



A computer vision model for detecting road features in satellite images.

Problem Statement, Research Questions, and Objectives

Contextualization: “Public Road Mapping and Feature Extraction”

Research Questions:

- Is it possible to build a computer vision model, using data from a large city, that can generalize road segmentation across different cities?
- What are the limitations of this approach, and how can they be overcome to deliver a complete road mapping solution?

Objectives

Develop a tool capable of extracting information about road characteristics from satellite imagery, including features such as street and sidewalk length and width.

Contribution

The proposed tool addresses an existing gap by providing detailed road data that is currently not fully cataloged by platforms such as OpenStreetMap.

This innovation improves map accuracy and can support several urban planning applications, including traffic management, public transportation planning, emergency response, and event planning.

OpenStreetMap



“OpenStreetMap is developed by a volunteer community of mappers who contribute to and maintain up-to-date data about roads, railways, cafés, train stations, and much more across the world.”



OpenStreetMap Road Information

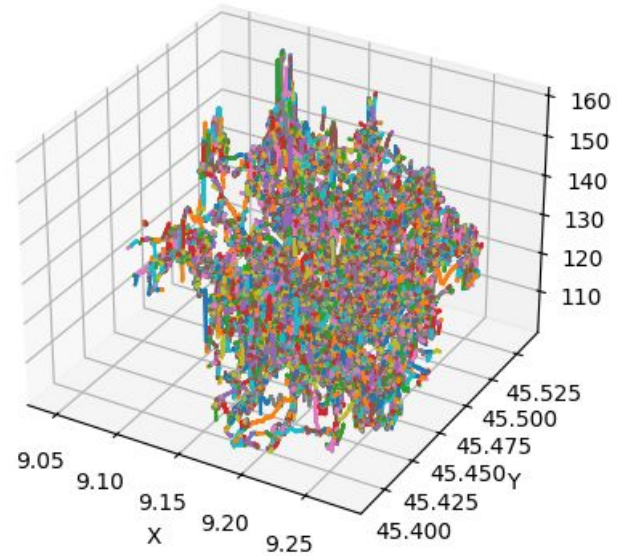


osmid	name	highway	oneway	reversed	length	lanes	geometry
31014839	Via Primo Bandini	residential	FALSE	TRUE	4081		LINestring (10.3432476 44.7934719, 10.3431984 44.7934832)
47361101	Via Zarotto	tertiary	FALSE	FALSE	6835	2	LINestring (10.3432476 44.7934719, 10.343221 44.7934134)
722189115	Via Antonio Zarotto	tertiary	FALSE	FALSE	8354	2	LINestring (10.3432476 44.7934719, 10.3432801 44.7935434)
47361101	Via Zarotto	tertiary	FALSE	TRUE	6835	2	LINestring (10.343221 44.7934134, 10.3432476 44.7934719)
47361104	Via Luigi Boccherini	residential	FALSE	FALSE	4357		LINestring (10.343221 44.7934134, 10.3432742 44.7934029)
47361104	Via Luigi Boccherini	residential	FALSE	TRUE	767		LINestring (10.3432835 44.7934009, 10.3432742 44.7934029)
722189115	Via Antonio Zarotto	tertiary	FALSE	TRUE	8354	2	LINestring (10.3432801 44.7935434, 10.3432476 44.7934719)
31014839	Via Primo Bandini	residential	FALSE	FALSE	4029		LINestring (10.3431498 44.7934943, 10.3431984 44.7934832)
31014839	Via Primo Bandini	residential	FALSE	FALSE	4081		LINestring (10.3431984 44.7934832, 10.3432476 44.7934719)
31014839	Via Primo Bandini	residential	FALSE	TRUE	4029		LINestring (10.3431984 44.7934832, 10.3431498 44.7934943)
47361104	Via Luigi Boccherini	residential	FALSE	FALSE	767		LINestring (10.3432742 44.7934029, 10.3432835 44.7934009)
47361104	Via Luigi Boccherini	residential	FALSE	TRUE	4357		LINestring (10.3432742 44.7934029, 10.343221 44.7934134)

