Global Journal of Engineering and Technology Advances

eISSN: 2582-5003

Cross Ref DOI: 10.30574/gjeta

Journal homepage: https://gjeta.com/

(RESEARCH ARTICLE)

Research on deep learning models and data preprocessing techniques for driver fatigue detection

Ngoc Bich Nguyen Thi * and An Hoang Thi

Faculty of Information Technology, Sao Do University, HaiDuong, Vietnam.

Global Journal of Engineering and Technology Advances, 2024, 19(03), 050–055

Publication history: Received on 25 April 2024; revised on 04 June 2024; accepted on 07 June 2024

Article DOI: https://doi.org/10.30574/gjeta.2024.19.3.0098

Abstract

Detecting signs of fatigue while driving is one of the most important and urgent issues in today's modern transportation system. Early detection and warning will significantly reduce the risk of traffic accidents. To accomplish this, several methods have been proposed. In recent years, methods based on deep learning techniques have attracted great attention due to their high efficiency and cost savings in detecting and warning about signs of fatigue while driving. In this paper, we study a method to improve the accuracy of deep learning models based on the data preprocessing technique, applied to detect signs of fatigue while driving. Data preprocessing helps the training images become more accurate, noise-reduced, diverse and rich instead of just using the main image. To ensure high reliability and accuracy, the methods are trained and tested on large datasets including thousands of images of all types. Experimental results show that the combination of the VGG-16 model and the data augmentation technique gives better results than the traditional deep learning model, with an accuracy of over 96%.

Keywords: Convolutional Neural Network; Preprocessing; Data Augmentation; Histogram Equalization; Detect Signs of Driving Fatigue

1. Introduction

Driving in a state of fatigue is a common phenomenon for many drivers when participating in traffic. This is a potentially dangerous behavior for both the driver and the surrounding vehicles. According to [1], in Vietnam, of the 3354 traffic accidents that occurred in the first 6 months of 2022, 1.21% of incidents occurred due to driving fatigue, an increase of 0.03% over the same period last year, causing thousands of deaths and injuries [1, 2]. In the United States, up to 100,000 traffic accidents caused by driving in a state of fatigue resulted in 400,000 injuries and 1550 deaths [3]. As a result, research and warning of signs of driving fatigue have become a common subject worldwide.

Previously, some methods based on the driver's physiological signals have been very effective such as electroencephalography (EEG), and electrocardiogram (ECG, EOG) [4, 5]. Methods based on vehicle state such as steering wheel manipulation [6, 7, 8], and vehicle movement in the lane [9] are also applied to detect signs of fatigue. The disadvantage of these methods is the high cost due to the installation of pressure sensors, angle converters, or expensive EEG receivers. The detection of fatigue based on the condition of the eyes is also researched and applied by researchers. Eriksson and Papanikolopoulos have proposed eye state monitoring based on pixel correlation [10]. In this method, the authors have used the concept of the axis of the face to divide and determine the eye's position. Pixel position and number of consecutive frames are applied to monitor the state of the eye.

In recent years, machine learning and deep learning models have been developed and popularly applied because of their high accuracy and effectiveness. In 2004, S. M. Shi et al. used BP networks and computer vision techniques to detect the "yawning" state of mouth to warn of signs of fatigue while driving [11]. Wanzeng Kong's team has proposed an

* Corresponding author: Ngoc Bich Nguyen Thi

improvement to the Adaboost model to identify faces more accurately, for fatigue detection using the additional left and right deflection classifications. If the front face detector fails, these two classifiers will be used to increase the accuracy of face identification [12]. Younes Ed-Doughmi and Najlae Idrissi have compared and evaluated LSTMs, LSTM-CNN, and MLP deep learning models in detecting signs of driving fatigue, experimental results showed that using LSTMs was the most effective with an accuracy of 92.71% [13]. In 2021, Burcu Kir Savas and Yasar Becerikili studied and evaluated the CNN deep learning model on 3 large datasets YawDD, Nthu-DDD, and KouBM-DFD. Experimental results showed that the CNN model with the YawDD dataset achieved the highest accuracy of 99.35% [14]. Combining a Kalman filter with a TLD tracker to detect tired features of the face, the CNN network was then used to detect "yawning" and alerts in real time [15].

