations, and regions inaccessible by human operators can also be covered. This results in the acquisition of on-demand and almost real-time data, which delivers high-resolution data. The method shows that networks trained on a task for

image recognition can potentially be used to solve the problem of classification. The approach also deduced that these methods for land cover classification are consistent on a same area and different time as the dataset used for training the model. This proposed approach is a breakthrough as significant unlabeled photogrammetric datasets can now be classified and categorized. Adapting a deep pre-trained network on a new dataset that has a limited number of labeled images to train quickly, learn and adjust the weights and biases of the network on the new dataset in effect delivers promising results.

The trained model in DL for the study areas 1 and 2 both orthophoto and DSM have higher precision of OA, AA, k, and PA than only orthophoto due to the height information. The values of land cover classification for train area were OA of 0.987, AA of 0.938 and k of 0.991 also test area from combining orthophoto with DSM with OA of 0.904, AA of 0.681, and K of 0.937 for study area 1 in which data was used for same time and same UAV. Similarly, for study area 2 the values for train area were OA of 0.994, AA of 0.979 and k of 0.992 and for test area with OA of 0.751, AA of 0.636, and K of 0.684 in which data for train and test were of different time and different UAV.

The deep learning model presented in this thesis showed that it is feasible to automate the detection and monitoring of reclaimed mines which does not require a long waiting period between data acquisition sessions at different intervals of time. However, this thesis took that concept and utilize automation in the monitoring reclaimed mine from UAV imagery. The automation process suggested in this study would reduce the time spent significantly. The trained model in this study showed promising performance. It was evaluated qualitatively and quantitatively to determine the feasibility of the proposed method. The final model also attests that an image classification problem can be solved effectively through the UNet architecture. Also, this thesis serves as a methodology and provides a concept to apply deep learning for monitoring and classifying land cover at reclaimed mine.

It is concluded that when implementing succession for restoration of vegetation the following points are considered evaluation of environmental site conditions; deciding whether spontaneous succession of vegetation at reclaimed mines area; prediction of successional development; monitoring of the results. The need for

interdisciplinary approaches and communication between engineers and decision-makers is emphasized. Land cover information at reclaimed mines plays a significant role in the earth system, which reflects the influence of human activities and environmental changes. Administrative and extractive companies are bounded by laws for carrying out the vegetation activities at the reclaimed mines area. Also, this approach helps engineers and decision makers to conclude the results at the reclaimed mines area.

Also, this research developed a UAV photogrammetric approach using the deep learning method for land cover classification. The land cover classes were classified and mapped with high accuracy and computational efficiency for the train and test areas. Below are the main conclusions emerging from this research.

- Effectiveness of the UAV photogrammetric approach was compared to the traditional monitoring methods for monitoring rehabilitation in the mines in terms of time consuming, less labour and cost.
- Land cover classification with deep learning using convolutional neural networks (CNN), where a group of pixels was used for labeling, and further incorporated the spatial context with high accuracy. Substantial progress in developing a method for extracting information from reclaimed mines has been made.

5.2 Limitations

The UAV imagery was limited by normal RGB camera from the UAV and the time series datasets for the study area as the research could be conducted at the same time later using the same UAV approach. Additionally, to some extent, shadows from the UAV platform may be responsible for the final results from the deep learning. In this study, different tree species were considered and analyzed. UAV-acquired data help provide low-cost, high-resolution, and continuously updated mine land cover information for sustainable development.

5.3 Future Works

The following suggestions for future work based on our experience of DL for mine reclamation: Future studies may explore the feasibility of using UAV systems to collect

multispectral data for monitoring. The collected information will be beneficial for various monitoring tasks because the UAV data captured in different wavelengths can reveal different features. This direction will result in a new set of data analytic methods, which will complement the proposed approaches in this work. Further multispectral UAV sensors need for more accurate land cover mapping at the rehabilitated mines area. Further works can be listed as:

- Future studies should conduct for identifying the types of trees at large areas to further improve and refine the proposed workflow for monitoring.
- Also, the data can be collected at the same time (October 2021) using UAV with the flight conditions same as in October 2020 and the results can be predicted using the model of October 2020.
- While DL is very promising and currently applied in a wide range of applications. It might be interesting to explore other classification methods like Random Forest models on this same dataset to make a comparison of the prediction results with the DL model

REFERENCES

Aasen, H., Burkart, A., Bolten, A., & Bareth, G. (2015). Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance. *ISPRS Journal of photogrammetry and remote sensing, 108*, 245-259.

Abd, H. A. A.-R., & Alnajjar, H. A. (2013). Maximum likelihood for land-use/land-cover mapping and change detection using landsat satellite images: a case study "South Of Johor". *International Journal of Computational Engineering Research, 3*(6), 26-33.

Adeline, K. R. M., Chen, M., Briottet, X., Pang, S. K., & Paparoditis, N. (2013). Shadow detection in very high spatial resolution aerial images: A comparative study. *ISPRS Journal of photogrammetry and remote sensing, 80*, 21-38. doi:https://doi.org/10.1016/j.isprsjprs.2013.02.003

Akar, Ö. (2017). Mapping land use with using Rotation Forest algorithm from UAV images. *European Journal of Remote Sensing, 50*(1), 269-279.

