International Journal of Scientific Research in Computer Science, Engineering and Information Technology


# Driver Emotional Status Recognition Using Artificial Neural Network

**Mohan Arava*[1], Prudhvi Ravi Raja Reddy Mallidi[2]**

[1]Assistant Professor, Department of IT, Aditya College of Engineering & Technology, Surampalem, Kakinada, East Godavari, Andhra Pradesh, India

[2]Assistant Professor, Department of CSE, Aditya College of Engineering & Technology, Surampalem, Kakinada, East Godavari, Andhra Pradesh, India

## ABSTRACT

Artificial neural network is one of the fascinating area of study, the proposed architecture is performs better feature extraction than earlier proposed Fuzzy logic and SVM algorithms. Day to day the automotive industries are actively supporting research and innovations related to safety issues, performance, and environment. A driver status assessment system constructed in two different modules: one is for driver fatigue detection based on captured images through Digital cameras. The fatigue is a percentage of eyes closure of the best indicator of fatigue for vision systems. And the second one is driver distraction system by using head, facial expressions, and body, a fusion strategy is to deduce the type's driver distractions. Of course many aspects are there but these two fatigue and distraction are only a fraction of all possible drivers' states of their dramatic impact on traffic safety.The standard databases are used for quantitative evaluation of the current state of the art approach using ANN achieves better accuracy then compared to Fuzzy and SVM. The proposed architecture endorses the efficiency and reliable usage of the work for real world applications.

**Keywords :** Artificial Neural Network, SVM Algorithms, Fuzzy Logic, GoogleNet, AAM

## I. INTRODUCTION

Industries are actively supporting research and development is to meet the safety issues, performance, and environmental protection. A novel Cutting edge technologies are designed to cope with intelligent transportation systems in safety. To protect human life's this application is enable intra and inter vehicle communications, critical information and driver status, road obstacles, or lane departure, has to be assessed. Ongoing Research in the field of robotics are especially in the field of humanoid robots are becomes interesting to integrate these capabilities into machines allowing more diverse and natural way of communication.

From last two decades many methods are rely on extraction of the facial regions. Methods are often involves the Facial Action Coding System, which describes the facial expression using Action Unit. The present approach uses Artificial Neural Networks (ANN). The trend of ANNs differ it trained on data with less need of manual interference. It is a special kind of algorithm and have been shown to work well as features extraction whenever we are using images/videos as input and are real time capable. GoogleNet has been introduced during the Image Net Large Scale Visual Recognition Challenge 2014. The objective of this challenge analyses the quality of different image classification algorithms are implemented by various research scholars. The challenge of object detection with additional training data GoogleNet has achieved about 44% precision. Finally it has been used as inspiration for the proposed architecture. The proposed ANN has been evaluated on the extended Cohen-Kanade Dataset and on MMI Dataset. The results of experiment on these datasets demonstrate the success of deep layered Artificial Neural Network Structure, with a 10-fold Cross validation of recognition accuracy 99% has been achieved.

## II. RELATED WORK

The detailed overview of driver distraction system for facial emotions recognition was given by Bettadapura. The literature survey of similar to the proposed methods are presented. Earlier Szegedy et al have proposed an architecture of GoogleNet. It is a Multilayer deep network of ANNs. For Any algorithm we give image as input, based on methodology which we are using it will localizes the face region. From the facial region, facial patches like the eyes, lips, or noseare detected and points are marked as key points. From those patches of key points are most variance between two images, features are extracted. Based on image dimensionality features are reduced and then given to support vector machine. With this the average accuracy is 94%.

Same as video based emotion recognition has been proposed by Byeon and Kwak. For this they developed a three dimensional ANN which uses group of 5 consecutive frames as input, with a database containing 15 persons has been used to achieve an accuracy of 95%.

Kanade dataset contains emotion annotations as well as action unit annotations. For classification they also have evaluated the datasets using Active Appearance Models in combination with SVMs. To identify the position and track of the face over different images they workout AAM which generates a Mesh/Mask on top of the face. With the mesh faces they are extracted two feature vectors. One is the normalized vertices with respect to rotation, translation, and scale. Second is gray scale image from the mesh data, input image has been extracted. This training process achieving an accuracy over 80%.

