LiDAR-based Lane Marking Extraction through Intensity Thresholding and Deep Learning Approaches: A Pavement-based Assessment

Yi-Ting Cheng [1, †], Ankit Patel [1, †], Darcy Bullock [1], Ayman Habib [1, *]

[1] Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Dr., West Lafayette, IN 47907, USA - (cheng331, pate1019, darcy, ahabib)@purdue.edu

Commission III, WG III/5

KEY WORDS: Asphalt Pavement, Concrete Pavement, Lane Marking Extraction, Intensity Normalization, Deep Learning, Automated Labeling

ABSTRACT:

With the rapid development of autonomous vehicles (AV) and high-definition (HD) maps, up-to-date lane marking information is necessary. Over the years, several lane marking extraction approaches have been proposed with many of them based on accurate and dense Light Detection and Ranging (LiDAR) point cloud data collected by mobile mapping systems (MMS). This study proposes a normalized intensity thresholding strategy and a deep learning strategy with automatically generated labels. The former extracts lane markings directly from LiDAR point clouds while the latter utilizes 2D intensity images generated from the LiDAR point cloud. Additionally, the proposed approaches are also compared with state-of-the-art strategies such as original intensity thresholding and a deep learning approach based on manually established labels. Finally, each strategy is evaluated in asphalt and concrete pavements separately to assess their sensitivity to the nature of pavement surface. The results show that the deep learning model trained with automatically generated labels performs the best in both asphalt and concrete pavement area with an F1-score of 84.9% and 85.1%. In asphalt pavement area, original intensity thresholding strategy shows a lane marking extraction performance comparable to the other strategies while in concrete pavement area, it is significantly poor with an F1-score of 65.1%. Between the proposed normalized intensity thresholding and deep learning model trained with manually labeled data, the former performs better in asphalt pavement area while the latter obtains better results in concrete pavements.

1. INTRODUCTION

With the advent of autonomous vehicles (AV) and advanced driver assistance systems (ADAS), high-definition (HD) maps with lane-level details such as pedestrian crosswalks, signalized intersection, and bike lanes are necessary for navigation and route planning. Lane markings form an integral part of such maps, and thus their extraction is essential. In addition, clearly identifiable lane markings are critical elements of traffic management systems and accident mitigation. Especially in populated urban areas, worn-out lane markings have led to many car accidents [1]. Therefore, it is required to provide detailed and up-to-date information about lane markings along the road surface.

A number of studies have detected lane markings through imagery or videos; however, these approaches cannot provide precise information about the reflective properties of lane markings. In order to accurately evaluate the quality of lane markings, LiDAR point clouds are chosen in this research since they can be obtained in a short interval of time with high density and accuracy without being affected by weather, lighting, or occlusions. Additionally, the intensity information provided by LiDAR can be used to assess the quality of lane markings by departments of transportation for road maintenance operations.

Since lane markings are retro-reflective materials painted on low albedo pavements (asphalt or concrete), the extraction using LiDAR point clouds mainly depends on intensity thresholding. Many researchers also rasterized the point cloud into an intensity image for lane marking extraction to reduce computations. Thus, LiDAR data-based strategies can be classified into two categories based on input data: (1) 3D LiDAR point cloud-based extraction, and (2) 2D LiDAR intensity image-based detection. For extraction of lane markings directly from a point cloud, Yu et al. [2] at first partitioned the road surface point cloud into multiple blocks across the driving direction to account for varying point density and intensity distribution. Subsequently, Otsu's threshold was applied to extract candidate lane markings. Finally, false positives were eliminated by first defining a spatial density filter. Points with a spatial density less than a threshold would be considered as noise and hence removed. Yan et al. [3] divided the LiDAR data into scanlines which are computationally less expensive to process. Non-lane marking points were then removed through intensity-based filtering that preserved lane marking points at the edge. Lastly, all points located between edge points were extracted as lane marking points. Yang et al. [4] implemented an adaptive block and multi-threshold detection method to account for intensity variation in the point cloud due to scanner-to-object distance and incidence angle. The road surface point cloud at first was segmented along the driving direction. Then, each segment was divided into several blocks across the driving direction. For each block in each segment, Otsu's threshold and a gradient threshold were calculated. The latter depends on the average intensity value of the center block, maximum block width, and the number of blocks in each section. The greater of these two thresholds was considered optimal to binarize point clouds in a block for lane marking extraction.

For lane marking extraction from LiDAR intensity images, Guan et al. [5] used an Inverse Distance Weighting (IDW) strategy to calculate pixel values in intensity images generated from road surface point clouds. Multiple scanning-distance-based thresholds were then applied to extract lane markings in

---

* Corresponding author
† Indicates equal contribution

the intensity images. The markings were further refined by application of Otsu's thresholding and morphological operations. Kumar et al. [6] began by generating two raster images based on intensity and range values. Thereafter, a range and cross-slope value-based threshold was applied for lane marking extraction. Morphological operations were then utilized to remove outliers. Soilán et al. [7] first hypothesized a Gaussian Mixture Model for intensity distribution. They defined two classes – a class with a larger fraction of points and low-intensity values belonging to pavement and a class with high-intensity values and a smaller number of points belonging to lane markings. Each point is assigned to a class based on higher posterior probability. Intensity images were then created from the high-intensity class to ensure minimum points are processed. Finally, the lane markings are extracted through application of Otsu's threshold and an area-based filter. Yao et al. [8] proposed another LiDAR intensity image-based lane marking extraction approach. The image pixel values were calculated based on IDW interpolation where the chosen pixel size is close to point spacing. An adaptive thresholding strategy was applied to extract lane marking by first generating an integral image from the original intensity image [9]. In an integral image, each pixel value is a sum of the top-left rectangular area of the original image. Then, the original image was binarized based on the sum of pixel values in its neighborhood in the integral image.

