CEST 2023

18th International Conference on Environmental Science and Technology

Athens, Greece, 30 August to 2 September 2023

Predicting Land Cover Map Changes in the Philippines for use in LULC-based Carbon Capture Monitoring using Deep Learning

CAMACLANG R.[1], BALLESTEROS F.[2], LABAO A.[3,*]

[1]University of the Philippines Los Baños, Laguna, Philippines
[2]University of the Philippines Diliman, Quezon City, Philippines
[3] University of the Philippines Diliman, Quezon City, Philippines

*corresponding author:
e-mail: alfonso.b.labao@gmail.com

Abstract In the area of Carbon Capture and Storage (CSS), monitoring plays an important role not only to determine if there is any anomalies in the release of CO_2 in the atmosphere, but also to prepare for disasters and to plan better future developments in the industrial sector, transportation sector, real estate development, and other sectors. One way to monitor changes in the carbon cycle is by looking at Land Use and Land Cover (LULC) changes, since the primary methods of carbon capture and storage is by biological and geological sequestration. In this study, we designed a Deep Learning model that can predict land cover changes in the Philippine Land Cover Maps generated by the National Mapping and Resource Information Authority (NAMRIA). We evaluated our results and our model yielded a 78.64% overall accuracy and a Kappa coefficient of 0.725.
Keywords: Land Use and Land Cover (LULC), Land Cover Prediction, Carbon Capture and Storage, Deep Learning

1. Introduction

The Philippine islands, with a land area of approximately 299,764 km2 (Doroteo, 2015) and more than 7600 islands, constitute one of the largest archipelagos in the world. The country has a tropical climate and is greatly influenced southwest and northeast monsoons. Its wet season extends from June to November while the dry season is from December to May (Pang, et al., 2021). Due to its geographical location, the country is also exposed to some of nature's hazards. High incidence of tropical storms, tsunamis, earthquakes, volcanic eruptions, landslides and droughts are observed in the country (Doroteo, 2015). In terms of population, according to the recent census of the Philippines Statistics Office, the country has now approximately 109 million (PSA,2020) and still is steadily growing. These two factors, natural calamities and demographics, along with other factors contribute to the changes in the land cover and land use in the country.

1.1 Land Cover of the Philippines

According to Liu, et al. (2022), Land use and land cover (LULC)" is the most intuitive and widespread representation of surface systems". Land cover is the composition and characteristics of land surface elements. It is an important determinant of land use and thus of value of land to the society (Cihlar, 2000). It also shows the spatial distribution of the different covers on the Earth's surface (Feng & Li, 2020).

The National Mapping and Resource Information Authority (NAMRIA) released Land cover map of the Philippines in 2010 and 2015 (See Figure 1). The Land cover mapping project is a nationwide assessment of land cover using a satellite-based images, more specifically the LandSat 8 (30meter resolution) images taken from 2014 to 2016.

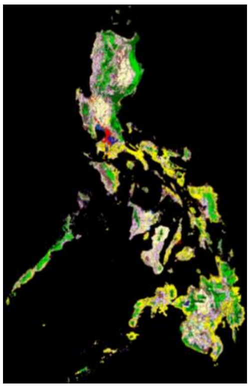

Figure 1: 2015 Land Cover Map and classes (NAMRIA).

1.2 Carbon Capture and Land Use and Land Cover(LULC)

The study of Carbon Capture (or sequestration) and its storage is important to control and manage the carbon cycle, especially when talking about energy consumption. The Carbon storage techniques that are in use can be classified into 3 main categories: geological storage, ocean storage or mineralization (Pires, et al., 2011). In line with the discussions on biological and geological carbon sequestration techniques, one aspect that is worth looking at is the effects of changes in land use and land cover on carbon sequestration. LULC covers four types of carbon storage carriers, namely, above-ground biomass, below-ground biomass, soil biomass, and dead biomass. All these contribute to carbon storage and carbon emissions (Liu, et al., 2022).

