Performance Comparison of Machine Learning and Deep Learning While Classifying Driver's Cognitive State

Rahul Bhardwaj,[a,b] Swathy Parameswaran[a] and Venkatesh Balasubramanian[a]
[a]RBG Labs, Department of Engineering Design, IIT Madras, Chennai 600036, India
[b]Harita Seating Systems Limited, Belagondapalli, Thally Road, Hosur 635114, Krishnagiri District, Tamil Nadu, India
rb@haritaseating.com, swathy0406@gmail.com, chanakya@iitm.ac.in

***Abstract*—Driver fatigue is a major cause of the road accidents that occur throughout the globe. It has been observed that among total number of accidents, 20% are contributed from driver fatigue. Acknowledging the existing data it is clear that a notification system for driver fatigue is of at most importance. Over the past a large number of strategies have been tested out and among them EEG based systems have shown to be the most accurate and reliable to estimate driver's cognitive state. The direct relation of brain activity to EEG signal explains its high accuracy in a fatigue detection system. Current researches in machine learning as well as deep learning have shown a new perspective in EEG data analysis. This work proposed a highly accurate, EEG based driver fatigue classification system which can reduce the rate of fatigue related road accidents using machine learning and deep learning algorithms. The results showed that the relative power of theta, alpha, beta and delta showed significant correlation to driver fatigue. The selected features were trained and evaluated using 20 well established classifiers in the field of driver fatigue. Among all the classifiers tested, the Fine Tree, Subspace KNN, Fine Gaussian SVM, and Weighted KNN were performed to the highest accuracy levels. Different performance metrics are used for this work and Deep Autoencoder and KNN are identified as the best suitable Deep learning and Machine Learning Algorithms for driver fatigue prediction with an accuracy of 99.7% and 99.6 % respectively.**

Keywords—Driver fatigue, Machine learning, Deep learning, EEG, Driving simulator, Cognitive fatigue.

I. INTRODUCTION

Road accidents are one of the major safety concerns as it cause huge number of death across the globe. As per the most recent report released by the Ministry of Road Transport and Highways [1], an average 1317 accidents and 413 deaths occur every day on Indian roads which further interprets into 17 deaths and 55 road accidents every hour in 2016, one of the highest throughout the globe. In addition to that the rate of people killed per 100 road accidents has become more severe in 2016 compare to previous years. As reported by World Heath Organization (WHO) published by International Road Federation, Geneva, India shows higher incidence of deaths per 100,000 when compare to other countries such as France, China, Australia, Japan, Canada, U.S.A, Republic of Korea, Portugal ,Germany, Poland etc. except Russian Federation [2]. According to Ministry of Road Transport and Highways (MoRTH, 2016), 86.5 per cent of share in road accidents is accounted by top 13 States in India. Every year in India a young people in the productive age group lose their lives because of road accidents. These premature deaths lead to the substantial loss of productivity to the nation.

Every year a large number of road accidents reported due to driver fatigue. It isn't potentially feasible to identify the correct number of fatigue related road accidents even though many researches stated that driver fatigue may be contributing up to 20% of road accidents. Driver fatigue may be defined as the deterioration in driver performance that leads to increase in reaction time and decrease in driver's cognitive ability [3].

Driver Fatigue is generally classified as physical and mental/ cognitive fatigue [4], [5]. The temporary inability to continue optimal cognitive performance is known as mental fatigue. Cognitive fatigue also reduces the physical performance of the driver. Regardless, this can be hazardous when performing driving tasks that demands constant concentration. Many factors can influence to driver fatigue such as, lack of quality sleep, poor vehicle design, stress driving, driving after heavy physical task etc. Several outcomes as a result of fatigue can be observed: (i) reduced vigilance–participants perform the worst on attention-demanded tasks, (ii) Reduced information processing–fatigued driver might not remember the previous few minutes of driving, (iii) Slower reaction times–The time required to respond in an emergency situation is increased by fatigue.

