App. Envi. Res. 43(4) (2021): 39-54

Applied Environmental Research

Journal homepage : http://www.tci-thaijo.org/index.php/aer

Land Cover Classification Based on UAV Photogrammetry and Deep Learning for Supporting Mine Reclamation: A Case Study of Mae Moh Mine in Lampang Province, Thailand

Tejendra Kumar Yadav[1], Polpreecha Chidburee[2*], Nattapon Mahavik[1]

[1] Department of Natural Resources and Environment, Faculty of Agriculture Natural Resources and Environment, Naresuan University, Phitsanulok, Thailand

[2] Department of Civil Engineering, Faculty of Engineering, Naresuan University, Phitsanulok, Thailand

[*] Corresponding author: polpreechac@nu.ac.th

Article History

Submitted: 14 February 2021/ Revision received: 7 June 2021/ Accepted: 17 June 2021/ Published online: 29 September 2021

Abstract

Detailed, accurate, and frequent mapping of land cover are the prerequisite regarding areas of reclaimed mines and the development of sustainable project-level for goals. Mine reclamation is essential as the extractive organizations are bounded by-laws that have been established by stakeholders to ensure that the mined areas are properly restored. As databases at the mines area become outdated, an automated process of upgrading is needed. Currently, there are only few studies regarding mine reclamation which has less potential of land cover classification using Unmanned Aerial Vehicle (UAV) photogrammetry with Deep learning (DL). This paper aims to employ the classification of land cover for monitoring mine reclamation using DL from the UAV photogrammetric results. The land cover was classified into five classes, comprising: 1) trees, 2) shadow, 3) grassland, 4) barren land, and 5) others (as undefined). To perform the classification using DL, the UAV photogrammetric results, orthophoto and Digital Surface Model (DSM) were used. The effectiveness of both results was examined to verify the potential of land cover classification. The experimental findings showed 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 of 0.904, Average Accuracy of 0.681, and Kappa index of 0.937. Our experiments showed 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. The procedure resulting from this research constitutes approach that is intended to be adopted by government organizations and private corporations so that it will provide accurate evaluation of reclamation in timely manner with reasonable budget.

Keywords: Land cover classification; Unmanned Aerial Vehicle (UAV); Photogrammetry; Deep learning (DL); Convolution neural network (CNN)

https://doi.org/10.35762/AER.2021.43.4.4

Introduction

Mine reclamation can be defined as the technique designed to assist governments reshape the terrain and plantation of the trees [1]. Mine reclamation is important because it helps to minimize destruction and avoids harmful environmental effects in mined areas, to return the soils to their state of pre-mining, and to protect the landscape [2]. Previously, monitoring techniques relied on visual perception or computer-related classification techniques [3] thorough mining surface cover mapping while enhancing the reliability of the mining area. Traditional methods of measuring cannot achieve high precision; additionally, the monitoring can only capture images at the point level and cannot monitor an entire surface [4]. Therefore, land cover mapping is a fundamental task for monitoring in mine reclamation and management.

UAV technology can acquire image data with high spatial and temporal resolution in a scalable manner at local scales, compared to conventional Photogrammetric works. UAV has recently been widely used in various photogrammetric applications, such as the management of natural resources [5–6]. Using UAV images combined with ground-based measurements, [7] investigated the extent of mine reclamation. The development of UAV photogrammetry has replaced conventional approaches and has become a modern technology for monitoring in mining areas [4]. UAV technology is being developed aggressively (over the years) regarding mapping applications. Moreover, UAV as a low budget process that fewer time constraints and less manpower rather than satellites or manned aircraft that insure expensive flight costs, are time-consuming, weather dependent data collection, restricted workability, limited flying time, and low ground resolution mapping process [8–11]. UAV photogrammetry is thus an attractive approach for mine reclamation monitoring.

A classification of land cover in a mine includes land cover details and the forms of human activity involved in land use [12]. Land cover is classified into two methods, pixel-based and object-based methods [13]. Mapping of land cover follows the fulfillment of a shared vision of mining development, where existing and potential land-use values and preferences are estimated, reported, and reviewed in a clear, transparent, and inclusive manner. Some research has focused on studying thorough mining surface cover mapping while enhancing the reliability of the mining surface area [3]. Monitoring the effects of the landscape in reclaimed mines provide useful information regarding the long-term geological and environmental consequences of mining operations [14]. Thus, the classification of land cover of a mine includes the complete land cover details.

Deep learning (DL) is the application of machine learning techniques that have been applied in recent years to the tasks involved in the classification of images, which provides information from images efficiently and the derived features are often very effective for processing images [15]. The visual recognition of deep learning is based on convolution neural networks CNN [16]. Recent development in CNN has allowed it to perform better with improved models in different domains, such as object detection, dense semantic labeling, and image classification [16–17]. The previous result published regarding the image classification in deep learning was provided by [18–19]. U-net is a specific model used in the Fully Connected Network (FCN), where the output segmented image is generated by a CNN. U-net was first proposed by [20] for biomedical segmentation by a special form of fully convolutional network. DL does not require feature engineering, nor does it need an automatic process, however it appeals to researchers from the remote sensing community for testing of its usability for land-cover mapping [21–23]. Therefore, environmental protection and monitoring of the reclaimed mines area is essential for concerned authorities,

so the requirements for updating of the database using the automated method of classification must be a priority.

The objective of the research is to pro-vide a potential algorithm based on deep learning for the classification of land cover at the reclaimed mine based on the UAV photogrammetric results. Extractive organizations are expected to provide definitive decision-making activities, those impacting the environment, that are obstructed by the management so to make sure that the company complies with the law and ensure that the mined areas are properly restored. Previous reviews indicate that monitoring the reclaimed mines area for land cover classification in Thailand DL has not been applied. This was a result of insufficient development of research, and the DL based on CNN with UAV was proposed to carry out land cover classification at the reclaimed mines area at low altitude.

The following section will include: Section 2 presents the study area and data-sets; Section 3 presents the method of DL with CNN; Section 4 presents the results and assessment of the system compared to two methods: orthophoto and combination of orthophoto with DSM; Section 5 present the discussion; and Section 6 focuses on the conclusion and next stages of monitoring mine reclamation.

