CLASSIFICATION OF MOBILE TERRESTRIAL LIDAR POINT CLOUD IN URBAN AREA USING LOCAL DESCRIPTORS

M. Zaboli[1], H. Rastiveis[1,*], A. Shams[2], B. Hosseiny[1], W. A. Sarasua[3]

[1]Department of Photogrammetry and Remote Sensing, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran - (m_zaboli, hrasti, ben.hosseiny) @ut.ac.ir
[2]Advanced Highway Maintenance & Construction Technology (AHMCT) Research Center, University of California, Davis, CA, USA - asham@ucdavis.edu
[3]Glenn Department of Civil Engineering, Clemson University, Clemson, SC, USA - sarasua@clemson.edu

Commission II, WG II/3

KEY WORDS: Mobile Terrestrial LiDAR, Point Cloud, Classification, Geometric Features, Photogrammetry

ABSTRACT:

Automated analysis of three-dimensional (3D) point clouds has become a boon in Photogrammetry, Remote Sensing, Computer Vision, and Robotics. The aim of this paper is to compare classifying algorithms tested on an urban area point cloud acquired by a Mobile Terrestrial Laser Scanning (MTLS) system. The algorithms were tested based on local geometrical and radiometric descriptors. In this study, local descriptors such as linearity, planarity, intensity, etc. are initially extracted for each point by observing their neighbor points. These features are then imported to a classification algorithm to automatically label each point. Here, five powerful classification algorithms including *k*-Nearest Neighbors (*k-NN*), Gaussian Naive Bayes (GNB), Support Vector Machine (SVM), Multilayer Perceptron (MLP) Neural Network, and Random Forest (RF) are tested. Eight semantic classes are considered for each method in an equal condition. The best overall accuracy of 90% was achieved with the RF algorithm. The results proved the reliability of the applied descriptors and RF classifier for MTLS point cloud classification.

1. INTRODUCTION

A vehicle-mounted Light Detection and Ranging (LiDAR), known as Mobile Terrestrial Laser Scanning (MTLS), can provide an efficient and practical solution for acquiring three-dimensional (3D) point clouds along a roadway corridor. The sensor is scanning the building, landscapes, and other features at highway speed while recording positional data using a Global Positioning System (GPS), Inertial Measurement Unit (IMU), a Distance Measurement Instrument (DMI), and digital cameras simultaneously. The resulted point cloud contains highly accurate 3D locations of highway assets and nearby areas (Shams et al., 2018). However, automated extraction of useful information from the large volume of the ponts within a point cloud has been a challenging topic in Photogrammetry and Computer Vision.

Numerous algorithms for point cloud classification in urban areas have been developed to automatically extract objects from the derived point clouds. In this case, Yang and Dong (2013) proposed a shape-based segmentation method based on geometric descriptor calculations in optimal neighborhood size. The obtained segments were classified using a Support Vector Machine (SVM) algorithm to extract pole-shape objects, and an overall accuracy of 95% was reported. (Lehtomäki et al., 2015) developed a workflow to classify roadside objects including the removal of the ground and buildings through segmentation, classification, and object location estimation and achieve 88% accuracy. Li et al. (2017) proposed a novel method to recognize road furniture using their logical relations and functionalities. Their work achieved a strong performance in the interpretation of road furniture resulting in an accuracy in identifying 93.3% of poles, 94.3% of street light heads, and an overall accuracy of 76.9% was reported.

Deep learning based methods for object classification has recently become a trend in pattern recognition. In this regard, Huang and You (2016) introduced an innovative 3D point cloud labeling approach using 3D convolutional neural network (CNN). In their approach, the segmentation step was removed and, therefore, the prior knowledge of labeling was not necessary. Zhang and Zhang (2017) proposed a deep learning framework that includes a 3D-CNN, a deep Q-network, and a residual recurrent neural network for the efficient semantic segmentation of large-scale 3D point clouds. This network provides an automatic method for mapping a primary point cloud into considered categories to generate lightweight building models. Recently, Liu et al. (2018) compared the performance of two deep learning networks including fully CNN and patch-based CNN with two conventional classification methods (RF and SVM algorithms). The results of their experiments showed the superiority of the fully CNN in comparison with other methods. Although CNN-based approaches have shown promising results in classification, the extensive training required and numerous hyper-parameters to work with are more challenging in using these methods. Thus, researchers still use more conventional classification methods for point cloud processing and continued research is needed to improve on the methods especially in complicated areas.

