

Feature Based Techniques for a Driver's Distraction Detection using Supervised Learning Algorithms based on Fixed Monocular Video Camera

Syed Farooq Ali¹ and Malik Tahir Hassan²

^{1,2}Department of Software Engineering, SST, UMT,
C-II Johar Town, Lahore, Pakistan,

[¹email: farooq.ali@umt.edu.pk]

[²email: tahir.hassan@umt.edu.pk]

*Corresponding author: Syed Farooq Ali

*Received August 10, 2017; revised January 30, 2018; accepted February 22, 2018;
published August 31, 2018*

Abstract

Most of the accidents occur due to drowsiness while driving, avoiding road signs and due to driver's distraction. Driver's distraction depends on various factors which include talking with passengers while driving, mood disorder, nervousness, anger, over-excitement, anxiety, loud music, illness, fatigue and different driver's head rotations due to change in yaw, pitch and roll angle. The contribution of this paper is two-fold. Firstly, a data set is generated for conducting different experiments on driver's distraction. Secondly, novel approaches are presented that use features based on facial points; especially the features computed using motion vectors and interpolation to detect a special type of driver's distraction, i.e., driver's head rotation due to change in yaw angle. These facial points are detected by Active Shape Model (ASM) and Boosted Regression with Markov Networks (BoRMaN). Various types of classifiers are trained and tested on different frames to decide about a driver's distraction. These approaches are also scale invariant. The results show that the approach that uses the novel ideas of motion vectors and interpolation outperforms other approaches in detection of driver's head rotation. We are able to achieve a percentage accuracy of 98.45 using Neural Network.

Keywords: Distraction Detection; Safety Driving; Neural Network; Computer Vision; Driver Monitoring.

1. Introduction

Drowsiness and distraction are two major reasons of car accidents. Drowsy is a state of a driver when he is half-sleeping for a long period of driving. However, driver distraction is a diversion of attention away from activities crucial for safe driving. According to current estimates, more than 30% of the deaths in car accidents are due to drowsy driving. National Highway Traffic Safety Administration (NHTSA) estimates that 100 000 police reports on vehicle crashes were direct outcomes of driver drowsiness resulting in 71000 injuries, 1550 deaths and \$12.5 billion in monetary losses [1]. Due to this reason, the car manufacturers in the world are having a keen desire to develop a system that can prevent drowsy driving.

Statistical data show that the driving-related accidents and human casualties are increasing in the U.S. and around the globe [2]. The majority of the accidents involve maximum chances of fatalities in Europe and U.S while driving on highways due to just focusing off the road for few seconds which makes all passengers' lives in danger. It is passengers' common practice that in the start of their journey they switch on the radio, program their GPS location for better guidance or check their makeup in the mirror exhibiting partial or complete distraction.

The U.S. Department of Transportation, National Highway Traffic Safety Administration (NHTSA) reports confirmed that in 2012, there were more than 3,228 people killed in all fatal crashes, and driver's distraction was the leading cause of these accidents [1]. The driver's distraction is mainly caused by the 3D head rotation of the driver that can further be categorized into three types, i.e., rotation caused due to change in yaw angle, pitch angle and roll angle (Fig. 2 b). This study proposes one of the most common types of 3D head rotation, i.e., rotation due to change in yaw angle that is a main cause of many accidents.

The AAA foundation for Traffic Safety published a study about some major levels of distraction which involves plenty of control systems. The innovative iOS application called Siri is found to be one of the most dangerous voice assistant. David Strayer inferred this by testing vehicles via simulator on residential streets, and measured brain focus, heart rates and response time of 45 different drivers throughout the experiment [3].

Problems like lack of attention and illness lead a driver to a serious distraction due to losing focus and control on vehicle. Likewise by using cell phones while driving is the serious cause of distraction. In 2010, the NHTSA conducted a telephone survey of 6,002 drivers from all 50 States and the District of Columbia. The report confirms that driving becomes more dangerous when the drivers take their eyes off the road for few seconds or more due to use of cell phones [4].

Effective automatic methods for drivers' distraction detection can help save human lives. Some applications of Distraction Detection are shown in Table 1.

Table 1. Some applications of distraction detection

Applications	Description
Long Haul Transportation	Monitor fatigue, drowsiness and distraction which cause fatal accidents.
Driving Training Centres	Training of beginners and removing their weaknesses.
Safety and Awareness to elderly people	Provides complete safety to people and give awareness to the elderly people.
High speed racing	Le- Mans, a driver can drive in a more focused way. It removes distraction and improves their lap time.

The organization of this paper is as follows. Section 2 describes the previous work mainly related to driver's distraction. In Section 3, data set used in this research is explained. Section 4 proposes different feature based driver distraction techniques containing different features for the detection of driver's head rotation (due to change in yaw angle) using different classifiers including Neural Network, Naïve Bayes, SVM, Adaboost, and Decision Trees. In Section 5, results are presented in terms of percentage accuracy and time efficiency. Section 6 provides the conclusion while Section 7 explains the future directions in which open issues and challenges related to driver monitoring system especially distraction detection for safe driving are presented.

2. Related work

Driver's Distraction is one of the leading causes of accidents in United States [5]. In the last decade, the number of improvements for driver monitoring to ensure safe driving has been more than ever yet a significant number of serious accidents occur around the globe. Literature available in this regard broadly classifies the techniques of driver monitoring into drowsiness and distraction as the two major subjects of study [6]. Techniques used for analysis and detection of drowsiness can be broadly categorized into visual feature based techniques using computer vision approaches and non-visual feature based techniques. Visual feature based techniques include facial expression analysis, mouth and yawning analysis, eye blinking analysis and eye state analysis. Non-visual feature based techniques include vehicle parameter analysis and driver physiological analysis. The distraction detection systems are also based on computer vision approaches that mostly compute the head movement and gaze direction for an early detection of distraction to avoid accidents.

Distraction can be defined as anything that diverts the attention of a driver either in the form of cognitive distraction or visual distraction (head pose and gaze distraction) [7]. A layered algorithm using Dynamic Bayesian Network (DBN) and Supervised Clustering is used to detect cognitive distraction using eye movement and driving performance measures with high efficiency [8]. Head pose and gaze direction of a driver can be measured by applying computer vision techniques. Localized gradient histogram and support vector regressors (SVRs) are used to determine the awareness level by using head pose information [9-11] that is also used to provide the distraction level of a driver. Head pose estimation is even important in a situation when the eyes of driver are not visible. In fact, in order to estimate the gaze information of a driver, his head pose estimation is also required [12]. In [13], Driver Attention Guard is developed that applies novel learning algorithms for tracking the driver's head position and rotation using an array of cameras. An inexpensive vision-

based system is presented that tracks facial features, provides geometric reasoning and estimates head pose and gaze direction to detect driver distraction [5]. The head pose and face detection problems include in-plane head rotation, out-of-plane head rotation, illumination conditions, wearing hat i.e., covering full or part of the head and real time processing [14].

There exist many approaches for driver's head movement including Appearance Template methods, Non Linear Regression Methods, Geometric Methods, Tracking Methods and Feature based methods using Machine Learning techniques [15]. Different statistics show that Feature based methods including shape features (with or without eye position), texture features or hybrid features (consisting of both shape and texture features) are being used for detection of driver's head movement since 2008 and have shown experimental results of around 80% [16-20].

Different features were used to calculate driver's head movement using infrared sensors that can be used for distraction detection and also for drowsiness detection [16]. Different features calculated using Controller Area Network (CAN) and motion sensors (accelerometer and gyroscope) are used to detect distractions caused by leg and head movements with high accuracy of over 90% [17]. Long-term eye closure and gaze of driver are used as features to calculate "attention budget" of driver. If the value of the attention budget is lower than some specific threshold then system generates a warning that the driver is distracted [18]. The temporal relationship between the driver's eye gaze and the peripheral vehicle behaviour can be used as a feature for the detection of driver's cognitive distraction using naïve Bayesian Framework [19]. Different features were used to calculate head rotation (panning, shaking and tilting) and gaze direction that help in the detection of distracted and drowsy driver. Optical flow of facial features acquired using corner detection algorithm was used to detect the distraction caused by the driver's head movement [20]. In the system proposed by Wang *et al.* [14], the position of the head was determined by the vertical projection of the gray level image. An embedded system having feature based algorithm implemented on Raspberry Pi is developed that uses different features including eye detection, eye gaze direction, facial feature tracking and head pose estimation for distraction detection [21]. These features were trained and tested using Support Vector Machine (SVM) classifier. A real time driver detection system is developed that uses semi-supervised methods to alleviate the cost of labelling training data [22] but it was aimed at detecting cognitive distraction. The features consisting of eye and head movements are used to classify the driver's distraction.

