Road Extraction from RoadNet-based Satellite Imagery Using Classical Computer Vision Techniques
Final Project Report

Enzo PALOS
CentraleSupélec – ESSEC
95 rue Falguière, Paris
enzo.palos@student-cs.fr

Allister KOHN
CentraleSupélec – ESSEC
95 rue Falguière, Paris
allister.kohn@student-cs.fr

Emma SCHAPIRA
CentraleSupélec – ESSEC
95 rue Falguière, Paris
emma.schapira@student-cs.fr

Abstract

Road extraction from satellite imagery is a critical task with applications in urban planning, disaster management, and autonomous navigation. While deep learning approaches have demonstrated strong performance, their reliance on extensive labeled datasets, high computational costs, and limited interpretability present significant challenges. In contrast, classical computer vision techniques offer efficient, interpretable alternatives that do not require large-scale annotated data.

In this project, we propose a rule-based pipeline for road extraction using the RoadNet-based dataset. Our method integrates image enhancement (CLAHE, histogram equalization), edge detection (Sobel, Canny), and morphological operations to refine connectivity and suppress noise. Additionally, Hough Transform is employed to detect linear road structures. Preprocessing techniques, such as bilateral filtering and anisotropic diffusion, help reduce noise while preserving road boundaries.

In the classical approach, we primarily rely on qualitative analysis through visual inspection. In the machine learning approach, we compare predictions against ground truth using Intersection over Union (IoU). We further assess model reliability through qualitative analysis, including visual comparison of predictions with annotated maps. By demonstrating the effectiveness of traditional techniques, this work highlights the feasibility of lightweight and explainable solutions for road extraction, particularly in contexts with limited computational resources and annotated data.

1. Introduction

Road extraction involves identifying and delineating road networks from high-resolution satellite imagery, a process that plays a fundamental role in geographic information systems (GIS), urban planning, disaster management, and autonomous navigation. The ability to accurately extract roads is essential for updating transportation maps, optimizing traffic flow, planning infrastructure, and supporting emergency response efforts. With the growing availability of high-resolution satellite data from sources such as commercial Earth observation satellites and governmental agencies, there is an increasing demand for automated and computationally efficient road extraction methods.

Deep learning-based approaches have gained prominence in road extraction due to their ability to learn complex patterns and adapt to diverse environmental conditions. Convolutional Neural Networks (CNNs) and their variants have demonstrated strong performance in extracting road structures from satellite imagery. However, these methods require large-scale annotated datasets, high computational resources, and often function as black-box models, making them difficult to interpret. Additionally, the manual annotation of training datasets is a time-consuming process, and deploying deep learning models remains challenging in low-resource environments.

To address these challenges, this project investigates the feasibility of using classical computer vision techniques for road extraction. Traditional methods, based on image processing and feature extraction, provide advantages in terms of interpretability, efficiency, and adaptability in scenarios with limited computational resources. By leveraging edge detection (Sobel, Canny), morphological operations, and road segmentation techniques, we aim to develop a structured pipeline capable of extracting road networks with high precision while maintaining computational efficiency.

Our proposed approach consists of multiple processing steps, including:

- Contrast enhancement (using CLAHE and histogram equalization) to improve road visibility.
- Edge detection (via Sobel and Canny) to highlight road structures.
- Morphological operations (e.g., closing, dilation…) to refine connectivity and reduce noise.

- Hough Transform to detect and enhance linear structures corresponding to roads.
- Filtering of non-road regions (e.g., vegetation and buildings) using color segmentation in the HSV space.

The effectiveness of our approach will be evaluated using the RoadNet-based dataset, comparing extracted road networks against ground truth annotations. The classical approach will be evaluated through qualitative visual inspection, while for the machine learning approach, Intersection over Union (IoU) will be used as the primary quantitative metric. By demonstrating the feasibility of classical computer vision techniques, this work highlights a lightweight and interpretable alternative to deep learning for road extraction, particularly in resource-constrained environments.

2. Motivation

Our primary motivation for this project was to apply and deepen our computer vision skills by tackling a real-world problem—road extraction from satellite imagery. This task presents significant challenges, such as handling image noise, occlusions, and fragmented road structures, which require a solid understanding of image processing techniques. Unlike deep learning methods, classical approaches demand careful tuning of filters, edge detectors, and morphological operations to achieve robust results.

Throughout our course, we have studied image filtering, feature extraction, segmentation, and morphology, which provided the foundation for our pipeline. This project allowed us to integrate these concepts by experimenting with different contrast enhancement techniques (CLAHE, histogram equalization), edge detection methods (Sobel, Canny, Gabor filters), and road segmentation strategies (Hough Transform, HSV color filtering, morphological operations). By iteratively refining these steps, we developed an efficient and interpretable approach to extract road networks from satellite imagery.

