DEEP LEARNING PROJECT PRESENTATION-GROUP 12

FACE RECOGNITION WITH OCCLUSION AND ORIENTATION

UNDER THE GUIDANCE OF

DR.MRINAL KANTI DAS

ABHISHEK S MAYYA(142302014)

SHREYANSH ACHARYA(142302013)

PROBLEM STATEMENT

- Investigate the performance of an existing face recognition systems under combined challenges of occlusion and orientation variations, aiming to enhance robustness in real-world scenarios where faces may be partially covered and captured from different angles.

DATASET

- We have used Brazilian face dataset.
- Brazilian face database that contains a total of 2800 images, 14 images for each of 200 individuals and also there are 2 separate frontal images for each of the 200 individuals.
- Each image of the same individual is consistently labeled with a name that begins with the same number, we have used this as a FaceID.

PROJECT DESCRIPTION

- This project aims on enhancing face recognition systems to effectively identify individuals in images focusing more on occlusions and variations in facial orientation.
- Leveraging existing face detection and recognition models to optimize performance.

Starting Point:

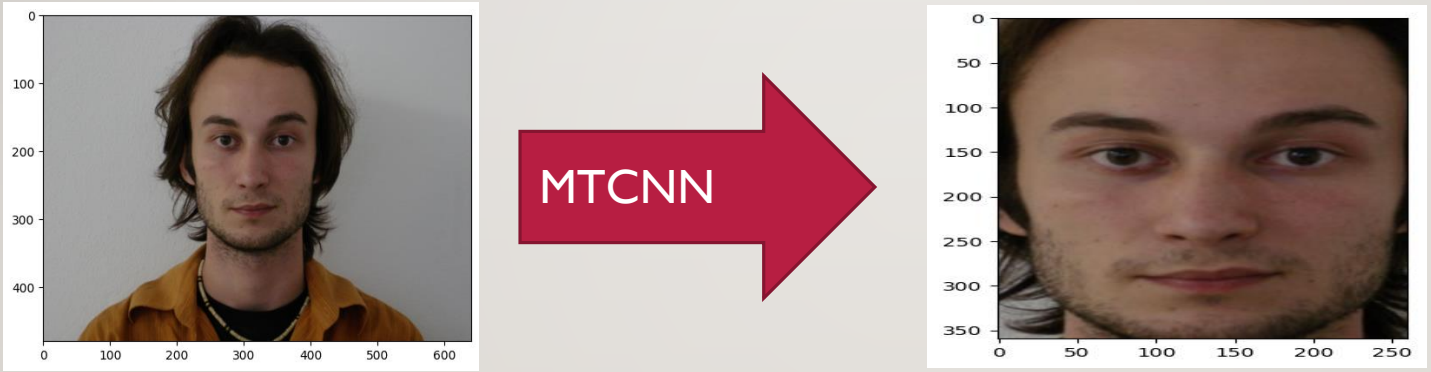
- The baseline for this project consists of established face detection and recognition models.

1. Face detection : MTCNN(Multi-task Cascaded Convolutional Networks)
2. Face recognition : VGG-face (Visual Geometry Group)

