Image Dataset for Persian Road Surface Markings

Seyed Hamid Safavi, Mohammad Eslami, Aliasghar Sharifi Najafabadi, Amirhosein Hajihoseini,
Mohammadreza Riahi, Maryam Rekabi, Sadaf Sarafan, Rahman Zarnoosheh,
Ehsan Khodapanah Aghdam, Sahar Barzegari Banadkoki, Seyed Mohammad Seyedin Navadeh,
Farah Torkamani-Azar
Digital Signal Processing LaboratorY (DiSPLaY),
Department of Electrical Engineering,
Shahid Beheshti University, Tehran, Iran
Email: {h_safavi, m_eslami, f-torkamani}@sbu.ac.ir

***Abstract*—Self-driving and autonomous cars are hot emerging technologies which can provide enormous impact in the near future. Since an important component of autonomous cars is vision processing, the increasing interest for self-driving cars has motivated researchers to collect different relative image datasets. Hence, we collect a comprehensive dataset about the road surface markings which are available in Iran. In addition, we evaluate the conventional recognition rate. In this paper, we present a novel and extensive dataset for Persian Road Surface Markings (PRSM) with ground truth labels. We also hope that it will be useful as a Persian benchmark dataset for researchers in this field. The dataset consists of over 68,000 labeled images of road markings in 18 popular classes. It also contains road surface markings under various daylight conditions. Our dataset with further details is available online at: http://display.sbu.ac.ir/databases.**

Keywords—Autonomous Vehicle, Dataset, Road Surface Markings, Persian.

I. Introduction

Over the last two decades, autonomous navigation and road intelligence are shown to be a challenging problem for car industries. In addition, understanding the surrounding environment is important to autonomous vehicles. One of the popular research topics in the context of Autonomous Driver Assistance Systems (ADAS) [1] is Road Surface Marking Recognition (RSMR). It is evident that road surface marking has been the focus of attention due to its capability in managing and controlling traffic activities, guiding traffic routes and detecting potential road safety hazards.

Road markings refer to the symbols or texts which are painted on the road surface with the aim of traffic guidance for vehicles and pedestrians. Common road markings include lane indication arrows, crosswalk, caution, words, speed limits and etc. These markings are as important as traffic signs at the side or on top of the roads, as they enable better understanding for autonomous vehicles about their surrounding environments. Moreover, to increase the safety of the vehicle and the passengers, these markings should be recognized in different challenging situations.

Iranian road markings include some texts in the Persian language, and there is a need to collect a new benchmark dataset for it. Therefore, we propose the first Persian Road Surface Markings (PRSM) dataset which contains occluded, depreciated and high motion blurred signs while the quality of the signs are also degraded significantly. In this paper, we introduce a new large dataset with ground truth labels which is captured using a GoPro camera attached to the hood or mounted on a roof rack of a vehicle. It also contains road surface markings under various daylight conditions such as sunny, sunset and night time. Furthermore, this dataset contains images of marking signs in three different qualities, excellent, fair and poor. The poor images have serious occlusion, motion blur or are depreciated. To the best of our knowledge, there is not any dataset which includes this volume of markings with different qualities, since collecting of this kind of datasets needs more time and work.

We expect that the proposed PRSM dataset will be a benchmark dataset for future works in this area. The dataset consists of over 68 thousand labeled images of road markings and includes 18 popular classes. The possible applications of this dataset are numerous such as recognition, detection, and classification of road markings which can be used in ADAS and autonomous vehicles. Our ultimate aim is to explore the challenges of the different applications of the Persian road surface marking. Here, as a first step, we just investigate the performance of the classification using well-known feature extractors and popular classifiers on collected PRSM dataset. The PRSM dataset and source codes are available online [1].

The rest of this paper is structured as follow. Section II introduces the related datasets. Section III provides the details of our PRSM dataset and collection procedure. Section IV presents the classification strategy and section V reports the performance results. Finally, section VI concludes the paper.

