

## WorldView-2 Satellite Image Classification using U-Net Deep Learning Model

Ilyas<sup>1,2</sup>, Lalu Muhamad Jaelani<sup>1\*</sup>, Muhammad Aldila Syariz<sup>1</sup> and Husnul Hidayat<sup>1</sup>

<sup>1</sup>Dept. of Geomatics Engineering, Institut Teknologi Sepuluh Nopember (ITS), Kampus ITS, Surabaya, Indonesia

<sup>2</sup>Dept. of Geomatics Engineering, Institut Teknologi Sumatera (ITERA), Lampung Selatan, Indonesia

\* Corresponding author e-mail: [lmjaelani@geodesy.its.ac.id](mailto:lmjaelani@geodesy.its.ac.id)

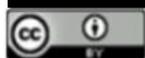
Received: July 08, 2021

Accepted: September 08, 2021

Published: September 08, 2021

Copyright © 2021 by author(s) and  
Scientific Research Publishing Inc.

Open Access



### Abstract

Land cover maps are important documents for local governments to perform urban planning and management. A field survey using measuring instruments can produce an accurate land cover map. However, this method is time-consuming, expensive, and labor-intensive. A number of researchers have proposed using remote sensing, which generates land cover maps using an optical satellite image with various statistical classification procedures. Recently, artificial intelligence (AI) technology, such as deep learning, has been used in multiple fields, including satellite image classification, with satisfactory results. In this study, a WorldView-2 image of Terangun in Aceh Province, which was acquired on Aug 2, 2016, was classified using a commonly used deep-learning-based classification, namely, U-net. There were eight classes used in the experiment: building, road, open land (such as green open space, bare land, grass, or low vegetation), river, farm, field, aquaculture pond, and garden. For comparison, three classification methods: maximum-likelihood, random forest, and support vector machine, were performed compared to U-Net. A land cover map provided by the government was used as a reference to evaluate the accuracy of land cover maps generated using two classification methods. The results with 100 randomly selected pixels revealed that U-Net was able to obtain a 72% and 0.585 for overall and kappa accuracy, respectively; whereas, overall accuracy and kappa accuracy for the maximum likelihood, random forest and support vector machine methods were 49% and 0.148; 59% and 0.392; and 67% and 0.511; respectively. Therefore, U-Net outperformed those three of classification methods in classifying the image.

**Keywords:** artificial intelligence, deep learning, landcover, sustainability, U-net, WorldView-2

### 1. INTRODUCTION

An accurate land cover map is a critical asset for land management, environmental development, natural resource estimation, and other applications on different geographical scales ranging from local to regional (Belward and Skøien, 2015; Gómez et al., 2016; Xing et al., 2017). A land cover map is considered accurate if it is spatially corrected relative to the actual condition of land cover conditions. Field surveys can provide satisfactory results; however, they require a large number of human resources and are time-consuming, thus rendering them impractical. Recently, optical satellite images have been used to alleviate the burden of field surveys using a statistical classification technique. (Mohajane et al., 2018) used the Landsat image series from the multispectral scanner, enhanced thematic mapper plus, thematic mapper, and operational land imager (OLI) sensors to generate time-series land cover maps in the Middle Atlas, Morocco, between 1987 and 2017. In another study, (Nguyen et al., 2020) used Sentinel-2 images for land classification over

Dak Nong Province, Vietnam. However, the spatial resolution of the optical images was medium (30–10 m), whereas, for spatially accurate mapping, higher-resolution images are necessary, such as those from Quickbird (2.4 m) and IKONOS (3.2 m). Classification techniques rely on the variation of radiance or reflectance values at different wavelengths of a pixel, known as pixel-based classification, or a group of pixels, known as object-based classification (Memarian et al., 2013). The maximum likelihood classifier is extensively used and considered the most accurate method in pixel-based classification in which certain pixels are classified into a corresponding class in which their spectral shapes or signatures have similar patterns (Sun et al., 2013). A limitation of the maximum likelihood classifier, however, is that the “salt-and-pepper problem” may arise because the spectral shape of an individual pixel does not represent the characteristics of the surface object (Stoian et al., 2019). To address this problem, the information of neighboring pixels should be

considered during classification, which is known as convolutional processing. Therefore, with the rapid development of deep learning, many algorithms based on convolutional processing for classification have been utilized such as U-Net (Kohl et al., 2018; Ronneberger et al., 2015; Xu et al., 2020), mask regional convolutional neural network (R-CNN)(He et al., 2020), and feature pyramid network (FPN) (Zhang et al., 2018). From all algorithms, U-net is one of the most well-recognized image segmentation algorithms and implemented on ArcGIS Pro (Esri, n.d.). However, U-Net has not been extensively implemented in land use or land cover classification using high-resolution images. Therefore, this study proposes a U-Net model for land cover map generation in Terangun, Aceh Province. The proposed model contains 19 convolutional layers, similar to the number of layers in the original U-Net proposed by (Ronneberger et al., 2015). However, the number of filters in each layer is smaller than the original number of filters to reduce the processing load. Moreover, a WorldView-2 image with a multispectral spatial resolution of 1.84 m (at nadir) that was acquired on Aug 2, 2016, was used over the study area. For training the proposed model, ground truth data obtained from Terangun's land cover were used. Moreover, a land cover map generated using the maximum-likelihood method as well as two well-known machine learning techniques, namely random forest and support vector machine, was produced and compared to the map generated by the proposed model.