Study area and datasets

The study site is located at Mae Moh mine in Mae Moh District, Lampang Province, Thailand. The Mae Moh mine is the largest Lignite mine in Thailand, it is one of the many locations where lignite has been producing fuel for the power plants managed by the Electricity Generating Authority of Thailand (EGAT). The terrain has a vegetated area of 135 km^2. The selected reclamation area of the mine was planted with a variety of trees for five years, such as the Silver Trumpet tree (*Tabebuia aurea*). The average height of the tress is between 1 and 2 meters. The study area is located in between the 18°20'07" N - 18°20'13" N latitude and 99°45'08" E - 99°45'13" E longitude, and covers an area of 9,744.66 m^2. The area of study in this research is shown in Figure 1.

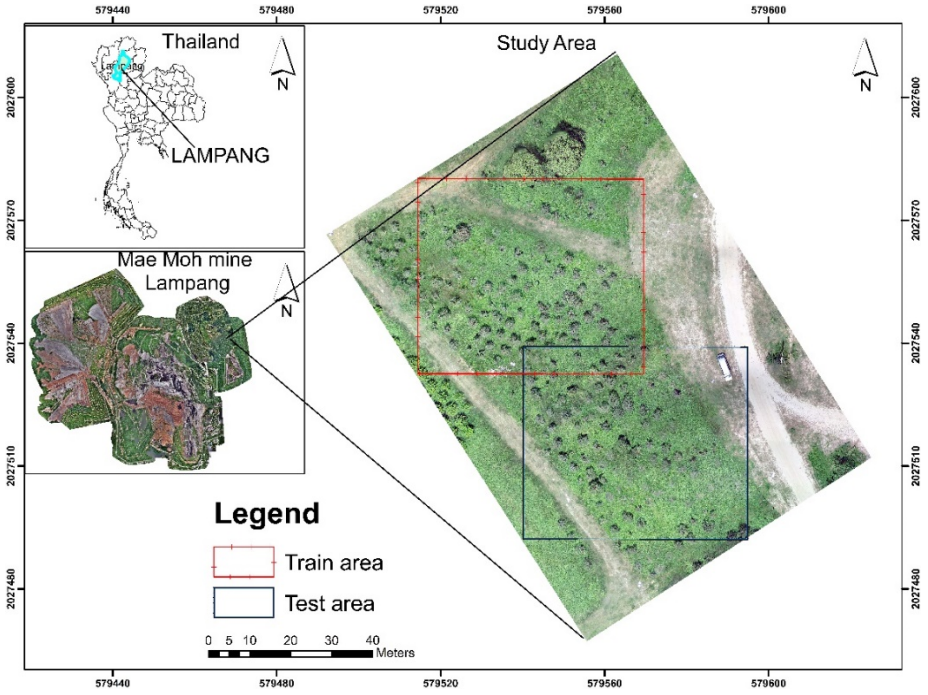

Figure 1 The study area in this research: the red and blue boxed areas are used for train and test datasets of DL, respectively.

1) Data acquisition

The DJI Phantom 4 of the UAV platform equipped with a normal digital camera was used to collect the imagery of the reclaimed area. The flight planning of DJI Phantom 4 for automatically controlling the UAV route was performed using the Pix4Dcapture mobile application on a smartphone. The UAV imagery was acquired on the 2nd October 2020 with a flying height of 50 m above the ground resulting in the ground sample distance (GSD) of 2.58 cm pixel^{-1}. For UAV flight planning parameters based on the pix4D capture of mobile phone application, double grid mission with camera angle of 45 degree, 80% forward overlap and 70% side-overlap were suitable for complex elevation of this area. The size of each image was 5472×3648 pixels (approximately 20 MP) and the UAV image set comprised 372 images. For geo-referencing, the placement of five targets over the experimental area were used as ground control points (GCPs) of the UAV photogrammetry. The locations of GCPs were observed by Real-Time Kinematic (RTK) technique of Global Navigation Satellite System (GNSS) surveying at centimeter-level accuracy with the coordinate system based on the World Geodetic System (WGS) 1984.

2) Data processing for UAV photogrammetry

For photogrammetric processing, firstly, the points measurement (also known as tie points) on the UAV imagery and geo-referencing with ground control points are used in bundle block adjustment for image orientation. Secondly, the dense image matching helps to generate a dense number of points as point cloud. Thirdly, the outcome of the point cloud is used for generating DSM. Finally, an orthophoto generated from UAV imagery is geometrically corrected such that the scale of the photograph is uniform and utilized in the same manner as a map. Orthophotos have the advantages of a high-detail, precise analysis combined with the advantages of a map, including a uniform size and true geometry. Orthophotos display all the valuable information of a photograph [24].

In this study, the Pix4D software was used to generate a dense point cloud, DSM, and orthophoto from a set of photos that overlap and include the geo-referenced details [25]. The images were matched for the orientation using the following functions: align photos and optimize cameras. Regarding the UAV photogrammetric processing, the Root Mean Square Errors (RMSE) of georeferencing with five GCPs was 2.23 cm. The point cloud was created using the orientation and position of the camera. The UAV outcomes obtained from Pix4D software were a single RGB image of orthophoto and DSM at the same resolution of 5 cm, as shown in Figure 2.

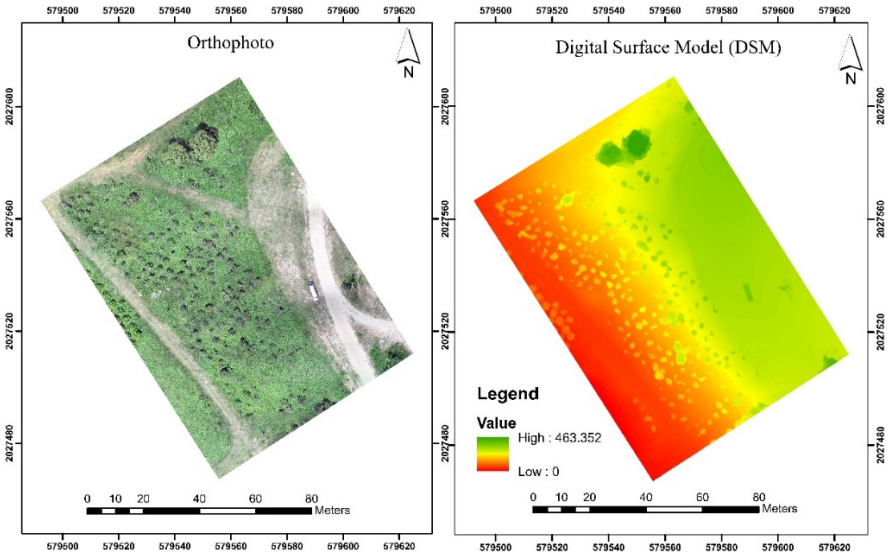

Figure 2 UAV photogrammetric results used in this work (a) Orthophoto and (b) DSM.

