EXTRACTING URBAN STREET FEATURES USING STREET LEVEL LiDAR DATA FOR CONNECTED VEHICLE APPLICATIONS

Rakesh Nune, Corresponding Author
District of Columbia Department of Transportation
55 M Street, SE, Washington DC 20003
Tel: 571-265-4047; Email: Rakesh.Nune@dc.gov

Kaiqun Fu
Department of Computer Science
Virginia Polytechnic Institute and State University
7054 Haycock Rd, Falls Church, VA 22043
Tel: 703-303-8419; Email: fukaiqun@vt.edu

Jason X. Tao
District of Columbia Department of Transportation
55 M Street, SE, Washington DC 20003
Tel: 202-671-1489; Email: jason.tao@dc.gov

Word count: 3061 word text + 14 tables/figures x 250 words (each) = 6561

ABSTRACT

The connected vehicles and autonomous vehicles have grasped a lot of public attention recently. One of the critical functions for connection vehicle applications is to transmit the roadway MAP data from the Dedicated Short Range Communications (DSRC) roadside equipment to the on-board unit within vehicles. Collection of the MAP data with traditional methodology is very tedious and costly. In this research, we developed an algorithm employing deep learning techniques to extract the lane markings and the stop bar locations from the existing panoramic LiDAR data and images. Previous research work in this direction includes model development based on Bayesian inference road lane detection. As the present work is one of the very first to use street level LiDAR data, a new algorithm is developed based on color clustering and deep learning techniques to identify, classify and geolocate pavement markings. The deep neural network model achieved an accuracy rate of 71.2% in identifying the pavement markings at intersections. The detected pavement markings can be geolocated to generate GIS files which can used to generate MAP messages. The research work has shown that the proposed approach is promising for the lane detection and road sign recognition thus mapping the city precisely for advancement of connected vehicles.

Keywords: MAP message, DSRC, deep learning, LiDAR, neural network

INTRODUCTION

Connected vehicles have the potential to greatly improve the safety and efficiency of the surface transportation systems. The connected vehicle technology, which includes Vehicle to Infrastructure (V2I) and vice versa communication, provides the functionalities to facilitate autonomous vehicles and driverless operations. In the connected vehicles world, the vehicle On-Board Units (OBU) exchange real-time information with the infrastructure through the DSRC radio for navigation, safety and other operations. In order to get prepared for the coming age of connected vehicle driving, many cities are working towards establishing required infrastructure for enabling smooth transition for connected vehicles (*1*).

The content of communication between the vehicles and the infrastructure includes the Signal Phasing and Timing (SPAT), Traveler Information Message (TIM), and the MAP message. The MAP message is a critical component that gives the user information about the geometry of the intersection (*2*). A typical MAP message contains the data of lane marking, centerline and stop bar locations. Traditionally, the MAP messages are generated manually by using the GIS systems. The methodology requires employing field visits and extensive verifications which is a slow and costly process (*3*).

Various early works based on computer vision techniques were proposed for exploring solutions to research challenges of such kinds. For instance, kernel based classification algorithms such as Support Vector Machines (SVM) was raised for road sign recognition (*4*); other works proposed models based on Bayesian inference were addressed for road lane detection (*5*).

The present research work, a methodology using artificial intelligence algorithms like deep learning is proposed to automatically extract pavement marking locations to generate MAP messages for urban areas.

Street Level LiDAR Data

Light Detection and Ranging (LiDAR) can be categorized into two classes: the airborne LiDAR (ALS) and the terrestrial LiDAR (TLS). The dominant type of LiDAR technology utilized in the studies of urban planning and geomorphology is the topographic LiDAR. The infrared laser light of this type is emitted from a fixed-wing aircraft. The terrestrial LiDAR techniques, on the other hand, usually mounted on moving vehicles, provide much more accurate point collection procedures and can be used to manage facilities, conduct highway and rail surveys.

