RoadCraft: Automated Road Segmentation using Aerial Images

Rishika Hazarika[1], Hillol Pratim Kalita[1]

[1]Jorhat Engineering College

March 31, 2025

HACKVITA 3.0 HACKATHON 2025

***Abstract*— Road segmentation in aerial images is a crucial task for urban planning, autonomous navigation, and infrastructure monitoring. Existing mapping solutions rely heavily on manual annotations, making frequent updates resource-intensive. This project presents a deep learning-based approach for road segmentation from aerial images using the Massachusetts Roads Dataset. Our model uses the U-Net architecture, a convolutional neural network designed for semantic segmentation and accurately identifies road networks in satellite imagery. This automation enhances road mapping processes. The fine trained model produces binary masks denoting road presence, enabling efficient and scalable road mapping solutions.**

I. INTRODUCTION

Semantic segmentation is the process of classifying each pixel of an image into distinct classes using deep learning. Automated road segmentation from aerial images has significant applications in smart city development, disaster response, and autonomous systems. Current methodologies depend heavily on human annotators, which is both time-consuming and labor-intensive.

Accurate and efficient road detection from aerial imagery remains a critical challenge in urban planning, disaster response, and autonomous navigation. Traditional methods often rely on manual annotation or computationally expensive techniques, limiting scalability and real-time applicability.

To address these challenges, we developed RoadCraft, a web-based application for road segmentation from aerial images using the encoder-decoder-based deep learning model U-Net. Built with FastAPI and a responsive Bootstrap front-end, it allows users to upload images and visualize road annotations in real-time. The project leverages the Massachusetts Roads Dataset and integrates advanced segmentation techniques for real-world applications.

Despite extensive research on road information extraction from aerial imagery, the task remains challenging due to noise, complex image backgrounds, and occlusions. Current technologies like Google Maps rely on manual labor to annotate and maintain their maps, making the process costly and time-consuming. Our project seeks to automate road detection in such maps, ensuring frequent updates and verification with minimal human intervention. By leveraging deep learning, RoadCraft enhances efficiency and accuracy in road segmentation, providing a scalable and accessible solution for real-time road mapping.

II. DATASET

The Massachusetts Roads Dataset serves as the primary dataset for model training and evaluation. This dataset comprises 1,171 high-resolution aerial images (1500x1500 pixels) with corresponding binary road masks. The dataset was divided into a training set (90%) and a validation set (10%) to effectively assess model performance.

III. TOOLS AND LIBRARIES

- **TensorFlow/Keras:** Model development and training.
- **Albumentations:** Data augmentation techniques.
- **Segmentation Models:** Implementation of advanced loss functions.
- **OpenCV, Numpy, Pandas, Matplotlib:** Supporting data preprocessing and visualization.
- **FastAPI:** Backend framework for wev deployment.
- **Bootstrap:** Front-end framework for responsive UI design.

IV. METHODOLOGY

A. Data Acquisition and Preprocessing

Our input data was aerial images paired with binary road masks. Preprocessing steps included:

- Resizing images to 512x512 pixels to optimize computational efficiency.
- Normalizing pixel values to a [0,1] range.
- Applying thresholding to binarize road masks for enhanced segmentation accuracy.
- Augmenting data with random horizontal flips and Gaussian blur to improve model generalization.
- Using `tf.data.Datasets` for efficient data loading and batching.

B. Model Architecture

We implemented the U-Net-based fully convolutional network for segmentation:

- **Encoder:** Downsampling using convolutional layers, batch normalization, and ReLU activation.
- **Decoder:** Upsampling via transposed convolutions with skip connections.
- **Final Layer:** Sigmoid activation function for binary segmentation mask output.

C. Feature Extraction and Training

- **Loss Function:** Dice coefficient loss, which helps mitigate class imbalance by focusing on overlap between predicted and ground-truth segmentation.
- **Evaluation Metrics:** Dice coefficient and Intersection over Union (IoU).

$$DiceCoefficient = \frac{2|y_{pred} \cap y_{true}|}{|y_{pred}| + |y_{true}|}$$

where $|y_{pred} \cap y_{true}|$ is the intersection of predicted and ground truth pixels, and $|y_{pred}| + |y_{true}|$ is the sum of their pixel counts.

$$DiceLoss = 1 - DiceCoefficient$$

- **Optimizer:** Adam optimizer with an initial learning rate of 5e-4, dynamically adjusted.
- **Training Configurations:**
  - Epochs: 15
  - Batch Size: 4
  - IoU score: **Training - 0.497**, **Validation - 0.415**

$$IoU = \frac{|y_{pred} \cap y_{true}|}{|y_{pred} \cup y_{true}|}$$

where y_{pred} is the predicted mask, y_{true} is the ground truth mask, $|y_{pred} \cap y_{true}|$ is the number of pixels in the intersection, and $|y_{pred} \cup y_{true}|$ is the number of pixels in the union.

