Journal of Computing and Applied Informatics (JoCAI) Vol. 04, No. 1, 2020 | 16 – 26

talenta PUBLISHER

DATA SCIENCE
Journal of Computing and Applied Informatics

Classification for Driver's Distraction and Drowsiness Through Eye Closeness Using Receiver Operating Curve (ROC)

Muharram Anis Hazirah Rodzi**, Zalhan Mohd Zin, and Norazlin Ibrahim

Universiti Kuala Lumpur Malaysia France Institute, Selangor, Malaysia

Abstract. In Malaysia, driver inattention and drowsiness becomes one of the causes of road accidents which sometime could lead to fatal ones. From the data provided by Malaysian Police Force, Polis Di Raja Malaysia or PDRM in 2016, deaths from road accidents increased from 6,706 in 2015 to 7,512 in 2016. Some accidents were caused by human factor such as driver's inattention and drowsiness. This problem motivates many parties to look for better solution in dealing with this human factor. Some of the car manufacturers have introduced to their certain models of car with an assistant system to oversee driver's condition. The assistant system is in fact part of the main safety system known as Advanced Driver Assistance Systems (ADAS). The kind of system has been developed to strengthen vehicle systems for safety and conducive driving. The system has been contemplated to congregate accurate input, rapid processing data, precisely predict context, and respond in real time. In addition to that, suitable method is also needed to detect and classify driver drowsiness and inattention using computer vision as the latter become more and more important in any intelligent system development. In this paper, the proposed system introduces a method to classify drowsiness into three different classes of eye state; open, semi close and close. The classification has been done by using feature extraction method, percentage of eye closure (PERCLOS) technique and Support Vector Machine (SVM) classifier. The performances of the methods have been then measured and represented by using confusion matrix and ROC performance graph. The results have show that the proposed system has been able to achieve high performance of distraction and drowsiness detection according to driver's eye closeness level.

Keyword: Classification, Distraction, Drowsiness, Driver, Advanced Driver Assistance Systems, Receiver Operating Curve, PERCLOS, SVM.

Received 10 September 2019 | Revised 15 December 2019 | Accepted 29 January 2020

1. Introduction

Human factor is one of the main factors that contribute to the high rate of road accidents. Being human, feeling abnormally sleepy or fatigue during the day may lead to additional symptoms, such as forgetfulness or falling asleep at inadmissible times [1] and unreasonable situation [2]. Corresponding to recent provisional data provided by Polis Di Raja Malaysia (2016), deaths from

*Corresponding author at: Universiti Kuala Lumpur Malaysia France Institute, Selangor, Malaysia

E-mail address: anis.rodzi@s.unikl.edu.my

ISSN: 2580-6769 | e-ISSN: 2580-829X | DOI : 10.32734/jocai.v4.i1-3516
Journal Homepage: https://jocai.usu.ac.id

road accidents in 2015 is 6,706 compared to 2016 is 7,512, which is the highest number of deaths recorded over previous years [3]. Drowsiness might be caused due to fatigue, lack of sleep, medication consuming, or routine related with vehicle driving [4]. It leads to a vital recession in driver's abilities of perception and vehicle control. Therefore this will threaten safe driving and will increase the possibility of road accidents which sometime may be fatal.

2. Methodology

Figure 1 shows the models of this research fundamentally started with the acquisition of the input images from two conditions of the images which are the images with glasses and without glasses. The process continues to interest point detection, feature extraction, vector quantization and lastly histogram of features frequencies.

