Road Extraction and Routing from Satellite Imagery by Image Segmentation using Deep Learning

Mohammed Arebi[1]

[1] Furtwangen University of Applied Science
mohammed.fateh.arebi@hs-furtwangen.de

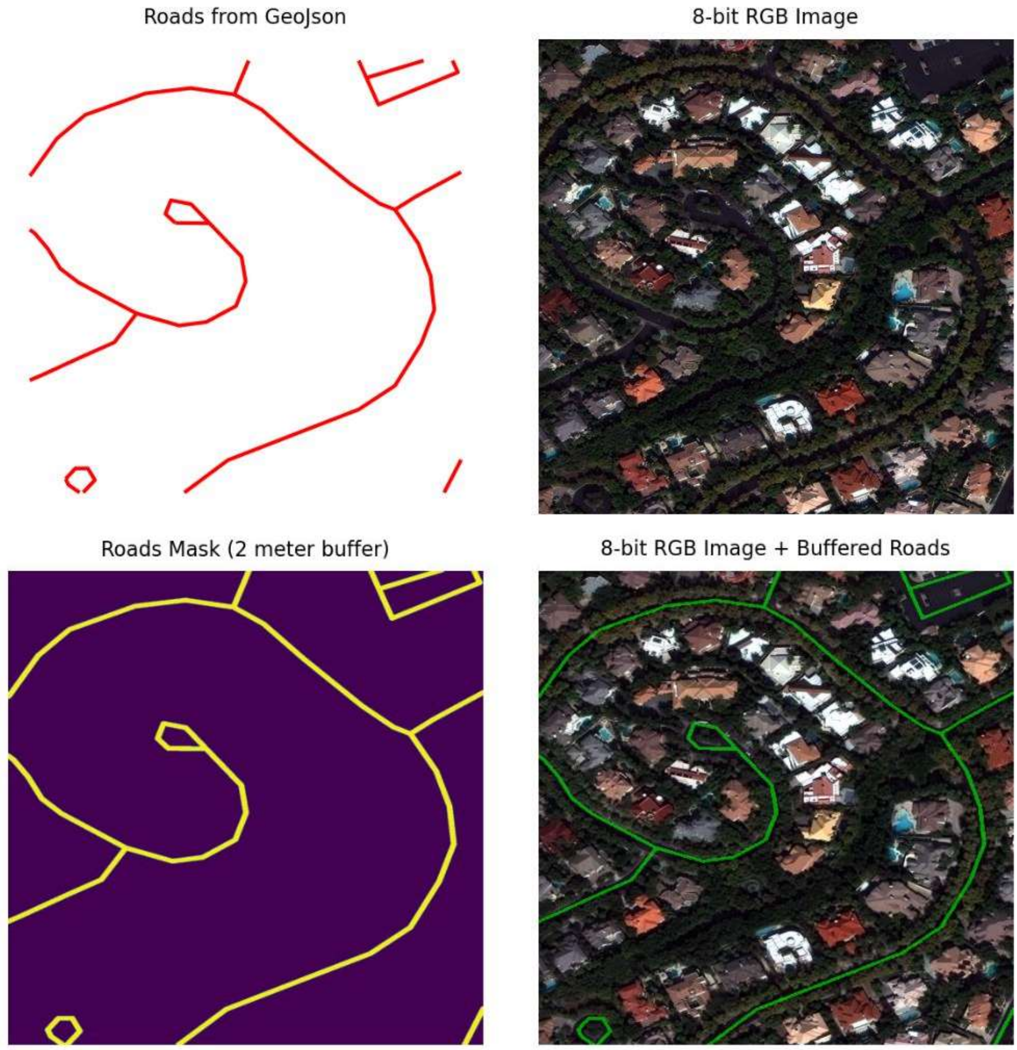
Abstract. In this thesis, we address the challenging task of interpreting large-scale satellite imagery by developing an automated system for generating semantic road maps and road graphs with speed limit predictions to enable efficient routing. We explore various convolutional deep neural networks, such as ResNet34, ResNet50, SeResNetX50, and InceptionV3, and conduct extensive studies on hyperparameters and loss functions to optimize the road extraction process. Our pipeline includes image pre-processing algorithms to handle varying image qualities, a model for road segment prediction, and post-processing techniques for graph extraction while retaining geographic information. The results demonstrate the effectiveness of our approach, showcasing the importance of appropriate model selection and optimization. The integration of graph extraction and geographic information enhances the routing process. Overall, this research contributes valuable insights into road extraction and routing from satellite imagery using deep learning, laying the groundwork for future advancements in this field.