Azman *et al.* [23] proposed an approach using correlation between geometric measurements of eye and mouth to detect driver's distraction. Bergasa *et al.* [24] proposed six parameters including face pose to monitor the driver vigilance. Various geometric measurements of a mouth including maximum length, maximum width and normal (current) width of mouth is used to detect driver's distraction [25]. Geometric features of nostril and mouth are also used in facial expression recognition as well [26]. Diaz-Chito *et al.* used various geometric features to predict head pose estimation and also generated its results on their own data set, i.e., DrivFace [27]. Ma *et al.* in 2015 proposed a novel approach based on Fisher vector of local descriptors for the estimation of head rotation due to change in yaw angle [28]. A concise 9-dimensional local descriptor for each pixel is computed that consists of pixel coordinates, magnitude of gradient, intensity, orientation of gradient, first derivative and second derivative. Besides performing facial feature tracking, eye gaze, eyes off the road and fatigue, Dongare *et al.* in 2016 also estimated 3D head pose estimation using POSIT

algorithm [29]. Billah et al. [30] proposed an algorithm that tracks the key body parts of a driver and detects the distraction of a driver using k-nearest neighbour classifier.

It can be observed from this literature survey that although different geometric, appearance and shape based features are present yet there exists a need to combine different features together and investigate their percentage accuracies and time efficiencies for detecting driver's distraction using different classifiers. The study proposes distraction detection approaches containing combination of existing and novel features to detect a special type of driver's head movement, i.e., rotation due to change in yaw angle on a self-generated data set (FT-UMT data set) as well as standard/public datasets. The frame containing such type of driver's head rotation is categorized into two classes, i.e., Distraction (D) or No Distraction (ND).

3. Data Set

Our data set consists of four videos containing frames having Distraction (D) or No Distraction (ND). The first video contains total number of 418 frames, out of which 191 frames contain Distraction (D) while other frames contain No Distraction (ND) as shown in Table 2. Similarly, the second, third and fourth videos contain total of 409, 390 and 397 frames respectively, out of which 221, 215 and 209 frames contain distraction (D) respectively while the rest of frames contain No Distraction (ND).

Fig. 1 shows different frames that contain No Distraction (ND) and also different types of Distraction D. It should be noted that the classification of No Distraction (ND) and Distraction (D) were based on subjective measure of human experts that resulted in the generation of ground truth. After observing various drivers, this subjective measure can be roughly translated into the rotation angle of driver's head. If the frame contains the driver looking straight down towards the camera without any rotation, it means that driver's head has got approximately 0 degree to 20 degrees rotation around y-axis (see Fig. 2 (a) for axis orientation) in anti-clockwise direction or clockwise direction, hence the frame contains No Distraction (ND). However, if the frame contains the driver head rotation of more than 20 degrees to 90 degrees around y-axis in anti-clockwise direction or clockwise direction, it means the frame contains Distraction (D).

Table 2. Video Specifications.

Video No.	Total no of frames	No of Frames with distraction (D)	No of Frames without distraction (ND)
1	418	191	227
2	409	221	188
3	390	215	177
4	397	209	188

The parts a, d, g, h, j and k of the Fig. 1 contain No Distraction (ND) and the rest contain Distraction (D). The parts b and c contain Extreme Right Distraction (ERD), part f contains Extreme Left Distraction (ELD), parts e and l contain Slight Left Distraction (SLD). The part i contains Slight Right Distraction (SRD).

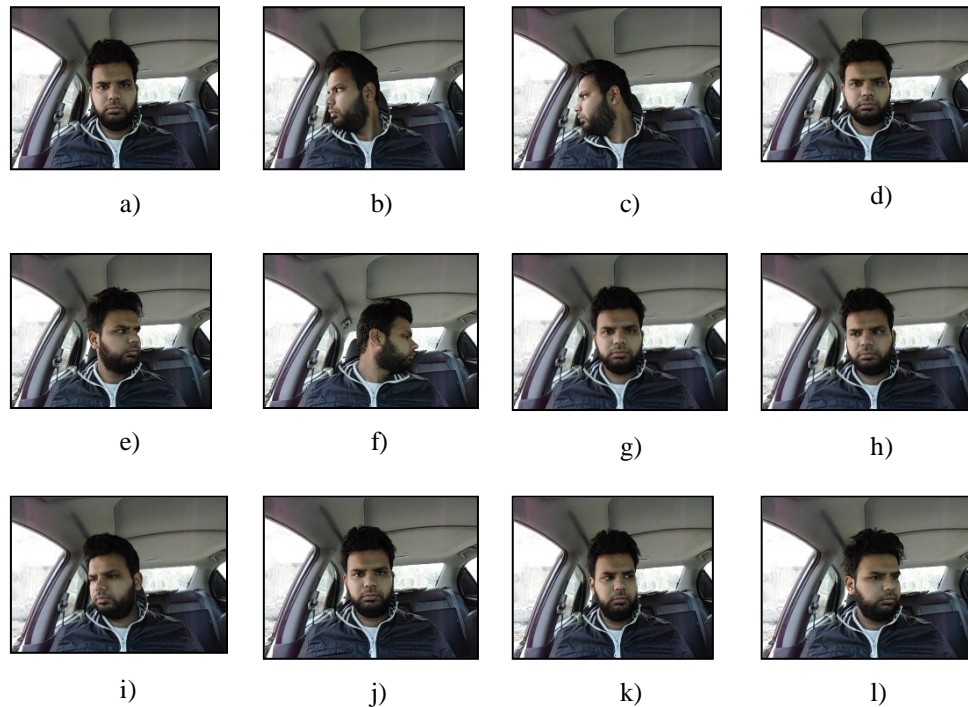


Fig. 1. Data frames having Distraction a) ND b) D c) D d) ND e) D f) D g) ND h) ND i) D j) ND k) ND l) D.

The direction of axis used in this study and three degrees of freedom of a human head are given in **Fig. 2**. It can be observed that the head rotation around y-axis, x-axis and z-axis in part **Fig. 2** (a) corresponds to head rotation caused due to the change in yaw angle (horizontal axial rotation, i.e., right to left rotation of the head), change in pitch angle (sagittal flexion and extension i.e., forward to backward movement of the neck) and change in roll angle (a frontal lateral bending i.e., right to left bending of the neck) respectively in **Fig. 2** (b) [15]. The study proposes a system to detect distraction caused by rotation of a driver's head around y- axis, i.e., driver's head rotation due to change in yaw angle.

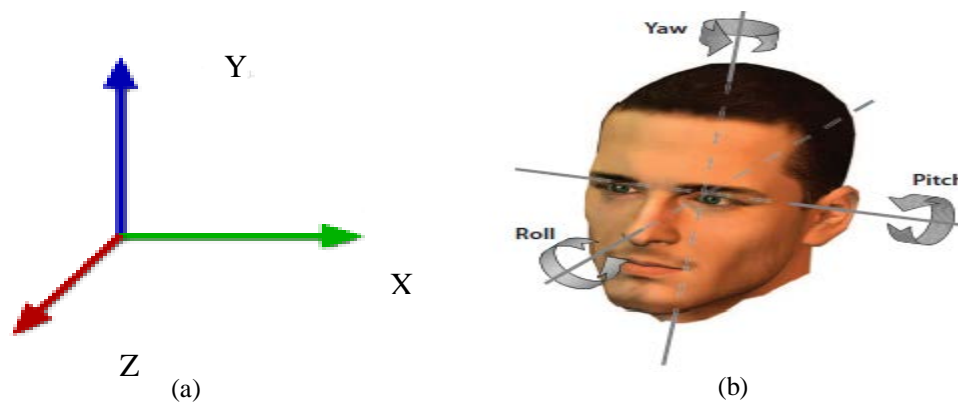


Fig. 2. (a) 3D coordinate system used in this study (b) The three degrees of freedom of a human head can be described by egocentric rotation angles pitch, roll and yaw [15].

4. Our Methodology

We wish to calculate the 3D rotation vector R corresponding to change in yaw angle of the driver's head in a sequence of 2D images that are spread across the temporal dimension. 3D rotation vector R is given by the Eq. 1.

$$R = R_Y(\theta) \quad (1)$$

where $(\theta) = (\text{yaw angle})$

If the driver's head rotation around y-axis reaches above certain threshold, it means frame contains Distraction (D) and hence any serious accident can occur. Otherwise frame contains No Distraction (ND). If the driver's head has got approximately 0 degree to 20 degrees rotation (threshold=20 degrees) around y-axis (see Fig. 2 (a) for axis orientation) in anti-clockwise direction or clockwise direction, then the frame contains No Distraction (ND). However, if the frame contains the driver head rotation of more than 20 degrees to 90 degrees around y-axis in anti-clockwise direction or clockwise direction, it means the frame contains Distraction (D). The problem can be formulated as:

If for any frame $R \geq \text{Threshold}$, Frame contains Distraction (2)

Else $R < \text{Threshold}$, Frame does not contain Distraction

where

$$R = R_Y(\theta) = f(\text{Feature Vector})$$

It can be inferred from the Eq. 1 and Eq. 2 that the driver's head rotation around y-axis is the function of the features computed for that frame. Hence, different feature based approaches are presented in this paper.