II. Ralated Datasets

To the best of our knowledge, there are a few datasets of road markings which are proposed for different purposes like recognition, detection, etc. However, there isn't any Persian road markings dataset. In the following, we briefly review the main international road marking datasets.

ROMA (ROad MArkings) image database [2] was collected in 2008. It comprises more than 100 original images of various road scenes. Moreover, the authors in [3] gathered a new dataset for road marking detection and classification. It consists of over 1400 labeled images of road markings with bounding boxes showing the location of the markings.

[1] http://display.sbu.ac.ir/databases

Fig. 1: Example images from various related datasets: a) Road Marking [3], b) KITTI Raw [4], c) Malaga 2009 [5], d) Malaga urban [6], e) KITTI road/lane [7], f) ROMA [2], g) Reading the Road [8].

The KITTI Vision Benchmark Suite is the Karlsruhe institute of technology and Toyota technological institute (KITTI) dataset and [4], [7] are comprehensive datasets. Also, they provide a benchmark for various autonomous vehicle applications. More clear, KITTI suite includes images and other information for different tasks such as stereo, optical flow, visual odometry, 3D object detection and 3D tracking. The road and lane estimation benchmark [7] consists of 289 training and 290 test

Fig. 2: GoPro camera is used in our experiment and attached to the hood of the car.

images. In [4], the dataset comprises the following information, captured and synchronized at 10 Hz: Raw color and gray images, 3D pointclouds, 3D GPS and IMU data, different calibrations and 3D object tracklet labels (cars, trucks, trams, pedestrians, cyclists). This dataset is categorized in 6 different scenes as follow. City, Residential, Road, Campus, Person and Calibration.

A Collection of Outdoor Robotic Datasets with centimeter-accuracy Ground Truth is produced in Malaga 2009 dataset [5]. That creates a 6D comprehensive benchmark for robotic vision and SLAM (Simoultaneous Localization & Mapping) operations as similar as KITTI. Malaga urban dataset [6] was gathered entirely in urban scenarios with a car equipped with several sensors, including one stereo camera (Bumblebee2) and five laser scanners.

Furthermore, the authors in [8] created a benchmark ground truth class annotated dataset containing 2068 images spanning city, residential and motorway roads and over 13099 unique annotations. Figure 1 shows example images of the captured ones in each dataset.

III. Dataset Generation

We collected our data using GoPro Hero camera that is mounted on a roof rack or attached to the hood of the car and definitely facing forwards. Fig. 2 shows the used camera and car in our experiments. Both GoPro Hero3 and Hero4 were employed and videos are recorded in 25 frame per second, resolution 1080p with wide setting of lens. The speed of the vehicle varied between 30 to 70 km/h. The urban roads of Tehran, Iran was considered under various lighting, and road conditions, including challenging situations such as sunny, sunset, night time and occluded markings. It should be noted that we have gathered a lot of challenging road surface markings with poor conditions that have shadowed or faded.

We extracted over 100 thousand frames from the captured videos. Then, since the number of frames in our dataset was huge and manually drawing bounding boxes is a tedious task, we only draw bounding boxes for one in every 4 frames and the rest of them were interpolated. When all of the frames are ready, we labled their signs using a MATLAB-based Graphic User Interface (GUI). Fig. 3 represents the schematic of this GUI. Furthermore, along with assigning the labels for each

TABLE I: Class distribution and showing the number and proportion of samples in each captured class.

