# CLASSIFICATION OF ROAD USERS TRACKED WITH LIDAR AT INTERSECTIONS AND EVALUATION OF THE TRACKING AND CLASSIFICATION RESULTS

Cristhian Lizarazo
Purdue University, West Lafayette, Indiana, USA, clizaraz@purdue.edu

**Abstract**

Data collection is a necessary component for characterizing the existing transportation infrastructure. Manual data collection methods have proven to be inefficient and limited in terms of the data required for comprehensive traffic and safety studies. Automatic methods are being introduced to characterize the transportation system more accurately and are providing more information to better understand the dynamics between road users. Video data collection is an inexpensive and widely used automated method, but the accuracy of video-based algorithms is known to be affected by obstacles and shadows and the third dimension is lost with video camera data collection. The impressive progress in sensing technologies has encouraged development of new methods for measuring the movements of road users. Professor Andrew Tarko at Purdue University proposed the application of a LiDAR-based algorithm for tracking vehicles at intersections from a roadside location. LiDAR provides a three-dimensional characterization of the sensed environment for better detection and tracking results. The feasibility of this system was analyzed in this thesis using a novel evaluation methodology to determine the accuracy of the algorithm when tracking vehicles at intersections. According to the implemented method, the LiDAR-based system provides successful detection and tracking of vehicles, and its accuracy is comparable to the results provided by frame-by-frame extraction of trajectory data using video images by human observers. The second component of this thesis focused on proposing a classification methodology to discriminate between vehicles, pedestrians, and two-wheelers. Four different methodologies: the K-Nearest Neighbor (KNN) classification algorithm, a multinomial logit discrete choice model, classification trees, and C.5.0 boosting, were evaluated to identify the best method for implementation in the existing tracking algorithm. The assumptions on which the techniques are built were discussed from the classification performance point of view to better understand the sources of the differences. The KNN algorithm, which is capable of creating adaptive decision boundaries based on the characteristics of similar observations, provided better performance when evaluating new locations. The multinomial logit model did not allow the inclusion of collinear variables into the model. Overfitting of the training data was indicated in the classification tree and boosting methodologies and produced lower performance when the models were applied to the test data. Despite ANOVA analysis not supporting superior performance by a competitor, the objective of classifying movements at intersections under diverse conditions was achieved with the KNN algorithm and was chosen as the method to implement with the existing algorithm.

**Keywords:** wide-area detection; tracking evaluation; classification model; KNN Classification; multinomial logit model; classification tree; C5.0 Boosting.