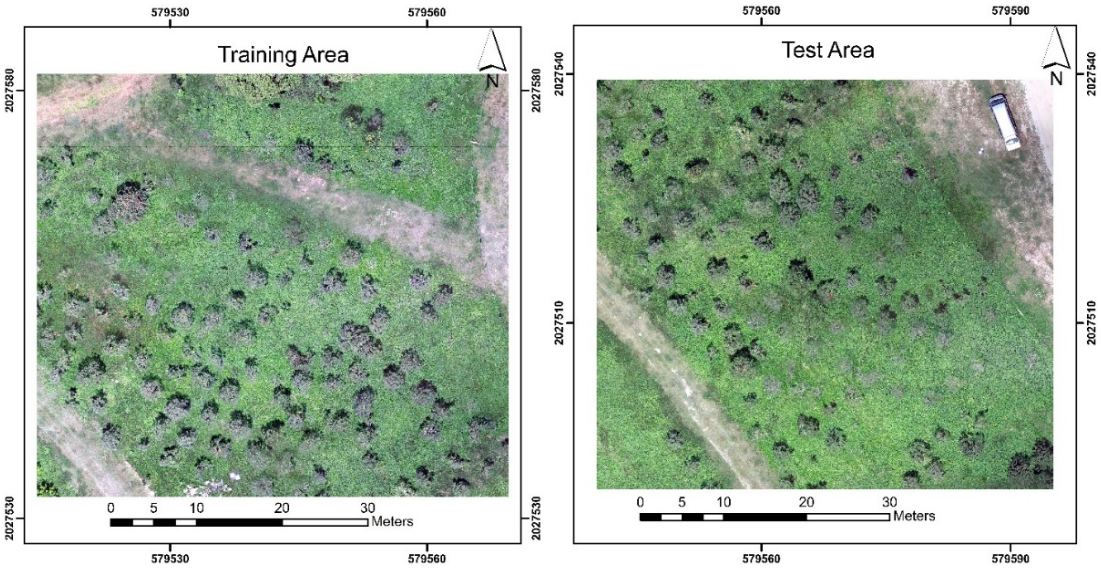

Figure 3 The two datasets for DL: (a) training and (b) test area subset from orthophoto.

To prepare the dataset for DL, both orthophoto and DSM of the UAV photogrammetric results were selected into two areas; train and test, as shown in Figure 1. Both areas were taken into two subsets from orthophoto and DSM using ArcGIS at a resolution of 1,101×988 and 1,100×985 pixels, as shown in Figure 3 (a) and (b), respectively.

Methodology

1) Pre-processing

Before the model training with DL, both datasets of orthophoto and DSM should be in the same type of grid data. The digital number (DN) in each band of RGB image for orthophoto is based on the data type integer and the value of DN is between 0 and 255. However, the value of gird data in DSM is the elevation as a decimal in centimeter-level. Since two datasets have unique attributes, the rescaling of data was implemented using normalization. Normalization refers to scaling the values of the DSM as the same data type of orthophoto.

2) Model training

In deep learning techniques, patch level analysis is used to resolve challenges raised by distortion and optimization for segmentation, which problems are associated with the extraction of pixel-level and object-level features [26]. During the training process, the module can autonomously learn the parameters required for spatial transformation and does not need to add any additional supervisory processing to the training. Patches are the sub-boxes of an image that is used at one time of a convolutional layer which is divided into the block of p×p (patch size) and each block is independently considered. Based on the resolution of the RGB image and/or DSM and the expected size of the objects in the image, the size of the image patch is used to train the model. The patch size used in this study was 512×512 pixels.

The accuracy and loss were calculated during the training process. The accuracy and loss curves show a large alternation phenomenon at the beginning of training which is more evident in training dataset. This observation is related to the small mini-batch size set in this study. Also, the scale of network oscillations in the late training period is even smaller than that in the early training period because of the strategy having the learning rate every 10 epochs. The model with the highest accuracy in the training dataset is selected as the final model and is applied to the test dataset.

2.1) Convolution neural network (CNN)

CNN is one of the most commonly used deep learning algorithms, which has recently achieved a significant interest in image processing [27]. It is also primarily designed to deal with inputs in the form of multidimensional arrays [28]. They were effective in tasks with multiple arrays configured in 2D, such as object recognition in images [16]. CNN includes three types of layers 1) convolutional layers, 2) pooling layers and 3) fully-connected layers, as follows:

1) Convolutional layer

The Convolution layer is the main component of the CNN with local relations and weights of mutual characteristics. The main task of the Convolution layer is to learn the feature representations. Each neuron of the same map characteristics is used to extract local features at different positions in the former layer, but its extraction for single neurons is local characteristics of the same positions in the former map [29].

2) Pooling layer

The goal of pooling layers is to gradually decrease the representation dimensionality and thus further minimize the number of parameters and the computational complexity of the model. In the input, the pooling layer operates over each activation map, and scales, using the "MAX" function [30]. To boost performance, pooling was used with non-equal filters and measured overlaps of the areas.

3) Fully connected layers

The fully-connected layer is close to how neurons in a CNN are organized. Neurons in the fully-connected layer, are directly linked to without being connected to any layer, the neurons in the two adjacent layers [31].

The network consisted of one 2D convolution layer in the description for this analysis, learning 64 kernels with a filter size of 3×3. The convolution layer passes through a rectified linear unit (ReLU) layer before being down-sampled with a filter size of 2×2 through max pooling. A dense layer is followed by ReLU activation with 128 neurons following final softmax dense layer with five neurons indicating the number of classes. The network was configured using a classifier from Adam with a batch size of 32, and 15 epochs. The loss for the training was calculated using categorical cross-entropy.