Class name	Number of training samples	Number of testing samples	Total number of samples in each class	Proportion
Caution Text	3441	1474	4915	9.80 %
Caution Symbol	1277	547	1824	3.64 %
Yield line or Shark's teeth	2463	1056	3519	7.01 %
Crosswalk	19525	8368	27893	19.93 %
Crosswalk Caution Text	163	70	233	0.46 %
Crosswalk Caution Symbol	826	354	1180	2.35 %
Forward	4023	1724	5747	11.45 %
Forward and Turn Left	674	289	963	1.92 %
Forward and Turn Right	1462	627	2089	4.16 %
School	760	325	1085	2.16 %
Slow	2737	1173	3910	7.79 %
Speed Bump	3758	1610	5368	10.70 %
Speed Limit	135	58	193	0.38 %
Stop	1241	532	1773	3.53 %
Stop Line	3931	1684	5615	11.19 %
Strain Speed	657	281	938	1.87 %
Turn Left	142	61	203	0.40 %
Turn Right	438	187	625	1.25 %
Total Number of Samples	47653	20420	68073	100 %

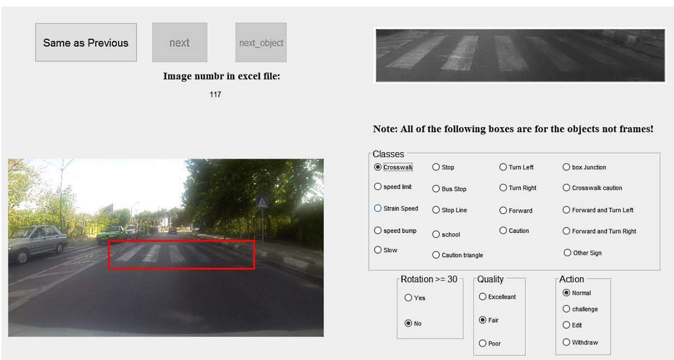


Fig. 3: Designed MATLAB-based GUI for labeling the frames.

sign encompassed in bounding box, its quality and rotation are dedicated.

Fig. 4 illustrates some examples of the captured road marking classes and Table I indicates the class distribution of our dataset and shows the number and proportion of samples in each class. As it can be seen, we considered over 68 thousand labeled images of road markings which consist of 18 popular classes with the option of labeling different qualities such as excellent, fair and poor. Moreover, we considered the rotation above 30 degree of each road surface marking in the GUI. Consequently, our dataset include an excel file with the above-mentioned labels. Finally, we hope that this huge dataset will be a useful Persian benchmark for the road surface marking with the aim of helping researchers in this field.

IV. Recognition Investigation Strategy

The recognition strategy for evaluating this dataset and overcoming the challenges is described in this section as follows.

A. Preprocessing

To increase the contrast of the input images, we use histogram equalization as a preprocessing step. Using this method, the recognition of the road markings can be facilitated. Even though the gray scale invariance can be achieved using histogram equalization, but it cannot cover local variations. On the other hand, road surface marking and traffic sign detection methods usually use inverse perspective mapping to create a perspective free birds-eye-view image. It is evident that IPM would be useful for road markings detection. However, in our investigation for recognition purpose, this preprocessing is not applied. In other words, the features are extracted directly from original cropped signs by the popular feature extraction methods. We are trying to recognize the signs with lower computational cost.

B. Feature Extraction

Any road markings dataset include markings with different scales, rotations and illumination variations. Hence, a good feature extraction method is needed to be invarient with respect to these conditions. Therefore, we have tested three popular feature extraction methods such as Local Binary Pattern (LBP) [9], Histogram of Oriented Gradient (HOG) [10], and Patterns of Oriented Edge Magnitudes (POEM) [11].

It is known that LBP feature is originally proposed for texture description, but, it is also widely used in other applications. It is gray scale invariant and locally rotation invariant. Moreover, LBP and also it's variations are sensitive to the lighting conditions since they have only illumination information. They also achieve the gray scale invariance by thresholding the neighborhood of each pixel with the center pixel value and considering the result as a binary number. Recently, POEM descriptor is proposed by applying self-similarity based structure on oriented magnitudes. That is calculated by accumulating a local histogram of gradient orientations (as like as applying LBP on gradient information) over all pixels of image cells, centered on the considered pixel. On the other hand, based on the literature [3], [12], widely used HOG method for feature extraction of symbolic road markings encourages us to apply it in our work. The basic idea of HOG is that the distribution of the local intensity gradients or edge directions may be sufficient for discrimination [10]. Consequently, based on our observation, HOG is one of the best approaches for extracting the features of road markings.