Beyond the technical challenge, road extraction has numerous practical applications. It plays a key role in updating GIS databases, supporting autonomous navigation systems, and aiding disaster response and urban planning. Given the computational constraints of certain organizations (e.g., government agencies, non-profits, and research institutions), a lightweight, rule-based method can provide a valuable alternative to deep learning approaches, which require extensive labeled datasets and high-performance hardware. Our approach demonstrates that traditional computer vision techniques can still deliver meaningful results in data-scarce environments.

Moreover, this project has been a valuable exercise in critical thinking, problem-solving, and algorithmic optimization. By decomposing the problem into modular steps, systematically tuning hyperparameters, and evaluating our results using Intersection over Union (IoU) and qualitative analysis, we enhanced our ability to design structured and efficient pipelines. These skills are highly transferable and applicable in various industries, including healthcare, industrial automation, and robotics, where classical vision techniques remain relevant.

Ultimately, this project has served as both a technical challenge and an opportunity to explore the practical impact of computer vision. By demonstrating the effectiveness of interpretable, computationally efficient road extraction methods, we have strengthened our expertise in applied image processing and algorithmic design.

3. Problem Definition

3.1 Definition

Given high-resolution satellite images, our objective was to accurately extract and delineate road networks using classical computer vision techniques. The key challenges we addressed include:

- **Continuity and Connectivity:** Ensuring that the extracted road network remains continuous and well-connected, especially in urban environments where occlusions (e.g., shadows, buildings, or vegetation) can disrupt road detection. To handle this, we applied morphological operations (e.g., closing, dilation…) to refine road connectivity.

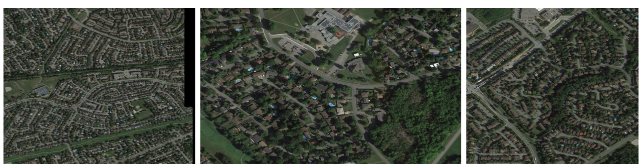
- **Noise and Variability:** Satellite images often contain noise due to compression artifacts, varying illumination, and occlusions (e.g., clouds, shadows, or tree canopies). We addressed this challenge using noise reduction techniques, such as bilateral filtering, median filtering, and anisotropic diffusion, to enhance feature detection.

- **Edge Discontinuities:** Roads may appear fragmented due to weak edge responses, which can make their detection difficult. To bridge these gaps, we applied edge detection methods (e.g., Sobel, Canny) combined with Hough Transform for structured road representation.

- **Parameter Sensitivity:** Classical methods are often highly sensitive to threshold values and filter parameters. To improve adaptability, we experimented with dynamic thresholding techniques, such as CLAHE-based contrast enhancement and adaptive binarization, to fine-tune the pipeline based on image characteristics.

- **Scalability and Efficiency:** The proposed method is designed to be computationally efficient, ensuring that large-scale satellite imagery can be processed without requiring extensive computational resources. Our pipeline optimizes image preprocessing and feature extraction steps to make it applicable in real-world GIS applications.

By addressing these challenges, our goal was to develop a robust and interpretable pipeline capable of extracting roads accurately from diverse satellite imagery datasets. This project was particularly challenging due to the need for handcrafted feature engineering, adaptive thresholding, and morphological processing to ensure robustness against variations in lighting, occlusions, and structural discontinuities, all while maintaining computational efficiency for large-scale datasets.



Some images of the dataset with varying formats and zoom levels

3.2 Related Work

Traditional Approaches

Early research on road extraction focused on handcrafted pipelines using morphological operations, edge detection, and linear connectivity analysis. For instance, Mena and Gamba (2007) employed mathematical morphology to progressively isolate candidate road segments, refining them through connectivity constraints in challenging urban settings. Mena and Bostrom (2010) further combined threshold-based intensity selection with geometric constraints to ensure road network continuity. Syed and Naeem (2014) provided a broad overview of classical strategies—ranging from Canny edge detection to region-growing techniques—for handling issues like shadows, occlusions, and fine-scale variations in aerial imagery. Chen et al. (2016), while proposing a framework called RoadNet, also underscored the potential of purely feature-based approaches—such as contrast adjustment and line detection—for addressing noise in high-resolution images. Despite their efficiency and interpretability, these traditional methods often exhibit high sensitivity to parameter choices (e.g., thresholds, filter sizes) and can struggle in diverse environments with fragmented or partially occluded roads.

