Proceedings of the 3rd International Conference on Inventive Computing and Informatics (ICICI-2025)
DVD Part Number: CFP25L34-DVD; ISBN: 979-8-3315-3829-3

Land Cover Classification and Mapping of Philippine Satellite Images using RAU-Net

Joseph Jessie S. Oñate
College of Computer Studies/AI Research Center for Community Development
Camarines Sur Polytechnic Colleges
Nabua, Camarines Sur, Philippines
josephjessie97@cspc.edu.ph

***Abstract*—Land Cover maps are one of the crucial descriptors in land use planning, management, and development. However, there is a rapid change in land cover due to urbanization and agricultural land development. In this paper, a Residual Attention U-Net Model was trained to classify and map the Land Cover in the Philippines using Satellite Images. The satellite image was given by DOST DATOS - Remote Sensing and Data Science Help Desk which provides free satellite Images for academic use. The image dataset was labeled with eight (8) land cover classes. Data pre-processing and augmentation were used to improve the image and mask dataset. The trained model was evaluated by segmentation metrics namely, Jaccard index and Dice coefficient. After tedious training, the model achieved a performance of 62% dice coefficient and 45% jaccard index with a focal loss of 16%. This shows that the model can predict the majority of the classes however, there are some classes that are predicted incorrectly which is understandable given the performance of the model. A large-scale satellite image was also fed into the model which automatically produced land cover segmentation.**

***Index Terms*—Land Cover Mapping, Philippine Satellite Images, Large-Scale Image Segmentation, Residual Attention U-Net**

I. Introduction

A key attribute of the Earth's surface is its land cover, and the supply of spatially represented data on land coverage and summary statistics is a requirement for many resource management decisions at the local, national, and international levels [1]. To extract this information, land cover can be developed and mapped. Land cover mapping plays an important role in the assessment of the environment and land use. It is also being used in monitoring environmental development for planning and could serve as an aid in the understanding of land usage issues, and in the formulation of policies and programs necessary for development planning. However, according to [2], the Earth's surface undergoes constant transformation due to a combination of human activities and natural processes. Over the last five decades, the predominant catalyst for changes in land use and land cover has been the expansion of agricultural land. Additionally, rapid urbanization since 1992 has played a significant role in the ongoing shifts in land use and land cover. Today, half of the world's habitable land is devoted to agriculture [3].

Today, Remote sensing is being used as a primary technique in land cover mapping. Remote sensing data plays a crucial role in ascertaining land cover, serving as a fundamental component for various operational mapping and reporting initiatives, in addition to supplying essential data to bolster scientific research endeavors [4]. However, good image classification algorithms are required for correctly and effectively mapping land cover using remote sensing. Moreover, numerous other elements, such as image resolution and atmospheric conditions, may have an impact on the categorization algorithms' efficacy and accuracy [5]. Also, according to [4] the accuracy of land cover maps, such as the classification algorithm and the degree of land cover variability, is strongly interconnected with the calibration (or training) data, predictors, supplementary data used in the classification model, and the manner in which the classification is executed. Enhancing calibration data can be achieved through various methods, such as refining training data by incorporating independent datasets before the mapping process.

In the Philippines, there is an existing land cover map available provided by Geoportal Philippines. However, the recent land cover map available is outdated. Hence, the main goal of the study is to train a model that would classify and map the land cover classes using Residual Attention U-Net in Philippine satellite images. This study could help in updating land cover maps in the Philippines. Generally, this study intends to develop a model using Residual Attention U-Net that classifies the land cover of Philippine Satellite Images. More specifically, this aims to gather and label the Philippine Satellite Images for land cover classification; train a Residual Attention U-Net model that classifies land cover on Philippine satellite images; evaluate the performance of the trained U-Net model in classifying land cover on Philippine satellite images.; and test the trained Residual Attention U-Net model by mapping land cover classes in some Philippine locations using large-scale satellite images. This study will only focus on the Land cover classification of Philippine satellite images. Residual Attention U-Net model will be trained automatically to classify the land cover in satellite images. The developed land cover map will only be based on the trained model. Satellite image of Nabua, Camarines Sur will serve as a test location for the study.

The remaining sections of the paper are structured as follows: In Section 2, we provide an overview of previous research in the field of land cover classification and offer a brief introduction to the Residual Attention U-Net. Section 3 elaborates on the methodologies and techniques employed in this study. Section 4 delves into the experimental results of our model. Finally, in Section 5, we summarize the overarching objectives and outcomes of the study, along with recommendations for potential areas of improvement.

