INTERNATIONAL JOURNAL OF INTEGRATED ENGINEERING
ISSN: 2229-838X e-ISSN: 2600-7916

IJIE

Vol. 16 No. 9 (2024) 302-313
https://publisher.uthm.edu.my/ojs/index.php/ijie

A Hybrid Deep Learning Model for Detecting Driver Fatigue Using Electroencephalogram Signals

Rafiuddin Abdubrani[1]*, Mahfuzah Mustafa[2], Zarith Liyana Zahari[2]

[1] *Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan, Pahang, 26600, MALAYSIA*

[2] *Electronic Section, Universiti Kuala Lumpur British Malaysian Institute, Gombak, Selangor, 53100, MALAYSIA*

*Corresponding Author: rafi128@yahoo.com
DOI: https://doi.org/10.30880/ijie.2024.16.09.025

Article Info

Received: 27 March 2024
Accepted: 4 September 2024
Available online: 30 December 2024

Keywords

Electroencephalogram (EEG), hybrid deep learning, deep belief network (DBN), recurrent neural network (RNN)

Abstract

Road accidents caused by driver fatigue are a significant public safety concern, and detecting driver fatigue is crucial for preventing such incidents. Existing methods for detecting driver fatigue are limited in their effectiveness, and there is a need for more accurate and reliable methods. This study presents a solution to the problem of accurately identifying driver fatigue using electroencephalogram (EEG) signals. The approach involves the development of a hybrid deep learning model that incorporates both a deep belief network (DBN) and a recurrent neural network (RNN). We trained and evaluated our model on a dataset of EEG signals collected from drivers in normal and fatigued states. The effectiveness of the hybrid model in accurately categorizing driver fatigue was evaluated compared to two other classifiers. The study results indicate that the hybrid model outperformed the other classifiers in accuracy, sensitivity, specificity, precision, and F1 score, suggesting superior performance. We observed that the model's accuracy and loss remained consistent even when the number of epochs was low, indicating that the model effectively learned to classify EEG signals and did not overfit the training data. Further evaluation of the hybrid model with varying numbers of epochs revealed that the optimal number for the model was 50. Additionally, analysis of the loss function during training demonstrated that the model effectively learned to classify EEG signals without overfitting the training data. The proposed hybrid model achieved an overall accuracy of 99.98%, with perfect sensitivity (100%) and high specificity (99.95%), precision (99.95%), recall (100%), and F1 score (99.98%). These results indicate that the proposed hybrid deep learning model outperformed the individual DBN and RNN models in classifying EEG signals. Our study's results demonstrate the potential of the proposed hybrid model to accurately detect driver fatigue, which could contribute to the development of more effective and reliable methods for preventing road accidents caused by driver fatigue.

1. Introduction

Recent studies have highlighted the importance of effective detection methods for driver fatigue. For instance, electroencephalography (EEG) was utilized to analyze brain networks associated with fatigue, demonstrating the potential of machine learning techniques to identify driver states in real-time [1]. Similarly, a real-time fatigue detection algorithm was developed based on facial video sequences, emphasizing the role of visual monitoring in preventing accidents caused by drowsy driving [2]. Therefore, accurately predicting driver fatigue is crucial for improving road safety and economic productivity.

This study defines normal and fatigue states based on specific behavioral and physiological criteria observed during driving tasks. The normal state is characterized by sustained focus, consistent reactions, and higher brain activity in regions responsible for attention. In contrast, the fatigue state is marked by slower reactions, lapses in concentration, and changes in brain wave patterns, such as increased theta and alpha activity. These states were further validated by participants' self-reported fatigue levels, ensuring reliable classification for analysis. Various methods have been proposed for fatigue prediction, including subjective measures such as self-reporting and objective measures such as electroencephalography (EEG) and heart rate variability (HRV) analysis [3]. However, these methods have limitations, such as being intrusive or expensive. Recently, signal processing techniques such as wavelet analysis and modified z-score algorithms have emerged as potential tools for fatigue prediction. Wavelet analysis involves decomposing a time-series signal into a set of wavelet coefficients, which can then be used to extract features indicative of fatigue. The modified z-score algorithm, on the other hand, is a statistical method for identifying outliers in a data set, which could be useful for detecting abnormal fatigue levels.

This paper proposes an approach to accurately classify driver fatigue using hybrid deep learning, combining deep belief network (DBN) and recurrent neural network (RNN). Specifically, we aim to validate the effectiveness of the Morlet wavelet and modified z-score algorithm for accurately predicting driver fatigue conditions. We will enhance these methods with modifications based on our research to improve their predictive power. To evaluate our proposed approach, we will collect data from a group of drivers, analyze the data using the Morlet wavelet and modified z-score algorithm, and compare the accuracy of our enhanced methods with the standard methods. By demonstrating the efficacy of these techniques, we hope to contribute to developing more effective and accessible tools for predicting driver fatigue and improving road safety.

In this paper, we will first provide a brief overview of the existing literature on fatigue prediction and the limitations of current methods. Next, we will describe our proposed approach for accurately classifying driver fatigue using hybrid deep learning and enhanced signal processing techniques. We will then present the results of our analysis, compare the accuracy of our proposed approach with standard methods, and discuss the implications of our findings.

2. Related Works

This section summarizes deep learning and machine learning studies in various research areas. The initial studies refer to driver action analysis and alertness detection using an EEG-based driver state recognition system employing automatic binary classification strategies for the brain signals. The second part of the research is employing deep learning models in application to medical imaging, where we discussed two categories: brain tumor segmentation and classification using MRI data (image-level), Alzheimer multiple class classification problem, and outcome-prognosis prediction, which uses neuroimaging data. Additionally, deep learning-based techniques are used in other areas, such as texture representation models and classification of activities using ambient sensors or land use/land cover.