Non-Deep Machine Learning

Subsequent works explored moderate machine learning techniques to improve adaptivity. Lu and Weng (2007) investigated supervised classifiers—such as Support Vector Machines (SVMs), decision trees, and k-nearest neighbors (k-NN)—for distinguishing roads from surrounding objects based on statistical texture and reflectance features, achieving more robust results than purely rule-based pipelines. Ensemble methods (e.g., Random Forest) proved effective in handling heterogeneous cityscapes, as they can combine multiple weak learners to accommodate varied lighting conditions and noise levels. Although these approaches are typically more data-driven than classical filtering, they still rely on carefully engineered feature sets and can require substantial tuning to maintain accuracy across different geographic regions.

4. Methodology

4.1 Images Collection

The dataset used in this project consists of 20 high-resolution satellite images from the Ottawa dataset, featuring variations in format, resolution, and zoom levels, along with corresponding ground truth road masks.

The dataset, which is 905 MB, can be downloaded [here].

The Ottawa dataset is organized into 20 numbered subfolders (1 to 20), each containing:

- A satellite image in TIFF format
- A ground truth road mask in PNG format, used for road detection validation

Each image was processed using OpenCV and NumPy, and the ground truth segmentation masks were used to evaluate road extraction performance.

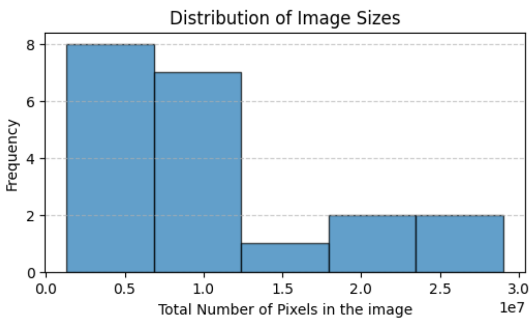
4.2 Image Analysis

Before designing the processing pipeline, a preliminary analysis of the satellite images was conducted to understand their characteristics. These images exhibit significant variations in resolution, contrast, and sharpness, which introduce challenges in road detection. Identifying and addressing these variations early on ensures a more effective and consistent preprocessing strategy.

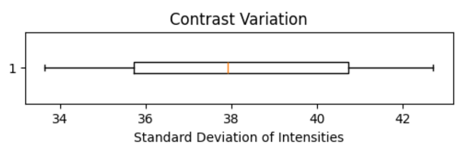
4.2.1 Resolution and Size Variation

Our analysis revealed several key aspects that need to be addressed in the preprocessing phase. The dataset contains images of different sizes, ranging from 1280x1024 pixels for the smallest to 5969x4864 pixels for the largest. This variability makes it necessary to apply normalization techniques to ensure consistency across images and optimize the performance of the detection models.

The distribution of image sizes is highly uneven, with a majority of smaller images but also a few very large ones, which could create biases in our road detection pipeline.

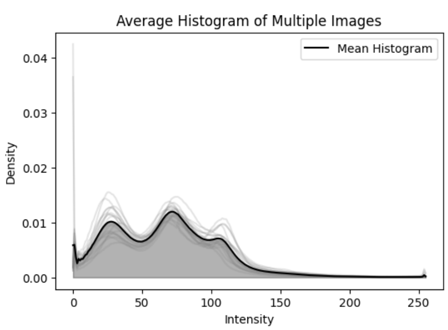


4.2.2 Contrast Distribution



The contrast variation, measured through the standard deviation of pixel intensities, indicates a limited range of lighting conditions and atmospheric effects.

4.2.3 Intensity Histogram Observations



Additionally, the average intensity histogram shows distinct peaks, suggesting that certain brightness levels are more common, which may reflect differences in urban landscapes, materials, or sensor properties.

4.2.4 Sharpness Evaluation

Another critical aspect is image sharpness, assessed through Laplacian variance, which highlights significant differences in clarity among the images. The Laplacian variance is a commonly used metric for sharpness assessment, where higher values indicate sharper images and lower values suggest blurriness.

From the distribution, we observe that:

- The majority of images have a Laplacian variance between 2500 and 3500, indicating a moderate level of sharpness.
- A few images exhibit significantly higher sharpness values (above 4500), suggesting well-defined edges.

4.2.5 Takeaways for Preprocessing

Overall, these observations emphasize the complexity of our dataset and the importance of careful preprocessing. Standardizing image sizes, adjusting contrast, and addressing sharpness issues will be key to ensuring that our methods can finally accurately and consistently detect roads across diverse urban environments.