   Source: Github link

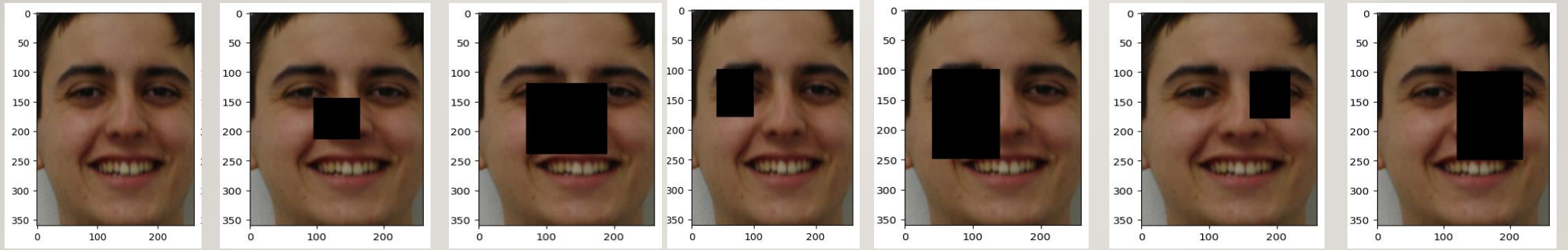
FACE EXTRACTION USING MTCNN

- MTCNN works best for extracting faces from the image.



APPLYING OCCLUSION

- We have used two types of occlusions, central and peripheral(left and right).
- We introduced occlusions of 5% and 15% by manually setting pixels to zero for rectangular shape in the images.



Original face

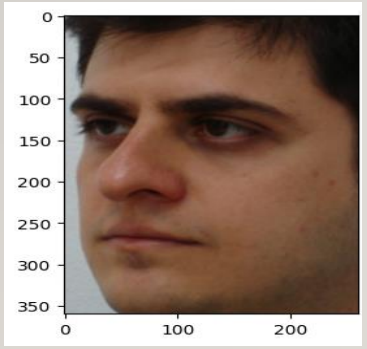
5% 15%
Central occlusion

5% 15%
Left occlusion

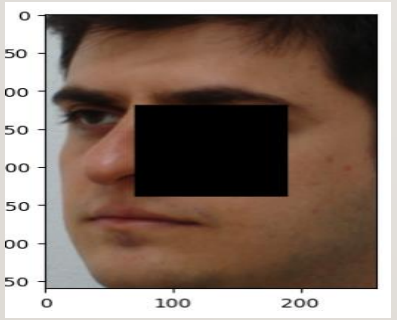
5% 15%
Right occlusion

APPLYING OCCLUSION WITH ORIENTATION

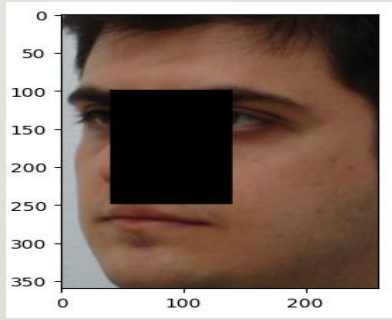
- We introduced occlusions of 15% by manually setting pixels to zero for rectangular shape in the images with different face orientation.



