POST-CLASSIFICATION ENHANCEMENT IN THE RESULT OF DEEP LEARNING LAND COVER CLASSIFICATION USING VERY-HIGH RESOLUTION SATELLITE IMAGERY

Yofri Furqani Hakim *[1] and Fuan Tsai [2]
[1]Ph.D Student, International Ph.D. Program in Environmental Science and Technology, National Central University
No. 300, Zhongda Rd., Zhongli District, Taoyuan City 32001, Taiwan
Email: yofrifh@gmail.com

[2]Professor, Center for Space and Remote Sensing Research, National Central University
No. 300, Zhongda Rd., Zhongli District, Taoyuan City 32001, Taiwan
Email: ftsai@csrsr.ncu.edu.tw

KEY WORDS: Land Cover Classification, Semantic Segmentation, Deep Learning, U-Net, Post-Classification Enhancement, Image Morphology.

ABSTRACT: Land cover is one of the fundamental data utilized in spatial analysis for a range of applications, including climate, environment, natural resources, agriculture, forestry, planning, health, and even social issues. The demand for land cover data is diverse, ranging from global scale, regional scale, and detailed scale. Furthermore, data updates also become a necessity for users. The current trend is the requirement for more accurate and up-to-date land cover data. The growing demand for accurate and up-to-date land cover data is in sync with improvements in satellite imagery data acquisition technology, which now offers improved spatial resolution and more effective satellite imagery data processing. Satellite imagery data processing technology is also growing rapidly with the advance of artificial intelligence for semantic classification or segmentation. This research uses a deep learning approach to classify land cover on Pleiades very high-resolution satellite imagery. A post-classification enhancement is also carried out to improve the consistency and accuracy of the deep learning classification results. The preliminary research results show that post-classification enhancement with the algorithms proposed in this study can increase the accuracy of classification results using deep learning-based approaches by approximately 2%. The original Overall Accuracy and Kappa of the deep learning classification results were 0.84 and 0.79. After the post-classification enhancement process, the Overall Accuracy and Kappa values increased to 0.86 and 0.81.

1. INTRODUCTION

1.1 Background

Spatial-based development planning is a popular subject in Indonesia. Moreover, after the government implemented new approaches in development planning, namely: holistic-thematic, integrative, and spatial. Planning with a holistic approach implies considering all aspects of the activity or program, rather than only concentrating on one aspect. Planning with a thematic approach while planning means focusing on particular topics or concepts that are relevant to activities or programs. Planning with an integrative approach in planning means integrating all relevant aspects in planning activities or programs. And, planning with a spatial approach means considering spatial and geographical aspects in planning activities or programs.

The spatial approach that is applied encourages the increasing need for base maps by the national and local governments. The increase in the need for base maps is mainly for large-scale base maps. Large-scale base maps are used by governments to prepare detailed spatial plans. The base map, also known as a topographic map, consists of 8 layers, namely: coastlines, hydrography, hypsography, land cover, building and public facility, transportation and utility, administration boundary, and toponym.

One of the layers that is often needed from this base map is land cover. The most proven method to get land cover data is using photogrammetry. At present, the advancement of remote sensing technologies is currently accelerating. There are many different types of remote sensing data, including optical and active sensors, multispectral and hyperspectral sensors, also low and high spatial resolution. This remote sensing data provides an alternative for providing land cover data other than aerial photo data. This research use Pleiades Satellite Imagery, which is one of very-high-resolution satellite imagery with 0.5 m resolution.

The advancement of science and technology is speeding up right now. Artificial intelligence is being employed in a variety of fields, including remote sensing data processing. In order to extract data on land cover from images, Deep Convolutional Neural Networks (DCNNs) have emerged as one of the most competitive deep learning algorithms. One of the DCNN architectures that is suitable for remote sensing data processing is U-net. U-net was first introduced by

Ronneberger et.al (2015) for medical purposes. This research uses U-net architecture to derive land cover data from very high-resolution satellite imagery.

