FAST AND ACCURATE DETECTION OF ROAD TO CONTROL THE VEHICLE MOVEMENT USING HAAR LIKE FEATURES AND YOLO CLASSIFIER

J.Jenifer John
Assistant professor,
Department of Electronics and Communication Engineering,
Jayaraj Annapackiam CSI College of Engineering
johnjjesun@gmail.com

Dr.S.Allwin
Professor,
Department of Computer Science and Engineering,
Infant Jesus College of Engineering
allwinstephen@gmail.com

ABSTRACT

Autonomous driving is the most focused research problem in real world to ensure the transportation services to the people. In this research work, autonomous driving is focused by introducing the techniques which can ensure the uninterrupted vehicle movement in the road even with presence of obstructs. This is ensured by introducing the technique Fast and Accurate Road Detection System (FARDS). In the proposed research method, initially preprocessing is done on the traffic surveillance video to remove the noises by applying the Hybrid Median Filter method. And then background extraction is done by using markov random field process to extract the required objects. After extracting the required objects from the videos, Haar like Feature extraction is done to detect the objects present in the videos. Based on those feature road and objects detection is performed dynamically by applying Yolo classifier. This information is utilized to adjust the vehicle speed, acceleration and steering values to ensure the uninterrupted vehicle movements without any accidents. The overall evaluation of the research method is done in the matlab simulation environment from which it is proved that the proposed method FARDS leads to ensure the better outcome than the existing research techniques.

Keywords: Vehicle movement, haar like feature, Yolo classifier, moving objects, acceleration, steering

I. INTRODUCTION

Self-driving means the autonomous driving of a vehicle to a specific target in real traffic without the intervention of a human driver [1]. Such a vehicle gets its input data primarily from visual information sources that are also available to the driver. In the preliminary stages of autonomous driving, technology enhanced a driver's awareness by providing information that enabled him or her to make a decision and react quickly [2]. But when a vehicle reacts autonomously – without active intervention from a driver – through algorithms that force the vehicle to react in a specific way, one speaks of autonomous driving [3].

In the current discussion on vehicle authorization between the 'Automated driving' working group of the VDA [German Association of the Automotive Industry] [4] and the Federal Highway Research Institute (BASt) [5], German OEMs have agreed on the following basic stages of autonomous driving: The driver must continuously monitor the automatic functions and cannot perform any non-driving task [6]. This includes the driving assistance systems that Mercedes-Benz is offering as 'Intelligent Drive' in the new S, E, and C-Class models as well an in the new CLS [7].

Our research advances the knowledge and insights on autonomous-driving technologies set to transform mobility industries, technology, cities, and the environment, on a global scale [8]. As autonomous-vehicle (AV) technology progresses from needing driver assistance to having full autonomy, driverless cars are looking more likely to become a reality. With these come significant benefits, including increased personal safety, time saving for drivers, mobility for non drivers, decreased environmental harm, and reduced transportation costs [9]. It will also lead to drastic shifts in value chains, profit pools, and needed capabilities (for example, software expertise or cyber security) that could introduce entirely new industries within automotive.

While AV technology presents revolutionary change, its adoption will be evolutionary [10]. We expect Level 4 autonomy—operating within virtual geographic boundaries—to

be disruptive and available between 2020 and 2022, with full adoption coming later. Full autonomy with Level 5 technology—operating anytime, anywhere—is projected to arrive by 2030 at the earliest, with greater adoption by that time [11]. These advances will change how people think about mobility in urban environments, with the greatest impact seen in car ownership and public transportation use.

The overall organization of the research work is given as follows: In this section, detailed introduction about the autonomous driving and its characteristics has been given. In section 2, discussion about the various related research methodologies has been given. In section 3, proposed research methodology is discussed with suitable examples and explanation. In this section working flow of proposed method and their benefits has been given in detail. In section 4, discussion about the results and discussion of the proposed research methodology has been given. Finally in section 5, conclusion of the research work is given based on simulation outcome.

II. RELATED WORKS

Huang et al (2008) introduced a real-time object detection algorithm to be used for night-time visual surveillance in [12]. This algorithm is dependent on contrast analysis in which the contrast observed in local change over time is utilized for detecting the probable objects that are in movement. Then, motion prediction and spatial nearest neighbour data association are utilized for the suppression offake alarms. The experimental results on real scenes indicate that this novel algorithm is efficient for night-time object detection and tracking.