II. Review of Literature and Related Studies

A. Land Cover Classification

Numerous applications, including land resource management, urban planning, precision agriculture, and environmental protection, heavily depend on land-cover classification utilizing remote sensing (RS) imagery [6]. In this context, the task involves categorizing individual pixels within aerial or satellite images into specific land cover classes, presenting a semantic segmentation challenge. These are generally broad categories such as ”forest” or ”field” that characterize the earth’s surface [7]. The major source of thematic data important to GIS analyzes is remote sensing, particularly land use and land cover data. Satellite images by Aerial and Landsat are also widely utilized for assessing the distribution of land cover and updating existing geospatial information [8].

Several studies in land cover classification were conducted such as the study of Helber et al [9], using Sentinel-2 satellite images, patch-based land use, and land cover classification method were presented. In addition, A new scheme was proposed in the study of [6] which is a method for classifying unlabeled HRRS pictures using a deep model derived from a labeled land-cover dataset.

Deep Learning architecture and neural networks have been employed as techniques for land cover classification. Scott et al. utilized Deep Convolutional Neural Networks in their study, achieving accuracies of $97.8 \pm 2.3\%$, $97.6 \pm 2.6\%$, and $98.5 \pm 1.4\%$ when utilizing CaffeNet, GoogLeNet, and ResNet, respectively, with the UC Merced (UCM) Land Use dataset.

B. Land Cover Classification in the Philippines

Different studies in land cover classification already exist however, in the Philippines, there are only limited studies on land cover. In addition, these studies only focused on the specific domains such as *Disaster risk* supported by the study of Santillan et al [10] which focuses on the land cover mapping and modelling of Agusan Marsh in Mindanao, Philippines using maximum likelihood classification of Landsat 5, TM, ETM+ and OLI images and study of Shrestha et al [11] in assessing flash flood at the Marikina River Basin; *Deforestation* presented in the study of [12] which uses two models, the FOREST-SAGE and GEOMOD in simulating national-scale deforestation in the Philippines; and in *Urban Regions* through the integration of crowdsourced geographic data (CGD) with multi-temporal/multi-sensor remote sensing (RS) imagery, one can conduct assessments of land cover change (LCC) [13]; *Agirculture* portrayed in the study of [14] used Google Earth Engine’s quick cloud computing capabilities, the study generated five-year interval land cover maps in a steep agricultural terrain (the Ifugao rice terraces, Philippines) from 1990 to 2020 before examining the land cover transitions and paddy field dynamics.

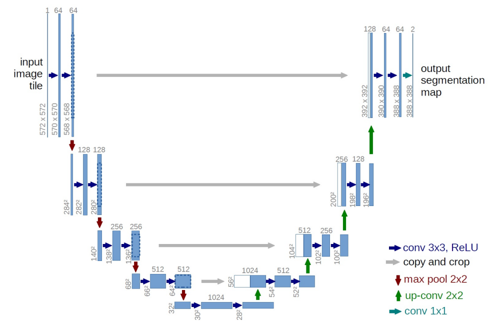


Fig. 1. U-Net Architecture

Given these studies on land cover classification in the Philippines, there is no exploration on the application of deep learning architecture in land cover classification in the Philippines which is also one of the major contributions of the study.

C. Residual Attention U-Net

Ronnerberger et al [15] developed the U-Net shown on Figure 1. The U-Net architecture is structured into two main pathways: the contraction path and the expansion path. In the contracting path, there is a series of standard convolutional operations, whereas the expansion path employs transposed 2D convolutional layers. The contracting path is divided into four blocks, with each block comprising 3x3 convolutional layers and a ReLU activation function, along with a 2x2 max-pooling layer. On the other hand, the expansion path also consists of four blocks. Each block is characterized by a deconvolution layer with a stride of 2, concatenation of information from the contracting path, two 3x3 convolutional layers, and a ReLU activation function. U-Net was originally developed for the segmentation of biomedical images however, in recent years U-Net has been used in other areas such as remote sensing [16].

Researchers have observed progress in the U-Net architecture through various studies, including Ni et al.’s [17] exploration, which introduced the Residual Attention U-Net Architecture. The Residual Attention U-Net combines the U-Net Architecture with both Residual and Attention modules.

