DL2023
Instructor: Dr. Mayank Vatsa

HUMAN FACE EMOTION RECOGNITION

DL LAB PROJECT REPORT

By-
Riyanshu Jain (B20AI060)
Ruthvik K (B20AI037)
Shubh Soni (B20BB039)

Abstract

Using the FER2013 dataset, we developed a method for recognising human facial expressions. The dataset includes images labelled with seven different emotion classes. To identify the best deep learning model for this task, we trained and compared the performance of several deep learning models, including custom model architecture, ResNet18, MobileNetV2, SqueezeNet1_0, and ShuffleNetV2_x1_0.

Introduction

Because of wide range of applications, such as surveillance, human-computer interaction, and medical diagnostics, human face recognition has been an active area of research in computer vision and machine learning. The accurate identification of facial expressions, which convey information about emotions, is one challenge in face recognition. The FER2013 dataset, which contains 48x48 pixel grayscale images of faces labelled with seven emotion categories, is a widely used dataset for facial expression recognition tasks.

Methods

We used PyTorch in this project to implement, train, and test various deep learning models. The dataset was divided into two parts: a training set and a test set, with data augmentation techniques like random cropping, rotation, and horizontal flipping applied to the training set to improve the model's generalisation capabilities. We tried out the following models:

- Custom Model: A custom CNN architecture was implemented, consisting of four convolutional layers followed by max-pooling and dropout layers, and two fully connected layers.
- ResNet18: A pre-trained ResNet18 model was fine-tuned for the facial expression recognition task.
- MobileNetV2: A pre-trained MobileNetV2 model was fine-tuned for the facial expression recognition task.
- SqueezeNet1_0: A pre-trained SqueezeNet1_0 model was fine-tuned for the facial expression recognition task.
- ShuffleNetV2_x1_0: A pre-trained ShuffleNetV2_x1_0 model was fine-tuned for the facial expression recognition task.

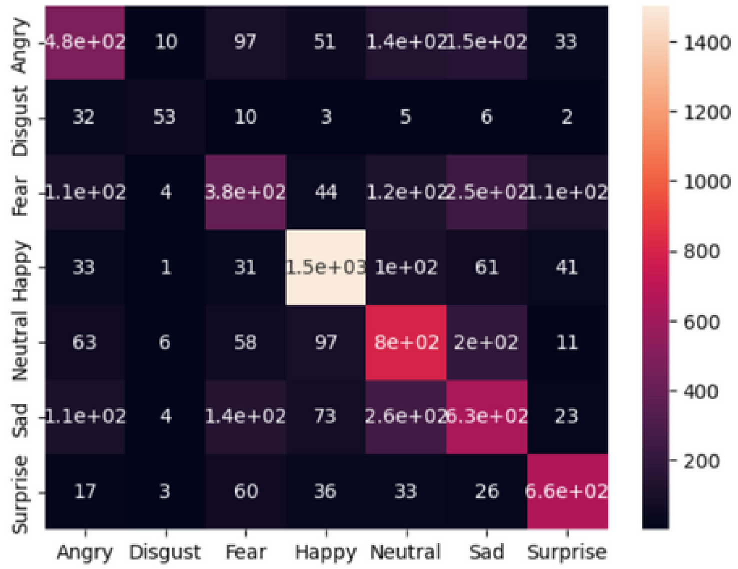
For each model, the optimizer used was Adam, and the loss function was CrossEntropyLoss. The models were trained for 50 epochs.