Al-Najjar, H. A., Kalantar, B., Pradhan, B., Saeidi, V., Halin, A. A., Ueda, N., & Mansor, S. (2019). Land cover classification from fused DSM and UAV images using convolutional neural networks. *Remote sensing, 11*(12), 1461.

Alshehhi, R., Marpu, P. R., Woon, W. L., & Dalla Mura, M. (2017). Simultaneous extraction of roads and buildings in remote sensing imagery with convolutional neural networks. *ISPRS Journal of photogrammetry and remote sensing, 130*, 139-149.

Anderson, K., & Gaston, K. J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment, 11*(3), 138-146.

Anderson, K., Ryan, B., Sonntag, W., Kavvada, A., & Friedl, L. (2017). Earth observation in service of the 2030 Agenda for Sustainable Development. *Geo-spatial Information Science, 20*(2), 77-96.

Arel, I., Rose, D. C., & Karnowski, T. P. (2010). Deep machine learning-a new frontier in artificial intelligence research [research frontier]. *IEEE computational intelligence magazine, 5*(4), 13-18.

Barmpounakis, E. N., Vlahogianni, E. I., & Golias, J. C. (2016). *Extracting kinematic characteristics from unmanned aerial vehicles*. Retrieved from

Ben-David, A. (2008). Comparison of classification accuracy using Cohen's Weighted Kappa. *Expert Systems with Applications, 34*(2), 825-832.

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.

Bergado, J. R., Persello, C., & Gevaert, C. (2016). *A deep learning approach to the classification of sub-decimetre resolution aerial images.* Paper presented at the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of photogrammetry and remote sensing, 65*(1), 2-16.

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., . . . Van Coillie, F. (2014). Geographic object-based image analysis–towards a new paradigm. *ISPRS Journal of photogrammetry and remote sensing, 87*, 180-191.

Boutaba, R., Salahuddin, M. A., Limam, N., Ayoubi, S., Shahriar, N., Estrada-Solano, F., & Caicedo, O. M. (2018). A comprehensive survey on machine learning for networking: evolution, applications and research opportunities. *Journal of Internet Services and Applications, 9*(1), 1-99.

Bradshaw, A., & Hüttl, R. (2001). Future Minesite Restoration Involves a Broader Approach. *Ecological Engineering - ECOL ENG, 17*, 87-90. doi:10.1016/S0925-8574(00)00149-X

Bratasanu, D., Nedelcu, I., & Datcu, M. (2010). Bridging the semantic gap for satellite image annotation and automatic mapping applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 4*(1), 193-204.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*(1), 5-32. doi:10.1023/A:1010933404324

Bryson, M., Reid, A., Ramos, F., & Sukkarieh, S. (2010). Airborne vision-based mapping and classification of large farmland environments. *Journal of Field Robotics, 27*(5), 632-655.

Chen, C., Zhang, B., Su, H., Li, W., & Wang, L. (2016). Land-use scene classification using multi-scale completed local binary patterns. *Signal, image and video processing, 10*(4), 745-752.

Chen, G., Weng, Q., Hay, G. J., & He, Y. (2018). Geographic object-based image analysis (GEOBIA): emerging trends and future opportunities. *GIScience & Remote Sensing, 55*(2), 159-182.

Chen, J., Li, K., Chang, K.-J., Sofia, G., & Tarolli, P. (2015). Open-pit mining geomorphic feature characterisation. *International Journal of Applied Earth Observation and Geoinformation, 42*, 76-86.

Cheng, G., Han, J., Guo, L., Qian, X., Zhou, P., Yao, X., & Hu, X. (2013). Object detection in remote sensing imagery using a discriminatively trained mixture model. *ISPRS Journal of photogrammetry and remote sensing, 85*, 32-43.

Coifman, B., McCord, M., Mishalani, R. G., Iswalt, M., & Ji, Y. (2006). *Roadway traffic monitoring from an unmanned aerial vehicle*. Paper presented at the IEE Proceedings-Intelligent Transport Systems.

Colomina, I., & Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of photogrammetry and remote sensing, 92*, 79-97.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*(1), 35-46.

Congalton, R. G., Oderwald, R. G., & Mead, R. A. (1983). Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogrammetric engineering and remote sensing, 49*(12), 1671-1678.

Dash, J. P., Watt, M. S., Pearse, G. D., Heaphy, M., & Dungey, H. S. (2017). Assessing very high resolution UAV imagery for monitoring forest health during a simulated disease outbreak. *ISPRS Journal of photogrammetry and remote sensing, 131*, 1-14.

Donnelly, L., Grant, O., & O'Reilly, C. (2017). Effect of deployment-type on stem growth, biomass partitioning and crown characteristics of juvenile Sitka spruce clones. *Silva Fennica, 51*(1).

Farabet, C., Couprie, C., Najman, L., & LeCun, Y. (2012). Learning hierarchical features for scene labeling. *IEEE transactions on pattern analysis and machine intelligence, 35*(8), 1915-1929.

Feng, Q., Liu, J., & Gong, J. (2015). UAV remote sensing for urban vegetation mapping using random forest and texture analysis. *Remote sensing, 7*(1), 1074-1094.

Flener, C., Vaaja, M., Jaakkola, A., Krooks, A., Kaartinen, H., Kukko, A., . . . Alho, P. (2013). Seamless mapping of river channels at high resolution using mobile LiDAR and UAV-photography. *Remote sensing, 5*(12), 6382-6407.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment, 80*(1), 185-201.