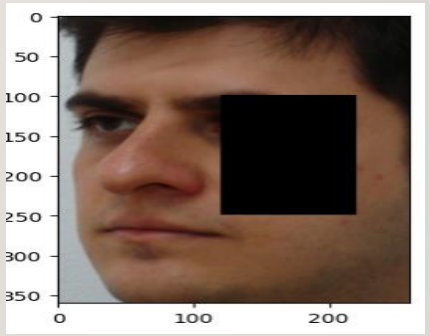
Original face



15%
Central occlusion



15%
Left occlusion



15%
Right occlusion

FACE REGISTER

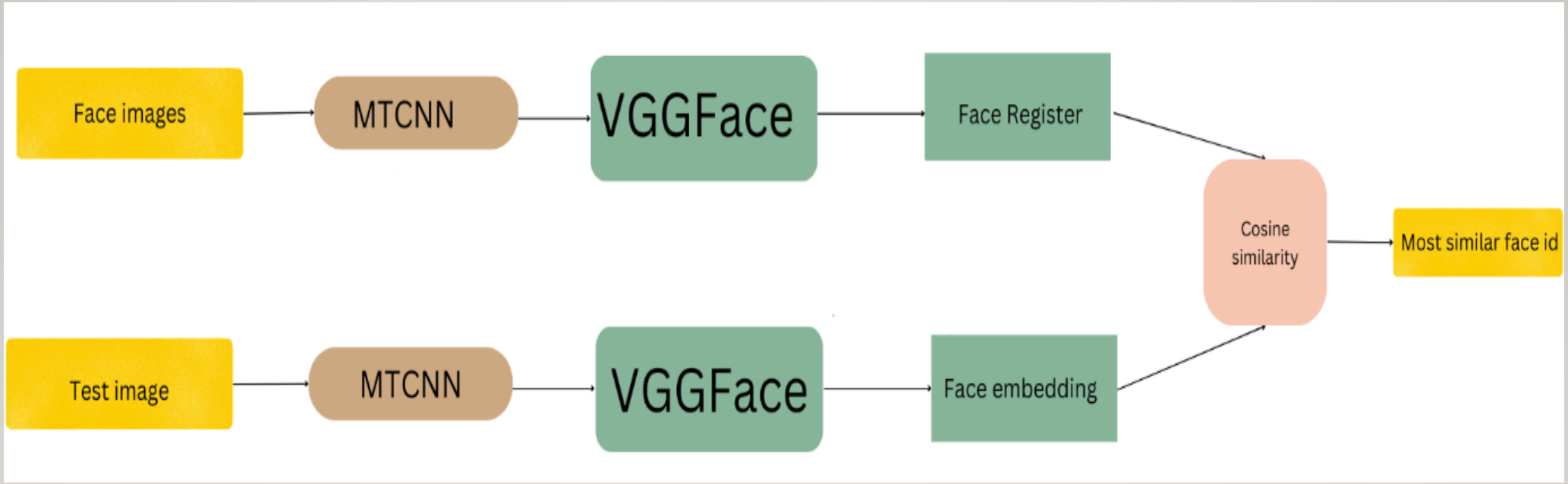
Register1

- Used 200 frontal faces for creating register, which are passed through MTCNN to extract only the faces.
- The extracted faces were then passed through the pretrained VGGFace model to obtain a 512-dimensional face embedding, which was stored in a dictionary, where the key is the faceID and the value is the embedding.

Register2

- Used 200 frontal faces , central occlusion faces, left occlusion faces and right occlusion faces each (in total 4 images for an individual)
- These faces passed through the pretrained VGGFace model to obtain **four 512 dimensional face embeddings** for each individuals.
- Then, the mean of the embeddings for an individual was taken, which was stored in a dictionary where the key is the faceID and the value is the average of embeddings.

FACE RECOGNITION MODEL



- Extracted faces from the image using MTCNN.
- Created face register(Register1 and Register2).
- Test images - images with orientation and different occlusions.
- Embeddings of test images is compared with the face register, face ID of most similar face from register is assigned.

RESULTS –PRETRAINED MODEL WITH REGISTER I

Type of data	Accuracy	Type of data	Accuracy
Orientation	94.5		
5% central occlusion	95.0	15% central occlusion	78.5
5% right occlusion	98.5	15% right occlusion	80.5
5% left occlusion	99.0	15% left occlusion	91.0