Previous studies (*6; 7*) proposed methods for tropical forest carbon mapping using the airborne LiDAR technologies. Sun et al. (*8*) proposed an automated method to create 3D watertight building models from airborne LiDAR point clouds. Zhang et al. (*9*) proposed an optimized segments and connected components classification process, based on the airborne LiDAR data. Although these airborne LiDAR techniques cover a wide ranges of certain regions, they fail to collect dense point clouds with high accuracies as constrained by various limitations (*8; 9*). On the other hand, a few previous works have addressed the feasibility of the terrestrial

LiDAR in research purposes. The terrestrial LiDAR techniques can facilitate rapid LiDAR calibration with minimal field data are developed. Recent years have seen the emergence of the researches that focus on the terrestrial LiDAR data (*10*). Research areas such as autonomous driving, especially in the direction of identification objects obscuring lane marks and road boundaries, exploited the feasibility of the terrestrial LiDAR data (*11*). The terrestrial LiDAR data has been used in resolving the problem of ground roughness estimation for road segmentations (*12*). Other works have explored the applicability of the terrestrial LiDAR data in curb and berm detection (*13; 14*).

In this research, autoencoders, a deep learning algorithm is used to train and recognize the pavement markings and lane information. Autoencoder is an unsupervised machine learning algorithm that is capable of extracting hidden features of given high dimension data (*15*). It is constructed under the scheme of the feedforward non-recurrent neural networks. It consists of an input layer, multiple hidden layers and one output layer. The difference between an autoencoder and a conventional neural network is that size of the output layer in an autoencoder is the same as the input layer; on the other hand, the output layer of a conventional neural network is usually smaller than the input layer. Encoders belong to unsupervised machine learning family unlike conventional neural networks. The two major components: the encoder and the decoder:

$$\begin{gathered} \boldsymbol{\phi}: \boldsymbol{\mathcal{X}} \rightarrow \boldsymbol{\mathcal{F}} \\ \boldsymbol{\psi}: \boldsymbol{\mathcal{F}} \rightarrow \boldsymbol{\mathcal{X}} \\ \underset{\boldsymbol{\phi},\boldsymbol{\psi}}{\mathbf{argmin}} \|\boldsymbol{X} - (\boldsymbol{\phi} \cdot \boldsymbol{\psi})\boldsymbol{X}\|^{\mathbf{2}} \end{gathered} \tag{1}$$

where $\boldsymbol{\phi}$ denotes the encoder, $\boldsymbol{\psi}$ denotes the decoder, and $\boldsymbol{\mathcal{X}}$, $\boldsymbol{\mathcal{F}}$ denote the input space of the data and the hidden features space respectively. Equation 1 is the objective function for minimizing the loss function. Street level Light Detection and Ranging (LiDAR) data is procured for this work which provides rich data attributes compared to 2D street level imagery.

The rest of the paper is organized as follows: the developed algorithm for extracting the street level data from LiDAR are summarized followed by deep learning technique used to identify the right pavement marking. The results are discussed in the conclusion section and authors discussed future work to expand the algorithm to cover other informational signs.

FEATURE DETECTION AND CLASSIFICATION OF STREET LEVEL DATA

The street level LiDAR data is collected by Washington DC department of transportation collected from a third party provider called Cyclomedia Inc[1]. The LiDAR data is overlapped with street view imagery providing a unique rich data reflecting real world attributes i.e. (X, Y, Z, R, G, B). Figure 1 shows a typical street level LiDAR image. The algorithm developed in the project can be divided in to three steps.