Foody, G. M. (2020). Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment, 239*, 111630.

Foody, G. M., & Mathur, A. (2004). A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on geoscience and remote sensing, 42*(6), 1335-1343.

Fu, G., Liu, C., Zhou, R., Sun, T., & Zhang, Q. (2017). Classification for high resolution remote sensing imagery using a fully convolutional network. *Remote sensing, 9*(5), 498.

Fung, T., & LeDrew, E. (1988). For change detection using various accuracy. *Photogrammetric engineering and remote sensing, 54*(10), 1449-1454.

Gago, J., Douthe, C., Coopman, R. E., Gallego, P. P., Ribas-Carbo, M., Flexas, J., . . . Medrano, H. (2015). UAVs challenge to assess water stress for sustainable agriculture. *Agricultural water management, 153*, 9-19.

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*: O'Reilly Media.

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1): MIT press Cambridge.

Greenfeld, J. (2001). Evaluating the accuracy of digital orthophoto quadrangles (DOQ) in the context of parcel-based GIS. *Photogrammetric engineering and remote sensing, 67*(2), 199-206.

Haala, N., Cramer, M., Weimer, F., & Trittler, M. (2011). Performance test on UAV-based photogrammetric data collection. *International archives of the photogrammetry, remote sensing and spatial information sciences, 38*(6).

Hartemink, A. E., & Minasny, B. (2014). Towards digital soil morphometrics. *Geoderma, 230*, 305-317.

Hastaoğlu, K. Ö., Gül, Y., Poyraz, F., & Kara, B. C. (2019). Monitoring 3D areal displacements by a new methodology and software using UAV photogrammetry. *International Journal of Applied Earth Observation and Geoinformation, 83*, 101916.

Hester, D. B., Cakir, H. I., Nelson, S. A., & Khorram, S. (2008). Per-pixel classification of high spatial resolution satellite imagery for urban land-cover mapping. *Photogrammetric Engineering & Remote Sensing, 74*(4), 463-471.

Heydari, S. S., & Mountrakis, G. (2018). Effect of classifier selection, reference sample size, reference class distribution and scene heterogeneity in per-pixel classification accuracy using 26 Landsat sites. *Remote Sensing of Environment, 204*, 648-658.

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., . . . Sainath, T. N. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine, 29*(6), 82-97.

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

Hu, F., Xia, G.-S., Hu, J., & Zhang, L. (2015). Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote sensing, 7*(11), 14680-14707.

Hu, S., & Wang, L. (2013). Automated urban land-use classification with remote sensing. *International Journal of Remote Sensing, 34*(3), 790-803.

Huval, B., Wang, T., Tandon, S., Kiske, J., Song, W., Pazhayampallil, J., . . . Cheng-Yue, R. (2015). An empirical evaluation of deep learning on highway driving. *arXiv preprint arXiv:1504.01716*.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112): Springer.

Kanistras, K., Martins, G., Rutherford, M. J., & Valavanis, K. P. (2013). *A survey of unmanned aerial vehicles (UAVs) for traffic monitoring*. Paper presented at the 2013 International Conference on Unmanned Aircraft Systems (ICUAS).

Khaliq, A., Musci, M. A., & Chiaberge, M. (2018). *Understanding effects of atmospheric variables on spectral vegetation indices derived from satellite based time series of multispectral images*. Paper presented at the 2018 IEEE Applied Imagery Pattern Recognition Workshop (AIPR).

Kobayashi, H., Watando, H., & Kakimoto, M. (2014). A global extent site-level analysis of land cover and protected area overlap with mining activities as an indicator of biodiversity pressure. *Journal of cleaner production, 84*, 459-468.

Korjus, H., Laarmann, D., Sims, A., Paluots, T., & Kangur, A. (2014). Assessment of novel forest ecosystems on post-mining restoration site in Aidu, Estonia. *Local and regional challenges of climate change adaptation and green technologies. Proceedings. University of West Hungary Press, Sopron*, 35-44.

Kraus, K. (2011). *Photogrammetry: geometry from images and laser scans*: Walter de Gruyter.

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.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84-90.

Kussul, N., Mykola, L., Shelestov, A., & Skakun, S. (2018). Crop inventory at regional scale in Ukraine: developing in season and end of season crop maps with multi-temporal optical and SAR satellite imagery. *European Journal of Remote Sensing, 51*(1), 627-636.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Längkvist, M., Kiselev, A., Alirezaie, M., & Loutfi, A. (2016). Classification and segmentation of satellite orthoimagery using convolutional neural networks. *Remote sensing, 8*(4), 329.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nat. 521, 436–444.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature 521.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278-2324.

Li, M., Stein, A., & Bijker, W. (2016). *Urban land use extraction from very high resolution remote sensing images by bayesian network*. Paper presented at the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).

Li, S., Hao, Q., Gao, G., & Kang, X. (2018). The Effect of Ground Truth on Performance Evaluation of Hyperspectral Image Classification. *IEEE Transactions on geoscience and remote sensing, 56*(12), 7195-7206. doi:10.1109/TGRS.2018.2849225

Li, Y., Zhang, H., & Shen, Q. (2017). Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote sensing, 9*(1), 67.

