Classifying Urban Land in Indonesia Using Convolutional Neural Networks

Final Report

Rebekah Adams & Tiffany M. Tran

Department of City and Regional Planning, University of Pennsylvania

MUSA 6500 Geospatial Machine Learning in Remote Sensing

Dr. Guray Erus

April 2023

1. Introduction

Indonesian cities are rapidly expanding. Between 2000 and 2014, Indonesian cities expanded by 3.9%, or 6,904 km^2—more than the land area of Bali (as cited in Coalition for Urban Transitions, 2021). Nearly three-quarters of this expansion was onto cultivated land, consuming cropland and protective ecosystems, such as mangroves. While urbanization in Indonesia is associated with economic growth and poverty reduction, urban expansion has negative implications for both climate and development. Deforestation has resulted in a significant loss of green land cover and biodiversity, depleting carbon sinks in Indonesia. Urban expansion will also increase emissions from urban service provision and increase travel distances between people and economic opportunities, especially in growing cities.

If smaller cities are urbanizing at faster rates than large cities, managing the associated effects of urban expansion, including distances required for urban service delivery, could hold much potential to prevent carbon emissions while supporting economic growth. Density and compact urban form are especially important in Indonesia's cities with populations under 1 million, where the majority of infrastructure needs to be built and carbon intensive spatial patterns have yet to be locked in. Over three-quarters of Indonesia's urban abatement potential to 2050 is estimated to be in these secondary cities (Coalition for Urban Transitions, 2021). Compact and transit-oriented development in secondary cities—especially in developing countries—are therefore important policy levers to maximize climate action and economic development. Yet benchmark datasets for land cover often do not include spatial patterns of developing countries and require high computational resources to manipulate, creating challenges to tracking urban expansion in rapidly urbanizing areas. To address these data limitations, we create a novel dataset and pilot three machine learning algorithms to classify urban land in the secondary city of Cirebon, Indonesia.

2. Data

For our testing and training data, we create a novel dataset of 1,424 tiles using Sentinel-2 satellite imagery georeferenced to Cirebon (Figure 1). Each tile is 64x64 pixels with a spatial resolution of 10 meters per pixel. To create the dataset, we first download Sentinel-2 satellite imagery over the entire metropolitan area of Cirebon, taken with 3% cloud cover in 2022. We next generate a grid with 64x64 pixel tiles in QGIS and overlay the grid over the satellite imagery. We then manually label each tile as either urban ($\hat{y} = 1$) and or non-urban ($\hat{y} = 0$) depending on its land cover class. To determine the boundary between urban and non-urban, we use a simple "intersects join" to combine the 2014 urban expansion polygon from the Atlas of Urban Expansion dataset for Cirebon with the grid (NYU, 2016).

The resulting dataset contains 114 urban samples and 1310 non-urban samples. Because models using this initial dataset resulted in skewed predictions, we create a balanced data set with the 114 urban images and a random selection of 114 non-urban images, resulting in a dataset of 228 images (Figure 2). We then use data augmentation to expand the number of urban and non-urban images and create a larger and balanced final dataset.