Milan Road Data



Road polygons



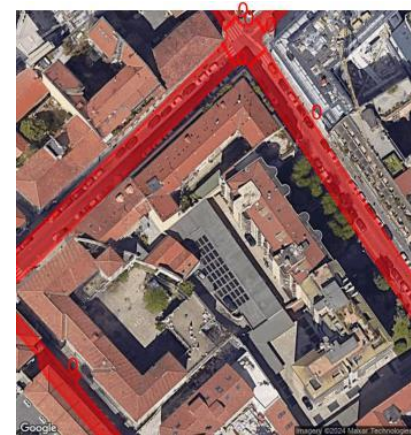
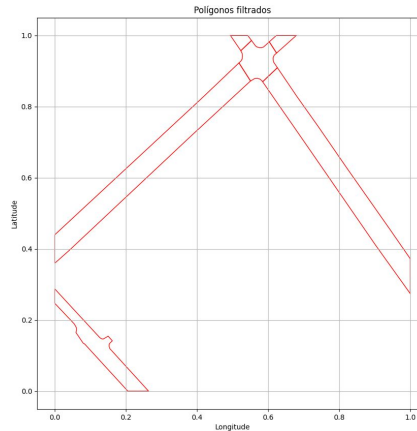
Satellite-Acquired Data

- 18,850 images from Milan and surrounding regions;
- 9,639 images containing points associated with road polygons.



Overlay Process

- Select the image;
- Generate the image's geographic coordinates;
- Search for and crop polygons within the image's geographic area;
- Save the filtered polygons in a format compatible with segmentation models.



Overlay Results





Polygon Overlay Classification

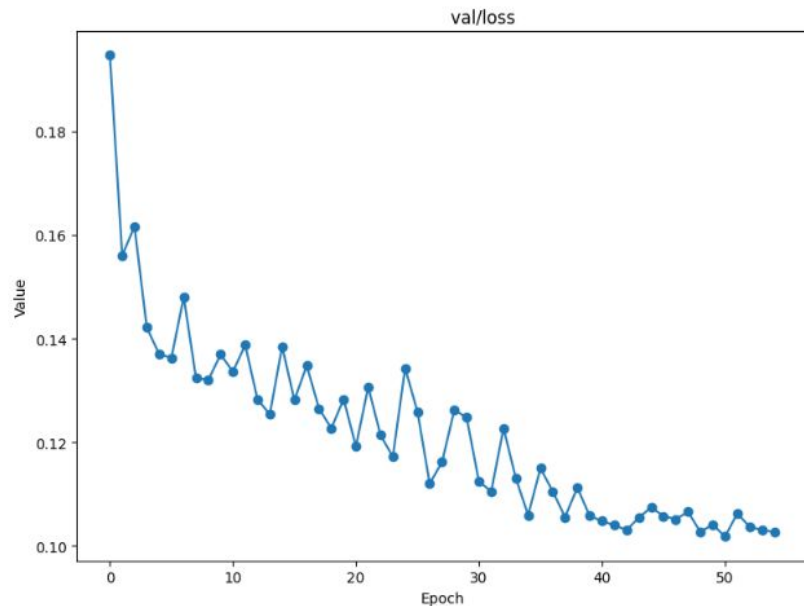
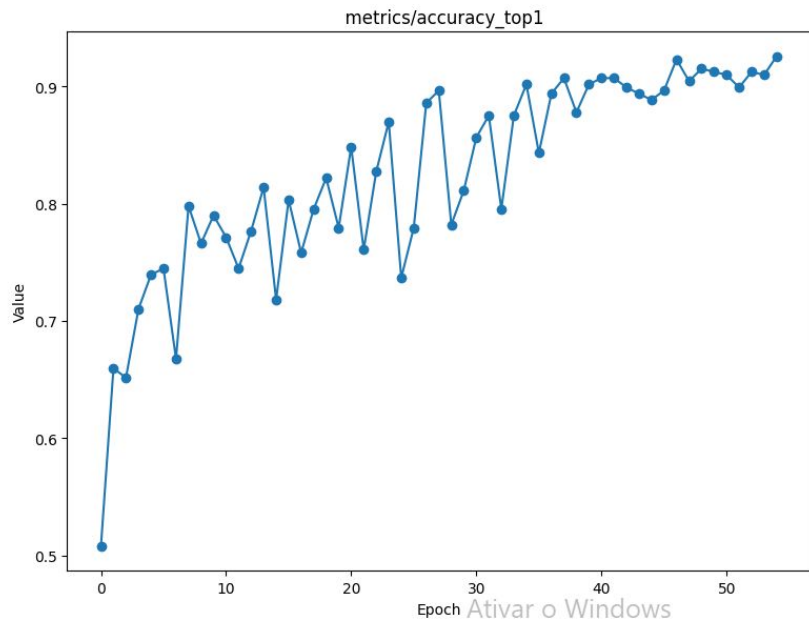
Classification Dataset