Li, Z. (1993). Theoretical models of the accuracy of digital terrain models: An evaluation and some observations. *The Photogrammetric Record, 14*(82), 651-660.

Liu, Q., Hang, R., Song, H., & Li, Z. (2017). Learning multiscale deep features for high-resolution satellite image scene classification. *IEEE Transactions on geoscience and remote sensing, 56*(1), 117-126.

Liu, X., He, J., Yao, Y., Zhang, J., Liang, H., Wang, H., & Hong, Y. (2017). Classifying urban land use by integrating remote sensing and social media data. *International Journal of Geographical Information Science, 31*(8), 1675-1696.

Long, Y., Gong, Y., Xiao, Z., & Liu, Q. (2017). Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Transactions on geoscience and remote sensing, 55*(5), 2486-2498.

Lucieer, A., Jong, S. M. d., & Turner, D. (2014). Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography. *Progress in Physical Geography, 38*(1), 97-116.

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017). A review of supervised object-based land-cover image classification. *ISPRS Journal of photogrammetry and remote sensing, 130*, 277-293.

Ma, X., Wang, H., & Wang, J. (2016). Semisupervised classification for hyperspectral image based on multi-decision labeling and deep feature learning. *ISPRS Journal of photogrammetry and remote sensing, 120*, 99-107.

Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). High-resolution aerial image labeling with convolutional neural networks. *IEEE Transactions on geoscience and remote sensing, 55*(12), 7092-7103.

Marcos, D., Volpi, M., Kellenberger, B., & Tuia, D. (2018). Land cover mapping at very high resolution with rotation equivariant CNNs: Towards small yet accurate models. *ISPRS Journal of photogrammetry and remote sensing, 145*, 96-107.

Marmanis, D., Schindler, K., Wegner, J. D., Galliani, S., Datcu, M., & Stilla, U. (2018). Classification with an edge: Improving semantic image segmentation with boundary detection. *ISPRS Journal of photogrammetry and remote sensing, 135*, 158-172.

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing, 39*(9), 2784-2817.

McRoberts, R. E. (2014). Post-classification approaches to estimating change in forest area using remotely sensed auxiliary data. *Remote Sensing of Environment, 151*, 149-156.

Ming, D., Li, J., Wang, J., & Zhang, M. (2015). Scale parameter selection by spatial statistics for GeOBIA: Using mean-shift based multi-scale segmentation as an example. *ISPRS Journal of photogrammetry and remote sensing, 106*, 28-41.

Ministry of Industry Thailand (2009). *Mining in Thailand*. Thailand: Thai government.

Mnih, V. (2013). *Machine learning for aerial image labeling*: University of Toronto (Canada).

Mossa, J., & James, L. (2013). 13.6 Impacts of Mining on Geomorphic Systems.

Movia, A., Beinat, A., & Crosilla, F. (2016). Shadow detection and removal in RGB VHR images for land use unsupervised classification. *ISPRS Journal of photogrammetry and remote sensing, 119*, 485-495. doi:https://doi.org/10.1016/j.isprsjprs.2016.05.004

Myint, S. W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment, 115*(5), 1145-1161.

Nogueira, K., Penatti, O. A., & Dos Santos, J. A. (2017). Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognition, 61*, 539-556.

Ockendon, N., Thomas, D. H., Cortina, J., Adams, W. M., Aykroyd, T., Barov, B., . . . Brombacher, M. (2018). One hundred priority questions for landscape restoration in Europe. *Biological Conservation, 221*, 198-208.

Oliva-Santos, R., Maciá-Pérez, F., & Garea-Llano, E. (2014). Ontology-based topological representation of remote-sensing images. *International Journal of Remote Sensing, 35*(1), 16-28.

Osterkamp, W., & Joseph, W. (2000). Climatic and hydrologic factors associated with reclamation. *Reclamation of drastically disturbed lands, 41*, 193-215.

Othman, E., Bazi, Y., Alajlan, N., Alhichri, H., & Melgani, F. (2016). Using convolutional features and a sparse autoencoder for land-use scene classification. *International Journal of Remote Sensing, 37*(10), 2149-2167.

Otukei, J. R., & Blaschke, T. (2010). Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation, 12*, S27-S31.

Park, S., & Choi, Y. (2020). Applications of unmanned aerial vehicles in mining from exploration to reclamation: A review. *Minerals, 10*(8), 663.

Patino, J. E., & Duque, J. C. (2013). A review of regional science applications of satellite remote sensing in urban settings. *Computers, Environment and Urban Systems, 37*, 1-17.

Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., . . . Lu, L. (2013). A global human settlement layer from optical HR/VHR RS data: Concept and first results. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6*(5), 2102-2131.

Pingel, T. J., Clarke, K. C., & McBride, W. A. (2013). An improved simple morphological filter for the terrain classification of airborne LIDAR data. *ISPRS Journal of photogrammetry and remote sensing, 77*, 21-30.

Pontius Jr, R. G., & Millones, M. (2011). Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing, 32*(15), 4407-4429.

Puri, A. (2005). A survey of unmanned aerial vehicles (UAV) for traffic surveillance. *Department of computer science and engineering, University of South Florida*, 1-29.

Qin, F., Guo, J., & Sun, W. (2017). Object-oriented ensemble classification for polarimetric SAR imagery using restricted Boltzmann machines. *Remote sensing letters, 8*(3), 204-213.