One of the problems in deep architecture is that the training is dawdling and a degradation problem may be encounter hence, the Residual unit shown in Figure 2 was introduced by He et al [18] and was used by Zhang et al [19] in Road Extraction to improve the training and address the problem in degradation.

Attention Gate shown in Figure 3 were also integrated with U-Net to draw attention to important features that are transferred through the skip connections. In the attention gate, the input features (x^1) are scaled using the attention coefficients (α) computed in AG. Activations and contextual data from the coarser-scaled gating signal (g), which is used to select spatial regions. Trilinear interpolation is utilized in grid resampling of attention coefficients. This study will use both Residual and Attention modules to better improve the performance of the segmentation.

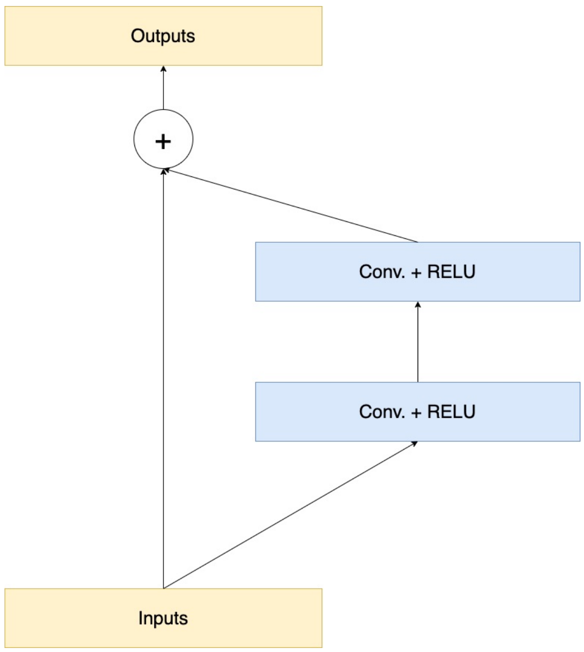


Fig. 2. Residual Block

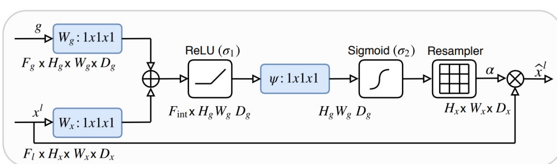


Fig. 3. Attention Gate

III. Methodology

This study followed the process which the phases shown in Figure 4 and were discussed formally below:

A. Data Gathering

Satellite images were gathered on from Philippine Earth Data Resource Observation Center (PEDRO) of the Department of Science and Technology - Advanced Science and Technology Institute. The satellite image data is a PlanetScope 4-band multispectral basic and orthorectified scenes of Nabua, Camarines Sur, Philippines.

B. Data Pre-processing and Labelling

The satellite image is a large-scale image and cannot be fed in the model as a whole hence, satellite image were processed into 256 x 256 x 3 image patches to ensure that it will be accepted by the model architecture for training. Subsequently, after pre-processing, the image data was labeled with features such as buildings, cropland, paved, wetland, ground, water, and trees using Labelbox. After the data was labeled, it was also structured according to the Residual Attention U-Net Architecture's input.

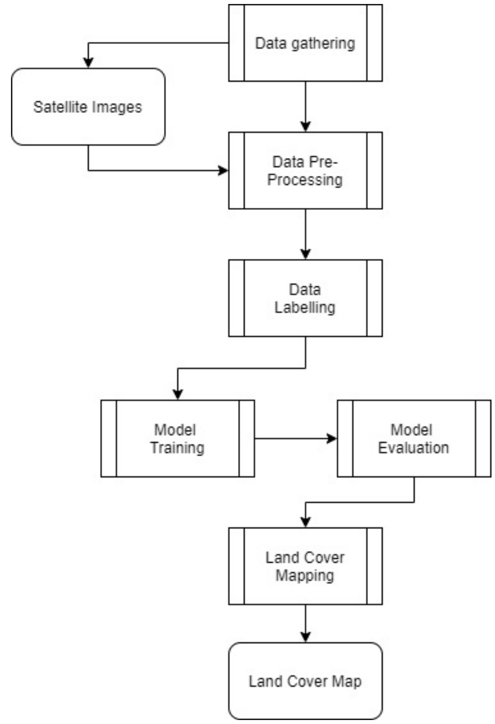


Fig. 4. Process Diagram

C. Model Training

The dataset was experimentally trained using the Residual Attention U-Net Architecture. Supervised learning was used to train the model. The training seeks to discover weights and biases that have minimal losses on average across all classes. Tensorflow will be used to build the architecture of the model and Keras as the backend.

