International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-5, June 2019

Driving Pattern using Sensor Generated Data

Rohan Kanade, Shweta Gawahale, Shivraj Jadhav, Reeba Benny, Ajit Kumar Shitole

Abstract: Rash driving has become a trend in today's world. Speed limit violation has largely increased the rates of mishaps happening. The drivers are carelessly ignoring the road safety rules that cause a threat to the surrounding society by their rash driving. A reliable method to take care of this kind of ignorant behavior and to reduce the accident rate, is being introduced in this paper, which is done by monitoring the driver's driving style in their learning phase. This method intends to check the eligibility of the driver to gain a license by classifying them in groups of a SAFE or UNSAFE DRIVER which can help to filter harsh drivers to a large extend. To be eligible for the license the driver should pass the test by overcoming all the constraints set by the system. The introduced method targets in reducing the number of careless drivers which may naturally help with lowering the accidents caused by improper or harsh driving. It concludes with the discussions of the challenges, results and future works of the proposed technique.

Index Terms: Accelerometer, C4.5, Decision Tree, GPS Pattern Recognition, Smartphone.

I. INTRODUCTION

The increasing rate of accidents and the damage caused in the society due to the careless behavior of the drivers can be reduced to a great extent by the system introduced in this paper. This system assess the driving styles of the driver in the learning phase itself before gaining a permanent license which may help to distinctly classify safe and unsafe drivers. This method helps to monitor the driving skills of the driver using various smartphone sensors. The categorization will be performed by analyzing the information that is generated by the accelerometer sensor and GPS present in the smartphones. The intention of this method is to help the Driving License departments to provide a license to only those drivers that are classified as safe drivers. This classification of the drivers will be done with the help of a fixed system which will evaluate each applicant with equal importance and each will be given a fair chance, as the system will be automatic it will help in eliminating any illegal activities. The classification of an unsafe driver will be observed over many factors such as driving over the speed limit, swift turns, harsh breaking, sudden acceleration, sudden breaking etc. This information will be captured continuously in real time. The method proposed in this paper can be effectively used to filter harsh drivers and to provide a safe environment for other drivers as well. The system will capture and analyze the data in real time hence the man power and time required to generate a license will be reduced. It can help the Driving License departments to evaluate the drivers over a period of time rather than a single session which will be more effective and promising [6].

Manuscript published on 30 June 2019.
***** Correspondence Author (s)
Rohan Kanade*, Computer Engineering Department BE, International Institute of Information Technology, Hinjawadi, India.
Shweta Gawahale, Computer Engineering Department BE, International Institute of Information Technology, Hinjawadi, India.
Reeba Benny, Computer Engineering Department BE, International Institute of Information Technology, Hinjawadi, India.
Shivraj Jadhav, Computer Engineering Department BE, International Institute of Information Technology, Hinjawadi, India.
Prof. Ajitkumar Shitole, Associate Professor, Computer Engineering Department, International Institute of Information Technology, Hinjawadi, India.

II. MOTIVATION

The motivation behind introducing a method that could potentially classify safe drivers is that it could help decrease the risks of accidents, and enhance the security of the pedestrians. Rash drivers can cause nuisance on roads by driving rashly and disobeying the traffic rules. Therefore a method to give a license by checking the driver's eligibility as a safe driver is being introduced in this paper.

III. GOALS

This system was built to achieve the goals mentioned below:

- ***Frauds***

Reduce the frauds that are happening at the Driving License departments.

- ***Time Consumed***

Reducing the time consumed by the current license generating system.

- ***Driver Skill***

As the driver will be evaluated over the training period he/she will try to keep their performance good in order to get the Permanent License which will automatically improve the skills of the driver.

- ***Insurance Companies***

Help the Insurance companies get an estimate of the skills of the driver so that they can use the data to decide whether or not to give the driver Insurance Policy.

- ***Mishaps***

To reduce the mishaps happening due to lack of driving skills of the drivers, simultaneously increasing the safety of the drivers as well as the pedestrians.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

IV. SMARTPHONE SENSORS

This section gives a brief description about the sensors present in the smartphone that are being used in this system:

- **Accelerometer**

It is an electromechanical device that is present in the smartphone out of the box. It is used to measure the acceleration forces. These forces may be static or dynamic in nature, for e.g. gravitational force is static and forces caused by moving or vibrating are dynamic in nature. It measures Proper Acceleration, i.e. acceleration of the body in its own rest frame [1].

- **Global Positioning System**

GPS is a satellite based Navigation system, it gives us the positions on the map where you have been using the values of longitude, latitude.

- **Camera**

Camera is a device that is used to capture images and Videos, it is an important device in Face Recognition applications.

V. RELATED ALGORITHMS

A. *Haversine*

Haversine is a formula that is used to calculate the great-circle distance between two points on a sphere, given their latitude and longitude. It is a special case of more general formula in spherical trigonometry. It calculates the shortest distance between the two points on the sphere [3].