The common problem in the land cover classification process is that there are many noises in the result of classification. In several cases, the noise can significantly decrease the accuracy of the result. So, it is important to be more aware of this noise. After the classification process, the post-classification enhancement process is carried out to remove the noises. This research focuses on extracting land cover data from very-high-resolution satellite imagery using deep deep-learning approach and improving the accuracy of the result by implementing post-classification enhancement.

1.2 Study Area

The city of Mataram is located in Lombok Island, Indonesia. It is the capital of Nusatenggara Barat Province. The city is mainly covered by built-up areas, agricultural land, and plantations. The study area of this research takes place in this city which can be seen at Figure 1.

Figure 1. Study area at City of Mataram, Indonesia.

2. METHODOLOGY

2.1 Dataset

The dataset used in this research is Pleiades Satellite Imagery. This very-high-resolution imagery data consists of 5 bands with 0.5 m resolution for the panchromatic band, and 2 m resolution for multispectral bands (Red, Green, Blue, and Near Infrared). The acquisition year is 2015. As ground truth data, the research uses land cover data from the visual classification of the Pleiades Satellite Imagery. The delineation of the ground truth data was guided with the digital topographic map at a scale of 1:5.000.

2.2 Data Preparation

This study uses 7 bands of image data as input in the deep learning classification process. The 7 bands are Red, Green, Blue, Near-Infrared, NDVI (Normalized Difference Vegetation Index), 1st PCA (Principal Component Analysis), and 2nd PCA of GLCM-based texture features. All of the bands are in 0.5 m resolution. The NDVI measures the difference between the red band, which vegetation absorbs, and the near-infrared band, which vegetation strongly reflects to quantify the vegetation. This band is expected to improve the sensitivity of the model to the vegetation. The PCA bands are obtained from the texture analysis as described below.

Texture provides information about colors or intensities of image in spatial arrangement. In this study, the texture is used to enrich the original band of Pleiades imagery by exploring the intensities. There are three techniques to define the texture. The first is the modelling technique which involves constructing models to specify textures. The second is the structural technique which texture is treated as a set of primitive texels in some regular or repeated relationship. The third is the statistical technique in which texture is treated as a quantitative measure of the arrangement of intensities in a region and this set of measurements is called a feature vector. This study applies statistical techniques as an approach to get texture information from the images. Haralick et al. (1973) introduced a number of methods for extracting textural information from Gray level co-occurrence matrices (GLCM), these methods are also known as spatial co-occurrence matrices. GLCM is a statistical technique to get texture. Five GLCM formula is used: homogeneity, dissimilarity, contrast, entropy, and ASM (Angular Second Moment). The GLCM matrix dimension should be determined to ensure the best result. Semi-variance analysis is carried out to determine the GLCM matrix dimension.

The principal component analysis (PCA) is based on the fact that adjacent bands of the texture analysis result are highly correlated and frequently transmit nearly identical information about the feature. The analysis is used to alter the original data in order to reduce the correlation between the bands. The optimum linear combination of the original bands accounting for the variance of pixel values in an image is discovered during the procedure. PCA is carried out after the texture analysis. The result of GLCM is treated as input to the PCA process. This PCA process extracts texture information from the 5 GLCM formula, which are: homogeneity, dissimilarity, contrast, entropy, and Angular Second Moment into 2 PCA bands.

2.3 Deep Learning Classification Model

This study used one of the DCNN architectures, U-net, to train and classify Pleiades imagery. The U-net design is made up of two major components: encoder and decoder (Figure 2). Restnet34 is utilized as a backbone in the Encoder section, where it applies convolutional blocks and is followed by a maxpool down sampling to encode the input image into feature representations at various levels. The Decoder section uses up-sampling and convolution techniques to project semantic information from feature representation into high-resolution pixel space for dense categorization.