1.3 Deep Learning

Deep Learning is a branch of machine learning that has become very popular in the recent years. Deep Learning, according to LeCun, et al. (2015), "allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction". It has been successfully applied to areas such as speech recognition (Deng, et al, 2013), object detection (Zhao, et al, 2019), machine translation (Singh, et al., 2017), among many others.

2. Problem Statement

Carbon Sequestration monitoring is an important area in the field of Carbon Capture and Storage. When there is a monitoring system in place, it is easier to plan for future developments, to mitigate the effects of impending problems, and to prepare in case of disaster. One way of effectively monitoring Carbon capture is by looking at land use and land cover changes, since the primary methods of carbon capture and storage is by biological and geological sequestration. In this study, we plan to address the following:

1. Develop a prediction model for Land Cover changes in the Philippines using the latest machine learning technique, namely Deep Learning, that could prospectively be used in computing future carbon capture computation of the country;
2. Produce a working dataset from the NAMRIA Land Cover maps of the Philippines for training and testing the prediction model;
3. Assess the usefulness of the prediction model, including the prediction results and future directions of this study.

With these percentage of changes in the land cover classes, we can the compute for future carbon capture estimates using any LULC-based model of carbon capture computation. An example of these models is the InVEST Carbon Storage and Sequestration model designed by the Natural Capital Project of Stanford University.

3. Methodology

3.1 Experimental Setup

For this study, we used a machine with the following specifications:

- Intel I7 Processor; 32 GB RAM; 3080 GPU
- OS: Windows 10 OS
- Programming language: Python 3

3.2 Data

The data that we used for study came from the GeoPortal PH website. It is a website containing a collection of the different maps generated by various government agencies. In particular, we used the following maps:

- NAMRIA Baseline Map (for geographical reference)
- ArcGIS Map (for background)
- NAMRIA Land Cover Map (2010)
- NAMRIA Land Cover Map(2015).

We took screenshots of different areas in the Philippines and named each sample according to the most prominent Philippine province included in the image. We only used one zoom level for all the training and test sets. All in all, we were able to gather 15 land cover images and we divided it into training and test sets. We chose 10 images for training and the remaining 5 as test set.

3.3 Data Pre-processing

Color Correction

In order to correctly construct our prediction model, we need to make sure that the colors within a specific land cover area are uniform. If we zoom in to our image dataset, we can see different shades of a specific color per class. We corrected this using a Python Script.

Class Merging

In order to simplify our prediction task, we decided to combine come of the land cover classes we came up with 7 different classes as shown in figure 2. We did this using another Python script.

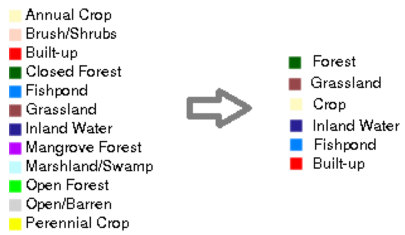

Figure 2: Land Cover classes merged to 7 classes(including void).

3.4 Prediction Model

The Deep Learning (DL) architecture that we used for the experiments is the Google Deeplab v3 segmentation network with a 50 Layer residual network trunk. It is a pre-trained network built-in in the PyTorch machine learning framework. We trained the model using a 3080 model graphics processing unit (GPU) for around 24 hours. We then tested the trained prediction model using our test set and evaluated the results using the following metrics.

3.5 Metrics for Analysis

First of all, we generate a confusion matrix of the prediction results. The columns represent the prediction results from the perspective of the reference class (true class) while the row represent it from the perspective of the predicted class. The diagonal of the matrix are the correctly classified pixels.

We then computed for the overall accuracy of the model which is simply the total number of correctly predicted pixels over the total number of pixels classified.

We also computed for the errors of omission (Type I) and commission (Type II). Error of omission refers to the true class that were left out(omitted) after the prediction. Error of commission, on the other hand, is the error that is incurred when there is an incorrect classification of the pixel.