[1] Cyclomedia: *http://www.cyclomedia.com/us/*

Figure 1. LiDAR Data at 4th St NW and Decatur St NW

Step 1: Preprocessing of LiDAR cloud and Images

Elevations vary significantly along roadway corridors to accommodate various requirements like drainage, America Disability Act safety requirements. Figure 2 shows heat map of elevation on a typical corridor in Washington DC ranging from lowest elevation shown in blue increasing to red. This presents a challenge in extracting ground LiDAR data that is needed for analysis. To address this issue, corridors are cut regular intervals depending on gradient to generate ground level data

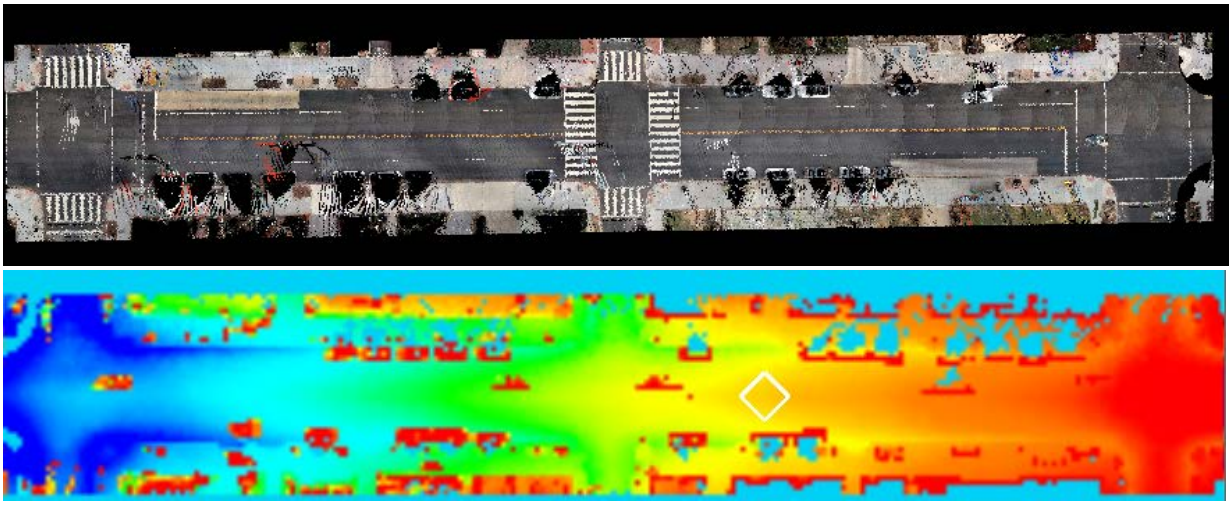

Figure 2: Typical Elevation of Street Profile

Depending on the angle of equipment, shadows and presence of other obstacles like vehicles, the quality of LiDAR cloud point data may vary at ground level which is the main interest of our research. Figure 3 shows LiDAR data at an intersection with spare points on one side.