2.2) U-net

The U-net is a deep learning network architecture, which is widely used for semantic segmentation tasks. It was first proposed by Ronneberger, et al. [20] for medical image segmentation. U-net model mainly relies on convolution operations to learn high-level features. Unlike a general convolutional network, U-net is a full-convolution network that does not include a fully connected layer and also not demanding on the amount of dataset. This network is simple, efficient, and easily used. Its architecture was mainly divided into two parts, namely, an encoder and a decoder. The encoder continuously sampled through multiple convolution layers to obtain different image levels. The decoder performed multilayer deconvolution on the top-level feature map and combined different feature levels in the down-sampling process to restore the feature map to the original input image size and complete the end-to-end semantic segmentation task of the image. In this study, the network was improved to make it suitable for the land cover classification at reclaimed mines area based on the original U-net architecture.

2.3) Ground truth data

Ground truth can be said as the representation of the real features over the Earth's surfaces and also is essential to assess the accuracy of the trained model from DL. This is used in statistical models to prove or disprove key findings. The ground truth data involves the process of collecting the right objective (verifiable) data for this test [32]. The processed ground image (ground truth image) dataset was manually labeled

by visual interpretation using the Image labeler application in MATLAB R2020b. Each image has a hand-drawn mask indicating the classified area as ground truth (GT). GT helps image data to be linked with the features to actual characteristics on the image with per-pixel GT for multiple classes. Additionally, imagery can be reviewed to verify ground cover and extract different features. The key aspect of the GT is the pixel resolution labeling.

In this research, five land cover groups were considered: (i) trees, (ii) shadow, (iii) grassland, (iv) barren land, and (v) others. The database provides GT 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 with area of 2,702.85 m^2, as shown in Table 1. Each class was labeled according to its respective class.

Table 1 Number of pixels per land cover class in the test area for the ground truth

Land cover class	No. of pixel	Areas (m^2)
Trees	138,452	346.13
Shadow	51,976	129.94
Grassland	725,547	1,813.87
Barren land	163,008	407.52
Others	2,156	5.39

3) Experimental design

The classification of reclaimed mine areas can be automatically performed using DL from outcomes of UAV photogrammetry (i.e. orthophoto or DSM). To assess the potential of land cover classification in this research, the efficacy of two environments was explored using the DL with CNN from 1) the orthophoto (RGB) imagery and 2) the combination of orthophoto with DSM. For the step of the training model, both the orthophoto imagery, and the combination of orthophoto with DSM were trained separately for 15 epochs, and the two model were generated. The two trained models were applied separately to the train and test area as the prediction results for the land cover classification of the reclaimed mines area. The classified results from both trained models were evaluated with the same ground truth data using the DL. Figure 4 shows the workflow and evaluation performed with programming on MATLAB R2020b software for automated processing.

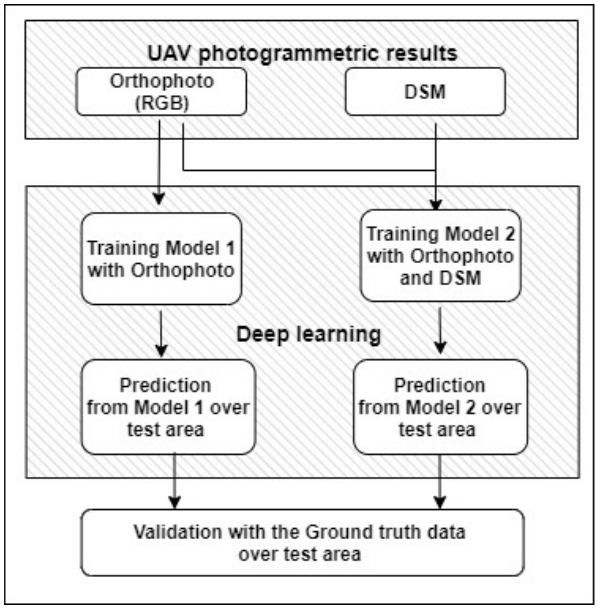

Figure 4 Workflows and evaluation of the proposed classification based on DL.

4) Accuracy Assessment

In the evaluation process, the variance between the outcome of classification and the reference data is due to a classification error. However, the error of the misclassified area should be investigated to provide a reliable report on the accuracy of classification. The confusion matrix table helps an algorithm to visualize its results [33]. The instances in a projected class are expressed in each row of the matrix, while each column represents instances in a real class (or vice versa) [34]. After the generation of a confusion matrix, other important accuracy assessment elements, including overall accuracy (OA), Average Accuracy (AA), per class accuracy (PA), and kappa coefficient, were derived.

In general, overall accuracy shows us what percentage was correctly mapped from all of the study areas. The percentage of total classified

pixels currently classified in the particular land cover groups was determined by OA and was computed by dividing the total correctly classified pixels (T_{mn} or the sum of major diagonal) by the total number of pixels (N) in the confusion matrix, as shown in Eq. 1, and PA was calculated using Eq. 2;

$$OA = \frac{\sum T_{mn}}{N} \quad \text{(Eq. 1)}$$

$$PA = \frac{\sum T_{mn}}{S_i} \quad \text{(Eq. 2)}$$

Where T_{mn} is the total number of the correctly classified pixels in row m and column n and S_i is the total number of pixels in the row. Average accuracy (AA) was computed as in Eq. 3 defining C as the number of the classes.

$$AA = \frac{\sum_1^m PA}{C} C \quad \text{(Eq. 3)}$$

Alternatively, the Kappa coefficient is a calculation of the actual and predicted values consistency of an error matrix that takes into consideration non-diagonal elements [35]. Kappa analysis is known as a powerful way to evaluate a single mistake matrix and comparison of the differences between error matrix [36]. The value of the kappa ranges from 0 to 1. The analysis of Kappa produces a K metric, a quantitative estimate of the properly labeled pixels and the degree of consensus or Landis and Koch [37]. The kappa was calculated as:

$$k = \frac{N \sum_{p,q=1}^m T_{pq} - \sum_{p,q=1}^m R_i, S_j}{N^2 - \sum_{p,q=1}^m R_i, S_j} \quad \text{(Eq. 4)}$$

Where m is the number of classes, T_{pq} is the number of correctly classified pixels in row p and column q, R_i is the total number of pixels in row p, S_j is the total number of pixels in column q and N is the total number of pixels.