D. Road Segmentation and Post-Processing

Predicted probability maps were thresholded to classify road pixels. Morphological operations were applied to remove noise and refine road segments. Connected components analysis was used to eliminate small, irrelevant detections, enhancing segmentation accuracy. This step improved the clarity of the road boundaries and reduced false positives.

V. Results and Discussion

The model was evaluted on a validation set:

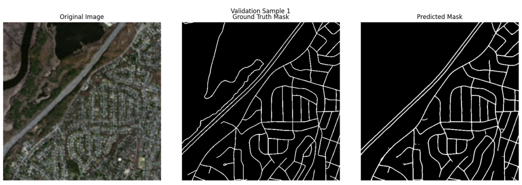

Fig. 1. Result 1

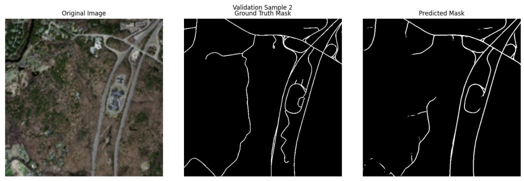

Fig. 2. Result 2

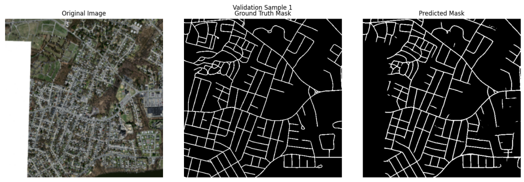

Fig. 3. Result 3

The results demonstrate that the model effectively segments major road structures, though it struggles with finer details and intersections, occasionally leaving smaller pathways incomplete. While the segmentation aligns well in most cases, minor deviations remain.

To assess training performance, we tracked loss and IoU over 14 epochs. The loss curve shows a significant decrease, stabilizing at around 0.3, while the Intersection over Union (IoU) score improves steadily, reaching approximately 0.5 for validation data. This indicates that the model generalizes reasonably well but still has room for refinement.

The following visualization represents the loss and IoU trends during training:

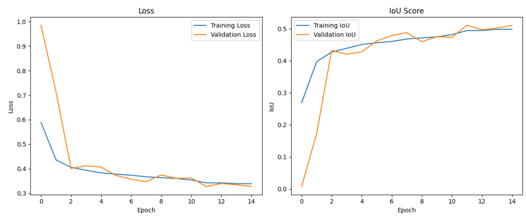

Fig. 4. Loss vs. IoU Score

Overall, the model achieves reasonable segmentation accuracy, with potential for improvement in detecting finer road structures.

VI. FastAPI Deployment for Road Segmentation

The Road Segmentation API is built using FastAPI, providing an endpoint to upload images and receive annotated road segmentations. Key features include:

- **Model Integration:** Loads the pre-trained U-Net model for segmentation.
- **File Handling:** Accepts image uploads, processes them, and stores results in structured directories.

- **Segmentation & Debugging:** Generates multiple versions of the segmented mask, including grayscale, binary, and colored overlays (yellow, blue, red, green) for visualization.
- **User Interface:** Serves a web-based UI using Jinja2 templates
- **Deployment:** Runs on a configurable port using Uvicorn.

The Road Segmentation API provides two key endpoints: `POST /predict/`, which accepts an image file, processes it using the pre-trained U-Net model, and returns an annotate segmentation overlay along with various debug images, and `GET /`, which serves the web-based user interface using Jinja2 templates. These endpoints ensure seamless image processing and visualization for road segmentation.

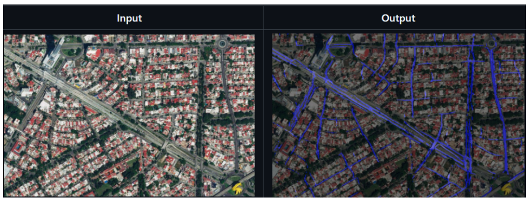

Fig. 5. Input vs output image

VII. CONCLUSION

This report presents a deep learning-based road segmentation model using the U-Net architecture. By leveraging aerial imagery and automated segmentation techniques, the model provides an efficient and scalable alternative to traditional road mapping methods. Despite the challenges posed by noisy backgrounds and occlusions, RoadCraft demonstrates the potential of deep learning in automating road detection. Future improvements could involve enhancing model robustness, incorporating additional datasets, and refining post-processing techniques to improve segmentation accuracy further.

REFERENCES

[1] U-Net: Convolutional Networks for Biomedical Image Segmentation - Olaf Ronneberger, Philipp Fischer, Thomas Brox (2015)
[2] Road Segmentation in Aerial Images by Exploiting Road Vector Data - Jiangye Yuan, Anil M. Cheriyadat (2013)
[3] Road Segmentation in SAR Satellite Images with Deep Fully-Convolutional Neural Networks - Corentin Henry, Seyed Majid Azimi, Nina Merkle (2018)