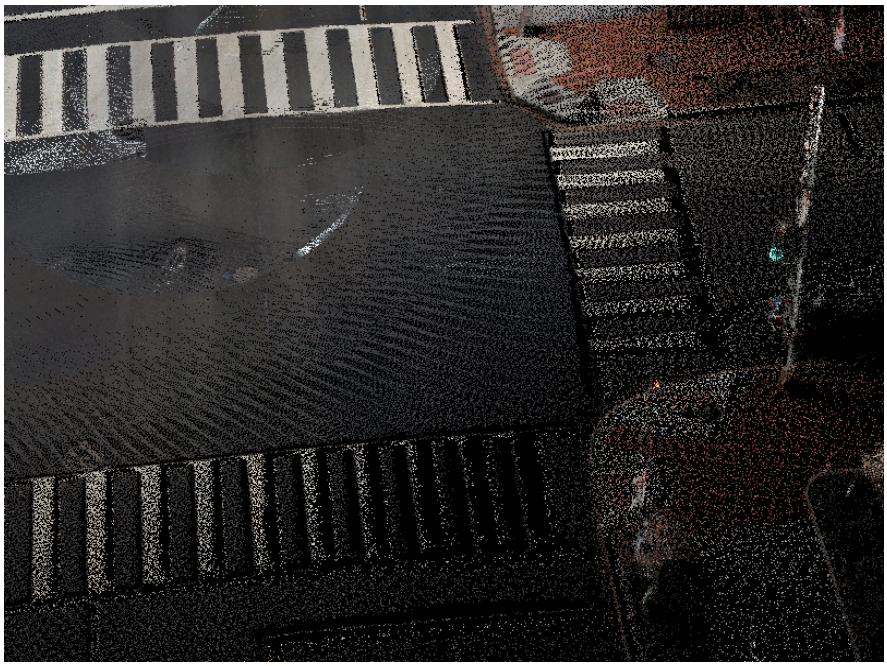

Figure 3. LiDAR picture generated at 16th and R St NW

Step 2: Clustering the data and image generation

In this step, data points are clustered using a distance based clustering method. It was expected that color aggregation would easily separate out pavement markings from background but in reality a huge spectrum of colors was found. Color spectrum occurs because of shadows, bright sunshine and deterioration of infrastructure named *Checker Shadow Illusion*[2]. In order to distinguish the pavement marks from the pavement background, a distance based two-clustering algorithm is used. Due to the extremely large size of the target LiDAR point clouds data set, it is not computationally efficient for the clustering algorithm to consider the entire data set as the input. To increase the efficiency, the Random Sample Consensus (RANSAC) paradigm is utilized with 10% sample rate of the original files. The Random Sample Consensus (RANSAC) algorithm was first proposed to serve as a solution towards the outlier detection problems. It estimates the parameters of a mathematical model from a set of observations which contains outliers in an iterative way. In the proposed coloring clustering task, the intuition of the RANSAC paradigm (*16; 17*) is referenced, namely, its capability of generalizing large noisy data sets with outliers.

K-mean clustering algorithm

The k-mean clustering algorithm is used on the cloud point with RGB. K-means clustering algorithm partitions the observations of the data points into k clusters, in which each point belongs to the cluster with the closest mean point. In the coloring based clustering scenarios, the data points (pixels) are transformed to the RGB color space, represented by vectors comprise of red, green, and blue channels of the given pixels. More specifically, given a set of

[2] Checker Shadow Illusion: *https://en.wikipedia.org/wiki/Checker_shadow_illusion*

LiDAR points (pixels)($\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \ldots, \boldsymbol{p}_N,$), where each observation of pixel is a 3-dimensional vector in RGB color space. The k-means clustering algorithm partitions the N observations into k sets: $\boldsymbol{S} = \{S_1, S_2, S_3, \ldots, S_k\}$ with regard to minimize the sum of squares within clusters:

$$\underset{\boldsymbol{S}}{\mathbf{argmin}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}} \sum_{\boldsymbol{X \in S_i}} \|\boldsymbol{X} - \boldsymbol{\mu_i}\|^{\boldsymbol{2}} \tag{2}$$

where μ_i is the mean of the points in S_i. In our coloring clustering scenario, under the assumption that the pixels of the pavement markings differ significantly from the other pixel points, the parameter k is set to be 2. Note that the target areas are limited to the surface of the streets in a small range.

The mean color calculated in RGB value is represented by a vector with three elements:

$$\boldsymbol{C_{mean}} = [\boldsymbol{R_{mean}}, \boldsymbol{G_{mean}}, \boldsymbol{B_{mean}}]^{\boldsymbol{T}} \tag{3}$$

the standard deviation is also represented by a vector of three elements:

$$\boldsymbol{C_{std}} = [\boldsymbol{R_{std}}, \boldsymbol{G_{std}}, \boldsymbol{B_{std}}]^{\boldsymbol{T}} \tag{4}$$

According to the calculated means and standard deviations, the filtering process filters the point cloud by applying the following threshold with a 3-std:

$$\boldsymbol{C_{range}}: [\boldsymbol{C_{mean}} - \boldsymbol{3C_{std}}, \boldsymbol{C_{mean}} + \boldsymbol{3C_{std}}] \tag{5}$$

