# Detecting Road Intersections from Satellite Images using Convolutional Neural Networks


**Fatma El-taher, Luis Miralles, Jane Courtney, Susan Mckeever**

**School of Computer Science, Technological University Dublin, Ireland**


SAC2023


## Abstract

The location of intersections is an important consideration for vulnerable road users such as People with Blindness or Visually Impairment (PBVI) or children. Route planning applications, however, do not give information about the location of intersections as this information is not available at scale. In this paper, we propose a deep learning framework to automatically detect the location of intersections from satellite images using convolutional neural networks. For this purpose, we labelled 7,342 Google maps images from Washington, DC, USA to create a dataset. This dataset covers a region of 58.98 km$^{2}$ and has 7,548 intersections. We then applied a recent object detection model (EfficientDet) to detect the location of intersections. Experiments based on the road network in Washington, DC, show that the accuracy of our model is within 5 meters for 88.6\% of the predicted intersections. Most of our predicted centres of the intersections (approx 80\%) are within 2 metres of the ground truth centre. Using hybrid images, we obtained an average recall and an average precision of 76.5\% and 82.8\% respectively, computed for values of Intersection Over Union (IOU) from 0.5 to 0.95, step 0.05. We have published an automation script to enable the reproduction of our dataset for other researchers.


## Background

Predefining intersection features plays a crucial role in different domains such as route planning and road network extractions. Intersection identification is an important task to extract road networks from satellite photos, especially in complicated regions. Predefining intersection features increases the flexibility of route planning as it allows users to add or eliminate crossroads in accordance with their preferences.

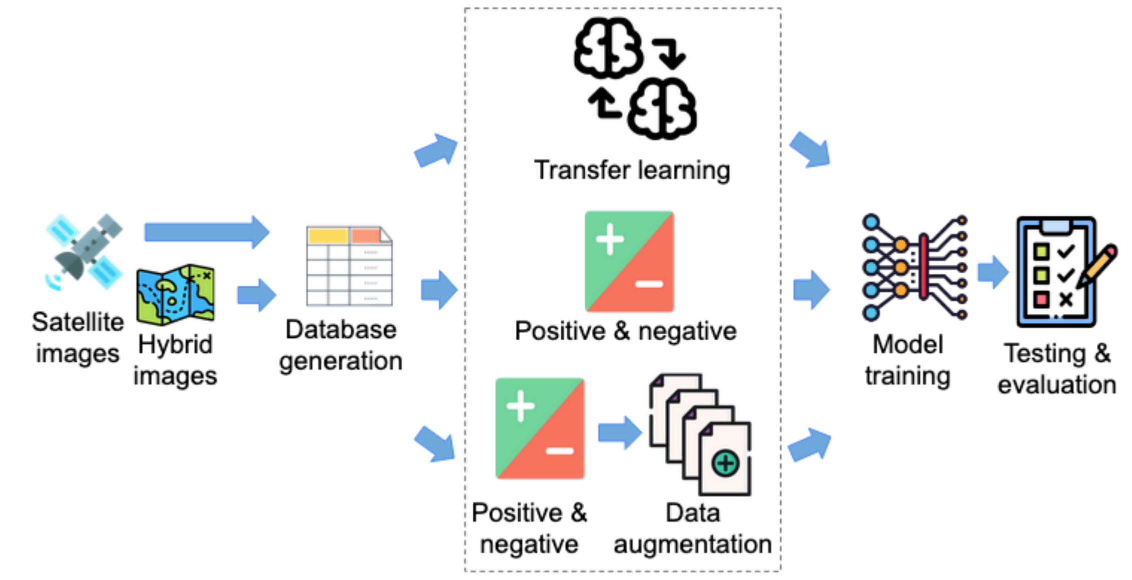
The majority of PBVI outdoor navigation systems use publicly available pathfinding services that generally do not provide information about the location of intersections. This insight is crucial for determining the safest route and avoiding crossings for PBVI. Three key benefits make satellite images a viable source of information for detecting intersections. First, they capture various types of intersection structures. Second, they offer extensive coverage (such as Google Maps photos), enabling broadly-applicable detection models. Thirdly, they can be utilized to extract map data during offline procedures. Most previous works have a limited size of remote-sensing datasets. Moreover, there are no labeled satellite-images datasets indicating the location of intersections.

To our knowledge, there are no other works in the literature estimating the location of intersections in navigation maps. This paper aims to address this gap by presenting a deep learning framework to automatically detect the location of intersections from satellite images. We first extended our prior dataset to incorporate the location of intersections and we added a bounding box around the intersections. Then, we trained an object detection model based on convolutional neural networks (CNN). We used the resultant output bounding box from the model to calculate the location of the center of each intersection (longitude and latitude). This is a required pre-step to enabling large-scale annotation of maps with an extra layer of information about intersections.

