

SCIENCE

Estimating Vehicular Traffic Intensity With Deep Learning and Semantic Segmentation

Student researcher: Logan Bradley-Trietsch, Junior

Semantic segmentation, a computer vision task that involves assigning a label to each pixel in an image, has many applications to topics such as medical imaging or autonomous driving. This project used the deep learning architecture Deeplabv3+ to analyze traffic video using semantic segmentation in order to quantify and make predictions about the intensity of vehicular traffic in traffic video.

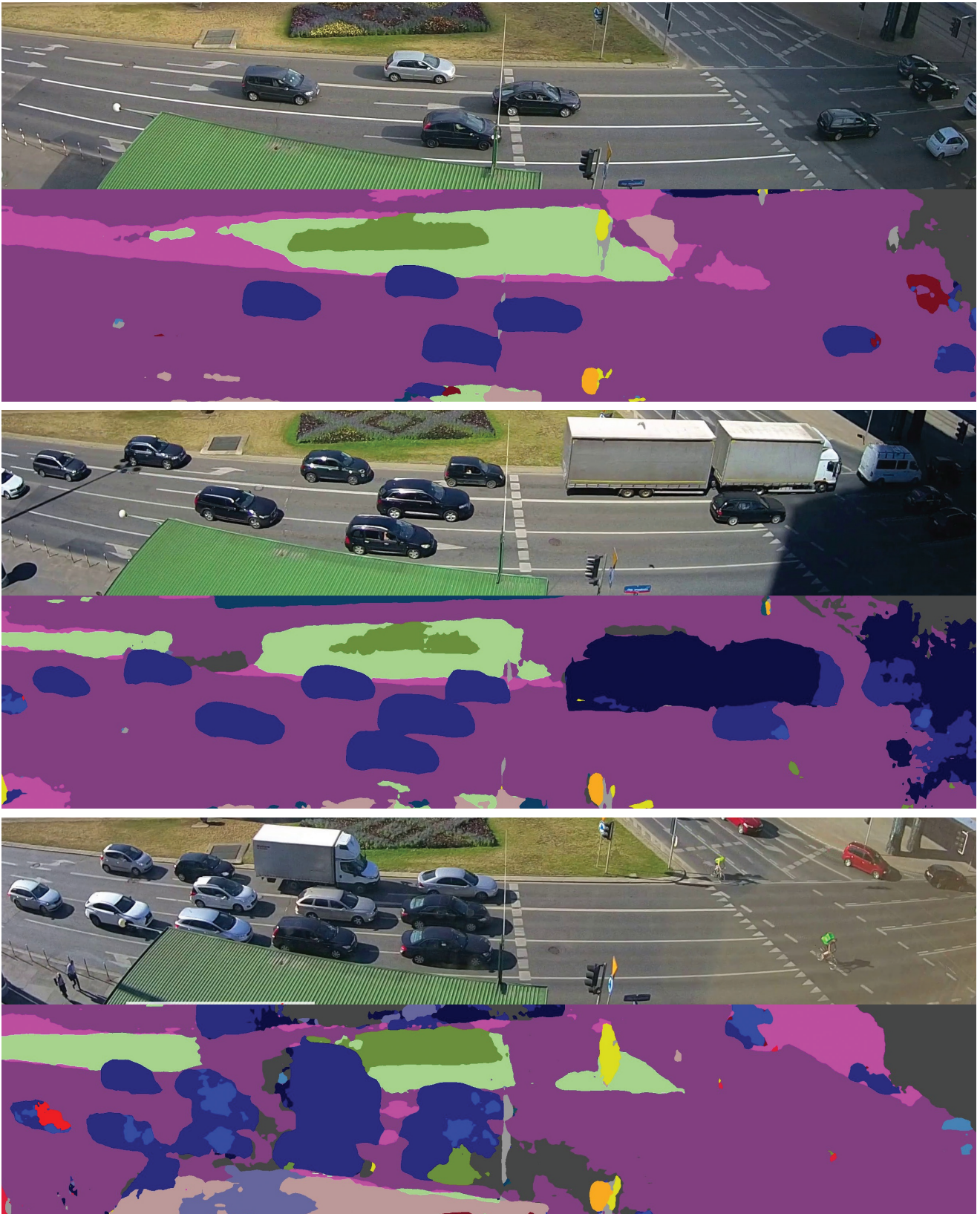
The model for Deeplabv3+ was trained on the Cityscapes and ImageNet datasets. Since there is a paucity of high-quality training data, finding real-world data with class distributions that closely mimic the training data was challenging. The real-world data that was used for this project was from an intersection in Warsaw. A metric called the “vehicle to total pixel ratio” was created in order to quantify traffic intensity, and it was defined as the total number of pixels that correspond to a vehicle class divided by the total number of pixels in an area of interest (e.g., an entire image or an individual lane), and is calculated for each frame in the video. The data pipeline involved downloading terabytes of traffic video, extracting and preprocessing individual frames from the video, and performing analysis using Deeplabv3+ to yield both visualized results (as seen in the figure) and the vehicle to total pixel ratio for each frame. An exponential smoothing model was fit for the vehicle to total pixel ratio versus time, which was selected based on the root-mean-square

error of holdout data. However, there was limited success in the model’s predictive ability due to shadows of vehicles being labeled as false positives, a level of class imbalance compared to the training data, and building a model on limited data since analyzing hundreds of thousands of images took significant computation time.

This unique approach to traffic estimation may have advantages over traditional methods of analyzing traffic video (e.g., object detection) in its ability to more accurately account for size differentials for vehicles within the same class. As the number of high-quality training datasets increases and as deep learning models become increasingly powerful and efficient, one would expect results to become more accurate and thus more conducive to quantifying and making predictions about traffic intensity. This project and its methods are one step toward the “smart city”, that is, a city that has a traffic system based on artificial intelligence that can collect and analyze immense amounts of traffic data in order to efficiently route traffic in real-time with the end goal of mitigating gridlock traffic and improving driver and pedestrian safety.

This research is supported by NSF Grant DMS-1246818 (Statistics Living-Learning Community).

Research advisor Xiao Wang writes: “Logan was able to perform the whole data analysis pipeline independently, which includes data preprocessing, analysis, visualization, and validation. This analysis of using deep learning to model traffic videos has a wide range of applications in traffic control and is a big step toward a smart city.”



The top images show (from left to right) representative frames from the morning, afternoon, and evening, respectively. Specifically, the images have the same vehicle to total pixel ratio as the average vehicle to total pixel ratio of their time of day. The bottom images show visualized results after analysis.