Measuring road network expansion and its effect

Hadrien Sevel [315940], Pietro Pecchini [377243], Matthijs Scheerder [377763]
CS-433 Machine Learning - Project 2 - EPFL

***Abstract*—Access to healthcare facilities is a significant challenge in underdeveloped countries, often due to the lack of paved roads connecting small villages to hospitals. Research, such as that cited in [1], indicates that this limited access contributes to higher postnatal child mortality rates, as many mothers give birth without specialist assistance. Expanding the road network could therefore reduce child mortality and improve overall living standards. This article presents a machine learning model, based on the U-Net architecture, designed to identify roads in satellite images (specifically, Landsat 8). Trained using Open Street Maps (OSM) data as labels, our model aims to detect the presence of roads in image pixels. However, the results of our model are inconclusive.**

I. Introduction

In recent years, the expansion of road networks in Sub-Saharan Africa (SSA) has become a critical factor in improving healthcare accessibility. Our project aims to develop a methodology using machine learning to identify and analyze these expansions. We utilize Open Street Map and satellite data to predict changes in road networks. Our primary goal is to compare different identification methods, assess their strengths and weaknesses, and then integrate this information with Demographic and Health Survey (DHS) data from a specific country. This integration will enable a comprehensive evaluation of the impact on health facility access.

A key element of our methodology is the use of U-Net architectures, inspired by their successful application in biomedical image segmentation as detailed in Ronneberger et al.'s seminal paper [2]. U-Nets are recognized for their efficiency and precision in segmenting images, even with limited data, making them an ideal choice for our task. Their structure, which includes a downsampling path to capture context and a symmetric upsampling path for precise localization, forming a "U" shape, is especially suited for the intricate task of identifying and analyzing road networks in the varied landscapes of SSA.

II. Initial Data Analysis

Since in the project description provided to us also other tasks were listed, we briefly summarize our results here. Because the actual machine learning part is the most interesting and is also what is expected from us from the course, the rest of the report is mainly focused towards that. The 5 tasks listed were the following:

1) Calculate the expansion of the road network based on the Open Street Map (OSM) data from 2013-2022
2) Explore the use of satellite data to identify road networks with machine learning methods.
3) Explore other ideas for identifying road network expansions.
4) Compare the different identification methods and discuss their strengths and weaknesses.
5) Aggregate predicted/measured road network expansion at Demographic and Health Survey (DHS) cluster level and merge with DHS data of a specific country to assess the effect on the likelihood of visiting a health facility.

Task 1

We were able to calculate the expansion of the road network in Ghana. We chose Ghana as a country as this seemed to have a more developed south compared to the north. Furthermore the road expansion over the years seemed reasonable. We calculated a road expansion based on OpenStreetMap (OSM) of 129% for the period 2017-2023. We could also conclude that GRoads data is only available up to 2010. The road expansion in the period 2010 to 2017 was calculated as 199%. Note that 2010 is thus GRoads and 2017 OSM. Tamale, the largest city of northern Ghana, appeared to be an ideal starting point for concentrating our efforts in task 2. This choice was influenced by the relatively low road density in the area, and its status as a large city within a predominantly rural region, making it a strategic hub for healthcare services.

Task 3

During the second task, we discovered that relying on satellite data makes task 2 significantly challenging. It is advisable to shift to other methods. Müller-Crepon et al. explored the use of using custom fully convolutional neural network to convert road network data from physical maps (Michelin maps) into a digital format, creating a time-variant digital road atlas for post-colonial Africa, similar to Google Maps[3][4]. The fact that these maps are usually released over 5 year intervals means that it could be used as the time-series data needed for task 5.

Task 4

Ideally we would use a fitting machine learning model. The disadvantage of using such model is that it is harder to segment roads in areas where roads are more covered and of very bad quality, i.e. tropical countries around the equator like Democratic Republic of Congo (DCR). A benefit of using physical maps is that it focuses less on the inner city roads but more on the connecting roads. Those are the most important roads for access to advanced infrastructure as health care facilities[5]. Moreover, it's worth noting that well-mapped areas, such as around large cities, do not resemble indistinct blobs as compared to using OSM.

Task 5

Due to our inability to develop a functional road segmentation model, we were regrettably unable to pursue this task. Alternatively, we could have attempted to address it solely with our OpenStreetMap (OSM) data. However, given our incapacity to differentiate between a road being genuinely new or merely recently mapped, we concluded that obtaining a dependable result was unlikely.