(a) Forward

(b) Turn Left

(c) Turn Right

(d) Forward and Turn Left

(e) Forward and Turn Right

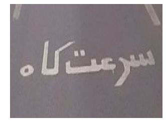
(f) Caution Symbol

(g) Caution Text

(h) Slow

(i) Stop

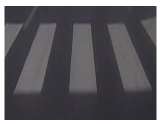
(j) Crosswalk caution text



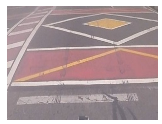
(k) Strain speed

(l) School

(m) Yield line or Shark's teeth



(n) Crosswalk



(o) Speed bump

(p) Speed limit

(q) Stop line

(r) Crosswalk caution symbol

Fig. 4: Classes of captured road surface marking.

C. *Classifier*

After extracting suitable features for recognition of the road surface markings, they are fed into the classifier. Here, we have started recognition of the road markings with famous classifiers that widely used in the computer vision community. Therefore, the performance of two popular approaches such as Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) are investigated in this paper.

V. Experiments

In this section, our aim is to evaluate the recognition performance of the different approaches on the proposed PRSM dataset. Note that since there is not any Persian road surface marking dataset, it is not possible to compare our founding with other works. We hope that this domestic dataset would be an opening of the research on the self-driving cars in Iran.

As it is already mentioned, the PRSM dataset includes the challenging conditions such as various daylight condition, occluded markings, and also poor quality markings that recognizing them are really hard. Therefore, to see how effective are the different approaches, we present the following experiments. The first step for evaluating the dataset is partitioning of the markings to the train and test categories. Hence, for each class, the number of markings in training set is chosen to be 70 percent of the total number of markings and the remaining 30 percent are chosen for the test set. Now, the question is that how we could choose these categories? Should the training set include just the markings with excellent quality or it could have both excellent and poor quality? To answer these questions and investigate the effect of poor quality markings in training set, we design two experiments: Scenario A: for each class, we choose the markings with better qualities as a training set (i.e. we exclude the poor markings) and as a result, the remaining markings which include poor qualities are chosen for the test set. Scenario B: the markings are chosen randomly for the training and test set. The shared dataset on the website obeys this two kind of separation. It should be also mentioned that since the crosswalk marking is mostly used in urban areas, the number of items in this class is higher than the other classes. So, in this paper, we just use only 10 thousand of this class in our experiments since this number is sufficient for the recognition purpose. But, the same numbers in Table I are used for the other classes.