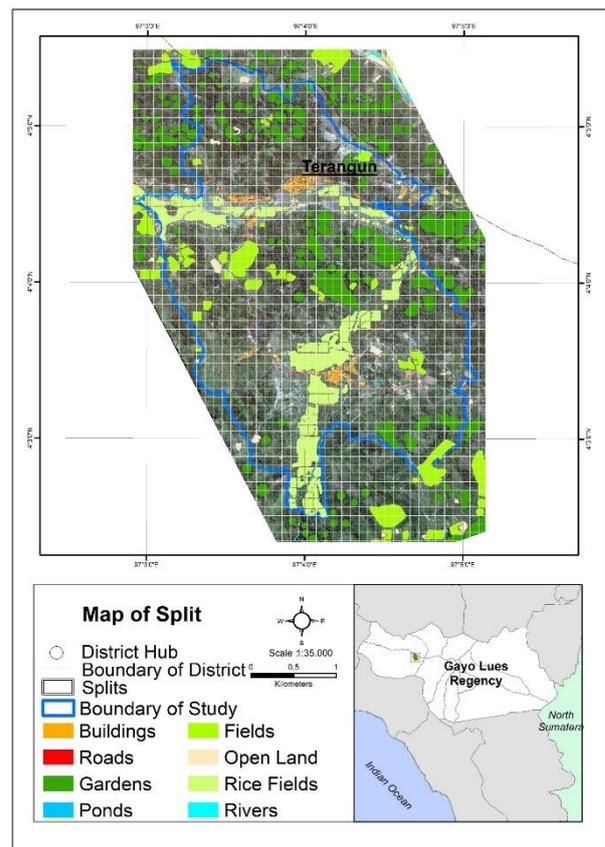
## 2. METHODOLOGY

### 2.1 Study Area

The study area was in Terangun, Aceh Province, Indonesia, located at  $04^{\circ}02'29''-04^{\circ}05'25''$  N and  $97^{\circ}02'56''-97^{\circ}05'06''$  E with a total area of 1211 ha (12.11 km<sup>2</sup>). Based on the local government's regulations on urban planning for 2012–2032, Terangun was determined to be a priority city for development because of its strategic location, sharing a boundary with the Aceh Barat Daya Region and Taman Nasional Gunung Leuser (Lues, 2013). Terangun was projected to be a transit city that could generate increased income for the government. Therefore, the selection of Terangun as the study area in this study was important to support the government's planning for the years 2012–2032. Figure 1 shows a map of Terangun released by the local government.

### 2.2 WorldView-2 Satellite Image

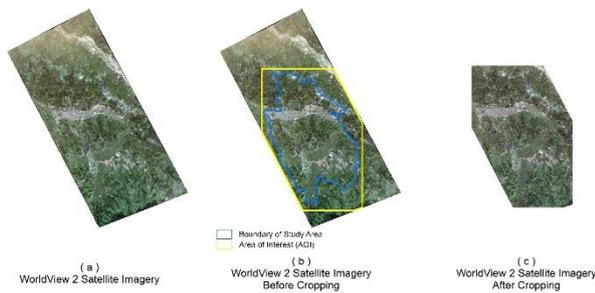
WorldView-2 satellite images are high-resolution satellite images that have panchromatic and eight multispectral bands with four standard colors ranging from blue to near-infrared wavelengths. The spatial resolution for the panchromatic band: 0.46 m (at nadir), 0.52 m (at 20° Off-Nadir); and for multispectral bands: 1.84 m (at nadir), 2.4 m (at 20° off-nadir)(Digital Globe, 2013, 2010). In this research, the WorldView-2 data is at near nadir (incident angle of 7.2°) that already pansharpened by vendor to produced multispectral resolution of 0.5 m. The multispectral and panchromatic band. With these specifications, the WorldView-2 image is sufficient and suitable for land cover mapping with a scale of 1:5000 (0.65 m resolution to fulfill the governmental regulation (Badan Informasi Geospasial, 2020)).



**Figure 1.** Map of Terangun City  
(Sources : Gayo Lues PUPR Agency)

One WorldView-2 satellite image covers the entire region of Terangun (acquired on August 2, 2016) was obtained and then subsetting following the administrative border of Terangun (Figure 2). Moreover, these data were used by the local government to produce Terangun's land cover map based on visual interpretation and delineation. This land cover map was used as reference data in this study. The satellite image was classified using semantic segmentation (pixel-based classification) using the U-Net method that was implemented on Esri ArcGIS Pro. Because the input to U-Net was a patch image with a size of  $128 \times 128 \times 4$ , the WorldView-2 image was preprocessed to generate several patch images. The generation of patch images was performed by selecting a pixel belonging to the Terangun region and expanding to all directions such that the spatial coverage became  $128 \times 128$  pixels. Then, all spectral information was extracted, which led to a patch image with a size of  $128 \times 128 \times 4$ . Once the patch image was obtained, a stride of 64 pixels, horizontally and vertically, was applied, which indicates that the next pixel had a distance of 64 pixels from the previous pixel. This procedure was repeated until the patch images belonging to Terangun were obtained. Figure 3 shows the patch images in which each patch is represented as a rectangle. More than 90 million pixels corresponded to Terangun, and 18,781 of which were used for the training dataset. Details of the training dataset are presented in Table 1. Furthermore, to increase the number of training data, image augmentation was performed using a rotation and zooming procedure. For rotation, new images were obtained by rotating each training image in four directions:  $-180^{\circ}$ ,  $-90^{\circ}$ ,  $90^{\circ}$ , and  $180^{\circ}$ . As for magnification, new images were obtained by magnification in on the training image