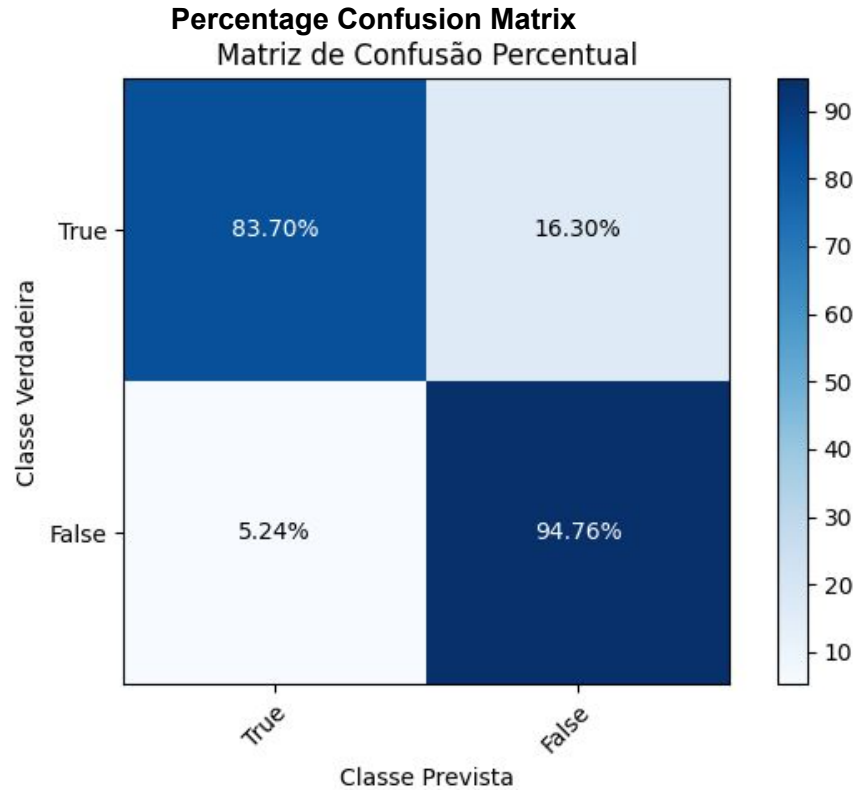
- A tool was developed to facilitate the manual analysis of images with polygon overlays;
- 2,500 images were selected to maintain a balanced dataset between good and poor overlay classifications.