Figure 1. Creating Image Tiles

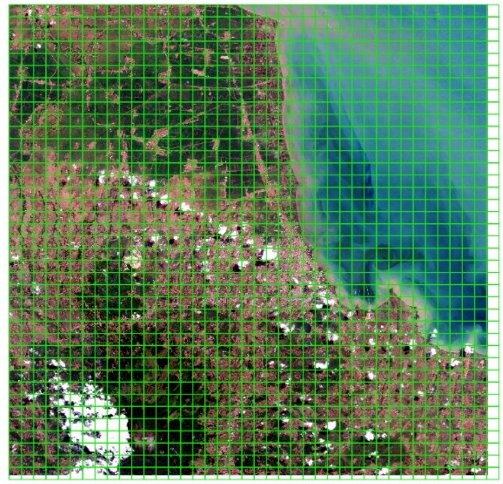
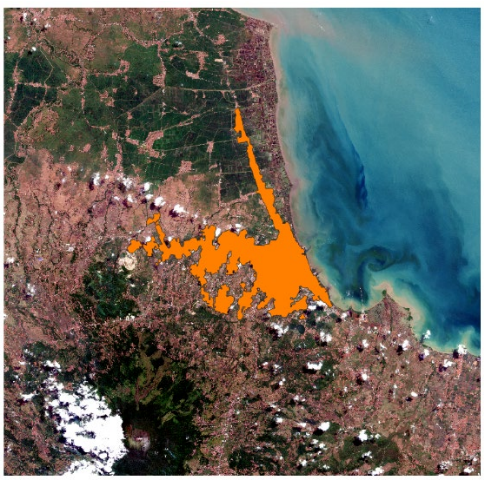
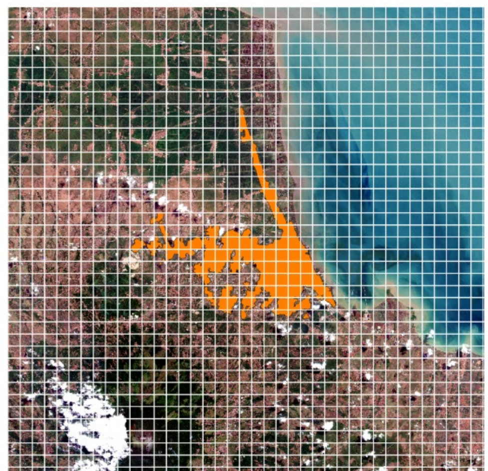
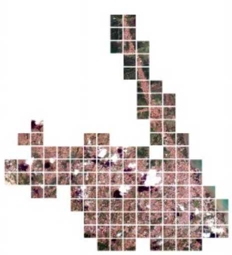
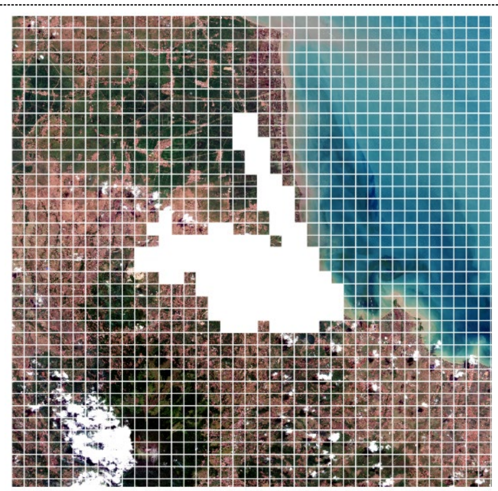
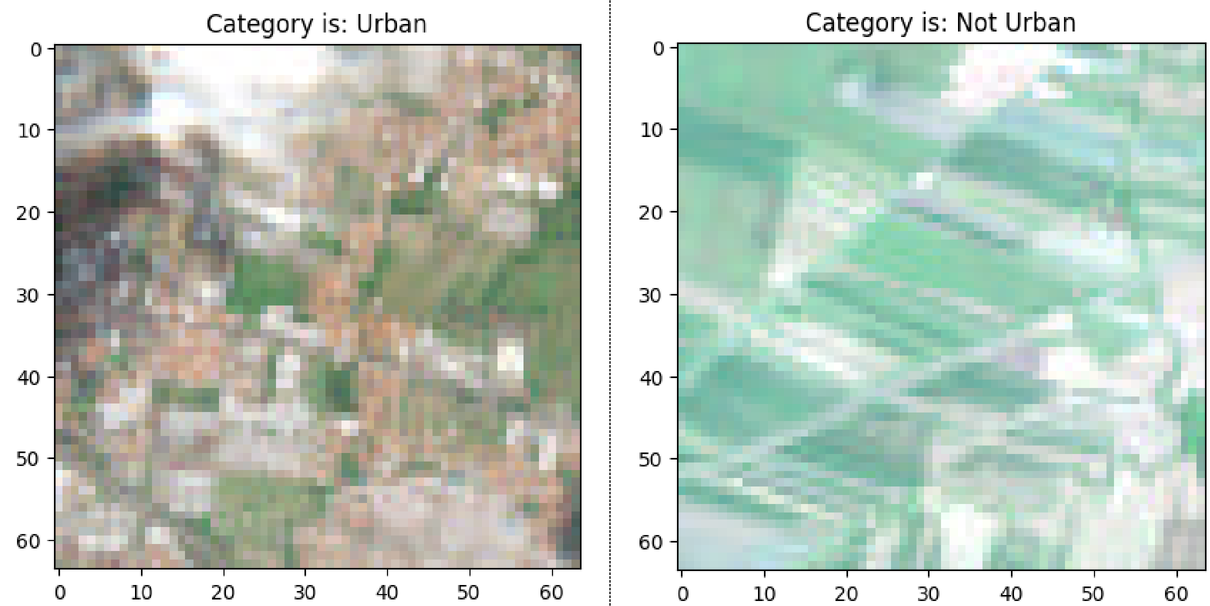
1. Download Sentinel-2 imagery	2. Overlay tile grid
	
