Driver Drowsiness Detection System

Tanisha Choudhary

Abstract

Driver Drowsiness is a leading stimulus of road accidents which in the worst case can result in loss of life and other important resources. With ever growing and advancing technology we have somewhat tried to find a solution to this and many other problems using well-developed detection systems. In this paper, I would be suggesting some CNN algorithms working on image based measurements which will help us to detect driver sleepiness. The models used are designed to generate real-time alerts to the driver when drowsiness is detected, allowing the driver to take time and situation appropriate actions. The models used in this paper result in a high accuracy for detecting drowsiness, which in turn will help us in monitoring driver fatigue more effectively.

1 Introduction

Driver drowsiness is among one of the foremost problem globally claiming approximately 1.35 million lives each year according to WHO. The traffic accidents between 20 and 30 percent are caused because of driver drowsiness. In addition, driver fatigue was claimed to be a cause of mounting traffic accidents in different countries.

Driver drowsiness can be understood as a car operator's reduced ability to maintain alertness and concentration while driving, leading to slower reaction times, reduced alertness, and impaired judgment which might lead to lane departures, erratic driving etc. Hence, the discovery of this issue's solution to prevent it is essential to ensure road safety.

Recently, various technologies and methods were developed to detect driver drowsiness, including physiological measures, on-vehicle measurements, and visual measures. These technologies range from wearable sensors to eye-tracking systems to machine learning algorithms. Physiological measures such as electroencephalography (EEG), electrooculography (EOG) and heart rate variability (HRV), are commonly used to detect changes in a car operator's physiological signals which might indicate drowsiness. Vehicle-based measures include steering wheel movements, vehicle speed and lane departure etc. to detect deviation in driving behavior that indicate drowsiness. Visual measures such as eye movements and facial expressions are used to detect change in the driver's visual attention that indicate drowsiness.

Despite the advances in driver drowsiness detection technologies, challenges remain which need to be addressed. For example, many current detection systems are prone to false positives alarms or are affected by external factors such as poor lighting conditions and weather. Therefore, there is an urgent need of models or systems that can detect the driver drowsiness in real-time.

2 Experiment

This study will help us compare the accuracy of various CNN models on a given dataset to distinguish the most accurate model for driver drowsiness detection. The implementation of these models involves collecting eye and mouth characteristic data to monitor changes in driver fatigue. The data collected is analyzed to determine the state of the eye which maybe open / closed and the yawning of the driver.

2.1 DATASET

Firstly I have selected a dataset that I will be using throughout this project and can be found on Kaggle with the name Yawn_Eye_Dataset_New. It is a collection of 2900 image files particularly designed for assessing the vigilance of the person.

The dataset contains pictures of both open plus closed eyes along with both yawning plus non-yawning stages images. We will be dividing the dataset into two parts/folders, namely the train and test folders. Each part/ folder contains four directories of labels, named as closed, open, no yawn, and yawn. The train folder consists of a total of 2467 files, with 617 closed eye images, 617 open eye images, 616 no yawn images, and 617 yawn images and the test folder contains 433 files, which consist of 109 closed eye images, 109 open eye images, 109 no yawn images, and 106 yawn images.

I am using this dataset because it is very well balanced for training as well as testing Convolutional Neural Networks (CNNs) for drowsiness detection in drivers as it provides a huge variety of images with appropriately defined eyes and mouth. Secondly, the partition of the dataset into train and test folders provides an effective way to gauge the working of different CNN models on Yawn_Eye_Dataset_New dataset. It contains the data of both men and women of different age groups and ethnicities.

Classes	**Testing**	**Training**
Closed	109	617
Open	109	617
Yawn	109	616
No Yawn	106	617
Total	433	2467

Table 1: Distribution of Dataset

2.2 MODEL-1

This model was replicated from Bercu Kir et al. and can be defined as a Sequential object, which means that the layers are arranged sequentially one above another. The input shape of the model is defined as (None, 256,256, 3), which means that the model can take in images of any height or width, as long as they have 3 color channels (RGB).

The architecture of this model is given as - First experiment: Convolution1(12,5,1) - A - Pooling1(2×2) - Convolution2(15,5,1) - A - Pooling2(2×2) - Convolution3(20,4,1) - A - Pooling3 - FullyConnected1 (512) - A - Fully Connected2 (128)

2.3 MODEL-2

In this model, there are 3 convolutional layers each followed by a max pool layer then a flatten layer, a dense layer, and an output layer. The input shape is set to (None, 256, 256, 3), which means that the network can accept images of any batch size with dimensions of 256x256 pixels and 3 color channels.

