MACHINE LEARNING PROJECT

EC20B1042 - Srinidhi.B
EC20B1065 - Bhavana.V
EC20B1099 - Bhavagna.K
EC20B1118 - Varshitha.Y

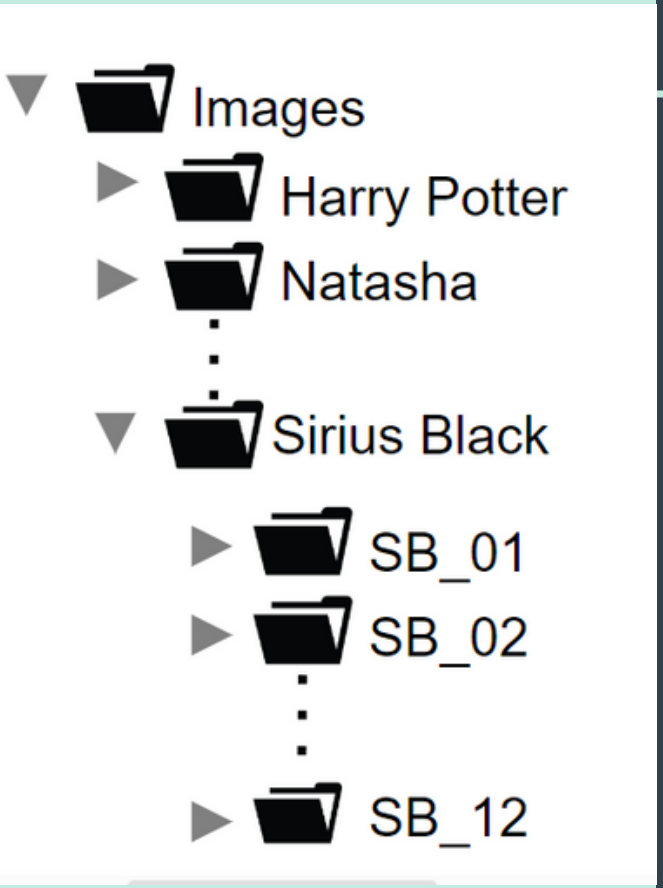
PART A - FACE RECOGNITION

- Dataset Preprocessing
- Process
  - Face detection
  - PCA and classification
- Results
- Our observations
- References