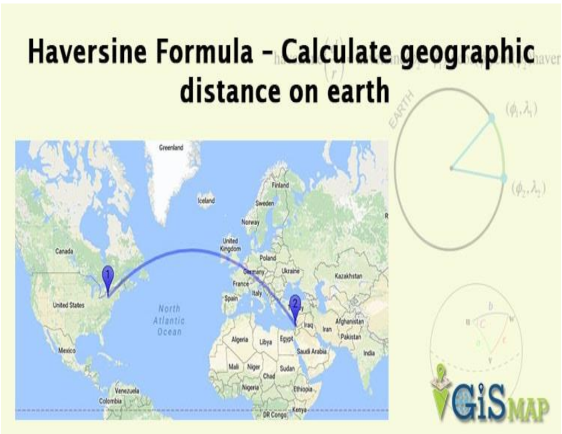

Fig 1: Haversine Co-ordinates capturing (src :: igismap.com)

Fig 1 shows how the geographical positions are captured on the great circle of the earth, these points are used to calculate the distance between them using the Haversine Formula.

Law of Haversine:

To derive law of Haversine one needs to start the calculation with spherical law of cosine [9].

i.e.

$\cos a = \cos b * \cos c + \sin b * \sin c * \cos A$

We can derive Haversine formula to calculate distance between two points as:

$$a = \sin\hat{}2(\Delta latDifference\ /\ 2) + \cos(lat1).\cos(lat2).\sin\hat{}2(\Delta lonDifference\ /\ 2) \quad (1)$$

$$c = 2.atan2\left(\sqrt{a}, \sqrt{(1-a)}\right) \quad (2)$$

$$d = r.c \quad (3)$$

Where,

$\Delta latDifference = lat1 - lat2$ (Difference of latitudes)
$\Delta lonDifference = lon1 - lon2$ (Difference of longitudes)
R is radius of earth i.e. **6371 KM or 3961 miles** and **d** is the distance calculated between two points.

Haversine Formula:

The Haversine formula is used to for the calculation of the distance between two points on the Earth's surface specified as longitude and latitude [3]. The latitude value ranges from 0^0 to 90^0 on either sides of the equator. The longitude value ranges from 180^0 East and West from Prime Meridian [4].

$$d = 2R\sin^{-1}\left(\sqrt{\left(\sin^2\left(\frac{\varphi_2-\varphi_1}{2}\right) + \cos\varphi_1.\cos\varphi_2.\sin^2\left(\frac{\Psi_2-\Psi_1}{2}\right)\right)}\right) \quad (4)$$

D is the distance between two points on earth's surface with longitude and latitude (ψ, φ) and *R* is the radius of the Earth.

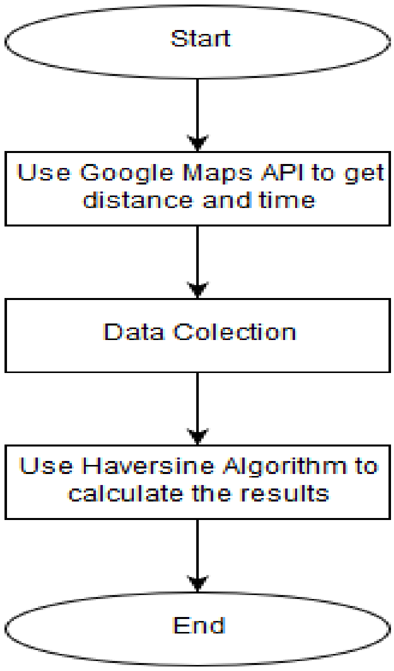

Fig 2 : Haversine Workflow Flowchart

B. *C 4.5*

C4.5 is an algorithm which is used to generate a decision tree. It is an extension of ID3 (Iterative Dichotomiser 3) algorithm. The decision trees that are generated by C4.5 can be used for classification of the data, and for this reason, it is often referred to as a statistical classifier.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

It is based on the Information gain ratio that is evaluated by entropy [7]. The Information gain ratio measure helps in selecting the test features at each node in the tree. Such a measure is referred to as a feature or attributes selection measure. The attribute is chosen as the test feature for the current node if its Information Gain Ratio is highest [10]. Let D be a set such that D= (D_1... D_j) data instances. Suppose the class label attribute has m distinct values each defining a distinct class, C_i (for I = 1... m). Let D_j be the number of samples of D in class C_i. The expected information needed to classify a given sample can given by