Results

After training and testing the models, the following accuracies were achieved:

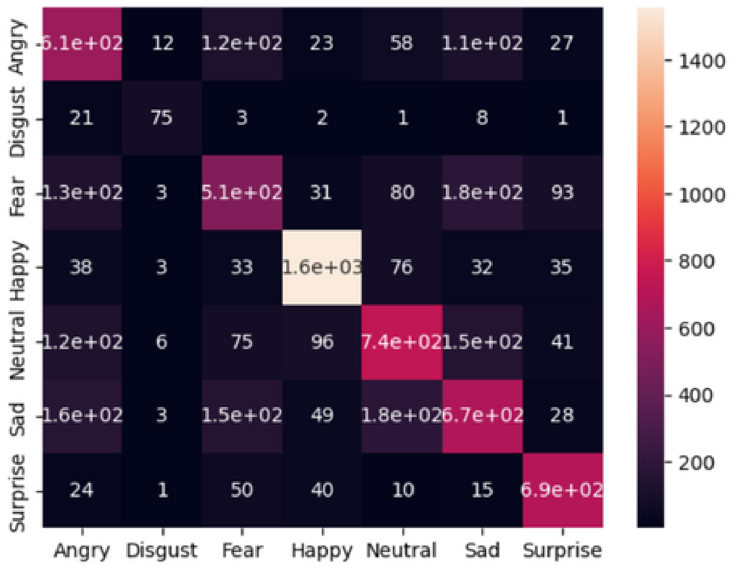
1. Custom Model: Accuracy: 62.76%
2. ResNet18: Accuracy: 67.54%
3. MobileNetV2: Accuracy: 66.95%
4. SqueezeNet1_0: Accuracy: 59.07%
5. ShuffleNetV2_x1_0: Accuracy: 67.69%

Confusion matrices were plotted to visualize the performance of each model across the different emotion classes.



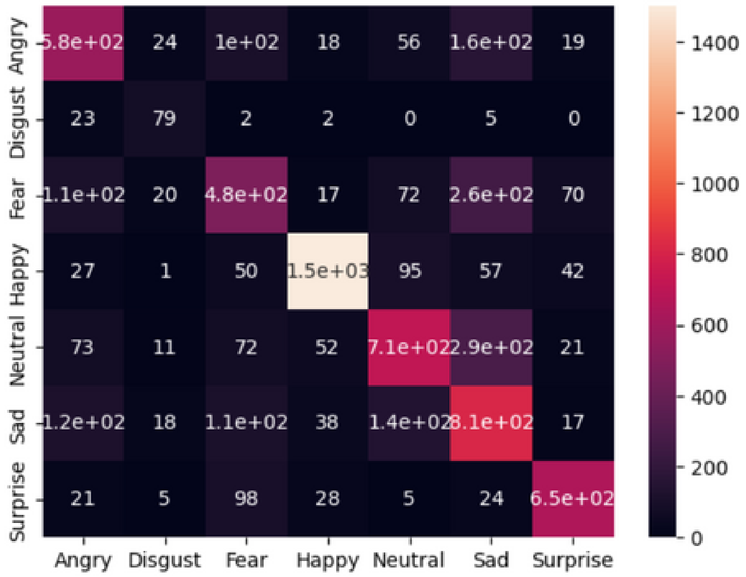
Custom Model

	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	4.8e+02	10	97	51	1.4e+02	1.5e+02	33
Disgust	32	53	10	3	5	6	2
Fear	1.1e+02	4	3.8e+02	44	1.2e+02	2.5e+02	1.1e+02
Happy	33	1	31	1.5e+03	1e+02	61	41
Neutral	63	6	58	97	8e+02	2e+02	11
Sad	1.1e+02	4	1.4e+02	73	2.6e+02	6.3e+02	23
Surprise	17	3	60	36	33	26	6.6e+02



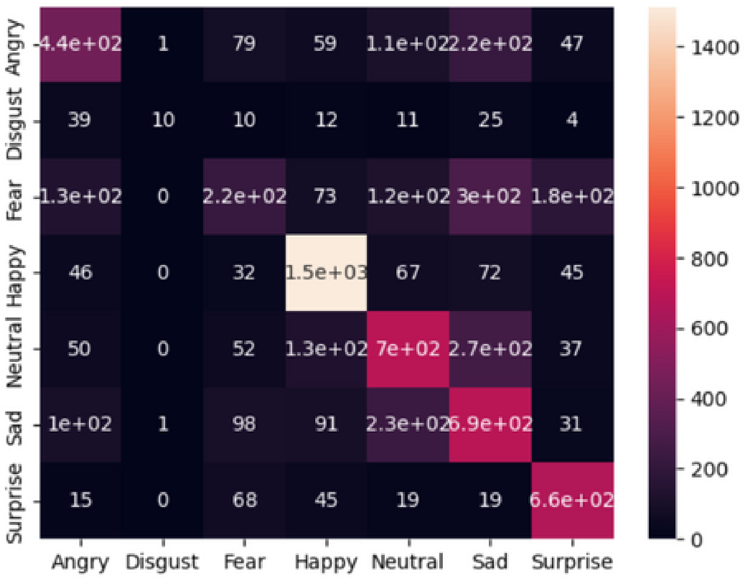
ResNet Model

	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	5.1e+02	12	1.2e+02	23	58	1.1e+02	27
Disgust	21	75	3	2	1	8	1
Fear	1.3e+02	3	5.1e+02	31	80	1.8e+02	93
Happy	38	3	33	1.6e+03	76	32	35
Neutral	1.2e+02	6	75	96	7.4e+02	1.5e+02	41
Sad	1.6e+02	3	1.5e+02	49	1.8e+02	6.7e+02	28
Surprise	24	1	50	40	10	15	6.9e+02



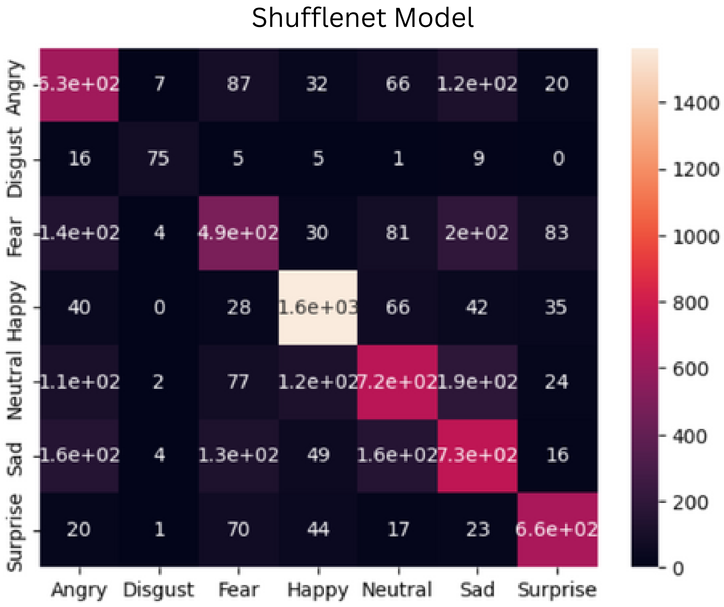
Mobilenet Model

	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	5.8e+02	24	1e+02	18	56	1.6e+02	19
Disgust	23	79	2	2	0	5	0
Fear	1.1e+02	20	4.8e+02	17	72	2.6e+02	70
Happy	27	1	50	1.5e+03	95	57	42
Neutral	73	11	72	52	7.1e+02	2.9e+02	21
Sad	1.2e+02	18	1.1e+02	38	1.4e+02	8.1e+02	17
Surprise	21	5	98	28	5	24	6.5e+02