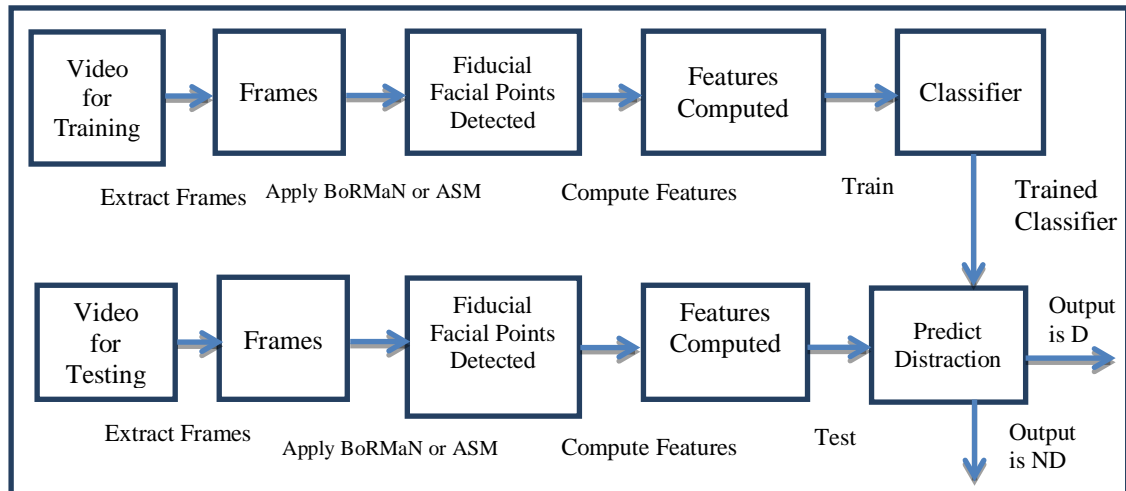


Fig. 3. Flow Chart showing the methodology

The features proposed in our approaches help in calculating the driver's head rotation around y-axis i.e., egocentric rotation caused due to change in yaw angle. These features are used to train and test the classifiers. In the testing phase of the classifier, if the values of the

features determine that head rotation around y-axis of the driver reaches above the threshold, then the classifier will automatically classify that frame as containing Distraction (D).

The study proposes four different approaches that extract features from the sequence of 2D frames extracted from a video to calculate the driver's head rotation around y-axis that result in distraction. Hence, the distraction can be avoided by alarming the distracted driver without wastage of any time that would in turn help in reducing serious injuries and mishaps. Fig. 3 gives the overview of our methodology.

Initially, four videos were acquired as input and frames were extracted from those videos. The 16 and 22 fiducial facial points for each frame were calculated using Active Shape Models (ASM) [31] and Boosted Regression Coupled with Markov Networks (BoRMaN) [32] respectively as shown in Fig. 4 (b, c). These fiducial facial points were used to calculate different features that were used to detect driver's distraction. These features and their class labels were given as input to a classifier for training. For testing purpose, different features of new images were given to the trained classifier. Distraction is detected on the basis of classifier output.



Fig. 4. a) Original b) Fiducial Facial Points Detected by ASM c) 22 Fiducial Facial Points Detected by BoRMaN.

Distraction detection of drivers through intelligent camera requires face processing as well. Neural networks are widely used for facial image processing and detection algorithms [33]-[37] that further motivated us to use neural network (Multilayer Perceptron (MLP)) for distraction detection of drivers, besides using Naïve Bayes, Support Vector Machine (SVM), Ada boost, Decision Table, J48 Decision Tree and Nearest-Neighbour-like Algorithm using Non-Nested Generalized Exemplars (Nnge) [38]-[43].

For detecting distraction of a driver, four approaches were adopted which are described in this section. First three approaches differ by the selection of features, however, the fourth approach has the same set of features as that of approach 3 but it also calculates the feature values of those frames in which first three approaches fail, i.e., frames having high rotation angle of driver's head.

Approach 1: Ratios of Facial Components (RFC)

In Approach 1, the ratios as features are calculated from the fiducial facial points given by BoRMaN and ASM algorithms. The following ratios are calculated for detecting distraction.

1. Ratio of length to width of lips
2. Ratio of length to width of left eye
3. Ratio of length to width of right eye
4. Ratio of areas of $\Delta 2,6,17$ & $\Delta 2,6,21$
5. Ratio of areas of $\Delta 2,6,17$ & $\Delta 18,20,17$
6. Ratio of areas of $\Delta 2,6,21$ & $\Delta 18,20,17$

The delta-notation Δ denotes the triangle. The points 2, 6, 17, 18, 20 and 21 are marked in **Fig. 5 (b, c)**. The details of these features are presented in Section 4.1.

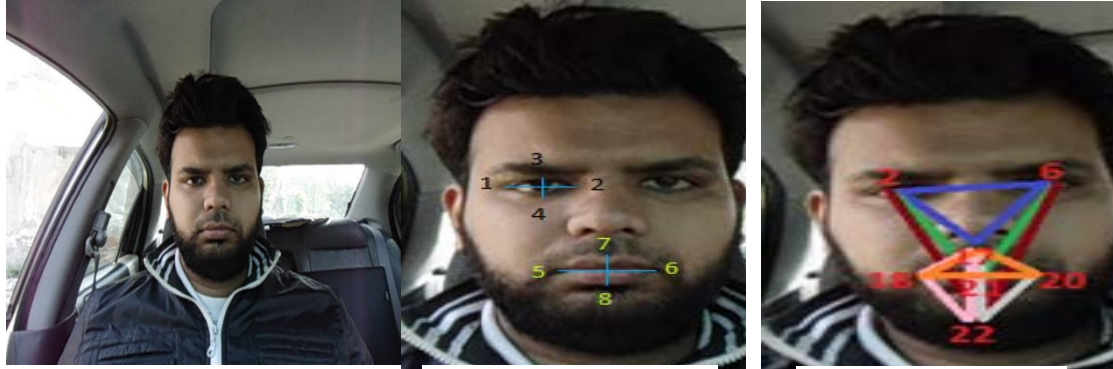


Fig. 5. (a) Original Image

(b) Ratios of length to width of eyes and lips

(c) Triangles and Angles

Approach 2: Ratios of Facial Components and Areas of Triangles (RFCA)

In this approach, the list of features in Approach 1: RFC is extended by adding the following area based features calculated from the facial points given by BoRMaN and ASM algorithms.

7. Area of $\Delta 2,6,17$
8. Area of $\Delta 2,6,21$
9. Area of $\Delta 18,20,17$

The points 2, 6, 17, 18, 20 and 21 are marked in **Fig. 5 (b, c)**. The details of these features are presented in Section 4.1.

Approach 3: Ratios of Facial Components, Areas of Triangles and Angles (RFCAA)

In this Approach, the following angle based features are added to the features of approach 2: RFCA.

10. Angle between 2,6,17 (i.e., $\angle 2,6,17$)
11. Angle between 2,6,21 (i.e., $\angle 2,6,21$)
12. Angle between 2,6,22 (i.e., $\angle 2,6,22$)
13. Angle between 17,20,22 (i.e., $\angle 17,20,22$)
14. Angle between 6,2,17 (i.e., $\angle 6,2,17$)
15. Angle between 6,2,21 (i.e., $\angle 6,2,21$)
16. Angle between 6,2,22 (i.e., $\angle 6,2,22$)
17. Angle between 17,18,22 (i.e., $\angle 17,18,22$)

The points 2, 6, 17, 18, 20 and 21 are marked in **Fig. 5 (b, c)**. The details of these features are presented in Section 4.1.