$$SplitInfo\ A(D) = -\sum(|D_j/D|) * \log((D_j/D)) \quad (5)$$

$$GainRatio(A) = {Gain(A)}/{SplitInfo\ A(D)} \quad (6)$$

Where,

$Info(D) = -\sum P_i \log_2(P_i)$ And,

$$Info\ A(D) = -\sum\left(\frac{|D_j|}{|D|}\right) * Info(D_j)$$

Where P_i = probability of distinct class C_i,

D = Data Set

A = Sub attribute from attribute

$\left(\frac{|D_j|}{|D|}\right)$ = act as weight of j^{th} partition

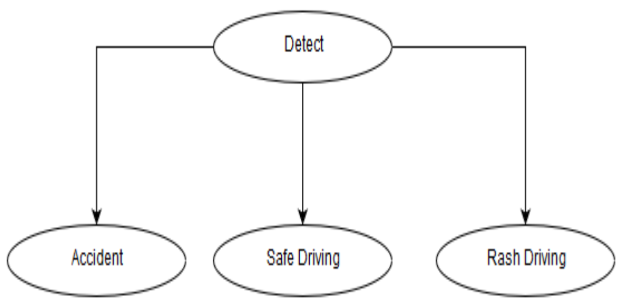

Fig 3 : Detection Graph

C. PCA

Dimensionality reduction is one of the preprocessing steps in many machine learning applications and it is used to transform the features into a lower dimension space [5]. PCA is a statistical approach for reducing the number of variables used in face recognition. In PCA, every image in the training set is represented as a linear combination of weighted eigenvectors also called as Eigen faces. These eigenvectors are obtained from the covariance matrix of the training image set. The weights are found out after selecting the most relevant Eigen faces. To recognize the test image it is spanned onto the Eigen faces subspace and then classification is done using the Euclidean distance measure [8].

Face Recognition Process

The Recognition Process involves two steps:

1) Initialization process
2) Recognition process

Initialization Process

The Initialization process consists of the following operations:

1) Acquire training set of images.
2) Calculate the Eigenfaces from the training set, keeping only the highest eigenvalues. These M images define the face space. As new faces are encountered, the Eigen faces can be updated or recalculated [11].
3) Calculate distribution in this M-dimensional space for each known person by projecting his or her face images onto this face-space.

These operations can be performed every now and then, whenever free excess operational capacity is available. These steps can be cached and used later to reduce the re-initialization overhead. Hence, the execution time is decreased and performance of the system increases.

Recognition Process

The Recognition process consists of the following operations:

1) Calculate a set of weights based by projecting the input image onto each Eigen face.
2) Determine if the image is a face by checking if the image is sufficiently close to a free space.
3) If it is a face, then classify the weight pattern as either a known person or as unknown person.

Eigen Face Algorithm

Let an image Γ(x, y) be a two dimensional m x n array of intensity values. In this project, we have used a set of 180 x 200 pixels image. An image may also be considered as a vector of dimension m x n, so that a typical image of size 180 x 200 becomes a vector of dimension 36,000 or equivalently a point in a 36,000 dimensional space.

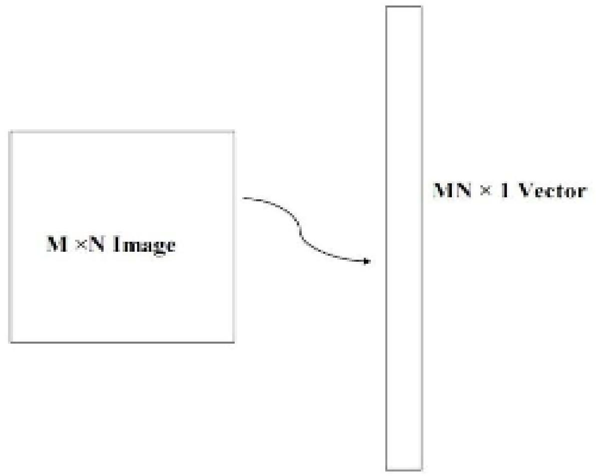

Fig 4 : PCA Vector Transformation

Step1: Prepare the training set.

Obtain face images I1, I2, I3, I4 . . . IM (training set). (The face images are required to be centered and must be of the same size.)

Step 2: Prepare the data set.

Each face image Ii in the database is transformed into a vector and placed in the training set Profile Picture. Each image is transformed into a vector of size mn × 1 and placed into the set.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

$S = \{\Gamma 1, \Gamma_2, \Gamma_3, \Gamma_4 \ldots \Gamma_m\}$

Step 3: Compute Average Face Vector

$$\Psi = \frac{1}{M} \sum \tau_n$$

Step 4: Subtract Average Face Vector from the original faces.

$$\Phi_i = \tau_i - \Psi$$

Step 5: Calculate the Covariance matrix

$$C = \frac{1}{M} \sum_{n=1}^{M} \Phi_n \Phi_n^T$$

$$= AA^T$$

Where,

$$A = [\Phi_{1,\ldots\ldots} \Phi_m]$$

Step 6: Calculate the Eigen values and Eigen vectors of the Covariance Matrix.

In this step Eigen faces (u_i) and Eigen vectors (v_i) of the covariance matrix are calculated. The Eigen faces must be normalized so that they are unit vectors, i.e. of length 1.