Ramoelo, A., Cho, M. A., Mathieu, R., Madonsela, S., Van De Kerchove, R., Kaszta, Z., & Wolff, E. (2015). Monitoring grass nutrients and biomass as indicators of rangeland quality and quantity using random forest modelling and WorldView-2 data. *International Journal of Applied Earth Observation and Geoinformation, 43*, 43-54.

Rathore, I., & Kumar, N. P. (2015). Unlocking the potentiality of uavs in mining industry and its implications. *International Journal of Innovative Research in Science, Engineering and Technology, 4*.

Reis, S., & Taşdemir, K. (2011). Identification of hazelnut fields using spectral and Gabor textural features. *ISPRS Journal of photogrammetry and remote sensing, 66*(5), 652-661.

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, H., Zhao, Y., Xiao, W., & Hu, Z. (2019). A review of UAV monitoring in mining areas: Current status and future perspectives. *International Journal of Coal Science & Technology*, 1-14.

Rosenfield, G. H., & Fitzpatrick-Lins, K. (1986). A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric engineering and remote sensing, 52*(2), 223-227.

Salehi, B., Ming Zhong, Y., & Dey, V. (2012). A review of the effectiveness of spatial information used in urban land cover classification of VHR imagery. *International Journal of Geoinformatics, 8*(2), 35.

Salvini, R., Mastrorocco, G., Seddaiu, M., Rossi, D., & Vanneschi, C. (2017). The use of an unmanned aerial vehicle for fracture mapping within a marble quarry (Carrara, Italy): photogrammetry and discrete fracture network modelling. *Geomatics, Natural Hazards and Risk, 8*(1), 34-52.

Sameen, M. I., Pradhan, B., & Aziz, O. S. (2018). Classification of very high resolution aerial photos using spectral-spatial convolutional neural networks. *Journal of Sensors, 2018*.

Sammut, C., & Webb, G. I. (2011). *Encyclopedia of machine learning*: Springer Science & Business Media.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks, 61*, 85-117.

Sharma, A., Liu, X., Yang, X., & Shi, D. (2017). A patch-based convolutional neural network for remote sensing image classification. *Neural Networks, 95*, 19-28.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., . . . Lanctot, M. (2016). Mastering the game of Go with deep neural networks and tree search. *nature, 529*(7587), 484.

Suk, H.-I., Lee, S.-W., Shen, D., & Initiative, A. s. D. N. (2014). Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage, 101*, 569-582.

Sun, Y., Zhao, L., Huang, S., Yan, L., & Dissanayake, G. (2014). L2-SIFT: SIFT feature extraction and matching for large images in large-scale aerial photogrammetry. *ISPRS Journal of photogrammetry and remote sensing, 91*, 1-16.

Suykens, J. A. K., & Vandewalle, J. (1999). Least Squares Support Vector Machine Classifiers. *Neural Processing Letters, 9*(3), 293-300. doi:10.1023/A:1018628609742

Tang, S., & Yuan, Y. (2015). *Object detection based on convolutional neural network*. Paper presented at the International Conference-IEEE–2016.

Tarolli, P. (2014). High-resolution topography for understanding Earth surface processes: Opportunities and challenges. *Geomorphology, 216*, 295-312.

Tarolli, P., & Sofia, G. (2016). Human topographic signatures and derived geomorphic processes across landscapes. *Geomorphology, 255*, 140-161.

Tassopoulou, M., Verde, N., Mallinis, G., Georgiadis, H., Kaimaris, D., & Patias, P. (2019). *Demonstrating the potential of remote sensing to support sustainable development goals implementation: case studies over Greece*. Paper presented at the Seventh International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2019).

Tian, J., Wang, L., Li, X., Gong, H., Shi, C., Zhong, R., & Liu, X. (2017). Comparison of UAV and WorldView-2 imagery for mapping leaf area index of mangrove forest. *International Journal of Applied Earth Observation and Geoinformation, 61*, 22-31.

Turner, D., Lucieer, A., & De Jong, S. M. (2015). Time series analysis of landslide dynamics using an unmanned aerial vehicle (UAV). *Remote sensing, 7*(2), 1736-1757.

Verburg, P. H., Neumann, K., & Nol, L. (2011). Challenges in using land use and land cover data for global change studies. *Global change biology, 17*(2), 974-989.

Vetrivel, A., Gerke, M., Kerle, N., Nex, F., & Vosselman, G. (2018). Disaster damage detection through synergistic use of deep learning and 3D point cloud features derived from very high resolution oblique aerial images, and multiple-kernel-learning. *ISPRS Journal of photogrammetry and remote sensing, 140*, 45-59.

Vidal, O., Goffé, B., & Arndt, N. (2013). Metals for a low-carbon society. *Nature Geoscience, 6*(11), 894.

Volpi, M., & Tuia, D. (2016). Dense semantic labeling of subdecimeter resolution images with convolutional neural networks. *IEEE Transactions on geoscience and remote sensing, 55*(2), 881-893.

Walde, I., Hese, S., Berger, C., & Schmullius, C. (2014). From land cover-graphs to urban structure types. *International Journal of Geographical Information Science, 28*(3), 584-609.

Wang, H., Wang, Y., Zhang, Q., Xiang, S., & Pan, C. (2017). Gated convolutional neural network for semantic segmentation in high-resolution images. *Remote sensing, 9*(5), 446.