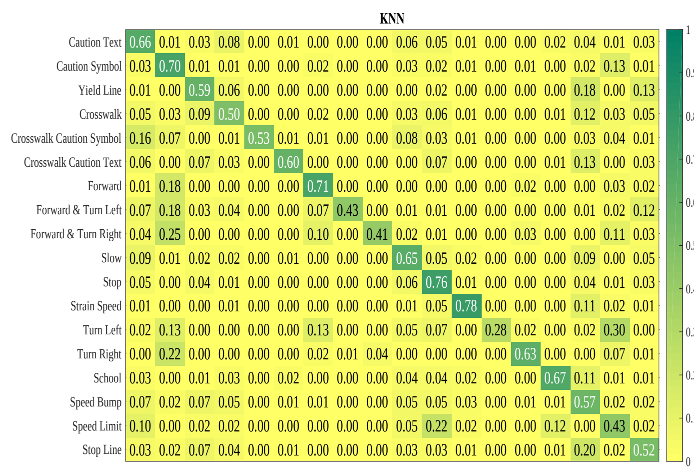
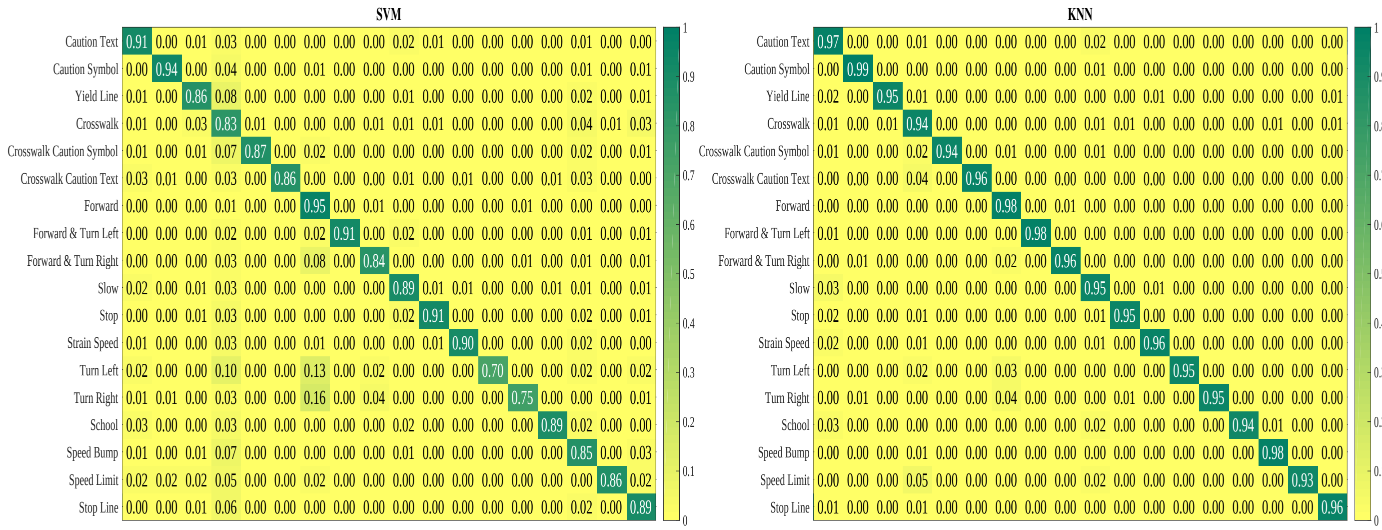


Fig. 5: Confusion matrix of the KNN classifier along with HOG feature extraction for the scenario A.

Since some of the markings are vertical, horizontal rectangular and some of them are close to square shape, our next



(a) SVM

(b) KNN

Fig. 6: Confusion matrix of classifiers along with LBP feature extraction for the scenario B: (a) SVM, (b) KNN

TABLE II: Accuracy Comparison of different Classifiers for the scenario B.

	SVM	KNN	Length of feature vector
LBP	0.87	0.97	6400
HOG	0.85	0.98	9216
POEM	0.89	0.93	8496

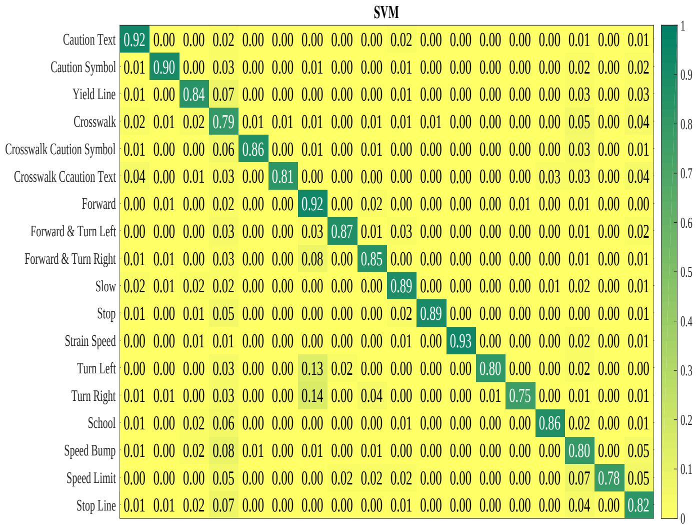
TABLE III: Precision and Recall Comparison of different Classifiers along with HOG feature extraction for the scenario B.