continuously on a scale of 0.005–0.450. In total, there were 11,893 training data that were further split into two sets, where one set containing 10,703 total samples was used to train the model, while the remaining 1,190 data were used for validation.



**Figure 2.** Original, (ii) boundary vector of Terangun City, and (iii) subsetted images

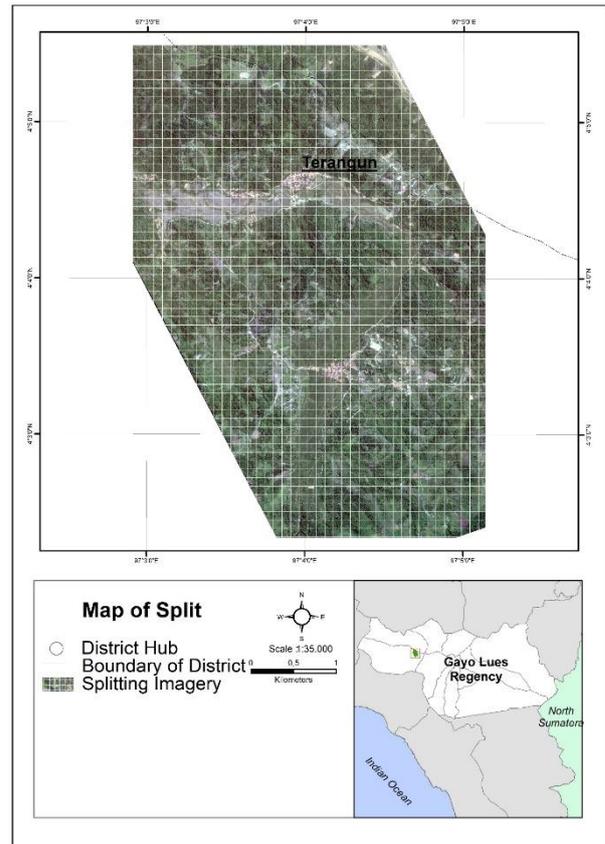
**Table 1.** Detail of training dataset

No	Landcover Class	# of polygon samples	# of pixel samples
1	Buildings	981	507,155
2	Roads	93	204,488
3	Open Land	40	376,617
4	Rivers	71	203,716
5	Rice Fields	130	4,259,372
6	Fields	68	3,719,385
7	Ponds	41	74,964
8	Gardens	153	7,574,063
<b>Total</b>		<b>1577</b>	<b>16,919,760</b>

Since the input to the U-Net is a patch image with a size of  $128 \times 128 \times 4$ , the images were then pre-processed in order to generate a number of patch images. The generation of patch image was by selecting a pixel that belongs to the Terangun City region and expanding to all directions such that the spatial coverage becomes  $128 \times 128$ . Then, all the spectral information of that spatial coverage was extracted which leads to a patch image with a size of  $128 \times 128 \times 4$ . Once a patch image was obtained, a stride of 64 pixels, horizontally and vertically, was applied, which means that the next pixel to conduct was the one with a distance of 64 pixels from the previous pixel. This procedure was repeated up until the patch images belong to the Terangun City were obtained. Figure 3 shows the patch images where each patch is represented as a rectangle.

There were more than 90 million pixels inside the Terangun City, and more than 1.5 thousand of them were used for the training dataset. Details to the training dataset are shown in Table 1. Furthermore, in order to increase the number of training, an image augmentation was conducted by using rotation and zooming procedure. For rotation, new images were obtained by rotating each training image in four directions, those directions were  $-180^\circ$ ,  $-90^\circ$ ,  $90^\circ$ , and  $180^\circ$ . As for zooming, new images were obtained by zooming in the training image continuously from 0.005 to 0.450 zooming scale. In total, there were 11,893 training data which further be split into two sets, where one set with total samples of 10,703 data was used to train the model

while the rest of the data as many as 1,190 data were used for validation.



**Figure 3.** Patch images representation by rectangles with training sample sets

### 2.3 U-Net Design

The U-Net model in this study contained 23 convolutional layers which were 16, 32, 64, and 128 features. Moreover, several max-pooling layers with a  $2 \times 2$  kernel filter and stride of 2 pixels were included in U-Net, and no padding was required. The activation function in all layers was set to the rectified linear unit (ReLU). Figure 4 shows the architecture of the U-Net model used in this study where there were 1,941,385 unknown parameters in total. Furthermore, as the image coverage was sized  $8,252 \times 11,639$  in terms of spatial resolution, training U-Net was time-consuming and required a high-performance computer. To address this challenge, the image was further patched following the procedure in Section 2.2 such that several image patches with a spatial size of  $128 \times 128$  were obtained. Therefore, input to the U-Net was a patch image with a size of  $128 \times 128 \times 4$ , while the output was a single image with a similar spatial size as that of the input where each pixel had the classification information. For hyperparameter adjustment in the training process, dropout was used to prevent the overfitting problem, and was set to 0.1 in each layer, which indicates that only 10% of the total unknown parameters in each layer were used to train the U-Net model. The loss function used in this study was weighted cross-entropy with a maximum of 50 epochs. To evaluate the capability of the model, the loss and mean intersection of union (IoU) value for the training and validation sample sets were measured, where a lower loss value and higher mean IoU indicated higher accuracy of the model and vice versa.

