Impact of *Data Augmentation* and *Transfer Learning* on the Accuracy of Models for Drowsiness Detection Model

Syeda Areesha Najam
Computer Science
Habib University
Karachi, Pakistan
sn05985@st.habib.edu.pk

Sana Fatima
Computer Science
Habib University
Karachi, Pakistan
sf06199@st.habib.edu.pk

***Abstract*—Machine learning approaches have been utilized to predict drivers' mental states based on bio-indicators, their behavior, and their facial expressions while driving. As an application of artificial intelligence, deep learning models are able to learn and improve without being explicitly programmed. This work proposes the study and investigation of the effect of data augmentation and transfer learning on the accuracy of the CNN-based drowsiness detection model. The inspiration for this endeavor comes from a number of previous studies in this field. These previous studies include the use of HAAR cascade, Viola-Jones algorithm, camshaft methods, ANNs, and other deep learning techniques and concepts for drowsiness detection models. This paper addresses the impacts of said strategies on the accuracy score of the model by evaluating the retrained model and comparing it to some state-of-the-art approaches, as well as the experiments and deductions that went along with them. The deductions are from these experiments are mentioned near the end of the report.**

Keywords—drowsiness detection, deep learning, transferred learning, data augmentation.

I. INTRODUCTION

Sleeping is a naturally occurring phenomenon that is needed to recharge your energy level and fuel you up with freshness. However, insufficient amounts of sleep will lead to many problems, including drowsiness which can be very dangerous if one drives a car while being drowsy. Drowsy driving increases the risk of car accidents as it makes a person lose attention. ‘The accidents caused by such driving are estimated to be 100,000 annually with approximately 1550 deaths, 71,000 injuries and 12.5 billion dollars of property damage’ [1]. According to the National Highway Traffic Safety Administration (NHTSA), in the year 2017, drowsy drivers caused 91,000 car crashes which included 800 deaths and approximately 50,000 injuries’ [1].

With advancements in design and computer-controlled automotive applications, the issues of safe driving, traffic flow, and driver comfort have become a focused topic. Given the rising incidence of car accidents, much work and study have gone into proposing and designing various sleepiness detection devices, and the research is still ongoing. One of the most common limitations of these systems is the limited datasets, due to which the efficiency of the model is also limited. Considering the said issue, in this paper, we have discussed the effect of the implementation of different techniques, i.e. data augmentation and transferred learning on the accuracy of the latest drowsiness detection model [11].

II. LITERATURE REVIEW

There are a number of studies and researches that are done to address the issue of detecting drowsy drivers (before and after the evolution of deep learning). Since 2013, people have been working on improving the dataset as well as the methodologies utilized in the drowsiness detection system.

A. *Drowsiness Detection for Driver Assistance [2]*

The following research was conducted in the year 2013, in which the researchers tried to solve the said issue by designing a model which used the Hidden Markov Models, a probabilistic model, and outlined the evolution of events that depend on hidden/internal factors.

For future analysis, the paper suggested improvising the dataset in terms of adding sources. In addition to this, it emphasized diversifying the applications of the said system and utilizing it to detect drowsy students to analyze its relation with their class performance. Moreover, it mentioned the importance of analyzing drowsiness based on more features, involving both, drowsy and non-drowsy expressions.

B. *Driver Drowsiness Detection To Reduce Major Road Accidents in Automotive Vehicles [3]*

This model was designed in 2015 and uses the technique of HAAR to detect drowsy drivers. HAAR features are basically similar to the kernel in Convolutional Neural Network but unlike kernel, HAAR features have to be determined manually and this is why there is a limit to which it can detect things.

Its future works suggested adding varying sensors with appropriate hardware units to improvise the efficiency of the model.

C. *Driver drowsiness detection using ANN image processing [4]*

The research conducted in 2017, addressed the same issue using the methodology based on artificial neural networks. The EEG and EOG signals were used as inputs to ANN. The said signals are cost and time inefficient.

For improving the efficiency of their model, they emphasized diversifying the dataset by adding different images, with different drivers, in different positions and lighting conditions.

