Deprived Area Mapping using Deep Learning on Sentinel-2 Imagery

Ryan Engstrom
Department of Geography and Data Science
George Washington University
Washington D.C, USA
rengstro@gwu.edu

Amir Jafari
Department of Data Science
George Washington University
Washington D.C, USA
ajafari@gwu.edu

Abdulaziz Gebril
Department of Data Science
George Washington University
Washington D.C, USA
gebril51@gwu.edu

Mojahid Osman
Department of Data Science
George Washington University
Washington D.C, USA
mojahid_osman@gwu.edu

Mina Hanna
Department of Data Science
George Washington University
Washington D.C, USA
minahanna@gwu.edu

***Abstract*—**Mapping of deprived areas has been a key research topic over the past years as it helps optimize the development planning and decision making by governments and other policy makers. This paper aims to utilize free to use satellite images to map deprived areas with a relatively smaller sized window of 100 m2.

This paper utilizes various geographic transformation techniques, image processing, classification algorithms and deep neural networks to classify and map deprived areas. First, cloud free maps were generated to obtain clear images for further processing. Second, image processing, data augmentation and other geographic transformations took place on training data maps and other data sources to build a comprehensive, balanced dataset for deep learning training. Third, various machine learning techniques were applied across different tracks to qualify the best models that will be combined to provide the best results of the mapping. We reached f1-score around 0.8 using low resolution satellite imagery. **(***Abstract***)**

Keywords—: image classification; deprived area; satellite imagery; informal settlements; convolutional neural networks; deep learning *(key words)*

I. Introduction

Mapping deprived areas from satellite images and other data sources will help governments, international and independent organizations to categorize the level of deprivation, prioritize projects and closely monitor the progress. This domain has been a key research area for the past decade with a lot of promising results supporting development efforts in line with various global initiatives to end poverty and build sustainable cities. The advancement in machine learning techniques and availability of Geo-data offers an opportunity to achieve efficient models.

According to an IdeaMaps review , Convolutional Neural Network and Classical Machine Learning algorithms are the most used techniques in mapping deprived areas in recent key publications, but they are mostly limited due to their high computational requirements. This research aims to apply different machine learning techniques on low resolution and free satellite images along with distance covariate and contextual data and other free datasets to detect deprived areas on a 100m2 level. This approach will provide adequate balance between having a reasonably computational requirement and not losing the ability to scale the solution.

This paper focuses on using labeled satellite images to train Convolutional Neural Network (CNN) and classify deprived areas. The labeled data was mainly provided by Ideamaps Network based on their ground observation work where deprived areas were mapped based on an agreed survey that was conducted by a local team in different cities.

II. Related Work

(Krishna et al., 2018) used Very High-Resolution Satellite images and applied CNN model to have a multi-class and Hierarchical class classifier able to classify informal areas with detailed labels like single/multi-story with overall accuracy reaching 0.71 and F-1 score of 0.70 for the multi-class classification.

(Monika et al., 2019) used Very High-Resolution Satellite images and applied a two-step CNN model (where one is trained on binary classification and the other added data from Data-Driven Index of Multiple Deprivation with image augmentation and reached an accuracy of 0.984.

(Michael et al., 2019) used transfer learning from a model trained on very high resolution imagery from QuickBird to Sentinel-2 and TerraSAR-X data. It was noticed that a significant increase for Sentinel-2 applying transfer learning can be observed (from 38 to 55% and from 79 to 85% for PPV and sensitivity, respectively).

III. Data

A Sentinel-2 Image was provided covering Lagos, Nigeria in a GeoTiff Format. The GeoTiff file included 47,560 labeled 100m2 locations. There are three types of labels which are built-up, non-built-up and deprived areas. the labels are not equally distributed where the non-built-up areas represented 67.5% of the total labels. The built-up areas and deprived areas represented 32% and 0.5% respectively. All labeling was conducted through ground work by IdeasMaps Network.

IV. Methodology

In this section, we describe different steps conducted to train various deep learning models classifying deprived areas in Lagos. Initially, Data Preprocessing was needed in order to generate model-ready data.

A. Data Preprocessing

The first step was to clip the images that correspond to the labeled locations and convert them to PNG format. During conversion, we kept the NIR (Near Infrared) band provided in the GeoTiff file and normalized the values of each pixel between 0-255 to fit the RGB standards.

In addition, Some satellite images were taken at a time where cloud presence disturbed the visibility of the image. Therefore, Cloud free images with NIR bands were generated using Google Earth. This resulted in much clearer images capturing more detailed information.

Figure 1. Extracting labeled data

B. Data Augmentation

Due to the complexity of mapping 100 m2 areas, the deprived area labels are significantly less than non-deprived labels and this required to generate more images from the minority class (deprived) to construct the required dataset for model training.