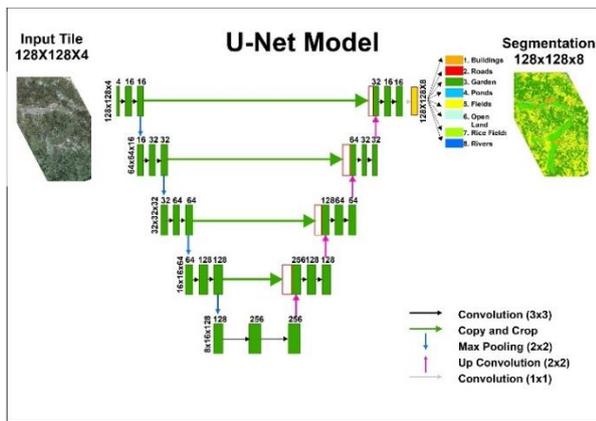


Figure 4. U-Net Architecture Model

## 2.4 Accuracy Assessment

Accuracy assessment was required to determine the accuracy of the classification results. A confusion matrix was used for assessment to determine the overall and kappa accuracy of classification by the U-Net and supervised maximum likelihood classification models, using the land cover map from the government for comparison. This assessment utilized 100 randomly distributed pixels, which were determined based on the percentage of area per land cover. Land cover with a smaller total area had fewer testing points. Moreover, land cover with a larger total area had more testing points.

## 3. RESULT AND DISCUSSION

### 3.1 Classification with U-Net

As mentioned in Section 2.3, the loss and mean IoU for the training and validation sample sets were measured to analyze the capability of U-Net in classification. For the training set, as shown in Figure 5, the loss and mean IoU at the first epoch were 0.8065 and 0.4768, respectively. In the 50th epoch, the loss decreased to 0.2732, while the mean IoU increased to 0.6592. This indicates that a longer duration of training led to higher accuracy.

For loss validation, the loss continued to decrease from 0.6387 to 0.2351. This reduction shows increasing accuracy of the data used for validation. Moreover, the initial mean IoU was only 0.4957, which then slowly increased with an increasing epoch. In the last epoch, the mean IoU was 0.6594. Figure 6 shows the loss and mean IoU for the validation set in the training process. Moreover, Table 2 summarizes the loss results for the validation set.

In the next stage, the trained U-Net model was used to classify the entire image. After classifying land cover with a deep learning model, a land cover map of Terangun was generated. Because the land cover classification results were still in raster format, conversion to vector format was required for further processing. The land cover map is shown in Figure 7.

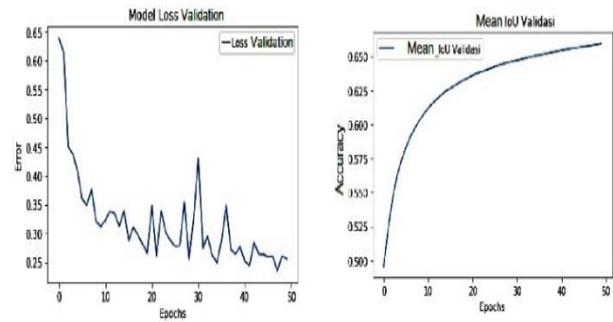


Figure 5. Loss value (Left) and Mean IoU (Right) for training sets at each epoch in the training process

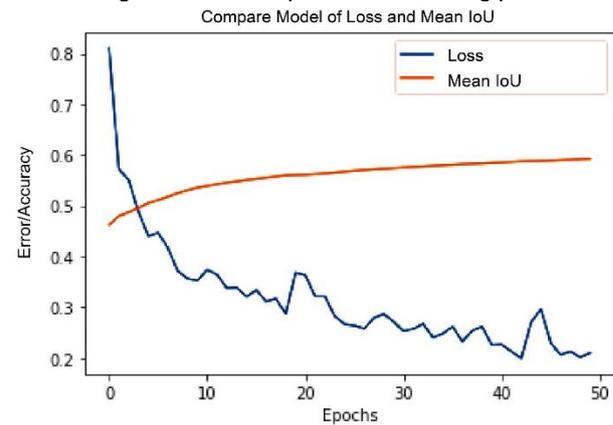


Figure 6. Loss value (blue) and Mean IoU (red) for validation sets at each epoch in training process

As shown in Table 3, land cover classes categorized as large areas were gardens (469.12 ha), fields (394.11 ha), and rice fields (264.63 ha). Other land cover classes such as open land (37.29 ha), buildings (26.7 ha), roads (11.42 ha), rivers (7.65 ha), and ponds (0.85 ha) were categorized as small areas with a total area under 50 ha.