D. *A Realistic Dataset and Baseline Temporal Model for Early Drowsiness Detection [6]*

The following research, conducted in 2019, focused on making a new dataset named Real-Life Drowsiness Dataset (RLDD) and used HM-LSTM - network to identify the temporal pattern in blinking i.e. limited to only closed and open state of eyes.

As this dataset only considered a single feature for detecting drowsy drivers, i.e. the blinking pattern, it suggested incorporating another parameter with the blinking parameter for better results.

E. *Bias Remediation in Driver Drowsiness Detection systems using Generative Adversarial Networks [7]*

This research further built on increasing the variation in the dataset using the Generative Adversarial Network to generate realistic images which are used in retraining a ResNet model and also prevented overfitting by changing the learning rates in each iteration.

For future work, it suggested that the training data should be built using various population groups so that this model becomes more efficient and diverse.

F. *Driver Drowsiness Detection [8]*

This paper, of 2020, mentions the past researches and varying techniques used to build Drowsiness Detection System. The paper talks about different algorithms including HAAR, Camshift, Perclos, Viola-Jones, etc.

The future works in the paper suggested to include not just the feature of yawning to detect the drowsy drivers but also to check the frequency of yawning which could reduce the error rate (e.g. laugh mistaken for yawn) and increase the overall accuracy of the model.

G. *A Compact and Interpretable Convolutional Neural Network for Cross-Subject Driver Drowsiness Detection from SingleChannel EEG [9]*

This research, of the year 2021, used the neural networks based on ResNet and MobileNet to design a model for drowsiness detection. But due to insufficient data, their work is still in process and is not ready for online application yet.

The current model has only been tested on the Oz channel on a single dataset with 11 subjects as an initial attempt. It suggests considering testing the model on various EEG channels and a larger dataset with more subjects in the future.

III. EXPERIMENTS AND METHODOLOGIES

After the extensive literature review, discussed before, the presented paper would answer the following questions: What is the effect of Data Augmentation and Transfer Learning on the accuracy scores of a model? Is Data Augmentation a good approach to be applied on datasets in order to increase the number of samples for training purposes? How does the use of a pre-trained model (Transfer Learning) contribute towards the relative increase in the accuracy of a model along with enhancing the training rate? The common limitations of previous researches and the suggestions for future work were mainly based on increasing the dataset which may improve the efficiency of the model as this will provide an exposure to the variety of samples. Considering this, a variety of samples has been introduced in the training dataset by integrating different datasets and performing data augmentation. In addition to this, the concept of transfer learning has also been implemented on a pre-trained model [11] with the purpose of investigating its effect on the accuracy of the model. Further, the experiments and results have been discussed in later sections which helped us to establish how the techniques mentioned above contribute towards the accuracy of drowsiness detection models.

For this research i.e. to analyze the effect of data augmentation and transfer learning on the accuracy of the model, we referenced a model named “Drowsiness Detection Model Tensor Flow” [11]. Initially, the model preprocessed the images to a size of 256×256, and ReLU is used as the activation function. It is suggested to use ReLU because it has been observed that training of deep neural networks using ReLU tends to converge more rapidly and reliably as it avoids the problems of vanishing gradients. At the output layer, softmax is applied to calculate the probability distribution of the output, predicted by the model. For the purpose of optimization, an Adams optimizer was used. In addition to this, the loss function used was “sparse categorical cross-entropy” for the classification of a single sample image to one of the 4 classes’ i.e. yawn, no yawn, closed eyes, or open eyes. The architecture of the model is given below in Fig. 1.1.

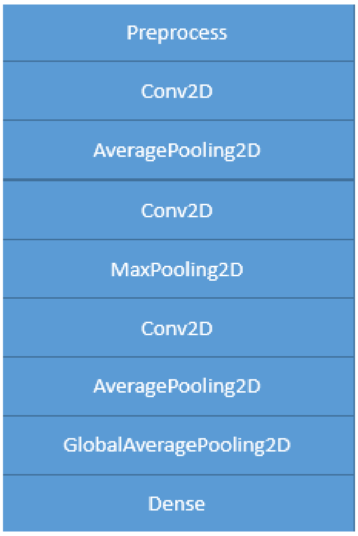

Fig 1.1: Architecture of Drowsiness Detection Model [11]

This dataset provided with the mentioned model had a total of 2467 training images and 433 testing images. The distribution of training and testing images into 4 different classes is given in Table 1.1. The same dataset was also used for the research focused on in this paper.