Class name	Precision		Recall	
	SVM	KNN	SVM	KNN
Caution Text	0.85	0.95	0.92	0.99
Caution Symbol	0.94	0.96	0.90	0.99
Yield line or Shark's teeth	0.86	0.98	0.84	0.99
Crosswalk	0.53	0.95	0.79	0.94
Crosswalk Caution Text	0.97	1	0.81	0.97
Crosswalk Caution Symbol	0.97	1	0.86	0.97
Forward	0.68	0.99	0.92	0.98
Forward and Turn Left	0.94	1	0.87	0.99
Forward and Turn Right	0.88	0.97	0.85	0.99
School	0.94	1	0.86	0.98
Slow	0.86	0.98	0.89	0.99
Speed Bump	0.67	0.97	0.80	0.99
Speed Limit	0.99	0.99	0.78	0.95
Stop	0.96	0.99	0.89	0.98
Stop Line	0.72	0.96	0.82	0.98
Strain Speed	0.97	0.97	0.93	0.99
Turn Left	0.98	0.99	0.80	1
Turn Right	0.98	1	0.75	0.98

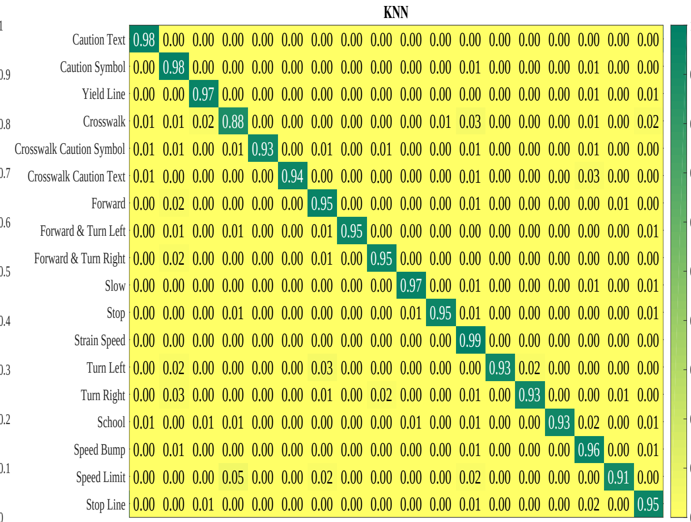
challenge in the experiments was choosing the markings size. Based on our investigations on different size of markings, since we use feature descriptors and the pixel values are not used directly, each rectangle of cropped signs is re-sized to 150×150 pixels. Following settings are also applied in our experiments:

- Local Binary Pattern (LBP) method; 30×30 non-overlap patches (25 patches), 8 sampling points on a circle of radius 4 and using no mapping table. The length of the feature vector is $25 \times 2^8 = 6400$.
- Histogram of Gradients (HOG) method; The number of cells in each block is $4 \times 4 = 16$, and the size of each Cell is 8×8. Also, the number of bins is assumed to be 9. Hence, the length of the feature vector is $16 \times 64 \times 9 = 9216$.
- Pattern of Oriented Magnitudes (POEM) method; The number of orientations is 4. The size of each cell around each pixel is 9×9. The diameter of the block for calculation of binary code in LBP is 14. The number of neighbors for calculation of binary code in LBP is 8 and 6×6 non-overlap patches (36 patches) for computing histograms. The length of the feature vector is 8496.
- For KNN algorithm, Euclidean distance metric is used with $k = 5$.

For evaluation of the first scenario in our experiments, we have plotted the confusion matrix for every feature extraction method along with the SVM and KNN classifiers. But, due to the page limit we just show the results of the HOG feature extraction and KNN classifier in Fig. 5. It can be observed that if the train set does not include the poor quality markings, then the recognition performance will be degraded. In addition, overall accuracy for this scenario is 0.58. It sounds reasonable since all of the poor quality markings are excluded from the training set and the recognition method did not have any sense from the types of markings occlusion, motion blur, and depreciation. Therefore, we believe that this type of separation has more challenges to overcome. In order to enhance the



(a) SVM



(b) KNN

Fig. 7: Confusion matrix of classifiers along with HOG feature extraction for the scenario B: (a) SVM, (b) KNN

performance, one approach could be including the poor quality markings in the training set. By this approach, the recognition method learns the type of depreciation and consequently, the performance will be improved. We consider this approach in our second scenario of experiments.