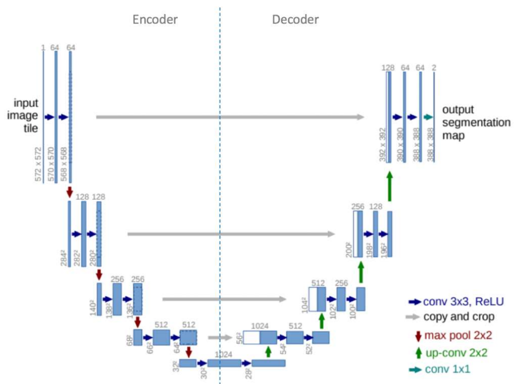


Figure 2. The U-net architecture.

The training samples for this model consist of 1.440 tiles of Pleiades imagery in the City of Mataram. Each tile has dimensions of 512x512 pixels. The Pleiades imagery consists of 7 bands which are Red, Green, Blue, NIR (Near Infrared), NDVI (Normalized Difference Vegetation Index), 1st PCA, and 2nd PCA. The model is trained to classify 6 classes of land cover which are bare land, building, plantation, agricultural land, road, and water. The epoch for the model training is 200 and the learning rate is set automatically.

After model training is done, the model is used to classify land cover at the test area. The test area covers 9.18 km x 6.90 km in the city of Mataram. The location of the test area can be seen in Figure 3. An evaluation is carried out after the deep learning classification procedures are done to figure out the accuracy of the classification result.

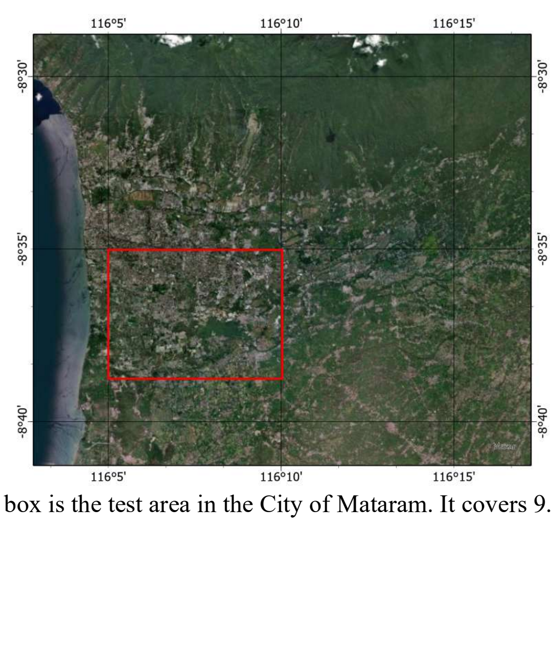


Figure 3. The red box is the test area in the City of Mataram. It covers 9.18km x 6.90km.

2.4 Post-classification Enhancement

Morphological image processing refers to image processing techniques concerned with the shape (or image morphology) of image features. This process is commonly used to remove imperfection shapes created during segmentation or classification. Morphological image processing is basically similar to spatial filtering. The structuring element is moved over each pixel in the original image to produce a pixel in the newly processed image. The value of this new pixel is determined by the procedure. Erosion and dilation (Figure 4) are the two most basic image morphological operations. Erosion will decrease the thickness of all raster features in the raster layer. Erosion can split apart joined features and strip away extrusions. Features may be erased if the threshold is set too high. Dilation will increase the thickness of all raster features in the raster layer. Dilation can repair breaks and intrusions in the feature. It can be used to add definitions to features before vectorization.

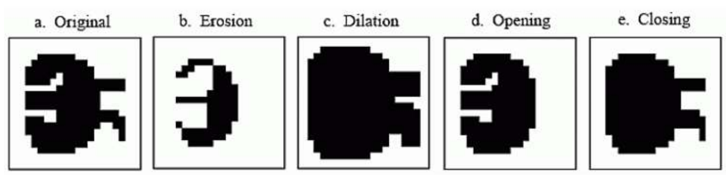


Figure 4. Illustration of erosion, dilation, closing, and opening from the original feature (Smith, S.W, 1999).