TABLE 1.1. YAWN EYE DATASET

Yawn Eyes Dataset		
Validation Dataset	Open	109
	Closed	109
	Yawn	106
	No_Yawn	109
Total Validation		433
Train	Open	617
	Closed	617
	Yawn	617
	No_Yawn	616
Total Train		2467

With the above mentioned architecture and dataset, the referenced model [11] achieved a test accuracy score of 90% on 100 epochs.

To analyze the accuracy, based on a larger dataset using different techniques, the dataset was revised and increased. For this purpose, two other datasets were combined, out of which, one consisted of only two categories, open and closed eyes [12]. The other dataset picked was licensed and consisted of grey scaled videos, out of which we extracted frames and categorized them into a yawn and no yawn [13]. After merging all three datasets, the final distribution of images was as mentioned in Table 1.2, from which it can be seen that our revised dataset is almost double, compared to the one we initially picked [11].

TABLE 1.2. REVISED DATASET

Final Dataset		
Validation Dataset	Open	100
	Closed	100
	Yawn	100
	No_Yawn	100
Total Validation		400
Train	Open	1216
	Closed	1217
	Yawn	1216
	No_Yawn	1113
Total Train		4762

We then moved on to test our hypothesis which stated that the accuracy of the model should improve by applying the concept of **data augmentation** and **transfer learning**.

A. Implementation of Data Augmentation

For data augmentation, we modified the model by adding convolutional 2d layers, changing the learning rate as well as the exponential decay rates of the optimizer to analyze which gives the best accuracy. After running the modified model, on the final dataset (without augmentation), for 100 epochs, the testing accuracy increased from 90% to 94% as can be seen in Fig 1.2.

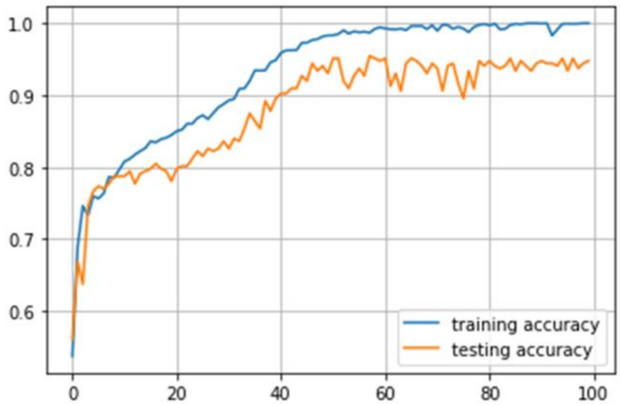

Fig 1.2: Training and Testing Accuracy of Revised Model

Now that accuracy of the revised model has been noted, our next goal was the augmentation of data and then using the augmented dataset for the training purpose of the model.

Data augmentation is a method that is used to expand the size of a training dataset by creating modified versions of images in the dataset either done manually or artificially. The paper, "A survey on Image Data Augmentation for Deep Learning" [14] mentions that the validation error must continue to decrease with the training error. The paper [14] suggested that data augmentation serves as a powerful technique to achieve this as the augmented data will be a more comprehensive set of possible data points, which will minimize the distance between the training and validation accuracy, even for different testing data [14]. In order to apply this method to the final training dataset, the below-mentioned filters were selected for data augmentation:

● Flipped images

● Images with noise

● Images with varying brightness levels

For the following purpose, the data was augmented with Gaussian noise, flipped images, and reduced contrast in images. Note that, randomly any image was picked from the source directory, and then all three filters of adding noise, reducing contrast, and flipping were applied on the same image. This method was repeated for 100 iterations on each category, which resulted in additional 300 pictures being added to each class.