In the second scenario of our experiments, we evaluate the performance of the random separation of train and test set. Table II compares the accuracy of the different classifiers together with the feature extraction methods. It can be seen that KNN approach with HOG feature extraction method achieves the highest accuracy. Hence, the precision and recall parameters are compared in Table III for just the HOG feature descriptor. Furthermore, Fig. 6, 7, and 8 illustrate the confusion matrix for the SVM and KNN approaches along with LBP, HOG and POEM feature extractors, respectively. It can be observed that KNN algorithm can be recognized road surface markings with higher recognition rate compared to SVM. Looking at the fourth column of the SVM confusion matrices (i.e. the “Crosswalk” road marking) reveals that most of the classes are confused with this road marking. One reason for this observation could be the existence of large number of marking in this class (about 20 percent of the whole dataset). Obviously, in most of the cities, there exist larger number of the crosswalk marking compared to the others. On the other hand, note that in our PRSM dataset, the number of road markings in some of the classes are low such as “Turn Left”, “Crosswalk Caution Text” and “Speed Limit”. It cal also affect the performance.

Moreover, to investigate the cases that the recognition of them is failed in the second scenario, Fig. 9 shows some of these cases that are not correctly recognized. As it can be seen, despite the good performance of KNN algorithm along with the mentioned features, there are still some shadowed or faded cases that didn’t recognized.

VI. Conclusion and Future Work

In this paper, we present a new large dataset for Persian Road Surface Marking (PRSM) with ground truth labels. The dataset includes over 68 thousand labeled images of road markings with bounding boxes showing the location of the marking. Moreover, it includes 18 popular classes which are used in Iran. It also contains road surface markings under various daylight conditions. Furthermore, we investigated the recognition challenges of the proposed PRSM dataset. The popular classifiers for recognition such as SVM and KNN are applied and their performance has compared. Experimental results showed that KNN algorithm has higher recognition accuracy compared to the SVM approach for our PRSM datset. Further work is underway to add popular detection methods along with the recognition approaches used in this paper. In addition to various daylight condition such as sunny, sunset and night, we will also extend our dataset to include additional challenging conditions such as twilight, cloudy, rainy and snowy weather.