Kotsia and Pitas detect the emotions by mapping s face mask with a low number of polygons on persons face. Initially on face grid is placed randomly on the image, so it is placed manually placed on the person face. Those emotions on the grid is tracked using a KanadeLucasTomasi tracker. Whatever the information is provided by the grid is used as feature vector for multiclass SVMs. Emotions are anger, fear, happiness, sadness, disgust, and surprise of various expressions based they evaluated with model on the Cohn-Kanade dataset and an Accuracy of 99% has been achieved.

LBP (Local Binary Patterns) are used to calculate over the facial region. With the extracted LBPs a feature vector is derived. Those features are depend on the position of the sub regions over witch LBP is calculated. And also used AdaBoost is used to find the sub region of images which contains the most discriminative information. With various classification algorithms has been evaluated of which an SVM with Boosted-LBP features performs the best with a recognition accuracy of 95% on various dataset.

## III. DESIGN AND IMPLEMENTATION

### a. Support Vector machine

Support Vector Machine (SVM) is a common kernel based algorithm in machine learning. SVM can be applied to classification and regression. The basic idea of SVM is that to find optimum separation hyperplane in order to separate data into two classes [10].

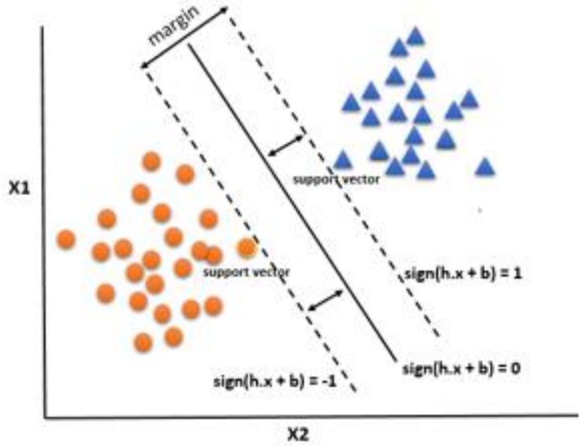


**Figure 1.** Separation Hyperplane of SVM.

Given training data **K**$\{x_i,y_i\}$,i = 1,2,...,n,, $x_i \in R^d y_i = \{1,-1\}$. SVM finds hyperplane which has the largest margin that separates two classes. All training data should satisfy $sign(x.h + b \geq 1)$ for $y_i$= 1 and $sign(x.h + b \leq -1)$ for $y_i$ = -1, the point which is plotted on hyperplane satisfies$sign(x.h + b = 0)$. Where $x$ is a input vector, $h$ is weight vector and b is bias. The data that have nearest distance to decision boundaries are called support vectors. In order to define separating hyperplane, the formulation can be written as

$$x.h^T + b \geq +1$$
$$x.h^T + b \leq -1$$

And the formations above can be combined in simple form as

$$y_i(x_i . h_i^T) \geq +1$$

***Optimum separation hyperplane, can be obtained by maximizing $\frac{2}{||h||}$. Because it is constraint optimization problem, it can be solved by the Lagrangian Multiplier***

$$min\ L_p = \frac{1}{2}\|h\|^2 - \sum_{i=1}^{l} a_i y_i (h_i . x_i + b) + \sum_{i=1}^{l} a_i$$

***After the derivative at min = 0, it can be expressed as***

$$h = \sum_{i=1}^{l} a_i y_i x_i, \sum_{i=1}^{l} a_i y_i = 0$$

***Therefore, the decision function for classification can be written as***

$$f(x) = \left( \sum_{i=1}^{l} a_i y_i x_i . h \right) + b$$

All the detail of SVM can be seen in [11]. In case of non-linear data, the data are mapped into high dimensional feature space using kernel function. In this study, we used several kernel functions such as, Polynomial, Radial Basis Function, and Linear.

**b. Artificial Neural Network**

Artificial Neural Network (ANN) algorithm that is inspired by biometric process of human facial expression recognition. It represents the process of computation using interconnected neurons which send signal among them. Basically ANN which is comprised by three layers, such as Input Layer, Hidden Layer and Output Layer is called Multilayer Perceptron.