Javed& Shah (2002) [13] explained about the problems thathave to benecessary to be solved before the design of entirely automated outdoor surveillance systems, and solutions for solving the surveillance based issues were presented. Any type of outdoor surveillance system has topossess the capability of tracking the objects, which are moving within its area of view, and then the classification of these objects and identification of some of their activities are performed. The authors developed a method for tracking and categorizing these objects in practical conditions. Object tracking using a single camera is performedwith background subtraction that is then followed by the region correspondence.

Ahn&Lee (2015) introducedefficient object identification and tracking method in [14], which uses a complex feature matching to be usedin practical time scenarios. This algorithm identifies an object that uses invariant features, and reduces the dimension of a feature descriptor to manage the reorganization problems faced in video surveillance in the cloud environment. The experimental results achieved from the method is found to be quicker and more robustcompared to the conventionalapproaches, and in addition to this, the method can perform the identification and tracking of a mobile object with accuracy in differentscenarios.

Zenget al (2015) suggested an idea to prevent the traffic accidents by using an object detection system in [15], which is neglected for the road traffic surveillance systems, which is based on the information obtained from three-dimensional image. In this research work, a Binocular Information Reconstruction and Recognition (BIRR) algorithm were utilized for the implementation of this concept. The suspected and neglected objects were detected by the static foreground region segmentation algorithm, which is dependent on the surveillance videoacquired from a monocularcamera.

Kanhereet al. [16]used the 3D world position associated with the points of the feature points of the vehicles and arranged the points altogether. The Scale-Invariant technique has been usedfor trailing by Choi et al. [17]. The drawback of feature-based trailing is the credit rate of the vehicles by using the two-dimensional image mechanism is very less, due to non-linear distortion because of perspective projection and the image variations because of motions proportional to the camera. In quantized RGB space.Chachichet al. [18]used color signatures for trailing the vehicles. Therefore in this research work, vehicle detections are associated with one another by using a hierarchical decision process, which consists of color information, arrival probability, and driver conducts.Using an immobile camera, the pattern-recognition associated feature to on-road vehicle identification has been evaluated with the add-on to the trailing vehicles. Just the tracking associated with color and pattern matching isn’t real.

III. VEHICLE MOVEMENT CONTROL THROUGH MOVING OBJECT DETECTION

In the proposed research method, initially preprocessing is done on the traffic surveillance video to remove the noises by applying the Hybrid Median Filter method. And then background extraction is done by using markov random field process to extract the required objects. After extracting the required objects from the videos, Haar like Feature extraction is done to detect the objects

present in the videos. Based on those feature road and objects detection is performed dynamically by applying Yolo classifier. This information is utilized to adjust the vehicle speed, acceleration and steering values to ensure the uninterrupted vehicle movements without any accidents.

3.1. PREPROCESSING

Suppose that the input representation in the form of video frames pixels as shown in the figure of 3x3 matrix, the 45^{o} neighbours median values shapes to an $'x'$and the 90^{o}neighbors of the representation pixels creating a $'+'$ are computed with the median value and the central pixel of that set is then fixed as the new pixel standard. Like in the case of median strain, the three pace position process does not enforce a solemn computational result. Each position process is a highly lesser number of standards compared to thatutilized in a square region of the same size. In the hybrid method, each one of the two clusters hasjust 5 pixels, and the last assessment occupies just three standards. The Hybrid Median filter method is quickercompared to the conservative median, with the extrachange done in the logic of values. This median strain leads to the tendency of the median and shortened median filters to remove lines that are slighter compared the half width of the neighbor and to the encircling corners. In comparison with other median filters, Hybrid Median Filter help in preserving the edges better, like it performs a three-step ranking operation, in which it categorizes the information from different spatial instructions individually. In Hybrid median filtering technique, anoverall of three median values are computed for example, HVspecifies the median of horizontal h and vertical v pixels, and D refers to the median of obliqued pixels and the strained value represents the median of the two median standards and the central pixel C, Median of the vertical and horizontal, diagonal and central corresponds to HV, D, C.

$$\left\{\begin{bmatrix} D & * & D \\ * & C & * \\ D & * & D \end{bmatrix}\begin{bmatrix} * & HV & * \\ HV & C & HV \\ * & HV & * \end{bmatrix}\begin{bmatrix} * & * & * \\ * & C & * \\ * & * & * \end{bmatrix}\right\} = \begin{bmatrix} D & HV & D \\ HV & DCHV & HV \\ D & HV & D \end{bmatrix} \quad (1)$$