This process involved different techniques which can be summarized in the following points:

1- Shifting images by 1 or 2 meters in 4 directions.
2- Detect any adjacent labeled areas and roll through them to generate intersecting new images.
3- Rotate the image in different directions.

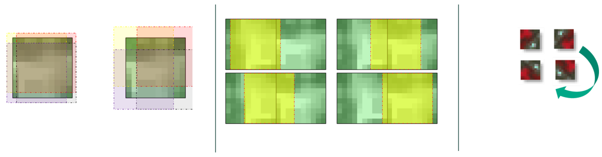

Figure 2. Techniques used to generate new deprived areas images

C. Modeling

Using a fully connected feedforward Convolutional Neural Network. Each Convolution layer will convolute the input image with a kernel and generate a feature map after applying an activation function.

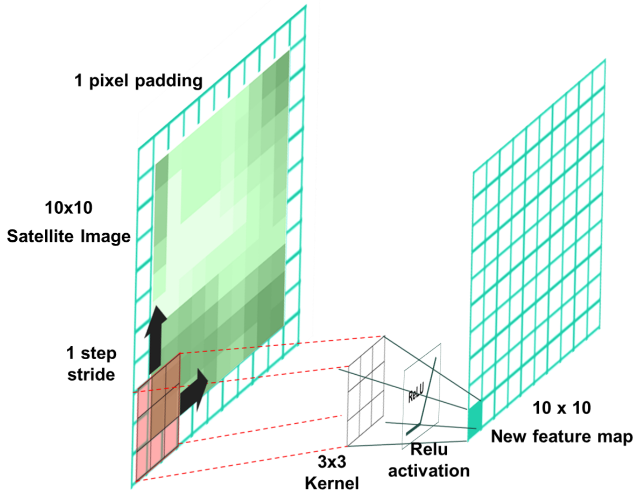

Figure 3. Convolution Layer Processing

Padding is set to be the same and the stride step size is set to one to ensure that the size (10x10) will go through the deeper layers. The kernel size of all layers is set to 3x3 and the activation function was set to “Relu”.

Pooling is used to reduce the dimension of the applied kernel and map it to a single point in the next layer and Batch Normalization layers are used to are used to keep output close to mean of 0 and variance of 1 which prevents the saturations of the activation function and help speeding up the training.

The following figure shows the Convolutional Neural Network architecture utilized in the model development.

Conv2D – 60
Batch Normalization
Conv2D – 120
Batch Normalization
Max Pooling
Conv2D – 200
Batch Normalization
Conv2D – 200
Conv2D – 100
Batch Normalization
Max Pooling
Linear Layer – 128
Dropout 25%
Linear Layer – 128
Dropout 25%

Figure 4. Convolution Neural Network Architecture

Multiple Convolutional Neural Networks training experiments were conducted The following was considered during training:

- Fattening layers will be applied. We also added a Dropout layer to reduce overfitting.
- The labels were reshaped into two outputs in order to utilize the softmax activation function which provided better results.
- We used 100 epochs with an early stopping option to avoid overfitting.
- GridSearch was used for hyperparameter tuning for different experiments.
- The use of different Pre-trained models was examined.
- Manually setting class weights for different labels was tested in order to penalize minority class misclassification. .
- Developing ensemble models based on the best trained models was tested.
- Our main metrics for the model were accuracy, precision, recall and F1-score.

V. Results & discussion

In this section, we discuss the results of the training experiments conducted in the previous section. Some observations were noticed. Using the Sentinel-2 images with Adamax, Adam and RMSProp the best accuracy achieved was with Adamax. Since non-built-up labels achieved with different methodologies (i.e. OpenStreetMap features or Google Open Building Dataset), it was decided to exclude the Non-built-up label and only use two labels in the classification (built-up and deprived).
The following are key highlights from the CNN modeling:

- Cloud free images with NIR band provided better results
- Small image size (10x10) could be the reason behind low performance of pre-trained models and dilation
- Class weights were used to penalize minority class misclassification but did not add significant improvements
- Augmentation was a key factor for improving model performance
- Hard-voting ensemble was used to combine the best performing models

M01	3 Labels | original image + class weights
M02	2 Labels | original image + class weights
M03	2 Labels | cloud free image + class weights
M04	2 Labels | cloud free image + class weights + augmt. (rotation)
M05	2 Labels | cloud free image + class weights + augmt. (rotation) VGG16
M06	2 Labels | cloud free image + class weights + augmt. (rotation) - blank images
M07	2 Labels | cloud free image + class weights + augmt. (rotation) - blank images VGG16
M08	2 Labels | cloud free image + class weights + augmt. (rotation) - blank images + dilation
M09	2 Labels | cloud free image + augmt.(rotation, rolling and shifting)
M10	Hard voting Ensemble model