With the ever-growing popularity of deep learning architectures, researchers have resorted to various image segmentation convolutional neural networks for extracting lane markings from intensity images. He et al. [10] proposed a lane marking detection strategy based on a Segnet inspired architecture which is a fully convolutional neural network (FCNN). They first normalized the intensity values in point cloud based on their mean and standard deviation, followed by their scaling to a range of 0 to 255. They generated 12,729 intensity images with cell size equivalent to an area of 1 cm^2. Out of these images, 2,729 were used for training the network. Wen et al. [11] also adopted an FCNN approach for lane marking extraction. They generated intensity images from road surface point cloud at a resolution of 4 cm. 3,000 images in highway and urban areas were manually labeled to train a U-net model. Another training dataset of 1,000 images was curated from an underground garage to train the second U-net model. While learning-based approaches are robust to intensity and point density variation, generating labeled training data is a major hurdle that prevents their ubiquitous adoption.

An important aspect of lane marking extraction with LiDAR intensity values is the nature of pavement surface. Adrian et al. [12] found that the average luminance of concrete pavements was 1.77 times that of asphalt pavements. This means that concrete surfaces, in general, are more reflective. Puttonen et al. [13] measured spectral and directional reflection properties of asphalt and concrete surfaces under sun exposure. Brightness was characterized by a bidirectional reflection factor which ranged between 0.14 to 0.41 for concrete surface and between 0.27 to 0.32 for asphalt surface over the electromagnetic spectrum of 400 nm to 2500 nm. Moreover, concrete surfaces showed higher reflectance in 1000 nm to 1500 nm spectrum which is typically the wavelength of laser beams in LiDAR scanners. These studies indicate that performance of lane marking extraction strategies based on LiDAR intensity values must be evaluated with respect to pavement surface.

Based on above literature, it is evident that there is a gap in analyzing the sensitivity of LiDAR-based lane marking extraction strategies to pavement material (asphalt and concrete) even though reflectivity depends, to a large extent, on nature of pavement surface. Further, intensity thresholding strategies

obtain a high number of false positives due to intensity variation caused by object-to-scanner distance and incidence angle. Even though adaptive thresholds or intensity calibration are sought as solutions, they require prior knowledge or assumptions for modeling intensity distribution. On the other hand, though deep learning approaches overcome such issues, they are still marred by the fact that a huge amount of data must be labeled manually for them to be effective. These challenges have been addressed in this paper whereby two approaches for lane marking extraction are proposed. In summary, the main contributions of this research are:

1. A 3D point cloud-based normalized intensity thresholding strategy is implemented for lane marking extraction where the normalization can be applied independent of the reference target.
2. A 2D intensity image-based deep learning strategy is also developed where labeled data is generated through an automated procedure. Thus, abundant training samples are generated in a short interval of time.
3. These approaches are compared with state-of-the-art strategies such as original intensity thresholding and deep learning based on manually labeled data over 3 datasets.
4. Based on the hypothesis that LiDAR data-driven lane marking extraction is sensitive to pavement surface nature, all four strategies have been separately evaluated in asphalt and concrete pavements.

2. MOBILE MAPPING SYSTEMS DESCRIPTION

A wheel-based MMS is used in the research, as shown in Figure 1. It is equipped with four 3D LiDAR scanners: three Velodyne HDL-32E and one Velodyne VLP-16 Puck Hi-Res. The HDL-32E scanner consists of 32 radially oriented laser rangefinders while VLP-16 has 16 of them. The specifications of both kinds of LiDAR scanners are list in Table 1. Additionally, this MMS is also mounted with three FLIR Grasshopper3 9.1MP GigE cameras (two forward-facing and one rear-facing). All the cameras are synchronized to capture RGB imagery with a maximum resolution of 9.1 MP at a rate of 1 frame per second per camera. The LiDAR and imaging sensors are geo-referenced by an Applanix POSLV 220 GNSS/Inertial Measurement Unit (IMU) navigation system. The GNSS collection rate is 20 Hz, and the IMU measurement rate is 200 Hz.

LiDAR Scanner	Field of View	Points Captured /sec	Range
Velodyne HDL-32E	**Vertical**: 40° (-30.67° to 10.67°) **Horizontal**: 360°	7×10^5	100 m (accuracy of ± 2 cm)
VLP-16 Puck Hi-Res	**Vertical**: 20° (-10° to 10°) **Horizontal**: 360°	3×10^5	100 m (accuracy of ± 3 cm)

Table 1. Specifications of LiDAR scanners mounted on the wheel-based mobile mapping systems (MMS)

To geo-reference point clouds from the different LiDAR scanners, mounting parameters between the onboard LiDAR scanners and GNSS/IMU unit are estimated through a system calibration strategy [14]. Simultaneously, a forward and backward projection between the reconstructed point cloud and RGB imagery can be achieved using the trajectory information and mounting parameters, which were estimated using the LiDAR-camera calibration procedure [15]. The projection facilitates the assessment of the performance of the different lane marking extraction strategies. An example is illustrated in Figure 2 where corresponding image and LiDAR point cloud are shown. The magenta circle in the former is projected onto

the corresponding LiDAR point cloud (displayed as a red dot). Hereafter, a red dot will represent a location in the LiDAR point cloud, while magenta circle will display the same location in RGB imagery.
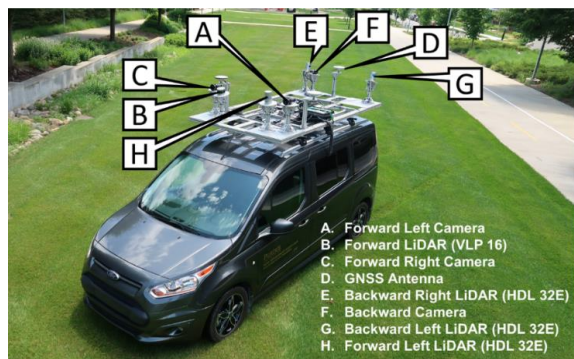