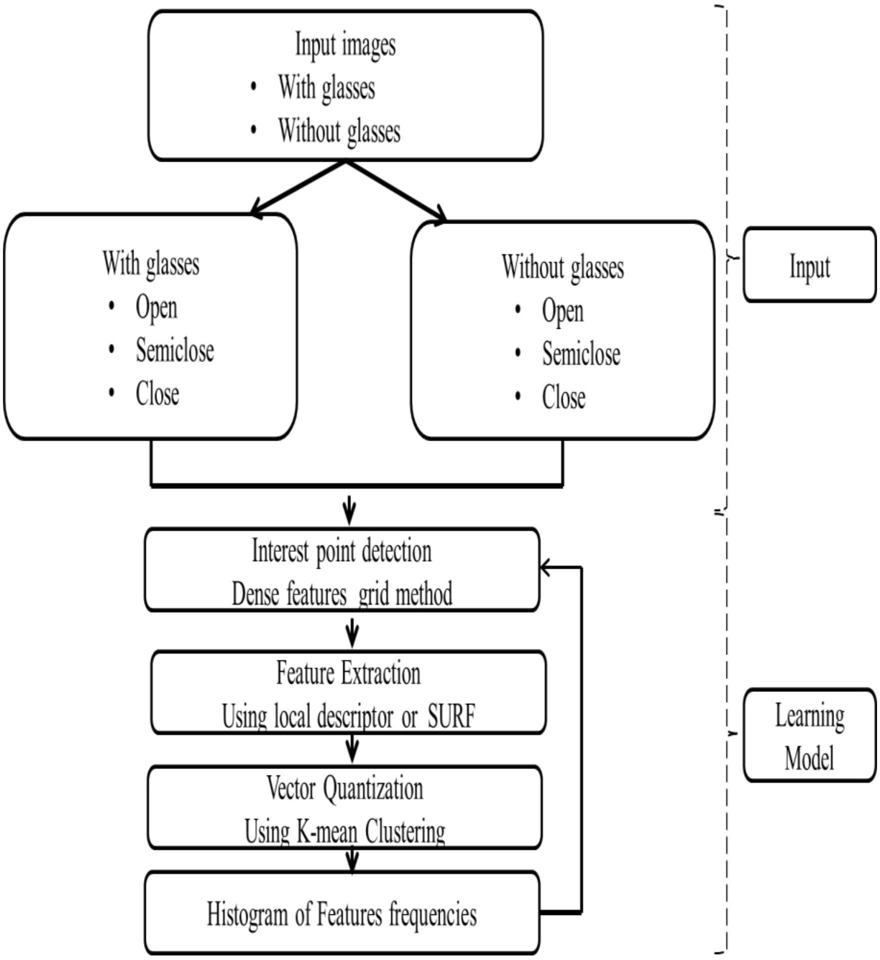


Figure 1. Work flow of this research

Three distinct classes of eye state images used to figure out driver inattention and drowsiness. Referring to Figure 3.2, the eye state is delineated according to the eye closeness degree which are open (a), semi close (b) and close (c). Both classes of semi close and close are considered as distraction or drowsiness. A set of M training images owned by each class are used to build the classifier. From all input images, 50% of them were assigned as the training images for each class while another 50% were used as testing images. The classifier contains the number of classes and the class labels (open, semi close and close) for the input images. The training of each class of image have been done using a learning model technique of Support Vector Machine (SVM), which is a multi-class classifier.

The grid method is used to select feature point and Speeded-Up Robust Feature (SURF) is used to extract features which have been selected. Within this step, feature vectors for each image from the training set are extracted with BlockWidth size [32, 64, 96, 128] pixels. Features are extracted from the images and the approximate nearest neighbour algorithm is used to construct features histogram for each class image. Visual vocabulary is built in form of feature vectors from representative images of each class. Visual vocabulary is defined by a group of selected feature vectors after using k-means clustering algorithm on the feature descriptors into k mutually exclusive clusters. For experimental setup, it is very important to determine the optimal cluster size, k so that the accuracy of classification is optimal. The range of suitable k is prescribed based on the baseline which is number of classes/categories[5]. Table 1 shows comparison of number of classes/categories and the cluster size used. The range of classes/categories is between 2 to 13 are using the k values starting from 160 to 1500. In this experiment, we use the value of k starting from 100 to 700 for 3 classes to determine the optimal accuracy.

The algorithm function trains a Support Vector Machine (SVM) multiclass classifier. Then, feature vectors for the images are represented by the histogram. Selected feature vectors are tested and evaluated using testing set image. Then, the clusters are compressed and grouped by similar characteristics and each cluster centre represents a potent feature vectors. Lastly, the image category classifier is used to predict the query images and determine its category according to class labels. The confusion matrix is used to portray the prediction. A perfect classification results in a normalized matrix containing 1s on the diagonal and an imperfect classification results in fractional value.

Table 1. Comparison of number of classes/categories and the cluster size, k

Previous work	Number of classes/categories	Cluster size, k
Luo et al.[6]	5/2	500
Lu and Ip [7]	3	100/200
Chen et al.[8]	8	1000
Bosch et al.[9]	6	700
Junsong et al.[10]	2	160/500
Monay et al.[11]	4	1000
Bosch et al.[12]	6/8/13	1500

3. Results and Discussion

We present the pre-experiment results or experimental setup for this research. Error! Reference source not found. shows the average accuracy with several values of cluster size, *k*. So, the optimal cluster size setting, *k* = 500 is chose according to the best value of accuracy which is 90%.