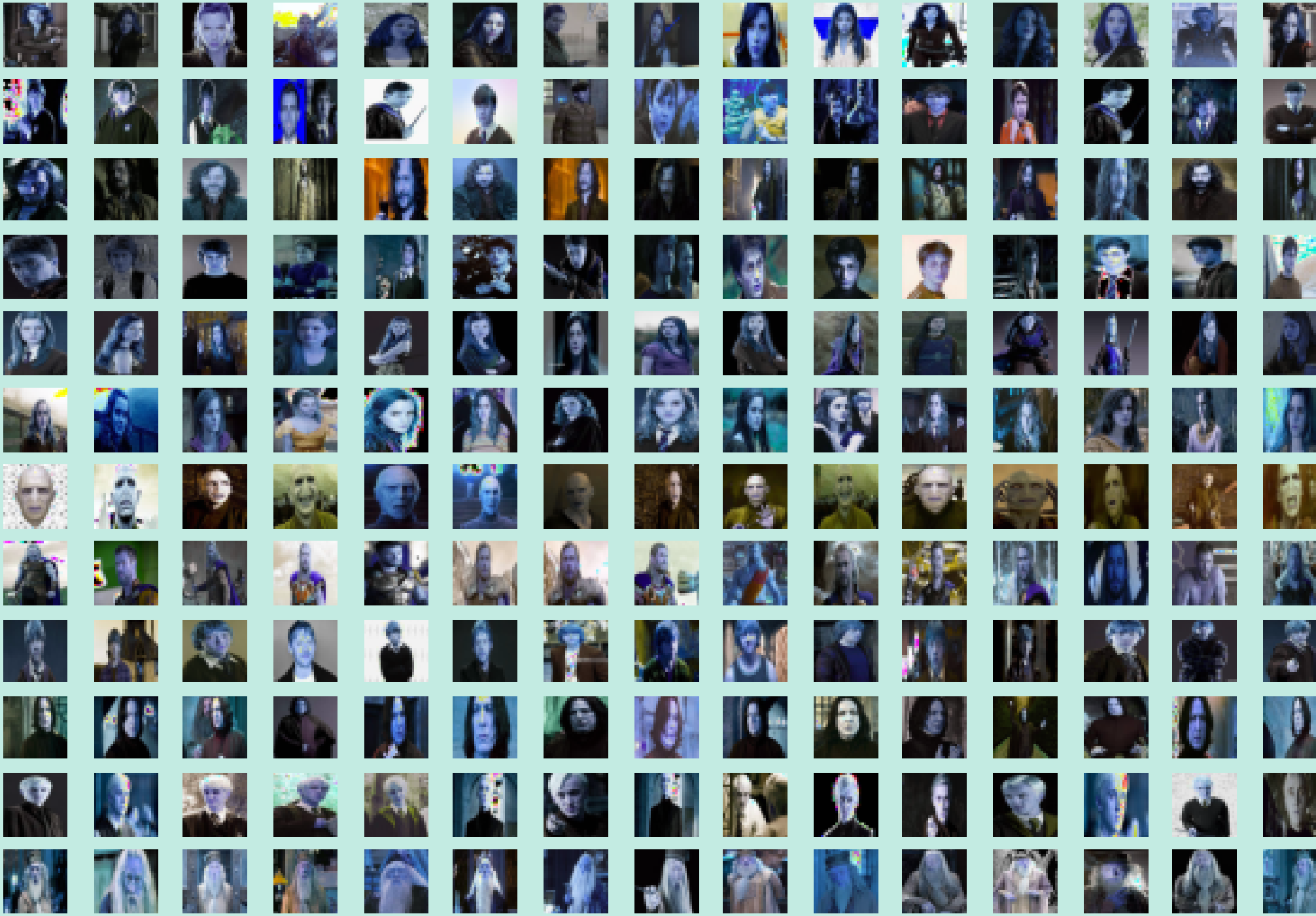
PART B - ALZHEIMER'S PREDICTION

- Problem statement
- Dataset
- Process
  - PCA
  - Classifications
- Results
- Our observations