Combinations of erosions and dilations can be used to achieve new morphological procedures. The new procedures are closing and opening (Figure 4). Closing is dilation followed by erosion using the same value. It can be used to smooth rough linear features in the image and fill in small gaps between image foreground features. Opening is erosion followed by dilation using the same value. It can be used to erase thin lines or features in the image.

This study uses morphological image processing to eliminate unwanted features or noises that appear after the deep learning classification procedures. After the morphological image processing is done, an evaluation is carried out to figure out the accuracy of the post-classification enhancement result. This accuracy after post-classification enhancement will be compared to the accuracy of the original deep learning classification result.

3. RESULTS

3.1 Texture Analysis and Principal Component Analysis

Semi-variance analysis is carried out to determine the dimension of the GLCM matrix. To apply the semi-variance analysis, this study creates samples from the study area for 6 land cover classes which are bare land, building, plantation, agricultural land, road, and water. Each class of land cover is represented with 2 samples. Figure 5 is an example of a semi-variance diagram of the "building".

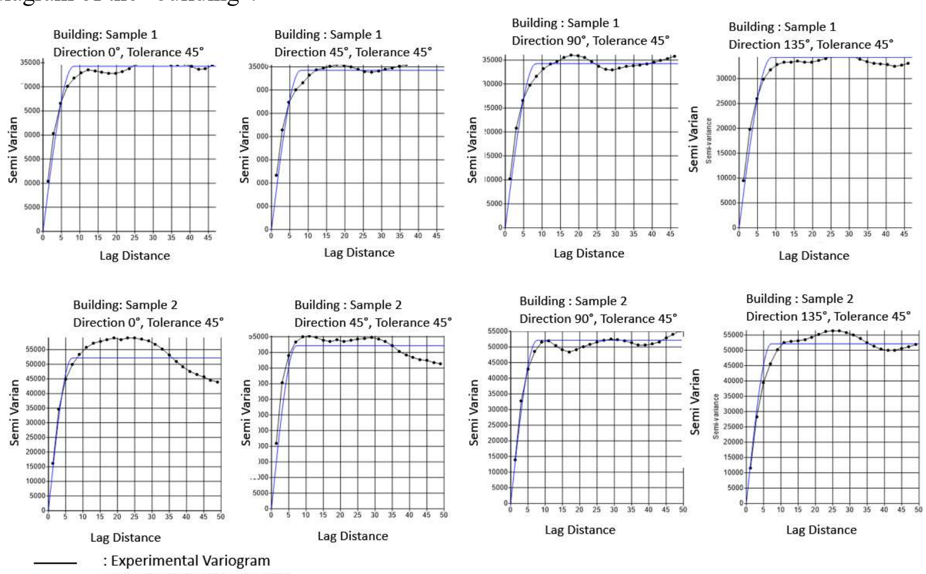


Figure 5. Examples of semi-variance diagrams for "building" sample 1 and sample 2.

Figure 5 describes the variogram of the building in sample 1 and sample 2. The dimension of the GLCM matrix can be indicated from the lag distance where the semi-variance value becomes stable. The black line is the experimental variogram and the blue line is the spherical model variogram. The resume of the semi-variance analysis using the spherical model variogram for all land cover classes can be seen in Table 1 and Table 2. Table 1 is the resume of semi-variance analysis result using a spherical model variogram in 4 directions (0°, 45°, 90°, 135°) for 6 land cover classes. From Table 1, the frequency for the lag distance, which indicates the matrix dimension, can be counted and presented in Table 2. Based on the semi-variance analysis, using the spherical model variogram, the recommended window size for the texture analysis is 5x5. So, the GLCM matrix dimension that used in this study is 5x5. Figure 6 presents examples of the GLCM textures from 5 formulas.