Webster, C., Westoby, M., Rutter, N., & Jonas, T. (2018). Three-dimensional thermal characterization of forest canopies using UAV photogrammetry. *Remote Sensing of Environment, 209*, 835-847.

Wu, J., Yu, Y., Huang, C., & Yu, K. (2015). *Deep multiple instance learning for image classification and auto-annotation.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Xia, G.-S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., . . . Lu, X. (2017). AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Transactions on geoscience and remote sensing, 55*(7), 3965-3981.

Xiang, J., Chen, J., Sofia, G., Tian, Y., & Tarolli, P. (2018). Open-pit mine geomorphic changes analysis using multi-temporal UAV survey. *Environmental earth sciences, 77*(6), 220.

Xu, Y., Xie, Z., Feng, Y., & Chen, Z. (2018). Road extraction from high-resolution remote sensing imagery using deep learning. *Remote sensing, 10*(9), 1461.

Yang, W., Yin, X., & Xia, G.-S. (2015). Learning high-level features for satellite image classification with limited labeled samples. *IEEE Transactions on geoscience and remote sensing, 53*(8), 4472-4482.

Yao, X., Han, J., Cheng, G., Qian, X., & Guo, L. (2016). Semantic annotation of high-resolution satellite images via weakly supervised learning. *IEEE Transactions on geoscience and remote sensing, 54*(6), 3660-3671.

Yucel, M. A., & Turan, R. Y. (2016). Areal change detection and 3D modeling of mine lakes using high-resolution unmanned aerial vehicle images. *Arabian Journal for Science and Engineering, 41*(12), 4867-4878.

Zhang, X., Chen, G., Wang, W., Wang, Q., & Dai, F. (2017). Object-based land-cover supervised classification for very-high-resolution UAV images using stacked denoising autoencoders. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10*(7), 3373-3385.

Zhao, B., Zhong, Y., & Zhang, L. (2016). A spectral–structural bag-of-features scene classifier for very high spatial resolution remote sensing imagery. *ISPRS Journal of photogrammetry and remote sensing, 116*, 73-85.

Zhong, Y., Zhu, Q., & Zhang, L. (2015). Scene classification based on the multifeature fusion probabilistic topic model for high spatial resolution remote sensing imagery. *IEEE Transactions on geoscience and remote sensing, 53*(11), 6207-6222.

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine, 5*(4), 8-36.

Zhu, Y., & Newsam, S. (2015). *Land use classification using convolutional neural networks applied to ground-level images*. Paper presented at the Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems.

APPENDIX

Appendix A

Co-ordinates of the ground control points

Point	Easting	Northing	Orthometric Height
		study area 1	
A1	579601.203	2027507.857	453.871
A2	579582.162	2027576.060	456.154
A3	579493.217	2027566.857	444.172
A4	579568.515	2027481.120	444.690
A5	579537.595	2027534.877	446.535
		study area 2	
B1	579718.878	2027263.832	443.748
B2	579756.781	2027191.277	444.362
B3	579790.363	2027128.612	444.624
B4	579739.465	2027127.284	442.509
B5	579665.410	2027238.617	442.565