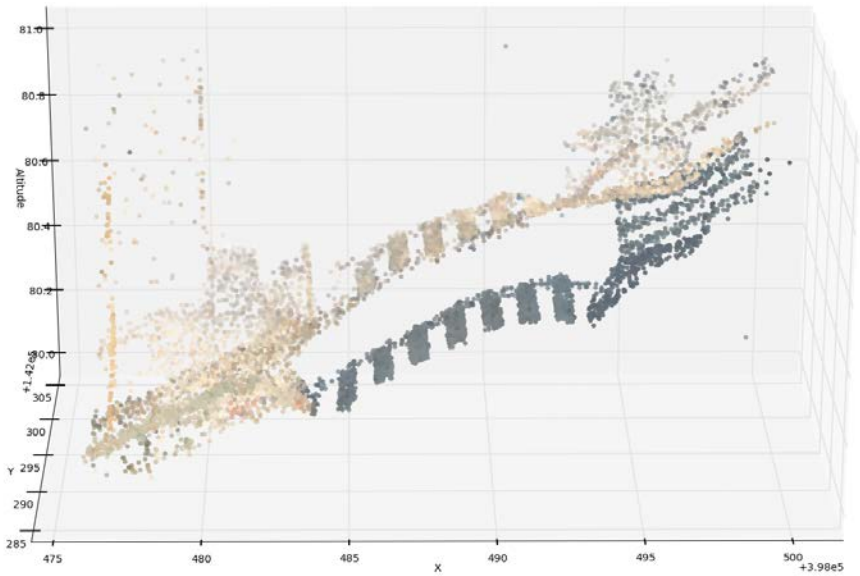

Figure 4. Clustering based on pixel colors

Upon clustering the points, DBSCAN (*18*), a density based spatial clustering algorithm, is performed on selected points to pinpoint the location and visualize the boundaries of the pavement marks.

Figure 5(a) shows clustering of LiDAR data based on coloring scheme. Figure 5(b) shows the clusters in two dimensions. The clusters then converted to images as shown in Figure 6. These images will be used as input for deep learning algorithms in step 3 to identify pavement marking from noise like curbs, cars and other infrastructure.