Ref	Test Acc.	Deprived Label			Macro Avg		
		Preci -sion	recall	f1-score	Preci sion	recall	f1-score
M01	0.63	0.02	0.78	0.03	0.55	0.64	0.46
M02	0.69	0.03	0.57	0.06	0.51	0.63	0.44
M03	0.72	0.04	0.72	0.08	0.52	0.72	0.46
M04	0.84	0.15	0.87	0.26	0.57	0.85	0.58
M05	0.85	0.15	0.76	0.25	0.57	0.81	0.58
M06	0.92	0.27	0.63	0.37	0.63	0.78	0.67
M07	0.87	0.18	0.67	0.29	0.58	0.77	0.61
M08	0.88	0.21	0.70	0.32	0.60	0.80	0.63
M09	0.96	0.32	0.63	0.43	0.66	0.80	0.70
M10	0.97	0.41	0.65	0.50	0.70	0.81	**0.74**

Tables 1,2 . key models' performance

	Built-up	Deprived
Built-up	1939	50
Deprived	19	35

Table 3. Confusion Matrix for Ensemble Model

VI. Model Experiment

Since the number of images kept aside for testing the model was relatively small (54 images based on the random stratified sampling used), it was essential to validate that model's performance to avoid inaccurate results due to under-representative data samples. To do so the models were trained using various random seeds to split the data differently and the model metrics were consistent. To further validate the numbers, the following experiment was conducted:

1- Annotation of new labels using manual reviews on high resolution google maps and street views
2- Split the expanded dataset which will have more test images
3- Run the model with more data points
4- Validate the KPIs

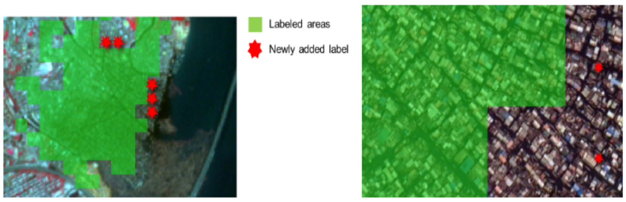

Figure 3. Using high resolution map to add more labels

Acknowledgment

We would like to acknowledge IdeaMaps for conducting all the groundwork necessary to provide us with the data needed for this research..

References

[1] Elena Ranguelova, Berend Weel, Debraj Roy, Monika Kuffer, Karin Pfeffer & Michael Lees (2019) Image based classification of slums, built-up and non-built-up areas in Kalyan and Bangalore, India, European Journal of Remote Sensing, 52:sup1, 40-61, DOI: 10.1080/22797254.2018.1535838

[2] Michael Wurm, Thomas Stark, Xiao Xiang Zhu, Matthias Weigand, Hannes Taubenböck,Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks,ISPRS Journal of Photogrammetry and Remote Sensing, Volume 150, 2019, ISSN 0924-2716, Pages 59-69, https://doi.org/10.1016/j.isprsjprs.2019.02.006

[3] Boris Babenko, Jonathan Hersh, David Newhouse, Anusha Ramakrishnan, Tom Swartz, Poverty Mapping Using Convolutional Neural Networks Trained on High and Medium Resolution Satellite Images, With an Application in Mexico, https://arxiv.org/pdf/1711.06323.pdf

[4] Christina Corbane, Vasileios Syrris, Filip Sabo, Panagiotis Politis, Michele Melchiorri, Martino Pesaresi, Pierre Soille & Thomas Kemper , Convolutional neural networks for global human settlements mapping from Sentinel-2 satellite imagery, https://link.springer.com/article/10.1007/s00521-020-05449-7

[5] Aaron E. Maxwell , Timothy A. Warner and Luis Andrés Guillén: Accuracy Assessment in Convolutional Neural Network-Based Deep Learning Remote Sensing Studies—Part 1: Literature Review

[6] Ajami, Alireza, Monika Kuffer, Claudio Persello, and Karin Pfeffer. 2019. "Identifying a Slums' Degree of Deprivation from VHR Images Using Convolutional Neural Networks" *Remote Sensing* 11, no. 11: 1282. https://doi.org/10.3390/rs11111282

[7] Krishna Karthik Gadiraju, Ranga Raju Vatsavai, Nikhil Kaza, Erik Wibbels, Anirudh Krishna. Machine Learning Approaches for Slum Detection Using Very High Resolution Satellite Images, 2018 IEEE International Conference on Data Mining Workshops (ICDMW), p.1397-1404, https://doi.org/10.1109/ICDMW.2018.00198