Several methods have been proposed in the past to estimate driver fatigue based on vehicle based parameters (deviation from lane position, steering wheel movement) [6], behavioral measures (yawning, eye movement and head position) [7] and physiological measures (electromyogram (EMG) [8], electroencephalogram (EEG) [9] electrocardiogram (ECG) [10], [9] and electrooculogram (EOG).

To differentiate state of the driver i.e to know whether driver is in fatigued stage or alert stage, several methods have been employed. Current researches in machine learning as well as deep learning have shown a new perspective in EEG data analysis. Machine learning can be characterized as the development of algorithms which have the ability to automatically detect the patterns in data.

Physiological feature based driver fatigue prediction are more flexible to classify the fatigue into different scales rather than binary classification (alert or drowsy). EEG signals for fatigue detection have been carefully reviewed, in this literature if a closer look is taken, one would notice that many papers have been published in this regard recently; which substantiates that this is an area of an increased interest [11], [12], [13]. In spite of having plentiful studies conducted in this field, there still occurs to be many shortcomings. Practically a larger sample size is advisable for this scheme. With a larger data at hand machine learning and deep learning ensures a better classifier performance which enhances its reliability in real time [14], [15].

Deep Learning is a new emerging area under Machine Learning. It guarantees powerful and fast machine learning, taking us one step closer to Artificial Intelligence [16], [17], [15]. Compare to older algorithms where the performance degrades as the data exceeds a limit whereas in deep learning performance can be improved by feeding more data set [18], [19]. Deep learning includes several levels, which vary depending on the amount of training data set [20]. Under generative model deep learning networks Autoencoder is the most popular classifier. An Autoencoder network is a type of ANN used for unsupervised learning applications. The core purpose of Autoencoder is to dimensionally reduce the representation (encoding) of data set. [15] Suggested a multimodal approach by combining forehead EOG and partial EEG to improve driving fatigue prediction. The results showed the best performance when the forehead EOG signal with the 6-channel combined temporal EEG signal.

In this study, we have classified and compared different level of fatigue based on machine learning algorithms and deep learning algorithm.

II. Materials and Methods

A. Subject Details

The total dataset collected from 15 male volunteers to reduce inter-subject differences. The average age, height, and weight of the subjects were 26.3±3.3 Years, 1.62±0.04 m, and 60.8±8.4 kg respectively. Participant selection criteria includes good health (questionnaires and physical examination), good sleep (by questionnaires), not a current smoker, valid driver's license , right handed drivers, not prone to simulator adaptation sickness and no shift work within one month of entering the experiment. Participants were compensated for their experimental time and provided with a written informed consent. The studies were conducted in RBG labs of IIT, Madras.

B. Simulation task

The static driving simulator for this experiment was developed in RBG lab, engineering design department of Indian Institute of Technology Madras as shown in Fig.1. The driving simulator consists of a Logitech G27 racing wheel, manual gear, gas petal. A monotonous driving environment with less traffic was created on a visual display (200 cmX200 cm). A driving situation is said to be monotonous when the stimuli change in a predictable pattern or remain unchanged over a period of time, which results in a sensory stimulation

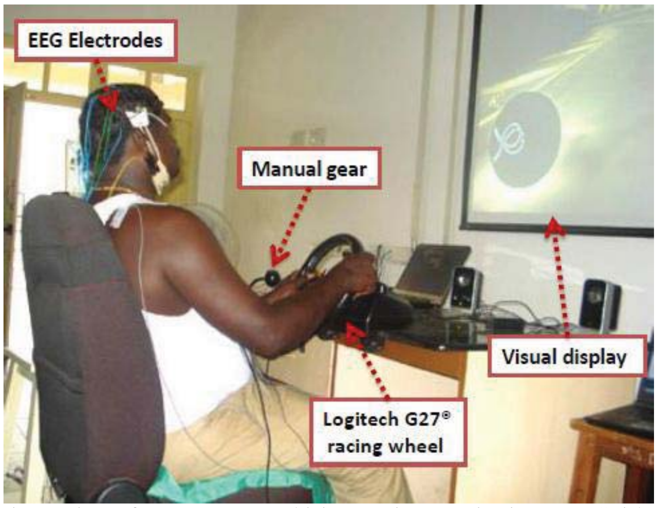


Fig.1. View of a monotonous driving session on simulator [RBG lab, Engineering design department of Indian Institute of Technology Madras]