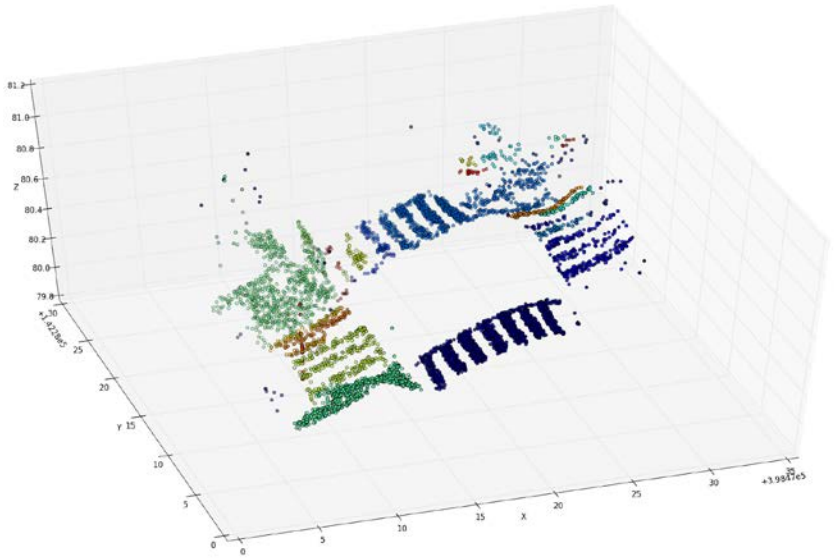

Figure 5(a). Spatial Clustering of LiDAR Data point data

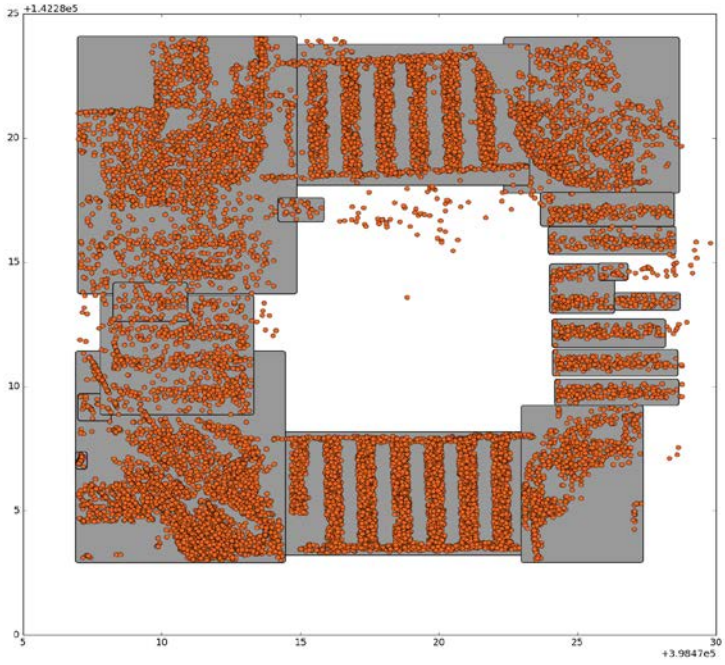

Figure 5(b). 2D clustering of data generated from 5(a)

Figure 6. Images generated from spatially clustered images

Step 3: Classification and Plotting of generated data using Deep Learning Techniques

The present work implemented autoencoder from deep learning techniques. Autoencoder is known for its promising performance in extracting and identifying features in high dimensional spatial data. The terrestrial LiDAR data is extremely dense and also has six dimensions which results in computationally intensive neural network model so for this work the dimension have been reduced to three i.e. just position data for training the model. The training data contains 128 road sections in Washington D. C., with an overall size of 18.1 GB. The 16th street NW corridor between T St NW and Church St NW is selected for testing the developed model. A total of 900 top view images of various intersections in Washington DC were used for training the autoencoder mode. Four intersections are selected for testing data generation consisting a total of 132 labelled data.

The algorithm is implemented on Ubuntu 14.04 system with a Python version 2.7, MATLAB 2015b[3]. The LibLAS[4] library is used for LiDAR data preprocessing and analysis. Python module PyProj[5] is utilized for cartographic transformations and geodetic computations. The autoencoder is constructed by using Neural Networks Toolbox provided by MATLAB 2015b. The layer of the autoencoder consists of 16384 feature nodes from the pavement markings proposal images. Two layers of encoder are implemented, the detailed configurations for each layers are listed below and shown in figure 7.

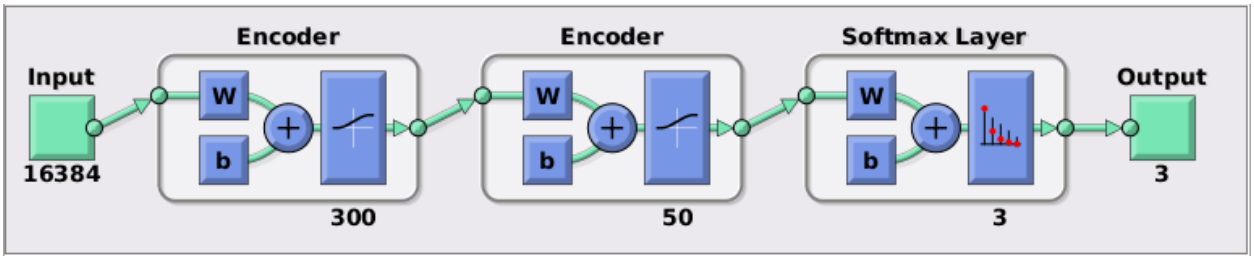

Figure 7: Neural Network Model Developed for Classification

[3] MATLAB: *http://www.mathworks.com/products/new_products/latest_features.html*
[4] LibLAS: *http://www.liblas.org/*
[5] PyProj: *https://jswhit.github.io/pyproj/*

RESULT DISCUSSION

The above autoencoder model has achieved an accuracy of 71.2% in classifying the data to right category which is plotted on map and shown in Figure 8(a). The green bounding boxes represent the pavement markings predicted by algorithm and Red boxes show noise recognized by the model between 16 St NW between N St and T St NW. Figure 8(b)-8(d) shows individual intersections.