This new filter is the transformedversion of the hybrid median filter described above. It functions on the sub-windows identical to hybrid median filter. The maximum value of the 45^{o}neighborhood pixels generate$'X'$ and the median value of the 90^{o}neighboring pixels generatea $'+'$, which will be match up with the central pixel and the median value of that specific set is then maintained to be the new pixel value. The course of action of hybrid median filter is given as below:

Step 1: Find the median HV of the horizontal and vertical pixels that islabelled as HV and the central pixel Cin the 3x3 matrix

Step 2: Find the maximum $D_{max}$of the diagonal pixels labelled as Dand the central pixel Cin the 3x3 matrix

Step 3: Finallycompute theNew Median Value M_{new}

$$M_{new} = median(HV, D_{max}, C) \quad (2)$$

Step 4: Filter median out value$I_{i,j} = M_{new}$; here i, jrefer to the indices of the spatial locations.

3.2. BACKGROUND SUBTRACTION

The MRF Background subtraction technique is utilized for thedetection of a moving object in the form of a foreground through its segmentation from a traffic scene of a traffic surveillance camera. The surveillance camera might be stationary or dynamic [19]. The process of Background subtraction attemptsidentifying the objects, which are in movement like vehicles moving in traffic scenes from the difference attained between the present frame and the reference frame in a pixel-by-pixel or block-by-block approach. The reference frame of the video file is usuallyknown as background image or background model. There is a variations in the dynamic traffic scenes; a good background model follows the dynamic scenes. The background information getting updated in regular intervals is capable of doing this [20], even though this can also be carried out without no update done on the background information [21]. Foreground objects in a video stream are identified with the help of Background subtraction technique. It is an extensively utilized real-time technique. It is the most significant stage in surveillance application. This work enforced the development of a background algorithm such as Markov Random Field (MRF), which aimsat enhancing the performance associated with the objects classification andidentification process. This section discusses about the MRF based background subtraction in traffic scenes in detail.

Background Subtraction Model: Consider the present noise independent video file image to be$O$ with s pixels from the Gaussian filtering technique with background representation to be$O^{b}$and foreground representation to be O^{f}, and $\Gamma_{SB}, \Gamma_{DB}$and Γ_{F}represent static and dynamic background and foreground with respect to regions correspondingly. The present noise free video file image as O can be represented in terms of noise-independent foreground pixel value m, mean of significant static background cluster μ_{SB} and significant dynamic background cluster μ_{DB} ,

$$O(s) = \begin{cases} \mu_{SB}(s) & s \in \Gamma_{SB} \\ \mu_{DB}(s) + v(s) & s \in \Gamma_{DB} \\ m(s) & s \in \Gamma_F \end{cases} \quad (3)$$

Let the measurement Ddenote the difference between the present noise independent image Img and background image B, and correlated-cluster-distance dtindicates the distance between two means of correlated clusters. The resultant measurement d is given as,

$$d_o(s) = \begin{cases} dt(s) & s \in \Gamma_{SB} \\ dt(s) + v(s) & s \in \Gamma_{DB} \\ dt(s) - m(s) - \mu_B(s) & s \in \Gamma_F \end{cases} \quad (4)$$

Wherecorrelated-cluster-distance dt_s is expressed as

$$dt(s) = \begin{cases} 0 & s \in \Gamma_{SB} \cup \Gamma_F \\ |\mu_{DB}(s) - \mu_B(s)| & s \in \Gamma_{DB} \end{cases} \quad (5)$$

The distribution of pixel corresponding to the static background can be considered to beentirely overlapped with the distribution of maintained background representation at the same position, i.e. $\mu_{SB} = \mu_B$,while the distribution of pixel corresponding to the dynamic background can be considered to be partially overlapped with the distribution of maintained background representation. As with the pixel corresponding to the foreground, its distribution is regarded to be uncorrelated with that of the maintained background representation; and in accordance, the correlated-cluster-distance is zero. Provided the measurement$d_o(s)$, and the two hypothesizes below are found: H_0: $s \in \Gamma_{SB} \cup \Gamma_{DB}$; H_1 : $s \in \Gamma_F$.