## Methodology

**Dataset:** To curate the dataset, we first downloaded two types of images (satellite and hybrid) from Google Maps, at zoom level nº19 which clearly provides intersection details. The satellite images are unprocessed pictures where roads and intersections may be blocked by other objects such as trees and buildings. Hybrid images are satellite images that Google has labeled to highlight roads and display street names.

To annotate images automatically, we used two available datasets from the Washington, DC, government website. The first one is the intersection points dataset which has the intersections' locations. The second one is the roadway intersection approach dataset, which contains road segments that lead to intersections. Each road segment is defined by several points using latitude and longitude. One of these points is the centre of an intersection.




**Pipeline of the steps involved in the proposed methodology for predicting the intersection location.**

**Approach:** We consider road intersections to be distinct identifiable objects, with consistent features (centre point, three or more branches) within an image. Thus, we applied an object detection approach. In the first step, we trained the state-of-the-art object detection model (EfficientDet) to detect the presence and location of intersections. To achieve the best results, we considered different decisions during the training phase (see the figure of methodology). Firstly, we compared two methods of training the intersection detection model; from scratch and by transfer learning methods. Secondly, we examined the effect of using images with no intersection (negative samples) in training the object detection model. Thirdly, we investigated the effect of using augmentation to extend the dataset.

## Results: Detecting intersection location

For the performance evaluation of object detection models, Average precision (AP) and Intersection Over Union (IOU) are used.

**Experiment I: Transfer learning versus training from scratch**
To investigate its impact on performance, we trained the model from scratch and with a transfer learning strategy. Using a hybrid dataset, the model was trained from scratch and using transfer learning, and it achieved AP@[50:5:95] of 67.93\% and 68.87\% respectively. Using a satellite dataset, the model was trained from scratch and using transfer learning, and it achieved AP@[50:5:95] of 64.90\% and 65.90\% respectively.

**Experiment II: Training using positive and negative images**
In "Positive object detector" the model is trained exclusively with positive cases and tested over both positive and negative examples. The "Classifier+positive object detector" approach consists of a classifier model as a pre-stage to classify images into intersection and no-intersection. The "Positive+Negative object detector" approach is based on training the object detection model using positive and negative images.

| Method | AP@[50:5:95] | AP@[50] | AP@[75] | AR@[50:5:95] |
|---|---|---|---|---|
| Pos obj detector | 67.93% | 84.00% | 75.90% | 82.13% |
| Classf+pos obj detr | 74.47% | 92.83% | 83.10% | 80.57% |
| Pos+Neg obj detr | 74.30% | 92.43% | 83.07% | 81.33% |

**Experiment III: The impact of using augmentation techniques**
The impact of using augmentation methods during training is examined in this paper. Various augmentation methods such as brightness and contrast, flipping, and Gaussian blur were tested with the EfficientDet model on both satellite and hybrid datasets.

| | Satellite images | | | |
|---|---|---|---|---|
| Augmentation | AP@[50:5:95] | AP@[50] | AP@[75] | AR@[50:5:95] |
| No | 65.43% | 87.63% | 73.30% | 74.50% |
| Yes | 68.80% | 89.53% | 76.07% | 76.97% |
| | Hybrid images | | | |
| Augmentation | AP@[50:5:95] | AP@[50] | AP@[75] | AR@[50:5:95] |
| No | 73.27% | 92.27 | 82.10% | 80.47% |
| Yes | 76.50% | 92.77 | 84.40% | 82.83% |

**Augmentation effect on object detection model's performance.**


**Examples of detected intersections within 2 meters.**

## Conclusion and future direction

Route planning applications do not provide users with information about intersection locations when embarking on a journey. Satellite images are a potentially global source for enabling the extraction of this information. We present a deep learning framework to detect the location of intersections. EfficientDet can detect the location of the intersection from satellite images with AP@[50:5:95] equals 68.80% and AR@[50:5:95] equals 76.97%. The performance is raised at hybrid images to equal 76.50\% AP@[50:5:95] and equal 82.83\% AR@[50:5:95].

Our model has been tested with images from the suburbs, it would be interesting in the future to test it in more dense areas of big cities such as New York or London. And also test it over areas covered with trees or with narrow paths. Additionally, one of the limitations of our methodology is that it finds it difficult to detect multiple intersections in one single image. To overcome this problem we would need to train the model with more images of this kind. For future work, we plan to extend our dataset to collect more images of other geo-areas with different road network styles and terrain.


## Acknowledgments

Insert This publication has emanated from research supported in part by a Grant from Science Foundation Ireland under Grant number 18/CRT/6222.