Step 6.1: Consider the matrix

$L = A^T A$ (M x M matrix)

Step 6.2: Compute Eigen vectors v_i of $L = A^T A$

Relationship between u_i and v_i:

$$A^T A\, v_i = \mu_i v_i$$
$$A^T A v_i = \mu_i v_i$$
$$A^T A v_i = \mu_i A v_i$$
$$AA^T A v_i = \mu_i A v_i$$
$$C A v_i = \mu_i A v_i \text{ Where,}$$
$$C u_i = \mu_i A v_i \text{ Where } u_i = v_i$$

Thus, $C = AA^T$ and $L = A^T A$ have the same eigenvalues and their eigenvectors are related as follows:

$$u_i = A\, v_i$$

Note 1: $C = AA^T$ can have up to N^2 eigenvalues and eigenvectors.

Note 2: $L = A^T A$ can have up to M eigenvalues and eigenvectors.

Note 3: The M eigenvalues of $C = AA^T$ (along with their corresponding eigenvectors) correspond to the M largest eigenvalues of $L = A^T A$ (along with their corresponding eigenvectors).

Where v_i is an eigenvector of $L = A^T A$. From this simple proof we can see that is an eigenvector of $C = AA^T$. The M eigenvectors of $L = A^T A$ are used to find the M eigenvectors of C that form our Eigen face basis:

$$u_i = \sum_{i=1}^{M} v_i \Phi_i$$

Where, u_i are the Eigen faces

Step 7: Project training sample into Eigen space

The feature weight of the training images can be calculated as:

$$\omega_i = u_i^T (\Gamma_i - \psi)$$

Where, u_i is the i^{th} Eigenface and i=1, 2, 3 . . . . . . K.

The weight is obtained as above form a vector:

$$\Omega_i^T = [\omega_{1\ldots\ldots} \omega_k]$$

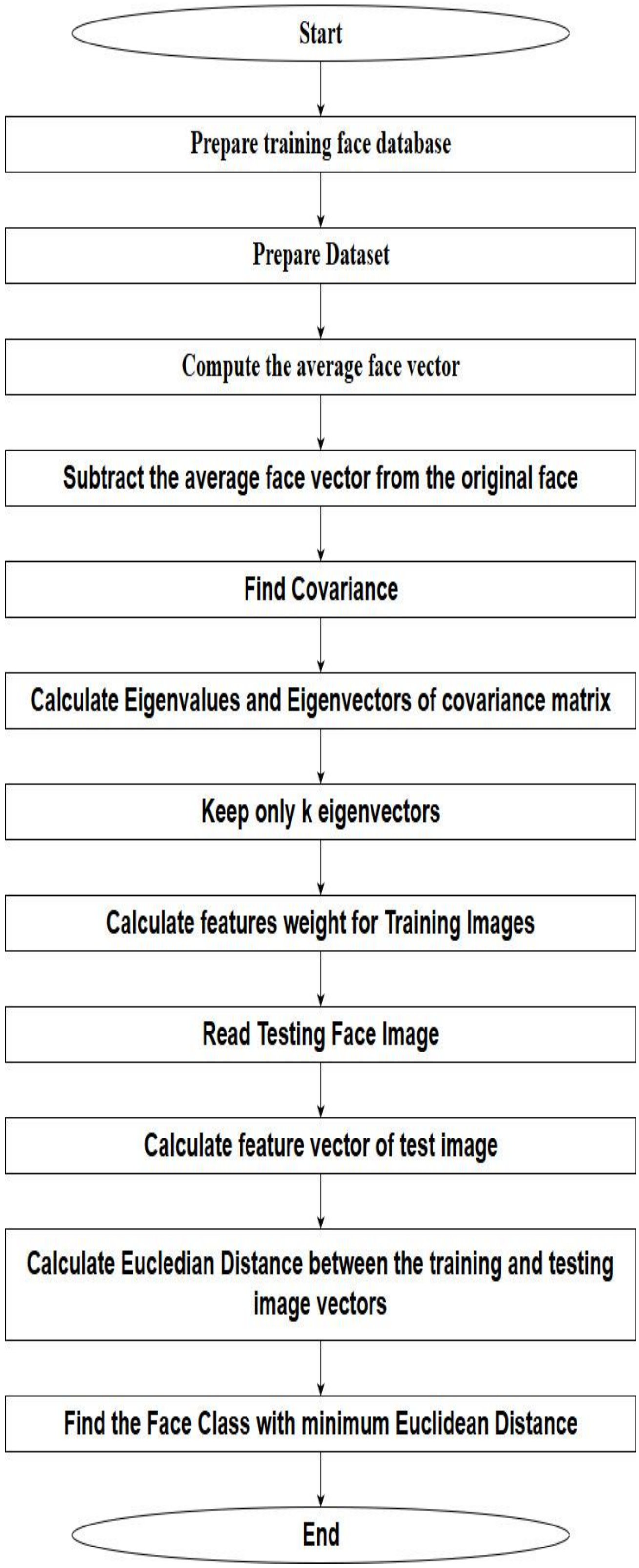

Fig 5: PCA Workflow Flowchart

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

VI. PROPOSED SYSTEM ACHITECTURE

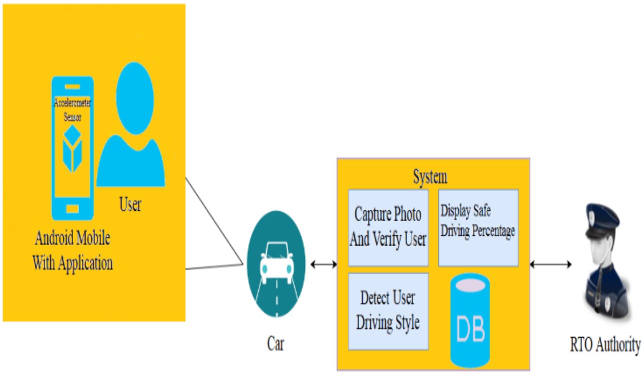

Fig 6: System Architecture

After learner's license test, the user will be given a unique login id and a password from the license department to login into their application to apply for a permanent license. After login in with the given credentials the user is verified as an authorized user with the data that is already present in the database. As soon as the user is known to be an authorized user they can apply for the permanent license. After applying for a permanent license the data of his driving styles will be captured whenever he drives and he needs to open the application to start. In this system the data of his driving styles is going to be taken with the help of accelerometer sensor which is present in the smartphones. The data which is captured is saved in the database to be evaluated later. Whenever the driver is driving, the system will also capture user's picture in every predefined minutes to check whether the data of driving styles being captured is of the same authorized user or not. To detect this image processing is done using Algorithm for facial parts detection to detect the faces in an image and locates the facial features in an image and is used for matching the captured images with the already stored image in the database, by comparing the images if the person driving does not match with the stored images, a message of unauthorized user will be shown. This will be done by using PCA algorithm. To make sure the assessment of the driving styles is completely of that person who is acquiring the permanent license, thus eliminating any unfair means. After the user has drove the standard distance which is required to be eligible and that is set by the RTO the user's data is send for further evaluation. The total distance travelled by the user is calculated by the Haversine formula. For the evaluation of the user the system compares the patterns of user's driving styles with the standard driving styles that are set by the RTO. This pattern matching is done by Decision Tree C 4.5 algorithm. Using these algorithms the authorized user is categorized as eligible driver for a permanent license or not. This categorization is done on the basis of whether he is a rash driver, if his driving styles matches with the patterns set as improper driving such as over speeding, swift turns, accidents etc. If not the user will be shown as eligible if he crosses the set eligibility percentage by the RTO. The RTO officer will analyze the user's data and confirm the permanent license generation based on the evaluation done by the system.

VII. RESULTS

A. *Admin System*

The Administrator is provided with a Web Application. The admin can login into the system using the given credentials so that he can access the data and perform various operations such as: New User Registration, Check Driving Result, Pass or Fail the applicant based on his Safe Driving percentage. Whenever the admin clicks on the Pass button an email is sent saying "Congratulations! You have passed the test", if admin clicks on Fail button an email is sent saying "Sorry! You have failed the test". The admin can also check if the user is the valid user or not, by submitting the applicants ID, after which an Image Processing algorithm is run in the background and the result is shown as in Fig 14, also the data regarding that user is displayed on the screen.

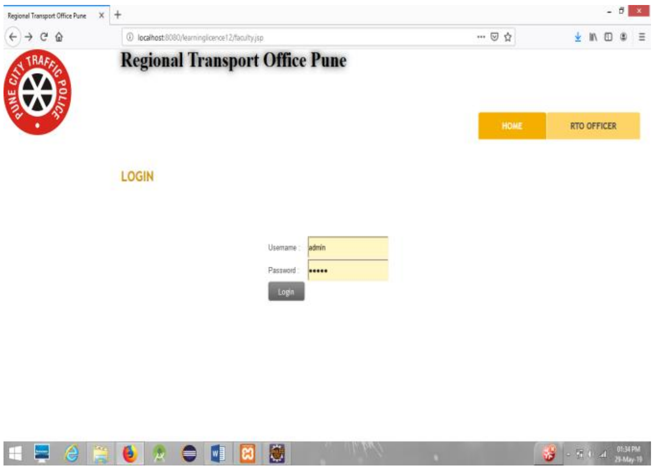

Fig 7: Admin Home Page

Fig 7 shows the Admin Home Page where the Administrator may login using his/her credentials.