III. Machine Learning Models and Methods

A. Data acquisition and data processing

The data used in this project is open source and readily accessible. To analyze road density, we extract geographic areas from OpenStreetMap (OSM) using its PBF database dumps. This extraction is performed with the Osmium Tool and is defined by a geojson file. The extracted data is then filtered to include only road information and converted into geojson format for compatibility with the Python library Geopandas. However, additional processing with Geopandas is necessary, as we found that Osmium sometimes retains non-road elements after filtering.

We also use Landsat 8-9 OLI/TIRS C2 L2 satellite imagery from the USGS Earth Explorer, specifically the Level-2 Surface Reflectance Bands. These bands are significant as Level-2 imagery provides atmospherically corrected surface reflectance, which is derived from the Level-1 data. A key selection criterion for these images is the cloud cover below 20%, ensuring clarity. These satellite images, which provide surface reflectance values across seven different bandwidths at a resolution of 30m x 30m per pixel, are stored in TIF format. For processing these images, we employ the Rasterio library.

To facilitate analysis, both Landsat data and OSM data are aligned to the same reference system (WGS 84 / World Mercator). This alignment allows us to effectively overlay the datasets. By masking the Landsat data with the OSM data, we create a label matrix that indicates which pixels are intersected by roads.

As the satellite images are quite large, we divide them into smaller patches of size 286x286 pixels. Originally, the U-Net architecture utilizes images of 572x572 pixels, but due to the lower resolution of the Landsat images, which makes road distinction challenging, this size was found to be too large. Consequently, we halved the dimensions. For a single satellite image focusing on the Tamale region, we generated 702 patches. An example of these patches is illustrated in Figure 1 (in red, the roads obtained from OSM).

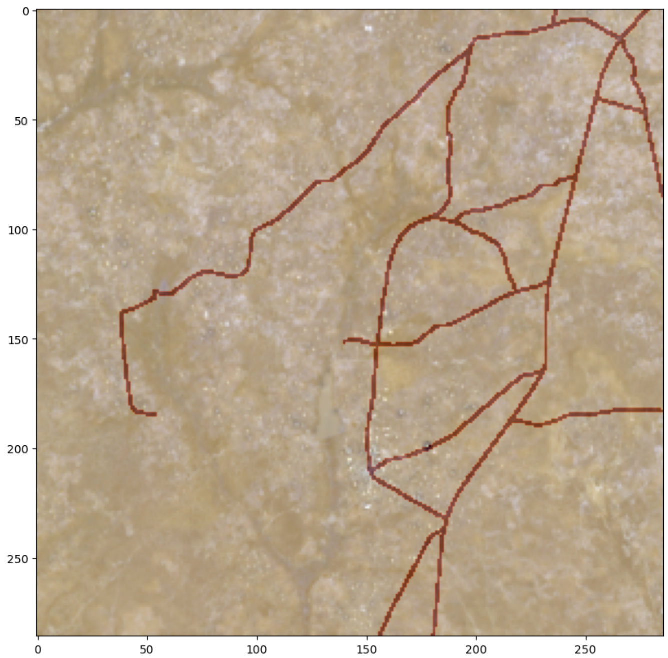

Figure 1. Patch showing OSM data on a satellite image

We also added a buffer to the road data from the OSM dataset to increase the apparent width of the roads, which were originally represented by a single pixel. A buffer of 10 pixels was applied, as this size most accurately represented the actual width of the roads.

To further enhance our dataset, we performed feature augmentation on the patches. This augmentation included rotations and stretching, with a stretching factor of 1.3. The model was then trained on this augmented dataset across various epochs. The purpose of these augmentations is to make the model more resilient to variations in angles and image deformations that might result from atmospheric disturbances.

B. Convolutional network architecture

As previously mentioned, the U-Net architecture we employ is adapted from Ronneberger et al.'s work [2]. The U-Net is essentially a fully convolutional neural network. Below is a brief overview of the architecture we used; for a detailed explanation, we refer readers to Ronneberger et al. [2].