4.3 Preprocessing

4.3.1 White Balancing

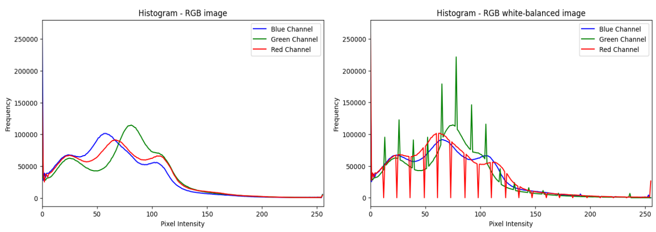
To ensure consistent color distribution across the dataset, we applied a **white balancing technique** based on the gray-world assumption. This step was essential to mitigate variations in lighting conditions and color discrepancies, which could otherwise impact road segmentation accuracy. By normalizing color representation, we ensured that roads remained distinguishable from their surrounding environment.
The **Gray-World Algorithm** assumes that in a well-balanced image, the average color across all channels (R, G, B) should be neutral gray. To achieve this, each channel's intensity $I_c'(i,j)$ is scaled based on the ratio of the global average intensity to the mean intensity of that channel:

$$I_c'(i,j) = \min\left(I_c(i,j) \times \frac{\mu_{\text{gray}}}{\mu_c}, 255\right)$$

where:

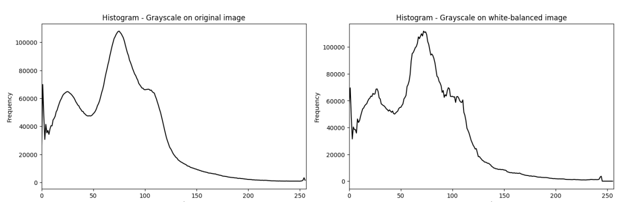
- $\mu_c = \frac{1}{N}\sum_{i,j} I_c(i,j)$ is the mean intensity of channel c across the image,
- $\mu_{\text{gray}} = \frac{\mu_R+\mu_G+\mu_B}{3}$ is the overall mean intensity across all three channels.

This scaling adjustment ensured that the color balance is more natural and consistent, reducing biases caused by illumination differences. As a result, roads appeared more uniform across different images, improving the effectiveness of subsequent segmentation and feature extraction steps.



4.3.2 Conversion to grayscale

Following white balancing, we converted each image to **grayscale** to simplify feature extraction. Grayscale conversion helped eliminate redundant color information, focusing the analysis on intensity variations, which are more relevant for edge detection and filtering operations. Applying the white-balancing method before grayscale conversion led to better image quality, even though the difference between the two grayscaled images was barely perceptible to the naked eye.



To further enhance contrast and improve feature visibility, we employed **histogram equalization**, which redistributed pixel intensities to enhance road structures against the background.

After applying white balancing, we converted each image to **grayscale** to simplify feature extraction. Grayscale conversion eliminates redundant color information, allowing the analysis to focus on intensity variations, which are more relevant for edge detection and filtering operations.
Applying white balancing before grayscale conversion resulted in better image quality, ensuring that brightness and contrast were more evenly distributed. Although the difference between the original grayscale image and the white-balanced grayscale image was subtle to the naked eye, this preprocessing step helped enhance the performance of subsequent image processing techniques.

4.3.3 Histogram Equalization

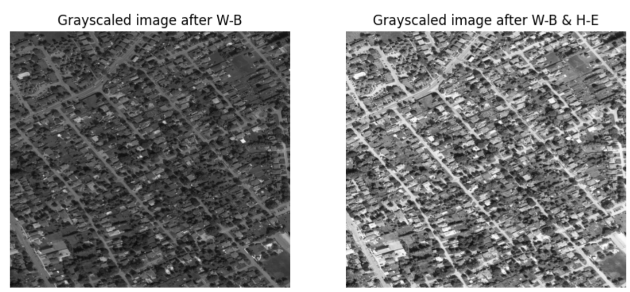
To further enhance contrast and improve feature visibility, we applied **histogram equalization**, a technique that redistributes pixel intensities to make road structures more distinguishable from the background.
Histogram equalization is a widely used method in grayscale image processing that adjusts pixel intensities to achieve a more uniform distribution. This transformation enhances important features—such as roads and boundaries—that might otherwise be lost in low-contrast or underexposed regions of satellite imagery. Formally, for each pixel intensity r in an image with a maximum gray level of $L - 1$, the mapping function is defined as:

$$s = T(r) = (L-1)\sum_{j=0}^{r}\frac{p_r(j)}{N}$$

where:

- $p_r(j)$ represents the number of pixels with intensity j,
- N is the total number of pixels in the image.