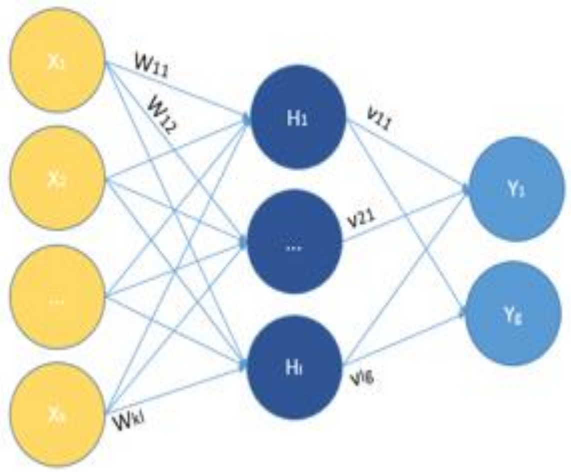


**Figure 2.** ANN Architecture.

The nodes in each layer are connected by weights. Given training data $D = (x_i, y_i)$, where $x_i = (x_1, x_2, ..., x_k) \in R^k$ and $y_i = (y_1, y_2, ..., y_g) \in R^g$. Input layer will receive values from x, and Hidden layer will calculate the result that can be expressed as

$$H_i = f\left( \sum_{k=1}^{D} w_{kl} . x_k - b_l \right)$$

Where $w_{kl}$is weight of input layer that is connected to any hidden node, l is the number of hidden nodes, and f is the than activation function that can be written as

$$f(x) = \frac{2}{(1 + exp^{-2x})} - 1$$

The calculation in output layer uses pureline activation function which can be expressed as

$$Y_i = s\left( \sum_{l=1}^{l} H_l . V_l - b \right)$$

Where $H_l$ represents a hidden node in hidden layer, $v_i$is the weight between hidden layer and output layer, b is the bias, and s represents pureline activation function that can be formulated as

$$s(x) = x$$

the error will be calculate by computed value of e, which is

$$e = Y_{actual} - Y_{desired}$$

In order to minimize the error, the computation in each layer will be repeated and all weights will be updated using back propagation algorithm. Several optimized back propagation algorithms such as Resilient Back propagation, Levenberg Marquardt, Scale Conjugate Gradient are used in this study. For further details of those techniques can be seen in [12], [13].

## IV. EXPERIMENTAL DISCUSSION

With the CKP database has been very well analyzed and the effective possible recognition accuracy has been achieved. For data preprocessing KNN imputation algorithm is used to fill the missing values of the dataset. In this case we set K=3. Data normalization [0,1] has been performed to continue variables.

After data preprocessing it divide the training and testing data using K-Fold Cross validation set value K=10. Training data will be used for training process of ANN, SVM, and Fuzzy. The aim of the training process is to train those algorithms to recognize the person emotions corresponds to surprise. Those images are depicts a person with a wide open mouth and eyes distractions. For ANN we try to add the odd number of hidden neurons between 49-101. And also applied various optimized back propagation algorithms are including Levenberg-Marquardt Resilient back propagation and Scaled Conjugate Gradient. Especially for SVM we train model using various kernel function, Gaussian and Linear.

In this study, Accuracy, Sensitivity and Specificity are used to measure the performance of each model of ANN, SVM and ELM. Accuracy, Sensitivity and Specificity can be expressed as first we calculate the false negative and second we calculate the false positive. Then we can go for Accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Spesificity = \frac{TN}{TN + FP} \quad (3)$$

Here TP is true positive means number of pixels that are detected as the changed with some reference value in dataset. TN is true negative it means to detect the unchanged reference values from dataset.

For this Application we need various system configurations like we need a Laptop processor core i7 $5^{th}$, RAM 8GB, system operating system windows 10, 64 bit. With MATLAB 2015.

## V. CONCLUSION AND FUTURESCOPE

In this paper we have built an inattention detection and recognition system based on with sensors, computer vision and Machine learning Algorithms ANN, SVM. Based on data collection with driving assistance system we are able to evaluate our work. And Compared to existing approaches aiming to detect inattention or provide a level of inattention of our system provides high-level information, which is more suitable for context aware human machine interaction. It can also use for driver safety, but also for long term statistics about driver habits. It is a neural network which is little computational effort then compared to Current state of the art SVM architectures. The result of the 10-fold cross validation is an average a recognition accuracy is 99% on the CKP dataset, it may vary the accuracy with other datasets. This paper presents a capable of competing with current state of the art approaches in the field of emotions recognition.