The goal of this paper is evaluate different methods to automatically classify an MTLS point cloud of an urban corridor based on local descriptors. The contribution of this paper is to explore and define a number of descriptors in a local neighborhood based on five powerful classification methods including MLP, RF, SVM, *k-NN*, and GNB. The obtained results from these methods are then compared and analyzed. The remainder of this work will present the methodology and theoretical backgrounds in Section 2. In Section 3, the

* Corresponding author

experimental results are provided. Finally, the concluding marks of this research are gathered in Section 4.

2. METHODOLOGY

Point cloud classification algorithms usually consist of two consecutive steps: 1) descriptor extraction; and 2) semantic labeling. These steps are described in the following sections.

2.1 Descriptor extraction

Descriptor extraction in point cloud classification includes a process of acquiring geometric, and radiometric, the information in each points' local neighborhood. There are numbers of geometric or radiometric descriptors which may be used for point cloud classification including linearity, planarity, scattering, omni-variance, local curvature, and others (Chehata et al., 2009). Additionally, normal vectors were selected as spatial descriptors by Rabbani et al. (2008). Some new descriptors such as sorted normalized eigenvalues were introduced based on the respectively derived eigenvalues (λ_i) within a certain neighborhood to explore and quantize local 3D shapes. The normalized eigenvalues (Yang and Dong, 2013) ($e_i, i = 1,2,3$) is presented as follows:

$$e_i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \lambda_3} \ , i = 1,2,3 \tag{1}$$

The implemented descriptors in this study are placed in four groups: 1) eigenvalues, 2) radiometric, 3) geometric, and 4) spatial. In this case, normalized values of the eigenvalues were calculated using the Equation 1. These values were also used to calculate the eigenvalue descriptors such as linearity, planarity, scattering, and local curvature. Normal vector components (spatial descriptors), the value of intensity (radiometric descriptor) and the mean height (geometric descriptor) value in a specified neighborhood radius (height descriptor) were implemented for the classification process. To extract these descriptors for each point, neighborhood points in a 0.5 m radius were considered. These descriptors were selected by visual inspection and the effective separation of the objects. Table 1 summarizes the extracted descriptors of this study.

Descriptors	Definitions
Normalized eigenvalue (max)	e'_1
Normalized eigenvalue	e'_2
Normalized eigenvalue (min)	e'_3
Linearity	$L'_\lambda = \frac{e'_1 - e'_2}{e'_1}$
Planarity	$P'_\lambda = \frac{e'_2 - e'_3}{e'_1}$
Scattering	$S'_\lambda = \frac{e'_3}{e'_1}$
Local curvature	$C'_\lambda = \frac{e'_3}{e'_1 + e'_2 + e'_3}$
Height mean	$\frac{1}{n}\sum_{i=1}^{n} Z_i$
Normal vectors	N_x
	N_y
	N_z
Intensity	I

Table 1. List of the applied descriptors in the point cloud classification process

2.2 Semantic Classification

Extracted descriptors for each point are imported to the classifier to detect its label. The classification algorithm, number of classes, and the training data are key factors in this step. One can select the number and the type of semantic labels according to the available dataset. Moreover, training data for each class can be collected, manually. In this paper, five of the most conventional classification algorithms are implemented. They are k-NN, GNB, SVM, MLP, and RF. Each of these algorithms are briefly described in the following subsections.

2.2.1 Gaussian Naive Bayes (GNB)

Naive Bayes classifiers are simple and powerful algorithms for predictive modeling. The GNB is used for continuous data where the mean and standard deviation of the training data are used to show the distribution of the data (i.e. Normal distribution). The GNB algorithm requires the same number of descriptors and predictors for linear parameters in a learning problem (Bishop, 2006; Gatziolis and Andersen, 2008; Jung et al., 2019).

2.2.2 K-Nearest Neighbors (k-NN)

k-NN classification is a non-parametric method that is an easy-to-implement supervised machine learning algorithm for classification and regression. In the k-NN method, the label of each point is detected based on the k closest training samples in the descriptor space (Therrien, 1989). A point is classified by the majority of votes in its neighbor, with the object being assigned to the most common class among its k nearest neighbors. When k = 1, the point belongs to the class of the closest training point (Schutt and O'Neil, 2013). Thus, the k-NN algorithm should be run several times with different values of k to select the optimum value in a manner to reduce errors while maintaining the algorithm's ability to accurately make predictions (Friedman et al., 2001).