The network comprises two primary components: the encoder (downsampling path) and the decoder (upsampling path), linked by a central bridge. The encoder features four sets of convolutional blocks, each containing two consecutive convolutional layers with ReLU (rectified linear unit) activation functions. This results in eight convolutional layers within the encoder. Max-pooling layers are integrated at each step to downsample the feature maps and capture critical information. The bridge, comprising convolutional layers with ReLU functions, connects the encoder to the decoder. The decoder mirrors the encoder's structure with four sets of layers, each with two ReLU-activated layers, totaling eight layers.

A key feature of the decoder is its use of skip connections from the encoder. These connections, added at each stage of the decoder, are vital for preserving spatial information. They are concatenated with the upsampled feature maps from preceding layers, ensuring accurate segmentation and context retention in the decoding process. Figure 2 shows the original U-Net architecture as described.

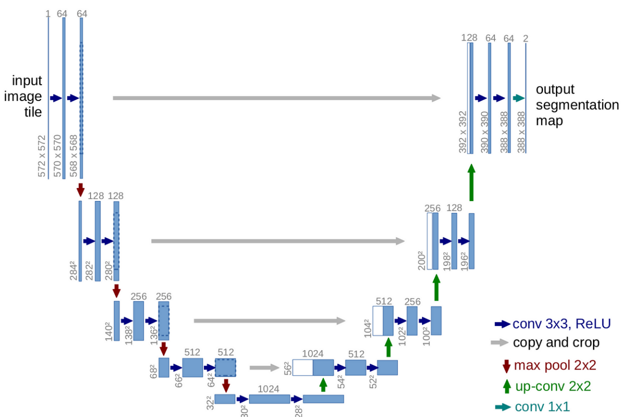

Figure 2. Original U-Net architecture (taken from U-Net paper [2])

For our specific application, several modifications were necessary. The model is fed with seven channels instead of the standard three, allowing it to process all seven bands of the satellite images. To maintain the original size of the input images and ease the masking process with our label matrix, we implemented padding of 1 in all convolutional layers. Additionally, optional batch normalization was added between the convolution and ReLU layers to normalize layer inputs and stabilize training.

C. Training

To enhance our model's detail capture capabilities, the original dataset is preprocessed into smaller patches as detailed in Section III-A. Each patch is sized at 286×286 pixels, covering approximately 74 km^2.

We split the dataset into a training set and a validation set as follows: 20% of the patches are allocated to the validation set to ensure the model is tested on unseen data during training. The remaining 80% form the training set, providing a broad range of data for the model to learn from.

Given that the algorithm's goal is binary pixel classification, we choose the Binary Cross Entropy loss function. To address the significant data imbalance (predominantly non-road pixels), we apply a weighted cross entropy loss. This weight is calculated based on the ratio of non-road to road pixels. For model parameter updates, the Adam optimizer is utilized for its effectiveness, aiding in more rapid convergence.

The training process incorporates the following hyperparameters:

- **Number of Epochs:** Set to 10, allowing the model to make ten complete passes through the entire training dataset. This gradual process aids in refining the model's accuracy over time.
- **Learning Rate:** Fixed at 10^{-4}. This value controls the optimization step size, has been choesn small enough not to diverge but still large enough to be able to converge at the minimum.
- **Weight Decay:** Also set at 10^{-4}. This regularization technique helps prevent overfitting by penalizing larger weight values, thereby encouraging a simpler model structure.

We conducted the training of our model using a satellite image of the Tamale region, choosing this smaller area to expedite the training process. Our goal was to test the effectiveness of our methodology on a smaller scale initially, with the intention of scaling up if the approach proved to be reliable.

IV. Results

In evaluating the performance of our model, we employed the F1-score as the primary metric, considering the imbalanced nature of our data. Additionally, we visually inspected the model's predictions by comparing the generated mask with the road data from OpenStreetMap (OSM).

Unfortunately, the results are inconclusive. The model struggles to accurately detect roads in the selected area. We observed that the use of batch normalization, as well as adjustments in hyperparameters, did not significantly impact the performance.

Figure 3 illustrates the result for the same patch discussed in Section III-A, highlighting the model's prediction in red and the OSM ground-truth in green. The outcome is unsatisfactory and not practically applicable. Despite various modifications to the model's parameters, we were unable to achieve improved results.

Figure 4 shows the loss and the F1-score for the training and the validation sets during 10 epochs. As we can see, the values are quite unstable.

The potential reasons for these shortcomings are further explored in the Discussion section V.