Keywords: Satellite Imagery, Road Extraction, Deep Learning, Graph Extraction, Routing

1 Introduction

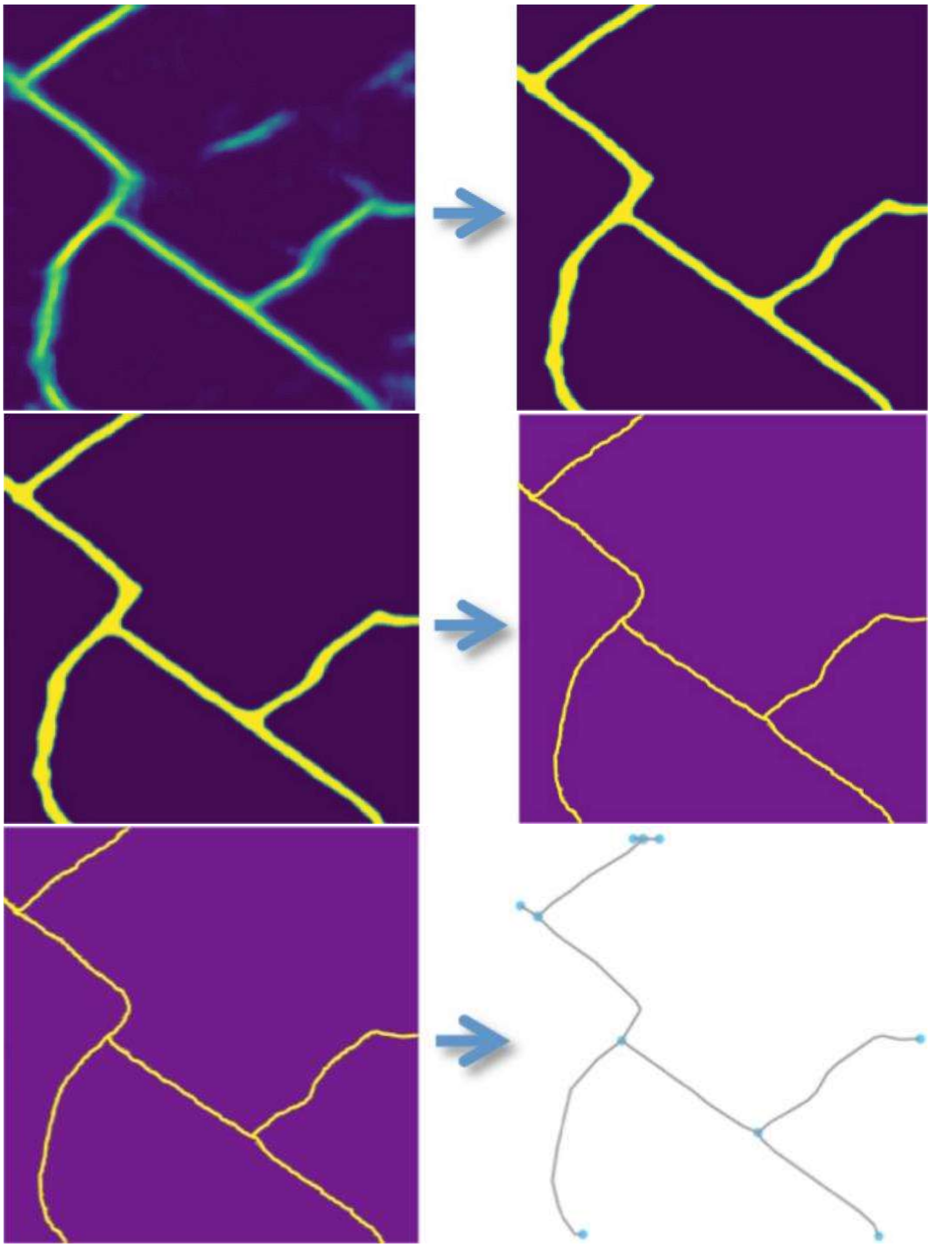
A stark increase in the amount of satellite imagery available in recent years has made the interpretation of this data a challenging problem at scale. Such images require a deep comprehension of the information contained in them to yield helpful insights. By creating an automated system for generating semantic maps of roads and highways and then further to road graphs with speed limit prediction, allows for routing in satellite images, this thesis investigates the aforementioned issue. Various convolutional deep neural networks were constructed, put into practice, and experimentally tested in order to solve the problem as a supervised machine learning task. We investigate the use of machine learning methods trained on aligned satellite images and possibly outdated maps for labelling the pixels of a satellite image with semantic labels as presented in Figure 1. For this, publicly accessible datasets and frameworks are employed. The resulting pipeline includes image pre-processing algorithms that allows it to cope with input images of varying quality, resolution, and channels, a model that predicts road segments from a satellite image and image post-processing algorithms to extract the road graph while retaining geographic information for efficient routing.