2.2.3 Support Vector Machine (SVM)

SVM is another classification method which produces input-output mapping functions from a group of labeled training data. This algorithm finds a hyperplane in an n-dimensional space (descriptor space) that distinctly classifies the data (Bishop, 2006). Mapping functions often transform the input data to a high-dimensional descriptor space by exploiting nonlinear kernel functions so the data in new space can be discriminated easier compared to the input space. The SVM kernel function accepts data as input and transforms it into the required form. Different SVM algorithms use various types of kernel functions (e.g. linear, nonlinear, polynomial, radial basis function (RBF), and sigmoid) (Therrien, 1989).

2.2.4 Multilayer Perceptron Neural Network (MLP)

MLP Neural Network is a feed-forward Artificial Neural Network (ANN) which consists of a minimum of three layers of nodes: an input layer, a hidden layer, and an output layer. The number of the hidden layers can be increased to make a model more complex (Looney, 1997). Except for the input nodes, each node is a neuron that uses a nonlinear activation function. The layers of an MLP are fully connected since each neuron in a layer is connected with a certain weight to all the neurons in the following layer. Each layer consists of independent units, where each unit has a unique weight. MLP utilizes a supervised learning technique for training called back-propagation. Using multiple layers and a non-linear activation function, MLP can easily classify data that is not linearly separable. Activation functions that introduce non-linearity into the network describe the input-output relations in a non-linear method. This provides a powerful model to be more flexible in describing optional relations. Popular activation functions are Sigmoid, Relu, Tanh and many more (Maiorov and Pinkus, 1999).

https://doi.org/10.5194/isprs-archives-XLII-4-W18-1117-2019

2.2.5 Random Forest (RF)

Random Decision Forests (RF) is an ensemble algorithm and learning method for classification that operates by constructing a multitude of decision trees as building blocks of the RF model at the training stage (Friedman et al., 2001). The output shows the label of the defined classes. The reason for the promising results of this method is that the trees protect each other from their individual. While some wrong trees are possible, some of the other trees will be right. Internal estimates check correlation, error, and strength are used to measure variable importance. These are used to show the reaction to increase the number of features used in the splitting. These ideas are also appropriate for regression (Breiman, 2001).

3. EXPERIMENTS AND RESULTS

3.1 Dataset

In this study, an urban roadway test section in Anderson, South Carolina, USA was selected to perform the research evaluation. The MTLS dataset is of a 600 m section of US Highway Route 76 (Clemson Blvd), is a 4-lane urban arterial beginning at Forest Hill Drive and ending at the intersection with East-West Parkway. From this dataset, a 61m section including over 3.5 million points was selected as test data. Figure 1 shows a 3D view of the selected section. As shown in this figure, the selected area has various objects including Powerlines, Poles, Buildings, and Trees.

(a)

(b)

Figure 1. The selected test section from a 600 m dataset. (a) The selected point cloud; (b) Street view of test area from Google Maps, 2019.

3.2 Results

In this study, the descriptors in Table 1 were extracted for all of the points of the dataset. Using these descriptors, eight semantic classes (i.e., Cars, Trees, Poles, Powerlines, Buildings, Billboards, Asphalt road, and Sidewalk) were considered in the classification methods. In the k-NN classification, k = 5 was the optimum value to increase the computation speed. In the MLP Neural Network, considering two hidden layers with 100 and 75 neurons, and a 0.000005 learning rate, the network was trained after 450 epochs. An RBF kernel was used in the SVM classification method, to increase the efficiency. In all tested methods, 70% of the input selected point cloud was used for training, while the other 30% for the evaluation. The training and test data were selected randomly at the beginning of the classification process and common points were used for all classification methods.

For better representation of the classification results by the selected methods, the different classes were separated by distinct colors and shown in Figure 2. The figure shows that the RF and k-NN methods seem to do well identifying Asphalt road, Powerlines, Buildings, Trees, Poles, and Cars but in Sidewalk class these methods performed poorly. The other methods seem less successful.