Model validation is used to determine how effective an estimator is on data that it has been trained on and generalizable to new input. The training and validation time on the horizontal axis and a consistency metric is plotted by validation curves. An accuracy curve is a standard measure of consistency derived from the matrix of uncertainty and displays both correct and incorrect classifications [38]. The accuracy is calculated as shown in Eq. 5.

$$Accuracy = \frac{\sum_{p=1} C_{pq}}{\sum_{p=1}^n \sum_{q=1}^n C_{pq}} \quad \text{(Eq. 5)}$$

Where C_{pq} 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 interpreting the potential of DL is the loss curve which demonstrate whether the process of optimization and relative improvement in learning enhances several epochs during the training [39]. The loss function is calculated using Eq. 6.

$$\text{Loss} = 1 - \text{accuracy} \quad \text{(Eq. 6)}$$

Results

After training of DL based on U-net model, land cover classification for the train and test dataset is predicted. During model training, the accuracy and loss curves were used to qualitatively test the outcomes and suggested that in the processing there was no indication of overfitting. Validation curves (accuracy and loss) of the two models evaluate the results of orthophoto and 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 10^{th} 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 curve was the highest when the iteration was stopped.

Regarding the comparison results for the predicted results from DL, Figure 5 and 6 show the classification of land cover from 1) orthophoto and 2) combination of orthophoto with DSM.

According to a visual interpretation between Figure 5 (c) and (d) with Figure 6 (c) and (d), both of them were similar in each class. It is likely that there is insignificantly identified for the type of land cover. In addition, the tree canopy for both the datasets was appropriately classified, and also shadow helps to detect the correct amount of area with the trees. In conclusion, the DL for land cover classification demonstrates an efficient approach for monitoring the vegetation (i.e. trees) at the reclaimed mines.

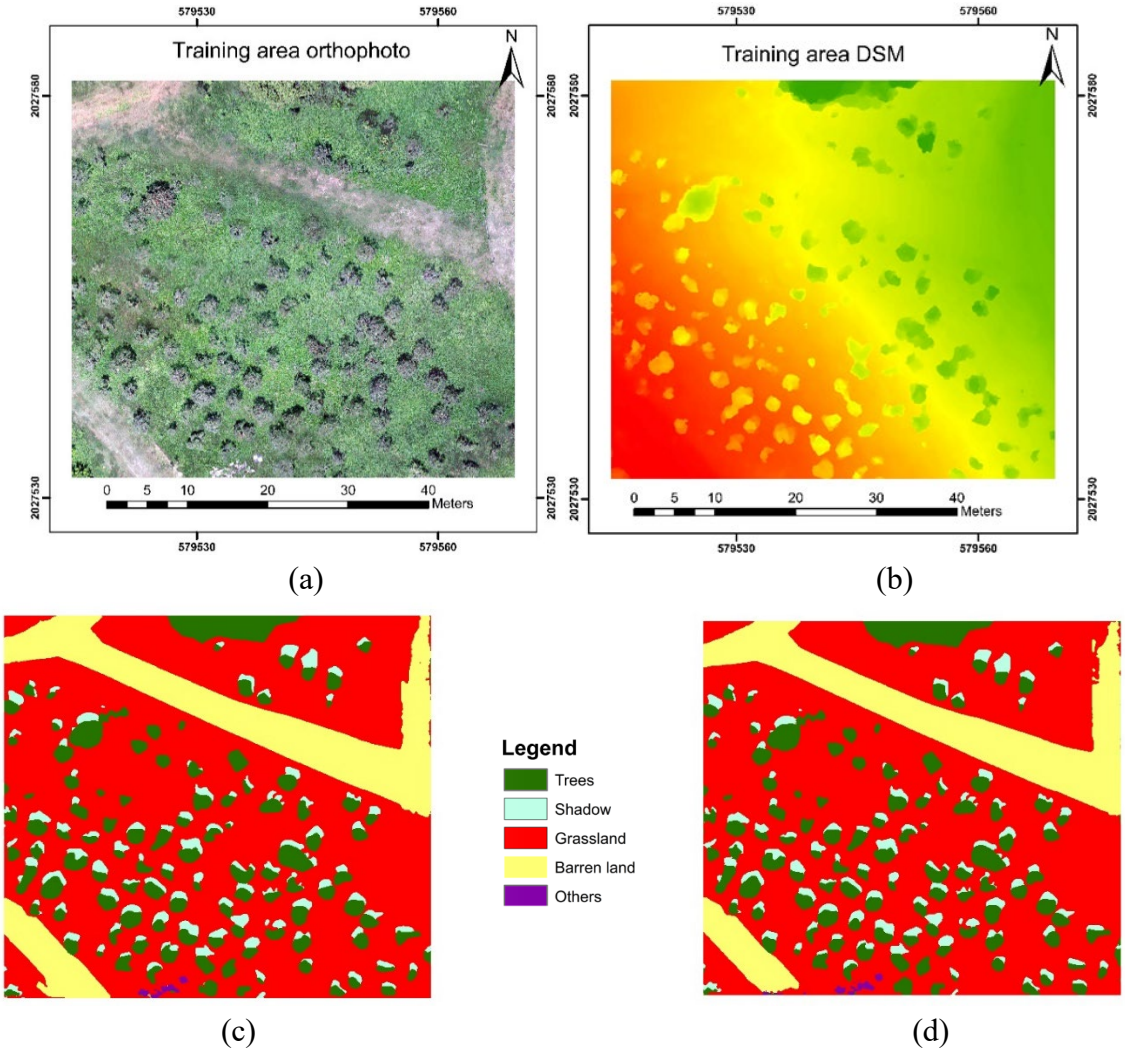

Figure 5 Training datasets and results after classification using DL (a) Orthophoto, (b) DSM, (c) land cover map from orthophoto and (d) land cover map from orthophoto and DSM with legend.