Approach 4: Ratios of Facial Components, Areas of Triangles, Angles and Interpolation using Motion Vectors (RI)

In first three approaches, the facial point detection algorithms including ASM and BoRMAN fail to detect fiducial points of non-frontal faces due to high rotation angle of a driver's head. This makes it impossible to calculate feature values of those frames. In order to solve this problem, Approach 4: RI uses interpolation technique based on motion vectors to calculate the missing values of features of those frames as shown in Fig. 6. Note that this approach uses the same list of features as Approach 3: RFCAA

Suppose we want to calculate the feature value of $(i+1)^{\text{th}}$ frame given that feature value of i^{th} frame and $(i-1)^{\text{th}}$ frame are already calculated. We assume that percentage change in feature value is proportional to the magnitude of motion vectors (M.V.). For calculation of motion vectors (i.e., mv_x and mv_y), block searching technique is used in which each frame is divided into small blocks of a fix size. The nose of driver's face is taken as a reference point and motion vector between any two consecutive frames (i.e., i^{th} and $(i-1)^{\text{th}}$ frame) is estimated by taking the difference of the nose block coordinates of i^{th} frame and nose block coordinates of $(i-1)^{\text{th}}$ frame. Then the magnitude of motion vectors is calculated using the Eq. 3.

$$\text{Magnitude of Motion Vectors} = M.V. = \sqrt{(mv_x)^2 + (mv_y)^2} \quad (3)$$

The percentage change of feature values between $(i-1)^{\text{th}}$ frame and i^{th} frames is calculated by the formula given in Eq. 4.

$$\text{Percentage change} = \left(\frac{\text{Value of current frame} - \text{Value of Previous Frame}}{\text{Value of Previous Frame}} \right) \times 100 \quad (4)$$

The percentage change of feature values between i^{th} frame and $(i+1)^{\text{th}}$ frame cannot be calculated using Eq. 4 as the feature value of $(i+1)^{\text{th}}$ frame is not known. As mentioned earlier, that the percentage change of feature value is proportional to the magnitude of motion vectors, hence percentage change between i^{th} frame and $(i+1)^{\text{th}}$ frame is calculated using motion vectors. This percentage change of feature value between i^{th} frame and $(i+1)^{\text{th}}$ frame, in turn, is used to calculate the feature value of $(i+1)^{\text{th}}$ frame.

For example, consider case 3 of Fig. 6, in which feature 'Area' is required to be calculated for the $(i+1)^{\text{th}}$ frame given that the Area for $(i-1)^{\text{th}}$ frame and i^{th} frame are 2000 and 1500 respectively. Moreover, the calculated magnitude of motion vectors between $(i-1)^{\text{th}}$ frame and i^{th} frame is 5 and the calculated magnitude of motion vectors between i^{th} frame and $(i+1)^{\text{th}}$ frame is 4. The percentage change in the area between $(i-1)^{\text{th}}$ frame and i^{th} frame is calculated using Eq. 4 and it came out to be 25. The percentage change in the area between i^{th} frame and $(i+1)^{\text{th}}$ frame is calculated using magnitude of motion vectors between i^{th} frame and $(i+1)^{\text{th}}$ frame and is came out to be 20. Using the area of i^{th} frame, and value of percentage change between i^{th} frame and $(i+1)^{\text{th}}$ frame, the area of $(i+1)^{\text{th}}$ frame is calculated using the percentage change formula given in Eq. 4 and it came out to be 1200.

In this way, all the features of non-frontal faces having high rotation angle of driver's head are calculated and then the features are passed to classifiers for training and testing.

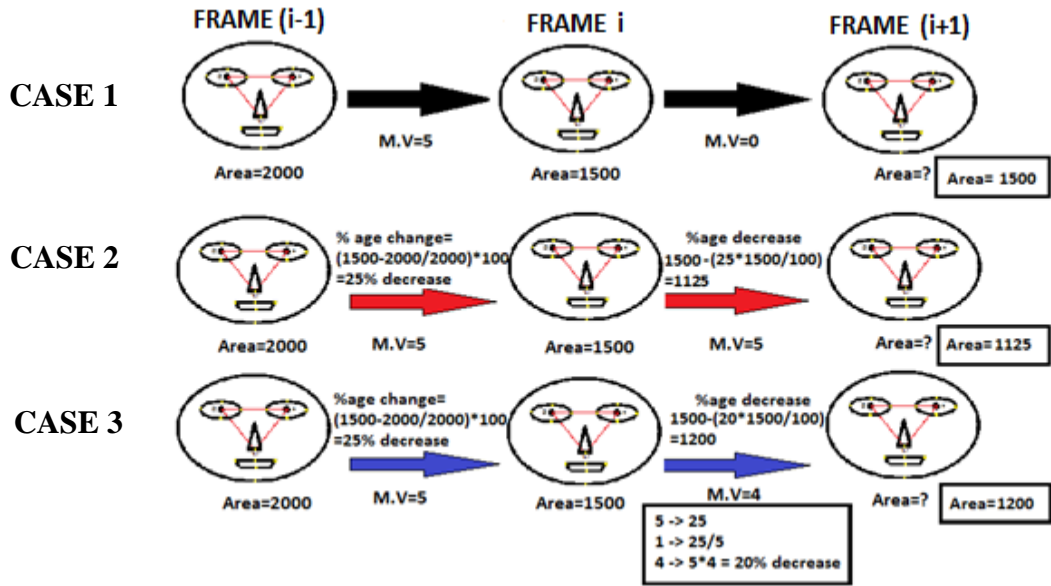


Fig. 6. Motion Vector and Interpolation Technique

The details of the features used in above approaches are presented in Section 4.1.

4.1 Features Description

The features being used for detecting distraction are explained as follows:

4.1.1 Ratio of Length to Width of Lips

The ratio of length to width of the lips is computed to detect distraction as its value varies significantly with distraction. Length of lips or width of lips alone can be taken as feature but in that case it would not be scale invariant. That is why the ratio of length and width of lips is taken as a feature to make it scale invariant as shown in Eq. 5.

$$\text{Ratio} = \text{Lips_Length} / \text{Lips_Width} \quad (5)$$

Where, length and width are calculated by using the Euclidean distance formula. To find the length of lips in Fig. 5 (b), the Euclidean distance is required to be calculated between the point 5 (x_5, y_5) and point 6 (x_6, y_6) as shown in Eq. 6.

$$d(\text{Lips_Length}) = \sqrt{(x_6 - x_5)^2 + (y_6 - y_5)^2} \quad (6)$$

Where, point 5 (x_5, y_5) and point 6 (x_6, y_6) are the points which are detected by ASM or BoRMaN.

Similarly to find the width of lips in Fig. 5 (b), the Euclidean distance is required to be calculated between point 7 (x_7, y_7) and point 8 (x_8, y_8) as shown in Eq. 7.

$$d(\text{Lips_Width}) = \sqrt{(x_8 - x_7)^2 + (y_8 - y_7)^2} \quad (7)$$

4.1.2 Ratio of Length to Width of Eyes

The ratio of length to width of the left eye and the ratio of length to width of the right eye are computed as shown in Eq. 8 and Eq. 9 respectively.

$$\text{Ratio for Left Eye} = \text{Left Eye_Length} / \text{Left Eye_Width} \quad (8)$$

$$\text{Ratio for Right Eye} = \text{Right Eye_Length} / \text{Right Eye_Width} \quad (9)$$

Left Eye_Length, *Left Eye_Width*, *Right Eye_Length* and *Right Eye_Width* are calculated using Euclidean distance formula similarly as shown in Eq. 6 and Eq. 7.

4.1.3 Area of Triangles

By calculating the areas of the triangles as shown in **Fig. 5 (c)**, distraction is detected. The area of the triangle is calculated by using the Hero's formula. If a driver is distracted either in the extreme right direction or extreme left direction, slightly right direction or slightly left direction then the area of the triangle varies that can be used for detecting distraction. The following are the three main triangles which are being used for calculating areas.

- $\Delta 2,6,17$ (triangle obtained by using the point 2, 6 and 17 of **Fig. 5 (c)**, where Δ represents the of triangle)
- $\Delta 2,6,21$ (the triangle obtained by using the point 2, 6 and 21 of **Fig. 5 (c)**)
- $\Delta 18,20,17$ (the triangle obtained by using the point 18, 20 and 17 of **Fig. 5 (c)**)

The area of $\Delta 2, 6, 17$ is calculated by using the Hero's formula which is given in Eq. 10.

$$\text{The area of } \Delta = \sqrt{(s \times (s - a)(s - b)(s - c))} \quad (10)$$

where s is the average length of three sides of triangle as shown in Eq. 11 while a , b and c are the lengths of the sides of a triangle calculated using Euclidean formula as shown in Eq. (12 – 14).

$$s = \frac{(a + b + c)}{2} \quad (11)$$

$$a = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (12)$$

$$b = \sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2} \quad (13)$$

$$c = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2} \quad (14)$$

4.1.4 Ratio of Areas

If a driver is distracted either in the extreme right direction or extreme left direction, slightly right direction or slightly left direction then the ratio of the areas of triangle varies that can be used for detecting distraction. As the ratios of areas of triangles is scale invariant hence this feature not only helps in detecting distraction but also contributes in making the distraction as scale invariant. The three triangles as shown in **Fig. 5 (c)** are used to calculate the ratios of the triangles $\Delta 2,6,17$, $\Delta 2,6,21$ and $\Delta 18, 20, 17$ as shown in Eq. (15- 16).

$$\text{Ratio of areas of triangles} = \frac{\text{Area of triangle 1}}{\text{Area of triangle 2}} \quad (15)$$

$$\text{Ratio of areas of } \Delta 2, 6, 17 \text{ \& } \Delta 2, 6, 21 = \frac{\text{Area of } \Delta 2, 6, 17}{\text{Area of } \Delta 2, 6, 21} \quad (16)$$

Similarly, the ratios of areas of rest of triangles are calculated in the similar way as of Eq. 18. The area of triangles can be calculated by using Hero's formula by using Eq. (10 – 14).