3.3 Accuracy Assessment

To evaluate the performance of the classification methods, all of the points in the dataset were manually divided into the selected classes. Comparing the manually divided points and the classification results, four accuracy assessment measures for each method were obtained. They are precision (Pr.), recall (R.), F1-score (F1), and overall accuracy (OA). By definition (Sun et al., 2018), these parameters are measured with true positive (TP), false positive (FP), true negative (TN), and false negative (FN) from the confusion matrix according to Equations 2-5.

$$precision = \frac{TP}{\mathrm{TP}\ + FP} \tag{2}$$

$$recall\ = \frac{\mathrm{TP}}{\mathrm{TP}\ +\ \mathrm{FN}} \tag{3}$$

$$F1 =\ 2 * \frac{precision * recall}{precision + recall} \tag{4}$$

$$ACC = \frac{TP + TN}{TP + FN + FP + TN} \tag{5}$$

The resulting accuracy assessment measures for all classification methods are presented in Table 2. Figure 3 shows a bar chart of the average values of all of the classes for each measure. The measures in figure 3 are distinguished by colors with taller bars representing better results. Figure 4 shows the evaluation of the applied classification methods for considered semantic classes.

Method	**Measures**	**Class**							
		Cars	Trees	Poles	Powerlines	Asphalt road	Sidewalk	Buildings	Billboards
k-NN	Pr.	0.58	0.63	0.75	0.83	0.99	0.96	0.85	0.77
	R.	0.56	0.61	0.73	0.86	0.99	0.96	0.85	0.77
	F1	0.57	0.62	0.74	0.84	0.99	0.96	0.85	0.77
GNB	Pr.	0.26	0.20	0.29	0.27	0.98	0.44	0.60	0.40
	R.	0.05	0.16	0.77	0.73	0.65	0.83	0.35	0.48
	F1	0.08	0.18	0.42	0.39	0.78	0.57	0.44	0.56
SVM	Pr.	0.47	0.53	0.81	0.60	0.96	0.93	0.66	0.66
	R.	0.13	0.24	0.45	0.64	0.99	0.85	0.90	0.65
	F1	0.21	0.33	0.58	0.62	0.97	0.89	0.76	0.66
MLP	Pr.	0.51	0.62	0.75	0.65	0.96	0.89	0.67	0.62
	R.	0.13	0.15	0.51	0.70	0.98	0.87	0.91	0.67
	F1	0.20	0.24	0.60	0.67	0.97	0.88	0.78	0.65
RF	Pr.	0.70	0.71	0.79	0.86	0.99	0.97	0.87	0.85
	R.	0.66	0.73	0.77	0.86	1	0.96	0.90	0.80
	F1	0.68	0.72	0.78	0.86	0.99	0.97	0.88	0.82

Table 2. Evaluation of the applied classification methods for MTLS point cloud classification.

https://doi.org/10.5194/isprs-archives-XLII-4-W18-1117-2019 | © Authors 2019. CC BY 4.0 License.