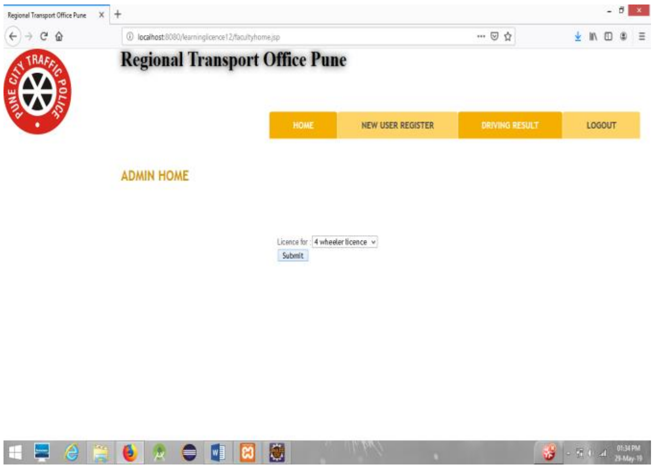

Fig 8: Admin Page where all operations are performed

Fig 8 shows the Admin Page which is redirected after the Admin logs in to the system here the admin may perform various operations such as New User Registration, Driving Result checking, Valid User check, Log off the system.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

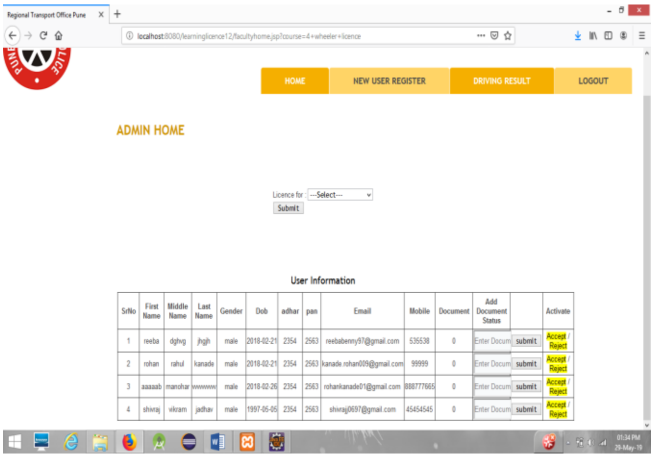

Fig 9: Driving result of the User based on User ID

Fig 9 shows the list of drivers that are registered under 4 Wheeler License category along with their Document status, if all the documents are correct the Admin may ACCEPT the application and the Username and Password is sent to the registered Email ID.

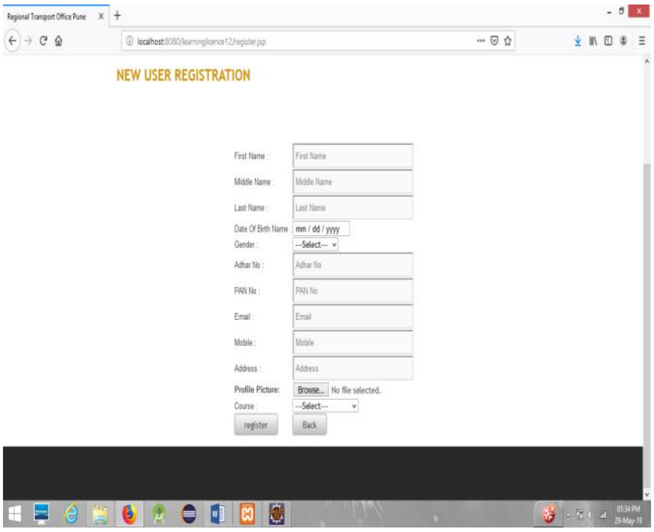

Fig 10: New User Registration Page

Fig 10 shows the page using which the Admin may register the user to the system after which the user is provided with an ID and Password.

B. *Driver Behaviour Analysis*

System continuously checks for patterns while driving based on live accelerometer readings from the accelerometer sensor of the phone which is mounted on the dash board of the driver's car. The accelerometer gives us three values: acceleration along x-axis, acceleration along y-axis, acceleration along z-axis [2]. By computing these values various constraints are checked and flags are set. Based on the flags the behavior is divided into three categories i.e. Safe Driving, Rash Driving, and Accident.

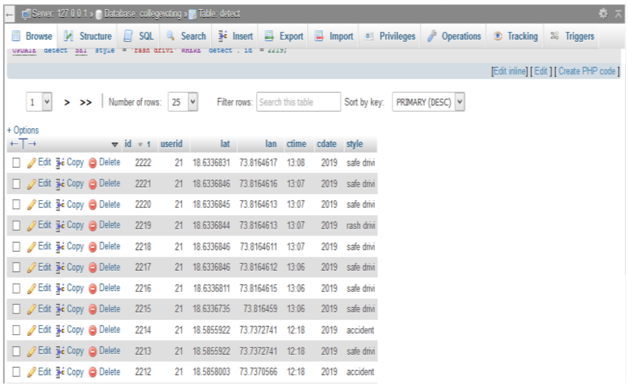

Fig 11: Driving pattern recorded in Database

Fig 11 shows how the data is recorded in the database after classification into either Accident, Safe Driving, Rash Driving. The database also stores the data about the location of the user after every interval alongwith the time and date the data was captured. The data about the location is used to calculate the distance of the journey, so that track of how much distance the driver has covered can be analysed.

C. *Distance of The Journey*

The location of the driver after every interval is captured and stored in the database. An android application has been provided to the user, which records all the data regarding the journey of a particular user. The location after pressing START and STOP button on the app is recorded and distance travelled between those positions is calculated and displayed.