By computing this cumulative probability distribution and assigning new intensity values s based on $T(r)$, the overall contrast of the image were improved, making roads and surroundings.

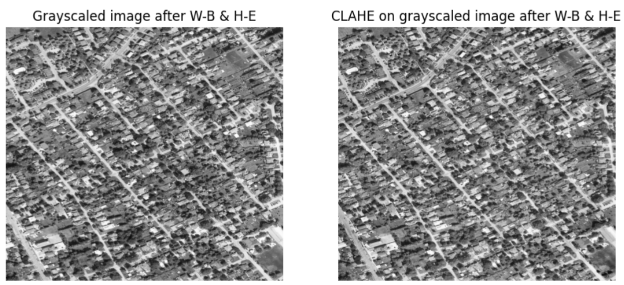


4.3.4 CLAHE

In addition to standard histogram equalization, we experimented with **CLAHE (Contrast Limited Adaptive Histogram Equalization)**, an advanced variation designed to enhance local contrast while preventing excessive noise amplification.

Unlike global histogram equalization, which applies a uniform transformation across the entire image, CLAHE works by dividing the image into small regions (tiles) and equalizing each one individually. This localized enhancement helps preserve details in darker and brighter areas, making it particularly effective for images with non-uniform lighting conditions. To prevent over-enhancement of noise, a clip limit is imposed on the histogram, ensuring that no single intensity range dominates the transformation.

In our dataset, many satellite images exhibited significant brightness variations due to shadows, atmospheric effects, or sensor limitations. By applying CLAHE, we improved road visibility while preserving fine details in urban landscapes. This enhancement made edge detection methods such as Sobel and Canny more effective, reducing the risk of missing faint road segments. Additionally, CLAHE helped mitigate issues related to washed-out or overly dark regions, ensuring that road boundaries remained clear and well-defined across diverse lighting conditions.



4.3.5 Anisotropic Diffusion / Median & Bilateral Filtering

To reduce noise while preserving road structures, we applied anisotropic diffusion filtering, a technique that smooths homogeneous regions selectively while retaining sharp transitions. This method was particularly effective in maintaining road continuity, especially in areas where roads appeared fragmented due to noise, occlusions, or environmental factors such as vegetation and cloud shadows.

In parallel, we applied bilateral and median filtering to further refine image quality:

- Bilateral filtering preserved edge integrity while reducing noise by considering both spatial proximity and intensity similarity of pixels.
- Median filtering effectively removed salt-and-pepper noise while maintaining the structure of linear features such as roads.

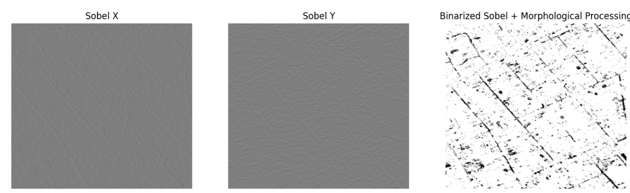
By combining these filtering techniques, we ensured that the images remained clean and well-defined, improving the performance of subsequent edge extraction steps.

4.4 Road Feature Extraction

To accurately delineate road structures in satellite imagery, we employed a combination of edge detection techniques, specifically **Sobel** and **Canny** edge detection. These methods allowed us to highlight road boundaries and refine the detected edges to achieve a clearer road network representation.

4.4.1 Sobel Edge Detection

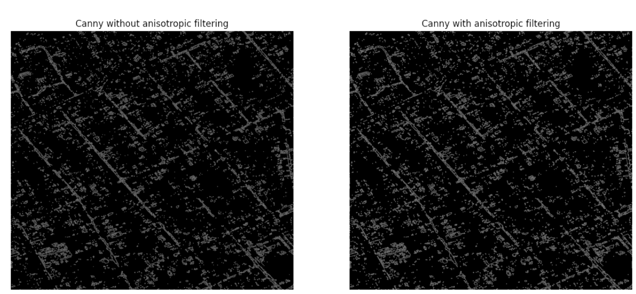
The **Sobel filter** is a gradient-based operator that computes intensity changes in both horizontal and vertical directions. It uses convolution kernels to detect regions of rapid intensity variation, which typically correspond to edges and boundaries in an image. In our application, Sobel filtering provided an initial approximation of road structures by enhancing edges where brightness changes significantly. However, due to its sensitivity to noise and texture variations, Sobel alone was insufficient for precise road extraction.