This model architecture was also replicated from the Burcu kir Savas et al. [12] where its architecture was given as - Convolution1(6,5,1) - A - Pooling1(2×2) - Convolution2(8,5,1) - A -Pooling2 (2×2) - Convolution3(10,4,1) - A - Pooling3 - Fully Connected1 (128).

3 Results

3.1 MODEL-1

```
Testing accuracy: 96.77%
              precision    recall  f1-score   support

  closed_eye       0.96      0.96      0.96       109
    open_eye       0.97      0.96      0.97       109
     yawning       0.96      0.99      0.97       109
     no_yawn       0.98      0.95      0.97       106

    accuracy                           0.97       433
   macro avg       0.97      0.97      0.97       433
weighted avg       0.97      0.97      0.97       433

Confusion matrix:
[[105   3   0   1]
 [  4 105   0   0]
 [  0   0 108   1]
 [  0   0   5 101]]
```

Figure 1: MODEL-1 Results

MODEL-1 that was evaluated on Yawn_Eye_Dataset_New dataset of 433 instances, with 109 instances each for the classes of "closed_eye", "open_eye", "yawning", and "no_yawn". The model attained an accuracy of 96.77%, representing that it correctly classified the highest number of instances.

The precision score of a class represents the proportion of instances that were truly positive among all instances that the model classified as positive.The recall value represents proportion of truly positive instances that the model correctly identified and F1-score is the harmonic mean of precision and recall which gives a single value to assess the performance of MODEL-1 for each class.

For the "closed_eye" class, precision of the MODEL-1 was 0.96, meaning that out of all instances that the model classified as "closed_eye", 96% were truly "closed_eye".Recall for this class was 0.96, which means that the model correctly identified 96% of all "closed_eye" instances and F1-score for this class was 0.96, indicating that the model achieved a high level of accuracy.

For the "open_eye" class, the model achieved a precision of 0.97, indicating that out of all instances that the model classified as "open_eye", 97% were truly "open_eye". Recall score for this class was 0.96 which means MODEL-1 correctly identified all "open_eye" instances. The F1-score for "open_eye" class was 0.97 explaining a high level of accuracy.

For the "yawning" class, MODEL-1's precision was 0.96, indicating that out of all instances that the model classified under the category of "yawning", 96% were truly "yawning". Recall for this class was 0.99 representing that the model correctly characterized 99% of all "yawning" instances and F1-score for this class was 0.97, indicating a good level of accuracy.

For the "no_yawn" class, the model achieved a precision of 0.98, indicating that out of all instances that the model classified as "no_yawn", 98% were truly "no_yawn". Recall for this class came out to be 0.95, representing that MODEL-1 correctly identified 95% of entire "no_yawn" instances and F1-score for this class was 0.97, indicating a good level of accuracy.

The confusion matrix provides a summary of the actual and predicted classifications for each class.

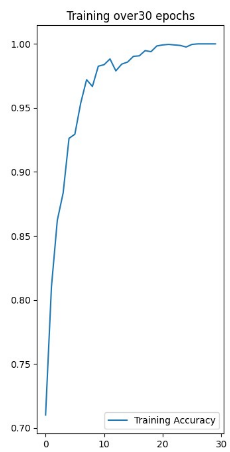

Figure 2: Training Accuracy Graph

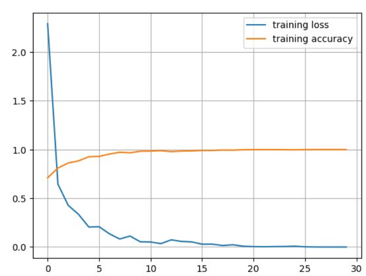

Figure 3: Training Loss and Accuracy

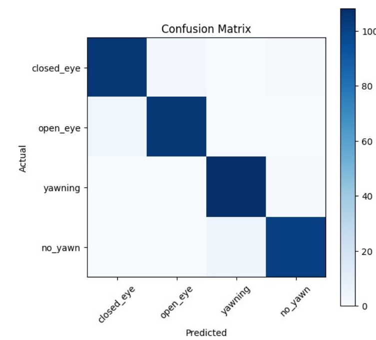

Figure 4: Heatmap-1

3.2 MODEL-2

The results of the classification model's testing accuracy indicate that MODEL-2 correctly classified 95.84% of the instances given in the test part of the dataset. The precision score for each class indicates the percentage of instances classified as that class were actually true positives, while the recall score indicates the percentage of true positives that were correctly classified. The f1-score is the harmonic mean of precision and recall, providing an overall measure of the model's accuracy. Looking at the precision, recall, and f1-score for each class, we can see that the model performed well across all classes, with scores ranging from 0.94 to 0.96. This suggests that the model is able to accurately identify closed eyes, open eyes, yawning, and no yawn conditions.