Table 2. Monitoring of loss value for validation set in model training process

#	Loss value (from)	Loss value (to)	#	Loss value (from)	Loss value (to)
1	undifined	0.63870	10	0.32060	0.31135
2	0.63870	0.61348	16	0.31135	0.28729
3	0.61348	0.45005	19	0.28729	0.28105
4	0.45005	0.43704	20	0.28105	0.26562
5	0.43704	0.40782	22	0.26562	0.26140
6	0.40782	0.35881	29	0.26140	0.25441
7	0.35881	0.34851	35	0.25441	0.24858
9	0.34851	0.32060	42	0.24858	0.24425

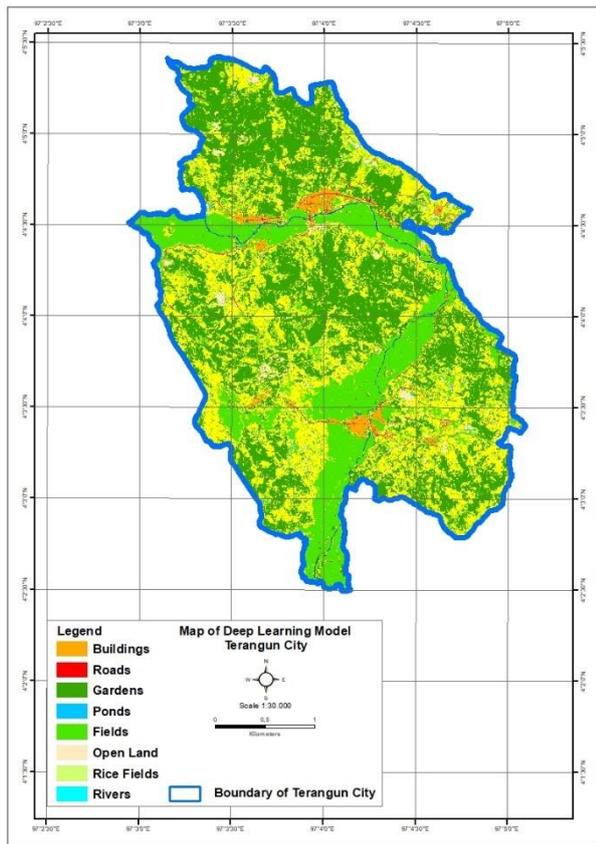


Figure 7. Classification map using U-Net model

Table 3. Area of each land cover using U-Net

No	Classes	Area (Ha)
1	Buildings	26.7
2	Roads	11.42
3	Open Land	37.29
4	Rivers	7.65
5	Rice Fields	264.63
6	Fields	394.11
7	Ponds	0.85
8	Gardens	469.12
	Total	1211.77

The results of the accuracy assessment show that the land cover of rivers, rice fields and ponds could be classified perfectly by deep learning models, whereas the building, roads, open lands, fields and garden classification results were not satisfactory. These fields were correctly classified with 73.68 % conformity, while the remaining 26.32% was classified incorrectly. These fields cover were classified as open lands and rice fields. This discrepancy was mostly caused by the fields, and open lands that were located next to the rice fields. Moreover, only 50 % of roads were classified as roads, other was classified as open lands.

Table 4. Confusion matrix of U-Net Deep Learning Classification.

Landcover	Buildings	Roads	Open Land	Rivers	Rice Fields	Fields	Ponds	Gardens	Total of raw	User's Accuracy
Buildings	2	0	0	0	0	0	0	0	2	100
Roads	0	1	0	0	0	0	0	0	1	100
Open Land	0	1	0	0	0	2	0	0	3	0
Rivers	0	0	0	1	0	0	0	0	1	100
Rice Fields	0	0	0	0	14	3	0	3	20	70
Fields	1	0	0	0	0	14	0	18	33	42.42
Ponds	0	0	0	0	0	0	1	0	1	100
Gardens	0	0	0	0	0	0	0	39	39	100
Total of Column	3	2	0	1	14	19	1	60	100	
Producer's Accuracy	66.67	50	0	100	100	73.68	100	65		

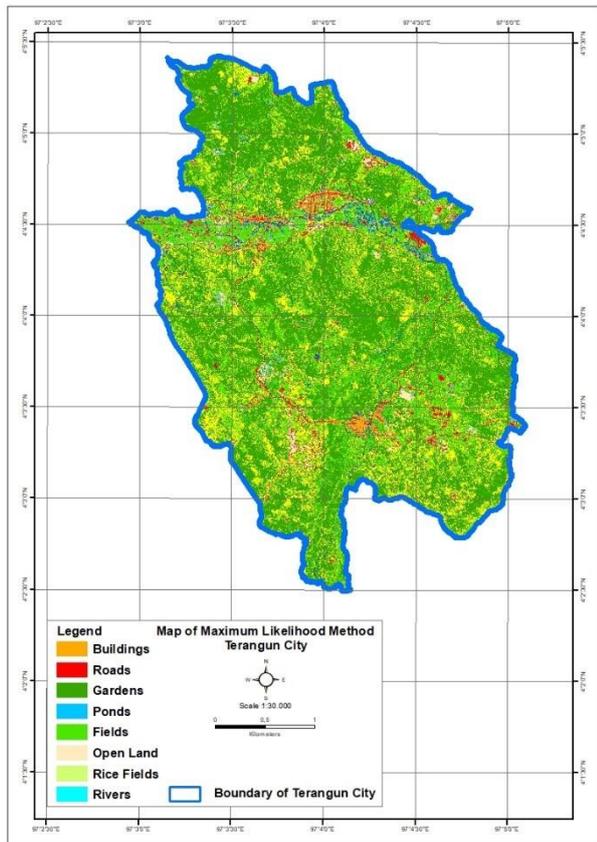
Misclassification of land cover as a open lands accident in 3 location, while for rice fields and fields occurred in 6 and 19 locations, respectively. Most errors occurred for the fields were attributed to similarities in color among the class of fields and gardens. Classification of the buildings, roads, rivers, pond and garden demonstrated 100% conformity.