Figure 1 The wheel-based mobile mapping systems (MMS) used in this research

Figure 2. Projection of a location (empty magenta circle) in a RGB imagery onto the corresponding LiDAR Point cloud (red dot) using the estimated LiDAR/camera/GNSS/IMU system calibration parameters

3. MOBILE MAPPING SYSTEMS DATASETS

All strategies in this research were evaluated on two datasets collected over an interstate highway and one acquired over a rural highway. The used sensor, length of asphalt or concrete pavements, average local point spacing (LPS) [16], and average driving speed are listed in Table 2. The location of each dataset along with regions of concrete pavement is shown in Figure 3.
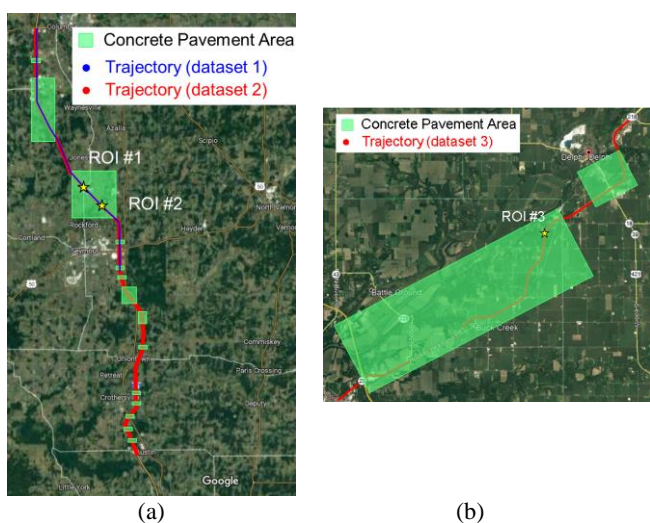

Figure 3. Trajectory, concrete pavement distribution, and regions of interest (ROIs) for generating intensity normalization maps of each LiDAR-based MMS dataset in (a) dataset 1 and 2, and (b) dataset 3

Dataset	Used Sensors	Length of pavement (mile)	Average LPS	Average Speed
1	HDL32E-F [1] HDL32E-L [1] HDL32E-R [1]	Asphalt: 15.55 Concrete: 2.49	3.11 cm	45.62 mph
2	HDL32E-F [1] HDL32E-L [1] HDL32E-R [1] VLP16	Asphalt: 27.65 Concrete: 6.22	3.19 cm	47.42 mph
3	HDL32E-F [1] HDL32E-L [1] HDL32E-R [1] VLP16	Asphalt: 2.23 Concrete: 13.06	3.16 cm	47.70 mph
[1] HDL32E-F, HDL32E-L, and HDL32E-R denote different LiDAR sensors of the same model				

Table 2. Description of the three LiDAR-based MMS point clouds

4. METHODOLOGY

Figure 4 presents the proposed framework for lane marking extraction in this research. The road surface point cloud is at first extracted from the MMS-based LiDAR point cloud [17]. Thereafter, the road surface point cloud is directly processed through original and normalized intensity thresholding strategies to obtain lane markings. On the other hand, for deep learning-based detection, intensity images are generated from the same road surface point cloud. Two U-net models are trained: one on manually-established labels and another on automatically-generated labels. The latter is obtained from the lane markings extracted through the normalized intensity thresholding strategy. Finally, the performance of these strategies is compared in asphalt and concrete pavement areas.
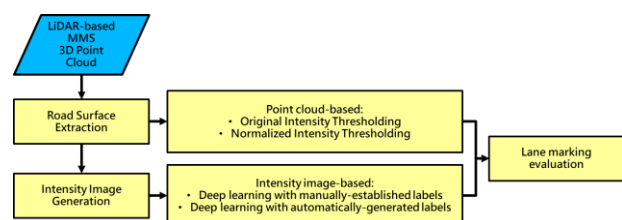

Figure 4. Flowchart of the proposed framework for lane marking extraction

The rest of this section describes lane marking extraction strategies implemented in this research Section 4.1 elaborates on 3D point cloud-based lane marking extraction: original and normalized intensity thresholding strategies, collectively referred to as "intensity thresholding approaches". Section 4.2 details the lane marking detection from 2D intensity images through deep learning strategies based on manually-derived and automatically-established labels, denoted as "deep learning approaches".

4.1 Point Cloud-based: Intensity Thresholding Approaches

In order to evaluate the performance of the proposed normalized intensity thresholding strategy, a state-of-the-art strategy – original intensity thresholding – is implemented [17]. This strategy extracts hypothesized lane markings by thresholding original intensity values using a 5th percentile threshold. Then, a noise removal strategy is applied to obtain final lane markings. However, this strategy obtains a large number of false positives, especially in highly reflective concrete pavement regions. On the other hand, for the normalized intensity thresholding strategy, the intensity values are normalized before applying a threshold to reduce false positives. This normalization procedure assumes the same objects should

exhibit similar intensity values acquired from different laser beams [18, 19]. Specifically, the normalized intensity value corresponding to an original value observed by a particular beam is calculated as the conditional expectation of intensity observed by other beams for the regions where that beam observed the given intensity value. The normalized intensity thresholding strategy includes the following five steps:

1. For a given dataset, a small section is randomly chosen from the road surface point cloud.
2. The LPS of the small road surface point cloud is evaluated for determining the cell size [16]. Then, the point cloud is gridded into cells. For each cell, intensity values and laser beam IDs of the points that lie within its bounding box are stored.
3. For a given laser beam b that recorded an intensity value i, all the cells containing the pair (b, i) are searched. The average intensity over these cells is computed, ignoring values observed by beam i. The resultant average is the normalized intensity of (b, i), and a lookup table (LUT) stores all the original and normalized intensity values.
4. The original intensity values of the whole dataset are then normalized using the LUT.
5. Hypothesized lane markings can be extracted from the normalized road surface point cloud using the 5th percentile intensity threshold.
6. The false positives of derived hypothesized lane markings are removed using the strategy proposed by Ravi et al. [17].