Table 1. Resume of semi-variance analysis result using spherical model variogram in 4 directions for 6 land cover classes.

Samples	**Lag Distance (ranges) in Direction**			
	0°	**45°**	**90°**	**135°**
Road S1	-	-	-	-
Road S2	10	10	15	15
Bare Land S1	5	5	5	5
Bare Land S2	30	20	20	-
Plantation S1	5	5	5	5
Plantation S2	5	5	5	10
Water Body S1	10	10	15	-
Water Body S2	30	30	-	-
Agricultural Land S1	25	-	-	30
Agricultural Land S2	30	30	-	-
Building S1	10	10	10	10
Building S2	5	5	5	5

Table 2. The frequency of matrix size. The matrix size is indicated by the lag distance.

Lag distance (ranges)	**Matrix size**	**Frequency**
5	5x5	15
10	9x9	9
15	15x15	3
20	21x21	2
25	25x25	1
30	31x31	6

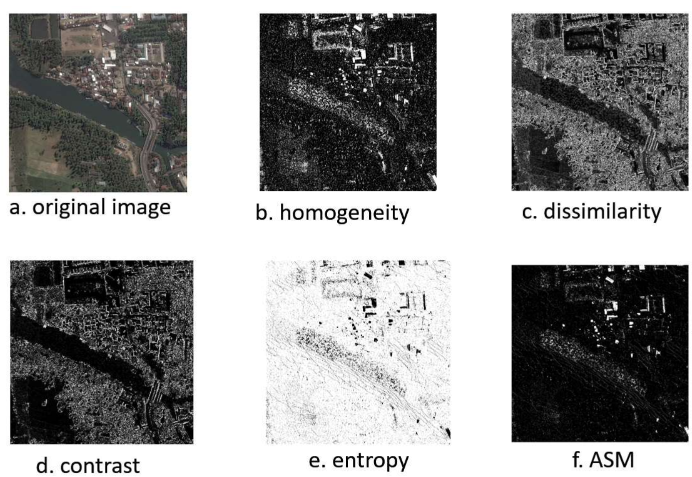
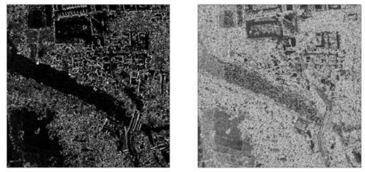


Figure 6. Images b, c, d, e, f are 5 results of the GLCM formula from the original image (a).



a. 1st PCA band b. 2nd PCA band

Figure 7. Example of 1st PCA and 2nd PCA band.

Principal Component Analysis (PCA) is carried out after the texture analysis. The 1st PCA and the 2nd PCA are taken. The example of 1st PCA and 2nd PCA can be seen at Figure 7. To figure out the best choice between the original texture band or the PCA band that will be included in the input of deep learning classification model, an experiment was conducted using a small part of the training set. The first training samples include 5 original texture bands, and the second training samples include 2 PCA bands. The result can be seen in Table 3 and Table 4.

Table 3. Precision, Recall, and F1-Score of Model Training including 5 texture bands: homogeneity, dissimilarity, contrast, entropy, ASM.

Land Cover Class	Bare Land	Building	Plantation	Agricultural Land	Road	Water
Precision	0.900678	0.868628	0.968492	0.454444	0.722694	0.988634
Recall	0.886911	0.870867	0.964743	0.474912	0.580912	0.988579
F1-Score	0.892131	0.869700	0.966595	0.462889	0.606345	0.988601

Table 4. Precision, Recall, and F1-Score of Model Training including 1st PCA and 2nd PCA.