### 3.2 Comparison with Maximum-Likelihood, Random Forest, and Support Vector Machine Classification

To compare the U-Net model with a familiar classifier, maximum likelihood (ML), random forest (RF), and support vector machine (SVM) (Gislason et al., 2006; Hogg, 1979;

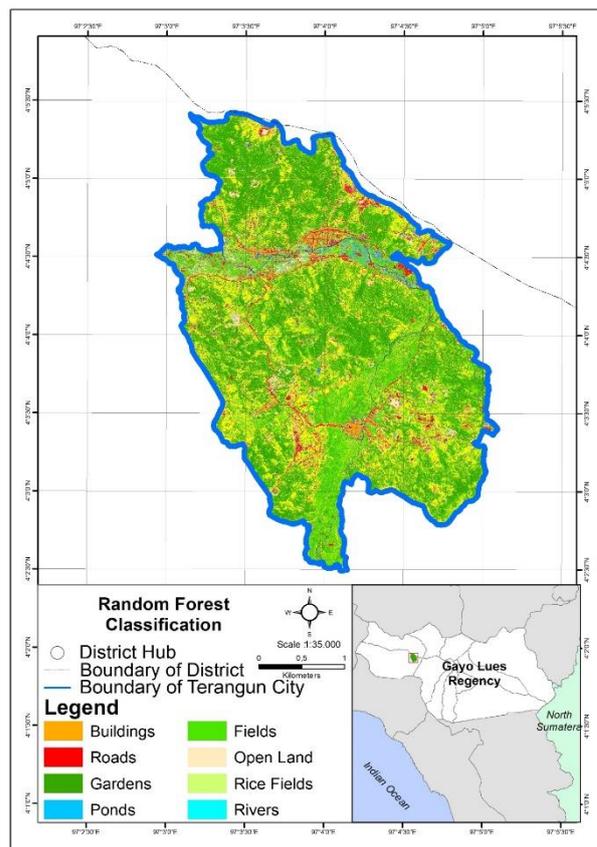
Huang et al., 2002; Pal and Mather, 2005) classification was performed using similar training sets. The results of this classification method were in raster data format; therefore, conversion to vector data was required, as performed for the deep learning model classification data. The results of classification using the ML, RF and SVM classification method are shown in Figure 8, 9 and 10, respectively. To compare the U-Net model's accuracy and three of classification methods, it was necessary to test the accuracy of them against Terangun's land cover. The assessment used comprises 100 pixel points that were randomly distributed, covering a building of 2 pixels, road of 1 pixel, open land of 3 pixels, river of 1 pixel, rice field of 20 pixels, field of 33 pixels, pond of 1 pixel, and garden of 39 pixels.

In U-Net model, only rivers, rice fields, and ponds classes could be classified appropriately, while buildings, roads, open land, fields, and gardens did not match the classification results. Buildings were incorrectly classified as fields; roads were incorrectly classified as open land; no pixel was classified as open land, field was incorrectly classified as rice fields, and gardens were incorrectly classified as rice fields. The overall accuracy was 72%, while the kappa accuracy was 0.585.