It is worth noting that the small road surface point cloud randomly selected to generate LUTs belonged to a concrete pavement area where high-intensity values are observed for both pavement surface and lane markings. It is also assumed that there is no negative effect on the intensity contrast between lane markings and asphalt pavements upon normalization even though the small segment of the road surface point cloud belongs to a concrete region. This is evident by the results presented later in section 5.1. Another important consideration in LUT generation is the designated cell size. In this research, the cell size is determined based on a multiplication factor threshold (Th_{MF}) and the LPS of the selected point cloud [16]. In addition, LUTs are generated for each dataset because the driving speed varies from one dataset to another. Additionally, the number of laser beams and scanning orientation changes the LPS of LiDAR scanners mounted on our MMS, so the selected point cloud is split according to the used scanners. The above steps in this strategy are applied to the selected point cloud of each sensor in each dataset.

4.2 Intensity Image-Based: Deep Learning Approaches

As stated earlier, two U-net models [20] are trained using manually-established and automatically-generated labels. These deep learning approaches include the following four steps:

1. Road surface point cloud is at first partitioned into point cloud blocks. Then, these blocks are rasterized into intensity images
2. Intensity images are manually labeled for the first U-net model (referred to as "U-net model 1"). On the other hand, for the second U-net model (referred to as "U-net model 2"), labels are generated automatically using the lane markings obtained from the normalized intensity thresholding strategy.
3. The manually-established and automatically-generated labels are utilized for training U-net model 1 and 2, respectively.
4. The lane markings in intensity images are detected through the trained U-net models.

For intensity image generation, the most important factor is the cell size. An optimal cell size avoids time-consuming calculations and maintains an adequate level of lane marking details in the image. In addition, the width of mapped roads, as well as the LPS of available data, are the other two important considerations in cell size selection. The width of highways surveyed in this research varied between 12 to 16 meters (including shoulder width). Thus, the road surface point cloud is partitioned into blocks of length 12.8 meters along the driving direction, as shown in Figure 5 (a) and required a fixed input image size of 256×256. Thus, resizing along both dimensions of the block is minimized without negatively impacting the level of detail in the intensity image. After the point cloud partitioning, the two-step enhancement is applied to generate intensity images. The first intensity enhancement is applied to the point cloud blocks by selecting a threshold (Th_{EN}) – 5th intensity percentile – where the intensity values greater than Th_{EN} are set to 255, while lower intensity values remain the same. Then, the enhanced point cloud block, as shown in Figure 5 (b), is converted into intensity images, as depicted in Figure 5 (c). A pixel value is calculated as the average of the intensity of all points falling within it. Finally, the second enhancement is applied to the intensity images – using a 5th intensity percentile threshold. The enhanced intensity image, as shown in Figure 5 (d), hereafter referred to as "intensity image," is used for labeling, training, and detection. This dual enhancement (for the point cloud block and the intensity image) amplifies the lane marking pixel values for easier detection by the U-net model.
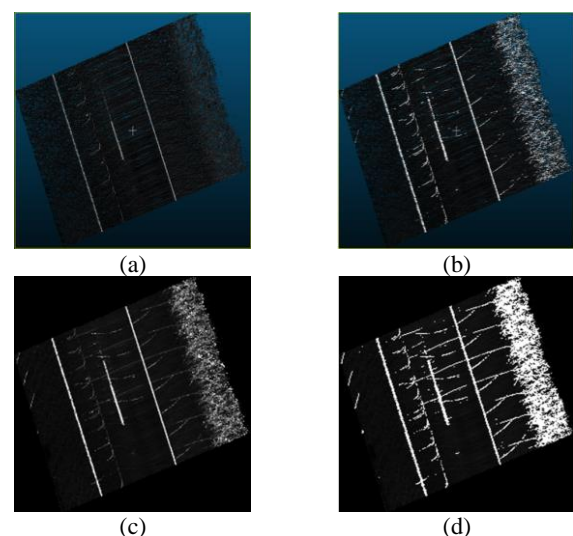

(a)	(b)
(c)	(d)

Figure 5. Illustrations of (a) original road surface point cloud block (b) enhanced point cloud block, (c) intensity image, and (d) enhanced intensity image

After the intensity image generation, some intensity images are manually labeled to train U-net model 1, while for U-net model 2, the lane markings derived through normalized intensity thresholding strategy are utilized for generating labels automatically. First, lane markings obtained from the normalized intensity thresholding, as presented in Figure 6 (a), are divided into 12.8-m-long blocks and converted into images with a 5cm pixel size and image size of 256×256. Then, in order to improve the spatial structure of lane markings, the boundary pixels of a lane marking are used to generate a bounding box, as displayed in Figure 6 (b). All the pixels within this bounding box are labeled as lane markings, as depicted by Figure 6 (c). These labeled images are utilized for U-net model 2 training. For robust training of both U-net models, data augmentation techniques are applied during each epoch such as: (i) clockwise

image rotation in the range from 0° to 180°, (ii) image resizing between 80% to 120% of its original size to simulate zoomed in and out images respectively, and (iii) horizontal flip.
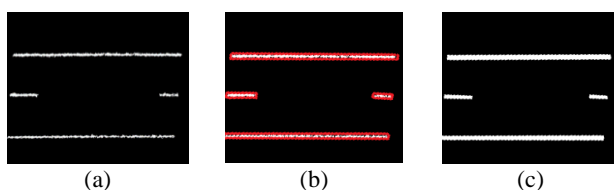

(a)　　　　　　(b)　　　　　　(c)

Figure 6. Automatic label generation: (a) normalized intensity thresholding lane markings, (b) bounding boxes (in red) enclosing lane markings, and (c) labeled image.