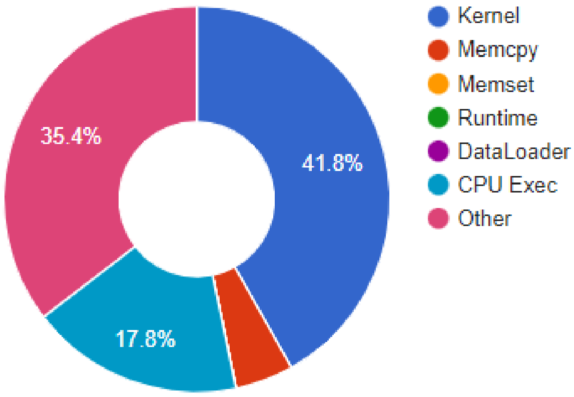
Squeezenet Model

	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	4.4e+02	1	79	59	1.1e+02	2.2e+02	47
Disgust	39	10	10	12	11	25	4
Fear	1.3e+02	0	2.2e+02	73	1.2e+02	3e+02	1.8e+02
Happy	46	0	32	1.5e+03	67	72	45
Neutral	50	0	52	1.3e+02	7e+02	2.7e+02	37
Sad	1e+02	1	98	91	2.3e+02	6.9e+02	31
Surprise	15	0	68	45	19	19	6.6e+02

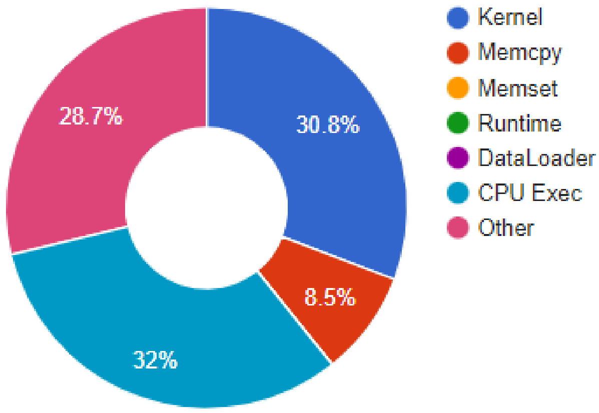


TensorBoard Visulisation Of Each Model

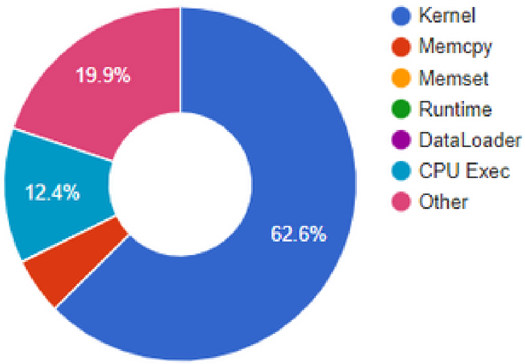
Resnet Model Average Step Time 496,908



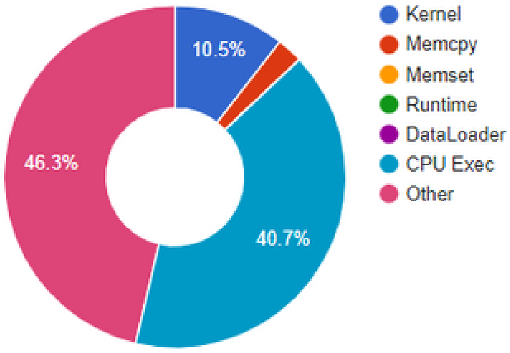
Mobilenet Model Average Step Time 496,908



Squeezenet Model Average Step Time 469,653



Shufflenet Model Average Step Time 988,294

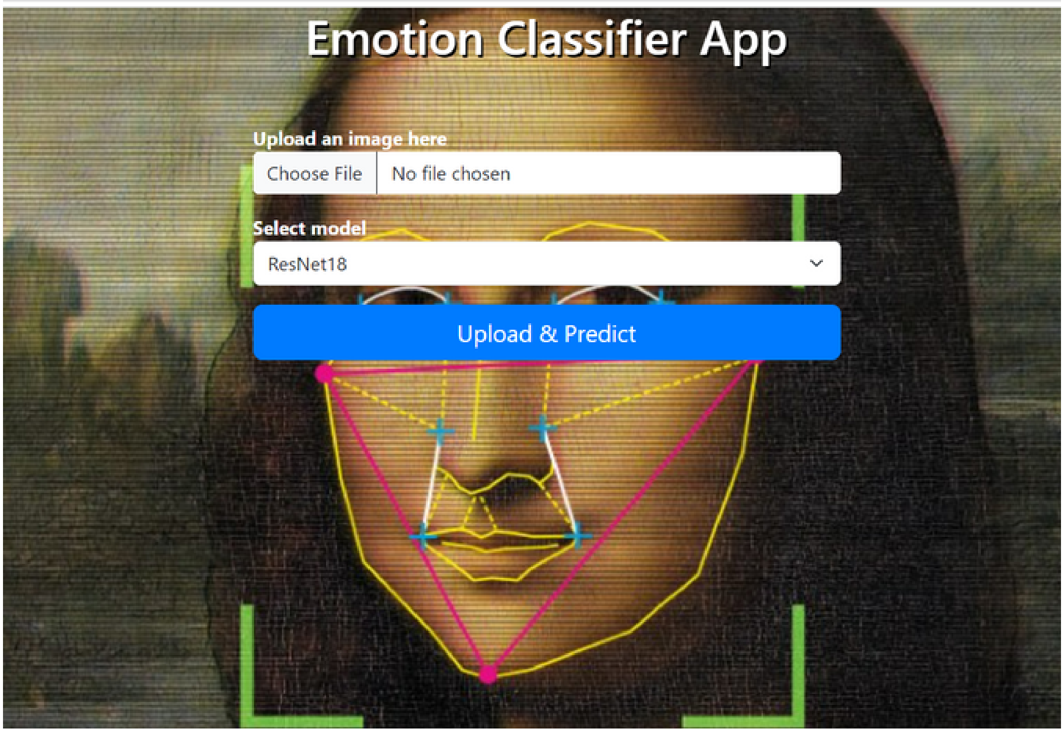


Conclusion

Using the FER2013 dataset, we investigated various deep learning models for human face recognition in this project. The results show that the ResNet18 model is effective at recognising facial expressions as it have higher accuracy and lower average step time. Future research should concentrate on improving model architectures and using larger datasets to improve the performance of facial expression recognition models.

Deploying our Model

- We have first deployed our model on local host using Flask later using StreamLit we have deployed our model on Huggingface.

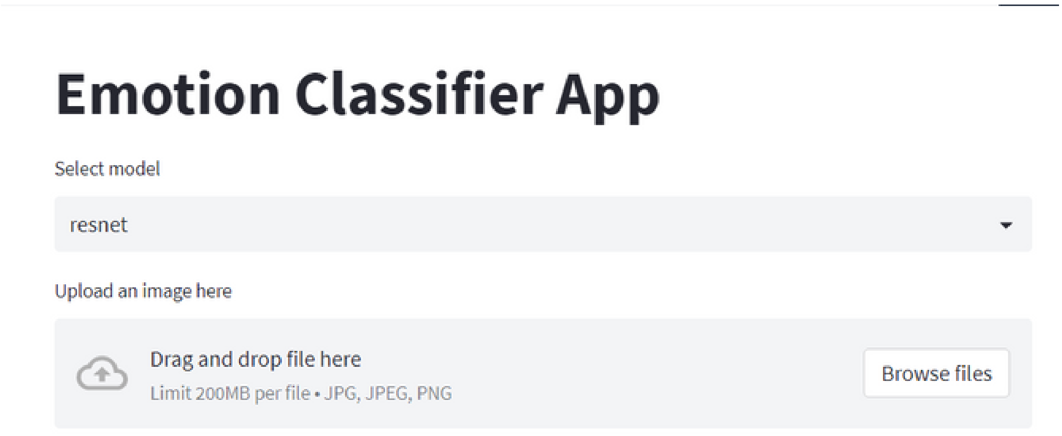


Emotion Classifier App

Upload an image here

Select model

ResNet18



Emotion Classifier App

Select model

resnet

Upload an image here

- Link where the model is deployed using HuggingFace

https://huggingface.co/spaces/ruthvik2/DLOps_Project_1