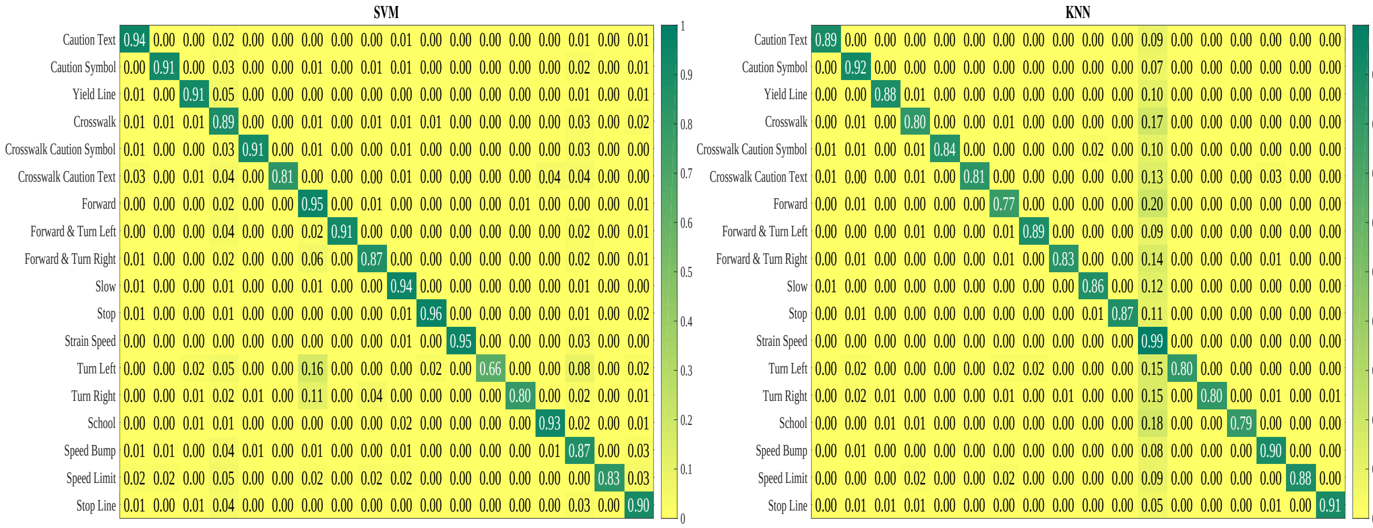
References

[1] R. Bishop, *Intelligent vehicle technology and trends*, 2005.

[2] T. Veit, J.-P. Tarel, P. Nicolle, and P. Charbonnier, “Evaluation of road marking feature extraction,” in *11th International IEEE Conference on Intelligent Transportation Systems (ITSC).*, 2008, pp. 174–181, available online: http://perso.lcpc.fr/tarel.jean-philippe/bdd/index.html.

[3] T. Wu and A. Ranganathan, “A practical system for road marking detection and recognition,” in *IEEE Intelligent Vehicles Symposium (IV)*, 2012, pp. 25–30, available online: http://www.ananth.in/RoadMarkingDetection.html.

[4] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, “Vision meets robotics: The kitti dataset,” *International Journal of Robotics Research (IJRR)*, 2013, available online: http://www.cvlibs.net/datasets/kitti/raw_data.php.



(a) SVM

(b) KNN

Fig. 8: Confusion matrix of classifiers along with POEM feature extraction for the scenario B: (a) SVM, (b) KNN

(a) (b) (c) (d) (e) (f) (g) (h) (i)

Fig. 9: Examples of incorrect recognition cases for the scenario B.

[5] J.-L. Blanco, F.-A. Moreno, and J. González, "A collection of outdoor robotic datasets with centimeter-accuracy ground truth," *Autonomous Robots*, vol. 27, no. 4, pp. 327–351, November 2009, available online: http://www.mrpt.org/robotics_datasets. [Online]. Available: http://www.mrpt.org/Paper:Malaga_Dataset_2009

[6] J.-L. Blanco, F.-A. Moreno, and J. Gonzalez-Jimenez, "The mlaga urban dataset: High-rate stereo and lidars in a realistic urban scenario," *International Journal of Robotics Research*, vol. 33, no. 2, pp. 207–214, 2014, available online: http://www.mrpt.org/MalagaUrbanDataset. [Online]. Available: http://www.mrpt.org/MalagaUrbanDataset

[7] J. Fritsch, T. Kuehnl, and A. Geiger, "A new performance measure and evaluation benchmark for road detection algorithms," in *International Conference on Intelligent Transportation Systems (ITSC)*, 2013, available online: http://www.cvlibs.net/datasets/kitti/eval_road.php.

[8] B. Mathibela, P. Newman, and I. Posner, "Reading the road: Road marking classification and interpretation," *IEEE Trans. Intelligent Transportation Systems*, vol. 16, no. 4, pp. 2072–2081, 2015.

[9] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[10] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, 2005, pp. 886–893.

[11] N.-S. Vu and A. Caplier, "Face recognition with patterns of oriented edge magnitudes," *Computer Vision–ECCV*, pp. 313–326, 2010.

[12] J. K. Suhr and H. G. Jung, "Fast symbolic road marking and stop-line detection for vehicle localization," in *Intelligent Vehicles Symposium (IV), 2015 IEEE*. IEEE, 2015, pp. 186–191.