Resultado Classificador



Classifier Results

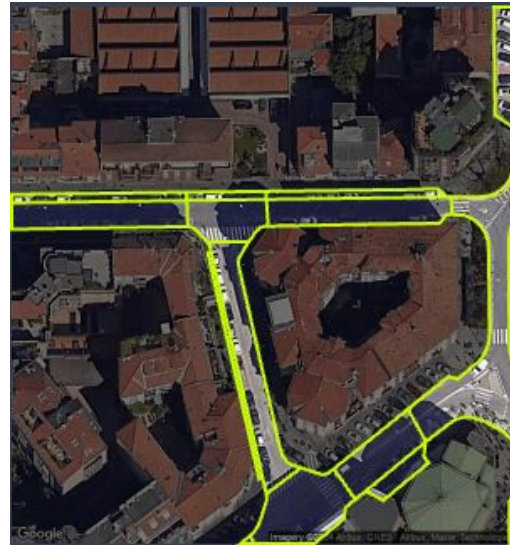




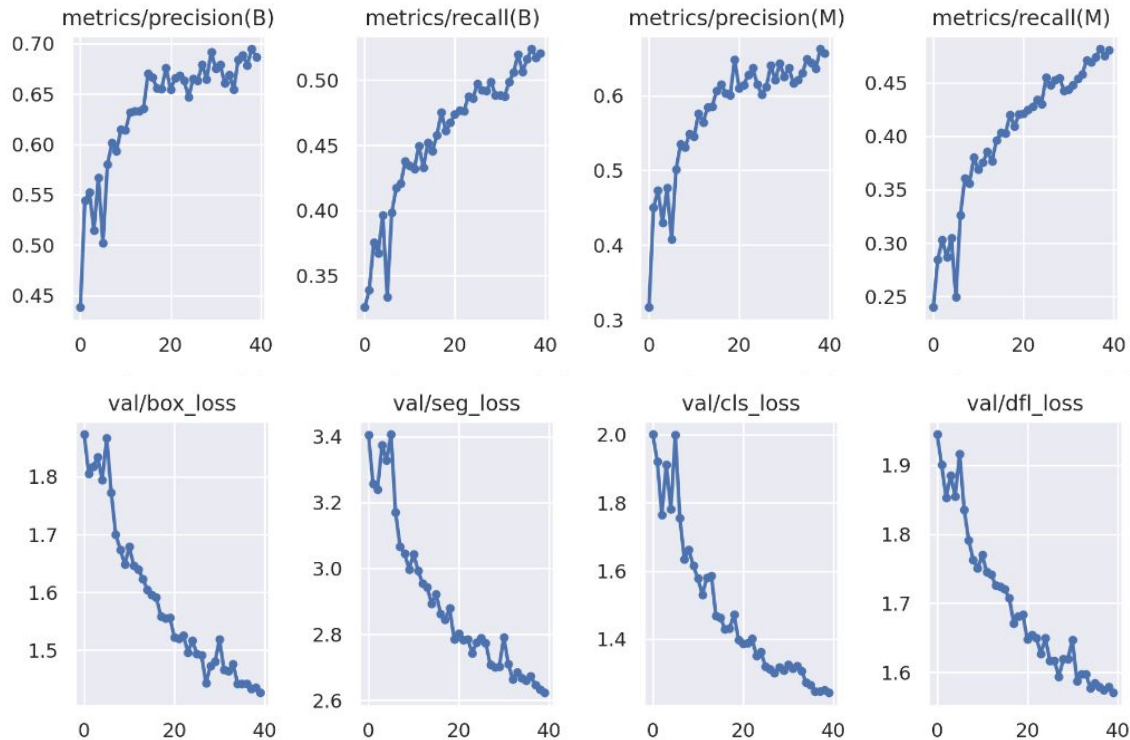
Segmentation

Dataset

- Built by combining manually validated data with the classifications generated by the previous model;
- 3,731 images;
- 38,113 polygons.