3. Overlay urban expansion polygon	4. Intersects join polygon to grid
	
5. 114 urban image tiles	6. 1,310 non-urban image tiles
	

Figure 2. Dataset Visualization



3. Methods

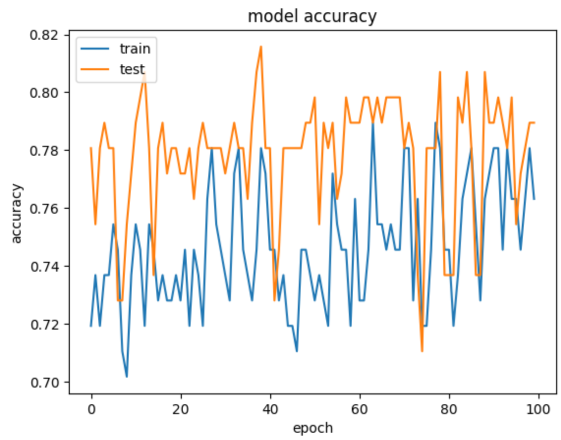
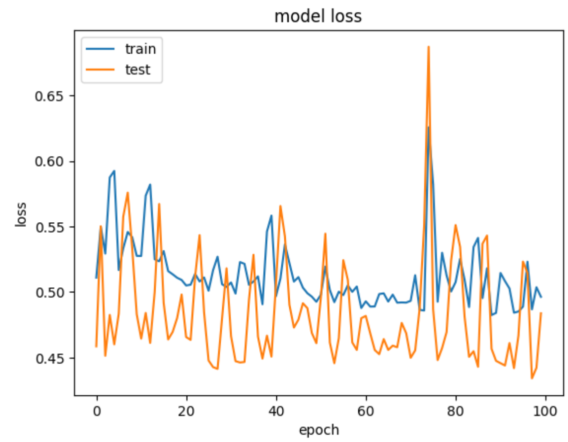
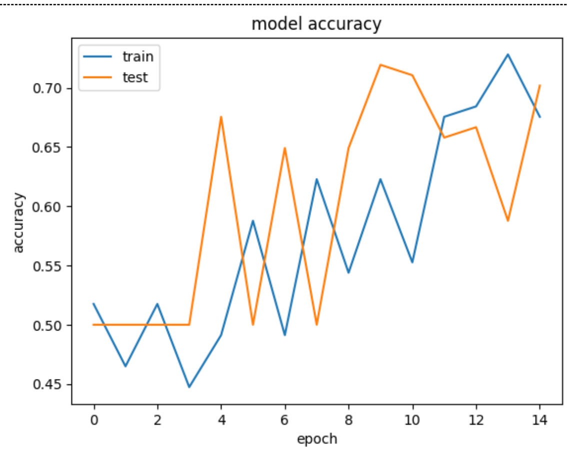
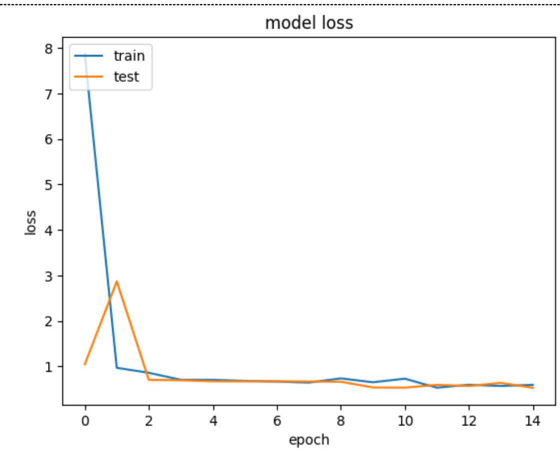
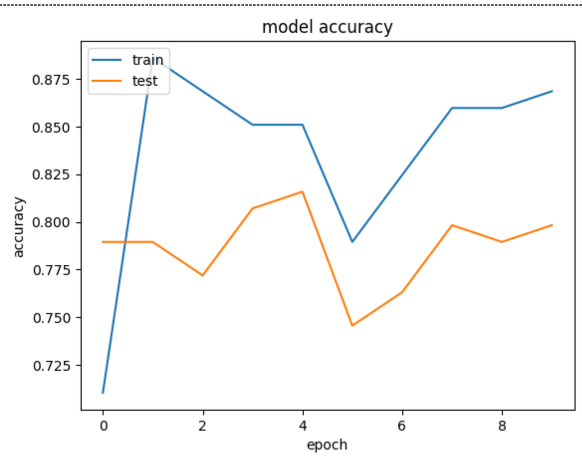
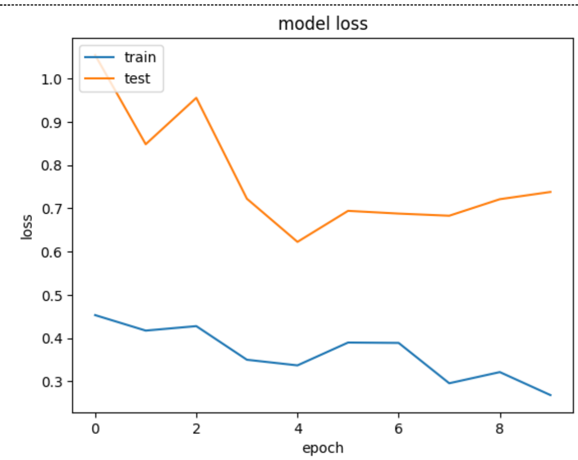
We train our augmented dataset using three CNN architectures, as CNN is well-suited to extracting local information from images for classification. We use a custom CNN as a base model and test it against VGG16 and EfficientNetV2 models, both pre-trained on ImageNet data, to see whether transfer learning improves accuracy scores. First, we build a custom CNN architecture which applies a 3x3 filter with 32 convolutions, activated by the ReLU function. It uses a 2x2 pooling window to reduce dimension and randomly drops 25% of the elements to prevent overfitting before flattening the data. It then passes the reshaped data into a dense layer which uses a softmax function for the output. Because we have low image resolution, we tune the kernel size and pooling size parameters to be small. In our second model, we add several layers to the VGG16 architecture: a dense layer with 50 neurons activated by the ReLU function; a dense layer with 20 neurons also activated by the ReLU function; and a dense layer activated by the softmax function for the output. We also randomly drop 20% of the elements to prevent overfitting. Lastly, we try EfficientNetV2, a new pre-trained CNN architecture that trains faster and is up to nearly seven times smaller than previous CNN models (Tan & Le, 2021). As with VGG16, we add the same custom layers. We compile all models using the Adam optimizer and a loss function that minimizes binary cross-entropy and then train over 100 epochs with a batch size of 100. To find the best model for classifying urban land, we compare their testing accuracy scores and AUCs.