4.1.5 Angles

The values of angles of different triangles as shown in **Fig. 5 (c)** vary significantly in case of driver's head rotation around y-axis hence; it can be used as a feature for detecting driver's distraction. The angles of the triangle are computed by using the Fundamental Law of Cosines which is Eq. (17 – 19).

$$\text{Alpha } (\alpha) = \cos^{-1} \left(\frac{b^2 + c^2 - a^2}{2 \times b \times c} \right) \quad (17)$$

$$\text{Beta } (\beta) = \cos^{-1} \left(\frac{a^2 + c^2 - b^2}{2 \times a \times c} \right) \quad (18)$$

$$\text{Gamma } (\gamma) = \cos^{-1} \left(\frac{a^2 + b^2 - c^2}{2 \times a \times b} \right) \quad (19)$$

The Eq. (17-19) are used to compute the angle based features for our approaches.

5. Results and Discussion

The results of four approaches in terms of percentage accuracy (percentage split = 64%, i.e., with 64 % training data and 36% testing data) and time efficiency are shown in **Table 3** and **Table 5** respectively. **Table 3** shows that the performance in terms of percentage accuracy of 'Approach 4: Ratios of Facial Components, Areas of Triangles, Angles and Interpolation using Motion Vectors (RI)' using ASM is better than rest of the three approaches for all the classifiers. Moreover, the average percentage accuracy of Approach 4: RI using ASM is significantly better than rest of the approaches as well. Possible reasons of better performance of Approach 4: RI includes addition of more features as compared to the rest of approaches and use of Interpolation technique based on Motion Vectors. The Neural Network classifier [44, 45] using Approach 4 and ASM algorithm outperforms the rest of classifiers including Decision Table, Ada boost, SVM, Naïve Bayes, J48 and Nnge. Moreover, the average percentage accuracy of Neural Network classifier is better than the rest of the classifiers. It is because of the fact that Neural Network can approximate any function, regardless of its linearity. Moreover, the Neural Network improves gradually by making discriminating weighted feature combinations.



Fig. 7. Non Frontal Faces a) Extreme Right Distraction (ERD) b) Extreme Left Distraction (ELD) c) Slight Right Distraction d) Slight Left Distraction

Table 3. Results of Approach 1: Ratios of Facial Components (RFC), Approach 2: Ratios of Facial Components and Areas of Triangles (RFCA), Approach 3: Ratios of Facial Components, Areas of Triangles and Angles (RFCAA) and Approach 4: Ratios of Facial Components, Areas of Triangles, Angles and Interpolation using Motion Vectors (RI) in terms of percentage accuracy with training (64): testing (36) percentage split.

Classifier Name		Approach 1: RFC		Approach 2: RFCA		Approach 3: RFCAA		Approach 4: RI		Avr. % Acc.
		ASM	BoR MaN	ASM	BoR MaN	ASM	BoR MaN	ASM	BoR Ma N	
Naïve Bayes	TP	0.93	0.85	0.91	0.89	0.92	0.86	0.96	0.87	
	FP	0.36	0.40	0.35	0.28	0.26	0.37	0.05	0.15	
	ROC	0.92	0.83	0.93	0.91	0.89	0.89	0.98	0.92	
	% Acc.	91.76	84.87	90.93	89.12	92.31	85.58	95.53	86.77	90.02
SVM	TP	0.91	0.84	0.91	0.84	0.91	0.91	0.96	0.90	
	FP	0.46	0.45	0.48	0.45	0.47	0.25	0.05	0.11	
	ROC	0.73	0.70	0.71	0.69	0.72	0.83	0.96	0.89	
	% Acc.	90.66	84.16	90.66	83.92	90.93	91.02	95.70	89.86	89.62
Ada- boost	TP	0.92	0.85	0.92	0.93	0.94	0.94	0.98	0.94	
	FP	0.32	0.43	0.32	0.17	0.15	0.15	0.02	0.07	
	ROC	0.93	0.84	0.93	0.93	0.97	0.83	0.996	0.98	
	% Acc.	91.76	84.87	91.76	93.14	93.96	93.85	97.59	93.99	92.62
Decision Table	TP	0.94	0.85	0.93	0.92	0.95	0.91	0.98	0.94	
	FP	0.16	0.35	0.25	0.19	0.15	0.18	0.02	0.06	
	ROC	0.91	0.79	0.92	0.93	0.93	0.93	0.98	0.97	
	% Acc.	94.23	85.34	93.13	91.96	94.78	90.54	98.11	94.16	92.78
Neural Network	TP	0.92	0.86	0.94	0.93	0.95	0.93	0.98	0.96	
	FP	0.41	0.30	0.24	0.15	0.16	0.09	0.02	0.04	
	ROC	0.63	0.82	0.85	0.95	0.90	0.98	0.98	0.99	
	% Acc.	91.48	86.29	93.41	93.14	95.33	93.38	98.45	96.39	93.48
J48	TP	0.94	0.86	0.93	0.92	0.94	0.92	0.98	0.95	
	FP	0.22	0.26	0.27	0.15	0.12	0.16	0.02	0.05	
	ROC	0.88	0.86	0.86	0.93	0.88	0.86	0.98	0.94	
	% Acc.	93.68	86.29	93.41	92.20	93.96	91.49	97.94	94.67	92.95
Nnge	TP	0.93	0.86	0.92	0.92	0.96	0.94	0.96	0.96	
	FP	0.18	0.28	0.19	0.15	0.07	0.10	0.04	0.04	
	ROC	0.88	0.79	0.86	0.89	0.94	0.92	0.96	0.96	
	% Acc.	93.13	86.05	92.31	91.72	95.60	93.62	95.70	95.70	92.98
Avr. % Acc.		92.4	85.4	92.2	90.8	93.8	91.4	97.0	94.1	

The second and third best percentage accuracies are shown by Decision Table and J48 classifiers respectively using Approach 4: RI and ASM algorithm as shown in **Table 3**. The possible reasons of better percentage accuracies of these classifiers include completeness property and ability to handle missing values. The completeness property guarantees checking of all combinations of condition values [42].

It can also be seen that in 'Approach 1: RFC', Neural Network does not outperform other classifiers as in RFC Approach there are lesser number of features, i.e., ratios of lengths and areas. But as the number of features gradually increases from Approach 2 till Approach 4, it can be observed that the performance of Neural Network in terms of percentage accuracy improves significantly and it reaches its best in Approach 4 (using ASM) i.e., 98.45%.

It can also be observed from the **Table 3** that ASM performs better than BoRMaN in most of the cases. One possible reason is the better robustness of ASM for non-frontal facial images (shown in **Fig. 7**) as compared to BoRMaN [46].

Table 4 and **Fig. 8** provides a comparison of existing approaches and proposed approach (RFCAA using NN) in terms of percentage accuracy with 10-fold cross validation on our own dataset (FT-UMT dataset as mentioned in Section 3), DrivFace database, Boston University (BU) dataset and Pointing'04 dataset using Neural Network (NN) classifier. The existing approaches include 'Existing Approach 1: Eye and Mouth based Statistics (EM-Stat)' [23], 'Existing Approach 2: Triangle Area (TA)' [24], and 'Existing Approach 3: Anthropometric measurement of a mouth (AM-Mouth)' [25].

Table 4. Results of Existing Approaches including 'Existing Approach 1: Eye and Mouth based Statistics (EM-Stat)', 'Existing Approach 2: Triangle Area (TA)', 'Existing Approach 3: Anthropometric measurement of a mouth (AM-Mouth)' and Proposed Approach i.e., 'Approach 3: RFCAA using NN' using ASM in terms of percentage accuracy with 10-fold cross validation on our own data set (FT-UMT dataset as mentioned in Section 3), DrivFace database¹, Boston University (BU) dataset² and Pointing'04 dataset³ using Neural Network (NN) classifier.
EA=Existing Approach, PA=Proposed Approach, ds=dataset, db=database

Data set	Existing Approaches			Proposed Approach
	EA 1: EM-Stat (2011) [23]	EA 2: TA (2006) [24]	EA 3: AM-Mouth (2004) [25]	PA3: RFCAA using NN
FT-UMT	95.25	90.10	92.68	95.25
DrivFace	90.78	89.56	89.74	92.35
BU ds	86.33	71.58	55.04	89.20
Pointing'0	82.82	67.40	72.24	82.82
Average	88.79	79.66	77.43	90.78

The 'Existing Approach 1: EM-Stat' is based on Azman et al. paper of 2011. The proposed approach outperforms this existing approach on all four datasets. The idea of using the area of triangle between left eye, right eye and nostril in 'Existing Approach 2: TA' came from Bergasa et al. paper of 2006, in which he calculated six parameters including face pose to monitor driver vigilance. The results of our proposed approach outperform 'Existing Approach 2: TA' on all the four datasets. The 'Existing Approach 3: AM-Mouth' of

¹The DrivFace database is available at <http://adas.cvc.uab.es/site/index.php/datasets>

²The Boston University (BU) dataset is available at <ftp://csr.bu.edu/headtracking/>

³The Pointing'04 dataset is available at <http://www-prima.inrialpes.fr/Pointing04/data-face.html>

Rongben et al. uses the anthropometric measurements of mouth including maximum length of mouth, maximum width of mouth and current width of mouth to detect a driver's distraction. Our proposed approach again yields better percentage accuracy than this existing approach on all the four datasets. Moreover, the average percentage accuracy of our proposed approach 'PA3: RFCAA using NN' outperforms the average percentage accuracies of all the existing approaches as mentioned in [Table 4](#).