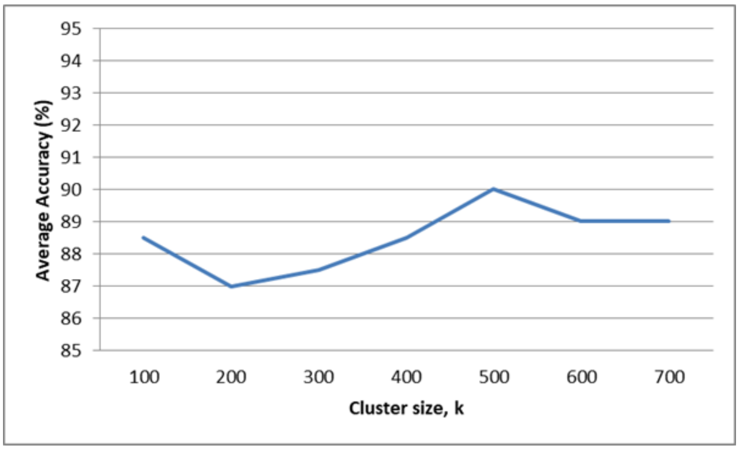


Figure 2. The best cluster size for three classes

The optimal value of k is needed to avoid poor result, possibly to the point of becoming unusable [13] Figure 3 shows the example of images of training set of the experiment A (Experiment A1: ((a), (b), (c)) and second experiment (Experiment A2: (d), (e), (f)). 500 feature vectors extracted for classifier model known as visual words.500 feature vectors establishing the visual words with the visual vocabulary size of 500 words considering all the three classes (open, semi close and close) which is visualize as visual word index.

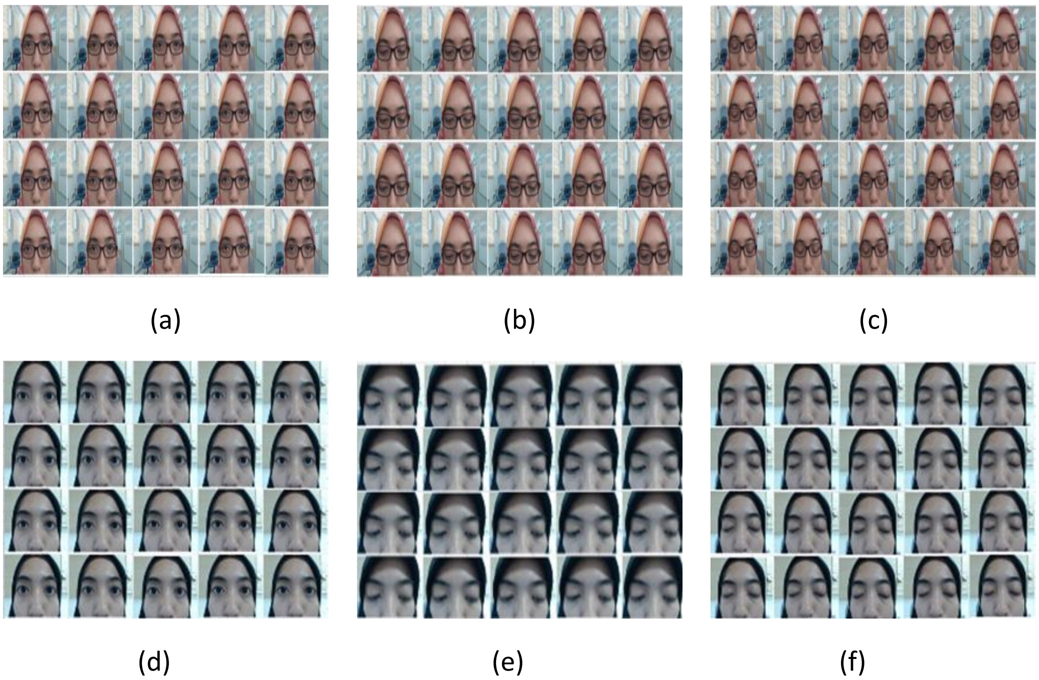
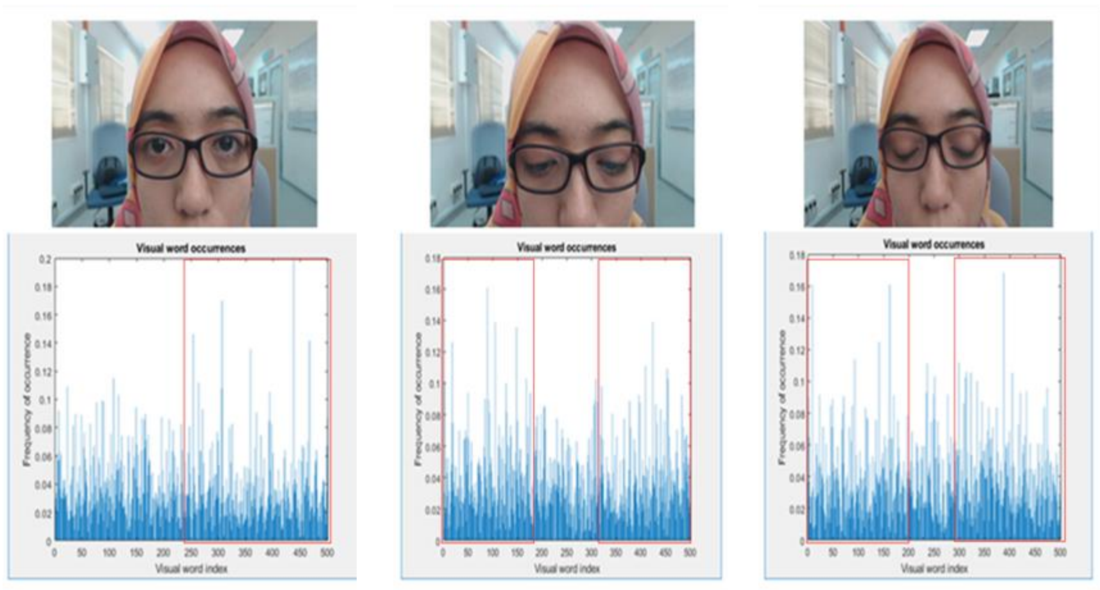