The process designates the label q_s to every pixel s form the binary label-set: $w = \{Fg, Bg\}$with respect to two probable classes: foreground (Fg) and background (Bg). EAch label q_s either takes up the value Bgin case the found measurement $d_o(s)$adopts the null hypothesis H_0, or the value Fg in case the observed value of $d_o(s)$ dose not follow this assumption, i.e. alternate hypothesisH_1. Hence, the background subtraction is equal to a global labeling$\phi = \{q_s | s\epsilon S\}$. Making use of Bayesian estimation, the label set is estimated so that it's a posterior probability $P(\phi|d)$provided the measurement object d is increased. In this newly introduced MRF background subtraction process, the Bayesian estimation is used for context distinct regions identical to local region processing methods. The context distinct region is represented as sub-image thathas significant measurement d to be found. The below section explains about the background subtraction employing MRF with Bayesian estimation.

Markov Random Field (MRF): Supposelabeling sets $\phi^s_{Bg} = \{q_i = Bg \,|i \neq s,\ i \in S\}$and $\phi^s_{Fg} = \{q_i = Fg \,|i \neq s,\ i \in S\}$ from Bg and Fg are known, the Bayesian likelihood ratio [22] of measurement d at pixel s can be found as,

$$\frac{P\left(\bar{d}_o(s)\middle|\phi^s_{Bg}\right)\ \ Bg}{P\left(\bar{d}_o(s)\middle|\phi^s_{Fg}\right)\ Fg} \lessgtr Th.\frac{P(\phi^s_{Bg})}{P\left(\phi^s_{Fg}\right)} \quad (6)$$

whereThrefers to a decision threshold, and $\overline{d_o}$, which includes the measurements lying within the structural window$w$, is represented$\overline{d_o}(s) = \{d_o(s)|s \in w_s\}$, $P\left(\bar{d}\middle|\Phi\right)$refers to the conditional pdf of the observed measurement image $d$provided global labelling ϕ . $P\left(\phi^s_{Bg}\right)$and $P(\phi^s_{Fg})$stands for priori probabilities for Bg and Fg respectively. As the Gaussian distribution has been performed on background model for every pixel, measurement can, therefore, be considered to follow zero mean Gaussian distribution after the removal of the contribution made by correlated-cluster-distancedt. The information obtained from two different models is integrated with the help of Bayesian estimation. The local visual observations at every node to be labeled provide label likelihoods. The resultant label likelihoods are integrated with a priori spatial knowledge about the neighborhood represented as an MRF.As a result, $P\left(\bar{d}\middle|\Phi\right)$leads to conditional Gaussianpdf$\eta(.)$ in structural window formulation using the conditionally free assumption,

$$\eta(\overline{d_o}(s)\Big|H_j) = \left(\left(\frac{1}{\sqrt{2\pi\sigma_j^2}}\right)^{w_s} \exp\frac{\sum_{s\epsilon w_s}(d_o(s) - dt(s))}{-2\sigma_j^2}\right)($$

The priori probability is modelled with the help of Markov random fields.Considering the above equation, the likelihood test equation (7) can be further modified to regional format. In this, pixels present in the same local region have the same variance and correlated-cluster distance for the likelihood test. Using the fact that background related distributions within a local region can be

regained from the multivariate Gaussian background model. The domain-oriented information can be used by encoding any related parameters regarding that region to choose the right regional variance and correlated-cluster distance.

Let σ_R^k and dt_R^kdenote the chosen variance of null hypothesis and correlated-cluster-distance of kth local region. Taking the variance associated with the foreground is much greatercompared to the variance associated with the camera noise or the dynamic background change, the Bayesian likelihood test can then be approximated by computing the logarithm on both sides, and the obtained equation is based on a priori measurement distributions been found in temporal domain. The σ_R^k further acts as a multiplication factor that adaptively increments/decrements the threshold with respect to the context of noticed background change in a local region.The MRF-based Bayesian estimation model for background subtraction initialization mechanismlets the training video file sequence to have foreground objects. This work have specified that the results of a good background model initialization has considerably improved foreground object detection, which leads for a superior object detection in surveillance system.