D. Model Evaluation

Intersection over Union (IoU), also referred to as the Jaccard index, stands out as a prevalent evaluation metric in various tasks like segmentation, object detection, and tracking. This metric, originally developed by Paul Jaccard [20], is formally applied by employing the Equation 1 for assessing segmented image predictions. In this equation, A represents the ground truth image, while B signifies the predicted image. In simpler terms, IoU is calculated by dividing the intersection of A and B by the union of A and B.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

Another widely used evaluation metric is the Dice Coefficient, also recognized as the F1 Score, initially introduced by Sørensen [21]. The Dice Coefficient bears similarities to IoU but differs in that it counts the intersection twice, as illustrated in the Equation 2.

$$DC = \frac{2|A \cap B|}{|A| + |B|} \tag{2}$$

For loss, the researchers used Focal Loss which was first introduced by [22] due to highly imbalanced classes on

Proceedings of the 3rd International Conference on Inventive Computing and Informatics (ICICI-2025)
DVD Part Number: CFP25L34-DVD; ISBN: 979-8-3315-3829-3

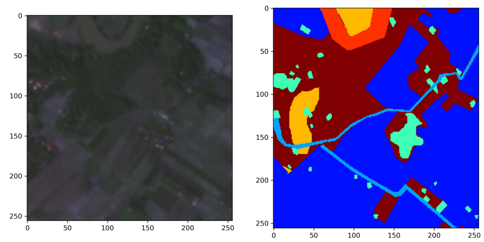


Fig. 5. Sample Image Dataset and Mask

the dataset. Focal loss modifies the cross entropy loss by adding modulating term in order to concentrate learning on challenging negative examples.

E. Land Cover Mapping

Once the model has achieved an acceptable or high accuracy, Land Cover Map will be developed by classifying a large scale satellite image by segmenting the smaller patches and coalesce back again the segmented images after the segmentation as a large scale segmented image.

IV. Results and Discussion

A. Dataset

The dataset is composed of 184 patches with (256 x 256 x 3) size and it was augmented by rotating and flipping the images. After data augmentation, the dataset increased to 1104. Shown on Figure is the sample image dataset with the corresponding mask dataset. The mask is labelled categorically with 8 classes which labelled as 0 is for null, 1 for cropland, 2 for paved, 3 for building, 4 for wetland, 5 for ground, 6 for water, and 7 for trees. The dataset was split into 80% training and 20% for testing.

B. Model Performance

The model was trained in Apple M1 Pro Macbook with 100 epochs with a batch size of 16 and with an elapsed training time of 50 minutes and 2 million plus trainable parameters. During the training, the model starts with loss of 0.3809, Dice coefficient of 0.1794 and jaccard index of 0.0988 and it ended with a loss of 0.0489, Dice coefficient of 0.7030, and jaccard index of 0.5423. After the training, testing dataset were fed in the model and achieved a performance of 62% dice coefficient, and 45% jaccard index with a focal loss of 16% as shown on Table I.

TABLE I
Model Performance

Performance Metric	***Score***
Dice Coefficient	62%
Jaccard Index	45%
Focal Loss	16%