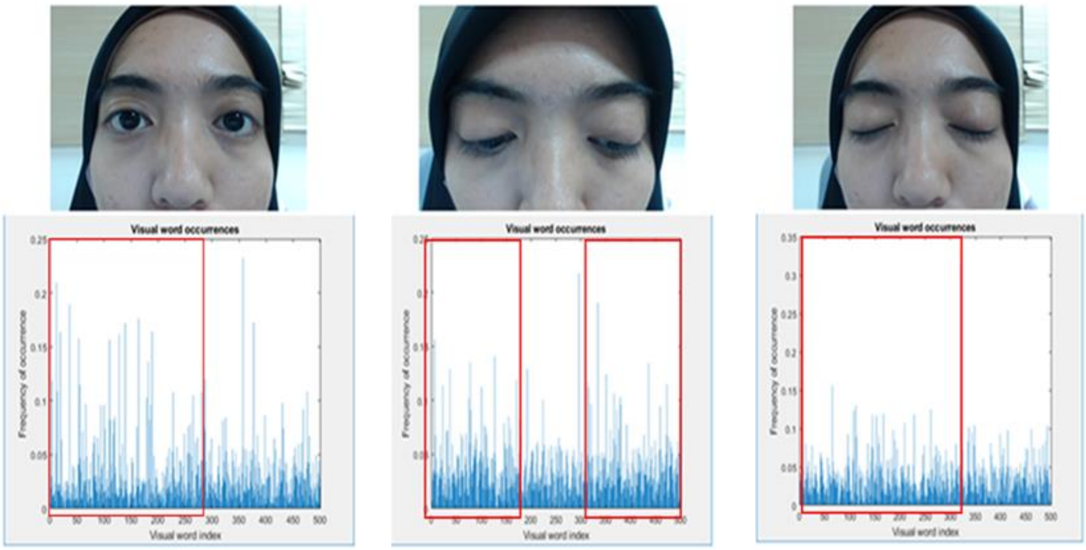
Figure 3. Examples of training /testing

Figure 6 shows the feature histogram of the classes of eye state. From the histogram, they just look similar to each other. However, each eye state has different histogram distribution. The frequency pattern distribution for open state, **Figure 4**(a) is obviously concentrated in the middle to the third quarter of the histogram specifically at visual word index from 250 to 500. Besides, for half close state in **Figure 4** (b), the frequency pattern distribution hoarded in the first and third quarter the histogram which is between visual word index 0 to 170 and 330 to 500. For close state in **Figure 4** (c), the frequency pattern distribution concentrated in first and third quarter of the histogram which is between visual word index 0 to 200 and 270 to 500.