To improve the accuracy of deep learning models to detect signs of fatigue while driving, in this paper we research the VGG16 model and then use data augmentation techniques to improve model performance. Experimental results show that the VGG16 model when combined with data augmentation techniques achieves an accuracy of up to 96.6%. In session 2, we present deep learning models that are applied to detect signs of driving fatigue. Session 3 covers the testing process and the results achieved. Finally, session 4 is conclusive.

2. Proposed Method

2.1. Data Preprocessing

In this paper, we use a dataset of 1448 photos of Serena Raju, including 725 photos of being alert while driving and 723 photos showing signs of fatigue. Figure 1 depicts some images in the dataset. The data is preprocessed and divided into two parts with 80% of the samples used for training, and 20% of the samples used for testing. The distribution of training and testing data is presented in Table 1. The training data will be trained and tested using the deep learning algorithm VGG16.

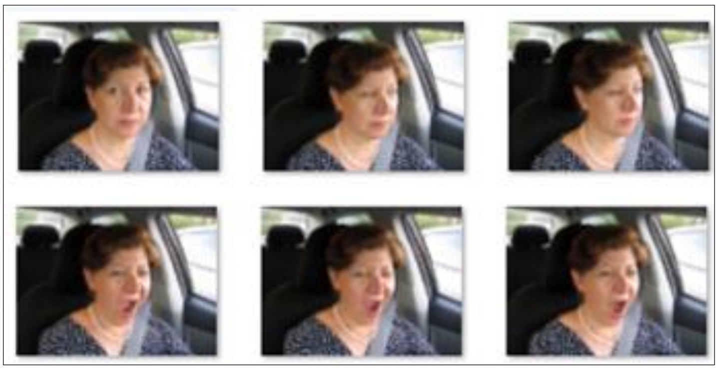

Figure 1 Some images in the dataset (Row 1 is an image belonging to the alert class, row 2 is an image belonging to the tired driving class.)

Table 1 Distribution of data in the paper

Data	Conscious	Signs of fatigue	Total
Training	580	578	1158
Testing	145	145	290
Total	725	723	1448

The training process is as follows: the image is preprocessed by converting it to grayscale to reduce the dimensionality of the input matrix and converting the image to the popular size 224x224. Additionally, aiming to better represent the diversity of images, we perform enrichment of the training dataset using the data augmentation technique. The images are duplicated, rotated 20 degrees, zoomed 20%, flipped, and then fed into the deep learning model for training. Figure 2 depicts some images after enrichment. At the end of the training process, we will have a model aimed at detecting signs of fatigue while driving.

Figure 2 Some images after data augmentation

Because our collected data has noise and brightness irregularities, therefore, the next step in preprocessing is Histogram Equalization. A Histogram of an image is a column chart that lists the number of times light levels appear in the image. This method adjusts the gray level evenly across all gray levels of the image histogram based on the probability distribution, thereby improving the contrast level of the image [17, 18]. The Histogram Equalization algorithm can be described by the formula below.

$$H': I'(x, y) = H'(I(x, y))$$

In which, $H'(i) = \sum_{j=0}^{i} H(j)$, is Histogram normalized to [0, 255], $H(i)$ is the Histogram of the original image, $I(x, y)$ is the original image, $I'(x, y)$ is the Histogram normalized image.