Lastly, we computed for the Kappa Coefficient which is a statistical measure to evaluate the accuracy of the prediction, that is, how well the prediction task performed compared to just randomly assigning values. This measure gives off a value from -1 to 1, where in a negative value close to -1 can be interpreted as worse than random while a value close to positive 1 means the model performed better than random (HSU, 2022). We used the *cohen_kappa_score()* function from the *sklearn.metrics* package in Python to generate this.

4. Experimental Results and Analysis

After the model was trained, we used it to predict the land cover changes in our test set and we got the following results (see Figure 3):

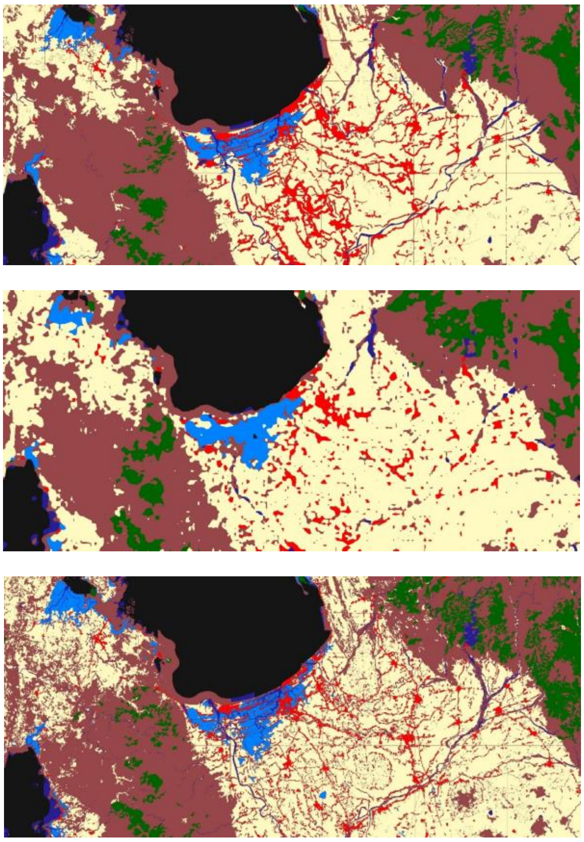

Figure 3: Prediction of Pangasinan Land Cover 2010 Actual (above) vs 2015 Predicted (middle) vs 2015 Actual (below).

4.1 Confusion Matrix

We begin analyzing our results by generating a confusion matrix of the prediction.

Table 1. Confusion Matrix of the prediction result (in pixel count).

	Void	Forest	Grassland	Crops	Inland Water	Fish ponds	Built-up	Total
Void	**2,410,825**	1,284	28,274	1,966	9,739	1,119	1,002	**2,454,209**
Forest	2,806	**586,381**	167,530	12,931	1,415	106	385	**771,554**
Grassland	20,390	139,848	**1,449,948**	549,567	22,545	22,378	52,406	**2,257,082**
Crops	590	7,406	248,172	**1,459,796**	3,279	6,893	45,431	**1,771,567**
Inland Water	17,335	3,504	42,368	24,374	**167,745**	24,877	4,839	**285,042**
Fish ponds	1,385	812	48,168	61,670	11,842	**188,107**	5,040	**317,024**
Built-up	1,155	1,178	47,284	126,569	1,764	8,257	**290,115**	**476,322**
Total	**2,454,486**	**740,413**	**2,031,744**	**2,236,873**	**218,329**	**251,737**	**399,218**	**8,332,800**

As we can see in table above, the diagonal refers to the correctly classified pixel count per class. When we look at the columns, we interpret it in reference to the ground truth of the specific class. For example, when we look at the void column, the item on the diagonal (in bold) is the number of pixels that were correctly identified as ground truth. Meaning, the other items on the column are missed (or omitted) ground truth. If we add all those missed pixels and divide it over the total number of ground truth pixels for that class, we get the omission error. In the case of the void class, we get 2% omission error. On the other hand, when we look at the table row-wise, we can refer to the item in the diagonal as the correctly predicted pixel and the other items in the row as the incorrectly predicted pixels. When we add those items in a particular row that are not part of the diagonal, and we divide it over the total number of predicted pixels in that class, we get the commission error (see table 2 for complete tally of the errors). This means that in the case of forest class, we have 24% commission error. The Kappa coefficient is at .725, which indicates that our prediction model performs much better than random and is thus robust(HSU, 2022).