Figure 1.3 shows the initial image which is picked randomly from the category of closed eyes from the training dataset.

Fig 1.3: Random Image from Source Directory

After adding noise, flipping it, and reducing contrast i.e. increasing brightness the following images were augmented:

Figure 1.4: Added noise (1), Flipped (2), Brightened (3)

The augmented dataset had the distribution as mentioned in Table 1.3.

TABLE 1.3. DATASET CONTAINING AUGMENTED IMAGES

Final Dataset		
Validation Dataset	Open	100
	Closed	100
	Yawn	100
	No_Yawn	100
Total Validation		400
Train	Open	1516
	Closed	1517
	Yawn	1516
	No_Yawn	1513
Total Train		6062

After execution of the revised model on the augmented dataset, we realized that our hypothesis, about the increase in accuracy using data augmentation, was disproved. The results and deductions will be discussed in a later section.

B. Implementation of Transfer Learning

Furthermore, the effect of transfer learning was investigated on the accuracy score of the model by performing comparative analysis on the accuracy of the pre-trained model [11] and accuracy after combining it with Transfer Learning. With this purpose, we will use a similar architecture and dataset, as of [11], however combining Transfer Learning on the top of the pertained model. The aim is to establish whether the accuracy score improves or not.

The concept of transfer learning in deep learning models has recently gained attention and has been successfully applied to various domains leading to remarkable results. It is a machine learning technique whereby a model is trained and developed for one task and is then re-used on a second related task [10]. A better explanation of the same concept could be that 'it refers to the situation whereby what has been learned in one setting is exploited to improve optimization in another setting' [10]. The work, "A Study on CNN Transfer Learning for Image Classification" highlights a system that uses a pre-trained model, the Inception-y3 model which was trained on a base dataset (ImageNet). It was then used to learn features by training the model again on the new dataset CIFAR-10 and Caltech Faces [10]. This gives a clear idea of how the said technique allows us to begin with the already learned features and adjust these features along with the structure of the model in order to re-train it on a new dataset instead of starting from scratch. For this project, the model [11] was trained and saved for later use. Later, for implementation of transfer learning, the convolutional layers of the pre-trained model were made untrainable and further a top convolutional layer was added in architecture. For better understanding, refer to Figure 1.5.

Figure 1.5: Architecture of modified model for transfer learning

The above-mentioned modified architecture has a total of 455, 236 parameters among which 361,988 are trainable parameters whereas 93,248 were non-trainable parameters. A significant improvement was noticed via transfer learning which also proved the half part of our hypothesis. The results and deduction are discussed in detail in the next section.

III. RESULTS AND DEDUCTIONS

In this section, the results achieved from conducted experiments are discussed followed by deductions made from them.

As per our hypothesis, the testing accuracy of the model is expected to increase after getting trained on augmented data. The experiment conducted for analyzing the proposed hypothesis has already been discussed earlier. Evidently, the result contradicted our hypothesis. Figure 1.6 shows that the training accuracy is converging to 100%, in about 20 epochs, but the testing accuracy is not improving.