Appendix B

MATLAB CODES

(a) CNN prediction using orthophoto

```
imageDir = tempdir;
load('rit18_data.mat');
train_data = switchChannelsToThirdPlane(train_data);
val_data   = switchChannelsToThirdPlane(val_data);
test_data  = switchChannelsToThirdPlane(test_data);
train_data_ortho = imread('train_image_ortho.tif');
figure
%CH3 = Red, CH2=Green, CH1 =Blue
montage(...
    {histeq(train_data_ortho(:,:,[3  2 1])), ...
   histeq(val_data(:,:,[3  2 1])), ...
   histeq(test_data(:,:,[3  2 1]))}, ...
   'BorderSize',10,'BackgroundColor','white')
title('RGB Component of Training Image (Left), Validation Image (Center), and Test Image (Right)')
% display the 3 figures in the 3 color bands separating the bands from the
% RGB image
figure
montage(...
   {histeq(train_data_ortho(:,:,1)), ...
   histeq(train_data_ortho(:,:,2)), ...
   histeq(train_data_ortho(:,:,3))}, ...
   'BorderSize',10,'BackgroundColor','white')
title('Channels 1 (Left), 2, (Center), and 3 (Right) of Training Image')
% defining the classes for the images to segment
%define_Class = char('0.Trees ', '1.Grass ','2.Road ','3.Other Class/Image Border')
```

```
define_Class = char('1.Trees', '2.Shadow','3.Grass', '4.Barrenland', '5.Others')

%create the vector name for the classes

%classNames = [ "Trees", "Grass", "Road", "Other Class/Image Border"]

classNames = [ "Trees", "Shadow", "Grass", "Barrenland", "Others"]

% creating the train label and adding the image to the train label

train_label_ortho = imread('train_label_ortho.png');

% Overlay the labels on the histogram-equalized RGB training image. Add a colorbar
to the image.

cmap = jet(numel(classNames));

B =
labeloverlay(histeq(train_data_ortho(:,:,1:3)),train_label_ortho,'Transparency',0.8,'Col
ormap',cmap);

figure

title('Training Labels')

imshow(B)

N = numel(classNames);

ticks = 1/(N*2):1/N:1;

colorbar('TickLabels',cellstr(classNames),'Ticks',ticks,'TickLength',0,'TickLabelInterp
reter','none');

colormap(cmap)

% Save the training data as a MAT file

save('train_data_ortho.mat','train_data_ortho');

%create .mat file for storing the training images

imds =
imageDatastore('train_data_ortho.mat','FileExtensions','.mat','ReadFcn',@matReader);

%Create a pixelLabelDatastore

pixelLabelIds = 1:5;

pxds = pixelLabelDatastore('train_Label_ortho.png',classNames,pixelLabelIds);

% ******Create a randomPatchExtractionDatastore (Image Processing Toolbox)
from the image datastore and the pixel label datastore
```

```
dsTrain =
randomPatchExtractionDatastore(imds,pxds,[256,256],'PatchesPerImage',16000);
% ************************************************
% correct matReader.m in four bands as the image
% ************************************************
%The random patch extraction datastore dsTrain provides mini-batches of data to the
network at each iteration of the epoch. Preview the datastore to explore the data.
inputBatch = preview(dsTrain);
disp(inputBatch)
% Create U-Net Network Layers
% change in flie createUnet.m for the number of classes which we changed
% from 18 classes to 5 classes as our requirement****
inputTileSize = [256,256,4];%<<<<4 is channels of imagery
lgraph = createUnet(inputTileSize);
disp(lgraph.Layers)
%Select Training Options
initialLearningRate = 0.05;
%maxEpochs = 150;*******original
maxEpochs = 15;
% Original
% minibatchSize = 16;
minibatchSize = 8;
l2reg = 0.0001;
options = trainingOptions('sgdm',...
    'InitialLearnRate',initialLearningRate, ...
    'Momentum',0.9,...
    'L2Regularization',l2reg,...
    'MaxEpochs',maxEpochs,...
    'MiniBatchSize',minibatchSize,...
    'LearnRateSchedule','piecewise',...
```

```
    'Shuffle','every-epoch',...
    'GradientThresholdMethod','l2norm',...
    'GradientThreshold',0.05,  ...
    'Plots','training-progress',  ...
    'VerboseFrequency',20);
%********Train  the Network**********************************
doTraining  = false;
if doTraining
    modelDateTime  = datestr(now,'dd-mmm-yyyy-HH-MM-SS');
    [net,info]  = trainNetwork(dsTrain,lgraph,options);
    save(['multispectralUnet-'  modelDateTime  '-Epoch-' num2str(maxEpochs)
'.mat'],'net','options');
else
    load('multispectralUnet-17-Nov-2020-12-13-11-Epoch-30.mat');
end
%*****************************************************
%**********Predict  Results on Test Data
predictPatchSize  = [512 512];
%**********create  and import  the image  for train data%
test_datanew = imread('test_image_ortho.tif');
segmentedImage  = segmentImage(test_datanew,net,predictPatchSize);
segmentedImage  = uint8(test_datanew(:,:,3)~=0)  .* segmentedImage;
figure
imshow(segmentedImage,[])
title('Segmented  Image')
segmentedImage  = medfilt2(segmentedImage,[7,7]);
imshow(segmentedImage,[]);
title('Segmented  Image  with Noise Removed')
B = labeloverlay(histeq(test_datanew(:,:,[3  2
1])),segmentedImage,'Transparency',0.8,'Colormap',cmap);
```

```
figure
imshow(B)
title('Labeled Test Image')
colorbar('TickLabels',cellstr(classNames),'Ticks',ticks,'TickLength',0,'TickLabelInterpreter','none');
colormap(cmap)
imwrite(segmentedImage,'results.png');
%create the ground truth from the test data (from image labeller) and predict the global
%accuracy%
pxdsResults = pixelLabelDatastore('results.png',classNames,pixelLabelIds);
pxdsTruth = pixelLabelDatastore('gtruth_test_area.png',classNames,pixelLabelIds);
ssm = evaluateSemanticSegmentation(pxdsResults,pxdsTruth,'Metrics','global-accuracy');
% Per class accuracy%%%
metrics = evaluateSemanticSegmentation(pxdsResults,pxdsTruth);
metrics
metrics.ClassMetrics
metrics.ConfusionMatrix
```