4.2 Overall Accuracy

The overall accuracy of the predictor model was computed as follows:

Number of correctly classified site:
2, 410, 825 + 586, 381 + 1, 449, 948 + 1, 459, 796 + 167, 745 + 188, 107 + 290, 115 = **6, 552, 917**
Total number of reference sites = 8,332,800
Overall Accuracy = 6,552,917/8,332,800 = 78.64%

4.2 Kappa Coefficient

The Kappa coefficient was generated by using the *cohen_kappa_score()* function from the *sklearn.metrics* package in Python.

Kappa: 0.7251390800877328

4.3 Omission and Commission Errors

Below are examples of how omission and commission errors are computed. You can find the tallied results in table 2 below:

Omission Error Example based on the above confusion matrix:

Void Class:
Incorrectly classified ground truth pixels:
2,806 + 20,390 + 590+ 17,335 + 1,385 + 1,155 = 43,661
Total # of reference sites = 2,454,486
Omission Error = 43,661/2,454,486 = 2%

Commission Error Example based on the above confusion matrix:

Forest class:

Incorrectly classified sites:

2,806 + 167,530 + 12,931 + 1,415 + 106 + 385 = 185,173

Total # of classified sites = 771,554

Commission Error = 185,173/771,554 = 24%

Table 2. Summary of Omission and Commission Errors.

Land Class	Omission Error	Commission Error
Void	2%	2%
Forest	21%	24%
Grassland	29%	36%
Cropland	35%	18%
Inland water	23%	41%
Coastal Water	25%	41%
Built-up	27%	39%

4.4 Percentage of Land Cover Change

We also computed for the percentage of the land cover class changes in the predicted map. We did this by tallying first the number of pixels per class in the test set in the year 2010 side by side with the predicted 2015 land cover classes (see table 3).

Table 3. Number of pixels per land cover class (2010 Actual vs. 2015 Predicted).

	Manila		Masbate		Palawan		Pampanga		Pangasinan	
	2010	2015	2010	2015	2010	2015	2010	2015	2010	2015
Void	24,850	419,930	904,693	911,683	847,870	852,338	24,943	24,578	246,667	245,957
Forest	177,773	175,478	454	-	321,838	291,219	152,152	176,305	80,371	97,411
Grassland	383,714	356,634	374,983	333,216	374,213	414,755	533,131	442,270	557,699	484,869
Crops	218,709	296,630	328,111	384,804	52,833	57,079	632,013	777,422	623,770	720,938
Inland Water	126,790	127,789	34,983	21,274	32,912	29,329	47,490	23,286	32,515	16,651
Fish ponds	92,390	59,801	20,662	14,338	34,974	21,751	138,701	112,111	40,796	43,736
Built-up	242,334	230,298	2,674	1,245	1,920	89	138,130	110,588	84,742	56,998

Next, we get the percentage of each land cover class from each sample in the test data (both 2010 Actual and 2015 Predicted). Then, we compute for the difference between the predicted and the actual and the result will give us the percentage of change in the land cover classes. See table 4 for the summary of percentage changes in the land cover classes.

Table 4: Percentage increase (or decrease) of land cover classes (Actual 2010 vs Predicted 2015).