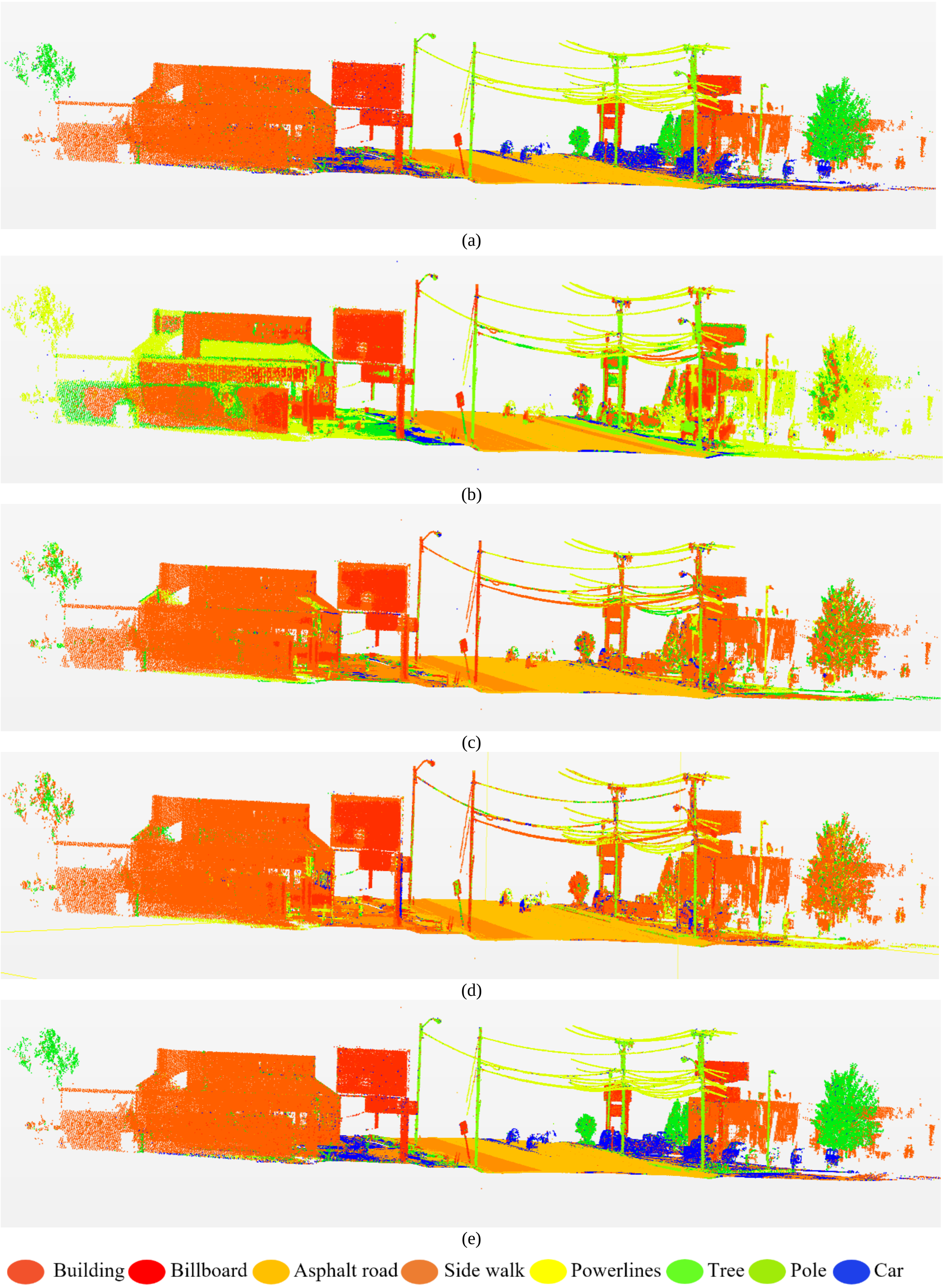

Figure 1. Classification Results of methods: (a) *k-NN*; (b) GNB; (c) SVM; (d) MLP; (e) RF.

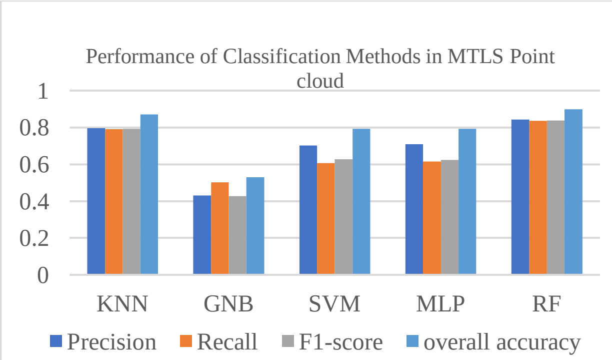

Figure 3. Evaluation of the applied classification algorithms for MTLS point cloud classification in urban area.

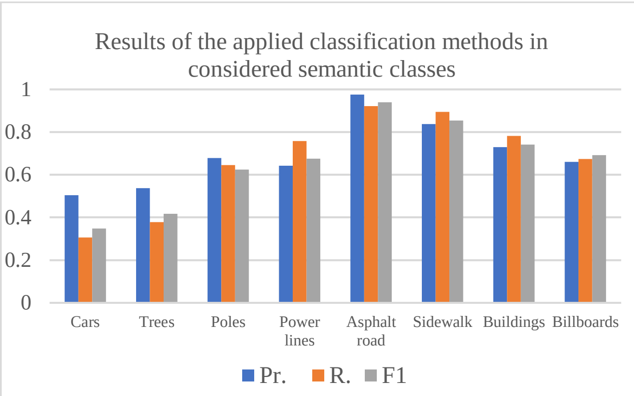

Figure 4. Evaluation of the applied classification methods for considered semantic classes

4. DISCUSSION

As shown in Figure 3, the RF and *k-NN* methods achieved good precision in separating classes from each other. The SVM and MLP methods achieved similar results in mean values and overall accuracies. The poor results of the GNB method in mean values and overall accuracy indicate that this method is inappropriate for the classification of this point cloud.

