Optimization of Deep Neural Networks for Modeling Traffic Data using GPS

A.Sampath Dakshina Murthy, Rudra Pratap Das, T.Karthikeyan

Abstract: As accident is increasing continually.GPS created traffic data need to optimal .Deep neural networks having more number of hidden layers are the rest portable to solution .Selecting proper DNNs the main objective . By trial and error proper the selective of DNN has been achieved for optimization of traffic data.

Index Terms: Back propagation, neural network, deep neural network, recurrent neural networks.

I. INTRODUCTION

The damage due to accidents on roads has very significant social impart in personal and national and environmental expenditure. As per WHO (World Health Organization), road fatalities have crossed one million from 2015, 3 years ago. It is pertinent to examine driving styles, risk patterns including the linkage of sensitivity of drivers to complex real life situations. The influence of driving pattern on environment and fuel use has been studied earlier [1].The insurance companies include pay-as-you-drive and pay-low-you-drive for billing which need to consider the aggressive quality, if any, of the person drivering.

GPS is the preferred device for actual position data [2]. The main advantages of GPS sensor data are 1) Large scale of sampling 2) Undeterred data collection 3) Continuous real time data set. Data fusing has been used to combing GPS data, gyroscope and accelerometer for the objective of signal processing and filtering [3].

Traditionally, statistical algorithms have been used to study the driving behavior. However, computer intelligence methods have primary depended on Deep learning methods. It is the objective of this research work to work on modeling through ANN for predicting driving patterns.

deep neural networks based on time and space with no supervisor for training can be developed learning drivers behavior patterns. This approach covers a) Features learning by auto encoders and b) mapping methods of clustering.

Revised Manuscript Received on 30 March 2019.

* Correspondence Author

A.Sampath Dakshina Murthy, Department of Electronics and Communication Engineering, Vignan's Institute of Information Technology, Visakhapatnam, A.P, India.

Rudra Pratap Das, Department of Electronics and Communication Engineering, Vignan's Institute of Information Technology, Visakhapatnam, A.P, India.

T.Karthikeyan, Department of Electronics and Communication Engineering, Vignan's Institute of Information Technology, Visakhapatnam, A.P, India.

The main advantages of GPS sensor are: a) Real time and continuous data. b) Large scale sampling .c) Data collection without obstruction .Two disadvantages are high level of noise and heterogeneous data source. Data fusing is possible for integrity GPS data, gyroscope and acceleration. For filtering and smoothening Smartphone data can also be combined. Objectives of Research work

- Design development of DNN for predicting driving patterns using GPS.
- To have a case study of such patterns with respect to metropolitan cities.
- To have simplified package for specific vehicles with GPS /Smartphone data.

II. PROPOSED METHODOLOGY OF NEURAL NETWORKS

In comparison with statistical averaging or regression by Kalman filtering [4] ANNs (Artificial Neural Networks) help in machine learning methods suitable for predicting traffic situations. Basically, ANNs learn from given examples of establish implicit functional relationships in parameters of data even though the internal relationships are unexplored on too hard to comprehend in figure 1. For problems whose solutions require information that cannot be easily specified, ANNs are most suitable even the unseen portions can be visualized in noisy information. To have maximum benefit it ANNs of is essential to have sufficient [Zhang, 1998] capauts of hidden layers.

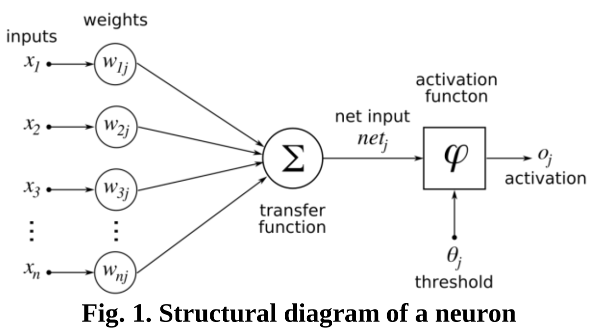


Fig. 1. Structural diagram of a neuron

Hidden layer

The hidden layers generate the weights of the connection to the output layer of the bias parameters bias during training. The hidden nodes allow the neural network to detect the future of capture pattern in the data suitable for performing non linear mapping. A single hidden layer is sufficient to handle any non linear function. The number of nodes in hidden layer in normally selected by experiments.

In many cases if is taken as half of sum of input neurons and output neurons. ANN (artificial neural networks) has a lot of important applications such as classification, associative memory, control and perception. Inspired by the biological neurons behavior in terms of dendrites, axon and change of state ANNs need to be properly implemented to attain level of human operations. A simple neuron can be described by the schematic shown in the figure 1.

The synaptic weights are represented by W's where as x_1, x_2……..x_n give the inputs. The condition of firing is decided by the inequality $\sum w_i\, x_i$>reference value of neuron. Beyond this the neuron sends out a signal equivalent to the signal through axon as in biological equivalent.

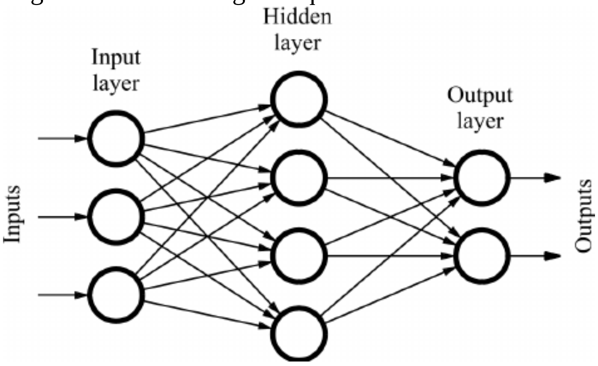


Figure. 2. Layered feed forward neural network

Figure 2 shows a typical case of layered feed forward neural network in which the hidden layer represents the additional neural structure involved in the operation. There can be many hidden layers as in deep neural network (DNN).

For utilizing ANNs in dealing with non linear structures, an activation function is supposed to be applied to the output of the neuron. There are many types of activation functions the popular choice being sigmoid function represented by output equal to inverse of $(1+e^{-x})$.

Back Propagation

As per the error between actual values of the solution of the output the weights of hidden-output, input-hidden layers have to be adjusted in backward movement. Taking w_{ji} as the synaptic weight connecting i^{th} input to the j^{th} neuron the objective of any program is to improve w_{ji} so that the mean square error is minimized. The average error ε_{avg} of the output of neuron j at iteration n for N no of examples in the training set is shown in figure 3.

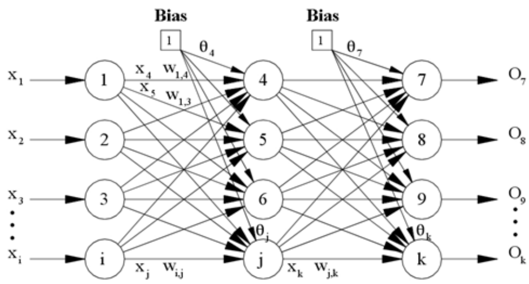


Figure 3 Back propagation neural network

$$\varepsilon_{avg} = \frac{1}{2N}\sum_{n=1}^{N}\sum_{j}(d_j(n) - y_i(n))^2 \tag{1}$$

Where $d_j(n)$ the desired output. The back propagation algorithm applies collective weight $\Delta\omega_{ji}$ which is proportional to the partial derivative. $\frac{\partial \varepsilon_{avg}}{\partial \omega_{ji}}$ of error with respective w_{ji}. Take η as the learning rate parameter $\Delta\omega_{ji} = -\eta \frac{\partial \varepsilon_{avg}}{\partial \omega_{ji}}$. the typical value of η are between 0.001 to 1.0. For increasing the rate of learning and avoiding instability. (Rumelhart 1986) suggested a momentum term.α.

$$\Delta\omega_{ji}(n) = \alpha\Delta\omega_{ji}(n-1) + \eta\delta_j(n)X_i(n) \tag{2}$$

The $\delta_j(n)$ gradient is defined by equation

$$\delta_j(n) = -\frac{\partial \varepsilon_{avg}}{\partial\left(\sum_{i=1}^{m} w_{ji}(n)X_i(n) + b_j\right)} \tag{3}$$

Neural network having more than one hidden layer is termed as Deep Neural Networks (DNN) figure 4. RNN (Recurrent Neural Networks) are capable of knowing dynamic classification of algorithm use for recognizing time series patterns in many domains in figure 5 [4]. Most successful deep learning neural networks use filter to find relationships between neighboring inputs [5]. ANNs work better when the models based on statistics are not valid. Unsupervised learning mechanism such as Auto encoder, Restricted Boltzmann machine and long short term memory network are capable of studying graphic related features [6].

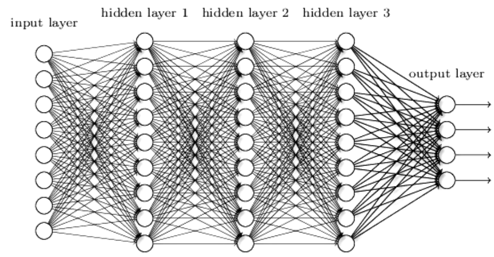


Figure 4. Deep Neural Networks

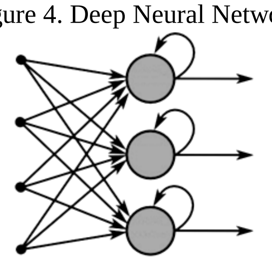


Figure 5 .Recurrent Neural Network

Design of input features

Deep Neural Networks based structure

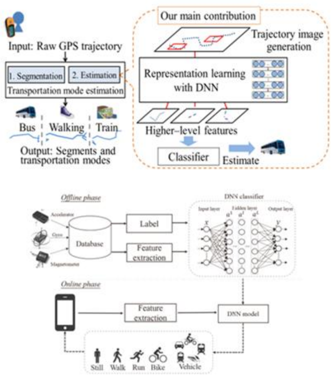


Figure 6. Deep Neural Networks based structure

From the figure 6 the objective of design is to have sequential information from multiple sensor to be given to DNN for densification .Estimation process has two phases: During offline phase .feature extraction is done by integrating sensor data into diversified features to train the DNN .In this online phase the testing data are converted into feature vectors .The previously trained DNN is fed with such data.

For DNN model with L hidden layers the output is

$$a^i = f(w^l a^{l-1} + b^l) \tag{4}$$

The ReLU function is adopted as the activation function

$$f(x)=\max(0,z) \tag{5}$$

Another function is placed on the top of the L^{th} hidden layer to perform classification (softmax function). A back propagation technique is used for further processing of weights.

It is evident that the driving patterns can be used to predict and classifier drivers as per the driving by using input data and algorithm [7,[8]. Classification can be done aggression, attention, drinking and sluggishness. Using a reference as preset driving profile unsafe drivers can be detected [9]. By extracting features in on supervised mode without early knowledge , the basic aim is to adopt multiplier encoder for transforming raw data into a code of low dimension which can be detected easily driving pattern with aggression are lack of attention are related to in consistence acceleration are reducing of speed .Risky speeding patterns and unwanted vehicle position are taken up as input data for GPS . Such data gives detail of vehicle orientation in degree/sec. Modeling by clustering methods can be used to transform features related to turning behavior.

Speed parameters

Parameters such as speed limits, mean speed and speed variants are examined to define underlying equations [10]. For $V_{max,}$ a speed v_f is defined as threshold speed when there is an over speed tendency where

$$\delta = \frac{T_f}{T_c} * 100\% \tag{6}$$

Where T_c the total travelling is time of this vehicle on road; T_f correspond to speed exceeds v_f. Moreover there is a clear correlation between road accident and speed variation. The mean and unbiased estimation of standard derivation of speed are covered in the correlation.

$$v_a = \frac{1}{n}\sum_{m=1}^{n} v_m \tag{7}$$

$$v_s = \sqrt{\frac{1}{n-1}\sum_{m=1}^{n} (v_m - v_a)^2} \tag{8}$$

Where v_m the momentary speed of the vehicle is collected by GPS sensor at the time m; n is the sample size.

Parameters related to Acceleration

The main parameters which need to incorporated related to acceleration, which are the unbiased estimation of standard deviation $a_s^+(a_s^-)$, positive (negative) standard deviation $a_m^+(a_m^-)$ of acceleration.

$$a_a = \frac{1}{n}\sum_{m=1}^{n} a_m \tag{9}$$

$$a_s = \sqrt{\frac{1}{n-1}\sum_{m=1}^{n} (a_m - a_a)^2} \tag{10}$$

$$a_m^{\pm} = \frac{v_m - v_{m-1}}{t} \tag{11}$$

$$a_a^{\pm} = \frac{1}{n}\sum_{m=1}^{n} a_m^{\pm} \tag{12}$$

$$a_s^{\pm} = \sqrt{\frac{1}{n-1}\sum_{m=1}^{n} \left(a_m^{\pm} - a_a^{\pm}\right)^2} \tag{13}$$

Where is the instantaneous acceleration of the vehicle collected by GPS at the time m; a is the average acceleration of the vehicles in the dataset.

Thus, on an experimental basis, we first selected eight features transformed from seque3nces of raw GPS data [11-13] to serve as the behavior-to-vec X of drivers, where X is constructed as follows:

$$X = [\delta, v_a, v_s, a_a, a_a^+, a_s^+, a_a^-, a_s^-] \tag{14}$$

Learning Features with Auto Encoder

The auto encoder perform equally well in principal component analysis and compression scheme for the cases with hidden layer smaller than IO/OP layer having linear activation [14].Current research as proved that nonlinear auto encoders can classify specific types of multimodal and nonlinear domains precisely, relatively deeper connection between variable are seen .

To enhance the performance of an auto encoder, the input vector needs to be element-wise normalized to

$$x_i = \frac{x_i - \bar{x}_i}{\sigma_i} \tag{15}$$

Where x_i is the average value of each feature and is the variance. We take the normalized vector X as the input of the auto encoder and define V as the behavior training set which is ready to feed into the auto encoder with a size of M.

$$V = \{X_1, X_2, \ldots X_N\}^T, n \in M \tag{16}$$

Where n is the size of the dataset. Design of the model based on multiple layer structure (>=3) as the basic objective of compressing and extracting latent features such as regression and classifications. First layer gives the input the intermediate layers is the hidden layers are hidden layers and finally last one will be output layer.

III. RESULTS AND DISCUSSION

Referring to figure 7, the inputs to ANFIS are used speed acceleration and auto encoder which is linked with DNN .The training data of both the inputs are available in table 1 and 2 .The start of training has in figure 7 (a) and (b), in figure 7(c) and (d) checking data figure 7 (e) error analysis of the behavior patterns.

The procedure includes the 18 fuzzy rules shown in figure 9 .Number of nodes is 58, linear parameters are 18, nonlinear parameters 24 .Number of training data pairs is 30 where checking data pairs are 17.Training is completed in epoch 2. The output of ANFIS is shown in figure 8.The choice of membership function is also found to be correct as less time the result has been obtained rms error in table 3 . It can be inferred that the basic objective of utilizing ANFIS using DNN has been successfully parried for driving behavior patterns.

Table 1 Data sheet of GPS Training data traffic speed acceleration and auto encoder values

Speed	Acceleration	Auto encoder	Projected value
0	0.2	0.5	96.2
0	0.3	0.5	87.7
0	0.4	0.5	74.9
0	0.5	0.5	60.3
0	0.6	0.5	58
0	0.7	0.5	60.1
0	0.8	0.5	101
0	0.9	0.5	93.75
0	1	0.2	85.98
0	0.2	0.2	75.9
0	0.4	0.2	91.64
0	0.5	0.2	89.46
0	0.6	0.2	70.44
0	0.9	0.3	73.44
1	1	0.3	93.4
1	0.2	0.3	79.926
1	0.3	0.3	78.676
1	0.4	0.4	77.426
1	0.5	0.4	76.176
1	0.6	0.4	74.926
1	0.7	0.4	73.676
1	0.8	0.4	72.426
1	0.9	0.5	71.176
1	1	0.5	69.926
1	0.2	0.5	68.676
1	0.3	0.5	67.426
1	0.4	0.5	66.176
1	0.5	0.5	64.926
1	0.6	0.5	63.676
1	0.8	0.5	62.426

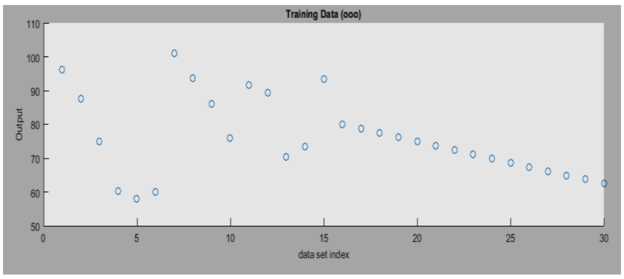
In the above table the driving behaviors for different parameters of the vehicles in traffic scenario.

Table 2 Data sheet checking data

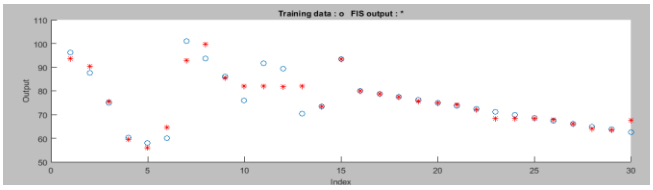
Speed	Acceleration	Auto encoder	Projected value
0	0.2	0.5	96.2
0	0.5	0.5	87.7
0	0.5	0.8	74.9
0	0.6	0.9	60.3
0	0.6	0.9	58
0	0.7	0.3	60.1
0	0.8	0.5	75
0	0.9	0.5	93.75
0	1	0.2	85.98
0	0.2	0.2	75.9
0	0.4	0.2	91.64
0	0.5	0.2	89.46
0	0.6	0.2	70.44
0	0.9	0.3	73.44
1	1	0.3	93.4
1	0.2	0.3	79.926
1	0.3	0.3	78.676

Table 3 representing the combination of ANFIS and DNN values

Parameters	ANFIS	DNN	ANFIS+DNN
Training data	4.3941	0.269	0.259
Testing data	4.48896	0.468	0.258
Checking data	4.3941	0.489	0.256
RMS error	3.585873	0.358	0.256

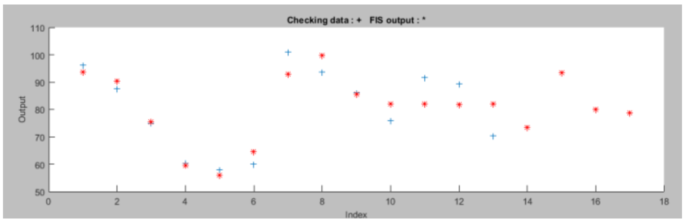


(a)

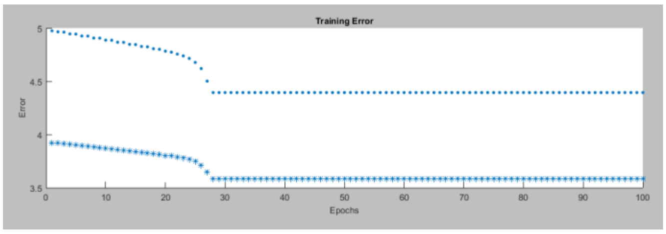


(b)

(c)



(d)



(e)

Figure 7 (a) Training data ANFIS+DNN (b) Training data DNN+ANFIS (c) Checking data ANFIS+DNN (d) Checking data DNN+ANFIS (e) ANFIS error data

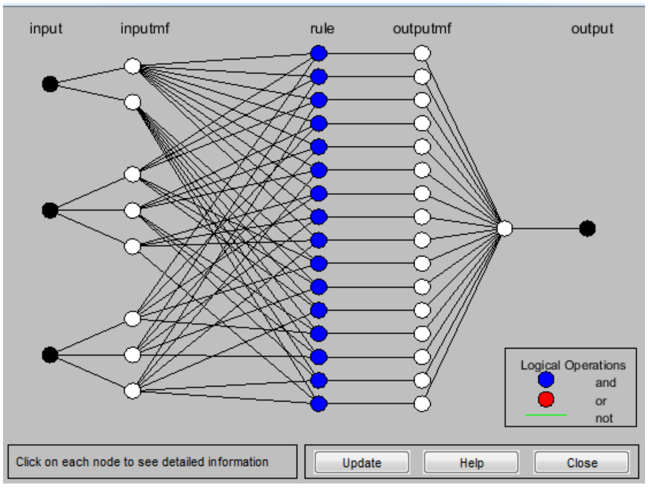


Figure 8 Structures of ANFIS

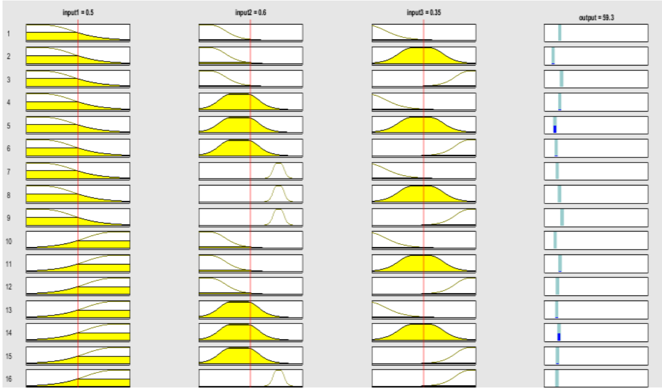
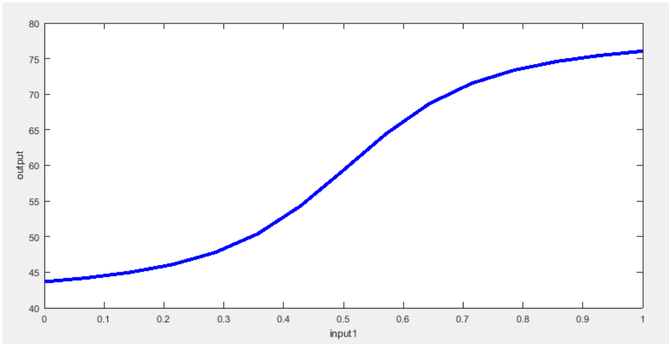
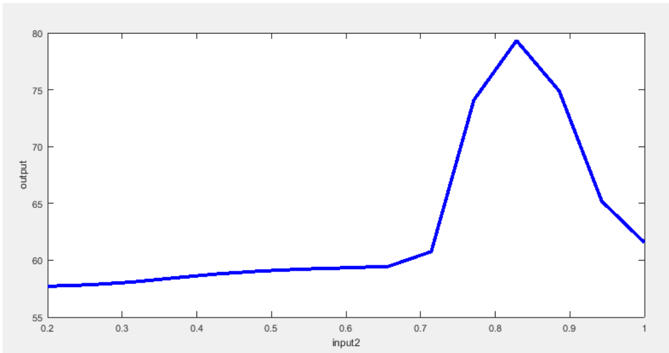


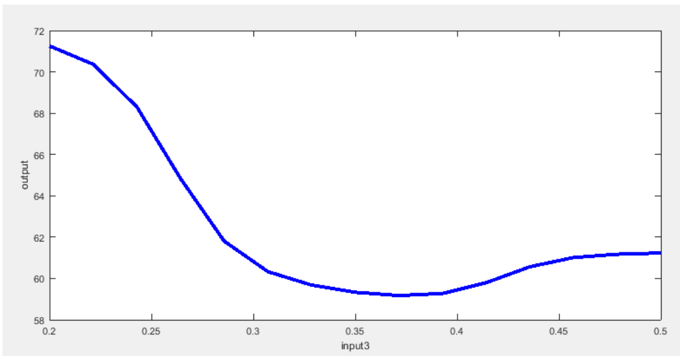
Figure 9 Rules viewer for driving behaviors



(a)



(b)



(c)

Figure 10 (a) speed (b) Accreleation (c) autoencoder

IV. CONCLUSION

As per the objective, design development of DNN suitable for driving pattern has been created out from the case studies. It has been found that auto encoder based compression scheme works equally for PCA .Overall; DNN method of optimization has been formed to be suitable for traffic data of GPS.

REFERENCES

1. 1.Bian, Yiyang, Chen Yang, J. Leon Zhao, and Liang Liang. "Good drivers pay less: A study of usage-based vehicle insurance models." Transportation research part A: policy and practice 107 (2018): 20-34.
2. 2.Liang, Yi, Zhipeng Cai, Qilong Han, and Yingshu Li. "Location privacy leakage through sensory data." Security and Communication Networks 2017 (2017).
3. 3.Yao, Shuochao, Shaohan Hu, Yiran Zhao, Aston Zhang, and Tarek Abdelzaher. "Deepsense: A unified deep learning framework for time-series mobile sensing data processing." In Proceedings of the 26th International Conference on World Wide Web, pp. 351-360. International World Wide Web Conferences Steering Committee, 2017.
4. 4.Figo, Davide, Pedro C. Diniz, Diogo R. Ferreira, and João M. Cardoso. "Preprocessing techniques for context recognition from accelerometer data." Personal and Ubiquitous Computing 14, no. 7 (2010): 645-662.
5. 5.LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521, no. 7553 (2015): 436.
6. 6.Ma, Xiaolei, Zhuang Dai, Zhengbing He, Jihui Ma, Yong Wang, and Yunpeng Wang. "Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction." Sensors 17, no. 4 (2017): 818.
7. 7.Ricardo, Leandro Jorge Correia. "Decision support system for city public transportation." PhD diss., Universidade de Aveiro, 2017.
8. 8.Meiring, Gys Albertus Marthinus, and Hermanus Carel Myburgh. "A review of intelligent driving style analysis systems and related artificial intelligence algorithms." Sensors 15, no. 12 (2015): 30653-30682.
9. 9.Wang, Jianqiang, Meng Lu, and Keqiang Li. "Characterization of longitudinal driving behavior by measurable parameters." Transportation Research Record: Journal of the Transportation Research Board 2185 (2010): 15-23.
10. 10. Cao, Lili, and John Krumm. "From GPS traces to a routable road map." In Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems, pp. 3-12. ACM, 2009.
11. 11. A.Sampath Dakshina Murthy, T.Syamala , S.Koteswara Rao, Dr. Rudra Pratap Das."Real Time Differential Global Positioning System Using Fuzzy Logic Extended Kalman Filter."Journal of Advanced Research in Dynamical and Control Systems" Vol. 9. Sp– 14 / 2017.
12. 12. A.Sampath Dakshina Murthy, S.Koteswara Rao ,"Application Of Fuzzy Logic Based Kalman Filter and Vehicle Rate Sensor in Optimizing Differential Global Position System". Journal of Advanced Research in Dynamical and Control Systems Vol. 9, Issue 2, OCT.2017.
13. 13. Dr. Rudra Pratap Das, G. Roja, A.Sampath Dakshina Murthy, Dr.S.Koteswarao Rao."Global Positioning System Object Tracking by Applying Fuzzy Logic Non Linear Techniques". Journal of Advanced Research in Dynamical and Control Systems Vol. 9. Sp– 14 / 2017
14. 14. Japkowicz, Nathalie, Stephen Jose Hanson, and Mark A. Gluck. "Nonlinear auto association is not equivalent to PCA." Neural computation 12, no. 3 (2000): 531-545.

AUTHORS PROFILE

Mr. A.Sampath Dakshina Murthy received his B.Tech in Electronics and Communication Engineering Degree in 2013 and M.Tech Degree in Digital Electronics and Communication Engineering 2015. He is currently working as assistant professor in Department of Electronics and Communication Engineering at vignan's institute of information technology, Visakhapatnam and also pursuing Ph.D in Electronics and Communication Engineering from Koneru Lakshmaiah Education Foundation Deemed University in Guntur, Andhra Pradesh. His research interest includes Image & Video Processing, Fuzzy Logic and Neural Networks. He has published 18 papers in International Journals. He is a Life member of I.S.R.S and AMIE. MIET.

Dr Rudra Pratap Das has extensive experience in the field of research, engineering and technological application. With over 20 years of academic and 11 years of industrial experience in India, He is currently working as professor in Department of Electronics and Communication Engineering at vignan's institute of information technology, Visakhapatnam he has contributed and applied technical acumen towards solving problems in different fields, ranging from Semiconductor Devices, Geo-physical Exploration to Neural Networks. Mr. Das, based on his technical experience and contributions intends to be successfully elevated to IEEE Senior Member. He published a book on "Neural networks and fuzzy logic".

Dr.T. Karthikeyan received his B.E. Degree in Electronics and Communication Engineering (2006) and M.E., Degree in Embedded System Technologies (2012) from Anna University and PhD in Wireless sensor Networks(2018) from KAHE, Coimbatore respectively. He has total of 11 Years of Experience out of which 2 years in the field of Software Testing. He is currently working as Associate Professor in the Department of Electronics and Communication Engineering (DST-FIST sponsored) at KLEF, Guntur. He is a life member of ISTE. He has published more than 8 research papers in the Leading Journals and Conferences respectively. His area of interest includes Wireless Sensor Networks, MANET, Wi-MAX, Wi-Fi and IoT.