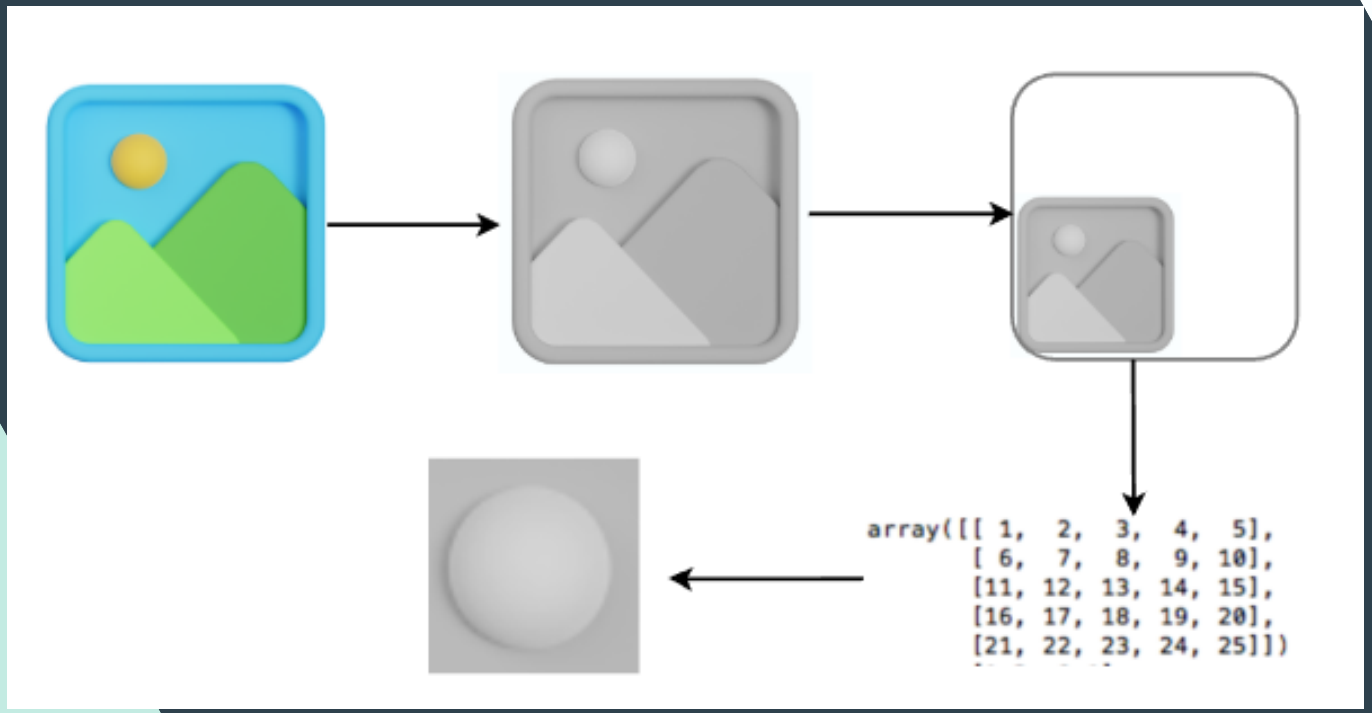
DATASET



- Images: Harry Potter & Avengers Characters
- 12 classes of 15 images each
- Google Images (https://images.google.com/)
- Scrapped manually