	Manila			Masbate			Palawan			Pampanga			Pangasinan		
	2010	2015	Diff	2010	2015	Diff	2010	2015	Diff	2010	2015	Diff	2010	2015	Diff
Void	25.5%	25.2%	-0.3%	54.3%	54.7%	0.4%	50.9%	51.1%	0.3%	1.5%	1.5%	0.0%	14.8%	14.8%	0.0%
Forest	10.7%	10.5%	-0.1%	0.0%	0.0%	0.0%	19.3%	17.5%	-1.8%	9.1%	10.6%	1.4%	4.8%	5.8%	1.0%
Grassland	23.0%	21.4%	-1.6%	22.5%	20.0%	-2.5%	22.5%	24.9%	2.4%	32.0%	26.5%	-5.5%	33.5%	29.1%	-4.4%
Crops	13.1%	17.8%	4.7%	19.7%	23.1%	3.4%	3.2%	3.4%	0.3%	37.9%	46.6%	8.7%	37.4%	43.3%	5.8%
Inland Water	7.6%	7.7%	0.1%	2.1%	1.3%	-0.8%	2.0%	1.8%	-0.2%	2.8%	1.4%	-1.5%	2.0%	1.0%	-1.0%
Fish ponds	5.5%	3.6%	-2.0%	1.2%	0.9%	-0.4%	2.1%	1.3%	-0.8%	8.3%	6.7%	-1.6%	2.4%	2.6%	0.2%
Built-up	14.5%	13.8%	-0.7%	0.2%	0.1%	-0.1%	0.1%	0.0%	-0.1%	8.3%	6.6%	-1.7%	5.1%	3.4%	-1.7%

5. Conclusion and Recommendation

5.1 Conclusions

Firstly, we were able to gather sufficient amount of land cover images to be used in training and testing our prediction model. We were able to set up a Deep Learning architecture specifically for the task of predicting Philippine land cover using the Google Deeplab v3 segmentation network with a 50 Layer residual network trunk. We have shown in this study that our model was able to predict future land cover maps using the NAMRIA Land Cover map data set. We evaluated the results of our prediction and we got 78.64% overall accuracy. The Kappa coefficient is at .725, which indicates that our prediction model performs much better than random. Lastly, we have identified the usefulness of this prediction model in computing for future carbon stocks using LULC-based carbon capture models.

5.2 Future Research Recommendations

We have identified several possible future studies related to this problem:

1. Test and compare other Deep Learning architecture to see if it will improve the prediction accuracy;
2. Use a different zoom level in collecting data from the Land Cover maps of the Philippines; the model might perform better in a higher zoom level. Also, to use samples with accurate measurement of the land area for better computation;
3. Use all the 14 land cover classes in training the model.

References

Land cover 2010 and 2015, https://www.geoportal.gov.ph/.

Philippine Statistics Authority. Household population, number of households, and average household size of the philippines (2020 census of population and housing).

HSU. Accuracy metrics. GSP 216 : Introduction to remote sensing, may 2, 2022[Online].

Doroteo, HJ. (2015). Philippines: disaster risk profile and disaster risk reduction (ddr) framework: natural calamities. *University of Oviedo—Department of Medicine Unit for Research and Emergency and Disaster, Asturias, Spain*,.

Pang, S. E., De Alban, J. D. T., & Webb, E. L. (2021). Effects of climate change and land cover on the distributions of a critical tree family in the Philippines. *Scientific Reports*, 11(1), 276.

Liu, Q., Yang, D., Cao, L., & Anderson, B. (2022). Assessment and prediction of carbon storage based on land use/land cover dynamics in the tropics: a case study of hainan island, China. *Land*, 11(2), 244.

Cihlar, J. (2000). Land cover mapping of large areas from satellites: status and research priorities. *International journal of remote sensing*, 21(6-7), 1093-1114.

Feng, M., & Li, X. (2020). Land cover mapping toward finer scales. *Sci. Bull*, 65(19), 1604-1606.

Pires, J. C. M., Martins, F. G., Alvim-Ferraz, M. C. M., & Simões, M. (2011). Recent developments on carbon capture and storage: An overview. *Chemical engineering research and design*, 89(9), 1446-1460.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444.

Deng, L., Hinton, G., & Kingsbury, B. (2013, May). New types of deep neural network learning for speech recognition and related applications: An overview. *In 2013 IEEE international conference on acoustics, speech and signal processing* (pp. 8599-8603). IEEE.

Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11), 3212-3232.

Singh, S. P., Kumar, A., Darbari, H., Singh, L., Rastogi, A., & Jain, S. (2017, July). Machine translation using deep learning: An overview. *In 2017 international conference on computer, communications and electronics (comptelix)* (pp. 162-167). IEEE.