4.4.2 Canny Edge Detection

To refine the detected edges, we applied the **Canny edge detector**, a multi-stage algorithm designed to produce thin and well-defined edge maps while minimizing noise and false edges. The Canny process involves:

1. Gradient Computation – Similar to Sobel, gradients are calculated to detect areas of intensity change.
2. Non-Maximum Suppression – Weak edges that do not form a significant boundary are removed.
3. Double Thresholding & Edge Tracking – Edges are classified as strong, weak, or non-edges. Weak edges are retained only if connected to strong edges, ensuring robustness against noise.



We experimented with Canny edge detection with and without anisotropic filtering to assess its impact on road extraction. **Anisotropic filtering** smooths an image while preserving edges, potentially reducing noise interference. However, we observed that Canny without anisotropic filtering performed better in our context. This was because anisotropic filtering tended to oversmooth fine road details, causing some thin roads and weaker edges to be lost. In contrast, applying Canny directly on the pre-processed image maintained sharper road contours and improved the accuracy of detected roads.

4.4.3 Gabor Filter

Additionally, we attempted to incorporate the Gabor filter, a method particularly well-suited for detecting elongated and linear features at different orientations and scales. Our goal was to enhance road detection regardless of their directional alignment in the image. However, this approach did not prove to be conclusive in isolating roads from other structures with similar textures, such as rooftops or parking lots.

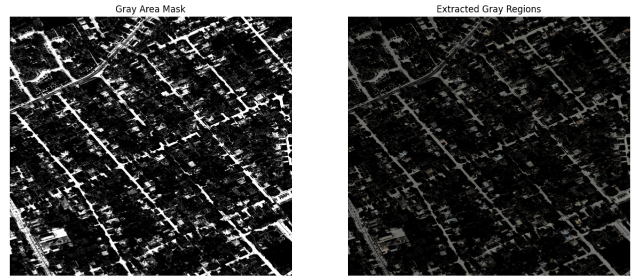
4.4.4 Masking techniques

To further improve road detection, we implemented masking techniques to isolate relevant regions of interest while filtering out non-road elements. The initial approach focused on preserving only grayscale regions, as roads often appear in shades of gray in satellite imagery. However, we observed that this mask also retained numerous buildings and rooftops, which introduced noise into the road detection process. To address this, we designed an additional house detection mask, allowing us to filter out buildings and enhance the clarity of detected roads.

A. Grayscale Mask for Road Extraction

Since roads predominantly exhibit gray tones, we applied a color thresholding approach in the HSV color space to retain only grayscale regions while removing vibrant or highly saturated areas such as vegetation and water bodies. This helped focus the detection process on potential road segments. However, despite improving the road visibility, this method also preserved gray rooftops and large concrete structures, which interfered with road continuity in our extraction pipeline.

We tested different value ranges to obtain the best possible mask. The goal of this step was to find the range that captures the maximum number of roads while minimizing the inclusion of houses. The range that seemed the most effective was 85–145.



B. House Detection Mask for Improved Filtering

To mitigate the issue of retained buildings, we introduced an additional mask specifically designed to detect and filter out houses and large structures. By leveraging color-based segmentation and morphological operations, we identified common rooftop colors (such as brown and certain shades of gray) and removed them from the grayscale mask. This step significantly reduced false positives, making it easier to isolate continuous road networks while preventing confusion between roads and rooftops.

4.5 Road Network Refinement

After extracting the primary road structures through edge detection and masking techniques, we applied additional refinement steps to ensure that the detected road network remained continuous, well-defined, and free of false detections. This phase was crucial in eliminating fragmented roads, reducing noise, and improving overall road connectivity. The two main techniques employed in this process were the Hough Transform for reinforcing linear structures and final thresholding & region-growing techniques to refine the extracted roads.

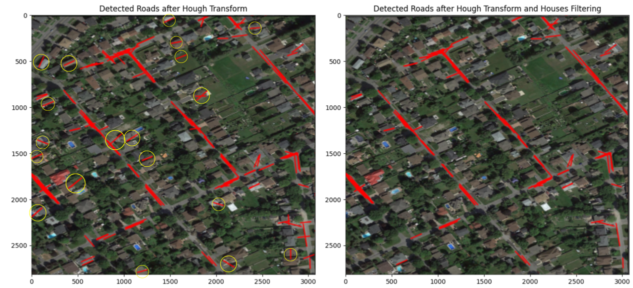
4.5.1 Hough Transform

The Hough Transform is a powerful technique used to detect linear features in images, making it particularly useful for reinforcing road segments that may have been partially detected due to noise, occlusions, or varying contrast levels. This method works by converting detected edges into a parametric space, where lines are identified based on the accumulation of votes. After applying Canny edge detection, the resulting edge maps contained numerous discontinuous road segments due to variations in lighting and occlusions. By applying the Hough Transform, we detected linear road structures and filtered out random edges that did not align with dominant road directions. This step significantly improved the consistency of detected roads, especially in regions where roads appeared as disjointed or faint lines in the edge-detected images. The application of the Hough Probabilistic Transform (`HoughLinesP`) was particularly useful, as it allowed us to identify and filter only meaningful road segments, ignoring short or insignificant lines that may have resulted from noise.