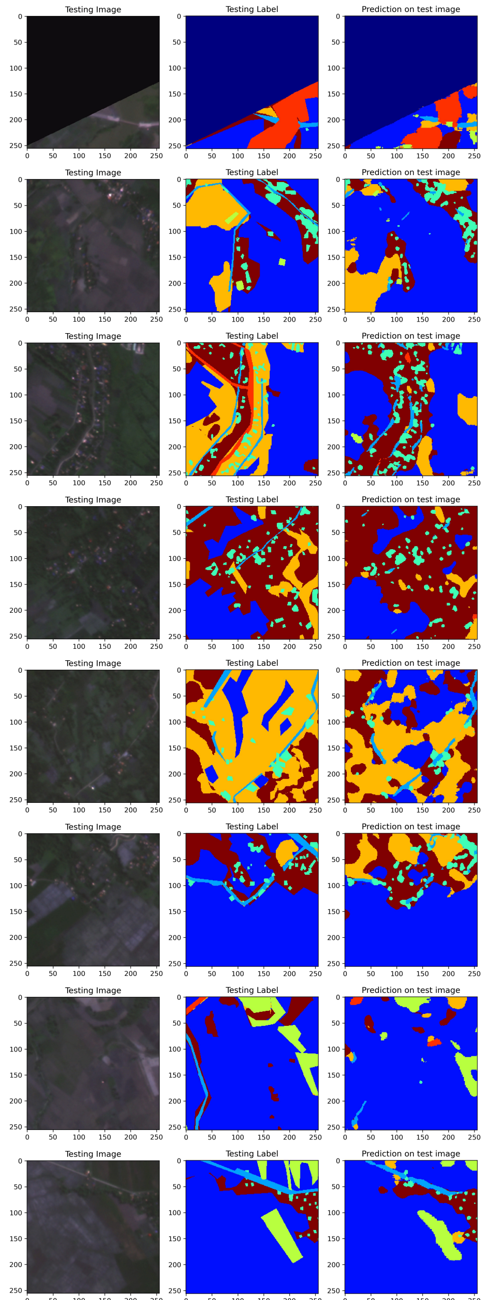
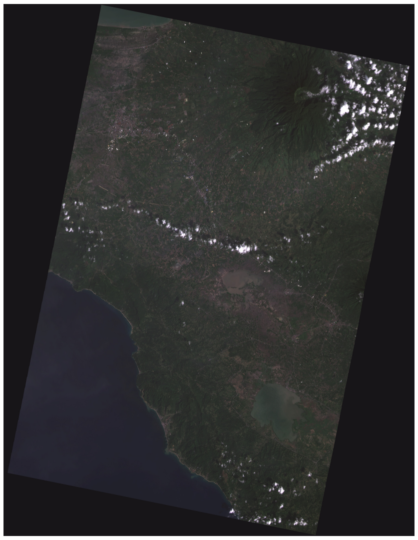


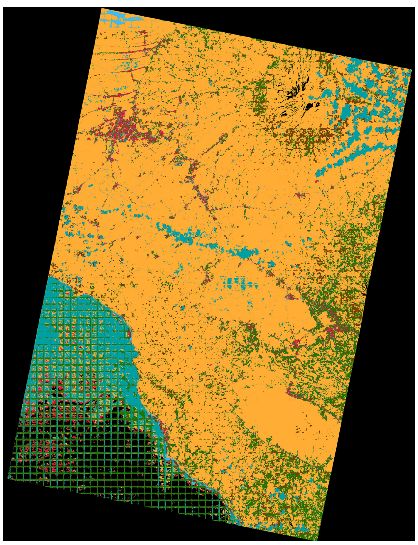
Fig. 6. Visualization of model segmentation result

C. Model Testing

Shown on Figure 6 is the visualization of some segmentation result of the testing dataset. The first column is the input satellite image, the second column is the labelled mask and the third column is the predicted mask. The result shows that the model can predict the satellite image correctly in majority of its assigned class such as buildings, water, wetland, trees, agriculture, and grounds. However, as for roads, the model sometimes can't segment the road and wetland properly. This is due to unbalanced dataset since there are lesser labels for these classes.