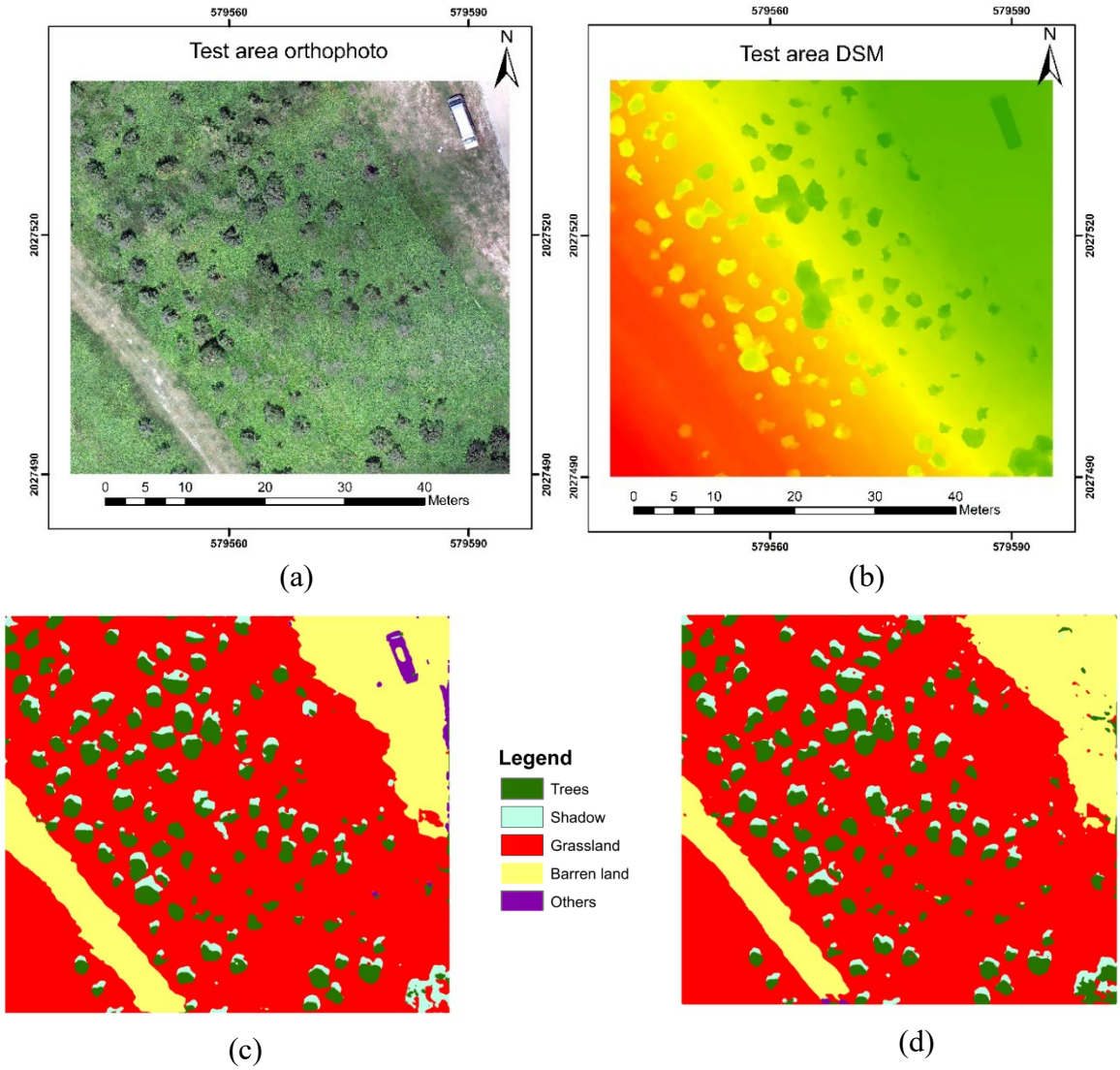

Figure 6 Test datasets and results after classification using DL (a) orthophoto, (b) DSM, (c) land cover map from orthophoto and (d) land cover map from orthophoto and DSM with legend.

In Figure 7, the number of pixels in the particular class of both datatsets were derived from the classification results of Figure 6 (c) and (d). The classificated results were approximately balanced for both datasets over test area. The number of shown pixels for the land cover classification obtained from both datasets range from a minimum of 0.1% and 0.3% (at the other class) to a maximum of 56.0% and 58.4% (at the grassland class). Therefore, the balanced distribution of pixels between land cover classes resulted in a simple equal distribution of training samples.

For validation of prediction results using DL, the evaluation criteria for OA, AA, and K were presented in Table 2. When the orthophoto was combined with DSM, the overall accuracy (OA) for training and testing datasets were 0.987 and 0.904, respectively. However, the OA and K of training and testing for the orthophoto and a combination of orthophoto with DSM are similar. 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 class, as illustrated in Table 3.

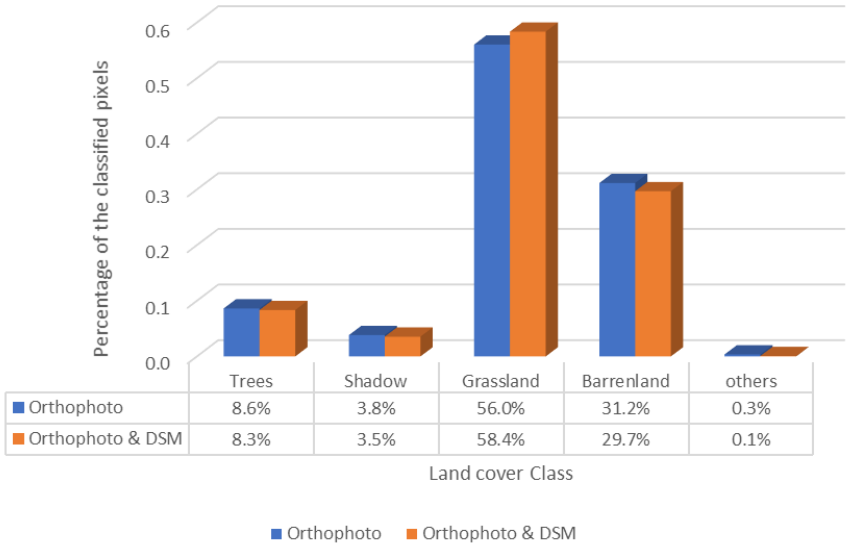

Figure 7 Percentage of classified pixels from test area for each land cover class.