According to Table 2, the RF method outperformed all of the other methods for almost every class and every measure. Almost all methods successfully extracted the Asphalt road and Sidewalk with a high level of accuracy except the GNB. Most of the methods failed to extract trees and cars while the GNB had the worst, and the RF and the *k-NN* had the best performance. According to Figure 4, the best result among the different classes was obtained for the Asphalt road class and the worst result belongs to the class of cars.

In Poles extraction, the best precision was obtained using the SVM method, however, the best overall result of this class was achieved using RF, *k-NN*, and MLP, respectively. Similar results were obtained for the Powerlines class, with the worst result again obtained using the GNB method. In Buildings class, RF method achieved the best precision, recall, and F1, while other methods except for GNB, obtained high recall in this class. The *k-NN* method achieved better results than SVM and MLP in this class. The GNB method had low recall and F1 in recognizing the class of Buildings. The same ranking with slightly different accuracy was obtained in the class of Billboards. Among all classes, the worst results were obtained in the Trees class using GNB method, while an almost perfect result was obtained in using RF method to extract Asphalt road. In general, RF method achieved the best results among the others in all classes and recognized Asphalt road, Sidewalk, Buildings, Powerlines, Billboards, Poles, Trees and Cars with the best precision, respectively. The *k-NN* was the runner-up with the best precision, recall and F1 using *k-NN* achieved in Asphalt road extraction.

Since the two methods of RF and *k-NN* have almost identical results, we compare the methods together. Asphalt road and Sidewalk classes were recognized with similar precision, recall and F1. RF method obtained significantly better results in the extraction of Billboards, Poles, Trees, and Cars, compared to the *k-NN* method. In the extraction of other classes such as Powerlines and Buildings, both methods achieved close results, though RF had better results.

SVM and MLP methods also obtained similar results in all classes with overall accuracy of 79% for both of them. These two classes also achieved the best precisions in the Asphalt road class and separated the Sidewalk class more accurately than the two previous methods. But these methods did not perform well in detecting Trees and Cars, and they recognized Powerlines, Poles and Billboards with precision of about 50%. Moreover, these methods achieved high recall value in Buildings class.

Like other methods, the Asphalt class has the best precision compared to other classes that are labeled by the GNB method and in the Buildings class obtained precision of about 50%. This method performed poorly in all other classes and achieved an overall accuracy of 53%. Therefore, this method is not suitable for classifying point cloud with a high number of classes~~,~~ and does not produce accurate results.

The results of the study showed that the use of selected methods to identify most descriptors from the point cloud of urban areas is feasible. However, all methods except RF and *k-NN* failed to detect Trees and Cars with relatively acceptable accuracy. It may be due to inadequate data from the mentioned classes in the dataset. For example, the points belonging to Asphalt road and Sidewalk were much higher than Cars and Trees. The other reason could be inequality of data, due to MLP and SVM methods be in sensitive to the unequal amount of training data in different classes. Therefore, when the numbers of points presenting different classes are close, the higher precision in classification could be expected. Also, the same constant values were used in applying different methods, so the variation of the values in the different method could improve the overall accuracy.