3.3. DETECTION USING COLOR MAP

A labelled image comprises of two arrays, which are an image matrix & a colormap. The colormap is basically a sorted set of values, which indicates colors in image. For every image pixel, image matrix has a value, which is an index into colormap. The colormap is actually an m-by-3 matrix of class double. Every row of colormap matrix indicates red, green, & blue (RGB) values for a single color:

$$color = [R\ G\ B]$$

R, G, & B refer to the real scalars, which range between 0 (black) and 1.0 (full intensity)

The pixels present in image are denoted by integers that are pointers (indices) to color values saved in colormap. The association between values in image matrix &colormapis dependent on whether image matrix belong to class double or uint8. If image matrix belongs to class double, value 1 leads to first row in colormap, value 2 leads to second row, & so on. If image matrix belongs to class uint8, an offset exists; value 0 directs to first row in colormap, value 1 directs to second row, & so on. The uint8 convention is also utilized in graphics file formats, & enables 8-bit indexed images to support up to 256 colors. In image shown above, image matrix belongs to class double, and therefore no offset is present. For instance, value 5 directs to fifth row of colormap.

First, $\mu(i, j)$ and $\sigma(i, j)$ matrices are thresholdedand the results are mapped on the image obtained by background subtraction, S(i, j) . Therefore, all the pixels with the following condition form the dark portions of the moving vehicle and are very possibly to be the strong shadow:

$$\mu(i,j) > mean\_threshold$$

$$\sigma(i,j) > std\_threshold$$

$$S(i, j) = 1$$

3.3. HAAR LIKE FEATURE EXTRACTION TO DETECT OBJECT

Haar-Like is a rectangular simple feature that is used as an input feature for cascaded classifier. In Figure 1, there are some filters based on Haar-Like feature. By applying every one of these filters into one special area of the image, the pixel sums under white areas are subtracted from the pixel sums under the black areas. That is the weight of white and black area can be considered as "1" and "-1", respectively.

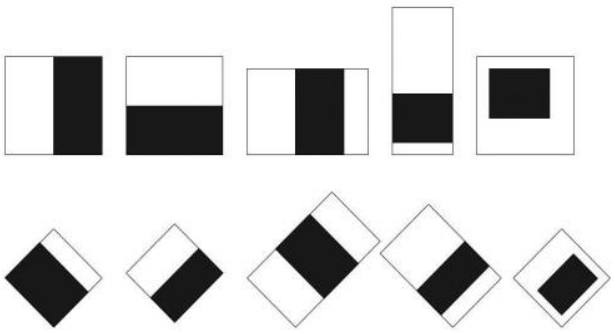

Figure 1. Different kinds of filters based on Haar-Like feature

For example, for using shadow under the vehicle and tire position feature, spatial filters are used to extract the best feature. In Figure 2, an example of these filters is shown.

Figure 2. The used vehicle's features by Haar-Like filter

The most important characteristic of the Viola-Jones framework is that the Haar like features can be computed rapidly at all scales in constant time by using the integral image. The Haar-like features of a gray-level image is intrinsically the gray contrast of local areas. The colorcontrast to skin color or the difference of gray gradient can be regard as another form of gray image, this kind of single valued feature represented in the form of gray image is called 'gray like image' in this paper. In this way, the Haar like features are not only extended in their forms but also in their physical natures, which add additional domain-knowledge to the learning framework and which is impossible to learn from gray-level image. All features represented in the form of gray like image can be used in the Viola-Jones framework, as they can be converted to Haar like features efficiently after computing their integral images.

Haar like features can be computed from gray like image, which may be gray image, gray gradient or color gradient, orientation, color difference and so on. From the point of view of filter design, the basic Haarlike feature is essentially a gradient filter, a first derivative, and gradient contrast is essentially a Laplacian filter, a second derivative.

3.4. YOLO CLASSFIER TO PREDICT THE OBJECT TYPE AND LOCATION

YOLO, a novel approach to object identification is presented. The earlier work on object detection re-uses classifiers to carry out detection. Rather, object detection is framedto be a regression problem for spatially differentiatedbounding boxes and related class probabilities. A single neural network does the prediction of bounding boxes and class probabilities directly from complete images in single evaluation. As the entire detection pipeline is a anindividula network, its optimization can be done end-to-end directly over detection performance.The convolutional layers are pretrainon the ImageNet 1000-class competition dataset. For the pretraining, the first 20 convolutional layers are used from Figure 3, which is then followed by an average-pooling layer and an entirely connected layer. This network is trained for nearly a week and attains a single crop top-5 accuracy of 88% on the ImageNet 2012 validation set, in comparison with the GoogLeNet models in Caffe's Model Zoo. The Darknet framework is used for all training and inference purposes.