Land Cover Class	Bare Land	Building	Plantation	Agricultural Land	Road	Water
Precision	0.928401	0.876124	0.970865	0.462578	0.710520	0.989666
Recall	0.874431	0.895280	0.970025	0.483561	0.681452	0.990206
F1-Score	0.897240	0.885482	0.970432	0.472089	0.693805	0.989933

Precisions of bare land, building, plantation, agricultural land, and water of model training that includes the PCA bands are better than model training that includes texture bands. Only the precision of the road does the opposite. The maximum increase of precision occurs in the precision of bare land which is 2.8%. Recall of building, plantation, agricultural land, road, and water of model training that includes the PCA bands are better than model training that includes texture bands. Only the recall of bare land does the opposite. The maximum increase of recall occurs in the recall of road which is 10.0%. F1-Score of all land cover classes of model training that includes the PCA bands are better than model training that includes texture bands. The maximum increase of F1-Score occurs in the F1-Score of road which is 8.7%.

3.2 Deep Learning Classification Result

To evaluate the deep learning classification result, the confusion matrix is carried out. The ground truth data is the land cover data obtained from visual interpretation of Pleiades imagery guided with the topographic map at a scale of 1:5.000. The test area covers approximately 63.3 Km^2 with more than 240 million pixels of ground truth data. The compositions of ground truth data are 24.57% of bare land, 21.95 of building, 19.55 of plantation, and 33.93% of agricultural land. The evaluation sample points are 10%, which is more than 24 million random sample points. It means that the sample point spreads out every 1-5 meters in the test area. The number of sample points for each class of land cover can be seen in Table 5. The evaluation is carried out in 4 land cover classes: bare land, building, plantation, and agricultural land.

Table 5. The number of random sample points.

No	Class	Number of sample points	Percentage
1	Bare land	6,280,537	25.47
2	Building	6,180,976	25.07
3	Plantation	6,026,986	24.44
4	Agricultural land	6,167,253	25.01
Total		24,655,752	100.00

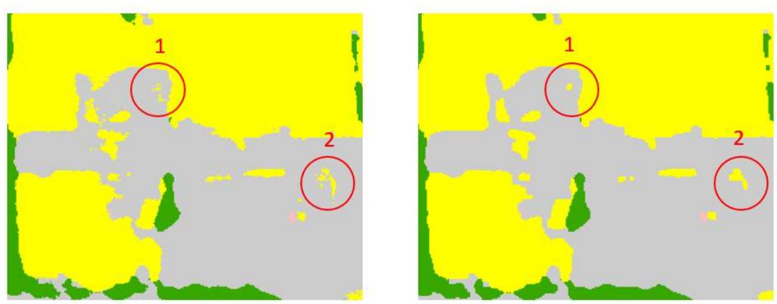
The result of the evaluation can be seen in the confusion matrix in Table 6. The overall accuracy of the deep learning classification result at the test area is 0.84, and the Kappa is 0.79. The producer accuracy and the user accuracy values range from 0.79 to 0.95 as seen in Table 6.

Table 6. Confusion matrix of the deep learning classification result.

Land Cover Class		GROUNDTRUTH				Total	User Accuracy	Kappa
		BARE LAND	BUILDING	PLANTATION	AGRICULTURAL LAND			
PREDICTION	BARE LAND	**22.37**	3.59	0.94	0.59	27.50	0.81	
	BUILDING	3.67	**19.88**	0.73	0.06	24.33	0.82	
	PLANTATION	1.04	1.45	**19.37**	0.56	22.42	0.86	
	AGRICULTURAL LAND	1.07	0.16	1.65	**22.87**	25.75	0.89	
Total		28.15	25.08	22.69	24.08	**100.00**	-	
Producer Accuracy		0.79	0.79	0.85	0.95	-	**0.84**	
Kappa								**0.79**

3.3 Post-Classification Enhancement Result

Post-classification enhancement is carried out by implementing morphological image processing. This process removes the unwanted features created during the deep learning classification process. Unwanted features could be a small group of pixels, a slight gap between land cover classes, or slight extrusions. Figure 8 is an example of the morphological image processing at the test area. In Figure 8, for example, some small-scattered groups of pixels in the red circles 1 and 2 transformed into a bigger group creating a more acceptable feature shape.



a. Original raster b. Result of morphological image processing

Figure 8. Example of morphological image processing.