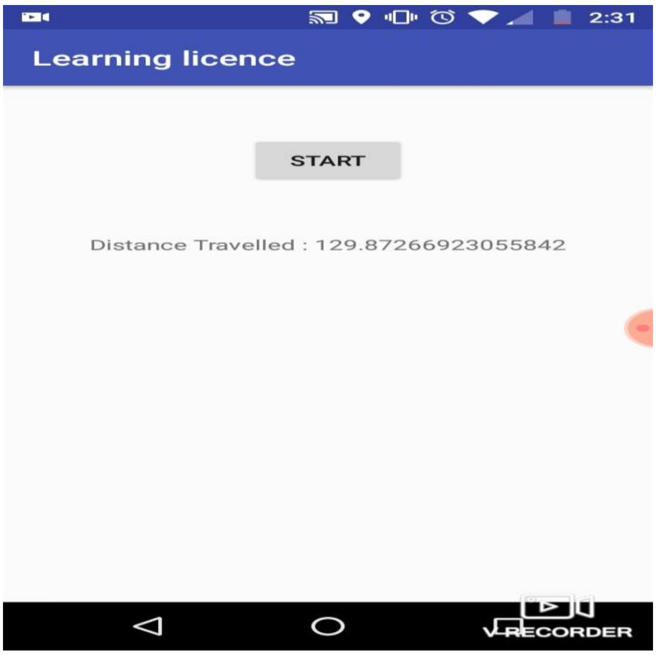

Fig 12: Application shows the distance travelled in meters.

Fig 12 shows the first page after the user logs into the system, the START button provided starts all the services that are required for evaluation of the applicant, Distance travelled in the journey is also displayed as in the fingure.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

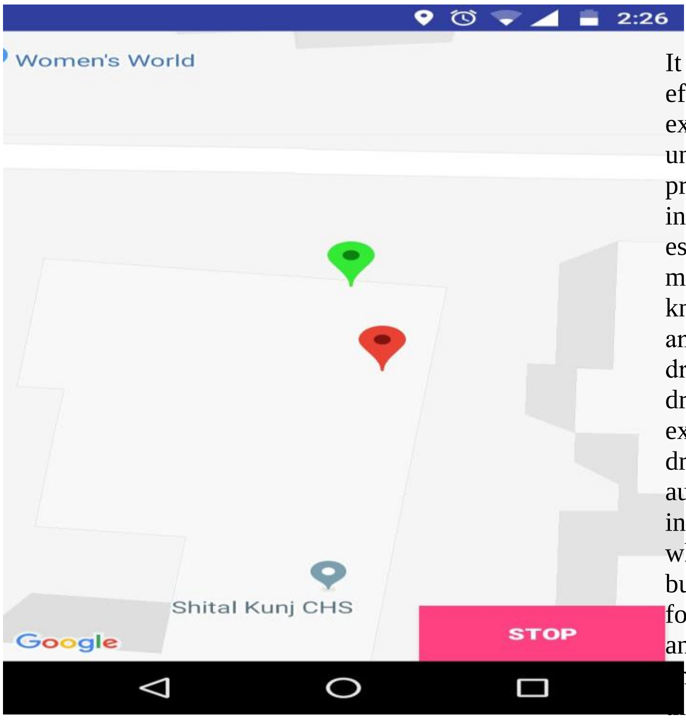

Fig 13: Final and Initial Location of the user.

Fig 13 shows the Initial Location and the Final Location of the journey. The Initial Location is marked using the Red marker and Final Location is marked using the Green marker.

D. *Driver Identification*

For driver identification and verification, Image Processing algorithm is used using Principal Component Analysis (PCA) in MATLAB. Images are captured in real time by the phone's front facing camera and stored on the server database over the duration of the test. These images are compared with the profile picture to find if it matches with the original profile picture of the driver. If the image matches then the application will generate a pop-up dialog box showing the Test Image and the Equivalent Image.

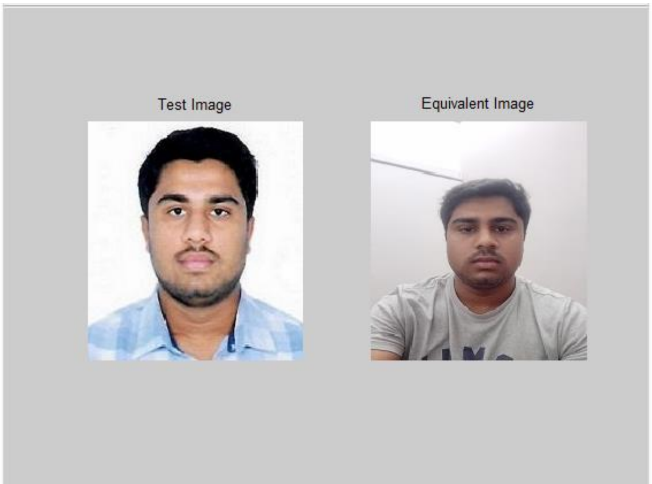

Fig 14: Images processed in Real Time

Fig 14 shows the result of the images matched after running an Image Processing module in the background, image in the left is the Profile Picture that is set during registration and on the right is the live captured image.