In addition to comparing the model performance, we also plot false positive samples to visualize urban expansion in Cirebon. We expect that some of the image tiles that were labeled as non-urban based on the 2014 urban boundary and misclassified by our models as urban are indeed new areas of urban expansion that lie beyond the 2014 urban expansion polygon boundary. These false positive samples should indicate areas of urban expansion between 2014 and 2022, if any.

4. Results

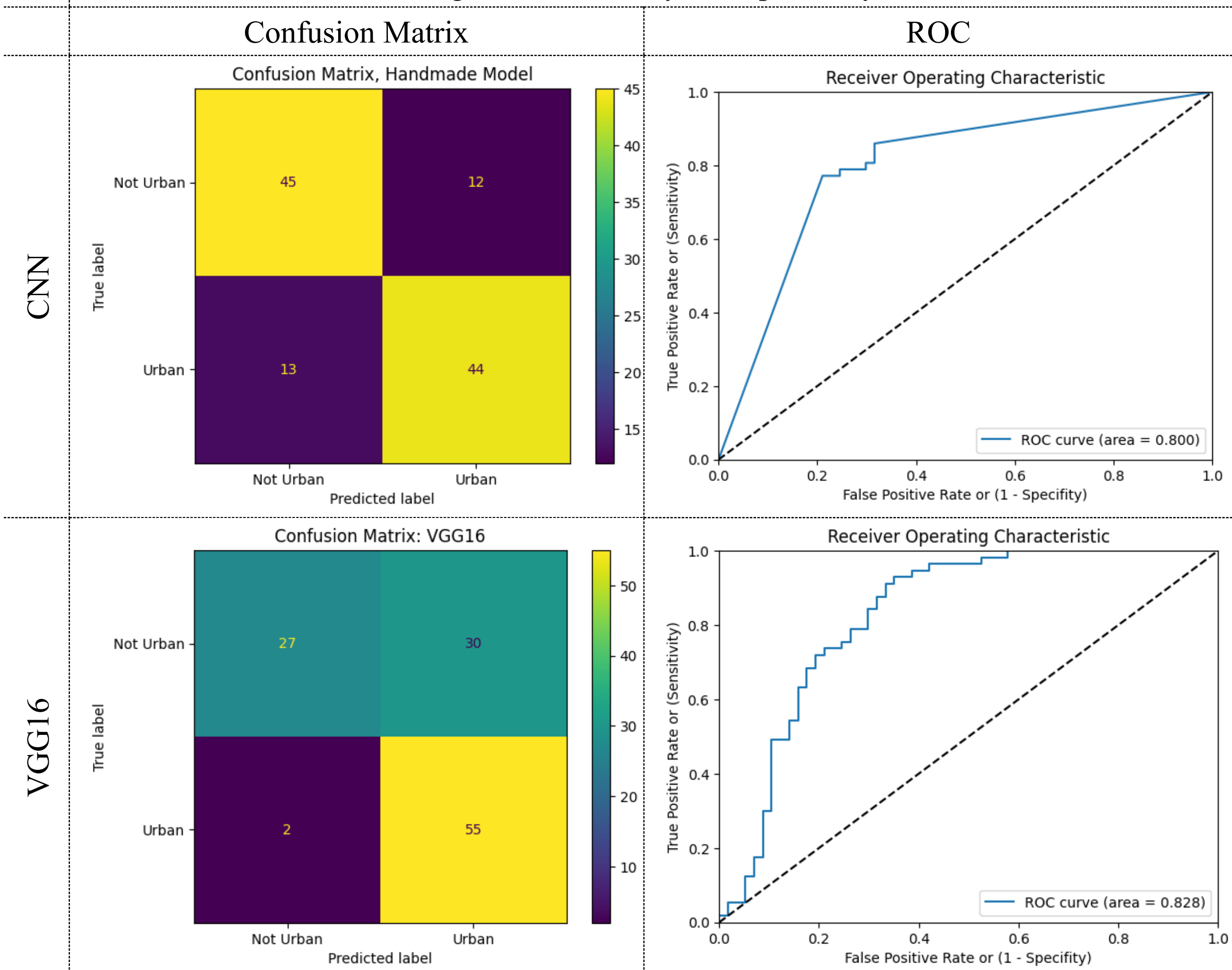
Comparing accuracy and loss, the customized EfficientNetV2 model outperforms the others, with a testing accuracy of 81.6% (Figure 3). The custom CNN model also performs well, with a testing accuracy of 78.1%. The customized VGG16 model has the lowest performance, with a testing accuracy of 71.9%.