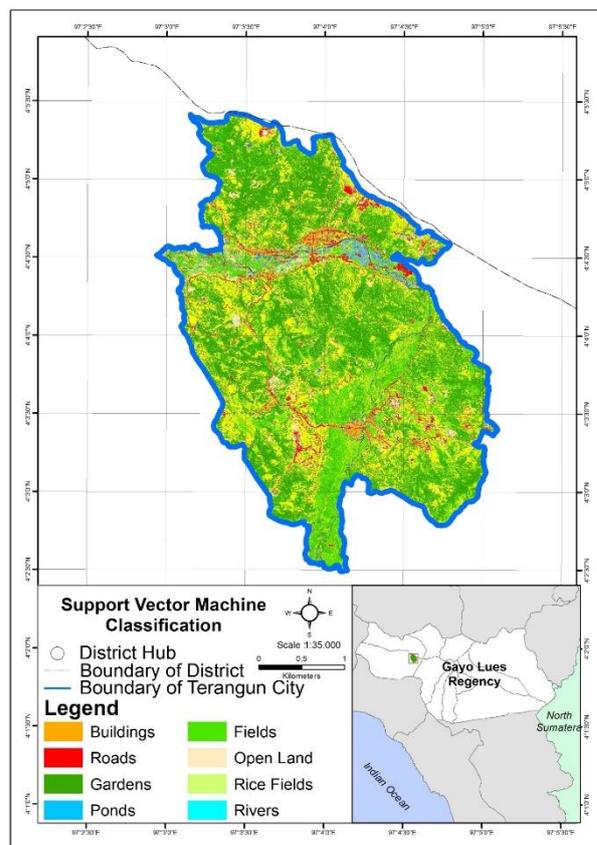


**Figure 8.** Classification map using the maximum likelihood

In supervised maximum likelihood, 100% of the ponds' classification result was correct, similar to U-Net model; water pixel is homogeneous and easy to classify. 59% of the classification results of gardens were valid, while the remaining 41% were incorrectly classified as fields. 55% of rice fields were correctly classified; however, the remaining 45% were incorrectly classified as fields and gardens; this occurred because the color of the adjacent pixels was both green. While other classes could not be classified as its factual land covers. Buildings, roads, open land, and field were characterized as mixed pixels and small objects that might not be easy to classify with this method. From the distribution of these test points, the overall accuracy was 49%, while the kappa accuracy was 0.148. Therefore, in this study, the supervised maximum likelihood classification's accuracy was lower than that of the deep learning model. In random forest classification, gardens, fields, roads, rice fields, and open lands were 71%, 53%, 50 %, 46%, and 33 % correctly classified.



**Figure 9.** Classification map using the random forest



**Figure 10.** Classification map using the vector machine classification

In comparison, rivers and fields could not be classified correctly. This classification method produces overall accuracy of 59% and kappa accuracy of 0.392.

In support machine classification, fields, gardens, rice fields, and open land were classified with 75%, 67%, 67%, and 57% correct classification. Whereas building, roads, rivers, and ponds could not be classified correctly. The overall and kappa accuracy for this method were 67% and 0.511.

The classification results using the U-Net model and three other classification methods indicate that there was a difference in the area of each class of land cover. Maximum likelihood produced the least area difference for open land, rice fields, fields, ponds, and gardens. U-net has the least difference for roads and rivers, whereas support vector machine for buildings.

**Table 5.** Estimated land cover class area using 4 classification methods

No	Classes	Referenced Area (ha)	Estimated Area (ha)			
			ML	U-Net	RF	SVM
1	Buildings	13.100	14.620	26.700	17.165	<b>14.437</b>
2	Roads	19.140	42.450	<b>11.420</b>	61.650	55.067
3	Open Lands	23.820	<b>30.870</b>	37.290	59.088	59.459
4	Rivers	7.520	22.240	<b>7.650</b>	13.513	12.223
5	Rice Fields	189.390	<b>231.740</b>	264.630	243.111	262.832
6	Fields	175.910	<b>236.750</b>	394.110	341.320	353.397
7	Ponds	4.200	<b>6.320</b>	0.850	24.802	17.710
8	Gardens	778.690	<b>626.780</b>	469.120	451.121	436.645
Total		1211.770	1211.770	1211.770	1211.770	1211.770

However, this data was not related to classification results' accuracy because pixel location was not considered. The estimated area of each land cover class was shown in Table 5. The classification results using the U-Net model and three other classification methods indicate that there was a difference in the area of each class of land cover. Maximum likelihood produced the least area difference for open land, rice fields, fields, ponds, and gardens. U-net has the least difference for roads and rivers, whereas support vector machine for buildings. However, this data was not related to classification results' accuracy because pixel location was not considered. The estimated area of each land cover class was shown in Table 5.

#### 4. CONCLUSIONS

This study performed deep learning classification using the U-Net model, which contained 19 convolutional layers and several max-pooling layers for land cover classification using a WorldView-2 image. Because of computational limitations, the image was further patched to a spatial size of 128 × 128 with a stride of 64 pixels. Therefore, the input to the U-Net was a patch image with a size of 128×128×4, while the output was a single image with a similar spatial size as the input, which corresponded to the classification map. There were ~11,000 training data, which were split into two sets: one set with 10,703 data that were used to train the model, 1,190 data that were used for validation, and another 100 pixels for testing. For performance comparison, the maximum likelihood classifier, random forest and support vector machine was used to evaluate the accuracy of the U-Net model. The results indicated that U-Net was superior to them in terms of classification accuracy. U-Net achieved an overall accuracy of 72 % and kappa accuracy of 0.585. Whereas, overall accuracy and kappa accuracy for the maximum likelihood, random forest, and support vector machine classification methods were 49% and 0.148; 59% and 0.392; and 67% and 0.511; respectively.