V. Discussion

As indicated in Section I, our project yielded inconclusive results. Our attempt to segment roads from Landsat satellite imagery using a U-Net-based convolutional neural network was more challenging than initially expected. Several hypotheses could explain these unexpected outcomes.

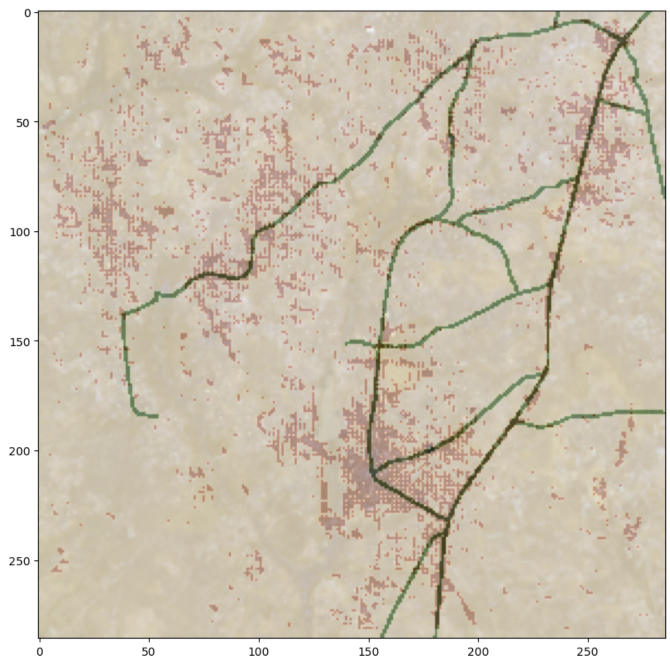

Figure 3. Result of the road segmentation with our model on one patch

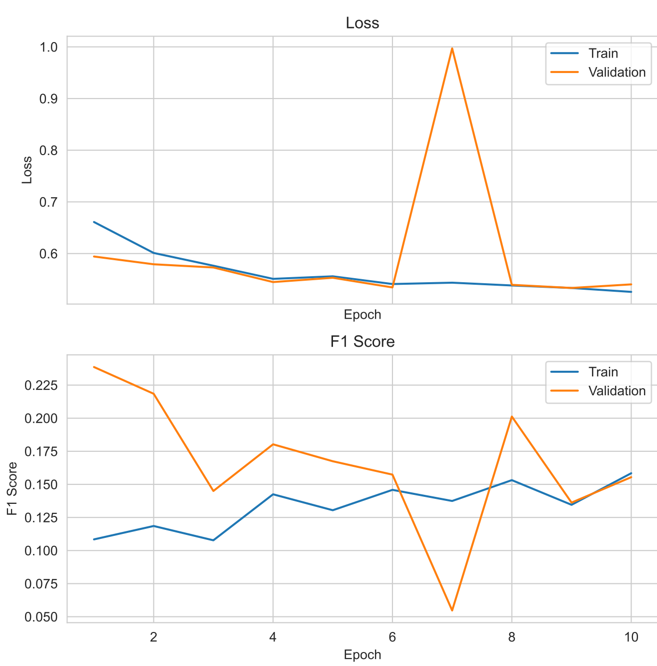

Figure 4. Loss and F1-score for training and validation sets during 10 epochs

Research in deep learning for road segmentation indicates that achieving effective results generally requires high-clarity satellite images with considerable zoom levels [6][7]. Most successful projects in this area have concentrated on urban settings with well-defined, asphalted roads, characterized by a distinct grid-like pattern and a high road-to-non-road ratio. This is significantly different from the conditions we encountered in our study.

For our project, we utilized raw satellite images where roads were often less distinct and sometimes barely visible. This lack of clarity added complexity to our task. While the time and effort invested in this project suggest that using Landsat imagery for road identification is more complicated than we initially thought, it doesn't necessarily mean that it's unattainable with this type of data.

Given these complexities, further investigations might include:

- Exploring higher resolution imagery: using satellite images with higher resolution may provide the additional detail necessary for more accurate road identification.
- Enhanced preprocessing techniques: developing more sophisticated image preprocessing methods to enhance road visibility in Landsat images could be beneficial.
- Algorithm modification: adjusting the neural network architecture or experimenting with different machine learning models might yield better results.
- Data fusion: combining Landsat data with other data sources, such as higher-resolution commercial satellite imagery or aerial photography, might improve the model's accuracy.