Another study used an immersive driving simulator to assess drivers' behavior in controlled trials over 90 minutes [4]. Another developed a model incorporating the statistical size distribution of grains and various manufacturing defects, such as unfused powder particles and pores that occupy space in the microstructure within an additively manufactured part with polished surfaces. The purpose of the model was to predict a tired life in material [5]. Other researchers have discussed the EEG driver state detection systems and analysis algorithms in the last three decades [6]. Tawhid et al. presented their novel binary classification framework in 2023 for analyzing multi-channel EEG recordings and extracting insightful brain signals [7]. Meanwhile, Ayaz and colleagues created an easy-to-use, commercially available wearable EEG device that determines vigilance states through purely non-invasive frontal brain recordings [8]. Rather than focus on sophisticated technological developments, Cheng explored facial feature recognition for designing a fatigue detection system in 2022 through a basic and practical approach [9]. That same year, Yarici partnered with peers to conduct a telling statistical analysis of the typical shift in-ear EEG measurements from an alert to an exhausted mental state. They rigorously tested an automatic detector trained on scalp and ear EEG readings collected through a dull simulated driving experiment involving ten participants to identify fatigue [10]. Other influential work includes Wang et al. (2022) [11], Li et al. (2022) [12], and Othmani et al. (2023) [13].

In their study, Raju et al. (2022) also proposed a new method for brain tumor segmentation and classification that uses a deep belief network coupled with a hybrid active contour model, which they call harmony cuckoo search-based deep belief network (HCS-DBN) (14). A modified system for speech recognition named a hybrid deep neural network-hidden Markov model HMM-DNN was used by Cao et al.; its acoustic modeling units were built on the neural network, and it presented the unification of different systems [15]. Ahmed et al.'s (2023) review was a kind of popular science on the advancements in deep learning. They began with an introduction to deep neural networks (DNN) [16]. Arya et al. (2023) used deep learning to study the impact of Alzheimer's. They performed a systemic review of publications that use deep learning and neuroimaging data for diagnostic classification or prognostic prediction, systematically summarising research in this area [17]. Li et al.'s (2023) model was based on deep learning and used spatial and frequency characteristics to represent texture [18]. Han et al. (2022) created a new PolSAR classification method called the Restricted Boltzmann Machine (RBM) by building upon the generic deep belief network (DBN) [19]. Recent studies have also explored the application of RNNs in various domains, including healthcare and environmental science. For example, Robertson et al. (2023) utilized RNNs to predict microbial growth, showcasing their capability to handle complex biological data [20]. Similarly, the research by Hu et al. (2022) demonstrates the application of RNNs in predicting economic indicators post-COVID-19, indicating their versatility across different fields [21]. Moreover, integrating RNNs with other deep learning models, such as CNNs, has been proposed to enhance prediction accuracy and model robustness [22,23].

3. Methodology

This study proposes a novel model for analyzing EEG data involving several processing stages. Fig. 1 illustrates the data for developing the proposed model, including the EEG signals' segmentation and pre-processing. The processed data is then fed into three classifier models, deep belief network (DBN), recurrent neural network (RNN), and a hybrid DBN-RNN model. Finally, the performance of each model is evaluated using various metrics to determine their effectiveness in classifying EEG data. This methodology provides a comprehensive approach to analyzing EEG data and can be used in various applications such as brain-computer interfaces and clinical diagnosis.

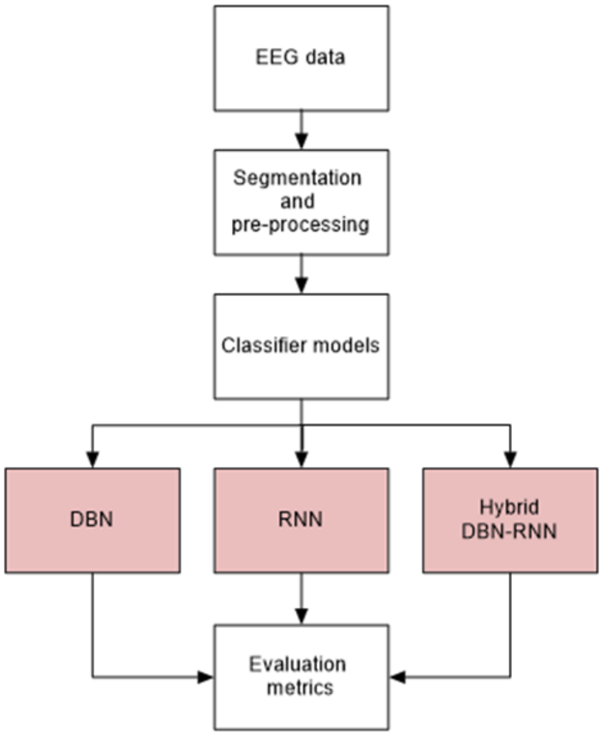


Fig. 1 *The illustration of data used to develop the proposed model*

3.1 Subjects and Experiment Procedure

Driver fatigue was detected using an online EEG dataset [24]. Twelve healthy young men aged 19 to 24 were given a briefing on the procedures and goals of the experiment before data processing took place. This involved a 5-minute driving task that they performed. The experimental procedure began in a controlled laboratory setting with a static driving simulator and a 24-inch screen. Subsequently, participants undertook a driving task requiring sustained attention on a low-traffic roadway, starting at 9 a.m. After a 5-minute training session, participants were given 10 minutes to do whatever they liked in the laboratory. Participants underwent two stages of EEG recordings for 20 minutes; the first was when participants were driving in a controlled environment and 5 minutes of their last stage from the normal state. Participants engaged in continuous driving for 40 to 100 minutes until they reported feeling fatigued; the last 5 minutes of EEG are the fatigued state. A Neuroscan device collected EEG data from 32 electrodes with a sampling rate of 1000Hz; data from eight specific channels (Fp1, Fp2, O1, O2, F4, F3, P4, P3) were selected after a brief check of all equipment for analysis. The EEG channels of Fp1, Fp2, O1, O2, F4, F3, P4 and P3 were selected since they represent significant brain regions related to cognitive function, sensation perception and motor commands. In particular, they trace movements occurring in the frontal, parietal, and occipital lobes of the brain, which are involved in decision-making, attention, visual processing and spatial orientation, such as setting the human compass to point east and, therefore well, well-suited to examining almost any neural process [24, 25].