The U-net architecture is shown in Figure 7. The left path of the network is called encoder and the right part is referred to as decoder. A loss function based on the Dice coefficient [21] is adopted because of the skewed distribution of lane marking and non-lane marking class in intensity images. The dice coefficient quantifies how well the two classes overlap. It is defined as in Equation (1), where y_{true} and y_{pred} represent the ground truth and predicted pixels for the lane markings. Each pixel can be classified as either non-lane marking or lane marking and takes value 0 or 1, respectively. The Dice coefficient value falls between 0 to 1, where the latter signifies a perfect overlap. The final goal is to predict lane markings as close to ground truth as possible, and the chosen loss function serves as an accurate proxy since minimizing it means maximizing the Dice coefficient or degree of overlap. Performance metrics such as precision, recall, and F1-score – represented by Equations (2 - 4) where TP, FP, and FN are the true positives, false positives, and false negatives, respectively – are reported for all strategies.

$$Dice \quad coefficient = \frac{2\Sigma_{pixel}\, y_{true}\, y_{pred}}{\Sigma_{pixel}\, y_{true} + \Sigma_{pixel}\, y_{pred}} \qquad (1)$$

$$Precision = \frac{TP}{TP + FP} \qquad (2)$$

$$Recall = \frac{TP}{TP + FN} \qquad (3)$$

$$F_1 \quad score = 2 \times \frac{Precision \times Recall}{Precision + Recal} \qquad (4)$$
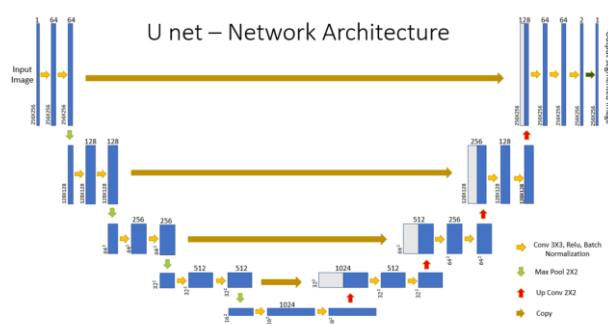

Figure 7. U-net architecture (adapted from Ronneberger et al. [20])

5. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we first discuss the LUTs generated from normalized intensity thresholding for each dataset and show how this strategy successfully reduces false positives in concrete pavement regions without affecting performance in asphalt areas. Then, performance metrics of each of the four strategies in both asphalt and concrete pavements are presented

and accounted for in general and in the context of pavement surface.

5.1 Effect of Pavement Surface and Intensity Normalization

At the intensity normalization stage, three small road surface point clouds located in ROIs 1, 2, and 3 (concrete pavement area) were selected for each dataset, as shown in Figure 3. The detailed information, including the length of ROIs, number of the sensors, driving speeds, and cell sizes, for generating LUTs are listed in Table 3. As mentioned previously, the total numbers of LUTs generated are 3, 4, and 4 for datasets 1, 2, and 3, respectively. From Table 3, relatively large cell sizes were determined for ROI 3 because of its faster driving speed. Also, the cell sizes of VLP16 are slightly larger due to the less points acquired by fewer laser beams for the same ROI. The resultant LUTs for one of the HDL32E LiDAR scanners in ROIs 1, 2, and 3 are shown in Figure 8. It is apparent from Figure 8 that normalized intensity values in ROI 3 are greater than ROIs 1 and 2. This is because datasets 1 and 2 belonged to the same interstate highway, while dataset 3 was acquired on a rural highway. Compared to a rural highway, wear and tear of pavement material in interstate highway is more gradual [22]. Thus, normalized intensity values are significantly impacted by the nature of pavements, which in turn affects the obtained LUT. After all the LUTs were generated for each LiDAR scanner in each dataset, the original intensity values of dataset 1, 2, and 3 were normalized. Figure 9 illustrates sample hypothesized lane markings derived through the original and normalized thresholding in dataset 3. As can be clearly observed, false positives are reduced in concrete pavement area after normalization. In addition, using the LUTs generated from concrete regions does not negatively impact the normalized intensity values in asphalt pavement regions.
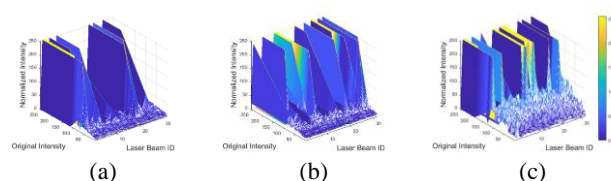

(a)　　　　　　(b)　　　　　　(c)

Figure 8. Lookup tables (LUT) generated for intensity normalization of an HDL 32E LiDAR unit in (a) ROI 1, (b) ROI 2, and (c) ROI 3.
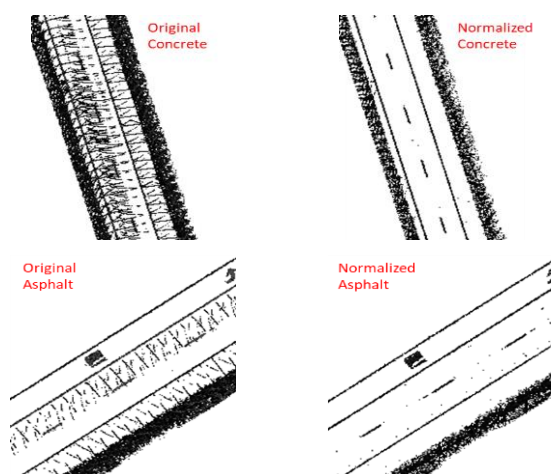