DATA PREPROCESSING



PROCESS - FACE DETECTION

Haar Cascade

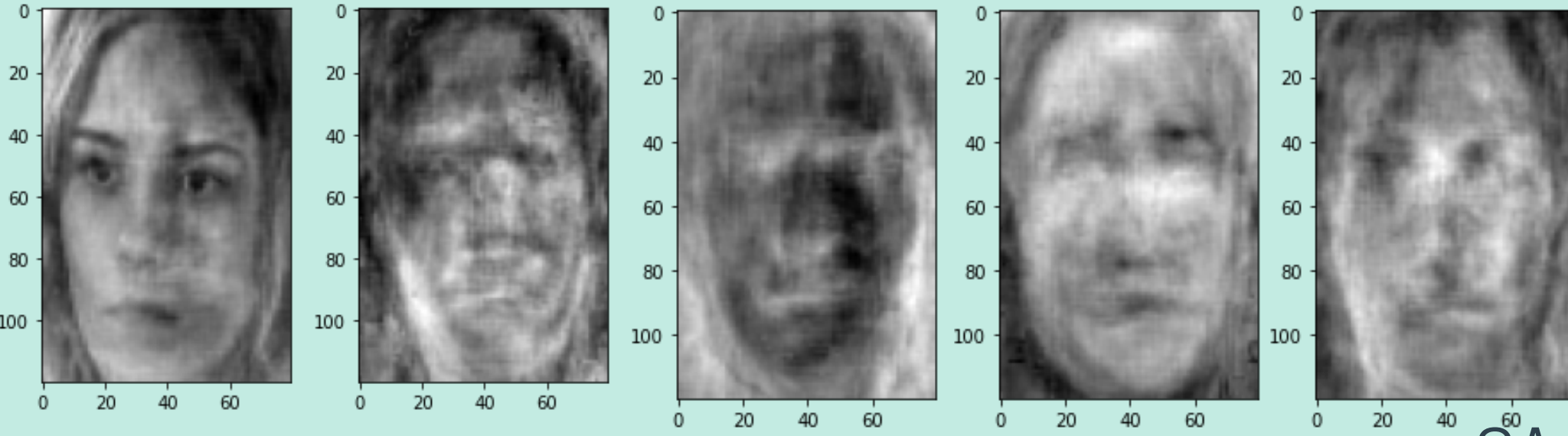
FaceNet

FACE DETECTION

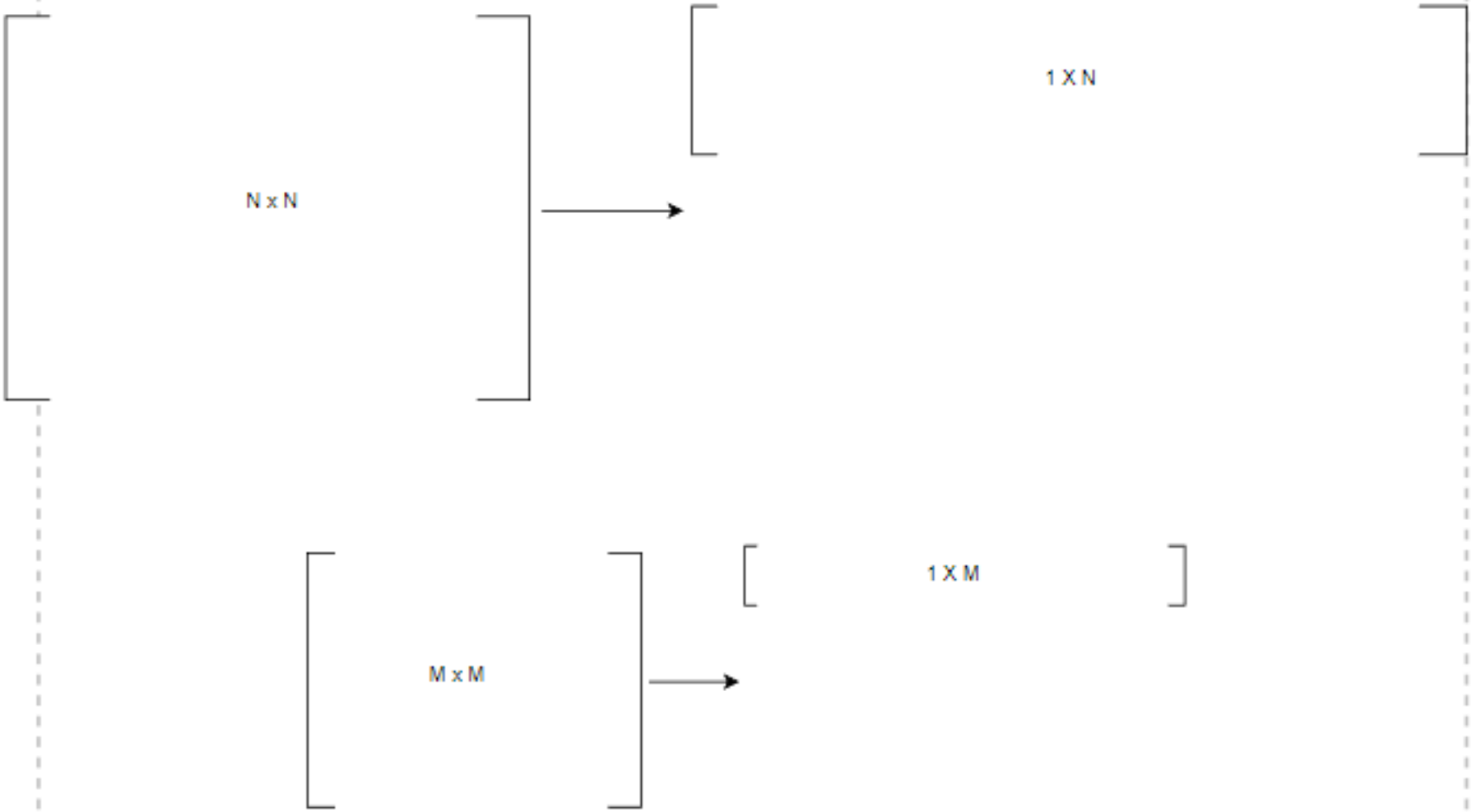
- Trained on ~100 manually labeled images
- Transfer learning on a pre-trained Model, built on a Faster R-CNN ResNet-50 FPN
- Detecto - Python Library