On the other hand, our use of OpenStreetMap (OSM) data to monitor changes in road networks over time has been effective. This approach is dependable and provides significant insights. However, it's important to note the possibility that roads identified as 'new' in OSM could be recently mapped rather than newly constructed. To fully verify the accuracy of this method, an in-depth analysis involving comprehensive statistical evaluation and validation would be required.

VI. Summary

The aim of this paper is to implement an effective model for road segmentation purposes, such as a UNet can be, and test it on a real dataset, in order to determine the road network expansion. Our dataset is based on a combination of satellite imageries and human-labelled roads in Ghana. Even though the model is widely used in literature, in our case it didn't perform properly and roads have not been accurately or precisely identified. Therefore the road expansion has been determined based on OSM data. We believe that by addressing specific problems regarding the input data and by refining the model, it could be possible to achieve satisfying results.

Acknowledgements

We thank Dr. Kahtrin Durizzo for the interesting project.

References

[1] S. Aggarwal, "The long road to health: Healthcare utilization impacts of a road pavement policy in rural india," *Journal of Development Economics*, vol. 151, p. 102667, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304387821000468

[2] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241.

[3] C. Müller-Crepon, P. Hunziker, and L.-E. Cederman, "Roads to rule, roads to rebel: Relational state capacity and conflict in africa," *Journal of Conflict Resolution*, vol. 65, no. 2-3, pp. 563–590, 2021, pMID: 33487734. [Online]. Available: https://doi.org/10.1177/0022002720963674

[4] C. Müller-Crepon, "African road network data — Carl Müller-Crepon — carlmueller-crepon.org," http://www.carlmueller-crepon.org/project/road_data/, [Accessed 10-12-2023].

[5] S. Aggarwal, "The long road to health: Healthcare utilization impacts of a road pavement policy in rural india," *Journal of Development Economics*, vol. 151, p. 102667, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304387821000468

[6] "GitHub - satellite-image-deep-learning/techniques: Techniques for deep learning with satellite & aerial imagery — github.com," https://github.com/satellite-image-deep-learning/techniques, [Accessed 02-12-2023].

[7] H. Ghandorh, W. Boulila, S. Masood, A. Koubaa, F. Ahmed, and J. Ahmad, "Semantic segmentation and edge detectionmdash;approach to road detection in very high resolution satellite images," *Remote Sensing*, vol. 14, no. 3, 2022. [Online]. Available: https://www.mdpi.com/2072-4292/14/3/613

Appendix

The algorithm developed aims for a more detailed road recognition and therefore a better estimation of the road expansion than the one that can be obtained with OSM data. This model represents a step of a wider research that aims to relate the roads expansion to the accessibility to hospitals and health facilities in african under-developed and developing countries; A similar investigation has been conducted in India in 2021 by Shilpa Aggarwal in [1].

The stakeholders in our project encompass the entire population of Ghana, with a particular emphasis on those who are less economically privileged and residents of small, remote villages that experience a greater impact from the construction and improvement of a more organized network.

Considering the broader project goal related to healthcare facilities and reducing postnatal child mortality, the consequences of a biased and unethical algorithm are exceedingly significant.

The ethics problem that thus persists in our project would be the overrepresentation of more populated areas in the data. This problem consists of two connected issues:

- **Focusing too much on cities:** Due to the prevalence of a significant number of roads in larger cities like Tamale, Ghana, the algorithm may exhibit a bias towards more efficient road recognition in these urban areas, potentially overlooking the expansion efforts in smaller villages. If this bias goes unaddressed and the algorithm's output is used to formulate a road expansion plan, it could lead to an undue emphasis on developing smaller villages whose growth potential was underestimated by the algorithm. Consequently, the resulting development plan may lack robustness and effectiveness.
- **Labels from OSM data:** Pixel labels are derived from OSM data, which are subject to user edits. This implies that in sparsely populated areas, certain roads might not be adequately documented in the database. This inherent bias in the algorithm could have adverse implications for future planning efforts.

  As mentioned before in the report it is hard to evaluate or measure directly how severe this issue of overrepresentation is. It is hard to obtain concrete data about this, which is the main barrier for taking this riks into account. We did not take this risk into account because we did not obtain a working model and therefore it would not be logical to focus our efforts on the ethical risks associated with our model.