Figure 3. Model Comparison

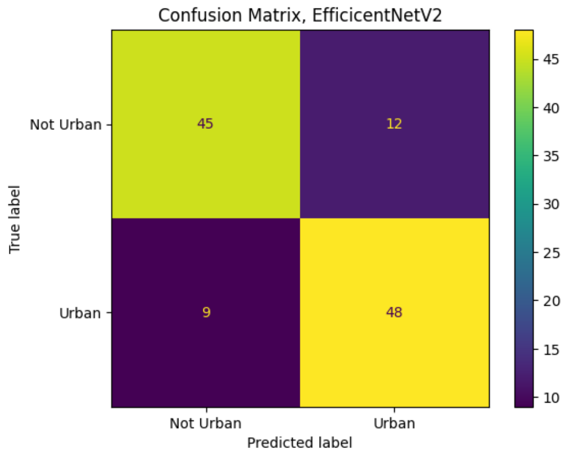
	Accuracy	Loss
CNN		
VGG16		
EfficientNetV2		

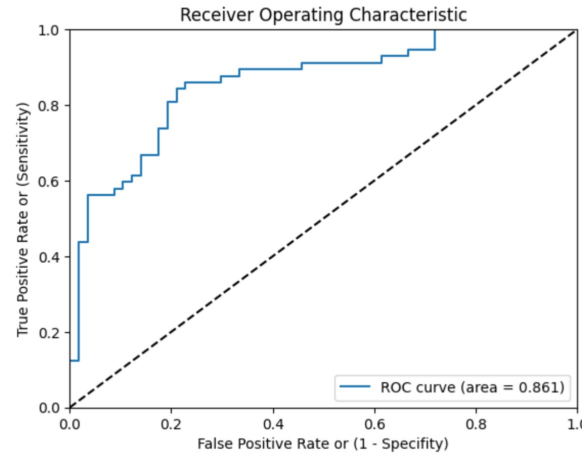
We also plot confusion matrices and ROCs to evaluate performance (Figure 4). For CNN, the true positive rate was 77.2% compared to a specificity of 78.9% (or false positive rate of 21.1%). VGG16 had a true positive rate of 96.5% and a specificity of 47.4% (or false positive rate of 52.6%), while EfficientNetV2 had a true positive rate of 84.2% and specificity of 78.9%. CNN and EfficientNetV2 were thus better at predicting non-urban images while VGG16 performed best at predicting urban images. Considering the tradeoff between sensitivity and specificity, the ROC plots show that the EfficientNetV2 model had the highest AUC at 0.861 and was thus the best classifier for distinguishing between urban and non-urban images overall.

Figure 4. Sensitivity and Specificity



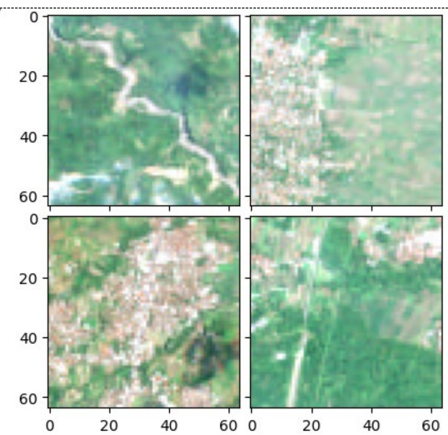
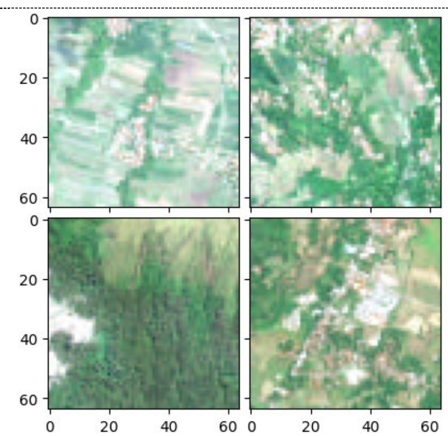
EfficientNetV2