Figure 9. Hypothesized lane markings derived through original and normalized intensity thresholding strategies in concrete and asphalt pavement area for dataset 3

ROI	1	2	3
# of sensors	3	4	4
Avg. speed (mph)	49	48	65
Length of ROI (m)	155	190	155
Cell Size of HDL32E sensors (m)	0.12	0.12	0.25
Cell Size of VLP16 sensor (m)	0.20	0.20	0.30

Table 3. Length of ROIs, number of the sensors, driving speeds, and cell sizes for generating LUTs of HDL32E and VLP16 LiDAR units in three datasets

5.2 Performance of Lane Marking Extraction Strategies

U-net model 1 is trained on 400 manually labeled intensity images and validated on 104 such images. On the other hand, for U-net model 2, a total of 1,183 automatically labeled intensity images are used as a training dataset, while another 238 automatically labeled images are used for validation. Datasets 1 and 3 are utilized to generate these images for each U-net model. In addition, the same hyperparameters were used for both of them. The learning rate, batch size, and epochs are 8×10^{-4}, 8, and 100, respectively. The learning rate is also diminished by a factor of 10 when there is no improvement in the validation loss from the current lowest value for 5 consecutive epochs. Additionally, early stopping is also implemented where the training is halted when the validation loss shows no improvement from the current lowest value for 15 consecutive epochs. The weights of the two models were updated by the Adam optimizer. Figure 10 shows the training loss (calculated on training data) and validation loss (calculated on validation data) plots for U-net models 1 and 2. These plots are produced by calculating training and validation loss at the end of each epoch. The model tries to learn the mapping from input to output based on the seen training data and performance on validation data indicates if the learning is useful or not when an unseen data is fed as input. U-net model 2 obtains the lowest validation loss of 0.12, while model 1 can only achieve a value of 0.17 for the same. U-net model 2 is trained on a larger training dataset, which helps it attain a lower validation loss than U-net model 1. For evaluating the performance of each strategy, a test dataset of 174 images is generated from dataset 2 by manual labeling, with 92 images in asphalt pavement areas and 82 in concrete pavement areas. For intensity thresholding approaches, the derived lane markings (point cloud) are rasterized into intensity images for subsequent performance evaluation. Table 4 and 5 show the performance metrics for the state-of-the-art strategies (original intensity thresholding and deep learning with manual labeling) and proposed approaches (normalized intensity thresholding and deep learning with automated labeling) in asphalt and concrete pavement areas.

Lane Marking Extraction Strategies	Precision	Recall	F1-score
Original intensity thresholding	83.8 %	68.3 %	72.9 %
Normalized intensity thresholding	81.1 %	78.5 %	77.9 %
Deep learning with manual labelling (U-net model 1)	57.5 %	99.3 %	72.0 %
Deep learning with automated labelling (U-net model 2)	80.3 %	92.0 %	84.9 %

Table 4. Performance metrics in asphalt pavements for lane marking extraction strategies

Lane Marking Extraction Strategies	Precision	Recall	F1-score
Original intensity thresholding	84.4 %	57.9 %	65.9 %
Normalized intensity thresholding	87.1 %	69.7 %	76.0 %
Deep learning with manual labelling (U-net model 1)	63.8 %	98.5 %	77.1 %
Deep learning with automated labelling (U-net model 2)	88.1 %	83.3 %	85.1 %

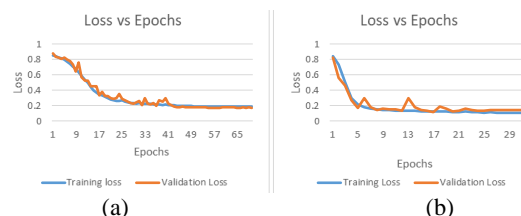Table 5. Performance metrics in concrete pavements for lane marking extraction strategies

Figure 10. Training and validation loss curves for (a) U-net model 1 and (b) U-net model 2

Based on performance metrics in Table 4 and 6, the results can be discussed in general and in the context of pavement surface:

1. **General trend:** In both pavement areas, U-net models obtain higher recall in comparison to intensity thresholding approaches. This means that the former can extract true lane markings better than the latter. Figure 11 displays an intensity image with low edge lane marking point density along with corresponding lane markings extracted by the different strategies. The deep learning approaches can extract lane markings with low point density, but intensity thresholding approaches miss them. Additionally, it is also observed that U-net model 1 shows a very poor precision metric. This means lots of positive predictions are false since it cannot distinguish well between actual lane markings and high-intensity non-lane marking points as illustrated in Figure 12. Overall, U-net model 2, which is trained on almost 3 times more training data than model 1, has the best performance in both pavement regions as evident by its high F1-score.

2. **Asphalt pavements:** Higher precision is obtained by the intensity thresholding approaches which means that less false positives are observed. This is because the outliers are eliminated to a great extent by the noise removal strategy. In addition, the normalized intensity thresholding strategy performs better than U-net model 1 based on the F1-scores. Lastly, the performance of the original intensity thresholding strategy is comparable to the other three strategies, which shows that its performance is not significantly impacted in asphalt pavements as shown in Figure 13.