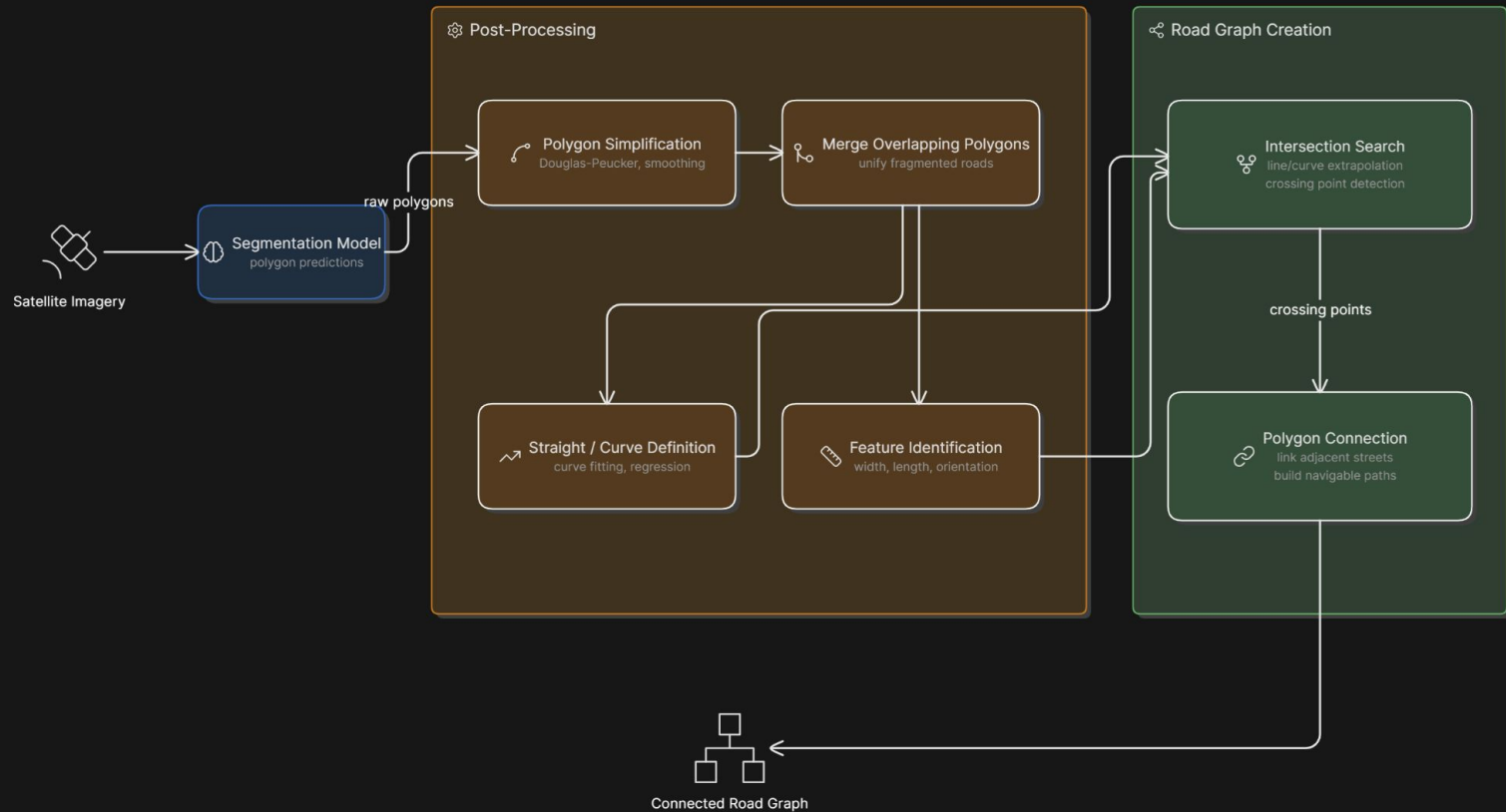
Resultados do Treinamento para segmentação