Figure 1. An example of the segmentation mask generating script's output. The raw GeoJSON label is shown in the top left, while the equivalent 8-bit RGB picture is shown in the upper right. The lower-left illustrates the output of the script: the pixel mask inferred from the GeoJSON label. The road mask is overlaid over the RGB picture in the lower right.



The study explores the efficacy of various U-Net variants with different backbone architectures, including ResNet34, ResNet50, SeResNetX50, and InceptionV3, among others to optimize the road extraction process. Additionally, a range of ablations studies and hyperparameter tests were conducted, examining for example different loss functions such as combinations of Binary Cross Entropy (BCE) loss and Dice coefficient, Focal loss, and Dice coefficient, as well as Weighted Tversky loss, mask road buffers, learning rate scheduling and network architectures. The application of graph extraction from the predicted segmentation map, along with the retention of geographic information, proved beneficial in the subsequent routing process. To achieve this, skeletonization and graph cleaning techniques were employed as shown in Figure 2. Speed prediction was also further developed on top of the road segmentation and graph extraction shown in Figure 3.

Figure 2. Prediction for a single mask channel before enhancement (left) and after enhancement (right) in the top row. Transformation of the enhanced prediction mask (left) to a skeleton (right) in the middle row. The skeleton (left) used to generate the graph structure (right) in the final row.



The study commenced with the collection and preprocessing of satellite imagery data. The U-Net model architecture was then implemented, serving as the backbone for road extraction. Various training procedures and hyperparameter optimization techniques were applied to enhance model performance. Evaluation metrics were carefully selected to assess the accuracy and efficiency of the road extraction process.

Figure 3. Left: sample training image in Vegas. Middle: typical binary training mask. Right: continuous mask where the mask value is proportional to the value of interest (in this case, speed limit).

Furthermore, the extracted road network underwent graph extraction, ensuring the preservation of critical geographic information. This facilitated efficient routing capabilities, empowering users to navigate the road network effectively. Skeletonization techniques were utilized to simplify the road network representation, while graph cleaning methods were employed to remove artifacts and enhance the accuracy of the road network.

The experimental results demonstrate the efficacy of the proposed approach. Comparative analysis of different backbone architectures revealed variations in performance, highlighting the importance of selecting appropriate models for road extraction. The hyperparameter tests and evaluation of loss functions, learning scheduling, classification thresholding and ensembling among others, provided insights into their impact on the quality of road extraction results. The utilization of graph extraction and geographic information retention significantly improved the routing process, demonstrating the practical applicability of the developed system. The skeletonization and graph cleaning techniques further enhanced the accuracy and reliability of the road network representation.

In conclusion, this research contributes to the field of road extraction and routing from satellite imagery by utilizing deep learning and image segmentation approaches. The findings underscore the importance of selecting suitable U-Net variants and backbone architectures, as well as optimizing hyperparameters and loss functions for accurate road extraction. The integration of graph extraction, geographic information retention, skeletonization, and graph cleaning techniques plays a pivotal role in achieving reliable road networks and efficient routing. Future research directions may involve exploring additional deep learning models and further refining the proposed methods to enhance performance and address emerging challenges. Figure 4 shows an example for the final pipeline for a sophisticated example image.