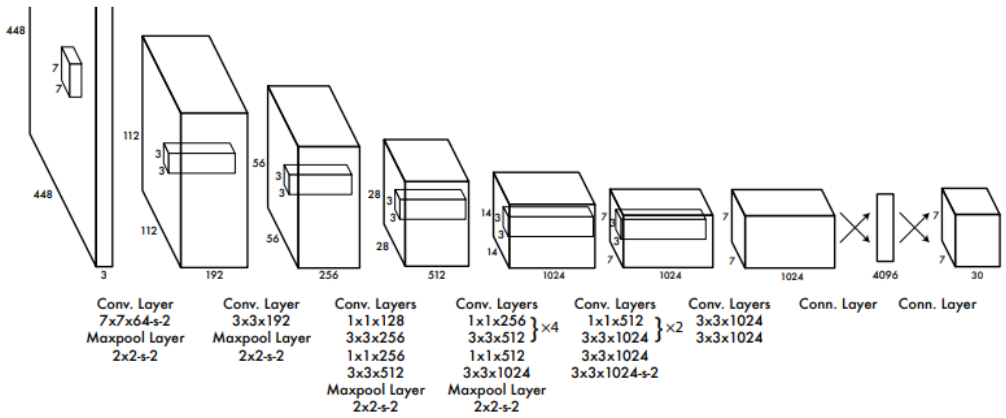

Figure 3. The Architecture. The detection network consists of 24 convolutional layers, which is followed by 2 fully connected layers. Alternating 1 × 1 convolutional layers decrease the features space from the earlier layers.

The convolutional layers are pretrainedon the ImageNet classification task at half the resolution (224 × 224 input image) and thereafter double the resolution for detection

Then the model is converted to carry out detection. Ren et al. showed that combining both convolutional and connected layers to pretrained networks can help in improving the performance. Adhering to their example, four convolutional layers and two completely connected layers are added with randomly initialized weights. Detection frequentlyneeds fine-grained visual information andtherefore the input resolution of the network is increased from 224 × 224 to 448 × 448. The final layer provides the prediction of both class probabilities and bounding box coordinates. The bounding box width and height are normalized by the image width and height such that they come between 0 and 1. The bounding box x and y coordinates are parametrizedto be offsets of a specific grid cell location such that they are also constrainedin between 0 and 1. A linear activation function is used for the last layer and every other layer make use of the below leaky rectified linear activation:

$$\Phi(x) = \begin{cases} x, & \text{if } x > 0 \\ 0.1x, & \text{otherwise} \end{cases}$$

Optimization is done sum-squared error in the result of themodel. Sum-squared error is usedas it is quite easy to be optimized, but it does not exactly align with theobjective of increasing the average precision. It weighs localization error equally with classification error that may not be an ideal condition. Moreover, in each image several grid cells do not have any object. This moves the "confidence" scores of those cells to zero, frequentlydominating the gradient from cells, which have objects. This can result in model instability, resulting in the training to diverge very early. To mitigate this, the loss from bounding box is increased to coordinate predictions and reduce the loss from confidence predictions for boxes, which don't have objects. Two parameters, λ_{coord} and $\lambda_{noobj}$are used forachieving this. $\lambda_{coord} = 5$ and $\lambda_{noobj} = .5$ is set.

Sum-squared error also equally weighs the errors in big and small boxes. The error metric needs todepict that tiny deviations in the big boxes do not matter much compared to in small boxes. In order to partially deal with this, the square root of the bounding box width and height is predicted rather than the width and height directly. YOLO does the prediction of thenumerous bounding boxes for each grid cell. During the time of training, just one bounding box predictor is needed to be accountable for every object. One predictor is assigned to be "responsible" for the prediction of an object depending on the prediction that has the greatest current IOU with the ground truth. This results in specialization between the bounding box predictors. Every predictor becomes better at the prediction ofparticular sizes, aspect ratios, or classes of object, increasing the overall recall.During training, the following, multi-partloss function are optimized:

$$\begin{aligned} &\lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{\text{obj}} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \\ &+ \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{\text{obj}} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \\ &+ \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{\text{obj}} \left(C_i - \hat{C}_i \right)^2 \\ &+ \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{\text{noobj}} \left(C_i - \hat{C}_i \right)^2 \\ &+ \sum_{i=0}^{S^2} 1_{i}^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2 \end{aligned}$$

where1_{ij}^{obj}represents if object shows up in cell i and 1_{ij}^{obj}represents that the jth bounding box predictor in cell i is "responsible" for that specific prediction. It is to be noted that the loss function penalizes classification error only when an object exists in that grid cell (therefore the conditional class probability explained before). It also just penalizes bounding box coordinate error when that predictor is "responsible" for the ground truth box (i.e. exhibits the highest IOU of any predictor in that specific grid cell).