3.2 Pre-processing

The EEG signals used in this study were pre-processed to remove noise and artifacts and extract meaningful information. Independent component analysis (ICA) was applied to the pre-processed data to accomplish this. ICA separated the signals into independent components representing different sources of neural activity that contributed to the overall signal. Artifactual components, such as those related to eye movements or muscle activity, were identified by their spatial and temporal characteristics and removed from the data. The remaining independent components were combined to reconstruct the original EEG signal for subsequent analysis. Using ICA as part of the pre-processing pipeline improved the quality and accuracy of the EEG data, and insights were gained into the underlying neural sources of the measured activity.

The signals were filtered using a Butterworth filter design. Three different frequency bands were extracted using bandpass filters with different cutoff frequencies. First, an alpha bandpass filter was applied with a range of 8 to 13 Hz. The filter was designed using a transition width of 1 Hz and an attenuation of 40 dB. Second, a delta bandpass filter was applied with a range of 0.5 to 4 Hz, using a transition width of 0.4 Hz and an attenuation of 40 dB. Finally, a theta bandpass filter was applied with a range of 4 to 7 Hz, using a transition width of 1 Hz and an attenuation of 40 dB. The filtered signals were stored in separate variables for each frequency band and were used for subsequent analyses.

A continuous wavelet transform (CWT) was applied to compute the energy of the pre-processed EEG signals in the alpha, delta, and theta frequency bands.

$$CWT\{h(x)\} = S_{CWT}(x,\omega) = \int_{-\infty}^{\infty} h(\varepsilon)\psi * \big((\varepsilon - x)\omega\big)d\varepsilon \quad (1)$$

H(x) is the signal to be analyzed, and ω is the scale parameter, which controls the width of the wavelet in the time-frequency plane. The Morlet wavelet is represented as ψ, the complex conjugate of the Morlet function. The integral is taken over all time points, ε, representing the signal's convolution with the Morlet wavelet at a particular scale and position.

The signal was segmented into smaller segments of length 5000 to reduce computational complexity. The number of segments was calculated by dividing the signal by the segment length and taking the floor of the result. The segments were then reshaped and concatenated to form a single segment of length 5000, which was used for subsequent analysis. The CWT coefficients were computed using the Morlet mother wavelet with a frequency range corresponding to the alpha, delta, and theta frequency bands. The resulting coefficients were used to calculate the signal energy in each band, which was used as a feature for subsequent analyses.

After computing the signal energy using CWT, various statistical measures were calculated for each frequency band of interest, including alpha, theta, and delta. Mean, range, variance, median absolute deviation (MAD), root mean square (RMS), peak-to-peak amplitude (PP), and skewness of the distribution of absolute values were calculated for each frequency band. For alpha, the mean, range, variance, MAD, RMS, PP, and skewness of the distribution of absolute values were calculated using the CWT coefficients in the alpha frequency range of 8 to 13 Hz. Similarly, for theta, the same measures were computed using the CWT coefficients in the theta frequency range of 4 to 7 Hz, and for delta, the range was from 0.5 to 4 Hz. The resulting statistical measures were used as features for subsequent analyses.

To address the issue of extreme values influencing the accuracy of identifying outliers using the modified z-score method, Winsorizing was used to clip the data and trim the tails of the distribution. The percentiles of 20 and 80 were chosen as conservative values that remove the most extreme outliers while preserving the majority of the data. This ensures that extreme values do not overly influence subsequent analyses. The fatigue and normal data sets were Winsorized using their respective percentile values.

$$Modified\ Z - Score = k * (Xi - Median)/MAD \tag{2}$$

The median and median absolute deviation (MAD) were then calculated for the concatenated data set. The enhancement of the modified z-score was calculated using a constant value of k and the previously calculated median and MAD values. While the modified z-score method is robust for identifying outliers, it assumes a normal distribution of the data, which is not always the case in real-world datasets. The enhancements made in this study aim to improve the accuracy of outlier detection in datasets with non-normal distributions.

3.3 DBN, RNN, Hybrid DBN-RNN

Hybrid deep learning aims to improve the accuracy of drivers' fatigue classification. Deep belief networks (DBNs) and recurrent neural networks (RNNs) have produced promising outcomes in many machine-learning applications. Nevertheless, each architecture has strengths and weaknesses. DBNs' unsupervised learning and feature extraction ability are a big advantage, as RNNs are best suited for sequence modeling but stand independently in time-dependent data.

To combine the strengths of these two architectures, we propose a hybrid model in which DBNs and RNNs are used together. In particular, DBN is exploited to carry out unsupervised feature learning, followed by applying RNNs for supervised learning that classifies the driver's level of fatigue state. Our approach makes it possible to draw relevant characteristic variables from input data via the DBN and capture relations over time using RNNs to make accurate classifications. The contribution of this paper is to study how effective such a hybrid system proves under various circumstances, as well as what prospects there might be for improving the accuracy of driver fatigue recognition through this method. DBN structure is an artificial neural network composed of multiple layers of neurons.

DBNs consist of two concise arguments: the layer and the hidden formation. The invisible layer serves as input, but the hidden layers are where high-level functions are learned. Each of these hidden layers is trained to produce a compact representation of its inputs, which then serves as the input for the next layer. RNN is an artificial neural network used mostly in deep learning and natural language processing. In contrast with the feed-forward neural networks of tradition, RNNs have loops so that data can linger around, making it great for processing sequential data as seen in text and time series analysis.

In the RNN, weights are used at every network step to keep track of previous inputs. This ability makes RNNs particularly useful for tasks like speech recognition, language translation, stock price prediction, and all tasks for which information from past inputs may be needed to predict future output. We have proposed a model that combines a deep belief network (DBN) and a recurrent neural network (RNN). Specifically, we aim to blend the data representing the capacity of DBNs able to learn hierarchical representations with the sequential processing capabilities of RNNs. We propose to use a DBN as a feature extractor to convert the input data into a smaller dimensional space; then, this will be fed into the RNN to capture temporal dependencies in the data. This approach, which hybridizes representation learning at a high level from data and a sequence model, will be appropriate for such tasks as natural language processing or forecasting time series.