3. **Concrete pavements:** In concrete pavement area, the original intensity thresholding strategy shows a substandard performance with an F1-score of only 65.9 %. Figure 14 compares lane marking results in a concrete region where all the strategies except the original intensity thresholding perform well. The normalized intensity thresholding strategy successfully suppresses high-intensity non-lane marking points. Thus, thresholding the normalized intensity values leads to less false positives and hence better lane marking extraction. On the other hand, the deep learning

models successfully learn to eliminate these high-intensity non-lane marking points lacking any spatial structure.
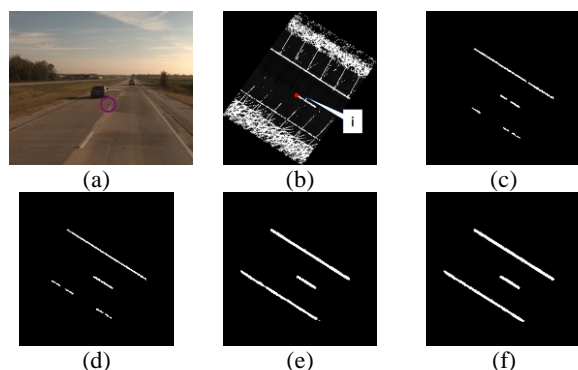

Figure 11. Worn-out yellow edge line area: (a) RGB imagery of location i, (b) original intensity image, predicted lane marking images of (c) original intensity thresholding, (d) normalized intensity thresholding, (e) U-net model 1, and (f) U-net model 2
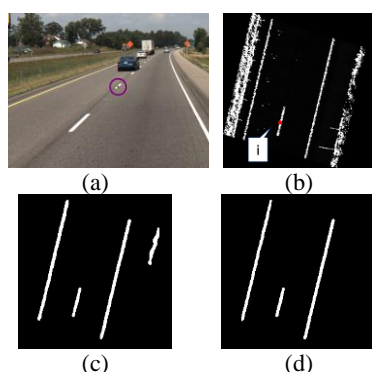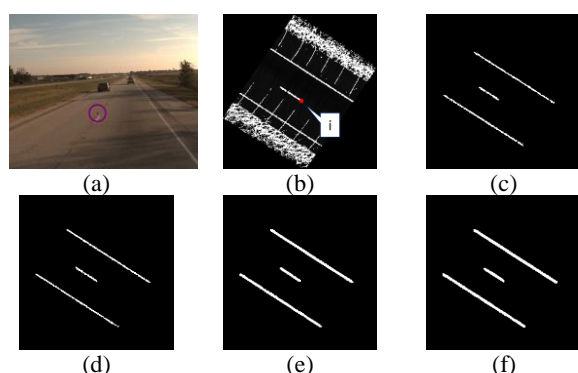

Figure 12. False positive problem: (a) RGB imagery of location i (b) original intensity images, and images with predicted lane markings from (c) U-net model 1 and (d) U-net model 2 in dataset 2

Figure 13. Asphalt pavement area: (a) RGB imagery of location i, (b) original intensity image and predicted lane marking images from (c) original intensity thresholding, (d) normalized intensity thresholding, (e) U-net model 1, and (f) U-net model 2.
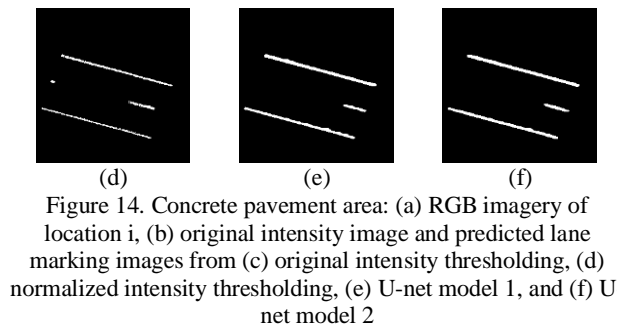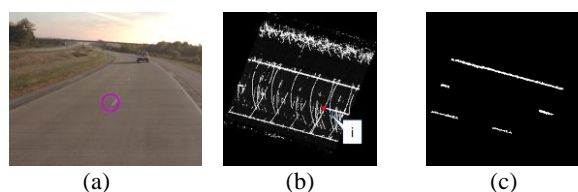

Figure 14. Concrete pavement area: (a) RGB imagery of location i, (b) original intensity image and predicted lane marking images from (c) original intensity thresholding, (d) normalized intensity thresholding, (e) U-net model 1, and (f) U-net model 2

In summary, it is observed that the lane marking extraction performance of deep learning approaches is superior to that of intensity thresholding approaches since the latter relies on content (intensity and point density), while the former's detections are based on both content as well as context (location of points). In asphalt pavement area, all the strategies obtain satisfactory results; however, in concrete pavement area, original intensity thresholding does not perform well and one must either normalize intensity to deal with lack of contrast between lane marking and pavement surface or train deep learning models that can learn complex mapping from input to output. Thus, the proposed normalized intensity thresholding and deep learning approaches are less sensitive to the nature of pavement surface. Finally, the U-net model trained on automatically generated labels outperforms the one trained on manually established labels which shows the robustness of automated procedure of label generation.

6. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

Even though the pavement surface plays an important role in lane marking extraction, there have not been any studies that evaluate the lane marking extraction in this context. In addition, the existing intensity thresholding strategies for lane marking extraction require prior knowledge for adaptive thresholding and noise removal. The learning-based strategies can get around those issues; however, curation of manually labeled data is a time-consuming step. Therefore, this paper seeks to tackle the above issues by proposing normalized intensity thresholding and automated label generation procedure-based deep learning strategy for lane marking extraction. The proposed strategies are compared with state-of-the-art strategies of original intensity thresholding and deep learning based on manually labeled training data. All of them are evaluated in asphalt and concrete regions separately.