IV. RESULTS AND DISCUSSION

In this research, proposed method has been implemented and evaluated in the OpenCV simulation environment. The experiments are conducted on the videos gathered from the social media.OpenCV built in mode helps in creating more positive images with the distortion on the actual positive image and using a background image. But, it does not permit doing this for a number of images. The performance metrics that are considered in this research method for the efficient implementation of the proposed and existing research methodologies are listed as follows: "Accuracy, Sensitivity,

Specificity, Precision, Recall, F-Measure, and detection rate".

The evaluation of the proposed method Fast and Accurate Road Detection System (FARDS)based on these performance metrics is done by comparing it with the existing research method namely Yolo classifier. The numerical evaluation of the proposed research method is conducted by comparing it with the existing research method which is shown in the following figure 4.

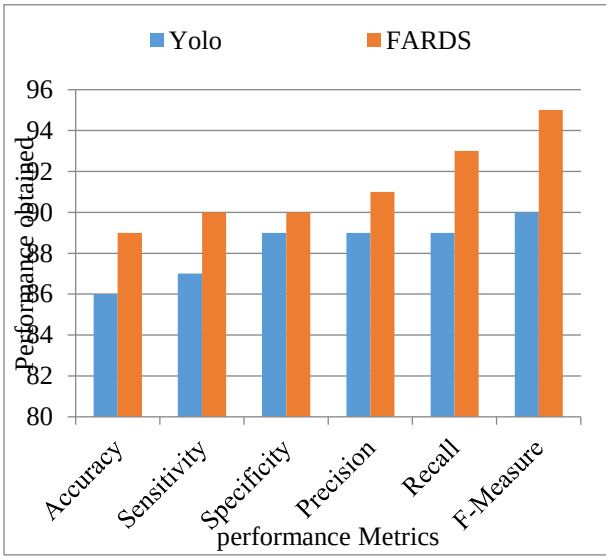

Figure 4. Numerical Comparison Outcome

From the figure 4, it can be concluded that the proposed research method leads to provide the improved performance than the existing research method by accurately retrieving the similar videos from the training database. From this outcome it can learnt that the proposed method FARDS shows 3.3% improved performance ratio than the existing research methodologies in terms of accurate detection of objects**.**

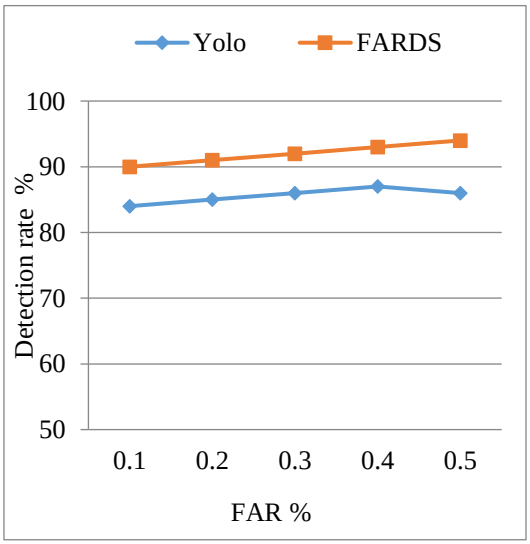

Figure 5. Detection Rate comparison

Figure 5 compares the detection rate of the multimodal biometric system by using FARDS approach and existing Yolo approach. In x-axis false acceptance rate is taken and y-axis detection rate is taken. From the graph results, it is observed that, the proposed FARDS is achieves better detection compared with existing methods.

V. CONCLUSION

In the proposed research method, initially preprocessing is done on the traffic surveillance video to remove the noises by applying the Hybrid Median Filter method. And then background extraction is done by using markov random field process to extract the required objects. After extracting the required objects from the videos, Haar like Feature extraction is done to detect the objects present in the videos. Based on those feature road and objects detection is performed dynamically by applying Yolo classifier. This information is utilized to adjust the vehicle speed, acceleration and steering values to ensure the uninterrupted vehicle movements without any accidents. The overall evaluation of the research method is done in the matlab simulation environment from which it is proved that the proposed method FARDS leads to ensure the better outcome than the existing research techniques.