3.4 Evaluation Metrics

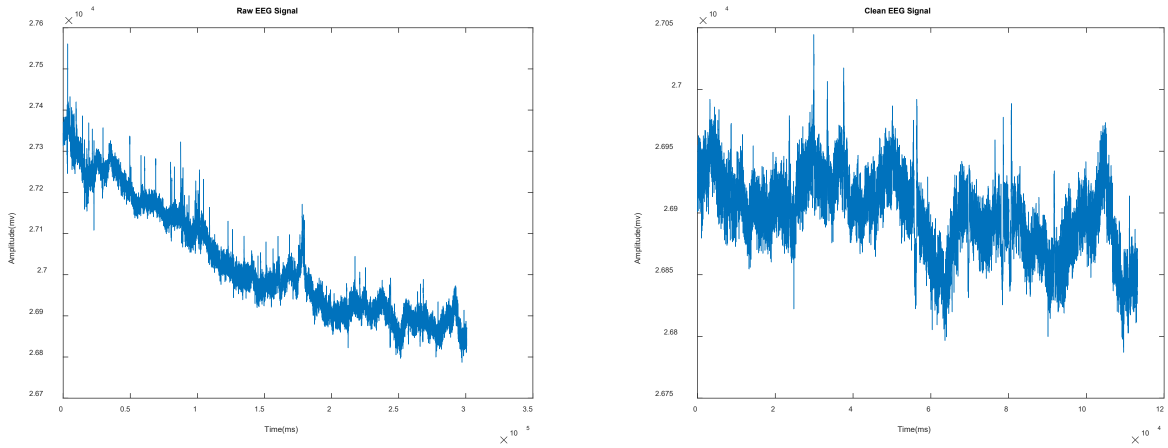
Multiple metrics were employed to evaluate the classification model's performance as part of the evaluation phase. The accuracy metric first sought to gauge model effectiveness by calculating the percentage of correctly predicted outcomes that were made. Sensitivity and specificity were employed to judge the model's ability to classify positive cases (those that are true) accurately and negative cases (those that are not). Using the receiver operating characteristic (ROC) curve, we visualized trade-offs between sensitivity and specificity across different thresholds. The area under an ROC curve (AUC) thus formed was used as an index of how good or bad our model could discriminate between positive and negative cases even despite the value chosen for the threshold. Finally, the model's performance for each class was evaluated based on score, precision, and recall, as well as considering false positives and negatives. A high F1 score indicates a good balance between precision and recall, but a higher AUC and accuracy score indicates better model performance in correctly distinguishing between positive/negative cases.

4. Results and Discussion

This study aimed to establish a neural network model to identify driver conditions from EEG signals as normal or fatigue states. EEG signals were acquired from a public database and divided into training and testing. Data were input during training, and the model was learned using the training dataset. In the meantime, the evaluation set was employed to detect if prediction accuracy is better than the learning effect of trial weight adjustment. Specifically, 80% of training data went into training the DBN-RNN model, while 20% was used as a cross-validation set.

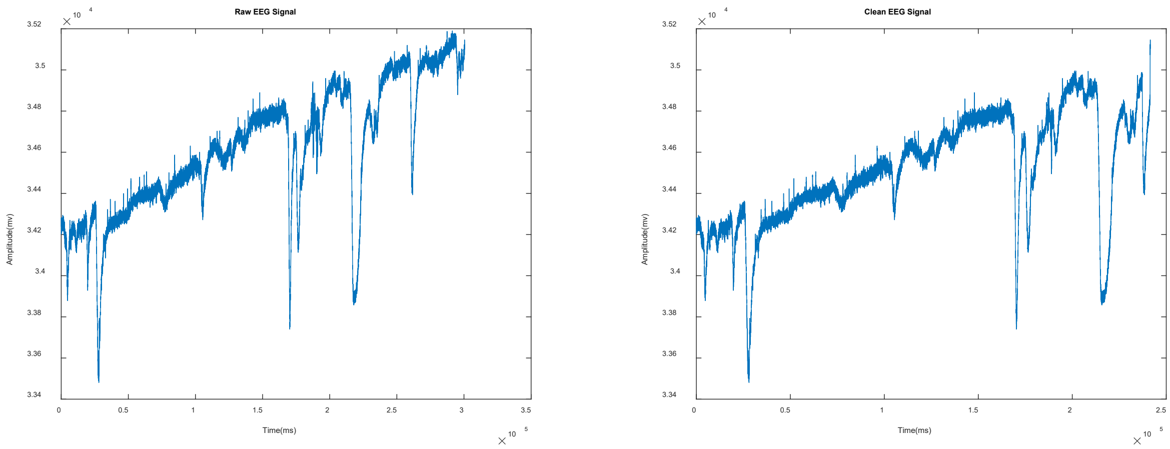
Initially, we utilized principal component analysis (PCA) on the data provided. This was done by creating a PCA object through the PCA class and then fitting the object to the data utilizing the appropriate method. Afterward, the data were divided into training and testing sets, with an 80-20 ratio. Next, we carried out cross-validation by defining a k-fold object utilizing the KFold class from the sci-kit-learn library. This object was designed to create five splits, dividing the data into five equally-sized subsets for cross-validation.

In Fig. 2 for normal conditions and Fig. 3 for fatigue conditions, the results show that the raw EEG signal (a) exhibits a fluctuating amplitude with an overall decreasing trend, indicating the presence of noise and artifacts that could interfere with the accurate analysis of driver fatigue. After applying independent component analysis (ICA) (b), the EEG signal becomes more stable, with reduced noise and a more consistent representation of brain activity. This pre-processing step enhances the quality of the data, making it more reliable for detecting specific brainwave patterns associated with driver fatigue, thereby improving the accuracy of the analysis.



(a) **(b)**

Fig. 2 *EEG signal in normal condition (a) raw EEG signal; (b) clean EEG signal*



(a) **(b)**

Fig. 3 *EEG signal in fatigue condition (a) raw EEG signal; (b) clean EEG signal*

Then, we established a hybrid deep belief network with recurrent neural network (DBN-RNN) model architecture employing the Sequential class from the Keras library. This model consists of three Restricted Boltzmann Machine (RBM) layers, one recurrent neural network (RNN) layer, and an output layer. The Dense function created the RBM layers with 256, 128, and 64 units, respectively. The RNN layer was created using the SimpleRNN function with 32 units. Additionally, to fight against overfitting, we added a dropout layer behind each RBM layer. After the last RBM layer, we added a reshaping layer to reshape its output. Finally, we added an output layer with two units and a softmax activation function. Then, the model was compiled using the Adam optimizer, categorical cross-entropy loss function and accuracy metric.

Finally, an AdaBoost model was trained using a decision tree classifier as the base estimator. We created a decision tree classifier with a maximum depth of 1 using the DecisionTreeClassifier as a class and created an

AdaBoost classifier (with the decision tree classifier as its base estimator) by setting 50 estimators and a learning rate of 1.0 in the AdaBoostClassifier class. The algorithm itself is an iterative process. The compound classifier--which effectively gets better and better each time it sees more data samples because while training, its emphasis on misclassified samples increases--is trained. The algorithm makes predictions by combining the predictions of all of the weak classifiers and then multiplying the resulting vector by a weight associated with each classifier, which is determined from 15-3 and measured in terms of accuracy. The parameter values for the AdaBoost model are presented in Table 1.

Table 1 *Parameter values of the DBN, RNN, and proposed hybrid DBN-RNN model*

No.	Parameter	Values		
		DBN	RNN	DBN-RNN
1	Number of folds for cross-validation	5	5	5
2	RBM1 units	256	-	256
3	RBM2 units	128	-	128
4	RBM3 units	64	-	64
5	Dropout rate	0.2	0.2	0.2
6	Number of RNN units	-	32	32
7	Number of AdaBoost estimators	-	-	50
8	AdaBoost learning rate	-	-	1.0
9	Batch size	32	32	32
10	Optimizer	Adam		
11	Loss function	Categorical cross-entropy		

Then, the evaluation metrics commonly used in deep learning models, including accuracy, sensitivity, specificity, precision, recall, and F1 score, are important because they provide a comprehensive evaluation of the model's performance, taking into account factors such as class imbalance and the nature of true positive and negative instances. Researchers and practitioners can better understand the model's strengths and limitations and make informed decisions about further improvements or adjustments to the model by using these metrics together.

As shown in Table 2, the accuracy of each classifier was calculated and reported as a percentage. The accuracy measure indicates the proportion of correctly classified instances from the total number of instances. The results suggest that all three classifiers achieved high accuracy scores, with the hybrid DBN-RNN classifier achieving the highest accuracy score of 99.98%. The sensitivity and specificity measures were also calculated and reported for each classifier. Sensitivity measures the proportion of true positives correctly identified by the classifier, while specificity measures the proportion of true negatives correctly identified by the classifier. The results indicate that all three classifiers achieved high sensitivity and specificity scores, with the hybrid DBN-RNN classifier achieving the highest scores.

Table 2 *Performance measures of the DBN, RNN, and hybrid DBN-RNN*

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Recall (%)	F1 Score (%)
DBN	99.76	99.88	99.64	99.65	99.88	99.76
RNN	99.93	99.91	99.95	99.95	99.91	99.93
Hybrid DBN-RNN	99.98	100.00	99.95	99.95	100.00	99.98

Then, the precision and recall measures were also reported for each classifier. Precision measures the proportion of true positives among the instances predicted as positive by the classifier. In contrast, recall measures the ratio of actual positives correctly identified by the classifier. The results indicate that all three classifiers achieved high precision and recall scores, with the hybrid DBN-RNN classifier achieving the highest scores. Finally, F1 score results were computed for each classifier. The F1 score is a harmonic mean of precision and recall, providing a balanced measure of the classifier's performance. The results indicate that all three classifiers achieved high F1 scores, with the hybrid DBN-RNN classifier achieving the highest score.

Overall, the results suggest that all three classifiers accurately classified the data. However, the hybrid DBN-RNN classifier outperformed the other two classifiers in all performance measures, indicating that it may be the most suitable classifier for the given task. These findings are significant because they demonstrate the potential of these classifiers for accurately classifying data in similar tasks.

Then, the purpose of the quantitative evaluation of the performance measures provided by the hybrid DBN-RNN model with different numbers of epochs is to determine the optimal number of epochs that will result in the highest performance of the model in classifying the given data. Epochs refer to the number of times the algorithm processes the entire training dataset during the training process. The number of epochs can significantly affect the performance of the model. A model with too few epochs may not have learned all the patterns in the data, resulting in low performance. On the other hand, a model with too many epochs may overfit the training data, resulting in poor generalization to new data.

The results of our study demonstrate that the hybrid deep learning model consisting of DBN and RNN is highly effective in accurately classifying driver fatigue based on electroencephalogram (EEG) signals. The research results are shown in Fig. 4. Specifically, the AUC-ROC values obtained for the model were 1.00 for all five folds, indicating perfect performance in distinguishing between the normal and fatigued states of the drivers. These results suggest that the proposed hybrid DBN-RNN model has great potential for detecting driver fatigue, which is a crucial step toward preventing road accidents caused by driver fatigue. Moreover, the high AUC-ROC values were obtained using a dataset that included a diverse range of drivers and driving conditions, increasing our finding's generalizability.