#### ACKNOWLEDGEMENT

This research was funded by the Indonesian Ministry of Research and Technology/National Agency for Research and Innovation, grant number 1377/PKS/ITS/2020

#### REFERENCES

- Badan Informasi Geospasial, 2020. Technical specifications of the base map for the preparation of the detailed spatial plan (spesifikasi teknis peta dasar untuk penyusunan rencana detail tata ruang) [WWW Document]. URL [http://tataruang.big.go.id/modules/pustaka/Spek\\_Pemeriksaan\\_Peta\\_RDTR/Sumber\\_Data\\_Citra\\_dan\\_Peta\\_Dasar/V3\\_Spek\\_Teknis\\_Data\\_Dasar\\_dan\\_Peta\\_Dasar\\_RDTR.pdf](http://tataruang.big.go.id/modules/pustaka/Spek_Pemeriksaan_Peta_RDTR/Sumber_Data_Citra_dan_Peta_Dasar/V3_Spek_Teknis_Data_Dasar_dan_Peta_Dasar_RDTR.pdf)
- Belward, A.S., Skøien, J.O., 2015. Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS J. Photogramm. Remote Sens.* 103, 115–128. <https://doi.org/10.1016/j.isprsjprs.2014.03.009>
- Digital Globe, 2013. Our Constellation - DigitalGlobe [WWW Document]. URL <https://www.digitalglobe.com/about/our-constellation#satellites&worldview-2> (accessed 10.26.16).
- Digital Globe, 2010. Radiometric Use of WorldView-2 Imagery Technical Note.
- Esri, n.d. How U-net works? | ArcGIS for Developers [WWW Document]. URL <https://developers.arcgis.com/python/guide/how-unet-works/> (accessed 9.7.20).
- Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification, in: *Pattern Recognition Letters*. <https://doi.org/10.1016/j.patrec.2005.08.011>
- Gómez, C., White, J.C., Wulder, M.A., 2016. Optical

- remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogramm. Remote Sens.* <https://doi.org/10.1016/j.isprsjprs.2016.03.008>
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2020. Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* <https://doi.org/10.1109/TPAMI.2018.2844175>
- Hogg, R. V., 1979. Statistical Robustness: One View of Its Use in Applications Today. *Am. Stat.* 33, 108. <https://doi.org/10.2307/2683810>
- Huang, C., Davis, L.S., Townshend, J.R.G., 2002. An assessment of support vector machines for land cover classification. *Int. J. Remote Sens.* <https://doi.org/10.1080/01431160110040323>
- Kohl, S.A.A., Romera-Paredes, B., Meyer, C., De Fauw, J., Ledsam, J.R., Maier-Hein, K.H., Ali Eslami, S.M., Rezende, D.J., Ronneberger, O., 2018. A probabilistic U-net for segmentation of ambiguous images, in: *Advances in Neural Information Processing Systems*.
- Lues, K.G., 2013. Qanun Kabupaten Gayo Lues Nomor 15 Tahun 2013 tentang Rencana Tata Ruang Wilayah Kabupaten Gayo Lues Tahun 2012-2032. Indonesia.
- Memarian, H., Balasundram, S.K., Khosla, R., 2013. Comparison between pixel- and object-based image classification of a tropical landscape using Système Pour l'Observation de la Terre-5 imagery. *J. Appl. Remote Sens.* 7, 073512. <https://doi.org/10.1117/1.jrs.7.073512>
- Mohajane, M., Essahlaoui, A., Oudija, F., El Hafyani, M., Hmadi, A. El, El Ouali, A., Randazzo, G., Teodoro, A.C., 2018. Land Use/Land Cover (LULC) Using Landsat Data Series (MSS, TM, ETM+ and OLI) in Azrou Forest, in the Central Middle Atlas of Morocco. *Environments* 5, 131. <https://doi.org/10.3390/environments5120131>
- Nguyen, H.T.T., Doan, T.M., Tomppo, E., McRoberts, R.E., 2020. Land use/land cover mapping using multitemporal sentinel-2 imagery and four classification methods-A case study from Dak Nong, Vietnam. *Remote Sens.* 12, 1367. <https://doi.org/10.3390/RS12091367>
- Pal, M., Mather, P.M., 2005. Support vector machines for classification in remote sensing. *Int. J. Remote Sens.* <https://doi.org/10.1080/01431160512331314083>
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, pp. 234–241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- Stoian, A., Poulain, V., Inglada, J., Poughon, V., Derksen, D., 2019. Land Cover Maps Production with High Resolution Satellite Image Time Series and Convolutional Neural Networks: Adaptations and Limits for Operational Systems. *Remote Sens.* 11, 1986. <https://doi.org/10.3390/rs11171986>
- Sun, J., Yang, J., Zhang, C., Yun, W., Qu, J., 2013. Automatic remotely sensed image classification in a grid environment based on the maximum likelihood method. *Math. Comput. Model.* 58, 573–581. <https://doi.org/10.1016/j.mcm.2011.10.063>
- Xing, H., Meng, Y., Hou, D., Song, J., Xu, H., 2017. Employing Crowdsourced Geographic Information to Classify Land Cover with Spatial Clustering and Topic Model. *Remote Sens.* 2017, Vol. 9, Page 602 9, 602. <https://doi.org/10.3390/RS9060602>
- Xu, W., Deng, X., Guo, S., Chen, J., Sun, L., Zheng, X., Xiong, Y., Shen, Y., Wang, X., 2020. High-resolution u-net: Preserving image details for cultivated land extraction. *Sensors (Switzerland)* 20, 1–23. <https://doi.org/10.3390/s20154064>
- Zhang, P., Ke, Y., Zhang, Z., Wang, M., Li, P., Zhang, S., 2018. Urban Land Use and Land Cover Classification Using Novel Deep Learning Models Based on High Spatial Resolution Satellite Imagery. *Sensors* 18, 3717. <https://doi.org/10.3390/s18113717>