Table 2 Output using standard accuracy metrics of the classification methods

Dataset	Model	Overall accuracy (OA)	Average accuracy (AA)	Kappa Index (K)
Training	Orthophoto	0.987	0.938	0.991
	Orthophoto and DSM	0.987	0.948	0.991
Testing	Orthophoto	0.900	0.837	0.935
	Orthophoto and DSM	0.904	0.681	0.937

Table 3 Training and test dataset with PA obtained by a different class

Class	Training		Test	
	Orthophoto	Orthophoto and DSM	Orthophoto	Orthophoto and DSM
Trees	0.978	0.983	0.731	0.770
Shadow	0.967	0.969	0.770	0.788
Grassland	0.992	0.994	0.920	0.933
Barren land	0.982	0.970	0.963	0.913
Others	0.772	0.824	0.800	0.000

Based on the training and test results for each class, Table 3 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 were relatively higher than the orthophoto. Also, from Table 3, in terms of the classification results of the test dataset, the PA of combination orthophoto with DSM was more accurate than of 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. It is likely that the classification results for the vegetation obtained from orthophoto and 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 class for training with DL (from Figure 7), dataset of the labelled pixel for GT was the least of all. It is noticed that a contributing factor may also be the lack of adequate sampling [40]. On the other hand, most of the sampled pixels in the training dataset were for grassland (approximately 58.5%); the highest precision of classification was predicted by grassland for both approaches.

Discussion

The findings of the study clearly show that the UAV approach for supporting mine reclamation provides photogrammetric results, consisting of orthophoto and DSM, that is possibly more useful information for monitoring land cover. Deep learning technique can employ the automatic extraction of land clover using the trained datasets from orthophoto and/or DSM at the pixel level (Figure not shown). UAV and deep learning approach provide land cover classification especially vegetation area of mine reclamation which is supporting measure and also is valuable tool for monitoring and managing the reclaimed mines. 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 accurate than orthophoto, especially the vegetation area (trees and grassland), as shown in Tables 3 and 4. 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 [41]. 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. [40]. 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 [32, 41]. 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 [42–43]. The results for the image classification as in Figures 6 and 7 analyzed shows precise results based on Tables 3 and 4 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 labelling has been added [16–17, 32]. 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 [15, 21]. 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 im-

provements 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 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 Li, et al. [44], 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^{-1}. The DL with U-net architecture showed land cover classification results in their mapping accuracies using between orthophoto and combination of orthophoto with DSM.

Based on deep learning of land cover classification, accuracy depends on several factors, such as the surface types, and the solar angle during image acquisition using UAV. The effect of the solar angle can be considered as 1) shadow, and 2) the intensity of light. Firstly, the presence of shadows due to solar angle is not able to visualize all the vegetation area [43]. Secondly, the colors of the visible spectrum, as well as band of the wavelengths, are present with relatively various intensity of sunlight [45] which may lead to the different color of the UAV imagery. These effects from solar angle should be considered in the future study.

Conclusions

Monitoring open-pit mine areas with UAV imagery and deep learning is effective and valuable as it offers timely and reliable information for mine reclamation. The resulting method can periodically and consistently sample the pixel size restoration at the region of interest for confined or unsafe areas to be monitored. It is capable to achieve detailed documents such as orthophoto, DSM, and land cover maps integrated with previous cartography and official databases. The introduction of emerging DL and Unmanned aerial systems (UAS) brings us exciting opportunities for monitoring the reclamation of the low-cost device and extensive analysis. The overall workflow was automated for land cover mapping also the particular emphasis on troublesome areas could be applied. However, it is convenient for the expertise to check the right implementation and assist with key decisions for land cover classification, such as the digitization of training and test areas. Moreover, relative to what would have been achieved in the traditional inspection of the reclaimed areas, the fieldwork duration in our case study was substantially shortened. To conclude, the time required to collect the ground-truth data and the UAV flight is enough to examine the reclaimed area, also the data processing can be carried out within a short time. From the experiment study at Mae Moh mine in Lampang, Thailand, the trained model in DL was applied to identify the vegetation areas for the test and train datasets from both orthophoto and DSM have higher precision of OA, AA, k, and PA than only orthophoto due to the height information. The precise values of land cover classification for test area from combining orthophoto with DSM with OA, AA, and K were 0.904, 0.681, and 0.937, respectively.

In future work, in terms of the monitoring mine reclamation on the vegetation area for land cover classification, applying trained model of classification for UAV dataset at different time will be evaluated. Also, it will offer a practical solution in mine reclamation for relevant authorities to make appropriate decision.

Acknowledgements

This research was supported by Thailand International Cooperation Agency (TICA) through the Naresuan University. We are appreciated to "Thailand International Cooperation Agency

(TICA)" for providing a research fund. Special thanks to the Faculty of Agriculture, Natural Resource and Environment for the research environment (Naresuan University) and the Electricity Generating Authority of Thailand (EGAT) at the Mae Moh mine for providing the study area and the Pix4D software for UAV photogrammetric processing. In addition, MATAB for student's version was used for the DL.

References

[1] Zhengfu, B., Inyang, H.I., Daniels, J.L., Frank, O., Struthers, S. Environmental issues from coal mining and their solutions. Mining Science and Technology (China), 2010, 20(2), 215–223.

[2] Yu, X., Mu, C., Zhang, D. Assessment of land reclamation benefits in mining areas using fuzzy comprehensive evaluation. Sustainability, 2020, 12(5), 2015.

[3] Chen, W., Li, X., He, H., Wang, L. A review of fine-scale land use and land cover classification in open-pit mining areas by remote sensing techniques. Remote Sensing, 2018, 10(1), 15.

[4] Ren, H., Zhao, Y., Xiao, W., Hu, Z. A review of UAV monitoring in mining areas: Current status and future perspectives. International Journal of Coal Science & Technology, 2019, 1–14.

[5] Boon, M.A., Greenfield, R., Tesfamichael, S. Wetland assessment using unmanned aerial vehicle (UAV) photogrammetry. 2016.

[6] Pashaei, M., Starek, M.J. Fully convolutional neural network for land cover mapping in a coastal wetland with hyperspatial UAS imagery. *In:* IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium, 2019, 6106–6109.

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

[8] Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., Bareth, G. Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. Remote Sensing, 2014, 6(11), 10395–10412.