4.5.2 Final Road Extraction / Post-processing

To obtain the final road network, we refined the extracted roads using previously generated masks and applied post-processing techniques to ensure accuracy and remove non-road elements.

- Since paved roads predominantly appear in shades of gray, we applied our grayscale mask to retain road-like regions while removing non-road elements such as vegetation and water bodies.
- However, buildings and rooftops also often appear in similar gray tones**,** leading to potential false positives. To address this, we incorporated the house detection mask**,** which allowed us to eliminate structures that might have been misclassified as roads**.**



4.5.3 Performance Considerations and Limitations

While our approach effectively extracted road networks from satellite imagery, it faced several challenges. The parameter sensitivity of classical computer vision techniques required careful tuning, as threshold values and filtering parameters needed to be adjusted dynamically for different image conditions. Additionally, occlusions and shadows occasionally disrupted road continuity, making it difficult to reconstruct certain road segments accurately. Although morphological operations helped bridge some of these gaps, the absence of contextual understanding in classical methods meant that certain roads could still be misclassified.

Another challenge was scalability, particularly when processing high-resolution satellite imagery. While our pipeline was optimized for efficiency, handling large datasets remained computationally demanding. Future improvements could involve integrating machine learning-based methods to enhance adaptability and robustness across diverse datasets.

4.6 Machine Learning Methods

To detect roads in satellite imagery, we developed a comprehensive methodology covering data preparation, the training of supervised and unsupervised learning models, and an evaluation of their performance.

4.6.1 Data Preparation

As a first step, each satellite image and its corresponding ground-truth mask was read and resized using a scale factor of 6. This reduces computational requirements while preserving sufficient detail for road detection. Any discrepancies in image dimensions or band counts were harmonized so that all data could be processed consistently.

Once standardized, the dataset was split into 75% for training and 25% for testing. In this setup, every pixel in an image is treated as a single sample, while the associated ground-truth mask (indicating whether the pixel is part of a road) provides the label for supervised learning. Although this per-pixel approach yields a large number of training samples, it also demands substantial data handling and processing power.

4.6.2 Model Training

We considered different machine learning approaches for road detection, but only the Random Forest model was fully trained and evaluated.

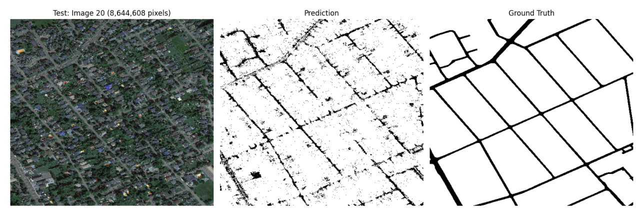
Supervised Learning:

- **Random Forest:** A powerful ensemble-based method that builds multiple decision trees and aggregates their predictions. Its resilience to high-dimensional input makes it well-suited to multispectral satellite data.
- **Support Vector Machine (SVM):** A kernel-based classifier (here, with an RBF kernel) that handles nonlinear relationships in the feature space. SVMs project pixels into a higher-dimensional space to enhance separability between road and non-road classes.
- **XGBoost:** A popular gradient boosting framework known for its speed and effectiveness on structured data. Various hyperparameters (such as tree depth and learning rate) were tuned to enhance its performance.

Unsupervised Learning:

- **DBSCAN (Density-Based Spatial Clustering of Applications with Noise)**: A density-based clustering method that detects high-density regions of pixels while treating outliers as noise. This can naturally reveal clusters that may correspond to roads.
- **Mean Shift:** A nonparametric clustering approach designed to locate density peaks within the feature space. It groups pixels around areas of high density, potentially highlighting continuous road networks.

Each model was trained and evaluated on the extracted pixel features. The supervised models used labeled ground truth data to learn patterns, while the unsupervised methods explored patterns without explicit annotations. By comparing the performance of these models, the most effective approach for road detection appeared to be Random Forest.