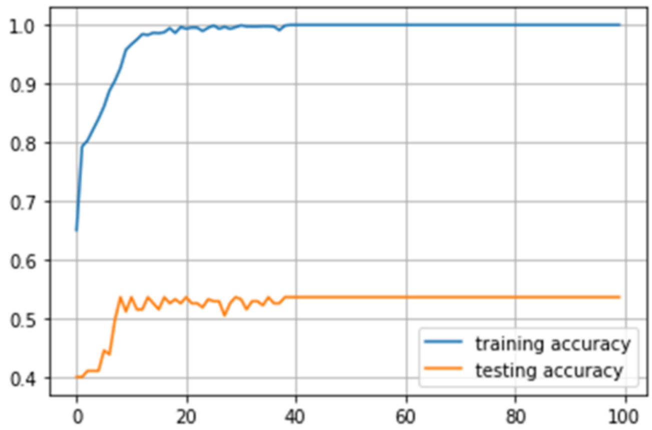

Figure 1.6: Training and Testing Accuracy after Data Augmentation

It was found that data augmentation decreases the validation accuracy as the training images do not match the real-world scenarios. The paper, "The Effectiveness of Image Augmentation in Deep Learning Networks for Detecting COVID-19" [15] highlights that 'the data augmentation is not a required step and actually harmed the deep learning model in this case, by exposing it to a large amount of distorted (or noise) images'. It also mentions that 'all augmentation methods scored lower validation accuracy than without augmentation.' This clearly shows that for the scenarios where data augmentation does not match the real-world scenarios, it is a bad practice to apply this technique as it lowers the validation accuracy of the model. The reason behind this failure is that the model shifts its focus towards augmented data and learns its patterns. While testing, the model expects augmented images to be part of the testing dataset which should not be the case in real-world scenarios and for real testing purposes. As augmented images essentially resulted in the prediction of an *incorrect probability distribution* for each category, as an end result we observed fluctuating validation accuracy.

Another part of our hypothesis, on the other hand, was proven, as the previously achieved testing accuracy of 90 % of the pre-trained model [11] was then increased to 93 % via transfer learning. At the same time, both the training and testing losses converged to 0.0035 and 0.4, respectively, along with appreciable training accuracy. Figure 1.7 demonstrates the said significant improvement in accuracy scores.

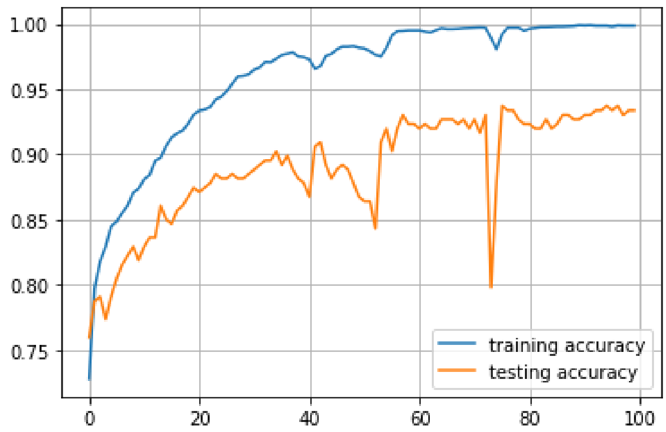

Fig 1.7: Training and Testing Accuracy after Transfer Learning

Besides the decreasing loss and increasing accuracy, we saw a considerable increase in the training rate of the model. The ETA (Estimated Time of Arrival) of 11 seconds was observed for each epoch in the pre-trained model. However, this was reduced to 2 seconds via the implementation of transfer learning. In addition, each batch took 31 milliseconds per step (batch) this time which is significantly lesser than 181 milliseconds served in the pre-trained model [11] per batch. Hence, we can deduce that transfer learning not only improves the accuracy but also improves the time taken per batch for training and ETA.

IV. CONCLUSION

The purpose of our research was to improve the Drowsiness Detection Model by increasing the existing accuracy of the model by implementing the techniques of **data augmentation** and **transfer learning.** We first revised the dataset by merging two different datasets to our initial one.

After doing so, we modified the model for data augmentation by adding more layers and changing the learning rate along with the beta values of the Adam optimizer. Furthermore, we augmented random images by flipping, increasing the brightness and adding noise to the original images. When we executed our revised model on the new augmented dataset, our hypothesis was proven wrong, as instead of increasing the validation accuracy, it resulted in decreased accuracy. This was because of the shift of model towards the augmented images which made it inefficient in predicting the real-life images by disturbing the probability distribution of different classes.

For transferred learning, we took the pre-trained model from [11] and then re-trained on our training data set which is essentially a combination of different data sets. For re-training purposes and with the expectation of observing increased accuracy we added more layers. Interestingly, with transfer learning, we observed increased accuracy up to 93%. Not only accuracy but the speed of the model accompanied by ETA and time taken per batch was also improved.