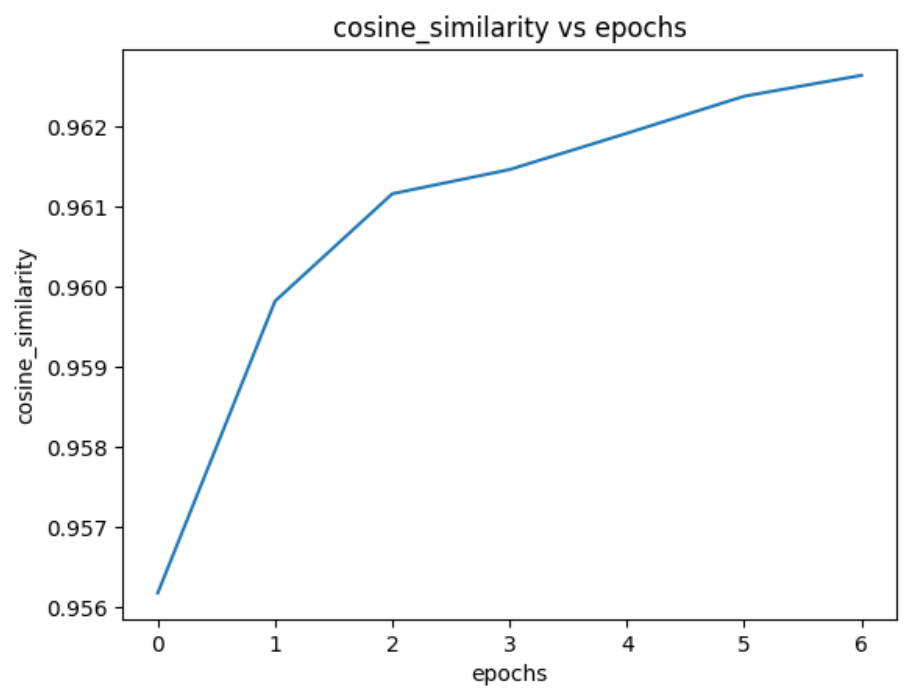
COMPARISON - PRETRAINED MODEL

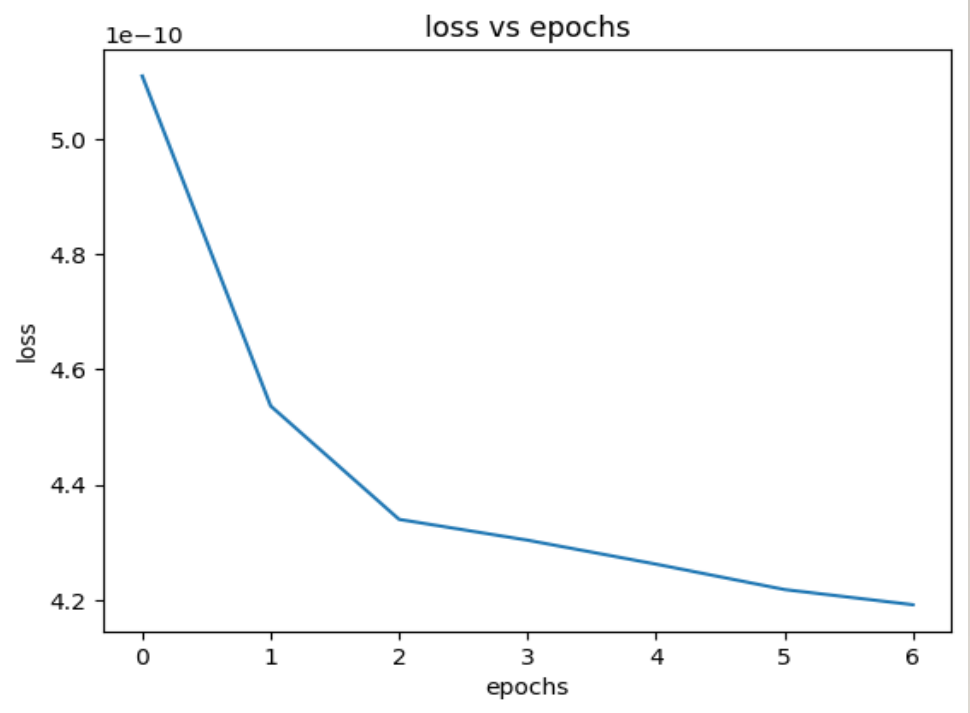
- We have used the pretrained VGGFace model for the face recognition.The results are as follows.

Type of data	Register1	Register2
	Accuracy	
Orientation	94.5	94.3
15% central occlusion	78.5	97.5
15% right occlusion	80.5	98.5
15% left occlusion	91.0	99.0
15% central occlusion+orientation	59.5	74.0
15% right occlusion+orientation	71.0	78.0
15% left occlusion+orientation	63.5	74.0

COMPARISON - TRAINED MODEL

- We have trained one convolution layer of the VGGFace with 400 images from the dataset, then the predictions are done.