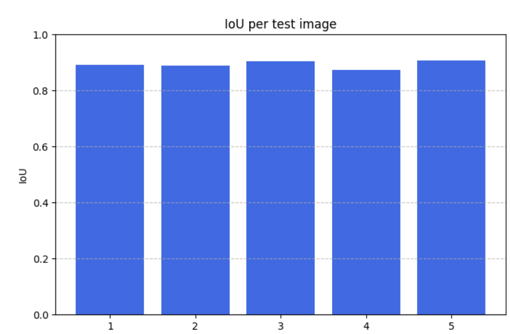
5 Evaluation and comparisons

5.1 Qualitative Assessment

We conducted a visual inspection of the segmentation results on the predictions made with the traditional computer vision methods to evaluate the model's performance. By overlaying the predicted masks on the original test images, we assessed their alignment with the ground truth. In most cases, the model successfully captured the primary structures of interest, demonstrating strong generalization across different test samples. However, we observed some misalignment in challenging areas, particularly in regions with low contrast or complex textures. These observations help us identify potential areas for improvement.

5.2 Quantitative Metrics

To quantify the machine learning model's performance, we computed the Intersection over Union (IoU) score for each test image. The results indicate that our model achieves consistently high IoU scores across the test set, with most values close to 0.9. This suggests a strong agreement between predictions and ground truth. The mean IoU across all test images was 0.8923, reflecting the overall accuracy of our segmentation approach.



5.3 Future Work

Although our model performs well on the given dataset, we have identified several areas for potential improvement:

- **Enhancing Data Augmentation**: Introducing more diverse augmentations could help improve generalization to unseen scenarios.
- **Refining Model Architecture**: Exploring more advanced architectures, such as transformer-based segmentation models, may further enhance accuracy.
- **Applying Post-Processing Techniques**: Leveraging morphological operations or CRF-based refinements could improve the precision of segmentation boundaries.
- **Expanding the Dataset**: Increasing the dataset size and diversity would provide a more robust evaluation and improve model generalization.

By implementing these improvements, we aim to further enhance the segmentation model's accuracy and robustness for real-world applications.

6 Conclusion

In this project, we demonstrated that a carefully designed classical computer vision pipeline can effectively extract roads from high-resolution satellite imagery, presenting an alternative to deep learning models that demand extensive labeled data and large computational resources. By combining image preprocessing (contrast enhancement, noise reduction, gray-scale masking) with edge detection, morphological operations, and the Hough Transform, we achieved robust results in qualitative evaluations. Additionally, our machine learning approach, based on Random Forest, achieved an average IoU close to 0.9.

A key advantage of our approach is its interpretability, as each stage of the pipeline can be tuned for specific imaging conditions without relying on large-scale training datasets. However, we also highlighted several challenges, including the sensitivity of classical methods to parameter choices (e.g., thresholds, kernel sizes), the lack of global contextual understanding, and scalability concerns when dealing with very large images. Future work could explore adaptive thresholding, graph-based connectivity refinements, or lightweight machine learning components to improve adaptability and reduce false positives. By striking a balance between efficiency, interpretability, and accuracy, this classical pipeline provides a viable option for road extraction in contexts where deep learning methods may be impractical or less transparent.

References

[1] Chen, D., Wang, C., Li, H., & Fu, P. (2016). *RoadNet: Learning to Comprehensively Analyze Road Networks in Complex Urban Scenes from High-Resolution Remotely Sensed Images*. IEEE Transactions on Geoscience and Remote Sensing.

[2] Mena, J. B., & Gamba, P. (2007). *Road extraction from aerial images by sequential process using mathematical morphology*. IEEE Transactions on Geoscience and Remote Sensing.

[3] Mena, J. B., & Bostrom, A. (2010). *Automated Road Extraction and Network Generation from High-Resolution Aerial Imagery*. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*.

[4] Syed, A., & Naeem, M. (2014). *A review on automated road extraction from high-resolution aerial imagery*. International Journal of Remote Sensing.

[5] Lu, D., & Weng, Q. (2007). *A survey of image classification methods and techniques for improving classification performance*. International Journal of Remote Sensing.

[6] Abraham, J., & Wloka, C. (2021). *Edge Detection for Satellite Images without Deep Networks*. arXiv preprint arXiv:2105.12633.

[7] Liu, Y., Yao, J., Lu, X., Xia, M., Wang, X., & Liu, Y. (2019). *RoadNet: Learning to Comprehensively Analyze Road Networks in Complex Urban Scenes from High-Resolution Remotely Sensed Images*. IEEE Transactions on Geoscience and Remote Sensing, 57(4), 2043–2056.

[8] Sanchez, B. (2022). *Automatic Detection and Estimation of the Area of Buildings*. Bachelor's Thesis, University of Twente, EEMCS.

[9] Van Etten, A. (2020). *City-scale road extraction from satellite imagery v2: Road speeds and travel times*. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* (pp. 1686–1694).