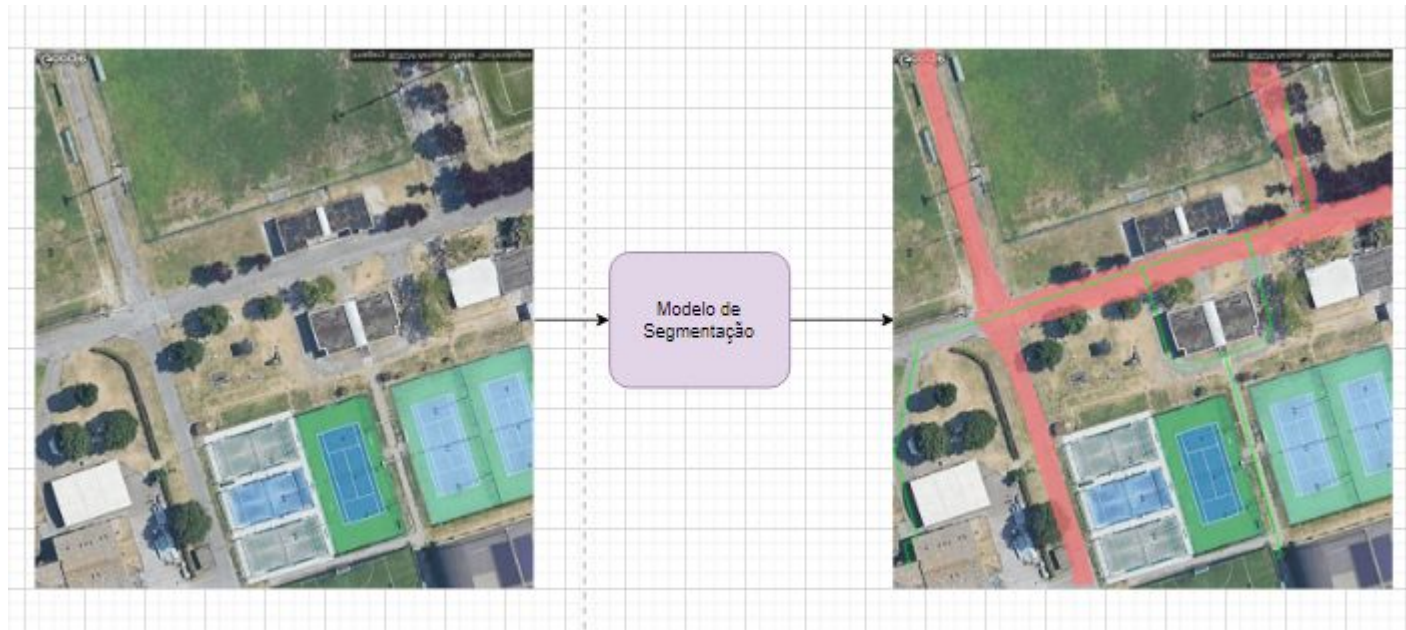


Tool

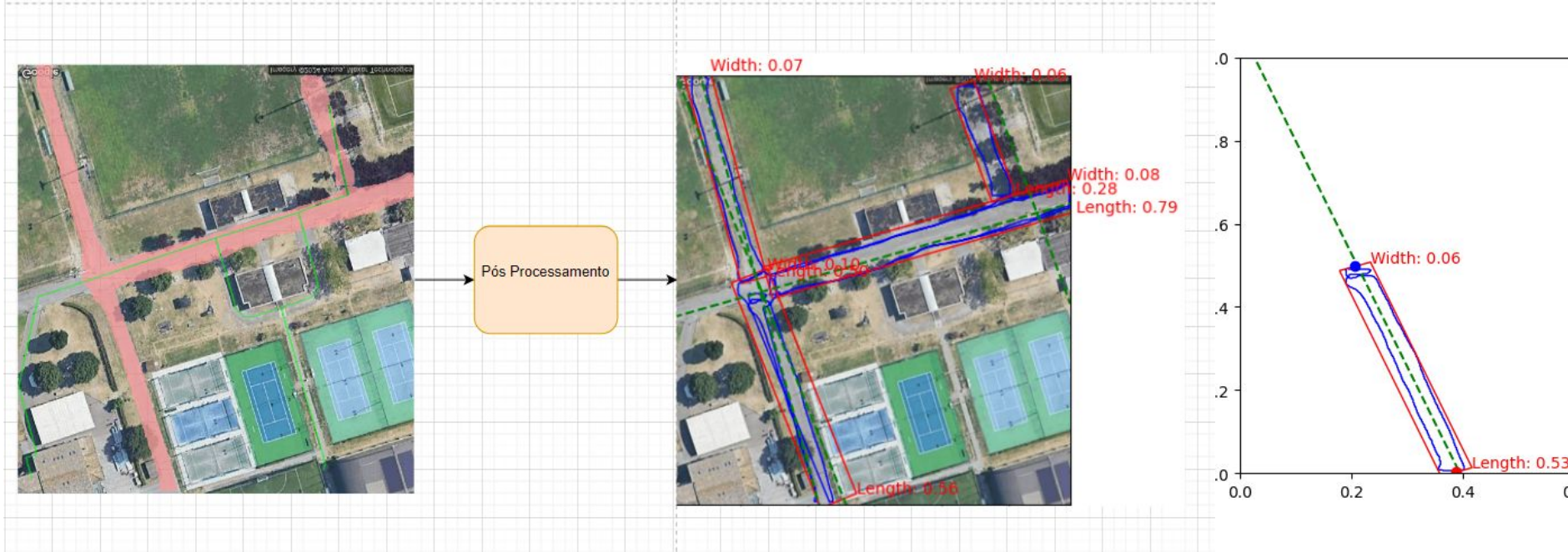
Road Segmentation Post-Processing Pipeline