Little bit different for not wearing glasses eye state in **Figure 4** (d) – (f). The frequency pattern distribution for open state in **Figure 4** (d) is obviously concentrated in the first to the middle quarter of the histogram specifically at visual word index from 0 to 270. Besides, for half close state in **Figure 4** (e), the frequency pattern distribution hoarded in the first and third quarter the histogram which is between visual word index 0 to 170 and 330 to 500. For close state in **Figure 4** (f), the frequency pattern hoarded from the first to the middle of the histogram which from visual word index 0 to 320. Overall, the frequency distribution pattern for both learning model (training and testing) is almost the same except for open state (**Figure 4** (a) and **Figure 4** (d)) and close state (**Figure 4** (c) and **Error! Reference source not found.**(f)) which are the frequency distribution are in different quarter of the histogram.



(a) (b) (c)



(d) (e) (f)

Figure 4. Eye state histogram for wearing glasses

The normalized confusion matrices shown in Table 1(a) and (b) summarize the comparison of eye state classification for the different learning model. A row represents a case of the actual class, while a column represents a case of the predicted class. Consequently, the degree of accuracy predicted classes is represented by the values of the diagonal elements. The 'confusion' is expressed by the false classified off-diagonal elements, since they are mistakenly labelled with another class by classifier. An interesting observation is that this method shows an impressive classification result between close and open state. Results show that proposed method for eye state classification can differentiate open and close state very well according to the confusion matrices.

*In the experiment, the classifier is tested and represented with confusion matrix for all query images in **Figure 5**. The normalize confusion matrices for Experiment A are shown in*

Table *2* –

Table 7 for each query image.

(a) (b) (c)

(d) (e) (f)

Figure 5. Types of query images

Table 2. Query image (a)

		Predicted		
		Close	Semiclose	Open
Actual	Close	1	0	0
	Semiclose	0	1	0
	Open	0	0	1

Table 3. Query image (b)

		Predicted		
		Close	Semiclose	Open
Actual	Close	1	0	0
	Semiclose	0	1	0
	Open	0	0	1

Table 4. Query image (c)

		Predicted		
		Close	Semiclose	Open
Actual	Close	1	0	0
	Semiclose	0	1	0
	Open	0.074074	0	0.925926

Table 5. Query image (d)

		Predicted		
		Close	Semiclose	Open
Actual	Close	1	0	0
	Semiclose	0	1	0
	Open	0.090909	0	0.909091

Table 6. Query image (e)

		Predicted		
		Close	Semiclose	Open
Actual	Close	0.909091	0.090909	0
	Semiclose	0	1	0
	Open	0	0	1

Table 7. Query image (f)

		Predicted		
		Close	Semiclose	Open
Actual	Close	0.9375	0.0625	0
	Semiclose	0	1	0
	Open	0	0	1

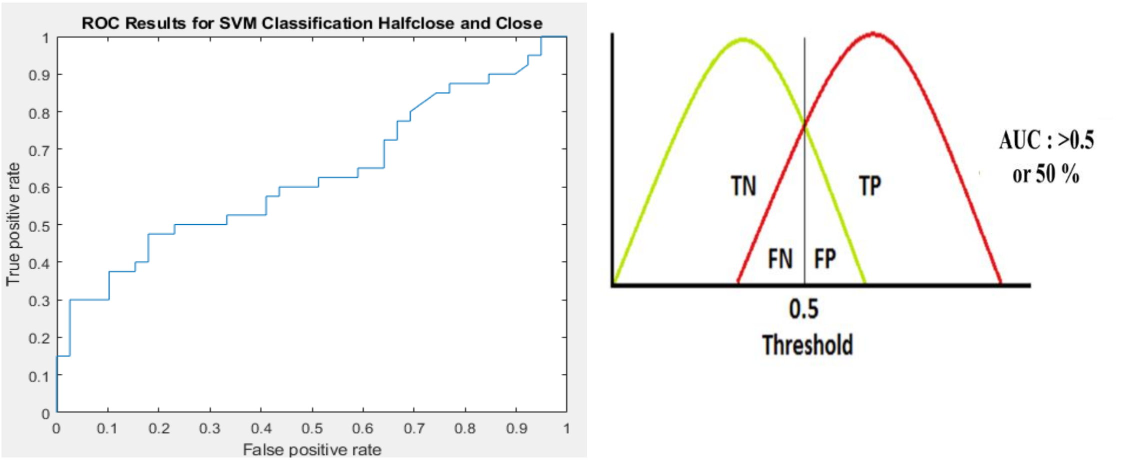
The diagonal elements represent the number of points for which the predicted label is equal to the actual label range between 0 and 1, whereas off-diagonal elements are those that are mislabelled by the classifier. The higher the diagonal values of the confusion matrix, the better indicating many correct predictions. By referring to Table 2 – Table 7, the conclusions can be made as follows:

a. There are no misclassified for all classes for Table 2 and Table 3.
b. There is 7.4% of misclassified Open as Close for Table 4.
c. There is 9.0% of misclassified Open as Close for Table 5.
d. There is 9.0% of misclassified Close as Semiclose for Table 6.
e. There is 6.3% of misclassified Close as Semiclose for Table 7.

The ROC curve shows the trade-off between specificity (1 – false positive rate (FPR)) and sensitivity or true positive rate (TPR). Classifiers that give curves nearer to the top-left corner shows a better performance. The closer the curve comes to the 45-degree of the ROC space, the more accurate the test. Besides, one of the most important evaluation methods for checking any classification model's performance is area under curve (AUC).

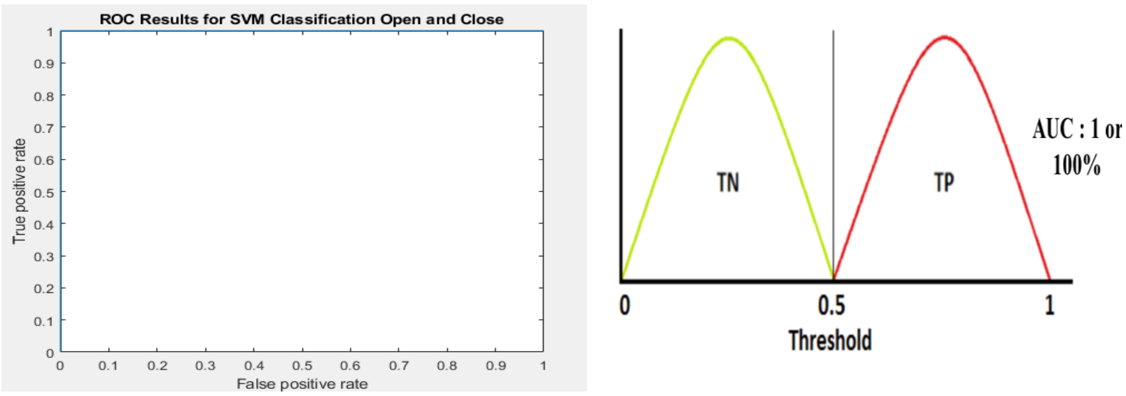
ROC is a probability curve and AUC represents the degree or measure of features distinction or separability. It tells how much model is capable of distinguishing between classes. Higher the

AUC, better the model is at predicting 0s as 0s and 1s as 1s. By analogy, higher the AUC, better the model is at distinguishing between eye's states. In ***Figure 6***(a), when distributions of two features overlap between semiclose and close, it means there is 64.06% chance that model will be able to distinguish.