[9] Dash, J.P., Watt, M.S., Pearse, G.D., Heaphy, M., Dungey, H.S. Assessing very high resolution UAV imagery for monitoring forest health during a simulated disease outbreak. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 131, 1–14.

[10] Tian, J., Wang, L., Li, X., Gong, H., Shi, C., Zhong, R., Liu, X. Comparison of UAV and WorldView-2 imagery for mapping leaf area index of mangrove forest. International journal of applied earth observation and geoinformation, 2017, 61, 22–31.

[11] Aasen, H., Burkart, A., Bolten, A., Bareth, G. 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, 2015, 108, 245–259.

[12] Rigge, M., Wylie, B., Gu, Y., Belnap, J., Phuyal, K., Tieszen, L. Monitoring the status of forests and rangelands in the western United States using ecosystem performance anomalies. International Journal of Remote Sensing, 2013, 34(11), 4049–4068.

[13] Li, D., Ke, Y., Gong, H., Li, X. Object-based urban tree species classification using bi-temporal WorldView-2 and WorldView-3 images. Remote Sensing, 2015, 7(12), 16917–16937.

[14] Padmanaban, R., Bhowmik, A.K., Cabral, P. A remote sensing approach to environmental monitoring in a reclaimed mine

area. ISPRS international journal of geo-information, 2017, 6(12), 401.

[15] Wu, J., Yu, Y., Huang, C., Yu, K. Deep multiple instance learning for image classification and auto-annotation. *In:* Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, 3460–3469.

[16] Krizhevsky, A., Sutskever, I., Hinton, G. E. Imagenet classification with deep convolutional neural networks. Communications of the ACM, 2017, 60(6), 84–90.

[17] Tang, S., Yuan, Y. Object detection based on convolutional neural network. *In:* International Conference-IEEE–2016, 2015.

[18] Chen, Q., Song, Z., Hua, Y., Huang, Z., Yan, S. Hierarchical matching with side information for image classification. *In:* 2012 IEEE conference on computer vision and pattern recognition, 2012, 3426–3433.

[19] Song, Z., Chen, Q., Huang, Z., Hua, Y., Yan, S. Contextualizing object detection and classification. *In:* Conference on Computer Vision and Pattern Recognition, 2011, 1585–1592.

[20] Ronneberger, O., Fischer, P., Brox, T. U-Net: Convolutional networks for biomedical image segmentation. *In:* Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Cham, 2015, 234–241.

[21] Alshehhi, R., Marpu, P.R., Woon, W.L., Dalla Mura, M. Simultaneous extraction of roads and buildings in remote sensing imagery with convolutional neural networks. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 130, 139–149.

[22] Ma, X., Wang, H., Wang, J. Semisupervised classification for hyperspectral image based on multi-decision labeling and deep feature learning. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 120, 99–107.

[23] Vetrivel, A., Gerke, M., Kerle, N., Nex, F., Vosselman, G. 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, 2018, 140, 45–59.

[24] Jicai, C., Wunian, Y., Xin, Y. The key technologies of the true orthoimage production based on DEM and DBM. Remote Sensing Technology and Application, 2013, 27(2), 168–172.

[25] Kalantar, B., Mansor, S.B., Sameen, M. I., Pradhan, B., Shafri, H.Z. Drone-based land-cover mapping using a fuzzy unordered rule induction algorithm integrated into object-based image analysis. International Journal of Remote Sensing, 2017, 38(8–10), 2535–2556.

[26] Karimi, D., Dou, H., Warfield, S., Gholipour, A. Deep learning with noisy labels: Exploring techniques and remedies in medical image analysis. Medical Image Analysis, 2020, 65, 101759.

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

[28] Coşkun, M., Yildirim, Ö., Uçar, A., Demir, Y. An overview of popular deep learning methods. European Journal of Engineering and Technology Research, 2017, 7(2), 165–176.

[29] Nair, V., Hinton, G.E. Rectified linear units improve restricted boltzmann machines. *In:* International Conference on Machine Learning, 2010.

[30] Koushik, J. Understanding convolutional neural networks. arXiv preprint arXiv: 1605.09081, 2016.

[31] Albawi, S., Mohammed, T.A., Al-Zawi, S. Understanding of a convolutional neural network. *In:* International Conference on Engineering and Technology, 2017, 1–6.

[32] Li, S., Hao, Q., Gao, G., Kang, X. The effect of ground truth on performance evaluation of hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(12), 7195–7206.

[33] Deng, X., Liu, Q., Deng, Y., Mahadevan, S. An improved method to construct basic probability assignment based on the confusion matrix for classification problem. Information Sciences, 2016, 340–341, 250–261.

[34] Powers, D. Evaluation: From precision, recall and F-factor to ROC, informedness, markedness & correlation. Machine Learning: Science and Technology, 2008, 2.

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

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

[37] Landis, J.R., Koch, G.G. The measurement of observer agreement for categorical data. Biometrics, 1977, 159–174.

[38] Abd, H.A.A.-R., Alnajjar, H.A. 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, 2013, 3(6), 26–33.

[39] Foody, G.M. Status of land cover classification accuracy assessment. Remote sensing of environment, 2002, 80(1), 185–201.

[40] Marcos, D., Volpi, M., Kellenberger, B., Tuia, D. Land cover mapping at very high resolution with rotation equivariant CNNs: Towards small yet accurate models. ISPRS journal of photogrammetry and remote sensing, 2018, 145, 96–107.

[41] Al-Najjar, H.A., Kalantar, B., Pradhan, B., Saeidi, V., Halin, A.A., Ueda, N., Mansor, S. Land Cover Classification from fused DSM and UAV Images Using Convolutional Neural Networks. Remote Sensing, 2019, 11(12), 1461.

[42] Movia, A., Beinat, A., Crosilla, F. Shadow detection and removal in RGB VHR images for land use unsupervised classification. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 119, 485–495.

[43] Adeline, K.R.M., Chen, M., Briottet, X., Pang, S.K., Paparoditis, N. Shadow detection in very high spatial resolution aerial images: A comparative study. ISPRS Journal of Photogrammetry and Remote Sensing, 2013, 80, 21–38.

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

[45] Walter, V., Saska, M., Franchi, A. Fast mutual relative localization of uavs using ultraviolet led markers. *In:* International Conference on Unmanned Aircraft Systems, 2018, 1217–1226.