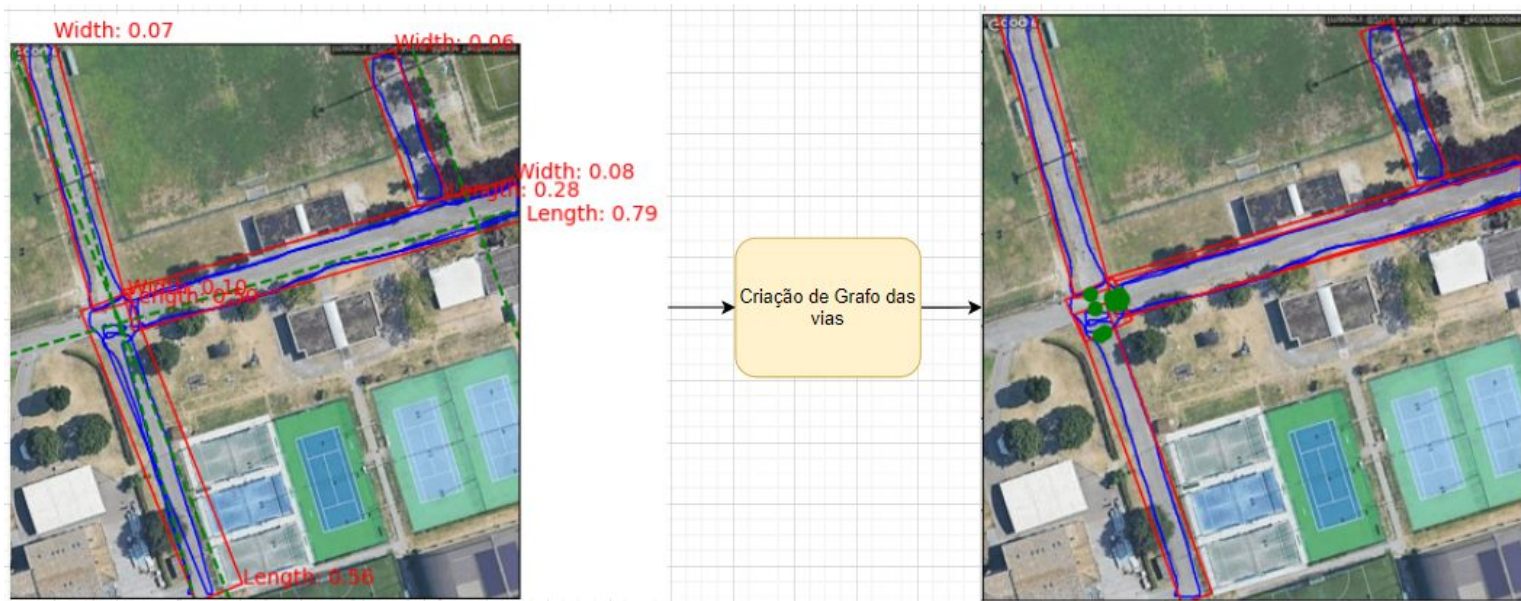
Segmentation model



Post-Processing



- Street-to-street connection and intersection connection;
- Every street has 2 connection nodes between streets and multiple intersection nodes.

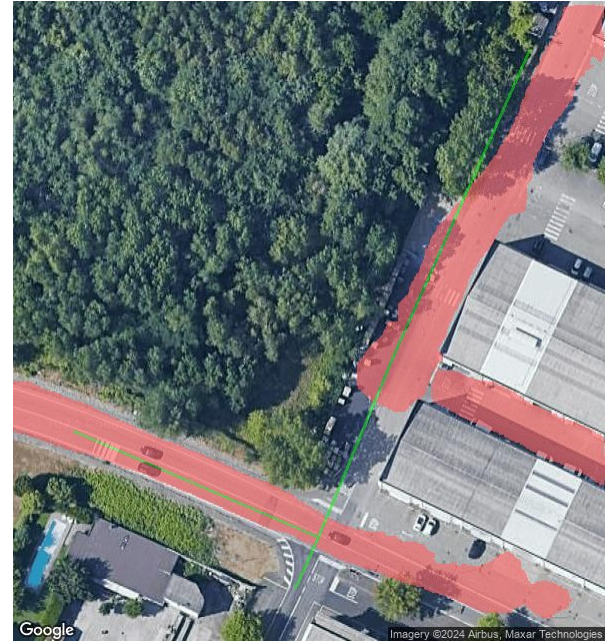
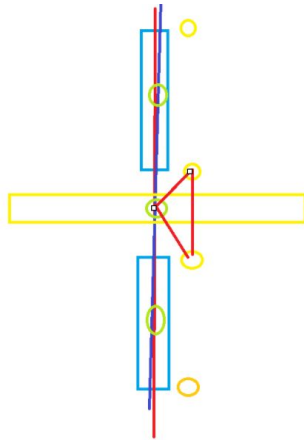




Challenges to Be Solved

Gaps Between Polygons

- Verify the similarity between the image area containing the polygons that should be connected and the gap between them.



Polygon Shapes

- Polynomial regression to identify the curve and perform extrapolation in order to detect nearby polygons.