V. FUTURE WORKS

Our modified model, as mentioned above, also works well with glasses and dim lighting at night. This system can be improved by increasing the variety in datasets and number of samples, as we did with data augmentation, but it is recommended that noise in datasets should be avoided because it shifts the focus of the model and causes fluctuations in graphical and numeric trends of accuracy. Furthermore, larger neural networks with the working principle of predicting human behavior by recognizing face attributes can be utilized to perform transfer learning, which can increase the model's accuracy and speed, as described previously. These solutions can be commercially generalized and successfully implemented in today's automobiles at a lower cost. When combined with automobiles, the technology has the potential to reduce traffic collisions caused by drowsy drivers.

REFERENCES

[1] Administration, N., 2021. *Home | NHTSA*. [online] NHTSA. Available at: <https://www.nhtsa.gov> [Accessed 5 November 2021].

[2] Shareok.org. 2021. [online] Available at: <https://shareok.org/bitstream/handle/11244/15141/Tadesse_okstate_0664M_13072.pdf;jsessionid=CEDFF0CFFDC78044FA4EFEF6798FB267?sequence=1> [Accessed 5 November 2021].

[3] Slideshare.net. 2021. *IRJET-Driver Drowsiness Detection to Reduce the Major Road Accidents ….* [online] Available at: <https://www.slideshare.net/irjetjournal/irjet-v2i163-47998149> [Accessed 5 November 2021].

[4] 2021. [online] Available at: <https://iopscience.iop.org/article/10.1088/1757-899X/252/1/012097/pdf> [Accessed 5 November 2021].

[5] 2021. *Computer Vision Project | Driver drowsiness detection | Full code explanation | OpenCV Python Dlib*. [image] Available at: <https://www.youtube.com/watch?v=ksi42rwGyas> [Accessed 5 November 2021].

[6] Paperswithcode.com. 2021. *Papers with Code - A Realistic Dataset and Baseline Temporal Model for Early Drowsiness Detection*. [online] Available at: <https://paperswithcode.com/paper/a-realistic-dataset-and-baseline-temporal> [Accessed 5 November 2021].

[7] Arxiv.org. 2021. [online] Available at: <https://arxiv.org/pdf/1912.12123v1.pdf> [Accessed 5 November 2021].

[8] Ijert.org. 2021. [online] Available at: <https://www.ijert.org/research/driver-drowsiness-detection-IJERTCONV8IS15008.pdf> [Accessed 5 November 2021].

[9] Arxiv.org. 2021. [online] Available at: <https://arxiv.org/ftp/arxiv/papers/2106/2106.00613.pdf> [Accessed 5 November 2021].

[10] Hussain, Mahbub & Bird, Jordan & Faria, Diego. 2018. A Study on CNN Transfer Learning for Image Classification. Available at: <file:///C:/Users/User/Downloads/UKCI-2018_paper_37.pdf> [Accessed 5 November 2021]

[11] Kaggle.com. 2021. Drowsiness Detection Model TensorFlow. [online] Available: https://www.kaggle.com/vanvalkenberg/drowsiness-detection-model-tensorflow/notebook [Accessed 5 November 2021].

[12] mahmoudreda55, "Closed open eyes," *Kaggle*, 10-May-2021. [Online]. Available: https://www.kaggle.com/mahmoudreda55/closed-open-eyes/data. [Accessed: 06-Dec-2021].

[13] "Driver drowsiness detection dataset," *NTHU CVlab - Driver Drowsiness Detection Dataset*. [Online]. Available: http://cv.cs.nthu.edu.tw/php/callforpaper/datasets/DDD/. [Accessed: 06-Dec-2021].

[14] Shorten, C. and Khoshgoftaar, T., 2019. A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(1).

[15] M. Elgendi, M. U. Nasir, Q. Tang, D. Smith, J.-P. Grenier, C. Batte, B. Spieler, W. D. Leslie, C. Menon, R. R. Fletcher, N. Howard, R. Ward, W. Parker, and S. Nicolaou, "The effectiveness of image augmentation in deep learning networks for detecting COVID-19: A geometric transformation perspective," *Frontiers in medicine*, 01-Mar-2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7956964/. [Accessed: 06-Dec-2021].