Process - PCA and Classification

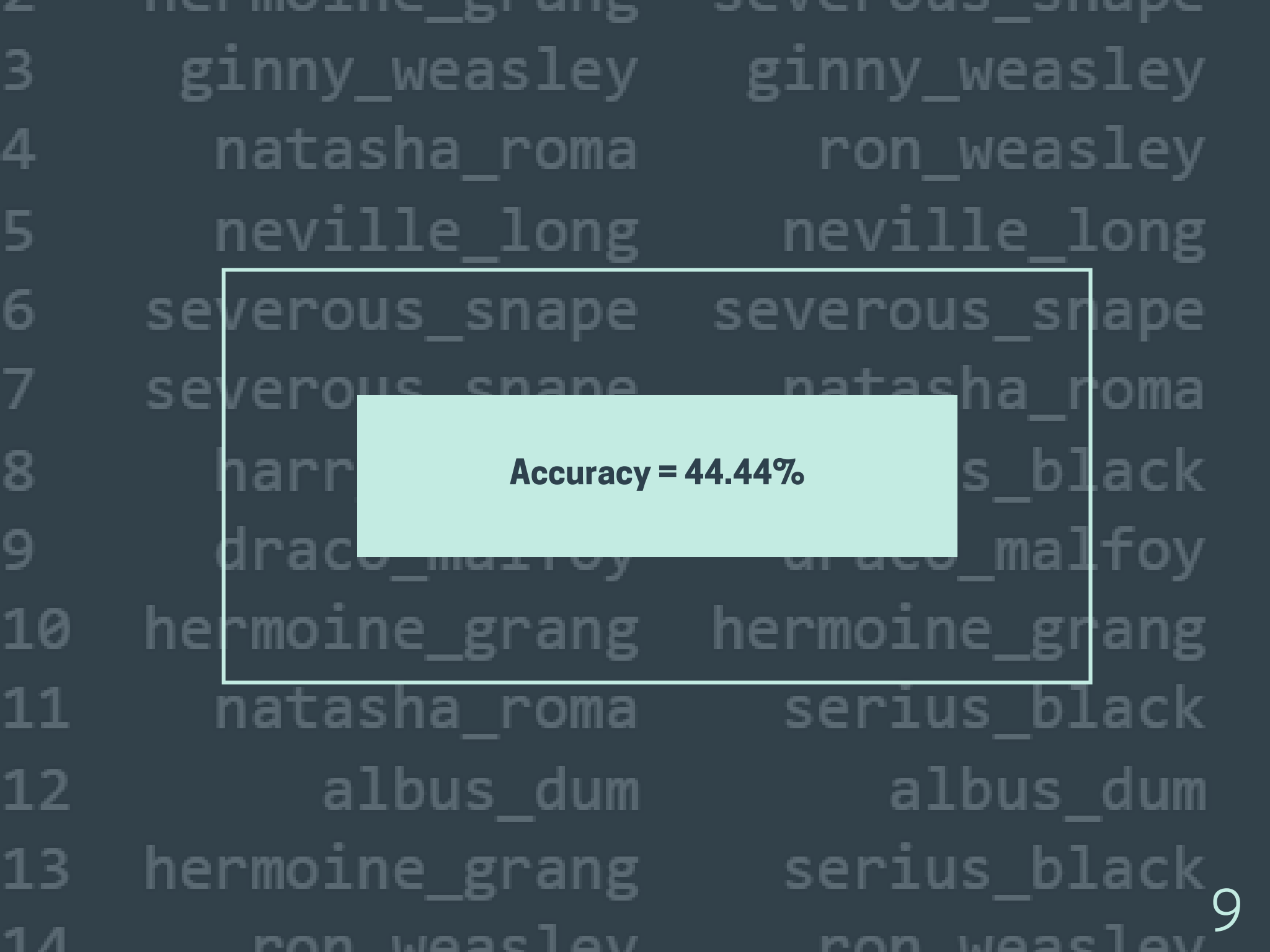
	w1	w2	w3	w4	w5
0	1.201372e+07	4.088297e+06	-7.934309e+06	3.305407e+07	2.523762e+06
1	2.842283e+06	-5.873095e+07	-3.798432e+07	-2.624902e+07	-5.985001e+07
2	-2.265907e+06	3.581680e+07	2.498598e+07	2.457728e+07	3.727135e+07
3	-4.229492e+06	-1.062019e+07	-3.838820e+06	1.592241e+06	-3.549609e+06
4	3.242093e+06	-3.985542e+07	-3.037811e+07	-7.261221e+06	-4.110285e+07
..	...	...	...	...	...
172	-5.093243e+06	-8.566256e+06	-1.590396e+06	-4.897470e+06	-4.668098e+06
173	-1.468696e+06	1.835450e+07	1.481507e+07	1.340280e+07	2.438023e+07
174	-9.032990e+04	1.323267e+06	-2.392218e+06	1.551071e+07	3.171846e+06
175	-7.159067e+06	1.577518e+07	1.281660e+07	8.772157e+06	2.104748e+07
176	-5.013973e+06	2.788356e+07	1.877450e+07	7.664817e+06	2.965710e+07