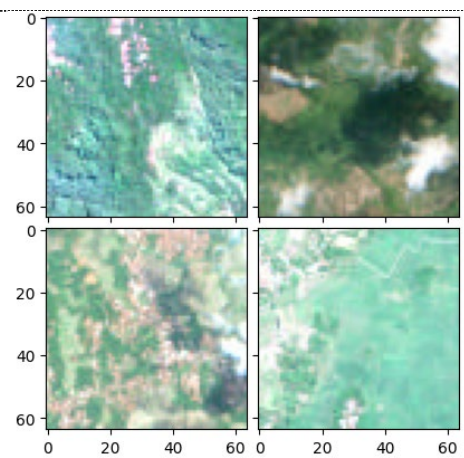


Examples of false positive samples for each model indeed contain areas of urban land as expected (Figure 5). These areas likely indicate new urban expansion since 2014 that is not captured within the urban expansion polygon boundary from the Atlas of Urban Expansion.

Figure 5. Examples of False Positive Classifications (Predicted = Urban)

CNN	VGG16
	

EfficientNetV2



5. Discussion

In this study, we attempt to classify urban land in Cirebon, Indonesia using three deep learning algorithms: CNN, VGG16, and EfficientNetV2. We find that the customized EfficientNetV2 model outperforms the others across all performance measures, with a testing accuracy of 81.6% and AUC of 0.861. While the custom CNN and VGG16 models also performed well, the former likely needed further training to compete with the pre-trained models, while the latter took the longest time to compile. For land cover classification, using transfer learning with the EfficientNetV2 algorithm may be an effective way to achieve good accuracy.

We also attempt to visualize urban expansion by plotting false positives “misclassified” as urban. We find that the urban extents of Cirebon have indeed expanded since 2014, with many false positive samples containing land that appears urban. These findings suggest that machine learning algorithms can be trained to identify urban land, including urban expansion, using novel and augmented datasets. This study contributes a simple and efficient methodology for understanding urban expansion in secondary cities, particularly in developing countries with limited data availability.

Our methodology has two limitations. Because we use a simple “intersects join” to combine the 2014 urban expansion polygon from the Atlas of Urban Expansion dataset for Cirebon with the grid, some tiles labeled as urban may actually be mostly non-urban, contributing to a higher misclassification error. Using zonal statistics and object detection to calculate the actual proportion of urban land within each tile would improve the overall accuracy of the labeling. Secondly, because the urban expansion polygon boundary is based on 2014 urban extents of the city, it likely underestimates urban land because of urban expansion within the past decade. While this allows us to plot the false positive samples to visualize newly urban areas, it likely contributes to lower testing accuracy scores. Mapping current urban extents of the city would improve model training and performance.

References

Coalition for Urban Transitions. (2021). Seizing Indonesia's urban opportunity: compact, connected, clean, and resilient cities as drivers of sustainable development. *Seizing the Urban Opportunity*. Coalition for Urban Transitions.

NYU Urban Expansion Program. (2016). *Cirebon*. Atlas of Urban Expansion. http://atlasofurbanexpansion.org/cities/view/Cirebon

Sengupta, D., Choi, Y. R., Tian, B., Brown, S., Meadows, M., Hackney, C. R., Banerjee, A., Li, Y., Chen, R., & Zhou, Y. (2023). Mapping 21st century global coastal land reclamation. *Earth’s Future*, *11*. https://doi.org/10.1029/2022EF002927

Tan, M. & Le, Q. V. (2021). EfficientNetV2: Small models and faster training. *Proceedings of the 38 th International Conference on Machine Learning*, PMLR 139. https://arxiv.org/abs/2104.00298