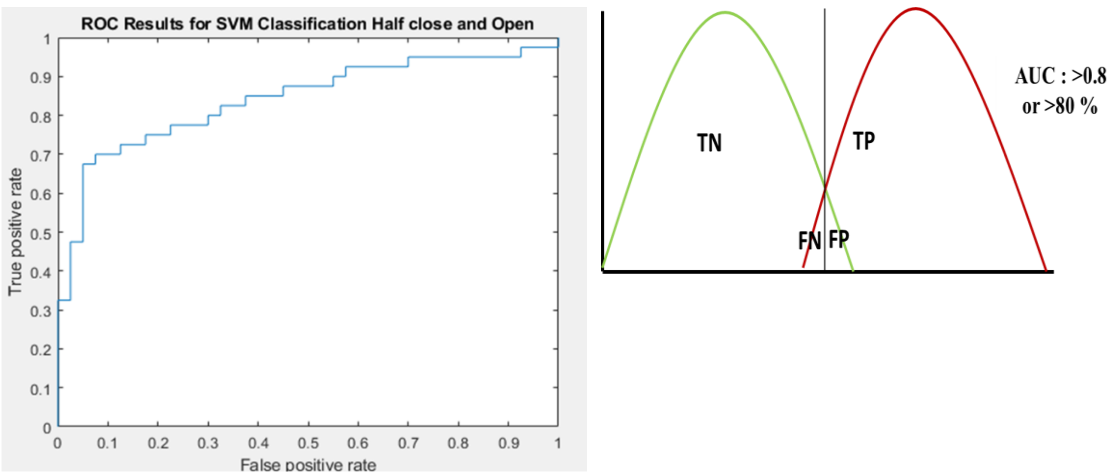
AUC = 0.6406 or 64.06%

(a)



AUC = 1 or 100%

(b)



AUC = 0.8413 or 84.13%

(c)

Figure 6. ROC results

between semiclose and close. However, in Figure 6(b), two curves do not overlap at all means model has an ideal measure of distinction and it is perfectly able to distinguish between open and close. Lastly, in Figure 6(c) distributions of two features overlap between half close and open, it means there is 84.13% chance of the model will be able to distinguish

4. Conclusion

The non-invasive method based on computer vison is applied to detect and classify driver drowsiness and inattention according to their classes. Large number of previous works comprised with one type of eye condition which is without glasses. Differently, in this work, all experiments have taken consideration of two eye conditions of eye with glasses and eye without glasses. The results of the experiments have shown that by using only one type of eye condition data training, the proposed system has been able to classify two types of eye condition according to three classes of eye states. The classification performances by using area under curve of ROC (AUROC) between open and semiclose has acheived more than 80%. This indicates good performance of detection classification of eye distraction and drowsiness. This is important because semiclose and close eye state classes are considered as distraction or drowsiness. It can be concluded that the proposed technique has been able to detect and classify distraction or drowsiness of a subject based on his or her eye state with very high accuracy and fast. However, the study was limited to controlled environment such as non-volatile illumination variation and very minimum movement by the subjects. Further research and study will be conducted using the proposed technique in the condition which is similar to real driving environment.

REFERENCES

[1] G. K. M. MBA, “Drowsiness:Causes,Symptoms and Diagnosis.” [Online]. Available: https://www.healthline.com/symptom/hemorrhage. [Accessed: 20-Nov-2017].

[2] Cholinesterase - Blood, “MedlinePlus Medical Encyclopedia: Drowsiness.” [Online]. Available: http://www.nlm.nih.gov/medlineplus/ency/article/003358.htm. [Accessed: 20-Nov-2017].

[3] MIROS, “2016 MIROS Annual Report,” 2017.

[4] T. Pradhan, A. N. Bagaria, and A. Routray, “Measurement of PERCLOS using eigen-eyes,” *4th Int. Conf. Intell. Hum. Comput. Interact. Adv. Technol. Humanit. IHCI 2012*, pp. 0–3, 2012.

[5] C.-F. Tsai, “Bag-of-Words Representation in Image Annotation: A Review,” *ISRN Artif. Intell.*, vol. 2012, pp. 1–19, 2012.

[6] K. Kesorn and S. Poslad, “An enhanced bag-of-visual word vector space model to represent visual content in athletics images,” *IEEE Trans. Multimed.*, vol. 14, no. 1, pp. 211–222, 2012.

[7] Z. Lu and H. H. S. Ip, “Image categorization with spatial mismatch kernels,” *2009 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work. 2009*, vol. 2009 IEEE, pp. 397–404, 2009.

[8] X. Chen, X. Hu, and X. Shen, “Spatial weighting for bag-of-visual-words and its application in content-based image retrieval,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5476 LNAI, pp. 867–874, 2009.

[9] A. Bosch, X. Muñoz, and R. Martí, “Which is the best way to organize/classify images by content?,” *Image Vis. Comput.*, vol. 25, no. 6, pp. 778–791, 2007.

[10] Y. Junsong, W. Ying, and Y. Ming, “Discovery of collocation patterns: From visual words to visual phrases,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2007.

[11] F. Monay, P. Quelhas, J. M. Odobez, and D. Gatica-Perez, “Integrating co-occurrence and spatial contexts on patch-based scene segmentation,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2006, 2006.

[12] D. Gökalp and S. Aksoy, “Scene classification using bag-of-regions representations,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2007.

[13] P. Eliasson and N. Rosén, “Efficient K-means clustering and the importance of seeding.”