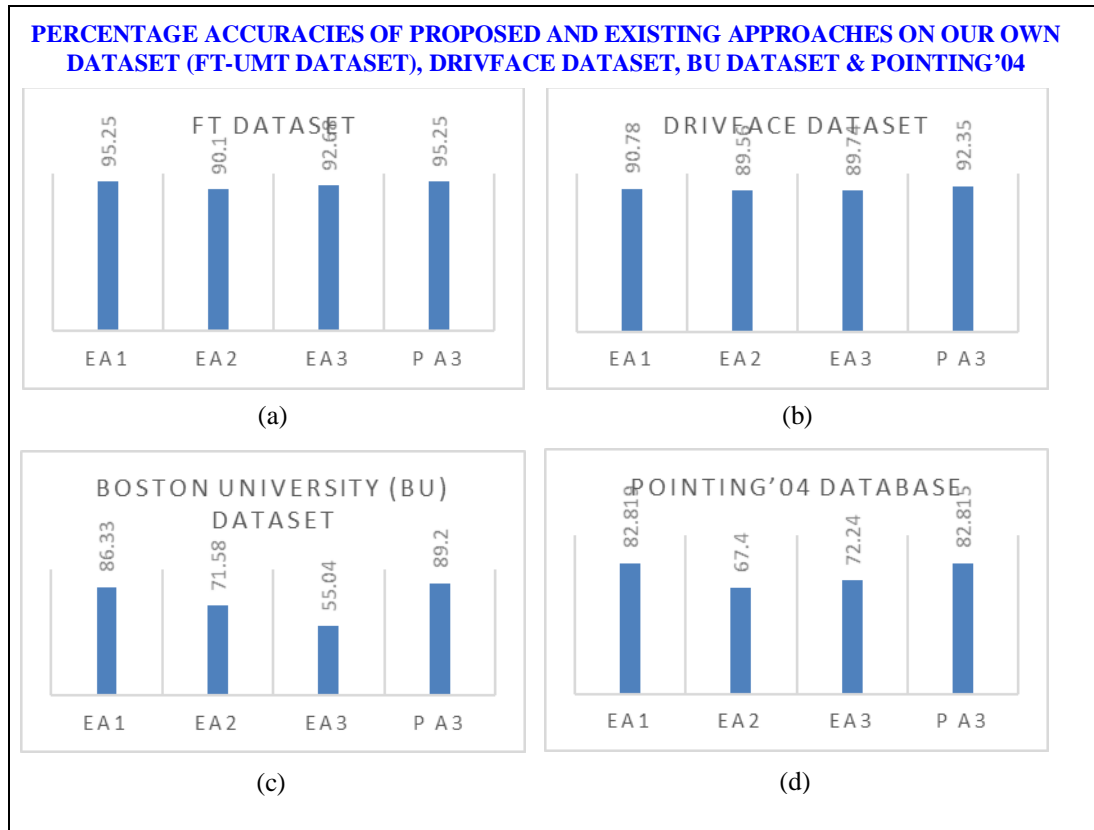


Fig. 8. Percentage accuracies of Existing Approaches including 'Existing Approach 1: EM-Stat, 'Existing Approach 2: TA', 'Existing Approach 3: AM-Mouth and Proposed Approach i.e., 'Approach 3: RFCAA using NN' with 10-Fold cross validation using ASM on a) Our Own data set (FT-UMT ds), b) DrivFace, c) BU dataset and d) Pointing'04 dataset using Neural Network (NN) classifier.

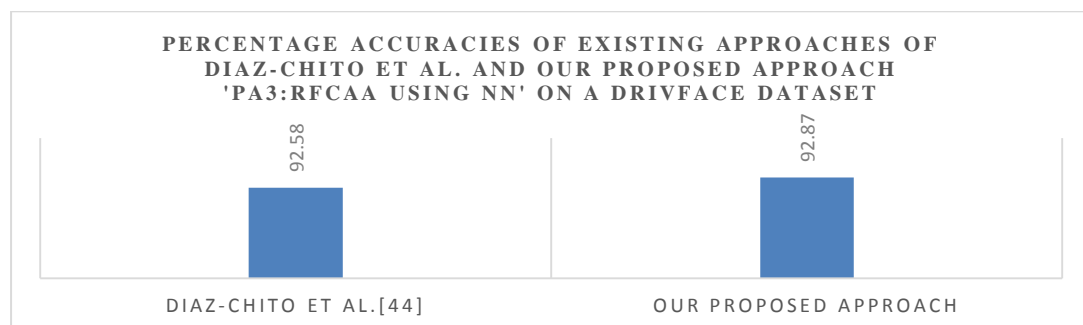


Fig. 9. Percentage accuracies of Diaz Chito et al. (2016) Approach [27] and our proposed approach 'PA3:RFCAA using NN' on DrivFace dataset.

It can be observed in [Fig. 9](#) that our proposed approach 'PA3: RFCAA using NN' (with 5- fold cross validation) outperforms the existing approach of Diaz-Chito et al. (2016) on a DrivFace dataset.

As can be seen in [Table 5](#), the time efficiency of Naïve Bayes for each approach using ASM or using BoRMaN is the highest as compared to the time efficiency of rest of the classifiers as Naïve Bayes takes least amount of time. Its efficiency is due to the fact that Naïve Bayes classifier builds a simple generative model using conditional independence assumption that leads to quicker convergence as compared to other classifiers [\[47\]](#). Unlike percentage accuracy of Neural Network that is highest for Approach 4: RI using ASM, the time efficiency of Neural Network is the lowest as they are computationally intensive to train.

The average time efficiency for Naïve Bayes is also highest as compared to the average time efficiency of rest of the classifiers. The second and third best average time efficiencies are shown by Ningi and J48. It should be noted that the J48 has also shown third best percentage accuracy for Approach 4: RI using ASM as can be observed in [Table 3](#).

Table 5. Results of Approach 1: Ratios of Facial Components (RFC), Approach 2: Ratios of Facial Components and Areas of Triangles (RFCA), Approach 3: Ratios of Facial Components, Areas of Triangles and Angles (RFCAA) and Approach 4: Ratios of Facial Components, Areas of Triangles, Angles and Interpolation using Motion Vectors (RI) in terms of total time (training and testing) with training (64): testing (36) percentage split on a self-generated data set (FT-UMT dataset).

Classifier Name	Time for Approach 1		Time for Approach 2		Time for Approach 3		Time for Approach 4		Average Time
	ASM	BoR MaN	ASM	BoR MaN	ASM	BoR MaN	ASM	BoR MaN	
Naïve Bayes	0	0.01	0.01	0.01	0.02	0.02	0.03	0.03	0.02
SVM	0.14	0.18	0.07	0.1	0.17	0.06	0.33	0.19	0.16
Adaboost	0.06	0.09	0.08	0.11	0.3	1.11	0.38	0.44	0.32
Decision Table	0.07	0.12	0.13	0.13	0.57	0.85	0.81	1.06	0.47
Neural Network	2.2	2.52	3.02	3.51	19.71	22.85	32.02	31.58	14.68
J48	0.03	0.05	0.04	0.05	0.18	0.27	0.2	0.27	0.14
Nnge	0.04	0.07	0.05	0.07	0.13	0.2	0.19	0.25	0.13
Average Time	0.36	0.43	0.49	0.57	3.01	3.62	4.85	4.83	

It can also be observed in [Table 5](#) that the average time efficiency of the ASM is better than the average time efficiency of BoRMaN in most of the cases. The possible reasons include BoRMaN uses Boosted SVR (i.e., Ada boost regression and subsequent cross-validation using SVRs) coupled with Markov Networks, however, in case of ASM

faces are represented by a set of fiducial points, which are found with feature detection methods [46]. Furthermore, ASM uses PCA (for regularizing the configuration shape variations for controlled face shape deformations) that makes it time efficient as compared to BoRMaN.

It should also be observed in **Table 5** that average time efficiency of Approach 1: RFC is highest as compared to the average time efficiency of rest of the approaches. As we move from Approach 1: RFC to Approach 4: RI, the average time efficiency starts decreasing. The average time efficiency of Approach 4: RI is the lowest. The reason of the decrease in average time efficiency is the increase in number of features as we move from Approach 1: RFC to Approach 3: RFCAA. Furthermore, Approach 4: RI uses the same set of features as that of Approach 3: RFCAA but it also includes the use of interpolation technique using motion vectors to find the missing features of non-frontal faces (due to high rotation angle of driver's head).