(b) CNN prediction using orthophoto with DSM

```
imageDir = tempdir;
load('rit18_data.mat');
train_data = switchChannelsToThirdPlane(train_data);
val_data   = switchChannelsToThirdPlane(val_data);
test_data  = switchChannelsToThirdPlane(test_data);
train_data_ortho = imread('train_image_ortho.tif');
figure
%CH3 = Red, CH2=Green, CH1 =Blue
montage(...
```

```
 {histeq(train_data_ortho(:,:,[3  2 1])), ...
 histeq(val_data(:,:,[3  2 1])), ...
 histeq(test_data(:,:,[3  2 1]))}, ...
 'BorderSize',10,'BackgroundColor','white')
title('RGB Component of Training Image (Left), Validation Image (Center), and Test Image (Right)')
% display the 3 figures in the 3 color bands separating the bands from the
% RGB image
figure
montage(...
 {histeq(train_data_ortho(:,:,1)), ...
 histeq(train_data_ortho(:,:,2)), ...
 histeq(train_data_ortho(:,:,3))}, ...
 'BorderSize',10,'BackgroundColor','white')
title('Channels 1 (Left), 2, (Center), and 3 (Right) of Training Image')
% defining the classes for the images to segment
%define_Class = char('0.Trees ', '1.Grass ','2.Road ','3.Other Class/Image Border')
define_Class = char('1.Trees', '2.Shadow','3.Grass', '4.Barrenland', '5.Others')
%create the vector name for the classes
%classNames = [ "Trees", "Grass", "Road", "Other Class/Image Border"]
classNames = [ "Trees", "Shadow", "Grass", "Barrenland", "Others"]
% creating the train label and adding the image to the train label
train_label_ortho = imread('train_label_ortho.png');
% Overlay the labels on the histogram-equalized RGB training image. Add a colorbar to the image.
cmap = jet(numel(classNames));
B =
labeloverlay(histeq(train_data_ortho(:,:,1:3)),train_label_ortho,'Transparency',0.8,'Colormap',cmap);
figure
```

```
title('Training Labels')

imshow(B)

N = numel(classNames);

ticks = 1/(N*2):1/N:1;

colorbar('TickLabels',cellstr(classNames),'Ticks',ticks,'TickLength',0,'TickLabelInterp
reter','none');

colormap(cmap)

% Save the training data as a MAT file

save('train_data_ortho.mat','train_data_ortho');

%create .mat file for storing the training images

imds =
imageDatastore('train_data_ortho.mat','FileExtensions','.mat','ReadFcn',@matReader);

%Create a pixelLabelDatastore

pixelLabelIds = 1:5;

pxds = pixelLabelDatastore('train_Label_ortho.png',classNames,pixelLabelIds);

% ******Create a randomPatchExtractionDatastore (Image Processing Toolbox)
from the image datastore and the pixel label datastore

dsTrain =
randomPatchExtractionDatastore(imds,pxds,[256,256],'PatchesPerImage',16000);

% ************************************************

% correct matReader.m in four bands as the image

% ************************************************

%The random patch extraction datastore dsTrain provides mini-batches of data to the
network at each iteration of the epoch. Preview the datastore to explore the data.

inputBatch = preview(dsTrain);

disp(inputBatch)

% Create U-Net Network Layers

% change in flie createUnet.m for the number of classes which we changed

% from 18 classes to 5 classes as our requirement****

inputTileSize = [256,256,4];%<<<<4 is channels of imagery

lgraph = createUnet(inputTileSize);
```

```
disp(lgraph.Layers)
%Select Training  Options
initialLearningRate  = 0.05;
%maxEpochs = 150;*******original
maxEpochs  = 15;
% Original
% minibatchSize  = 16;
minibatchSize  = 8;
l2reg = 0.0001;
options  = trainingOptions('sgdm',...
   'InitialLearnRate',initialLearningRate,  ...
   'Momentum',0.9,...
   'L2Regularization',l2reg,...
   'MaxEpochs',maxEpochs,...
   'MiniBatchSize',minibatchSize,...
   'LearnRateSchedule','piecewise',...
   'Shuffle','every-epoch',...
   'GradientThresholdMethod','l2norm',...
   'GradientThreshold',0.05,  ...
   'Plots','training-progress',  ...
   'VerboseFrequency',20);
%********Train  the Network**********************************
doTraining  = false;
if doTraining
   modelDateTime  = datestr(now,'dd-mmm-yyyy-HH-MM-SS');
   [net,info]  = trainNetwork(dsTrain,lgraph,options);
   save(['multispectralUnet-'  modelDateTime  '-Epoch-' num2str(maxEpochs)
'.mat'],'net','options');
else
```

```
    load('multispectralUnet-17-Nov-2020-12-13-11-Epoch-30.mat');
end
%**************************************************
%**********Predict  Results on Test Data
predictPatchSize  = [512 512];
%**********create  and import  the image  for train data%
test_datanew = imread('test_image_ortho.tif');
segmentedImage = segmentImage(test_datanew,net,predictPatchSize);
segmentedImage = uint8(test_datanew(:,:,3)~=0)  .* segmentedImage;
figure
imshow(segmentedImage,[])
title('Segmented  Image')
segmentedImage = medfilt2(segmentedImage,[7,7]);
imshow(segmentedImage,[]);
title('Segmented  Image  with Noise Removed')
B = labeloverlay(histeq(test_datanew(:,:,[3  2
1])),segmentedImage,'Transparency',0.8,'Colormap',cmap);
figure
imshow(B)
title('Labeled  Test Image')
colorbar('TickLabels',cellstr(classNames),'Ticks',ticks,'TickLength',0,'TickLabelInterp
reter','none');
colormap(cmap)
imwrite(segmentedImage,'results.png');
%create the ground  truth from the test data (from image labeller)  and predict  the
global
%accuracy%
pxdsResults  = pixelLabelDatastore('results.png',classNames,pixelLabelIds);
pxdsTruth  = pixelLabelDatastore('gtruth_test_area.png',classNames,pixelLabelIds);
```

```
ssm = evaluateSemanticSegmentation(pxdsResults,pxdsTruth,'Metrics','global-
accuracy');

% Per class accuracy%%%

metrics = evaluateSemanticSegmentation(pxdsResults,pxdsTruth);

metrics

metrics.ClassMetrics

metrics.ConfusionMatrix
```