Figure 8(a): Detected pavement marking on 16th St

Figure 8(b): Detected pavement marking on 16th St and Church St NW

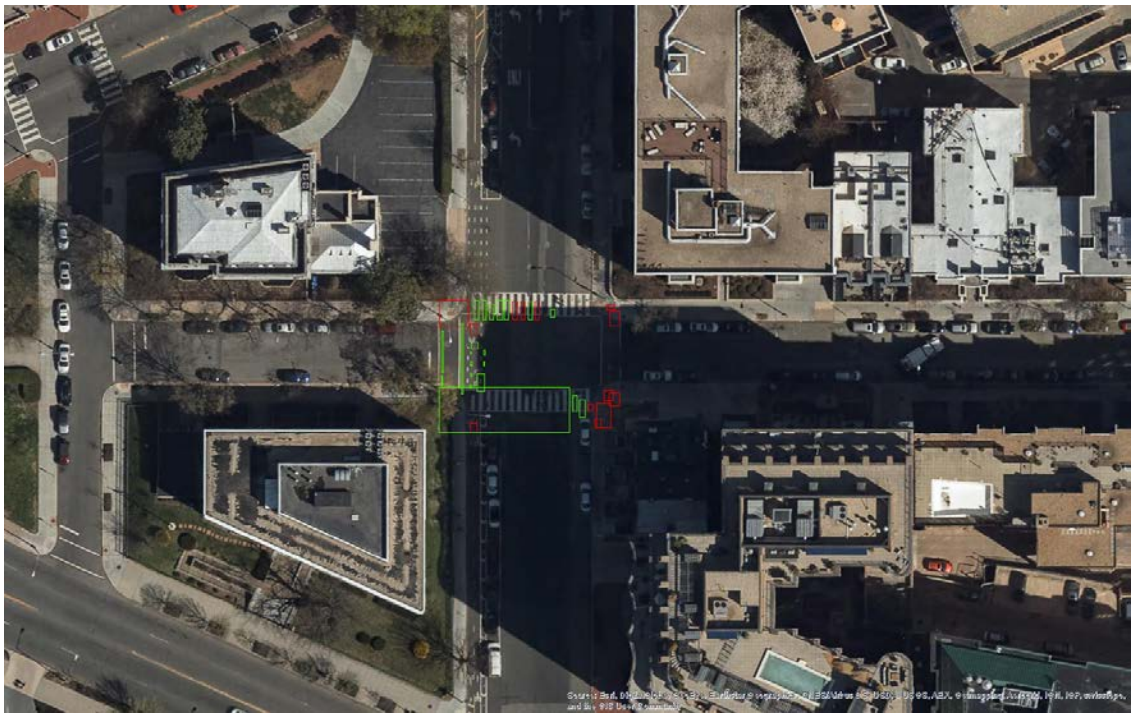

Figure 8(c): Detected pavement marking on 15th and N St NW

Figure 8(d): Detected pavement marking on 16th and Corcoran St NW

Accuracy rates can be further improved by training the model with more data and especially with better quality. For example Figure 9 shows an intersection at Church St with low data density which is reflected in the model prediction inaccuracy in Figure 8(b). Table 1 shows the total area to be recognized, area with bad data and actual recognized pavement markings by developed algorithm. It can be seen that good data is available only for approximately 77%.

Figure 9: Intersection level LiDAR data at 16th and Church St NW

Table 1: Intersection level LiDAR Data Statistics

Total Area Pavement markings area (Sq Ft)	Bad LiDAR data Area (Sq Ft)	Detected Area (Sq Ft)
9453	2178	5246

CONCLUSION AND FUTURE WORK

The research is pioneering work in using street level LiDAR data, as we continue working on algorithms and procure more data, accuracy will improve. Future work will concentrate on developing algorithms to generate missing LiDAR data which will improve accuracy. Immediate further work will focus on detecting road signs, road symbols and also reading the data. The generated data will be then converted into the MAP message and will be available for public usage to be implemented in connected vehicle demonstrations.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support received from District Department of Transportation (DDOT) for implementing the project especially from IT department for providing the data.