REFERENCE

1. Kyriakidis, M., Happee, R., & de Winter, J. C. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. Transportation research part F: traffic psychology and behaviour, 32, 127-140.

2. Katrakazas, C., Quddus, M., Chen, W. H., &Deka, L. (2015). Real-time motion planning methods for autonomous on-road driving: State-of-the-art and future research directions. Transportation Research Part C: Emerging Technologies, 60, 416-442.

3. Litman, T. (2017). Autonomous vehicle implementation predictions. Victoria, Canada: Victoria Transport Policy Institute.

4. Akpinar, M., &Vincze, Z. (2016). The dynamics of coopetition: A stakeholder view of the German automotive industry. Industrial Marketing Management, 57, 53-63.

5. Jung, W., Hajredini, B., &Zvereva, V. (2018). Fitness to drive with cardiovascular diseases: current guidelines of the German Federal Highway Research Institute. Herz.

6. Zeeb, K., Buchner, A., &Schrauf, M. (2016). Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. Accident Analysis & Prevention, 92, 230-239.

7. Grace, R. (2018). Concept Cars Offer a Peek into the Future of Mobility: As technologies enable a reality with driverless vehicles, automakers and suppliers are redesigning the very essence of what a car should be. Plastics Engineering, 74(3), 10-14.

8. Maurer, M., Gerdes, J. C., Lenz, B., & Winner, H. (2016). Autonomous driving. Springer Berlin Heidelberg, Berlin, Germany.

9. Brenner, W., & Herrmann, A. (2018). An overview of technology, benefits and impact of automated and autonomous driving on the automotive industry. In Digital Marketplaces Unleashed (pp. 427-442). Springer, Berlin, Heidelberg.

10. Claybrook, J., & Kildare, S. (2018). Autonomous vehicles: No driver… no regulation?. Science, 361(6397), 36-37.

11. Groshen, E. L., Helper, S., MacDuffie, J. P., & Carson, C. (2018). Preparing US Workers and Employers for an Autonomous Vehicle Future.

12. Huang, K., Wang, L., Tan, T., &Maybank, S. (2008). A real-time object detecting and tracking system for outdoor night surveillance. Pattern Recognition, 41(1). 432-444.

13. Javed, O., & Shah, M. (2002). Tracking and object classification for automated surveillance. In Computer Vision—ECCV, 2002, 343-357.

14. Ahn, H., & Lee, Y.H. (2015). Performance analysis of object recognition and tracking for the use of surveillance system. Journal of Ambient Intelligence and Humanized Computing, 1-7.

15. Zeng, Y., Lan, J., Ran, B., Gao, J., &Zou, J. (2015). A Novel Abandoned Object Detection System Based on Three-Dimensional Image Information.Sensors, 15(3).6885-6904

16. Kanhere, N., Birchfield, S., &Sarasua, W. (2006). Vehicle segmentation and tracking in the presence of occlusions. *Transportation Research Record: Journal of the Transportation Research Board*, (1944), 89-97.

17. Choi, J. Y., Sung, K. S., & Yang, Y. K. (2007). Multiple vehicles detection and tracking based on scale-invariant feature transform. In *2007 IEEE Intelligent Transportation Systems Conference* (pp. 528-533). IEEE.

18. Chachich, A. C., Pau, A., Barber, A., Kennedy, K., Olejniczak, E., Hackney, J.,&Mireles, E. (1997). Traffic sensor using a color vision method. In *Photonics East'96* (pp. 156-165). International Society for Optics and Photonics

19. Hu, W., Tan, T., Wang, L., &Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*,34(3). 334-352.

20. Lin,H.H., Liu,T.L., &Chuang, J.H., (2009).Learning a scene background model via classification. *IEEE Trans. Signal Process*. 57(5).1641–1654.

21. Du-Ming, T., &Shia-Chih,L., (2009). Independent component analysis-based background subtraction for indoor surveillance. *IEEE Trans. Image Process*. 18(1).158–167.

22. Zhang, W., Duan, P., Lu, Q., & Liu, X. (2014). A realtime framework for video object detection with storm. *Ubiquitous Intelligence and Computing, 2014 IEEE 11th Intl Conf on and Autonomic and Trusted Computing, IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops*