California State University, Northridge

Drowsiness Detection using InceptionV3 and DLIB

Student: Kunj Shah

Chair: Robert McIlhenny

Committee Members: Katya Mkrtchyan, Mahdi Ebrahimi

Outline

- Purpose
- Problem & Solution
- Different Approaches & Previous Studies
- Limitation
- Data Pre-processing
- Requirements
- Models & Experiments
- Assessment Metrics
- Conclusion
- Future Work

Mission

What's project's main objective or vision?

Why do we need it?

Problem

NHTSA estimates that drowsy driving causes car major accidents

- 100,000 police-reported, 71,000 injuries, and1,500 fatalities.

Solution

A system that detects drowsiness and alerts driver.

Combine DLIB for facial landmark identification and a CNN built on the Inceptionv3 architecture for feature extraction and detect Drowsiness of driver.

Why now?

- Advancements in computer vision and machine learning have made it possible to develop reliable drowsiness detection systems that can analyze facial expressions and other biometric signals to determine if a driver is becoming drowsy. These systems can alert the driver or even automatically take control of the vehicle if necessary.

Different Approaches to Drowsiness Detection.

Eye Aspect Ratio Based

The degree of eye closure can be estimated using the EAR, which is a measurement of the ratio between the length and width of the eye region in an image.

Machine Learning Based

Often involves collecting sensor data and using it to build a machine learning model that may later be applied to real-time drowsiness detection.

Electroencephalography

EEG signals approach to identify variations in patterns of brainwaves that signify drowsiness. These devices track changes in the brain's electrical activity, which can be evaluated to gauge how sleepy the driver is.

Studies previously done on drowsiness detection.

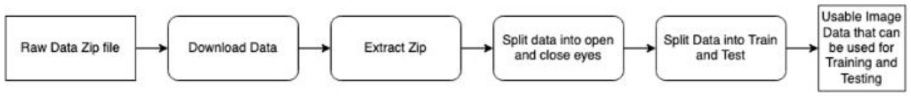
- Image processing-based methods
- Yawning Based Technique
- EEG based detection
- Eye Aspect Ratio
- Machine learning based methods

Limitation

- Image processing-based methods: Limited accuracy, Sensitivity to environmental factors, Limited applicability.
- Yawning Based Technique: Limited accuracy, Dependent on lighting conditions, Limited to frontal view.
- EEG based detection: Lack of accuracy, Limited information, Environmental factors.
- Eye Aspect Ratio: Limited accuracy, Computational complexity, Sensitivity to noise.

Data Pre-processing

- Getting Data
- Understanding data's structure
- Splitting data into two classes
- Splitting data into training data and testing data



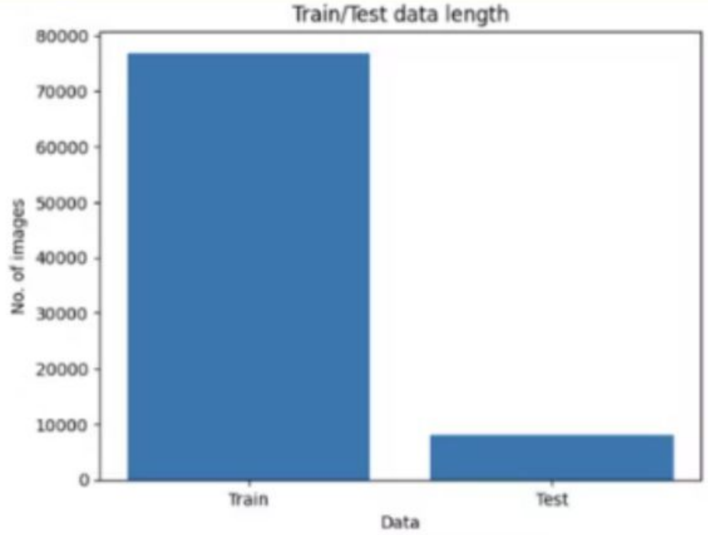
Data Pre-processing

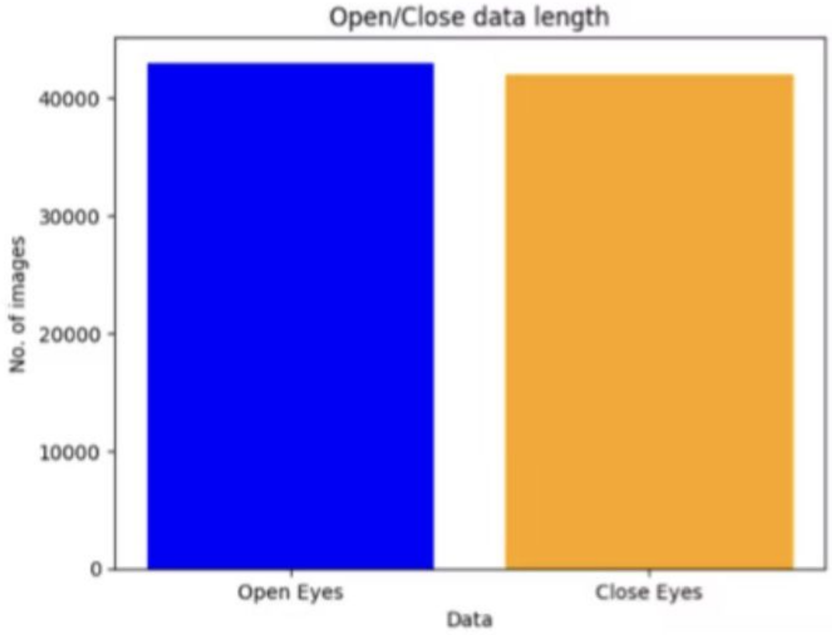
Splitting data into two classes

Shown are the statistics of data split in two classes, i.e. Open Eyes and close Eyes. Each class of data has 40000+ images. With Open Eyes having around about 2000 more images.

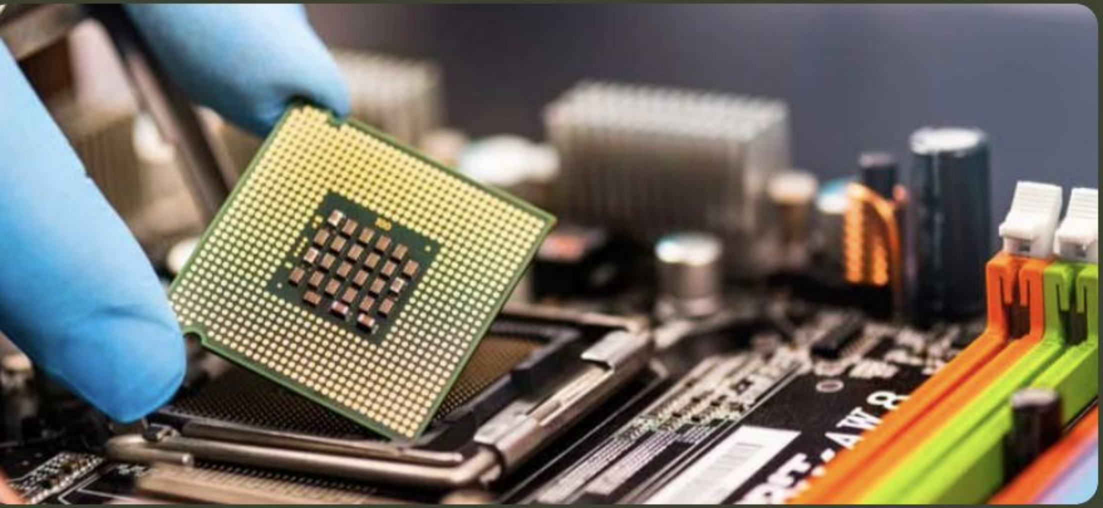
Splitting data into training data and testing data

As soon in the image besides, we have spliced the data into 1:9 ration. So around about 8500 images are used for testing and 85000 images are used in training.





Requirements



Hardware

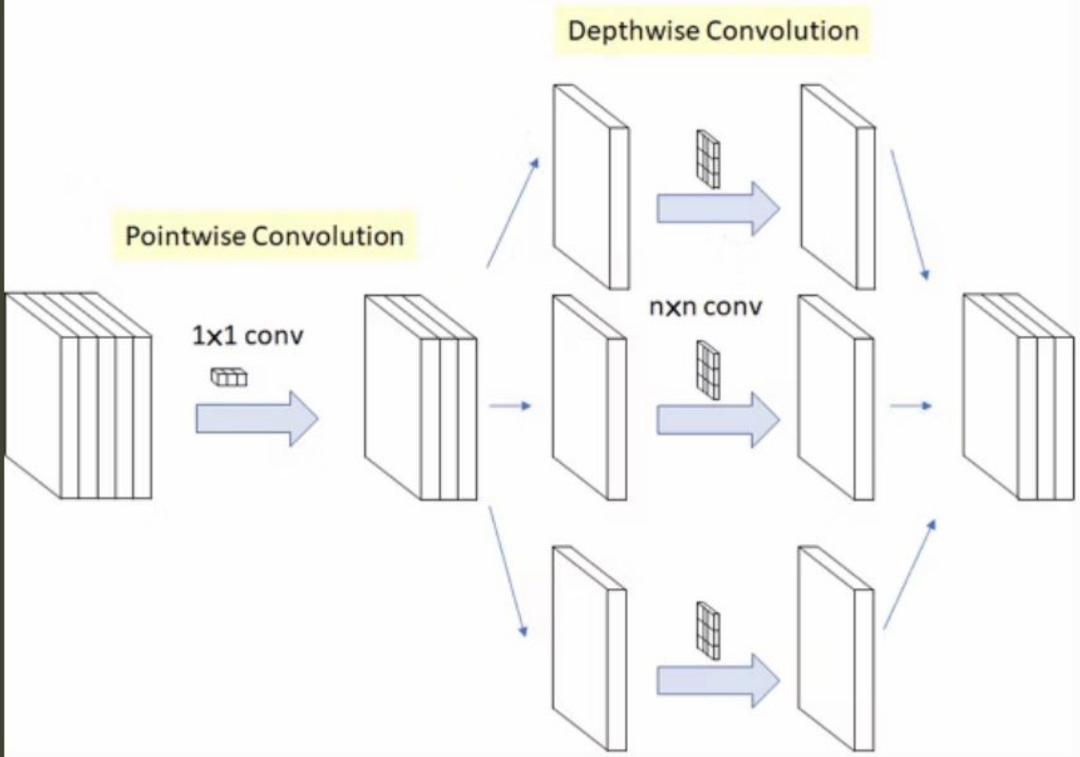
- Camera
- Processor - Raspberry Pi 4
- Memory - 8GB
- GPU (Required for training model)

Software

- Operating System
- Python
- Other required packages: numpy, pandas, scikit-learn, matplotlib, Imutils, OpenCV, Tensorflow, Dlib, Keras, Jupyter Notebook, VSCODE

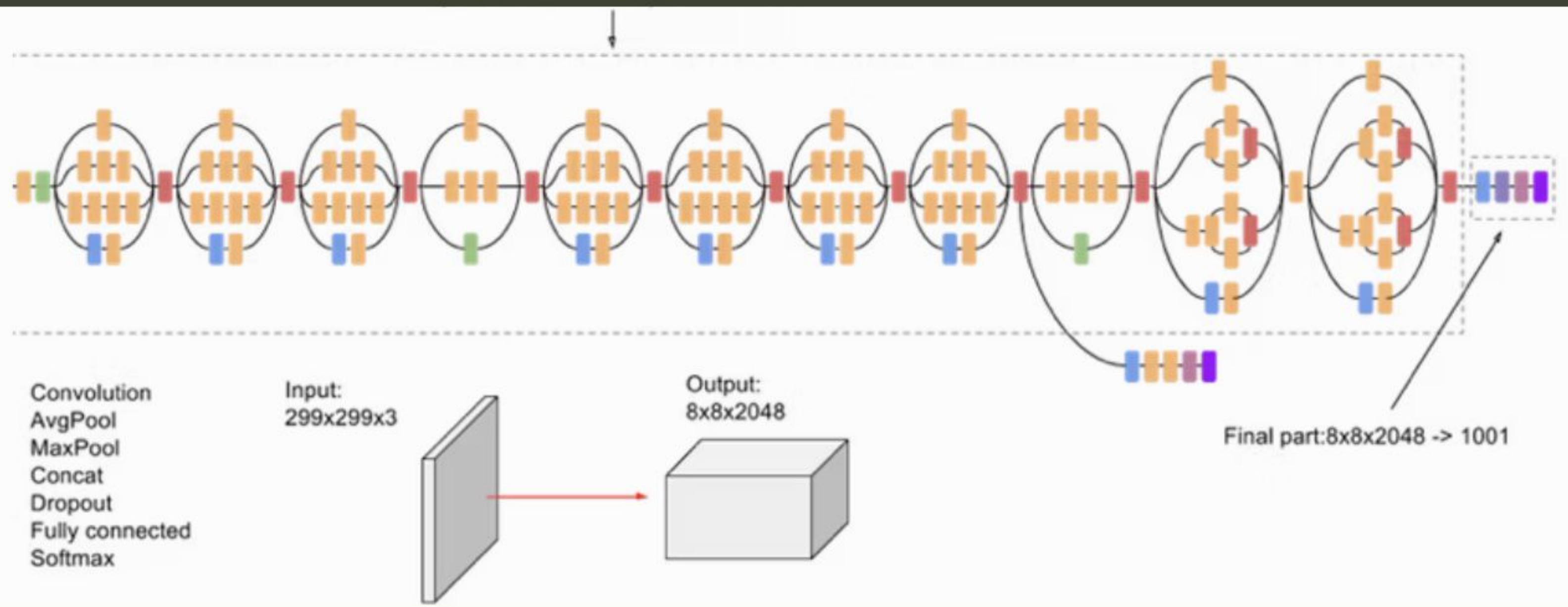
Xception Model

- Xception is a deep CNN architecture that uses depthwise separable convolutions instead of standard convolutions. This allows for a more efficient use of model parameters and faster training times.
- Xception has been shown to perform well on a variety of computer vision tasks, including image classification, object detection, and segmentation.
- Xception is pre-trained on the ImageNet dataset, which contains over a million labeled images. This pre-training allows the model to learn a rich set of features that can be fine-tuned for specific tasks with a smaller labeled dataset.



InceptionV3 Model

- InceptionV3 is a deep CNN architecture. It is an improvement over the previous Inception models, and is designed to be both more accurate and more computationally efficient.
- An inception module is a set of convolutional layers with different filter sizes, which are run in parallel and then concatenated to produce the module's output. This allows the network to learn features at multiple scales and helps to avoid the vanishing gradient problem.
- InceptionV3 has been trained on the ImageNet dataset, which contains over 1.2 million labeled images in 1,000 different categories. This large-scale training allows the model to learn a rich set of features that can be used for a variety of computer vision tasks, including image classification, object detection, and segmentation.



Xception Model.

- Batch size 32

Training Data: 67276 Validation Data: 16818 Testing Data: 805

Train loss: 0.1471 - accuracy:0.9394

Validation loss: 0.2586 - accuracy:0.8996

Test loss: 0.3041 - accuracy: 0.8957

- Batchsize 8

Training Data: 67276 ValidationData: 16818 Testing Data: 805

Trainloss: 0.1479 - accuracy: 0.9402

Validationloss: 0.2592 - accuracy: 0.8974

Test loss: 0.3041 - accuracy:0.8957

- Batchsize 16

Adam Learning Rate 0.001 Default

Training Data: 61516 Validation Data: 15378 Testing Data: 8005

Train loss: 0.1566 - accuracy: 0.9407

ValidationLoss: 0.2755 - accuracy: 0.8800

Test loss: 0.1778 - accuracy: 0.9389

- Batch Size 8

Adam learning rate 0.0001

Training Data: 61516 Validation Data: 15378 Testing Data: 8005

Train loss: 0.1591 - accuracy: 0.9401

Validationloss: 0.2728 - accuracy: 0.8832

Testloss: 0.1778 - accuracy: 0.9389

Experiments

InceptionV3 Model.

- Batch size 16

Training Data 61516 images Validation Data 15378 images Testing Data 8005

Train loss: 0.1506 - accuracy: 0.9406

Validationloss: 0.2452 - accuracy: 0.9011

Testloss: 0.1515 - accuracy: 0.9435

- BatchSize: 8

Training Data 61516 images ValidationData 15378 images Test Data 8005 images

- Train loss: 0.1217 - accuracy: 0.9685
- Validation loss: 0.1883 - accuracy: 0.9354
- Test loss: 0.1252 - accuracy: 0.9614

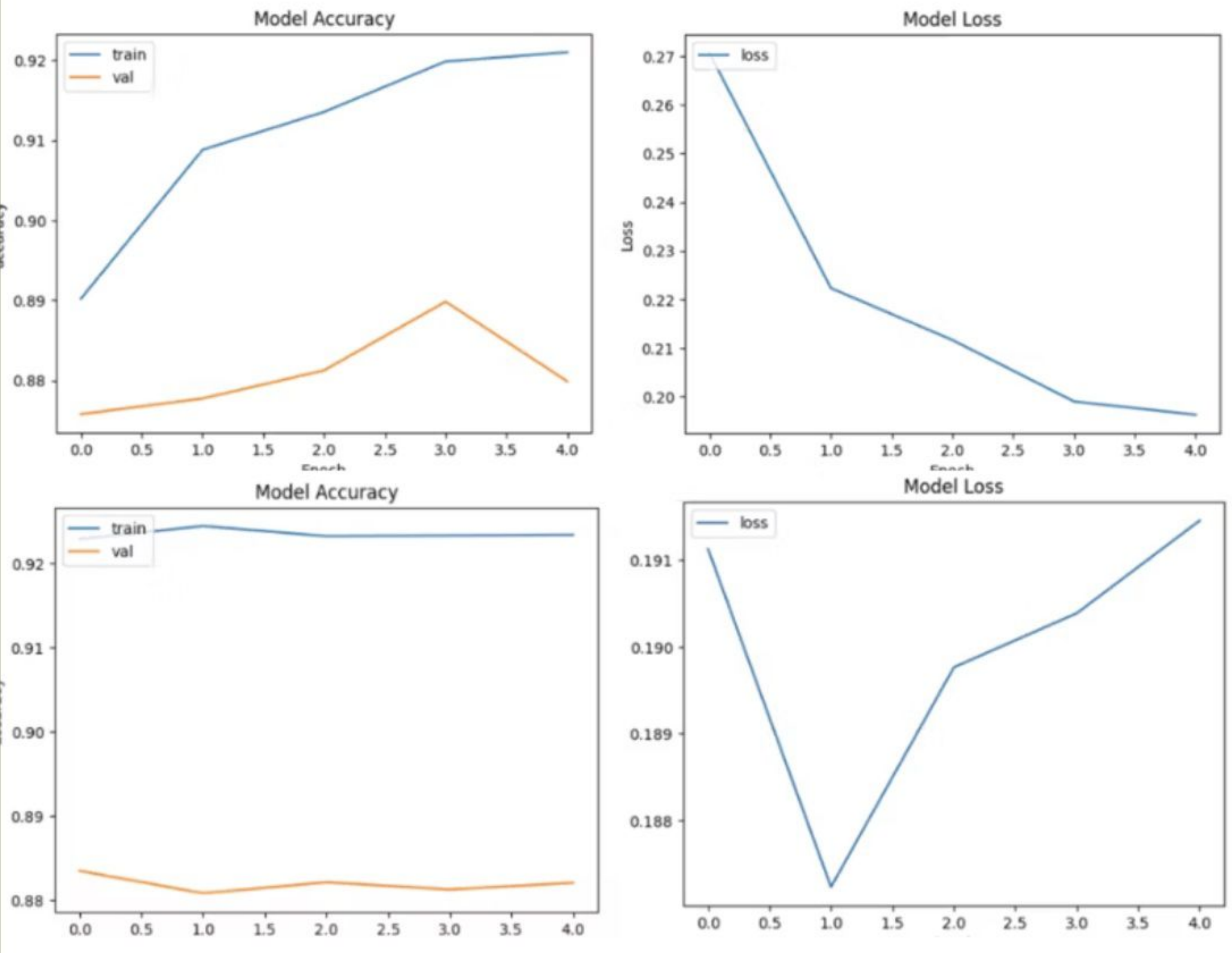
After trying different configuration such as batch size, learning rate, testing image size, we got best results with Batch Size 8, learning rate 0.0001 and testing data size as 10% of the data. With Accuracy of 96.14% on Testing data.

Xception

Given images besides represents Model Accuracy and Model less.

The images in the second half are the images after tweaking with the model.

Model using CNN's Xception did not show any significant change even after tweaking changes such as changing batch size and learning rate.



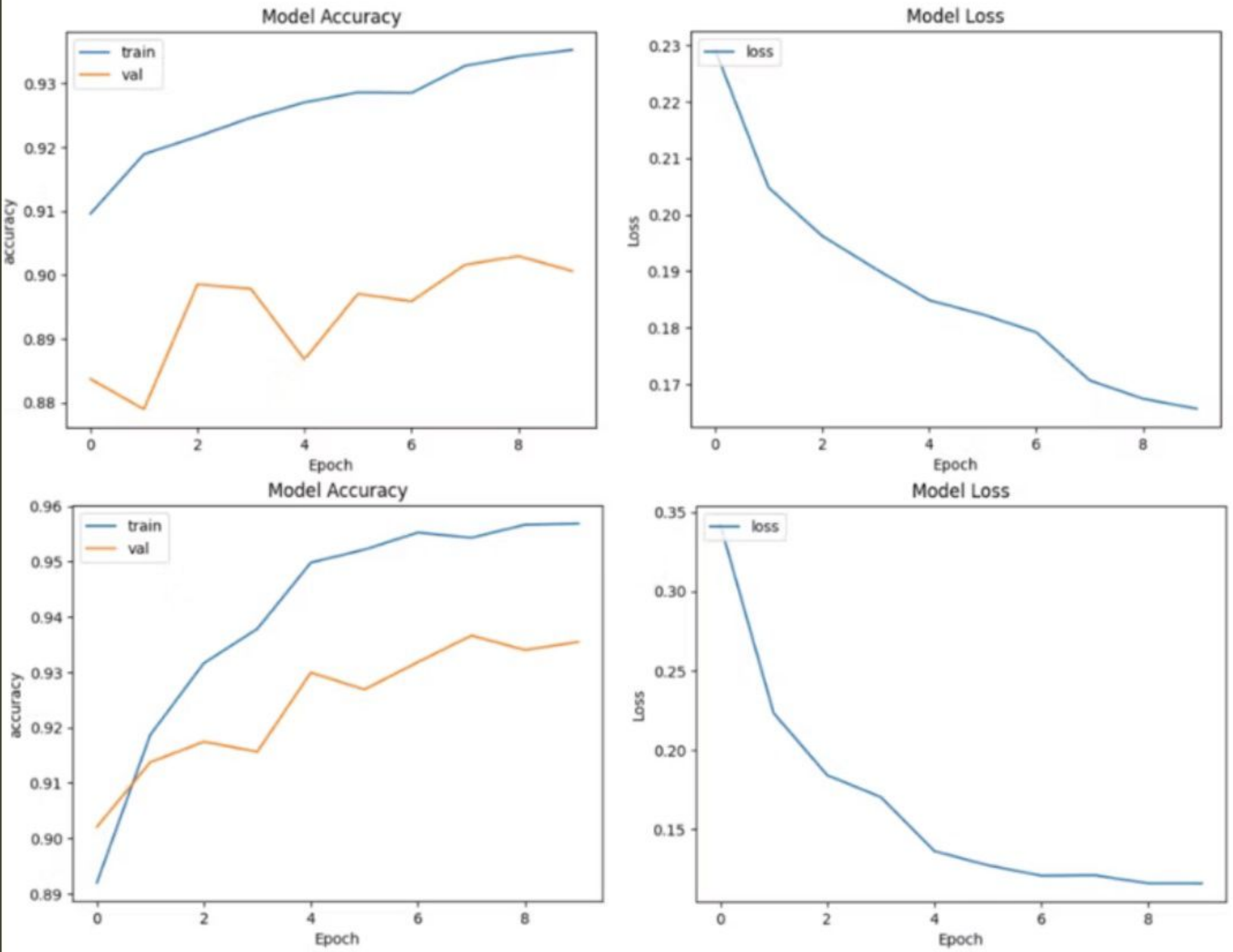
InceptionV3

Given images besides represents Model Accuracy and Model less.

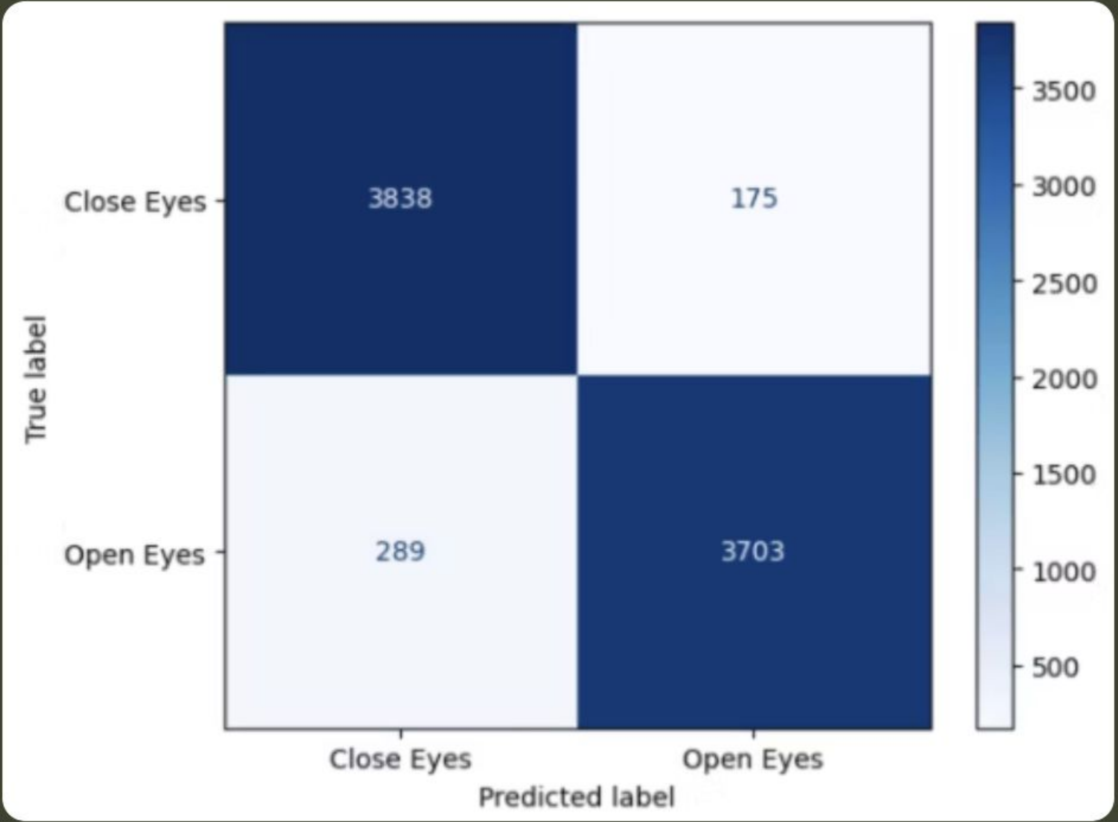
The images in the second half are the images after tweaking with the model.

Model using CNN's InceptionV3 did show change after tweaking changes such as changing batch size and learning rate.

Models accuracy changed from 93.5 to 96ish.



Assessment Metrics



Confusion Metrics (Xception)

Precision for class 0: 0.9299 and class 1: 0.9548

Recall: 0.9276052104208417

Accuracy: 0.9420362273579013

F1 score: 0.9410419313850064

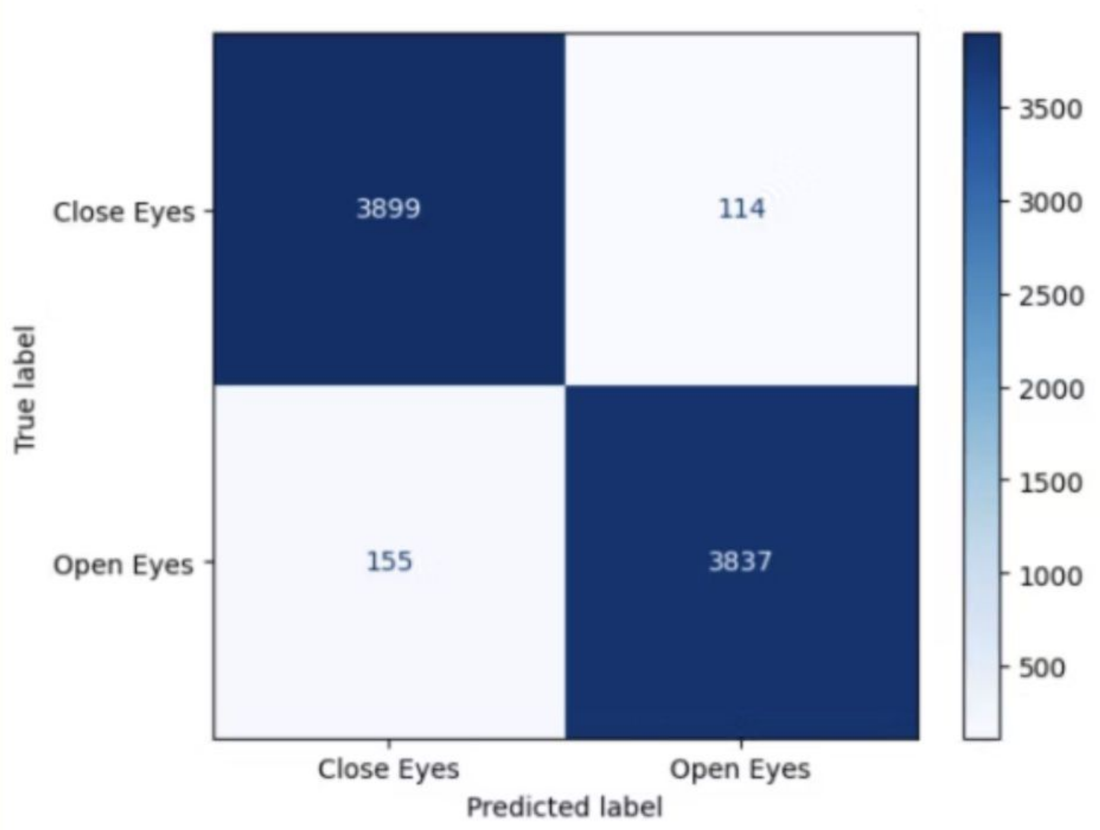
Confusion Metrics (Inceptionv3)

Precision for class 0: 0.9617 and class 1: 0.9711

Recall: 0.9611723446893787

Accuracy: 0.9663960024984385

F1 score: 0.9661337026312476



recall = TP / (TP + FN)

Accuracy = (TP + TN) / (TP + TN +FP + FN)

F1 score = 2 * (precision * recall)/ (precision + recall)

Conclusion

So, from the above information and experiments we got to explore,

- Effectiveness of two deep learning models, Xception and Inceptionv3, in drowsiness detection.
- Determined that Inceptionv3 outperformed Xception in terms of overall performance.
- Inceptionv3 showed better accuracy, precision, recall, and F-1 score compared to Xception.
- The architecture and design of the Inceptionv3 model, with its many inception modules and efficient feature extraction capabilities, are well-suited to the task of drowsiness detection.
- The study's findings suggest that Inceptionv3 can be a promising model for drowsiness detection.

Future Scope

- Multimodal data analysis. Multiple modalities, such as EEG, ECG, and facial expressions, can provide valuable information for drowsiness detection.
- Personalization Model improvement using current data. Personalized to each individual's sleep patterns, habits, and other individual factors
- Integration with other systems. Integrating with other systems, such as autonomous vehicles, health monitoring devices, and alarm systems
- Large-scale studies. Used a relatively small dataset for training and testing the deep learning models. Future research could focus on conducting large-scale studies to test the generalizability and scalability of drowsiness detection systems.

Reference

- *MRL Eye Dataset*. MRL. (n.d.). Retrieved April 6,2023, from http://mrl.cs.vsb.cz/eyedataset
- Arefnezhad, S., Hamet, J., Eichberger, A., Frühwirth,M., Ischebeck, A., Koglbauer, I. V., Moser, M., & Yousefi, A. (2022,February 16). *Driver drowsiness estimation using EEG signals with adynamical encoder–decoder modeling framework*. Nature News. Retrieved April 5, 2023, from https://www.nature.com/articles/s41598-022-05810-x
- Lee, S., & Lee,J. (2018). Non-intrusive driver fatigue monitoring system using yawningextraction. In 2018 2nd International Conference on BioSignal Analysis,Processing and Systems (ICBAPS) (pp. 1-5). IEEE. https://www.hindawi.com/journals/ijvt/2014/678786/
- Mehta, S., Dadhich, S., Gumber, S., & Jadhav Bhatt,A. (2019, June 14). *Real-time driver drowsiness detection system using eyeaspect ratio and eye closure ratio*. SSRN. Retrieved April 5, 2023, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3356401
- *Driver drowsiness detection using EEG power spectrumanalysis*. (n.d.). Retrieved April 6, 2023, from https://ieeexplore.ieee.org/abstract/document/6863035/
- DataTurks. (2018, July 12). *Face detection in images*.Kaggle. Retrieved April 6, 2023, from https://www.kaggle.com/datasets/dataturks/face-detection-in-images

Questions?

Thank you