2.2. VGG-16 Model

The VGG-16 model was founded by Karen Simonyan and Andrew Zisserman in Oxford University's Geometric Vision group in 2014 [16]. In terms of architecture, VGG-16 has several improvements including 13 2-dimensional convolutional layers, convolutional, blocks of CNN multi-tiered, and max pooling following instead of one CNN and max pooling alternating each other. VGG-16 only uses small-size filters of 3x3 that help reduce the number of parameters for the model and provide better computational efficiency [16]. The scheme of layers of VGG16 applied in the paper is depicted in Fig 3. Input images with dimensions of 224x224 are put through 2 convolutional layers including 64 filters with the size of 3x3 and padding = 1. Accordingly, the max pooling layer is used for reducing the input matrix size purpose. The next convolutional blocks include 2 Convolutional layers with 128 filters size 3x3, max pooling size 2x2, 3 Convolutional layers with 256 filters size 3x3, max pooling size 2x2, the next two convolutional blocks each consist of 3 Convolutional layers and 1 max pooling layer, the convolutional layer consists of 512 filters with size 3x3, max pooling size 2x2. Going through the convolutional blocks, the model will obtain the characteristic matrix of the image, then pass through 3 fully connected layers with the Relu activation function to flatten the matrix. Finally, we used the Sigmoid function to conduct binary layering with two cases of fatigue and soberness while driving.

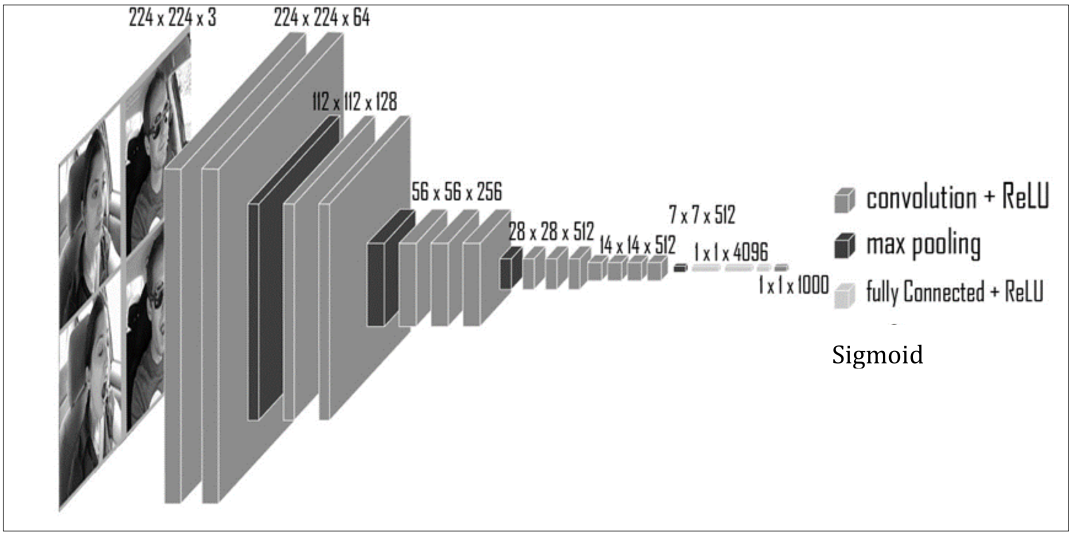

Figure 3 Fatigue detection using the VGG16 model

3. Experimental Results and discussion

3.1. Data Preprocessing

In this section, we present the experimental results of the proposed method for detecting signs of fatigue while driving. To demonstrate the effectiveness of the proposed model, we conduct experiments on unenriched data sets and enriched data sets. Our models are built on computers with configuration CORE I7-10700 2.9GHZ, 16 GB RAM, Windows 10 OS, Python 3.6, and TensorFlow with a Learning Rate of 1e-4, batch_size = 32, epochs = 100.

With the VGG-16 model, the results we obtained are quite high. Fig 4 depicts the accuracy of the VGG-16 algorithm with 100 epochs. In the first 20 epochs, the accuracy of the model was unstable but from the 21st epoch onwards, the accuracy of the model was relatively stable, the Test and Train values were relatively matched, and the model achieved an accuracy of 94.1%.

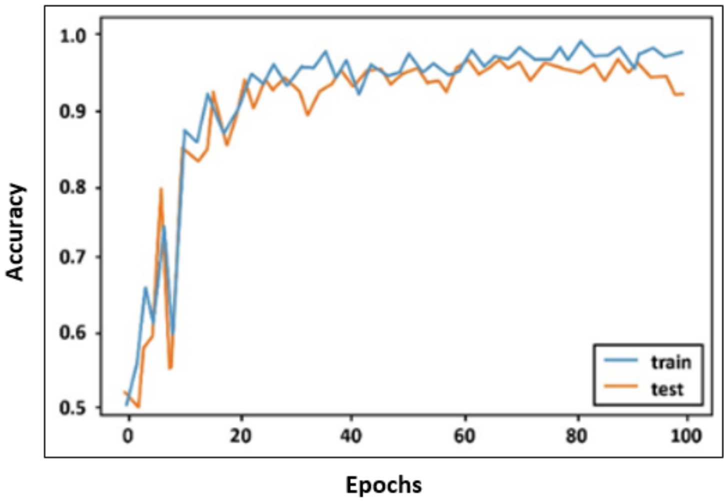

Figure 4 VGG16 model accuracy

When we applied the data preprocessing technique to the training dataset, the accuracy of the VGG16 model improved and achieved 96.6%. Fig 5 depicts the accuracy of the VGG16 model when applying the data augmentation technique. In the first 5 epochs, the accuracy of the model is unstable, but from the 6th epoch onwards the accuracy of the model is stable and gradually progresses to convergence, the Val values are smaller than the train values with a small difference. This proves the accuracy and reliability of the model when combining the data augmentation techniques and VGG-16.

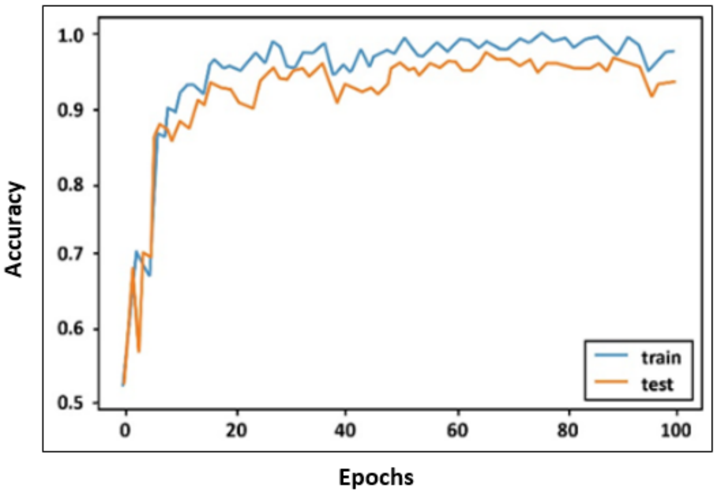

Figure 5 VGG16 model accuracy with data preprocessing technique

To prove the effectiveness of the proposed method, we perform statistics and compare four values after training the models including Precision, Recall, F1 Score, and Accuracy. Accuracy is the ratio between the number of correctly predicted data points and the total number of data points in the test set. Precision is the ratio of the number of correctly identified points in a class to the total number of points classified into that class. The Recall value is the ratio of the number of correctly identified points in a class to the total number of data points in that class. The quantity F1-score is the harmonic average determined based on two measures of precision and recall. The results of comparing the accuracy in detecting signs of fatigue while driving between algorithms are shown in Table 2.

Table 2 Compare the accuracy between algorithms

Models	Accuracy	Precision	Recall	F1-Score
VGG-16	94.13%	93.10%	95.07%	94.07%
Our Model	96.60%	97.24%	95.91%	96.57%

4. Conclusions

Detecting signs of driver fatigue when participating in traffic is an urgent need. This paper presents a method to improve the accuracy of deep learning models in detecting driver fatigue while driving. Data is preprocessed and enriched by rotating, flipping, zooming, etc., making the data richer, using Histogram Equalization to remove noise and balance brightness between images. This helps the training process achieve higher accuracy. The above combination is highly effective with 96.6% accuracy, about 2% higher than pure deep learning. In the future, we will develop and improve the deep learning network model to achieve even higher accuracy It can be effectively applied to detect and warn of driver fatigue in particular and support the field of traffic safety in general.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Yen Chi, Alarming for increased traffic accidents due to driver drowsiness and fatigue, Traffic Safety, 21/07/2022.

[2] Thanh Quang, The rate of deaths due to traffic accidents has increased in many localities, Lao Dong Newspaper, 16/07/2022.

[3] S. R. Paul, NHTSA's Drowsy Driver Technology Program, 2005.

[4] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness, Neuroscience and Biobehavioral Reviews, vol. 44, pp. 58–75, 2014.

[5] T. Myllylä, V. Korhonen, E. Vihriälä et al., Human heart pulse wave responses measured simultaneously at several sensor placements by two MR-compatible fibre optic methods, Journal of Sensors, vol. 2012, Article ID 769613, 8 pages, 2012.

[6] J. Krajewski, D. Sommer, U. Trutschel et al., Steering wheel behavior based estimation of fatigue, in Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp. 118–124, 2009.View at: Google Scholar.

[7] Y. Wu, J. Li, D. Yu, and H. Cheng, Research on quantitative method about driver reliability, Journal of Software, vol. 6, no. 6, pp. 1110–1116, 2011.View at: Publisher Site | Google Scholar.

[8] A. Eskandarian and A. Mortazavi, Evaluation of a smart algorithm for commercial vehicle driver drowsiness detection, in Proceedings of the IEEE Intelligent Vehicles Symposium, pp. 553–559, June 2007.

[9] C. X. Huang, W. C. Zhang, C. G. Huang, and Y. J. Zhong, Identification of driver state based on ARX in the automobile simulator, Technology & Economy in Areas of Communications, vol. 10, no. 2, pp. 60–63, 2008.

[10] M. Eriksson and N. P. Papanikolopoulos, Eye-tracking for detection of driver fatigue, in Proceedings of the IEEE Conference on Intelligent Transportation Systems (ITSC '97), pp. 314–319, 1997.

[11] S. M. Shi, L. S. Jin, R. B. Wang, and B. L. Tong, Driver mouth monitoring method based on machine vision, Journal of JiLin University, vol. 34, no. 2, pp. 232–236, 2004.

[12] Wanzeng Kong, Lingxiao Zhou, Yizhi Wang, Jianhai Zhang, Jianhui Liu,1 and Shenyong Gao, A System of Driving Fatigue Detection Based on Machine Vision and Its Application on Smart Device, Journal of Sensors, 2015.

[13] Younes Ed-Doughmi, Najlae Idrissi, Driver Fatigue Detection using Recurrent Neural Networks, NISS19, March 27–29, 2019, Rabat, Morocco.

[14] Burcu Kir Savas và Yasar Becerikili, A Deep Learning Approach to Driver Fatigue Detection via Mouth State Analyses and Yawning Detection, IOSR Journal of Computer Engineering (IOSR-JCE) , ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 23, Issue 3, Ser. I(May – June 2021), PP 24-30, 2021.

[15] Weiwei Zhang, Yi L. Murphey, Tianyu Wang, Qijie Xu, Driver yawning detection based on deep convolutional neural learning and robust nose tracking, 2015 International Joint Conference on Neural Networks (IJCNN).

[16] Karen Simonyan, Andrew Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, Computer Vision and Pattern RecognitionarXiv:1409.1556, 2014.

[17] B. Oktavianto, T. W. Purboyo, 2018. A Study of Histogram Equalization Techniques for Image Enhancement. International Journal of Applied Engineering Research, vol 13, no. 2, pp. 1165-1170.

[18] H. Ibrahim, N. S. P. Kong, et al, 2007. Brightness Preserving Dynamic Histogram Equalization for Image Contrast Enhancement. IEEE Transactions on Consumer Electronics, vol 53, pp. 1752-1758.