REFERENCES

[1] Vock, D. C. In Preparation for Driverless Cars, States Start Upgrading Roads.In *Governing*, governing.com, 2016.

[2] Kenney, J. B. Dedicated short-range communications (DSRC) standards in the United States. *Proceedings of the IEEE,* Vol. 99, No. 7, 2011, pp. 1162-1182.

[3] Xu, Q., T. Mak, J. Ko, and R. Sengupta. Vehicle-to-vehicle safety messaging in DSRC.In *Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks*, ACM, 2004. pp. 19-28.

[4] Bascón, S. M., J. A. Rodríguez, S. L. Arroyo, A. F. Caballero, and F. López-Ferreras. An optimization on pictogram identification for the road-sign recognition task using SVMs. *Computer Vision and Image Understanding,* Vol. 114, No. 3, 2010, pp. 373-383.

[5] Nieto, M., J. A. Laborda, and L. Salgado. Road environment modeling using robust perspective analysis and recursive Bayesian segmentation. *Machine Vision and Applications,* Vol. 22, No. 6, 2011, pp. 927-945.
[6] Asner, G. P., J. Mascaro, H. C. Muller-Landau, G. Vieilledent, R. Vaudry, M. Rasamoelina, J. S. Hall, and M. van Breugel. A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia,* Vol. 168, No. 4, 2012, pp. 1147-1160.
[7] Mascaro, J., M. Detto, G. P. Asner, and H. C. Muller-Landau. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment,* Vol. 115, No. 12, 2011, pp. 3770-3774.
[8] Sun, S., and C. Salvaggio. Aerial 3D building detection and modeling from airborne LiDAR point clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* Vol. 6, No. 3, 2013, pp. 1440-1449.
[9] Zhang, J., X. Lin, and X. Ning. SVM-based classification of segmented airborne LiDAR point clouds in urban areas. *Remote Sensing,* Vol. 5, No. 8, 2013, pp. 3749-3775.
[10] Huang, A. S., D. Moore, M. Antone, E. Olson, and S. Teller. Finding multiple lanes in urban road networks with vision and lidar. *Autonomous Robots,* Vol. 26, No. 2-3, 2009, pp. 103-122.
[11] Hernández, J., and B. Marcotegui. Filtering of artifacts and pavement segmentation from mobile lidar data.In *ISPRS Workshop Laserscanning 2009*, 2009.
[12] Urmson, C., J. Anhalt, D. Bagnell, C. Baker, R. Bittner, M. Clark, J. Dolan, D. Duggins, T. Galatali, and C. Geyer. Autonomous driving in urban environments: Boss and the urban challenge. *Journal of Field Robotics,* Vol. 25, No. 8, 2008, pp. 425-466.
[13] Kammel, S., and B. Pitzer. Lidar-based lane marker detection and mapping.In *Intelligent Vehicles Symposium, 2008 IEEE*, IEEE, 2008. pp. 1137-1142.
[14] Cremean, L. B., and R. M. Murray. Model-based estimation of off-highway road geometry using single-axis ladar and inertial sensing.In *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, IEEE, 2006. pp. 1661-1666.
[15] Vincent, P., H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research,* Vol. 11, No. Dec, 2010, pp. 3371-3408.
[16] Fischler, M. A., and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM,* Vol. 24, No. 6, 1981, pp. 381-395.
[17] Fontanelli, D., L. Ricciato, and S. Soatto. A fast ransac-based registration algorithm for accurate localization in unknown environments using lidar measurements.In *2007 IEEE International Conference on Automation Science and Engineering*, IEEE, 2007. pp. 597-602.
[18] Ester, M., H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise.In *Kdd, No. 96*, 1996. pp. 226-231.