Land Cover Class		GROUNDTRUTH				Total	User Accuracy	Kappa
		BARE LAND	BUILDING	PLANTATION	AGRICULTURAL LAND			
PREDICTION	BARE LAND	**22.69**	3.08	0.75	0.48	27.00	0.84	
	BUILDING	4.02	**19.84**	0.85	0.06	24.78	0.80	
	PLANTATION	0.91	1.32	**19.71**	0.43	22.36	0.88	
	AGRICULTURAL LAND	0.98	0.12	1.12	**23.64**	25.86	0.91	
Total		28.60	24.36	22.43	24.61	**100.00**	-	
Producer Accuracy		0.79	0.81	0.88	0.96	-	**0.86**	
Kappa								**0.81**

The result of the morphological image processing is then evaluated using a confusion matrix. The confusion matrix of the new land cover obtained from morphological image processing can be seen in Table 7. The overall accuracy of the new land cover is 0.86 and the Kappa is 0.81. The results show that post-classification enhancement using morphological image processing can increase the accuracy of classification results using deep learning-based approaches by approximately 2%. The original Overall Accuracy and Kappa of the deep learning classification results were 0.84 and 0.79. After the post-classification enhancement process, the Overall Accuracy and Kappa values increase to 0.86 and 0.81.

4. CONCLUSION AND FUTURE WORK

The preliminary result shows that the Principal Component Analysis (PCA) can be more effective and efficient than the original data. In this case, the 1st and 2nd bands of PCA deliver more accurate results compared to the original 5 texture bands (homogeneity, dissimilarity, contrast, entropy, and ASM). Less number bands but deliver better accuracy.

The deep learning classification process with U-net architecture promises acceptable accuracy for land cover classification using very high-resolution satellite imagery. The Overall accuracy of the deep learning land cover classification is 0.84 and the Kappa is 0.79.

The preliminary research results show that post-classification enhancement using morphological image processing can increase the accuracy of classification results using deep learning-based approaches by approximately 2%. The original Overall Accuracy and Kappa of the deep learning classification results are 0.84 and 0.79. After the post-classification enhancement process, the Overall Accuracy and Kappa values increased to 0.86 and 0.81.

For future work, it is important to explore another approach to improve the accuracy of deep learning classification results. The improvement approach can be done in the pre-classification process, during the classification, or post-classification process.

ACKNOWLEDGEMENTS

This research was supported in part by the Ministry of Interior, Taiwan under grant number 112PL024A. The Pleiades imagery used in this research was provided by Badan Informasi Geospasial (BIG), Indonesia.

REFERENCES

Hakim, Y. F., Tsai, F., 2023. Land Cover Classification Using Deep Learning to Support Large-Scale Topographic Map Production, Proceeding of International Symposium on Remote Sensing (ISRS).

Haralick, R. M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern*, *4*, 610–621

Kupidura, P., 2019. The Comparison of Different Methods of Texture Analysis for Their Efficacy for Land Use Classification in Satellite Imagery. Remote Sensing, Volume 11 (10).

Ning, H., Li, Z., Wang, C., Yang, L., 2020. Choosing an Appropriate Training Set Size when Using Existing Data to Train Neural Networks for Land Cover Segmentation. Annals of GIS, Volume 26, Page 329-342.

Rodarmel, C., Shan, J., 2002. Principal Component Analysis for Hyperspectral Image Classification. Surveying and Land Information Systems.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597.

Smith, S.W, 1999. The Scientist and Engineer's Guide to Digital Signal Processing, 2nd edition. California Technical Publishing

Tong, X. Y., Xia, G. S., Lu, Q., Shen, H., Li, S., You, S., Zhang, L., 2020. Land-cover Classification with High-Resolution Remote Sensing Images Using Transferable Deep Models. Remote Sensing of Environment, Volume 237, 111322.