The confusion matrix shows the number of instances that were classified as each class along with the number of TP , FP, FN and TN for each class.Moreover we can say that just like MODEL-1 our MODEL-2 has also performed quite well.

In conclusion, the results of the MODEL-2's accuracy along with its precision, recall, F1-score, and lastly a confusion matrix indicate that the classification model performed well in accurately identifying closed eyes, open eyes, yawning, and no yawn conditions. Although there were some misclassifications, MODEL-2 also performed well overall, suggesting that it could be useful for detecting eye and yawning behaviors.

```
Testing accuracy: 95.84%
              precision    recall  f1-score   support

  closed_eye       0.96      0.94      0.95       109
    open_eye       0.95      0.96      0.95       109
     yawning       0.96      0.96      0.96       109
     no_yawn       0.96      0.96      0.96       106

    accuracy                           0.96       433
   macro avg       0.96      0.96      0.96       433
weighted avg       0.96      0.96      0.96       433

Confusion matrix:
[[103   6   0   0]
 [  4 105   0   0]
 [  0   0 105   4]
 [  0   0   4 102]]
```

Figure 5: MODEL-2 Result

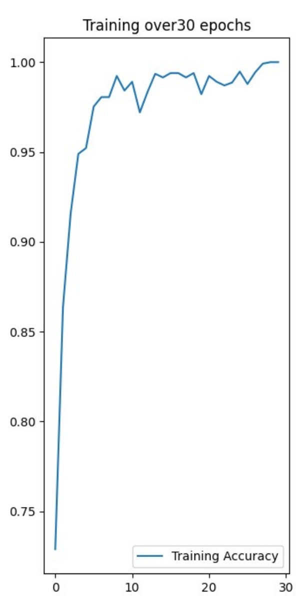

Figure 6: Training Accuracy Graph

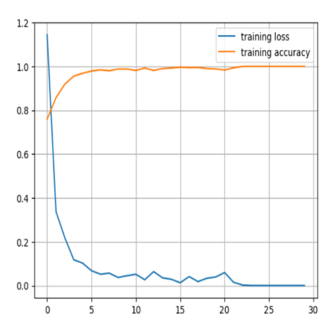

Figure 7: Training Loss and Accuracy

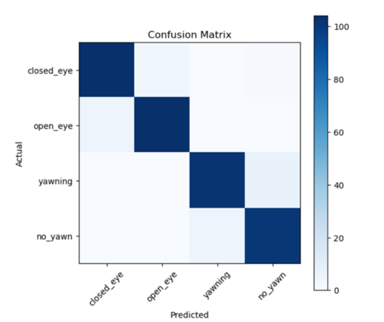

Figure 8: Heatmap-2

4 Conclusion

All the results report on a classification model that was evaluated on a dataset of 433 instances, with 109 instances each for the classes of ”Closed Eye”, ”Open Eye”, ”Yawning”, and ”No Yawn” gives us an accuracy measure of the model in use along with its precision, recall, and F1-score for each class. Additionally, a confusion matrix is included to summarize the actual and predicted classifications for each class. However, there are some differences in the details reported in each result. Here, we could see that the accuracy of Model 1 is 96.77% on test data and Model 2 has an accuracy of 95.84 %. However,still many more techniques could be used to improve upon the data , we could use different machine learning algorithms for feature extractions like facial landmark and can further move to train models on video datasets as well.

In conclusion, we can say that with the help of various image based measures we can identify the operator's state as drowsy / not-drowsy.Implementation of CNNs has shown a huge potential because it has the ability to learn and identify the patterns in images which in turn helps us to correctly identify them into the said or predefined categories.

5 Acknowledgements

It is important for me to thank everyone who made this research paper possible and helped me to understand the practical approah to CNN implementation.Mrs Lisa Meeden helped me with the source code for this experiment as it helped me work through competitive convolution in NEAT on a Pyrobot simulation. Additionally, I would like to extend my regards to Mr.Joe Boninger for making me understand the trigonometric algorithm which are used for robots to make optimal judgement. Lastly,I would like to extend my regards to Mr Ben Xie and Mr Tyler Zon as their contribution taught me to create different files that would run the tests outside of the simulation window allowing me to pactice more trials within a considerable period of time.

References

- https://www.google.com
- https://www.kaggle.com
- https://openai.com/chatgpt
- https://jupyter.org/
- 2012 - W. Zhang, B. Cheng and Y. Lin, ”Driver drowsiness recognition based on computer vision technology,”
- 2017 – B. Mandal, L. Li, G. S. Wang and J. Lin, ”Towards Detection of Bus Driver Fatigue Based on Robust Visual Analysis of Eye State,”
- 2018 - A. Kumar and R. Patra, ”Driver drowsiness monitoring system using visual behaviour and machine learning,”