The datasets processed covered both asphalt and concrete pavements. The proposed intensity normalization strategy significantly reduces false positives in concrete pavement area as opposed to thresholding of original intensity values. In addition, the U-net model 2 which is trained on automatically generated labels outperforms all other strategies in both asphalt and concrete areas. It had precision, recall, and F1-score of 80.3%, 92%, and 84.9% respectively in asphalt pavements while the same metrics in concrete pavements were 88.1%, 83.3%, and 85.1%. On the other hand, the intensity normalization strategy showed better performance than U-net model 1 in asphalt pavements and vice-versa in concrete pavements. We also observed that the original intensity thresholding strategy could provide reasonable lane marking extraction results in asphalt pavement area where the lane markings and pavement surface exhibit a significant intensity contrast. However, due to poor intensity contrast in concrete pavements, its performance suffers to a great extent, as shown by a poor F1-score of 65.9%. In such cases, the proposed

intensity normalization and deep learning approaches obtain better results indicating their robustness to varying intensity distribution due to the nature of pavement material. Lastly, we conclude that the deep learning approaches can detect lane markings with low point density unlike intensity thresholding approaches because of their ability to capture the content as well the context in the intensity image.

One of the directions for future work will explore the application of the current normalization algorithm to single-beam LiDAR scanners. As far as a deep learning approach is considered, transfer learning is another exciting direction for moving forward. The idea is to fine-tune the trained U-net model 2 on a training dataset that is captured by different LiDAR scanners or is captured in areas with different lane marking patterns than the one encountered in this research. This will save significant training time as the network is not trained from scratch. Preliminary results indicate that the fine-tuning encoder path of U-net leads to better performance than decoder. Additionally, one can also exploit RGB information along with the point cloud for improved lane marking extraction, particularly in areas where the markings are of poor quality.

7. ACKNOWLEDGMENTS

8. REFERENCES

1. Carnaby, B. *Poor road markings contribute to crash rates.* in *Australasian Road Safety Research Policing Education Conference, 2005, Wellington, New Zealand.* 2005.

2. Yu, Y., et al., *Learning hierarchical features for automated extraction of road markings from 3-D mobile LiDAR point clouds.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015. **8**(2): p. 709-726.

3. Yan, L., et al., *Scan line based road marking extraction from mobile LiDAR point clouds.* Sensors, 2016. **16**(6): p. 903.

4. Yang, M., et al., *Laser data based automatic recognition and maintenance of road markings from MLS system.* Optics & Laser Technology, 2018. **107**: p. 192-203.

5. Guan, H., et al., *Using mobile laser scanning data for automated extraction of road markings.* ISPRS Journal of Photogrammetry and Remote Sensing, 2014. **87**: p. 93-107.

6. Kumar, P., et al., *Automated road markings extraction from mobile laser scanning data.* International Journal of Applied Earth Observation and Geoinformation, 2014. **32**: p. 125-137.

7. Soilán, M., et al., *Segmentation and classification of road markings using MLS data.* ISPRS Journal of Photogrammetry and Remote Sensing, 2017. **123**: p. 94-103.

8. YAO, L., et al., *Automatic Extraction of Road Markings from Mobile Laser-Point Cloud Using Intensity Data.* International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2018. **42**: p. 3.

9. Bradley, D. and G. Roth, *Adaptive thresholding using the integral image.* Journal of graphics tools, 2007. **12**(2): p. 13-21.

10. He, B., et al. *Lane marking detection based on convolution neural network from point clouds.* in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC).* 2016. IEEE.

11. Wen, C., et al., *A deep learning framework for road marking extraction, classification and completion from mobile laser scanning point clouds.* ISPRS journal of photogrammetry and remote sensing, 2019. **147**: p. 178-192.

12. Adrian, W. and R. Jobanputra, *Influence of pavement reflectance on lighting for parking lots.* 2005, Portland Cement Association Skokie, IL.

13. Puttonen, E., et al., *Measurement of reflectance properties of asphalt surfaces and their usability as reference targets for aerial photos.* IEEE Transactions on Geoscience and Remote Sensing, 2009. **47**(7): p. 2330-2339.

14. Ravi, R., et al., *Bias impact analysis and calibration of terrestrial mobile lidar system with several spinning multibeam laser scanners.* IEEE Transactions on Geoscience and Remote Sensing, 2018. **56**(9): p. 5261-5275.

15. Ravi, R., et al., *Simultaneous system calibration of a multi-lidar multicamera mobile mapping platform.* IEEE Journal of selected topics in applied earth observations and remote sensing, 2018. **11**(5): p. 1694-1714.

16. Lari, Z. and A. Habib, *New approaches for estimating the local point density and its impact on LiDAR data segmentation.* Photogrammetric Engineering & Remote Sensing, 2013. **79**(2): p. 195-207.

17. Ravi, R., et al., *Lane Width Estimation in Work Zones Using LiDAR-Based Mobile Mapping Systems.* IEEE Transactions on Intelligent Transportation Systems, 2019.

18. Levinson, J. and S. Thrun. *Robust vehicle localization in urban environments using probabilistic maps.* in *2010 IEEE International Conference on Robotics and Automation.* 2010. IEEE.

19. Levinson, J. and S. Thrun. *Unsupervised calibration for multi-beam lasers.* in *Experimental Robotics.* 2014. Springer.

20. Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation.* in *International Conference on Medical image computing and computer-assisted intervention.* 2015. Springer.

21. Dice, L.R., *Measures of the amount of ecologic association between species.* Ecology, 1945. **26**(3): p. 297-302.

22. USGS. *Materials in Use in U.S. Interstate Highway.* 2016; Available from: https://pubs.usgs.gov/fs/2006/3127/2006-3127.pdf.