Figure 4. Example speed limit prediction incorporated in the predicted graph for an image in the test set, which is then overlaid on the test image.

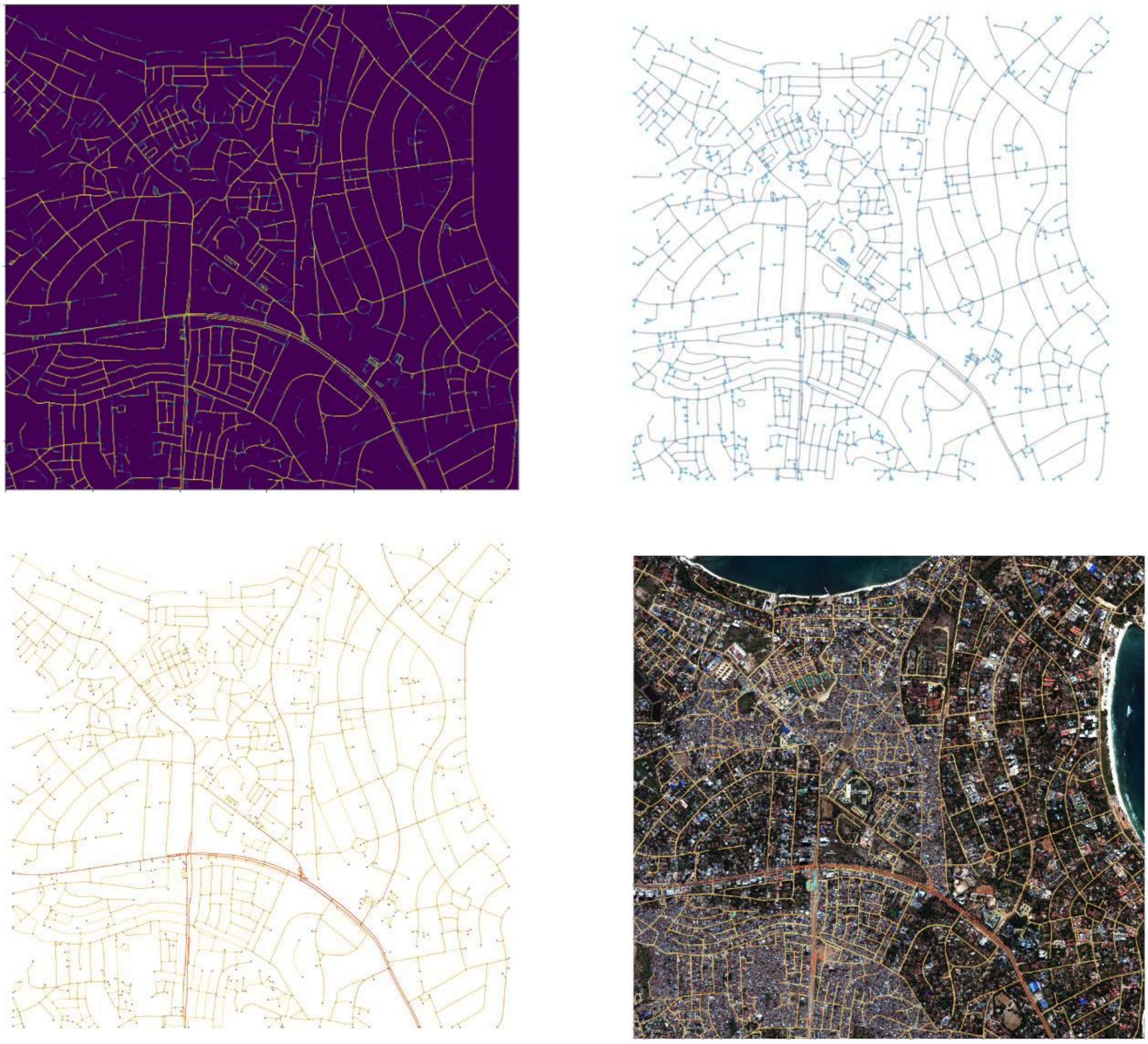


Figure 5. Optimal route based on either distance or travel time using the final proposed speed graph for the test image.