6. Conclusion and Future Work

Distraction detection systems can avoid serious injuries or mishaps while driving. Our proposed work relies on a single camera mounted on a car dashboard and uses feature based learning system to detect driver's distraction due to head rotation around y-axis (horizontal axial rotation, i.e., right to left rotation of head). These features include angles, triangles, ratios of length to width of eyes and lips, ratios of angles and ratios of triangles of facial components while the learning systems used were Neural Network, Naïve Bayes, Support Vector Machine (SVM), Adaboost, Decision Table, J48 Decision Tree and Nearest-Neighbor-like Algorithm using Non-Nested Generalized Exemplars (NNge). It can be concluded that average percentage accuracy of Approach 4: 'Ratios of Facial Components, Areas of Triangles, Angles and Interpolation using Motion Vectors' using ASM is significantly better than rest of the approaches while average percentage accuracy of Neural Network classifier is better than the rest of the classifiers. Moreover, ASM performance in terms of percentage accuracy is better than BoRMaN in most of the cases. The study shows that the average time efficiency of Naïve Bayes is highest as compared to the rest of the classifiers.

This work can be extended to further improve the percentage accuracy and time efficiency of detecting the distraction caused due to driver's head rotation around y-axis (horizontal axial rotation, i.e., right to left rotation of the head). Another extension could be the detection of distraction caused by head rotation around x-axis (sagittal flexion and extension, i.e., forward to backward movement of the neck) and head rotation around z-axis (a frontal lateral bending, i.e., right to left bending of the neck). Furthermore, the detection of head movement resulted from combination of rotations around x-axis, y-axis and z-axis may further be investigated while also maintaining the time efficiency of the system. A distraction detection system should demonstrate invariance to variety of image-changing factors including projective geometry, camera distortion, multi-source non-Lambertian lighting as well as facial expression, movement of facial muscles and jaw, biological appearance, use of cell phone, glasses and hats. The fiducial facial point detection algorithms including ASM and BoRMaN can be optimized for i^{th} Frame by incorporating the temporal information gained from Frame 1, Frame 2, ..., Frame (i-1). The proposed system should further be tested on fiducial facial point detection algorithms other than ASM and BoRMaN. The addition of night vision functionality to the proposed system can be an important feature of the system. Moreover, the detection of harmful facial expression of driver including anger,

fatigue, sad and fearful in the proposed system might bring fruitful results in reducing the car accidents and ensuring the safety of the driver.

Acknowledgements

We would like to thank Khawaja Ubaid Ur Rehman, Wasiq Maqsood, Junaid Jabbar Faizi, Hassaam Zahid and Muhammad Maaz Aslam who volunteered for data generation for this work. We are grateful to all the anonymous reviewers for their useful comments. This work was supported by the National ICT R&D under grant no. NICTDF/NGIRI/2013-14/Crops/2; and University of Management & Technology, Lahore, Pakistan.

References

- [1] The Dangers of Texting while Driving. (2014, 12 8). Retrieved from Federal Communications Commission: [Article \(CrossRef Link\)](#).
- [2] Oviedo-Trespalacios, Oscar, Md Mazharul Haque, Mark King, and Simon Washington, "Understanding the impacts of mobile phone distraction on driving performance: A systematic review," *Transp. Research Emerg. Tech*, vol.72, pp. 360-380, November, 2016. [Article \(CrossRef Link\)](#).
- [3] Rohl, Austin, Sven Eriksson, and David Metcalf, "Evaluating the effectiveness of a front windshield sticker reminder in reducing texting while driving in young adults," *Cureus*, vol. 8, no.7, July, 2016. [Article \(CrossRef Link\)](#).
- [4] L.E.Levine, B.M. Waite and L.L.Bowman, "Mobile media use, multitasking and distractibility," *Intern. J. Cyber Behav., Psych*, vol. 2, no. 3, pp. 15-29, July, 2012. [Article \(CrossRef Link\)](#).
- [5] Vicente, Francisco, Zehua Huang, XuehanXiong, Fernando De la Torre, Wende Zhang and Dan Levi, "D.Levi, Driver gaze tracking and eyes off the road detection system," *IEEE Trans. Intell. Transp. Sys*, vol.16, no.4, pp. 2004-2027, August, 2015. [Article \(CrossRef Link\)](#).
- [6] Mittal, Ajay, Kanika Kumar, Sarina Dhamija, and Manvjeet Kaur, "Head movement-based driver drowsiness detection: A review of state-of-art techniques," in *Proc. of the IEEE International Cone on Engineering and Technology*, pp. 903-908, March, 2016. [Article \(CrossRef Link\)](#).
- [7] Yin, Bao-Cai, Xiao Fan, and Yan-Feng Sun, "Multiscale dynamic features based driver fatigue detection," *J. Pattern RecogArtif. Intell* vol. 23, no.3 pp.575-589, May, 2009. [Article \(CrossRef Link\)](#).
- [8] Liang, Yulan and J.D.Lee, "A hybrid Bayesian Network approach to detect driver cognitive distraction," *Trans. Research Emer. Tech*. vol. 38, pp.146-155, January, 2014. [Article \(CrossRef Link\)](#).
- [9] Murphy-Chutorian, Erik, AnupDoshi, and Mohan Manubhai Trivedi, "Head pose estimation for driver assistance systems: A robust algorithm and experimental evaluation," in *Proc of 10th IEEE Conf on Intelligent Transportation Systems*, pp. 709-714, September, 2007. [Article \(CrossRef Link\)](#).
- [10] D'Orazio, Tiziana, Marco Leo, Cataldo Guaragnella, and Arcangelo Distanto, "A visual approach for drive inattention detection," *Pattern Recog. Vol.* 40, no. 8, pp. 2341-2355, September, 2007. [Article \(CrossRef Link\)](#).
- [11] Hansen, Dan Witzner, and Qiang Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *IEEE Trans. Pattern Anal. Mach. Intel.* Vol. 32, no. 3, pp. 478-500, March, 2010. [Article \(CrossRef Link\)](#).

- [12] Valenti, Roberto, Nicu Sebe, and Theo Gevers, "Combining head pose and eye location information for gaze estimation," *IEEE Trans. Img. Proc.*, vol. 21, no.2, pp. 802-815, Feb, 2012. [Article \(CrossRef Link\)](#).
- [13] Tawari, Ashish, Sayanan Sivaraman, Mohan Manubhai Trivedi, Trevor Shannon, and Mario Toppelhofer, "Looking-in and looking-out vision for urban intelligent assistance: Estimation of driver attentive state and dynamic surround for safe merging and braking," in *Proc of the IEEE International Conf on Intelligent Vehicles Symposium Proc*, pp. 115-120, June, 2014. [Article \(CrossRef Link\)](#).
- [14] Sigari, Mohamad-Hoseyn, Muhammad-Reza Pourshahabi, Mohsen Soryani, and Mahmood Fathy, "A Review on Driver Face Monitoring Systems for Fatigue and Distraction Detection," *Int. J. Adv.Sci.Tech*, vol. 64, pp.73-100, 2014. [Article \(CrossRef Link\)](#).
- [15] Murphy-Chutorian, Erik, and Mohan Manubhai Trivedi, "Head pose estimation in computer vision," *A survey, IEEE Trans. Pattern Anal. Mach. Intel.*, vol. 31, no.4, pp. 607-626, April, 2009. [Article \(CrossRef Link\)](#).
- [16] Lee, Dongwook, Seungwon Oh, Seongkook Heo, and Minsoo Hahn, "Drowsy driving detection based on the driver's head movement using infrared sensors," in *Proc. of Symposium of the International Conference on Universal Communication*, pp. 231-236, December, 2008. [Article \(CrossRef Link\)](#).
- [17] Sathyanarayana, Amardeep, Sandhya Nageswaren, Hassan Ghasemzadeh, Roozbeh Jafari, and John HL Hansen, "R.Jafari, John HL Hansen Body sensor networks for driver distraction identification," in *Proc. of the of IEEE International Conf on Vehicular Electronics and Safety*, pp. 120-125, September 2008. [Article \(CrossRef Link\)](#).
- [18] Zeng, Jinhua, Y.Sun and L. Jiang, "Driver distraction detection and identity recognition in real-time," in *Proc., of the IEEE International Second WRI Global Congress on Intelligent Systems*, pp. 43-46, December, 2010. [Article \(CrossRef Link\)](#).
- [19] Hirayama, Takatsugu, Kenji Mase, and Kazuya Takeda, "Detection of driver distraction based on temporal relationship between eye-gaze and peripheral vehicle behavior," in *Proc. of the IEEE International Conf on Intelligent Transportation Systems*, pp. 870-875, September 2012. [Article \(CrossRef Link\)](#).
- [20] Choi, In-.Ho and Yong-Guk Kim, "Head pose and gaze direction tracking for detecting a drowsy driver," in *Proc. of the IEEE International conf on Big Data and Smart Computing*, pp. 241-244, January 2014. [Article \(CrossRef Link\)](#).
- [21] Dongare, Harshada and Sanjeevani Shah, "Eye Gaze Tracking and Eyes off the Road Detection for Traffic Safety on Raspberry Pi," *Intern. J. Innov. Research Elect. Instr. Contr. Engr*, vol.6, no.4, pp.2321-2004, June, 2006. [Article \(CrossRef Link\)](#).
- [22] Liu, Tianchi, Y.Yang, G.B.Huang, Y.K.Yeo and Z.Lin, "Driver distraction detection using semi-supervised machine learning," *IEEE Trans. Intell. Transp. Sys.* vol.17, no.4, pp.1108-1120, 2016. [Article \(CrossRef Link\)](#).
- [23] Azman, Afizan, Qinggang Meng, Eran A. Edirisinghe, and Hartini Azman. "Non-intrusive physiological measurement for driver cognitive distraction detection: Eye and mouth movements," *International Journal of Advanced Computer Science*, vol. 1, no. 3, pp. 92-99, 2011. [Article \(CrossRef Link\)](#).
- [24] Bergasa, Luis Miguel, Jesús Nuevo, Miguel A. Sotelo, Rafael Barea, and María Elena Lopez, "Real-time system for monitoring driver vigilance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 1, pp. 63-77, 2006. [Article \(CrossRef Link\)](#).
- [25] Rongben, Wang, Guo Lie, Tong Bingliang, and Jin Lisheng, "Monitoring mouth movement for driver fatigue or distraction with one camera," in *Proc. of the 7th International Conference on Intelligent Transportation Systems. Proceedings. IEEE*, pp. 314-319, 2004. [Article \(CrossRef Link\)](#).