Possible improvements are background and environmental information have not been exploited yet and prove to be very useful in assessing the level of risk on the road. And also integrated with various additional sources related to driver (age, driving experience, fatigue) and the environment (time of day, road type, outside traffic, vehicle speed) in the inference process with dynamic Bayesian network. Moreover in recording session also used additional sensors such as a microphone, a heart rate monitor, and steering wheel signals. We will work in Future on integrating these signals into the driver assistance system. And also involving additional actions like as interactions and chatting with other passengers.

## VI. REFERENCES

[1]. K. Anderson and P. W. Mcowan, "A real-time automated system for recognition of human facial expressions," IEEE Trans. Syst.,Man, Cybern. B, Cybern, pp. 96-105, 2006.

[2]. I. Kotsia and I. Pitas, "Facial expression recognition in image sequences using geometric deformation features and support vector machines," Image Processing, IEEE Transactions on, vol. 16, no. 1, pp. 172-187, Jan 2007.

[3]. B. V. Kumar, "Face expression recognition and analysis: the state of the art," Course Paper, Visual Interfaces to Computer, 2009.

[4]. S. Happy and A. Routray, "Automatic facial expression recognition using features of salient facial patches," Affective Computing,IEEE Transactions on, vol. 6, no. 1, pp. 1-12, Jan 2015.

[5]. J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition," arXiv preprint arXiv:1310.1531, 2013.

[6]. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," CoRR, vol. abs/1409.4842, 2014. Online]. Available: http://arxiv.org/abs/1409.4842

[7]. LSVRC. (2014, Jun.) Results of the lsvrc challenge. Online]. Available: http://www.image-net.org/challenges/LSVRC/2014/ results

[8]. C.-D. Caleanu, "Face expression recognition: A brief overview of the last decade," in Applied Computational Intelligence and Informatics (SACI), 2013 IEEE 8th International Symposium on. IEEE, 2013, pp. 157-161.

[9]. V. Bettadapura, "Face expression recognition and analysis: the state of the art," arXiv preprint arXiv:1203.6722, 2012.

[10]. "SVM - Support Vector Machines." Online]. Available: http://www.support-vector-machines.org/. Accessed: 25-May-2016].

[11]. C. J. C. J. C. J. C. C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Data Min. Knowl. Discov., vol. 2, no. 2, pp. 121-167, 1998.

[12]. M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," IEEE Int. Conf. Neural Networks - Conf. Proc., vol. 1993-January, pp. 586-591, 1993.

[13]. M. T. Hagan and M. B. Menhaj, "Training Feedforward Networks with the Marquardt Algorithm," IEEE Trans. Neural Networks, vol. 5, no. 6, pp. 989-993, 1994.

## AUTHOR DETAILS:

**MOHAN ARAVA ,He is** working as an Assistant professor in Aditya College of Engineering & Technology in the department of Information Technology, Surampalem, Kakinada, Andhra Pradesh, affiliated to JNTU Kakinada and He completed his M.Tech (CST) from V.R.Siddhartha Engineering College, affiliated to JNTU Kakinada in the year 2013 and obtained his B.Tech (IT) from BAPATLA Engineering College, affiliated to ANU, Andhra Pradesh, India in the year 2010. He is interested in the fields of Image processing, Computer vision , Cloud computing, Computer Network& Information Security.

**PRUDHVI RAVI RAJA REDDY MALLIDI,He is** working as an Assistant professor in Aditya College of engineering & Technology in the department of Computer science & Enginnering, Surampalem, Kakinada, Andhra Pradesh, affiliated to JNTU Kakinada and he completed his M.Tech (CSE) from Aditya Engineering college, affiliated to JNTU Kakinada in the year 2011 and obtained his B.Tech (CSE) from Regency Institute of Technology, YANAM, affiliated to PONDICHERRY UNIVERSITY, Pondicherry, India in the year 2007. He is interested in the fields of Image processing, Cloud computing, Computer Network& Information Security