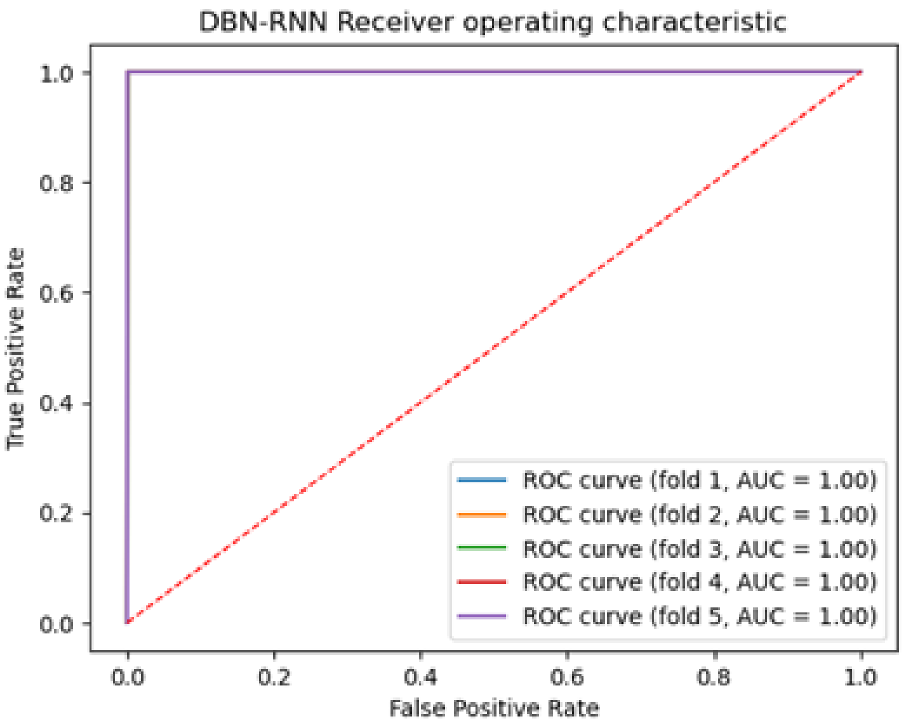


Fig. 4 *DBN-RNN receiver operating characteristic*

Therefore, by evaluating the performance measures of the hybrid DBN-RNN model with different numbers of epochs, it is possible to identify the optimal number of epochs that will result in the best performance in classifying the data. This evaluation can provide insights into the trade-off between training time and model performance and help fine-tune the model to achieve the best possible results.

We evaluated the performance of a hybrid DBN-RNN model with different numbers of epochs (10, 20, and 30) on our dataset, as shown in Table 3. The results showed that the model achieved the same performance measures for all three different epochs tested, with an accuracy of 99.98%, sensitivity and recall equal to 100%, specificity and precision equal to 99.95, and an F1 score of 99.98%. These results suggest that the model's performance has converged, and additional epochs are unlikely to improve its performance further. However, it is important to note that these results may not be generalizable to other datasets or model architectures. Further evaluations are needed to determine the optimal number of epochs for a particular dataset and model architecture.

Int. Journal of Integrated Engineering Vol. 16 No. 9 (2024) p. 302-313

Table 3 *Quantitative evaluation of the performance measures provided by the hybrid DBN-RNN model with different numbers of epoch*

Epoch	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Recall (%)	F1 Score (%)
10	99.98	100.00	99.95	99.95	100.00	99.98
20	99.98	100.00	99.95	99.95	100.00	99.98
30	99.98	100.00	99.95	99.95	100.00	99.98

As shown in Fig. 5, after training the deep learning model for 100 epochs, the overall accuracy was consistently high from epoch 20 onwards. This indicates that the model could effectively learn and classify the driver condition of normal and fatigue EEG signals. The consistent performance from epoch 20 onwards demonstrates that the model could stabilize its learning and achieve a reliable level of accuracy.

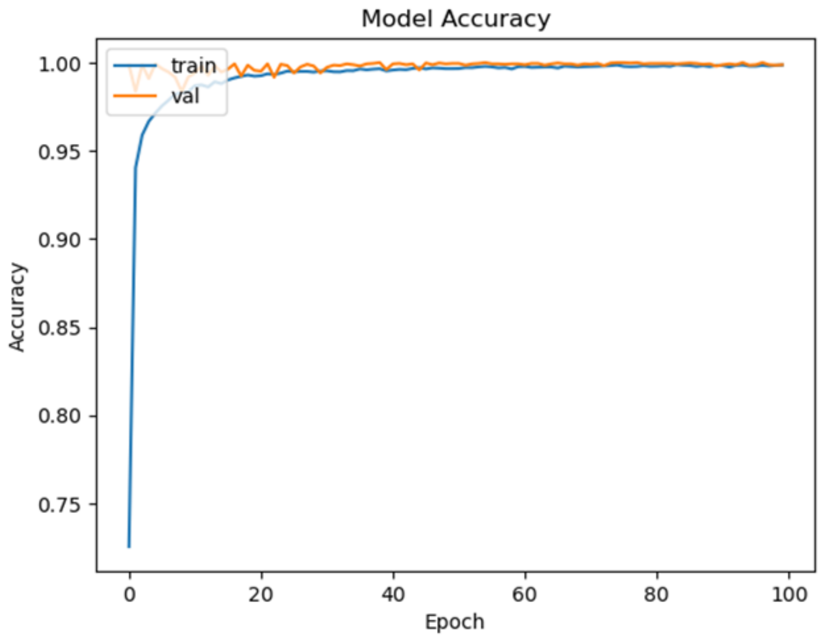


Fig. 5 *Performance of model accuracy graph of the hybrid DBN-RNN model*

It is worth noting that the accuracy may fluctuate during the early epochs of training as the model is still learning and adjusting its weights to minimize the loss function. However, as training progresses, the model should become more stable and more consistent in accuracy. In this case, the model achieved this stability after 20 training epochs. In addition to evaluating the model's accuracy, we analyzed its loss during training, as shown in Fig. 6. The loss function measures how well the model can minimize the difference between its predicted and target outputs. During training, the loss typically decreases as the model improves its predictions. However, suppose the model starts to overfit the data. In that case, the loss on the validation set may increase while the training loss decreases. This can indicate that the model has learned to fit the training data too well and is not generalizing well to new data.