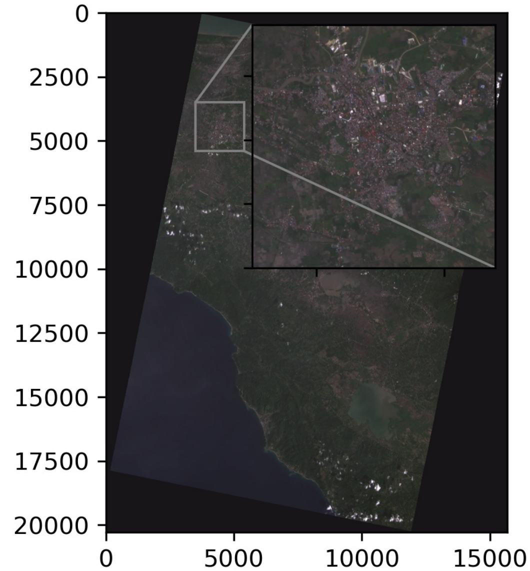


(a) Satellite Image

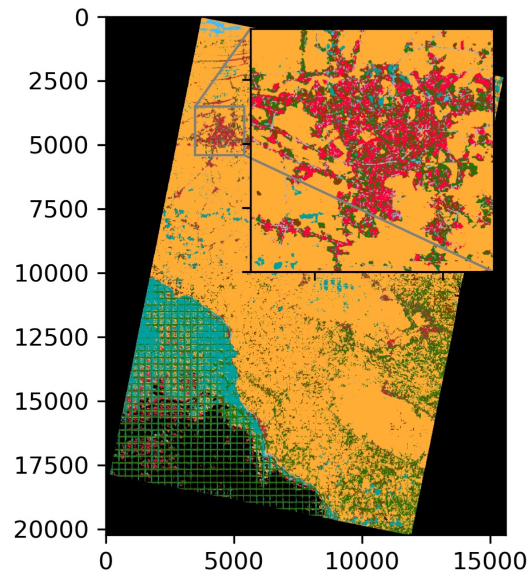


(b) Segmented Land Cover

Fig. 7. Model Prediction on Large-Scale Satellite Image



(a) Zoomed Portion



(b) Zoomed Portion of Segmented Land Cover

Fig. 8. Zoomed Portion of Model Prediction on Large-Scale Satellite Image

D. Model Prediction to Large-Scale Satellite Image

Shown in Figure 8 is the predicted large-scale satellite image of Rinconada which is the 5th District of Camarines Sur. It can be seen the various red spots which corresponds to buildings. Red spots shows the central location of the cities and municipalities. Majority of the Land Area is predicted as cropland which is visualized in orange pixels however, there are a large amount of areas which were wrongly predicted as cropland specially around the area of the volcano. The green spots are the trees which is the second noticeable in the segmented image. As for wetland, majority of the wetland were mislabeled, this is due to lesser number of labelled classes for wetland. Clouds in the satellite image were predicted as water and some part of the sea are predicted as water however, there are some part that were not predicted correctly. There were also parts of the sea where there it can be seen as squares. This is due to the prediction of small patches and reconstructed without smooth blending.

V. Conclusion and Recommendation

This study intends to develop a model using Residual Attention U-Net that classifies the land cover of Philippine Satellite Images. Large-scale satellite image was acquired from DOST - DATOS and were pre-processed to smaller patches in order to train in a Residual Attention U-Net Model and generated 184 patches. However, 184 patches are not enough for training and testing of the model. Hence, data augmentation were used to improve the number of training and testing dataset by flipping and rotating image patches. Data augmentation increased the number of patches to 1104. After which, patches were labelled in Labelbox with eight (8) classes to generate a mask for these patches.

A Residual Attention U-Net model underwent training with labeled patches serving as the input images. The performance of the model was determined during the training and testing using the Jaccard Index and Dice Coefficient. After a long hour of training, the model achieved a performance of 62% dice coefficient, and 45% jaccard index with a focal loss of 16%. This revealed that the model can predict the majority of the classes as presented on the predicted images in the result and discussion however, there are some classes which are predicted incorrectly. Nevertheless, given a dice coefficient of 62%, the model cannot guarantee a perfect prediction.

A large-scale satellite image of Rinconada, the 5th district of Camarines Sur, Philippines was also fed into the model and predicted majority of classes, especially building, cropland and roads. Central areas of the municipalities and cities in the Rinconada area can easily identified by looking at the large-scale predicted satellite image.

Given all the accomplished objectives of the study, further improvement shall be done such as improving the number of image dataset and its classes to achieve higher performance in segmentation. The image dataset can also be feed in other deep learning architectures to show the differences and similarities in terms of performance. Furthermore, smooth blending can also be used in the prediction of large-scale satellite image for a more cleaner and smooth prediction.

VI. Acknowledgments

This paper and the research behind it would not have been possible without the support of the Department of Science and Technology (DOST) DATOS - Remote Sensing and Data Science Help Desk for providing free satellite images. The BSCS undergraduate on-the-job trainees who served as data labeller for this project and last but not the least, I would like to thank my family, colleagues, and friends for supporting spiritually.

References

[1] M. A. Wulder, N. C. Coops, D. P. Roy, J. C. White, and T. Hermosilla, "Land cover 2.0," *International Journal of Remote Sensing*, vol. 39, no. 12, pp. 4254–4284, 2018.

[2] Q. Zhu, X. Guo, W. Deng, S. Shi, Q. Guan, Y. Zhong, L. Zhang, and D. Li, "Land-use/land-cover change detection based on a siamese global learning framework for high spatial resolution remote sensing imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 184, pp. 63–78, Feb. 2022.