VIII. CONCLUSION

It can be concluded that the proposed system can be highly effective when it comes to the safety of people as there is an extremely high rate of accidents taking place due to unprofessional drivers driving in rash manner. Road safety problems being a major area of concern in the transport industry as the rate of people driving vehicles is increasing especially in developing countries. The proposed system makes sure that this great number of people driving vehicles know how to drive in a professional and safe manner to avoid any risk of mishaps. The existing system doesn't check the driver's ability to an extent where it can be sure about their driving skills, as it is simple enough to get a license from the existing system without completely evaluating the driver's driving behavior. The proposed system allows the RTO authority to force people to follow all RTO rules and to drive in a proper manner for a period, which checks not only whether the driver is able to drive the car in a proper manner but also if he is able to make decisions in traffic while following rules properly. Proposed system makes sure each and every permanent license holder user is perfect in driving nd is following traffic rules. This will completely eliminate e false means that people use to get a license giving each individual a fair chance to prove their eligibility to gain one. This system can also be used if a person wants to improve their driving skills to learn driving in a better manner.

REFERENCES

1. Nidhi Kalra and Divya Bansal, "Analysing Driver Behaviour using Smartphone Sensors: A Survey", International Journal of Electronic and Electrical Engineering, ISSN 0974-2174 Volume 7, Number 7 (2014), pp. 697-702.
2. Armir Bujari, Bogdan Licar, Claudio E. Palazzi, "Movement Pattern Recognition through Smartphone's Accelerometer",Universitàdegli Studi di Padova Padova, Italy.
3. Prof. Nitin R.Chopde1, Mr. Mangesh K. Nichat2," Landmark Based Shortest Path Detection by Using A* and Haversine Formula", International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 2, April 2013.
4. Vinayak Hedge, Aswathi T S, Sidharth R, "Student Residential Distance Calculation using Haversine Formulation and Visualization through GoogleMap for Admission Analysis", 2016 IEEE International Conference on Computational Intelligence and Computing Research.
5. Alaa Tharwat ,Electrical Department, Faculty of Engineering, Suez Canal University, Ismailia, Egypt "Principal component analysis-a tutorial" International Journal of Applied Pattern Recognition 3(3) (2016) 197-240
6. Reeba Benny, Shweta Gawahale, Shivraj Jadhav, Rohan Kanade, Prof. Ajitkumar Shitole "Drivers Assessment using Smartphones" International Journal of Multidisciplinary, Vol. 4, Journal No. 44945
7. Badr Hssina, Abdelkarim Merbouha, Hanane ezzikouri, Mohammed
8. Erritali, "A Comparative Study of Decision Tree ID3 and C4.5",IJACSA
9. Dhiraj K. Das, " Comparative Analysis of PCA and 2DPCA in Face Recognition", International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, Volume 2, Issue 1, January 2012
10. https://www.igismap.com/haversine-formula-calculate-geographic-distance-earth/, Accessed 11th June 2019
11. Seema Sharma, Jitendra Agrawal, Sanjeev Sharma, "Classification Through Machine Learning Technique: C4.5 Algorithm based on Various Entropies", International Journal of Computer Applications (0975-8887), Volume 82–No 16, November 2013
12. Lata, Prof & Kiran, Chandra & Tungathurthi, Bharadwaj & Ram, H & Hanmanth, Ram Mohan Rao & Govardhan, Dr & L P Reddy, Dr. (2009). Facial recognition using eigenfaces by PCA. SHORT PAPER International Journal of Recent Trends in Engineering. 1.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

AUTHORS PROFILE

Rohan Kanade is working as Cloud Engineer Intern at BlazeClan Technologies, Pune. He is pursuing Computer Science Engineering from I^{2}IT, Hinjawadi, Pune. He has avid interest in Machine Learning and open source clouds. He has published a paper "Drivers Assessment using Smartphones". He is also an ACM(Association for Computing Machinery) member.

Reeba Benny is pursuing Computer Science Engineering at I^{2}IT Hinjawadi, Pune. She has published "Drivers Assessment using Smartphones".

Shivraj Jadhav is pursuing Computer Science Engineering at I^{2}IT Hinjawadi, Pune. He has published "Drivers Assessment using Smartphones".

Shweta Gawahale is pursuing Computer Science Engineering at I^{2}IT Hinjawadi, Pune. She has published "Drivers Assessment using Smartphones".

Ajitkumar S. Shitole is pursuing Ph.D. in CSE from Amity University Mumbai. He has published more than 20 research papers in various International Journals/ Conferences. His area of interest is Data Mining, Machine Learning, and Algorithms. Currently he is working as an Associate Professor in Computer Engineering Department at I^{2}IT Hinjawadi, Pune.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)