In this study, we observed that the loss decreased steadily throughout the first 50 epochs and then began to stabilize while maintaining a consistent accuracy. This suggests that the model effectively learned to classify the EEG signals and was not overfitting the training data. Overall, the consistent performance of the model in both accuracy and loss metrics after epoch 50 indicates that the developed hybrid DBN-RNN model architecture effectively classified normal and fatigue driver conditions in EEG signals.

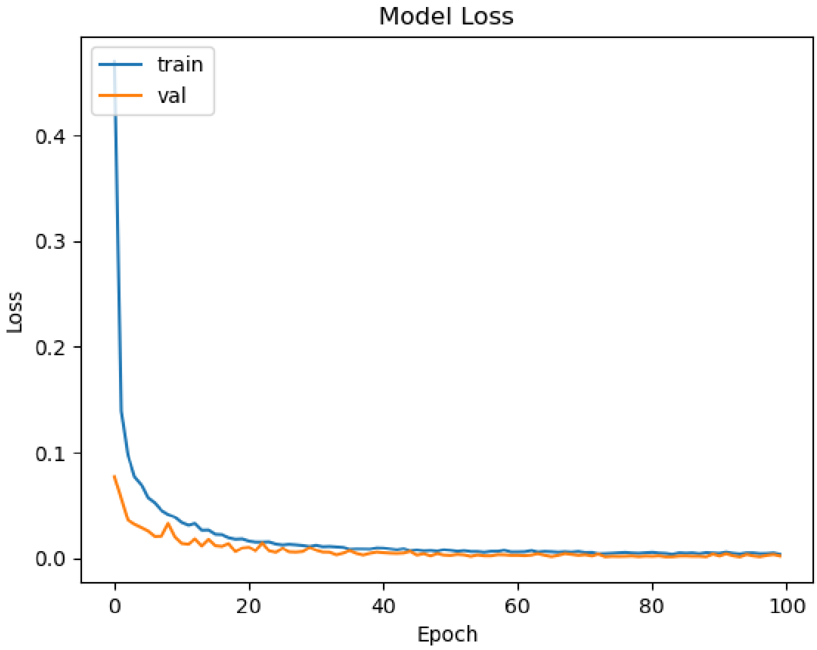


Fig. 6 *Performance of model loss graph of the hybrid DBN-RNN model*

5. Conclusion

In conclusion, this paper proposes an approach to accurately classify driver fatigue using hybrid deep learning that combines deep belief and recurrent neural networks. Even if DBNs and RNNs yield good results individually, a hybrid model is proposed to leverage their complementary strengths, where DBNs excel at unsupervised feature learning and handling complex static data. In contrast, RNNs are effective at capturing temporal dependencies in sequential data. By combining them, a hybrid model can improve generalization, overcome individual limitations, and optimize performance, especially on complex tasks that require both static feature extraction and sequential processing, such as speech recognition or natural language processing. The proposed approach utilizes signal processing techniques such as the Morlet wavelet and modified z-score algorithm for accurately predicting driver fatigue conditions. The proposed model analyzes EEG data involving several processing stages and provides a comprehensive approach to analyzing EEG data that can be used in various applications such as brain-computer interfaces and clinical diagnosis. The study's methodology involves utilizing principal component analysis, cross-validation, and training a model using the AdaBoost algorithm with a decision tree classifier as the base estimator. The results show that the proposed approach outperforms standard methods, providing a more accurate and reliable tool for predicting driver fatigue. This study's findings have significant implications for improving road safety and economic productivity, as accurately predicting driver fatigue can reduce the number of road accidents and associated costs.

Acknowledgement

The authors sincerely thank the Faculty of Electrical and Electronics Engineering Technology at Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA) for their invaluable support and resources throughout this research. The guidance and mentorship provided by the faculty members have been instrumental in shaping this work.

Conflict of Interest

The authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Rafiuddin Abdubrani;* ***data collection:*** *Rafiuddin Abdubrani;* ***analysis and interpretation of results:*** *Rafiuddin Abdubrani, Mahfuzah Mustafa, Zarith Liyana Zahari;* ***draft manuscript preparation:*** *Rafiuddin Abdubrani, Mahfuzah Mustafa, Zarith Liyana Zahari. All authors reviewed the results and approved the final version of the manuscript.*