- [26] Tang, Fangqi, and Benzai Deng, "Facial expression recognition using AAM and local facial features," in *Proc. of 3rd Conference on Natural Computation, ICNC. IEEE*, vol. 2, pp. 632-635, 2007. [Article \(CrossRef Link\)](#).
- [27] Diaz-Chito, Katerine, Aura Hernández-Sabaté, and Antonio M. López, "A reduced feature set for driver head pose estimation," *Applied Soft Computing*, vol. 45, pp. 98-107, 2016. [Article \(CrossRef Link\)](#).
- [28] Ma, Bingpeng, Rui Huang, and Lei Qin, "VoD: a novel image representation for head yaw estimation," *Neurocomputing*, vol. 148, pp. 455-466, 2015. [Article \(CrossRef Link\)](#).
- [29] Dongare, Harshada, and Sanjeevani Shah, "Eye Gaze Tracking and Eyes off the Road Detection for Traffic Safety on Raspberry Pi," *Eye 4*, no. 6, 2016. [Article \(CrossRef Link\)](#).
- [30] Billah, Tashrif, and SM Mahbubur Rahman, "Tracking-based detection of driving distraction from vehicular interior video," in *Proc. of 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 423-428. IEEE, 2016. [Article \(CrossRef Link\)](#).
- [31] Cootes, Tim, E. R. Baldock, and J. Graham, *Image Processing and Analysis*, Oxford University Press, 2000. [Article \(CrossRef Link\)](#).
- [32] Valstar, Michel, Brais Martinez, Xavier Binefa, and Maja Pantic, "Facial point detection using boosted regression and graph models," in *Proc. of IEEE Conf of Computer Vision and Pattern Recognition*, pp. 2729-2736, June, 2010. [Article \(CrossRef Link\)](#).
- [33] Sun, Yi, Xiaogang Wang, and Xiaoou Tang, "Deep learning face representation from predicting 10,000 classes," in *Proc. of the IEEE International Conf on Computer Vision and Pattern Recognition*, pp. 1891-1898, 2014. [Article \(CrossRef Link\)](#).
- [34] Taigman, Yaniv, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf, "Deepface Closing the gap to human-level performance in face verification," in *Proc. of the IEEE International Conf on Computer Vision and Pattern Recognition*, pp. 1701-1708, 2014. [Article \(CrossRef Link\)](#).
- [35] Li, Haoxiang, Zhe Lin, Xiaohui Shen, Jonathan Brandt, and Gang Hua, "A convolutional neural network cascade for face detection," in *Proc. of the IEEE International Conf on Computer Vision and Pattern Recognition*, pp. 5325-5334, 2015. [Article \(CrossRef Link\)](#).
- [36] Curran, K., X. Li, N. and Mc Caughley, "Neural network face detection," *J. Img. Sci*, vol.53, no.2, pp.105-115, June, 2005. [Article \(CrossRef Link\)](#).
- [37] Sun, Yi, Ding Liang, Xiaogang Wang, and Xiaoou Tang, "Deepid3: Face recognition with very deep neural networks," in *Proc. of the International conf arXiv preprint arXiv*, vol. 53, no.2, pp.1502-00873, June, 2015.
- [38] Kim, Gyu-jin, Taeki An, and Moonhyun Kim, "Estimation of Crowd Density in Public Areas Based on Neural Network," *KSII Transactions on Internet & Information Systems*, vol. 6, no. 9, 2012. [Article \(CrossRef Link\)](#).
- [39] Liang, Xiaolin, Hao Zhang, and T. Aaron Gulliver, "Energy Detector based Time of Arrival Estimation using a Neural Network with Millimeter Wave Signals," *KSII Transactions on Internet & Information Systems*, vol. 10, no. 7, 2016. [Article \(CrossRef Link\)](#).
- [40] Kerhet Aliaksei, Mirco Raffetto, Andrea Boni, and Andrea Massa, "A SVM-based approach to microwave breast cancer detection," *Engineering Applications of Artificial Intelligence*, vol. 19, no. 7, pp. 807-818, October, 2006. [Article \(CrossRef Link\)](#).
- [41] Wang, Xiang-Yang, Hong-Ying Yang, Yong-Wei Li, Wei-Yi Li, and Jing-Wei Chen, "A new SVM-based active feedback scheme for image retrieval," *Engineering Applications of Artificial Intelligence*, vol.37 pp. 43-53, January 2015. [Article \(CrossRef Link\)](#).
- [42] Parvin, Hamid, Miresmaeil Mirnabi Baboli, and Hamid Alinejad-Rokny, "Proposing a classifier ensemble framework based on classifier selection and decision tree," *Engineering Applications of Artificial Intelligence*, vol. 37, pp.34-42, January, 2015. [Article \(CrossRef Link\)](#).

- [43] Sarker, Mostafa Kamal, Sook Yoon, and Dong Sun Park, "A Fast and Robust License Plate Detection Algorithm Based on Two-stage Cascade AdaBoost," *KSII Transactions on Internet & Information Systems*, vol. 8, no. 10, 2014. [Article \(CrossRef Link\)](#).
- [44] Jia Yangqing, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell, "S.Guadarrama, T.Darrell, Caffe: Convolutional architecture for fast feature embedding," in *Proc. of the 22nd ACM international conf on Multimedia*, pp. 675-678, November, 2014. [Article \(CrossRef Link\)](#).
- [45] Smith and Steven, *Digital signal processing: a practical guide for engineers and scientists*, Newnes Press, 2013.
- [46] Çeliktutan, Oya, Sezer Ulukaya, and Bülent Sankur, "A comparative study of face landmarking techniques," *EURASIP Journal on Image and Video Processing*, vol. 13, no. 1, 2013. [Article \(CrossRef Link\)](#).
- [47] Caruana, R. and Niculescu-Mizil, A., "An empirical comparison of supervised learning algorithms," in *Proc. of the 23rd Int. conf. on Machine learning ACM*, pp. 161-168, June, 2006. [Article \(CrossRef Link\)](#).



Syed Farooq Ali did his BS from NUCES- FAST, Lahore, Pakistan and later earned an MS degree from LUMS, Lahore, Pakistan with Dean's Honor List. During his stay in MS, he was on a LUMS fellowship. He also completed his MS from Ohio State University, USA. Currently, he is a PhD candidate at University of Management and Technology, UMT, Lahore, Pakistan. His research interest includes computer vision, digital image processing and medical imaging.



Dr. Malik Tahir Hassan earned his MS and PhD degrees in Computer Science from Lahore University of Management Sciences (LUMS), Lahore, Pakistan. He has worked at Gwangju Institute of Science and Tehcnology (GIST), Gwangju, South Korea, as a Post-Doc fellow. Currently, he is working as an Assistant Professor at School of Systems and Technology (SST) at University of Management and Technology (UMT), Lahore, Pakistan. His research interests include pattern recognition, digital image processing, text mining, recommender systems and autonomic computing.