that is highly repetitive or constant. "Need for Speed" game was selected as the driving environment which was designed with the specifications such as, simple route, few scenery changes and moving objects. The test environment was maintained with minimal illumination except for the projected visual on the display thereby not providing any external visual stimuli to the participant.

C. Data acquisition, processing and feature extraction

The EEG equipment used was Mitsar-EEG 201 system–19 channel electrodes with a sampling frequency of 291 Hz as shown in Fig.1. The electrodes are positioned based on the international 10/20 method of [21]. To reduce the resistance electro-gel was applied while fixing the cap electrodes. EEG Bipolar recordings with 19 channels were used in order to reduce artifacts produced by heartbeat. EEG electrodes were placed on all 4 lobes of the brain they are occipital, frontal, parietal, and temporal lobes.

The raw EEG data measurement is always vulnerable to noise due to the driving motion artifacts. Therefore, various preprocessing techniques, such as band pass filter, notch filter and various other artifact correction procedures are used in order to eliminate this noise. The proposed model used WinEEG software that allows the recording, analysis and editing of continuously recorded EEG data using a Mitsar amplifier. The PCA or ICA decomposition and special filters that are present in the WinEEG software helps to increase quality of EEG by eliminating the artifacts. In addition to that a customized band pass filter with a pass band of 0.5-90 Hz and a notch filter with cutoff frequency of 50Hz implemented in Matlab R2017b to remove artifacts and power line interference respectively [9].

A number of time domain and frequency domain features are extracted from the EEG signal with the help of various feature extraction techniques, which includes Fast Fourier Transform (FFT), power spectral density (PSD) and Discrete Wavelet Transform (DWT) [12]. Currently, wavelet transform and wavelet-packet transform have been used widely for the

Part No: CFP1858A-ART; ISBN No:978-1-5386-8492-4/18/$31.00 ©2018 IEEE

EEG analysis. Wavelet transform only decomposes the low frequency components of the signals. With the help of wavelet packet transform more detailed information about EEG signal can be obtained. This study utilized the Daubechies wavelet of 6^{th} order because of its efficiency, which is of order 6 (db6) and five level decomposition was chosen.

EEG time domain analysis does not give clear information about the brain regions which are activated during fatigued and non-fatigued stage and their association with human physiological states. Hence more importance is given for frequency domain features over time domain features for classifying the data into different scales of fatigue. The frequency content of EEG signal can be separated, and it is usually divided into 5 different frequency bands such as alpha, beta, theta, delta and gamma. The EEG signal decomposed into different frequency bands which includes Delta (0.5 to 4 Hz), Theta (4 to 7 Hz), Alpha (8 to 13 Hz), Beta (14 to 30 Hz) and Gamma (>30 Hz) (Fig. 2). When investigating driver fatigue, some frequency bands are of a higher interest, especially the theta, alpha, and in certain cases beta bands.

D. *Data classification*

In this study, 60 min of data was split in four time zones i.e t0=1-15 min, t1=t0+15 min, t2=t1+15 min and t3=t2+15 min. The aim of the classifier in this study is to classify the EEG data into four output classes such as non-fatigue (t0), mild-fatigue (t1), moderate-fatigue (t2) and severe fatigue (t3) based on the EEG feature set (Fig.2). The proposed work has adopted two major EEG classification approaches. First classification approaches of the proposed work deals with machine learning techniques. The machine learning technique consists of three classifiers such as SVM, KNN, DT and ensemble. Second classification approaches of the proposed work, deals with deep learning techniques. This deep learning classification approach is implemented with a classifier called Autoencoder (Fig.3). Five widely used classifiers such as KNN, SVM, Ensemble, DT and Autoencoder were selected in this proposed work.

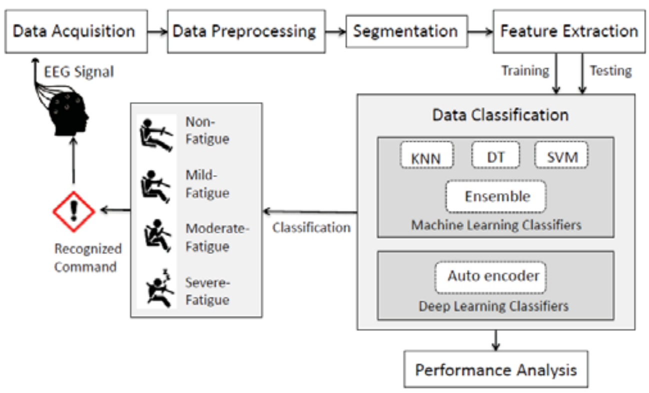


Fig.2. Overall system work flow of EEG based driver fatigue countermeasure Device.

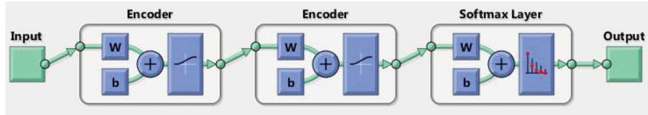


Fig.3. Autoencoder architecture generated in the Neural Network Toolbox to perform the deep learning.

III. RESULT AND DISCUSSION

A. *Selection of suitable machine Learning classifiers*

Three different types of classifiers performed under the class of decision trees. This includes Fine tree, medium tree and coarse tree classifiers with maximum number of splits 100, 20 and 4 respectively. From the observations Fine Tree or Complex tree classifier performed the best with an accuracy of 88.2 %. RUS boost, Boosted Trees, Subspace discriminant, Bagged Trees and Subspace KNN are the five different types of classifiers performed under the class of ensemble. Different learner types such as decision tree, discriminant and Nearest neighbors were adapted for each classifiers. The best performance is contributed by the subspace KNN classifiers with an accuracy of 87.5 %. From the six classifiers performed under SVM, Fine Gaussian SVM performed the best with an accuracy of 96.6 % and a Gaussian kernel function was used.

Among the six KNN classifiers were tested, fine KNN and weighted KNN performed best. Fine KNN uses a value of K = 1, hence the closest training sample will decide the class membership of the test sample. By only using the closest training sample as the predictor of class membership is not favored since it becomes more prone to noise.

Coming to the Deep Autoencoder, the autoencoder is designed to train with 2 hidden, one softmax layer and a linear transfer function used for the decoder. The L2 weight regularizer, sparsity regularizer and sparsity proportion were set to 0.001, 4 and 0.05 respectively. Once the features extracted from the first hidden layer it is used for training of second autoencoder without scaling the data. Then trained a softmax layer to perfrom classification, using the features extracted from the second autoencoder.

B. *Performance metrics for machine learning and Deep learning classifiers*

After identifying the best classifiers, the next step was to find out the effectiveness of each model based on some metric using test datasets. Different performance metrics are used to study different Machine Learning Algorithms. For classification problems confusion matrix and AUC (Area under Curve) are widely used. Another metric approach for evaluation of machine learning algorithms are precision, recall, specificity, F1 score etc. Receiver operating characteristics (ROC) graphs are helpful for visualizing classifier performance. The ROC curve is the plot between sensitivity and specificity for the selected trained classifier (Fig.4 and Fig.5).