5. CONCLUSION

In this study, five classification methods were implemented for assigning eight semantic classes (i.e., Cars, Trees, Poles, Powerlines, Asphalt road, Sidewalk, Buildings and Billboards) based on local descriptors of each point. These descriptors were selected based on how effectively they were able to distinguish between classes based on visual observation. In this work, a variety of radiometric descriptor, height descriptor, spatial descriptors, and eigenvalues descriptors were used to accurately classify the objects. Based on a neighborhood radius of 0.5 m for each point, the descriptors were calculated for all points in the dataset, and then the training data was prepared that consisted of 70% for training and the remainder for testing the selected methods. The classification results using the selected methods,

the values of overall accuracy, precision, recall and F1-score were calculated for all classes.
The best overall accuracy of 90% was reported using RF. Also, *k-NN* method achieved very good results and similar to the RF method in classifying the classes with an overall accuracy of 87%. SVM and MLP achieved approximately similar accuracy of 79% which is acceptable, however, GNB method with overall accuracy of 53% and it is not recommended as an appropriate method for classifying point cloud with a large number of classes. Asphalt road, Sidewalk, and Buildings classes were well detected by most of the algorithms except GNB. While considering the result of only GNB method shows higher accuracy in detecting Asphalt road and Sidewalk. Whereas, Cars and Trees were poorly detected using all algorithms.

Considering other geometric and radiometric descriptors as well as the other classification methods to achieve a better evaluation are strongly suggested. Furthermore, the performance of this method regarding computation cost and their robustness in relation to the size of training data is necessary for future studies.

REFERENCES

Bishop, C.M., 2006. Pattern recognition and machine learning. springer.

Breiman, L., 2001. Random forests. Machine learning 45, 5-32.
Chehata, N., Guo, L., Mallet, C., 2009. Airborne lidar feature selection for urban classification using random forests. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38, W8.

Friedman, J., Hastie, T., Tibshirani, R., 2001. The elements of statistical learning. Springer series in statistics New York.

Gatziolis, D., Andersen, H.-E., 2008. A Guide to LIDAR Data Acquisition and Processing for the Forests of the Pacific Northwest, United States Department of Agriculture.

Huang, J., You, S., 2016. Point cloud labeling using 3d convolutional neural network, 2016 23rd International Conference on Pattern Recognition (ICPR). IEEE, pp. 2670-2675.

Jung, J., Che, E., Olsen, M.J., Parrish, C., 2019. Efficient and Robust Lane Marking Extraction from Mobile Lidar Point Clouds. ISPRS Journal of Photogrammetry and Remote Sensing 147, 1-18.

Lehtomäki, M., Jaakkola, A., Hyyppä, J., Lampinen, J., Kaartinen, H., Kukko, A., Puttonen, E., Hyyppä, H., 2015. Object classification and recognition from mobile laser scanning point clouds in a road environment. IEEE Transactions on Geoscience and Remote Sensing 54, 1226-1239.

Li, F., Elberink, S.O., Vosselman, G., 2017. Semantic labelling of road furniture in mobile laser scanning data. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences 42, Volume XLII-2/W7, 2017.

Liu, T., Abd-Elrahman, A., Morton, J., Wilhelm, V.L., 2018. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. GIScience & remote sensing 55, 243-264.

Looney, C.G., 1997. Pattern recognition using neural networks: theory and algorithms for engineers and scientists. Oxford University Press, Inc.

Maiorov, V., Pinkus, A., 1999. Lower bounds for approximation by MLP neural networks. Neurocomputing 25, 81-91.

Rabbani, S.I., Devi, K., Khanam, S., 2008. Effect of thiazolidinediones on the erythropoeitic and germinal cells in the male wistar rats. Clinical medicine. Oncology 2, CMO. S678.

Schutt, R., O'Neil, C., 2013. Doing data science: Straight talk from the frontline. O'Reilly Media, Inc.

Shams, A., Sarasua, W.A., Famili, A., Davis, W.J., Ogle, J.H., Cassule, L., Mammadrahimli, A., 2018. Highway Cross-Slope Measurement Using Mobile LiDAR. Transportation Research Record 2672, 88-97.

Sun, Z., Xu, Y., Hoegner, L., Stilla, U., 2018. Classification of MLS point clouds in urban scenes using detrended geometric features from supervoxel-based local contexts. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences 4.

Therrien, C.W., 1989. Decision, Estimation, and Classification: An Introduction to Pattern Recognition and Related Topics. Wiley.

Yang, B., Dong, Z., 2013. A Shape-Based Segmentation Method for Mobile Laser Scanning Point Clouds. ISPRS Journal of Photogrammetry and Remote Sensing 81, 19-30.

Zhang, L., Zhang, L., 2017. Deep learning-based classification and reconstruction of residential scenes from large-scale point clouds. IEEE Transactions on Geoscience and Remote Sensing 56, 1887-1897.