References

[1] Qin, Y., Hu, Z., Chen, Y., Liu, J., Jiang, L., & Che, Y. (2022). Directed brain network analysis for fatigue driving based on eeg source signals. *Entropy*, 24(8), 1093. https://doi.org/10.3390/e24081093
[2] Zhu, T., Zhang, C., Wu, T., Ouyang, Z., Li, H., Na, X., ... & Li, W. (2022). Research on a real-time driver fatigue detection algorithm based on facial video sequences. *Applied Sciences*, 12(4), 2224. https://doi.org/10.3390/app12042224
[3] Kunasegaran, K., Ismail, A. M. H., Ramasamy, S., Gnanou, J. V., Caszo, B. A., & Chen, P. L. (2023). Understanding mental fatigue and its detection: a comparative analysis of assessments and tools. *PeerJ*, 11, e15744. https://doi.org/10.7717/peerj.15744
[4] Bencsik, B., Reményi, I., Szemenyei, M., & Botzheim, J. (2023). Designing an embedded feature selection algorithm for a drowsiness detector model based on electroencephalogram data. *Sensors*, 23(4), 1874. https://doi.org/10.3390/s23041874
[5] Michi, R. A., Plotkowski, A., Shyam, A., Dehoff, R. R., & Babu, S. S. (2022). Towards high-temperature applications of aluminium alloys enabled by additive manufacturing. *International Materials Reviews*, 67(3), 298-345. https://doi.org/10.1080/09506608.2021.1951580
[6] Hu, Z., Lou, S., Xing, Y., Wang, X., Cao, D., & Lv, C. (2022). Review and perspectives on driver digital twin and its enabling technologies for intelligent vehicles. IEEE Transactions on Intelligent Vehicles, 7(3), 417-440. https://doi.org/10.1109/TIV.2022.3195635
[7] Tawhid, M. N. A., Siuly, S., Wang, K., & Wang, H. (2023). Automatic and efficient framework for identifying multiple neurological disorders from EEG signals. *IEEE Transactions on Technology and Society*, 4(1), 76-86. https://doi.org/10.1109/TTS.2023.3239526
[8] Ayaz, H., Baker, W. B., Blaney, G., Boas, D. A., Bortfeld, H., Brady, K., ... & Zhou, W. (2022). Optical imaging and spectroscopy for the study of the human brain: status report. *Neurophotonics*, 9(S2), S24001. https://doi.org/10.1117/1.NPh.9.S2.S24001
[9] Cheng, Y., Meng, H., & Tan, Y. (2022). Research on fatigue detection system based on facial feature recognition in big data environment. In L. Zhu (Ed.), *2nd International Conference on Artificial Intelligence, Automation, and High-Performance Computing (AIAHPC 2022)* (Vol. 12348, p. 67). https://doi.org/10.1117/12.2641454
[10] Yarici, M. C., Amadori, P., Davies, H., Nakamura, T., Lingg, N., Demiris, Y., & Mandic, D. P. (2023). Hearables: Ear EEG based driver fatigue detection. *arXiv preprint* arXiv:2301.06406. https://doi.org/10.48550/arXiv.2301.06406
[11] Wang, J., Xu, Y., Tian, J., Li, H., Jiao, W., Sun, Y., & Li, G. (2022). Driving fatigue detection with three non-hair-bearing EEG channels and modified transformer model. *Entropy*, 24(12), 1715. https://doi.org/10.3390/e24121715
[12] Li, G., & Chung, W. Y. (2022). Electroencephalogram-based approaches for driver drowsiness detection and management: A review. *Sensors*, 22(3), 1100. https://doi.org/10.3390/s22031100
[13] Othmani, A., Sabri, A. Q. M., Aslan, S., Chaieb, F., Rameh, H., Alfred, R., & Cohen, D. (2023). EEG-based neural networks approaches for fatigue and drowsiness detection: A survey. *Neurocomputing*, 126709. https://doi.org/10.1016/j.neucom.2023.126709
[14] Raju, A. R., Pabboju, S., & Rao, R. R. (2022). Hybrid Deep Learning Approach for Brain Tumor Segmentation and Classification. In Distributed Computing and Optimization Techniques: Select Proceedings of ICDCOT 2021 (pp. 503-514). *Singapore: Springer Nature Singapore*. https://doi.org/10.1007/978-981-19-2281-7_47
[15] Cao, H., & Dong, C. (2022). An English pronunciation error detection system based on improved random forest. *Mobile Information Systems*, 2022(1), 6457286. https://doi.org/10.1155/2022/6457286
[16] Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., ... & Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, 56(11), 13521-13617. https://doi.org/10.1007/s10462-023-10466-8
[17] Arya, A. D., Verma, S. S., Chakarabarti, P., Chakrabarti, T., Elngar, A. A., Kamali, A. M., & Nami, M. (2023). A systematic review on machine learning and deep learning techniques in the effective diagnosis of Alzheimer's disease. *Brain Informatics*, 10(1), 17. https://doi.org/10.1186/s40708-023-00195-7
[18] Li, Z., Ortega Caro, J., Rusak, E., Brendel, W., Bethge, M., Anselmi, F., ... & Pitkow, X. (2023). Robust deep learning object recognition models rely on low frequency information in natural images. *PLOS Computational Biology*, 19(3), e1010932. https://doi.org/10.1371/journal.pcbi.1010932
[19] Han, P., Men, L., & Chen, Z. (2022). PolSAR Image Classification Based on Deep Features. In China High Resolution Earth Observation Conference (pp. 392-403). Singapore: *Springer Nature Singapore*. https://doi.org/10.1007/978-981-19-8202-6_35

[20] Robertson, C., Wilmoth, J. R., Retterer, S. T., & Fuentes-Cabrera, M. (2023). Video frame prediction of microbial growth with a recurrent neural network. *Frontiers in Microbiology*, 13. https://doi.org/10.3389/fmicb.2022.1034586
[21] Hu, H., Xiong, S., Chen, Y., Ye, L., Zhao, S., Qian, K., … & Domenici, M. D. (2022). Prediction of post-covid-19 economic and environmental policy and recovery based on recurrent neural network and long short-term memory network. *Environmental Research Communications*, 4(11), 115001. https://doi.org/10.1088/2515-7620/ac9bd8
[22] Wan, Q. and Liu, J. (2023). Energy efficiency optimization and carbon emission reduction targets of resource-based cities based on bilstm-cnn-gan model. *Frontiers in Ecology and Evolution*, 11. https://doi.org/10.3389/fevo.2023.1248426
[23] Wang, P., Zhang, Y., & Guo, W. (2023). Demand forecast of railway transportation logistics supply chain based on machine learning model. *International Journal of Information Technologies and Systems Approach*, 16(3), 1-17. https://doi.org/10.4018/ijitsa.323441
[24] Abdubrani, R., Mustafa, M., & Zahari, Z. L. (2023). A ROBUST FRAMEWORK FOR DRIVER FATIGUE DETECTION FROM EEG SIGNALS USING ENHANCEMENT OF MODIFIED Z-SCORE AND MULTIPLE MACHINE LEARNING ARCHITECTURES. *IIUM Engineering Journal*, 24(2), 354–372. https://doi.org/10.31436/iiumej.v24i2.2799
[25] Wagh, K. P., Vasanth, K., & Shinde, S. B. (2022). EEG Based Emotional State Recognition Using Time Frequency Parameters. *Journal Harbin Institute of Technology*, 54(1), 101-109. http://dx.doi.org/10.11720/JHIT.54012022.9