Part No: CFP1858A-ART; ISBN No:978-1-5386-8492-4/18/$31.00 ©2018 IEEE

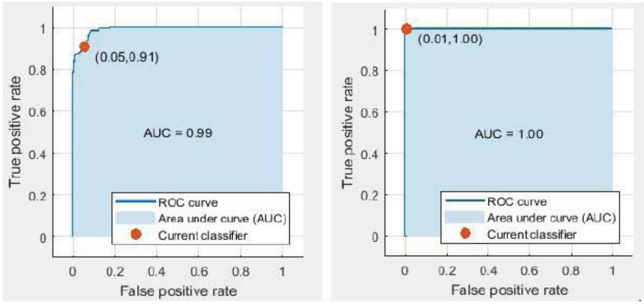


Fig.4. ROC plots of Decision (left) and SVM (right)

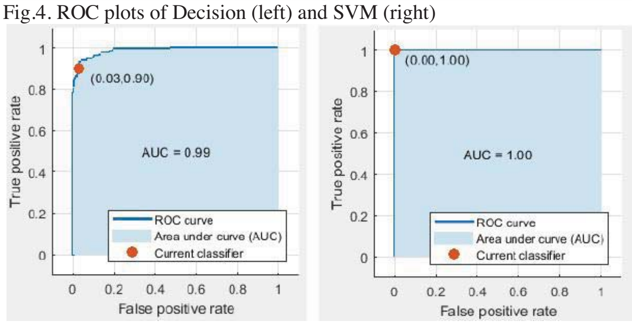


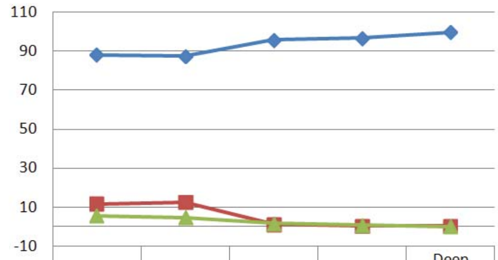
Fig.5. ROC plots of Ensemble (left) and weighted KNN (right)

C. Estimation of accuracy

Accuracy in classification problems is the number of correct predictions made by the model over total predictions made. Accuracy metric is a function of the false positives (FP), false negatives (FN), true positives (TP) and true negatives (TN). It can be calculated by the equation given below.

$$\text{Overall accuracy } = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad (1)$$

Fig.6. shows the accuracy for different machine learning classifiers, higher the accuracy better the performance of proposed detection system. Accuracy can consider as a good measure only when the target variable classes in the data are nearly balanced. But in majority of the practical scenarios data will be unbalanced therefore accuracy should never be used as a measure when the target variable classes in the data are unbalanced. For example, in our driver fatigue detection example with 100 people, only 10 people have experienced fatigue. Lets say our model is very bad and predicts every case as No fatigue.



	Tree	Ensemble	SVM	KNN	Deep Autoencoder
Overall Acuuracy (%)	88.2	87.5	95.8	96.6	99.7
Overall Error (%)	11.8	12.5	1.2	0.4	0.3
Training Time (Sec)	5.5	4.61	1.68	0.86	0.04

Fig. 6. Performance plot for different machine learning classifiers. Overall accuracy= tp+tn/(fp+fn+tp+tn) where where tp: the number of true positives, tn : the number of true negatives, fp : the number of false positives and fn : the number of false negatives.

In doing so, it has classified those 90 non fatigue subjects correctly and 10 fatigue subjects as Non-fatigue. Now even though the model is terrible at predicting fatigue, the accuracy of such a bad model is also 90%.

IV. CONCLUSIONS

In this paper, we attempt to use the different machine algorithms and deep learning algorithms to classify different cognitive state of the driver during the period of 60 min of simulated driving. This study identified the best EEG parameters that are highly correlated with driver fatigue. The results showed that the relative power of theta, alpha, beta and delta showed significant correlation to the driver fatigue. Therefore in proposed model, the selected features were trained and evaluated using 20 classifiers which belongs to the classifier groups such as ; SVM, KNN, Decision tree and ensemble. After identifying the best classifiers, the next step was to find out the effectiveness of each model based on some metric using test datasets. Based on different metrics, autoencoder (99.7%) of deep learning and KNN (96.6%) of machine learning methods were found to be most accurate while classifying different states of the fatigue.

V. REFERENCES

[1] "Ministry of Road Transportation & Highways (MoRTH)," *Road accidents in India,* 2016.