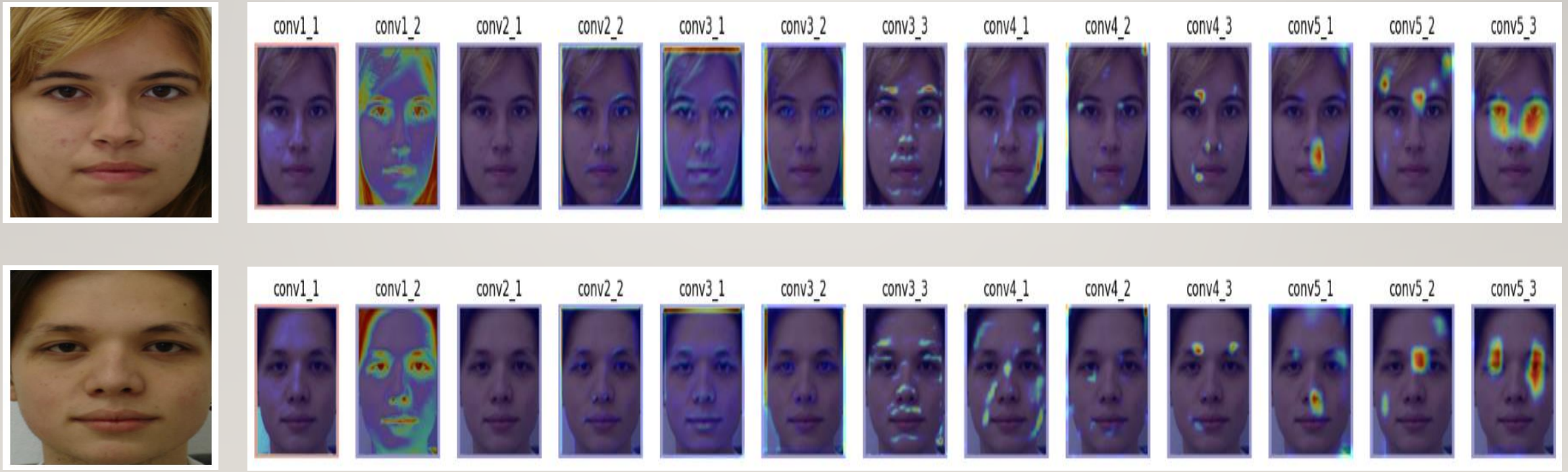


- The results are as follows.

Type of data	Register1	Register2
	Accuracy	
Orientation	92.0	93.0
15% central occlusion	85.5	98.5
15% right occlusion	85.0	100
15% left occlusion	92.0	99.5
15% central occlusion+orientation	56.0	72.5
15% right occlusion+orientation	69.0	77.0
15% left occlusion+orientation	58.5	70.5

HEATMAPS

- We can see that the deeper Convolution layers are focusing more on important features of the faces such as eyes , nose etc..



OBSERVATION

- Found that utilizing Register2, which combines both regular and occluded faces, consistently resulted in better accuracy, highlighting the importance of incorporating diverse data in training.

CHALLENGES

- **Dependency issues :** Installing VGGFace encountered various Library Compatibility, Version Conflicts etc..
- **GPU Limitations**: We faced challenges in training models with large images due to limitations in GPU capacity, affecting processing speed and scalability.
- **Quality vs. Size**: Reduction of image size to address GPU constraints led to a trade-off between maintaining quality and preserving crucial facial details necessary for accurate recognition.

CONCLUSION

- MTCNN works best for face extraction process.
- VGGFace is able to identify crucial facial characteristics, enhances its accuracy in face recognition tasks with occlusions.
- Achieved notable accuracy improvements, particularly with occluded and varied orientation images when considering Register2.
- Due to higher training time,we could only train the model for 7 epochs,training for more epochs could have resulted in much better results.

THANK YOU