Project Report

Facial Emotion Recognition

CSCI417/ECEN425: Machine Intelligence

ITCS School

Table of Content

Problem definition and requirements:

Facial expression recognition (FER) is a technology that can analyse facial expressions from both static images and videos in order to tell you what emotions the person is feeling. The technology that is used to create facial expressions can be used in many different ways. Facial emotion recognition is a technology that is used to analyse different sentiments from different sources, like pictures and videos. Affective computing is a type of technology that helps computers understand human emotions, like happiness or sadness. Facial expressions can tell us what people are feeling without having to talk. Decoding emotions from facial expressions is a topic of research in psychology. This is because it can help us understand people's feelings.

Face emotion recognition technology detects emotions and mood patterns invoked in human faces.

● This technology is used as a sentiment analysis tool to identify the six universal expressions, namely, happiness, sadness, anger, surprise, fear and disgust from a neutral face.

● Identifying facial expressions has a wide range of applications in human social interaction detection for industries like digital learning, market research and mental health.

We are going to train a deep learning model using 6 different models like CNN, VGG-16 , ResNet50V2, EfficientNetV2B0, EfficientNetB0, VGG-19 on 7 classes like, happy, fear, disgust, sad, angry, neutral, surprise for the purpose of comparing the accuracy of different models.

Dataset: we got our dataset form Kaggle: Face expression recognition dataset, link: https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset.

Related work:

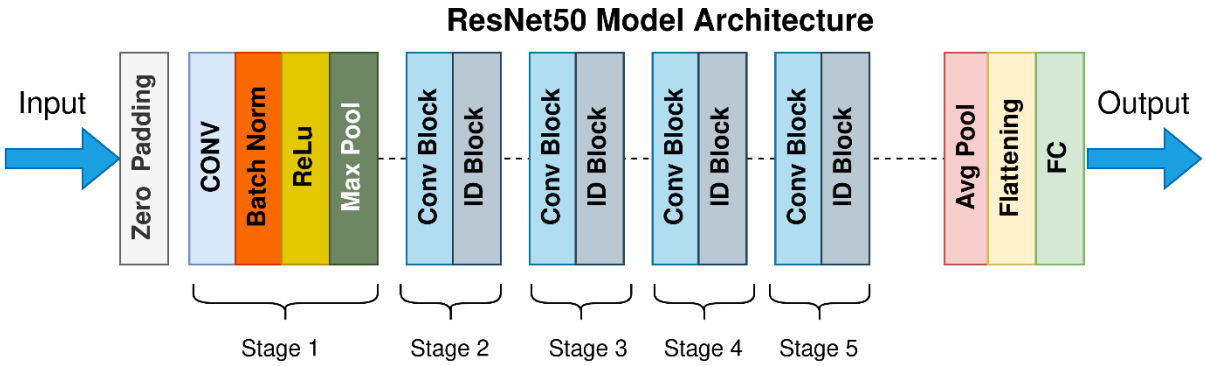
The computer can do many things, like storing information and making things happen on the screen. But it can also be used to help people interact with each other, like when people type things into the computer to get things done. In 2001, (Abdat et al.) studied a group of animals to see if they could tell the difference between different colors. They found that some animals could see colors better than others. (2011). Recently, new technology has made it possible

to detect certain characteristics in photos, like faces, with greater accuracy than ever before. Finally, the emotional state of the faces is classified according to how the features look. This is done by comparing the facial features of different emotional states, and using a variety of factors to decide which one is most likely. Face detection is the first step of the FER analysis. This process looks for faces in the image and records their features. Next, facial expression detection is carried out to determine what emotions the faces are showing. This is done by analyzing the shape and size of the facial features, as well as the position and movement of these features. People use certain facial landmarks to determine the emotions that a person is feeling. For example, the corners of the eyes might be raised in anger, or the mouth might be open in surprise. The end of your nose is called the nostril, and the part of your eyebrows above your eyes is called the eyebrows. Video analysis can help to identify contractions in a group of facial muscles. There are different ways to measure facial expressions, and scientists have found that there are a total of six basic emotions: happiness, sadness, anger, fear, surprise, and disgust. When someone is angry, they may feel disgust. When someone is scared, they may feel fear. When someone is happy, they may feel joy. When someone is sad, they may feel sadness. And when someone is surprised, they may feel surprise. When you're happy, you might feel happy, happy, or happy-sad. When you're surprised, you might feel surprised, surprised, or surprised-happy. When you're disgusted, you might feel disgusted, disgusted, or disgusted-happy. When you're fearful, you might feel fearful, fearful, or fearful-happy. When you're angry, you might feel angry, angry, or angry-happy. When you're surprised-happy, you might feel surprised-happy, surprised-happy, or happy-surprised. In 2014, there were many different kinds of animals. Some animals were big and some were small. Some animals had fur and others didn't. There were also a lot of different kinds of animals in 2014. Smiling means that someone is happy or content. Frowning means that someone is unhappy or has a problem. When you are tired or bored, you may feel like you can't keep your eyes open or your mind on anything. The images or videos that serve as input to FER algorithms come from different sources, such as surveillance cameras, cameras placed close to advertising screens in stores, and cameras placed on social media and streaming services, or from personal devices. Your FER can be used with your biometric identification to make sure you are who you say you are. The accuracy of the technology can be improved by analyzing different types of sources, such as voice, text, health data from sensors, or blood flow patterns inferred from the image. The FER insurance policy can be used for a variety of different purposes, like protecting you from financial losses in the event of a disaster, or helping you get a loan.

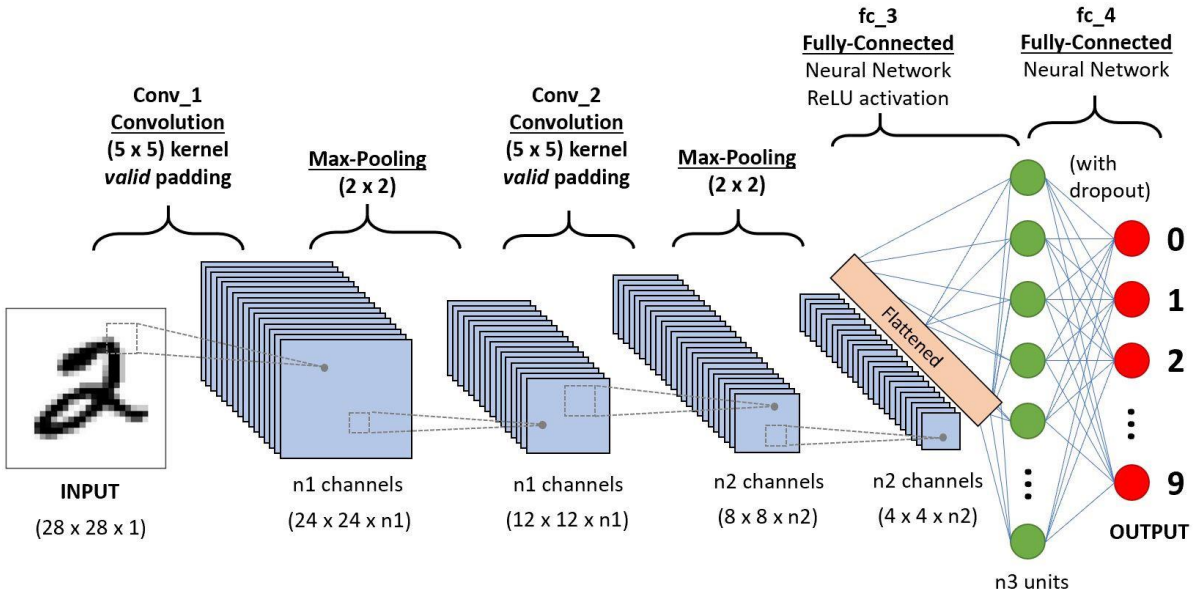
Many studies focused on just one model for emotion recognition but we wanted to use multiple models and compare between them.

Model architecture:

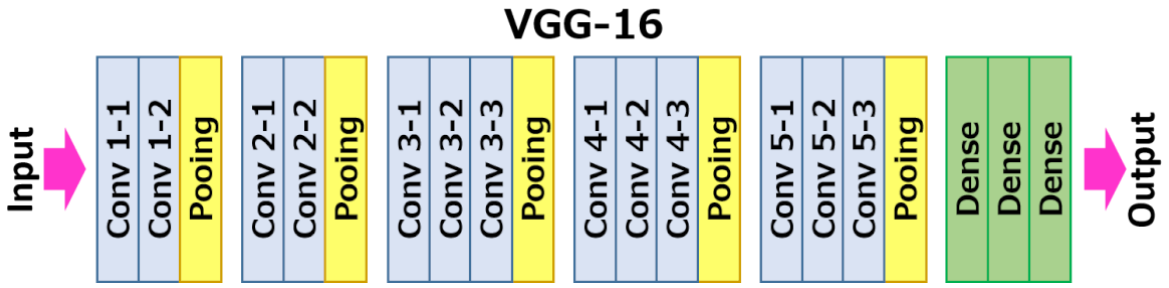
ResNet50:



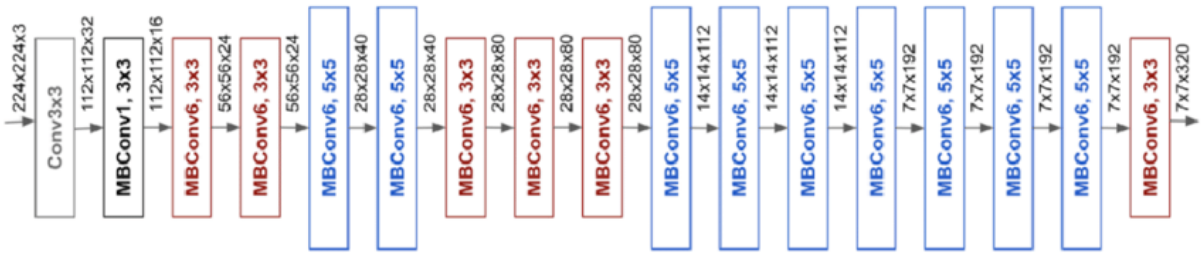
CNN architecture



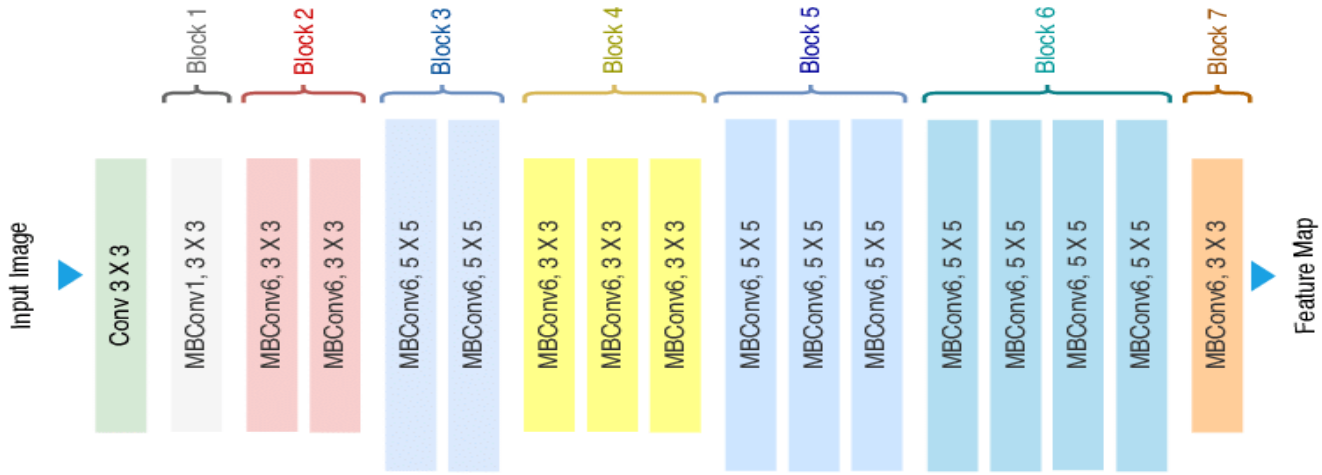
VGG-16 architecture



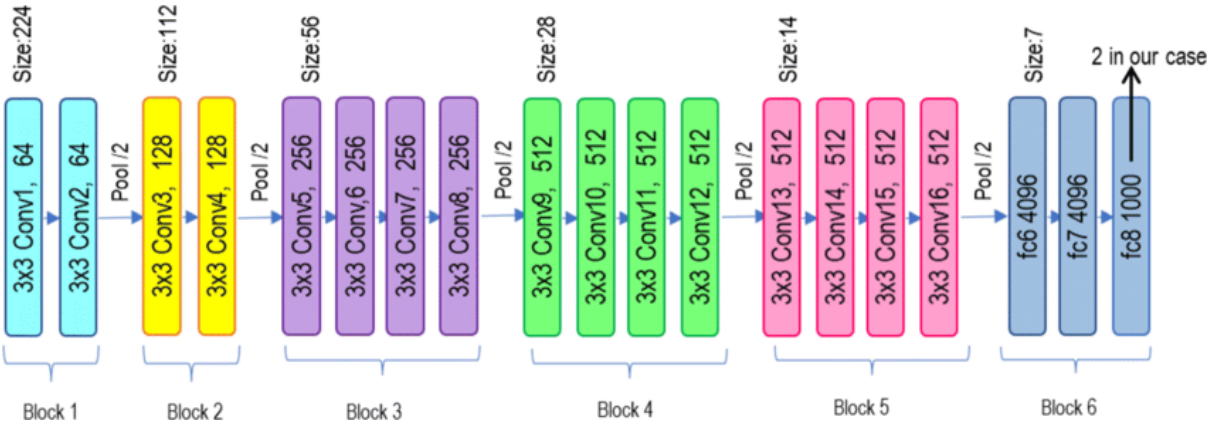
EfficientNetB0



EfficientNetV2B0



VGG-19



Applications:

It can be used in many different fields like:

Provision of personalised services

• analyse emotions to display personalised messages in smart environments

• provide personalised recommendations e.g. on music selection or

cultural material
• analyse facial expressions to predict individual reaction to movies

Customer behaviour analysis and advertising

• analyse customers' emotions while shopping focused on either goods or their arrangement within the shop
• advertising signage at a railway station using a system of recognition and facial tracking for marketing purposes

Healthcare

• detect autism or neurodegenerative diseases
• predict psychotic disorders or depression to identify users in need of assistance
• suicide prevention
• detect depression in elderly people
• observe patients conditions during treatment

Employment

• help decision-making of recruiters
• identify uninterested candidates in a job inter-
view
• monitor moods and attention of employees.

Education

• monitor students' attention
• detect emotional reaction of users to an educative program and adapt the learning path
• design affective tutoring system
• detect engagement in online learning

Model evaluation and results:

Data Augmentation:

- Image data augmentation is a technique that can be used

to artificially expand the size of a training dataset by creating modified versions of images.

- Augmentation techniques can create variations of the images that can improve the ability of the models to generalize what they have learned towards new images.
- Zoom, Horizontal Orientation and Brightness were the randomly augmented image features
- Apart from augmenting images, all the image pixel values in the dataset are rescaled from the [0, 255] range to the [0,1] range.

Evaluation Metrics:

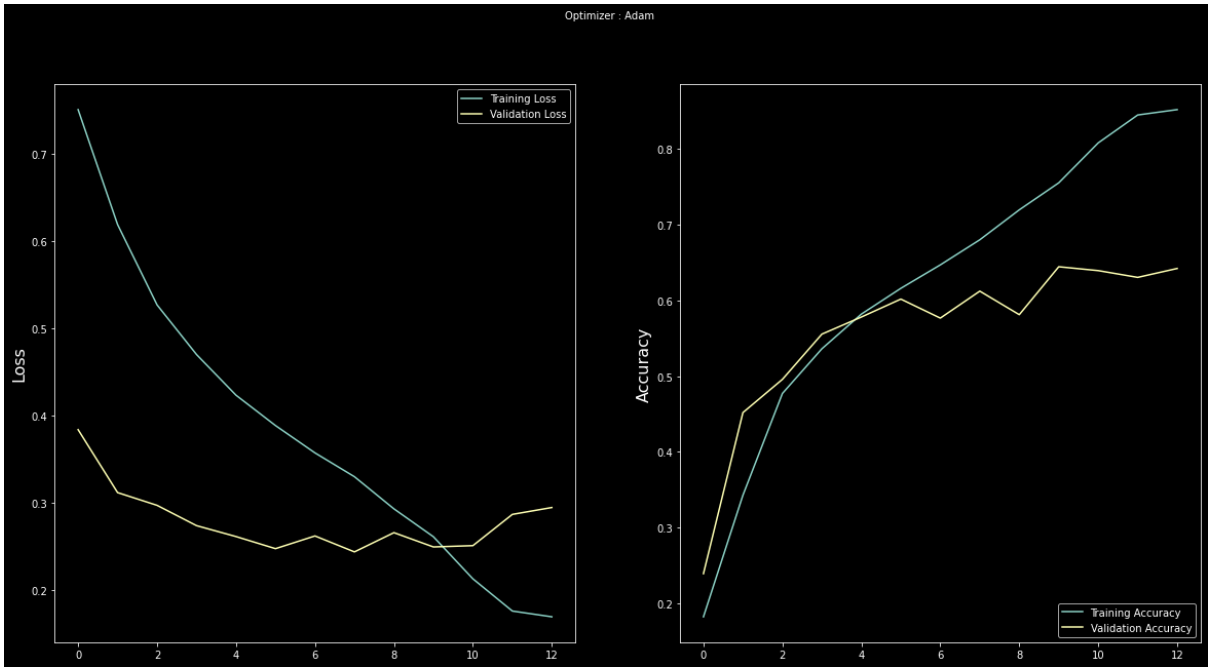
Categorical Cross entropy: is the sum of the product of the predicted value and the logarithm of the target value. Categorical Accuracy calculates the percentage of predicted values that match with actual values for one-hot labels.

Confusion Matrix: is the table that contains the list of true positives, true negatives, false positives and false negative occurrences

Graphs: containing training loss and validation loss and another Graph containing training accuracy and validation accuracy

For VGG-19:

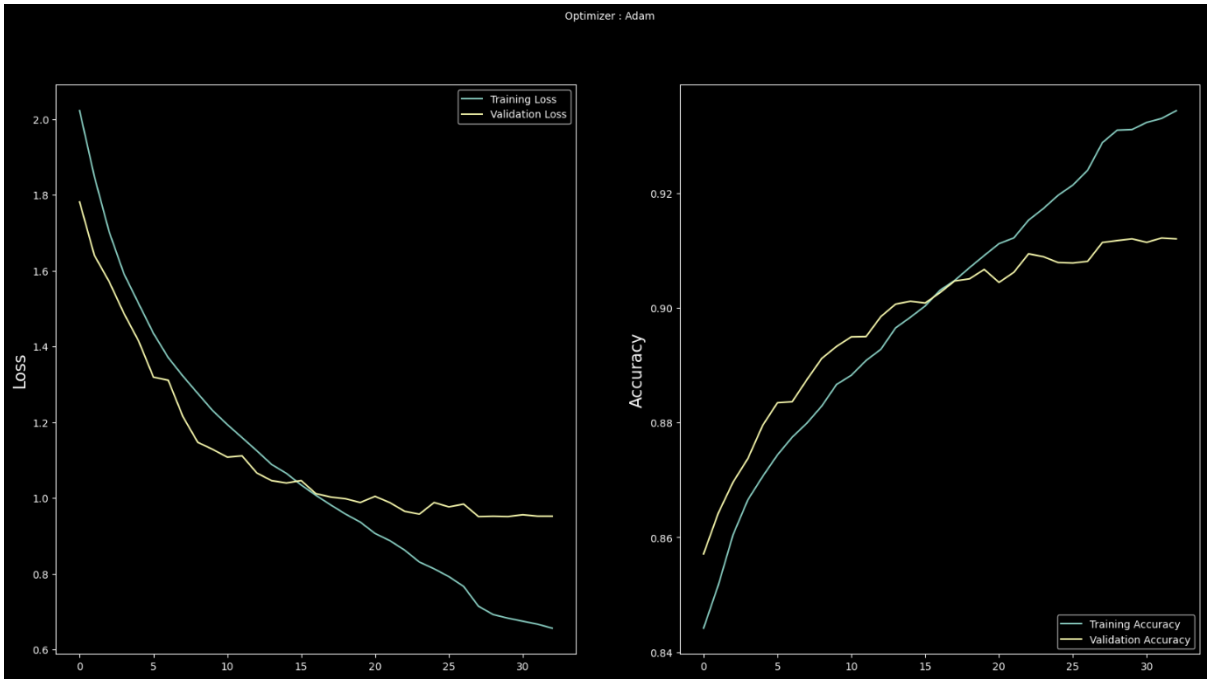
Loss: 0.1689 - accuracy: 0.8523 - val_loss: 0.2943 - val_accuracy: 0.6423


Optimizer : Adam
Training Loss
Validation Loss
Loss
Accuracy
Training Accuracy
Validation Accuracy


Normalized confusion matrix
angry 0.49 0.02 0.10 0.01 0.19 0.18 0.01
disgust 0.21 0.65 0.06 0.01 0.05 0.01 0.01
fear 0.10 0.01 0.42 0.02 0.16 0.23 0.05
happy 0.03 0.00 0.02 0.75 0.08 0.09 0.02
neutral 0.06 0.00 0.04 0.05 0.70 0.13 0.01
sad 0.10 0.02 0.09 0.02 0.25 0.52 0.01
surprise 0.02 0.00 0.14 0.06 0.05 0.03 0.69
True label
angry disgust fear happy neutral sad surprise
Predicted label

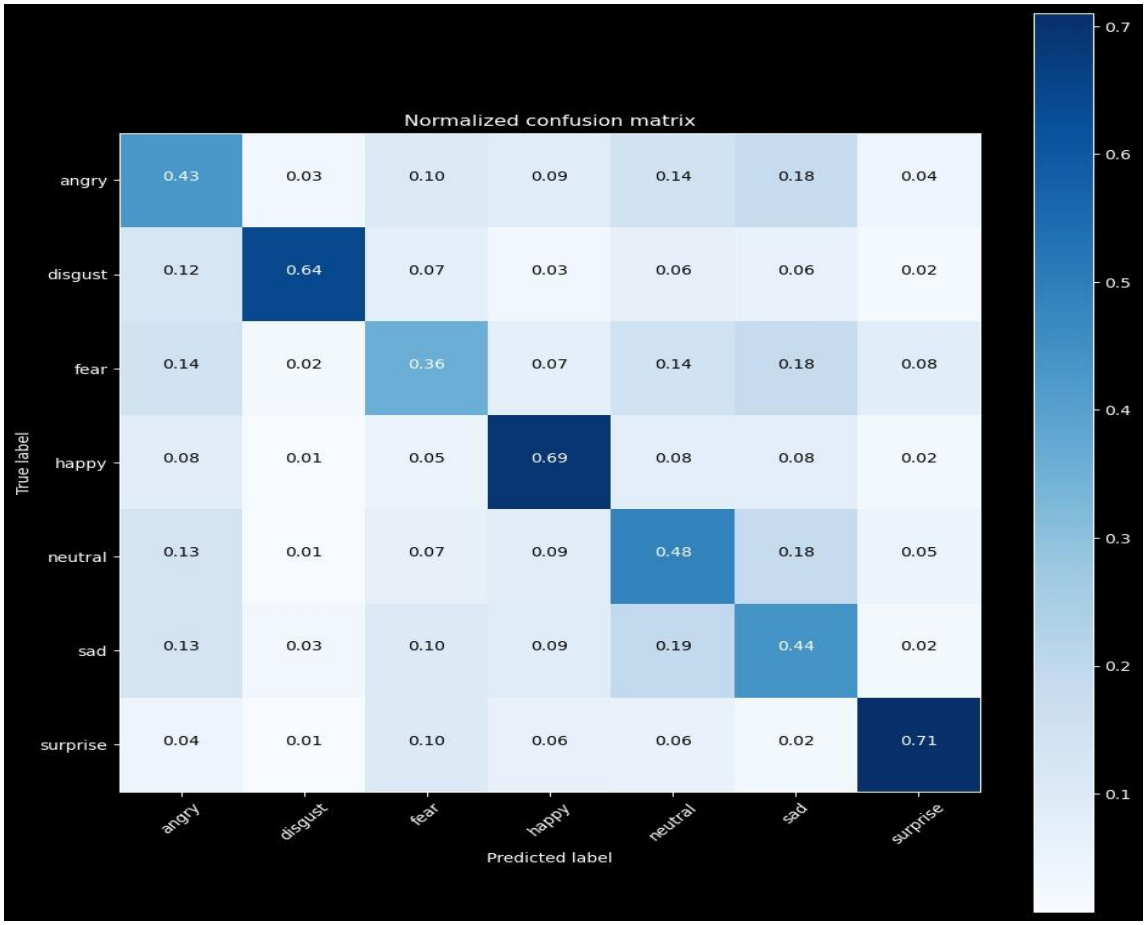
For CNN:

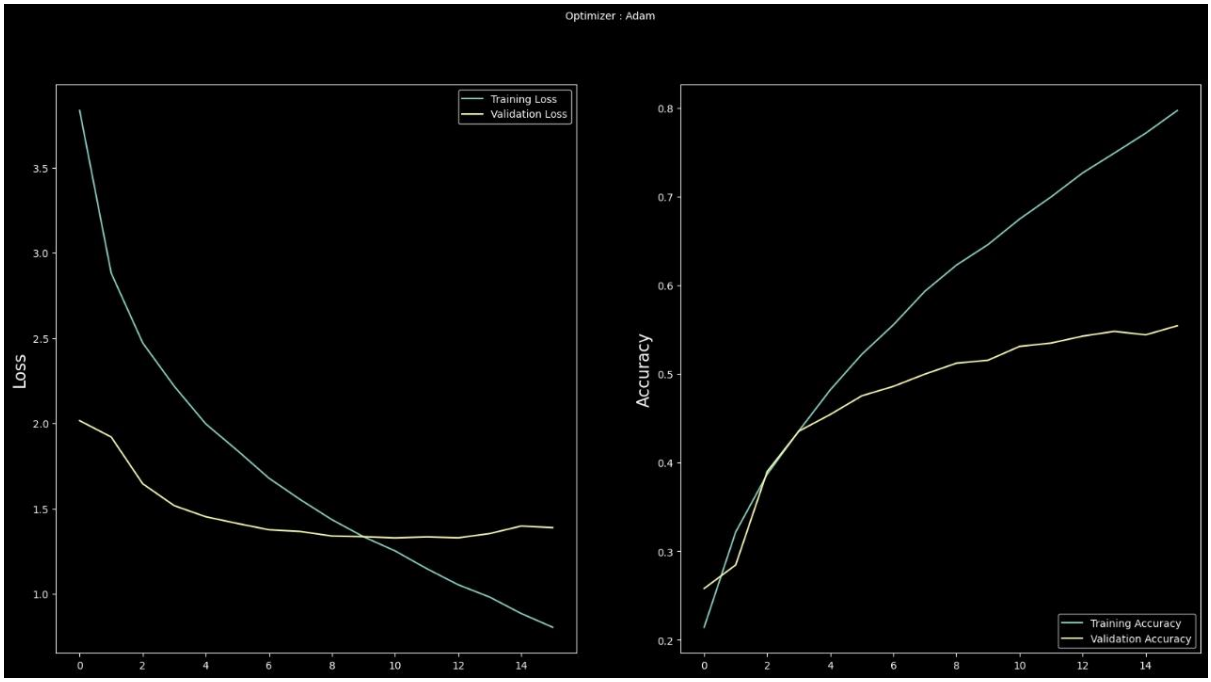
loss: 0.6565 - accuracy: 0.9343 - val_loss: 0.9521 - val_accuracy: 0.9120



For EfficientNetB0

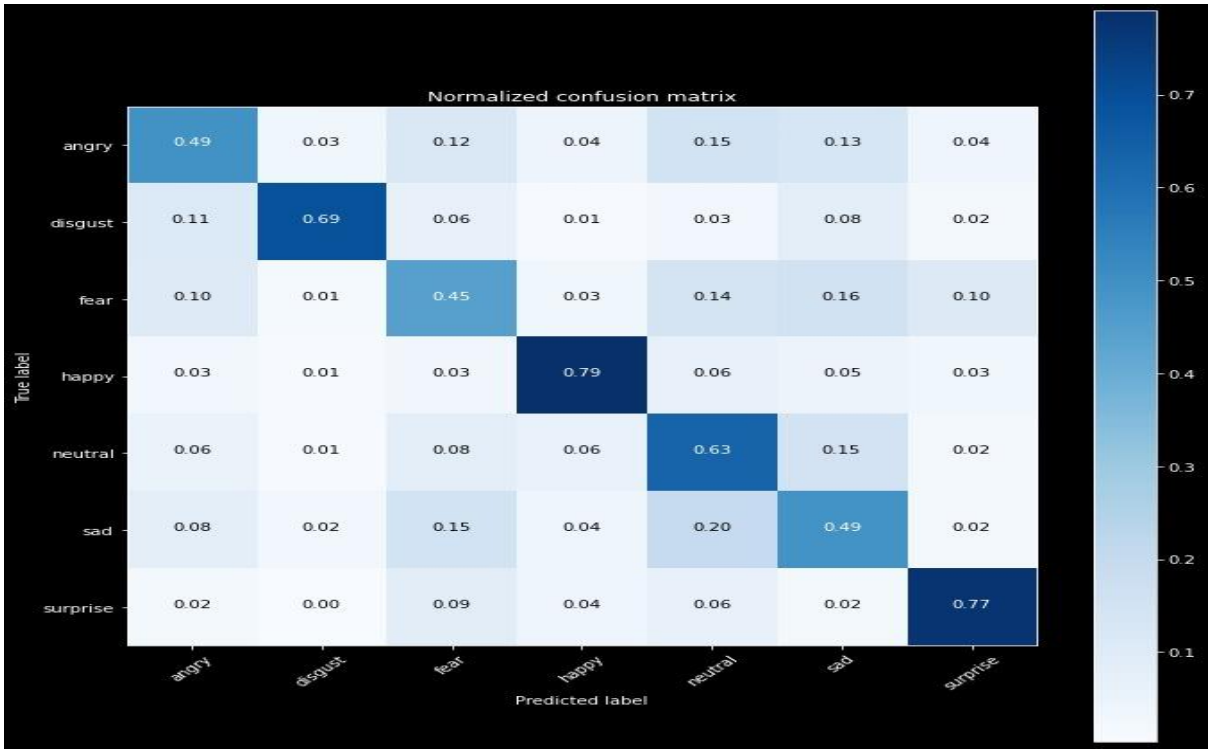
Loss 0.8045, Accuracy 0.7972, Val loss 1.3888, Val accuracy 0.5544

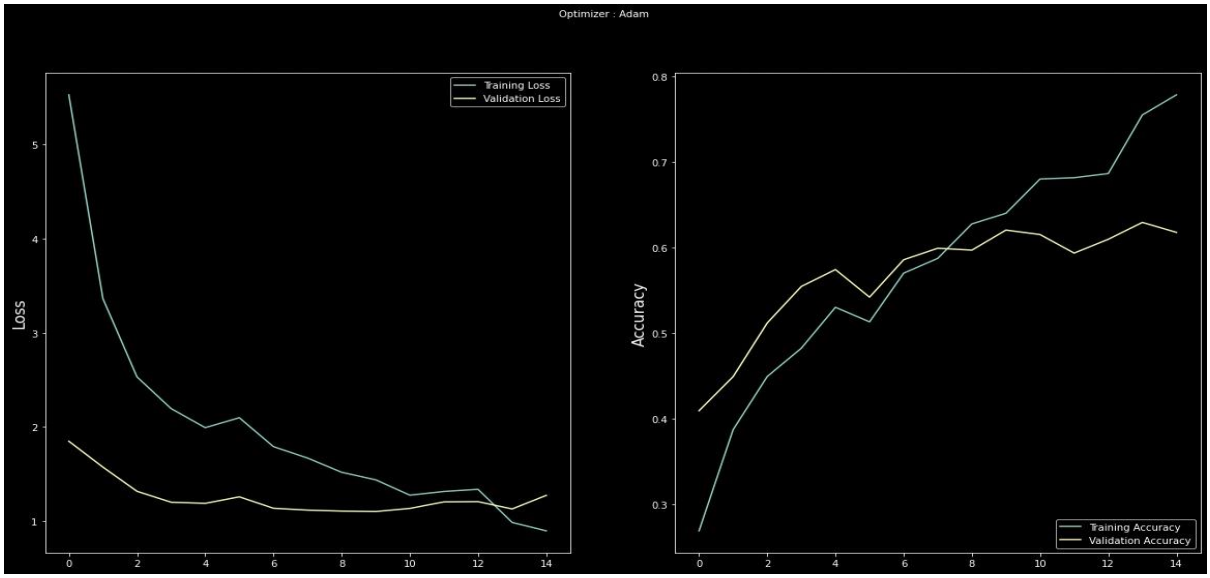




For ResNet50V2:

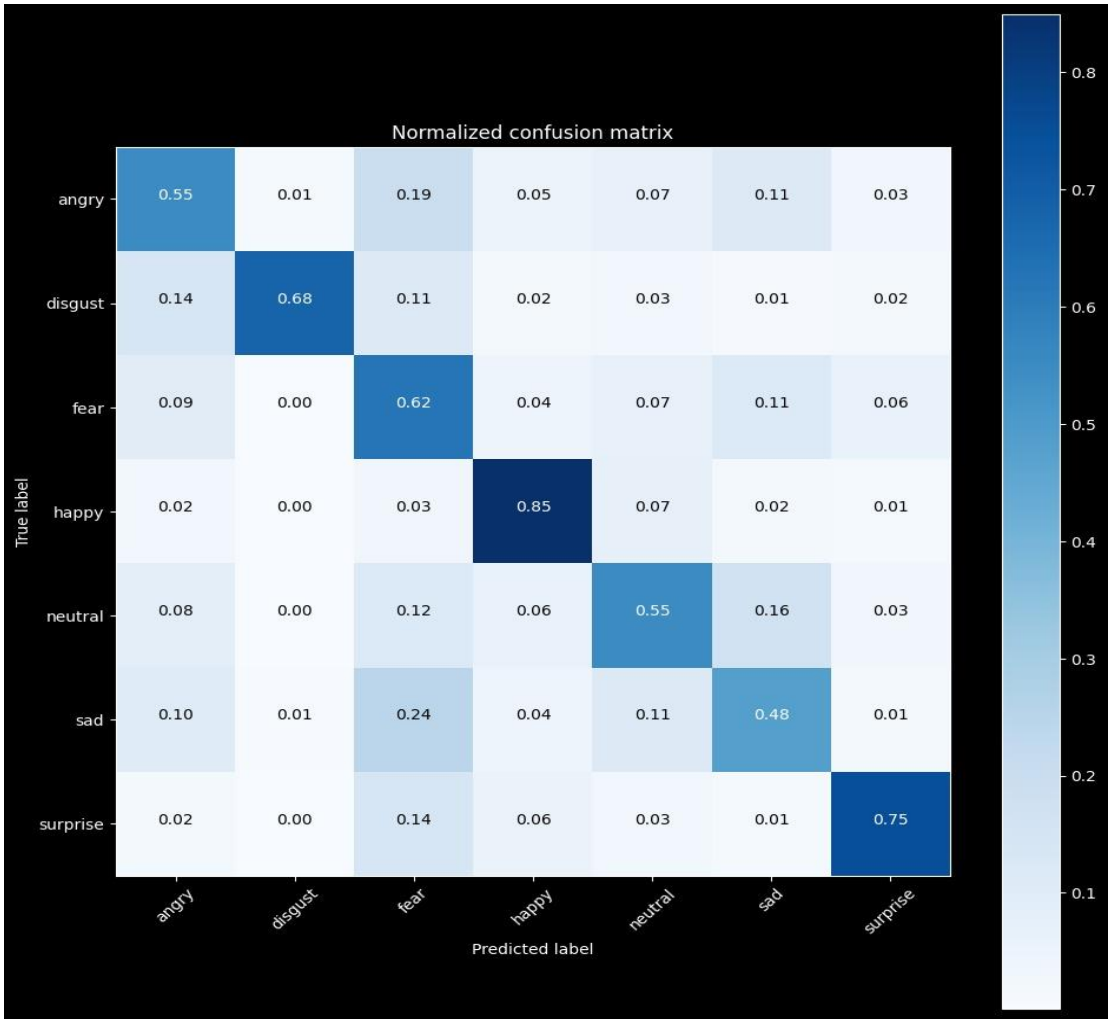
loss: 0.6714 - accuracy: 0.7600 - val_loss: 0.8906 - val_accuracy: 0.7046

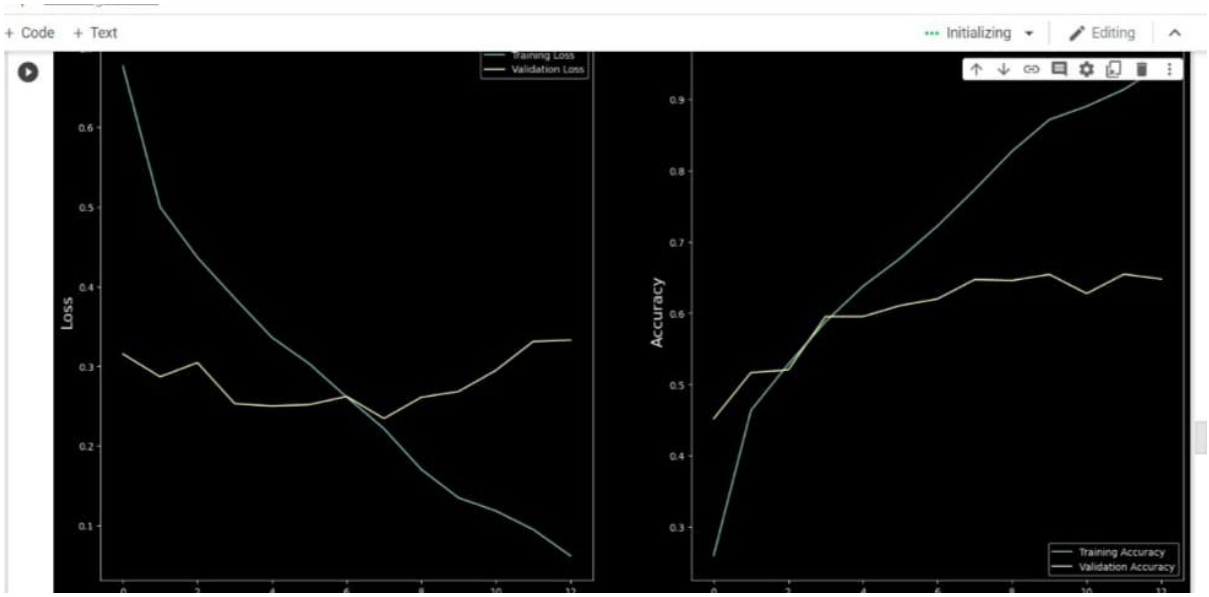




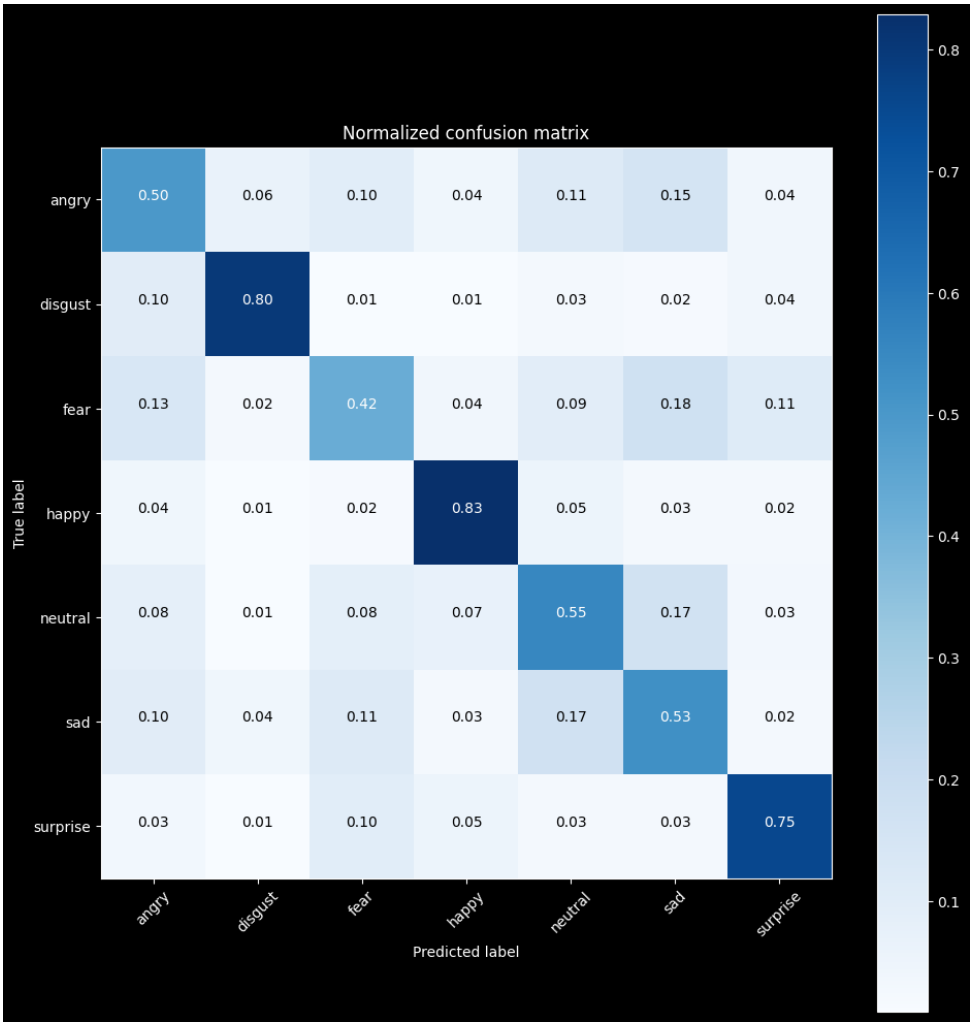
For VGG-16:

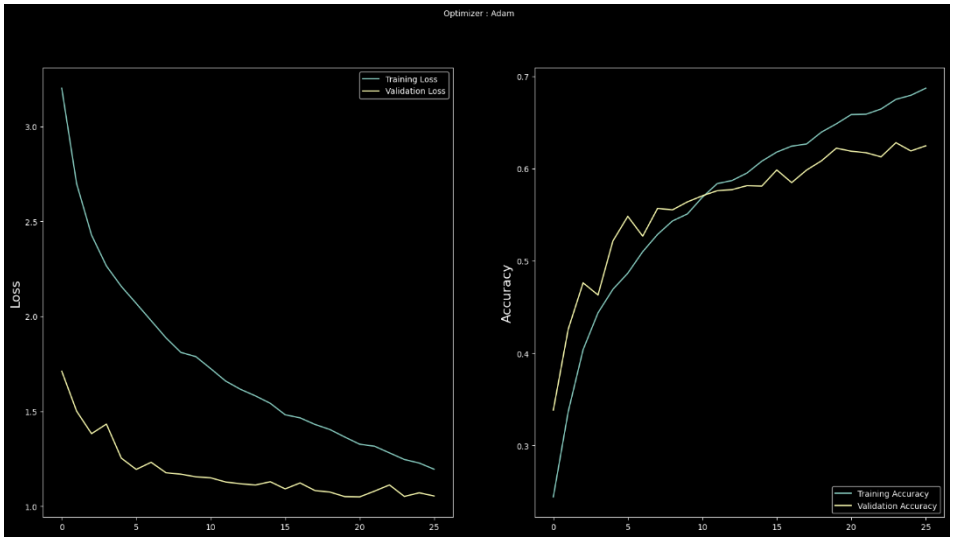
Loss: 0.0619- accuracy: 0.9467 - val_loss: 0.3329- val_accuracy: 0.6477


Normalized confusion matrix
angry 0.55 0.01 0.19 0.05 0.07 0.11 0.03
disgust 0.14 0.68 0.11 0.02 0.03 0.01 0.02
fear 0.09 0.00 0.62 0.04 0.07 0.11 0.06
happy 0.02 0.00 0.03 0.85 0.07 0.02 0.01
neutral 0.08 0.00 0.12 0.06 0.55 0.16 0.03
sad 0.10 0.01 0.24 0.04 0.11 0.48 0.01
surprise 0.02 0.00 0.14 0.06 0.03 0.01 0.75
True label
angry disgust fear happy neutral sad surprise
Predicted label
0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1


+ Code + Text
Initializing
Editing
Training Loss
Validation Loss
Loss
Accuracy
Training Accuracy
Validation Accuracy

For EfficientNetV2B0:





Evaluation Report

Model	Loss	Accuracy	Val_loss	Val_accuracy
CNN	0.6565	0.9343	0.9521	0.9120
VGG-19	0.1689	0.8523	0.2943	0.6423
EfficientNetB0	0.8045	0.7972	1.3888	0.5544
ResNet50V2	0.6714	0.7600	0.8906	0.7046
VGG-16	0.0619	0.9467	0.3329	0.6477
EfficientNetV2B0	1.1954	0.6871	1.0545	0.6247

References:

1-Facial emotion recognition - european data protection supervisor (2022). Available at: https://edps.europa.eu/system/files/2021-05/21-05-26_techdispatch-facial-emotion-recognition_ref_en.pdf

2- Kaggle Dataset: Face expression recognition dataset, link: https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset

3-"Facial expression recognition," *Papers With Code*. [Online]. Available: https://paperswithcode.com/task/facial-expression-recognition.