[2] W. H. Organization, *Global status report on road safety 2015*: World Health Organization, 2015.

[3] I. D. Brown, "Driver fatigue," *Human factors,* vol. 36, pp. 298-314, 1994.

[4] S. K. Lal and A. Craig, "A critical review of the psychophysiology of driver fatigue," *Biological psychology,* vol. 55, pp. 173-194, 2001.

[5] S. M. Marcora, W. Staiano, and V. Manning, "Mental fatigue impairs physical performance in humans," *Journal of Applied Physiology,* vol. 106, pp. 857-864, 2009.

[6] P. M. Forsman, B. J. Vila, R. A. Short, C. G. Mott, and H. P. Van Dongen, "Efficient driver drowsiness detection at moderate levels of drowsiness," *Accident Analysis & Prevention,* vol. 50, pp. 341-350, 2013.

[7] Z. Zhang and J. Zhang, "A new real-time eye tracking based on nonlinear unscented Kalman filter for monitoring driver fatigue," *Journal of*

Control Theory and Applications, vol. 8, pp. 181-188, 2010.

[8] V. Balasubramanian, & Bhardwaj, R, "Grip and Electrophysiological Sensor Based Estimation of Muscle Fatigue while Holding Steering Wheel in Different Positions," *IEEE Sensors Journal,* 2018.

[9] V. Balasubramanian and R. Bhardwaj, "Can cECG be an unobtrusive surrogate to determine cognitive state of driver?," *Transportation research part F: traffic psychology and behaviour,* vol. 58, pp. 797-806, 2018.

[10] R. Bhardwaj and V. Balasubramanian, "Driver's Cardiac Activity Performance Evaluation Based on Non-contact ECG System Placed at Different Seat Locations," in *Congress of the International Ergonomics Association*, 2018, pp. 278-285.

[11] P. Davidson, R. Jones, and M. Peiris, "Detecting behavioural microsleeps using EEG and LSTM recurrent neural networks," in *Proc. 27th Int. Conf. IEEE Eng. Med. Biol. Soc*, 2005.

[12] F.-C. Lin, L.-W. Ko, C.-H. Chuang, T.-P. Su, and C.-T. Lin, "Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system," *IEEE Transactions on Circuits and Systems I: Regular Papers,* vol. 59, pp. 2044-2055, 2012.

[13] A. M. Al-kaysi, A. Al-Ani, and T. W. Boonstra, "A multichannel deep belief network for the classification of EEG data," in *International Conference on Neural Information Processing*, 2015, pp. 38-45.

[14] K. T. Chui, K. F. Tsang, H. R. Chi, B. W. K. Ling, and C. K. Wu, "An accurate ECG-based transportation safety drowsiness detection scheme," *IEEE Transactions on Industrial Informatics,* vol. 12, pp. 1438-1452, 2016.

[15] L.-H. Du, W. Liu, W.-L. Zheng, and B.-L. Lu, "Detecting driving fatigue with multimodal deep learning," in *Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference on*, 2017, pp. 74-77.

[16] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science,* vol. 313, pp. 504-507, 2006.

[17] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning* vol. 1: MIT press Cambridge, 2016.

[18] K. Dwivedi, K. Biswaranjan, and A. Sethi, "Drowsy driver detection using representation learning," in *2014 IEEE International Advance Computing Conference (IACC)*, 2014, pp. 995-999.

[19] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431-3440.

[20] M. Hajinoroozi, T.-P. Jung, C.-T. Lin, and Y. Huang, "Feature extraction with deep belief networks for driver's cognitive states prediction from EEG data," in *Signal and Information Processing (ChinaSIP), 2015 IEEE China Summit and International Conference on*, 2015, pp. 812-815.

[21] G. H. Klem, H. O. Lüders, H. Jasper, and C. Elger, "The ten-twenty electrode system of the International Federation," *Electroencephalogr Clin Neurophysiol,* vol. 52, pp. 3-6, 1999.