[3] M. Azari, L. Billa, and A. Chan, "Multi-temporal analysis of past and future land cover change in the highly urbanized state of selangor, malaysia," *Ecological Processes*, vol. 11, Jan. 2022.

[4] T. Hermosilla, M. A. Wulder, J. C. White, and N. C. Coops, "Land cover classification in an era of big and open data: Optimizing localized implementation and training data selection to improve mapping outcomes," *Remote Sensing of Environment*, vol. 268, p. 112780, Jan. 2022.

[5] N. S.-N. Lam, "Methodologies for mapping land cover/land use and its change," in *Advances in Land Remote Sensing*, pp. 341–367, Springer Netherlands, 2008.

[6] X.-Y. Tong, G.-S. Xia, Q. Lu, H. Shen, S. Li, S. You, and L. Zhang, "Land-cover classification with high-resolution remote sensing images using transferable deep models," *Remote Sensing of Environment*, vol. 237, p. 111322, Feb. 2020.

[7] C. Robinson, L. Hou, K. Malkin, R. Soobitsky, J. Czawlytko, B. Dilkina, and N. Jojic, "Large scale high-resolution land cover mapping with multi-resolution data," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[8] S. S. Rwanga and J. M. Ndambuki, "Accuracy assessment of land use/land cover classification using remote sensing and GIS," *International Journal of Geosciences*, vol. 08, no. 04, pp. 611–622, 2017.

[9] P. Helber, B. Bischke, A. Dengel, and D. Borth, "EuroSAT: A novel dataset and deep learning benchmark for land use and land cover classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, pp. 2217–2226, July 2019.

[10] M. M. Makinano-Santillan and J. R. Santillan, "MAPPING LAND COVER CHANGE AND MODELLING ITS IMPACTS ON THE INUNDATION RESPONSES OF THE AGUSAN MARSH, MINDANAO, PHILIPPINES," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLIII-B3-2021, pp. 595–602, June 2021.

[11] D. Shrestha, B. Brebante, V. Jetten, E. Herrera, A. Blanco, R. Gonzalez, and D. Alkema, "Land cover change and the use of soilgrids data for assessing flash flood in marikina river basin, the philippines : abastract," 2018. AOGS-EGU Joint Conference : New Dimensions for Natural Hazards in Asia ; Conference date: 04-02-2018 Through 06-02-2018.

[12] J. M. Jardeleza, C. K. Gotangco, and M. A. L. Guzman, "Simulating national-scale deforestation in the philippines using land cover change models," *Philippine Journal of Science*, 2019.

[13] B. A. Johnson, K. Iizuka, M. A. Bragais, I. Endo, and D. B. Magcale-Macandog, "Employing crowdsourced geographic data and multi-temporal/multi-sensor satellite imagery to monitor land cover change: A case study in an urbanizing region of the philippines," *Computers, Environment and Urban Systems*, vol. 64, pp. 184–193, July 2017.

[14] I. Estacio, M. Basu, C. P. Sianipar, K. Onitsuka, and S. Hoshino, "Dynamics of land cover transitions and agricultural abandonment in a mountainous agricultural landscape: Case of ifugao rice terraces, philippines," *Landscape and Urban Planning*, vol. 222, p. 104394, June 2022.

[15] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science*, pp. 234–241, Springer International Publishing, 2015.

[16] J. J. S. Oñate and M. Ang-Tolentino, "Exploring RAU-net for semantic segmentation of philippines satellite images in identification of building density," *International Journal of Remote Sensing*, vol. 43, pp. 5738–5756, Nov. 2021.

[17] Z.-L. Ni, G.-B. Bian, X.-H. Zhou, Z.-G. Hou, X.-L. Xie, C. Wang, Y.-J. Zhou, R.-Q. Li, and Z. Li, "RAUNet: Residual attention u-net for semantic segmentation of cataract surgical instruments," in *Neural Information Processing*, pp. 139–149, Springer International Publishing, 2019.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, June 2016.

[19] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual u-net," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, pp. 749–753, May 2018.

[20] P. Jaccard, "THE DISTRIBUTION OF THE FLORA IN THE ALPINE ZONE.1," *New Phytologist*, vol. 11, pp. 37–50, Feb. 1912.

[21] T. Sørensen, *A method of establishing groups of equal amplitude in plant sociology based on similarity*, vol. 5. København: I kommission hos E. Munksgaard, 1948.

[22] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," 2017.