Some experiments



$$(1xM) \times (MxN) = (1xN)$$


Accuracy = 44.44%

PART B

ALZHEIMER'S DETECTION

PROBLEM STATEMENT

Alzheimer's disease is the most common form of dementia (60-70% of cases)

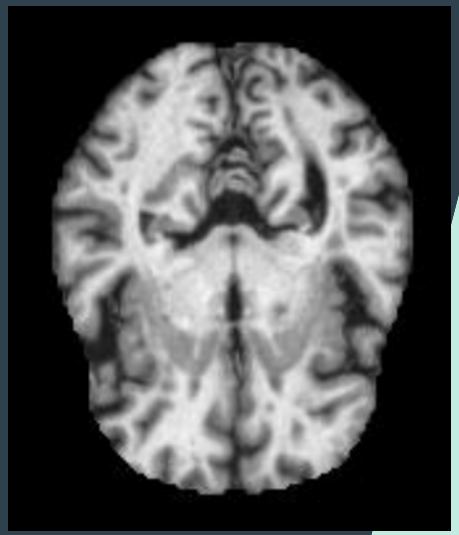
Predicting the severity of dementia using whole brain MRI scan images

DATASET:

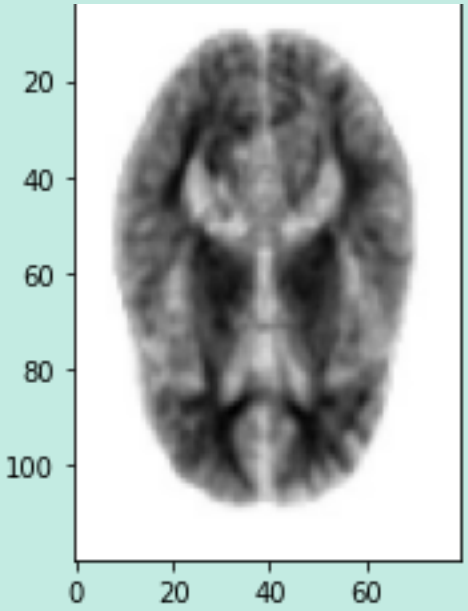
KAGGLE'S ALZHEIMER'S DATASET: (https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images)

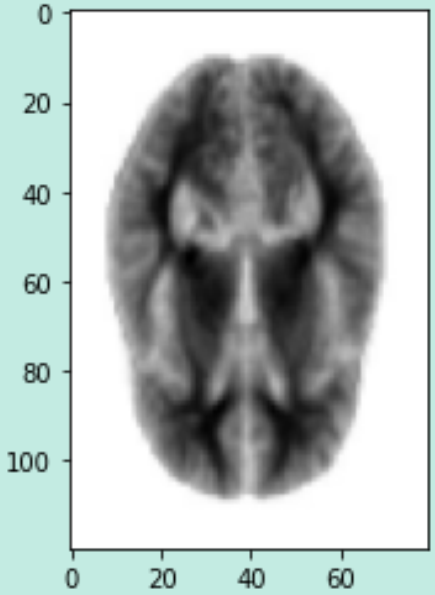
50 MRI scan images each for 4 Stages of Dementia:

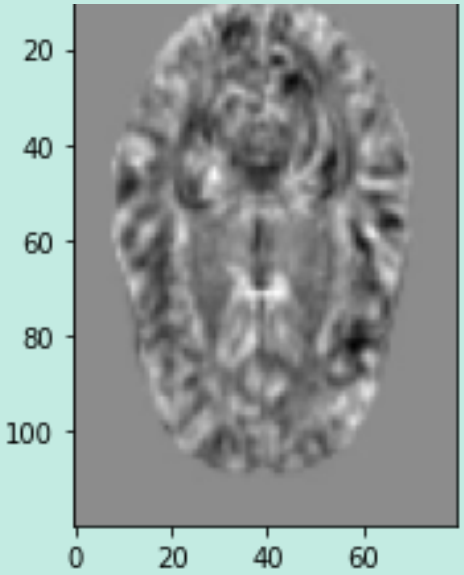
- Moderate Demented
- Mild Demented
- Very Mild Demented
- Non Demented

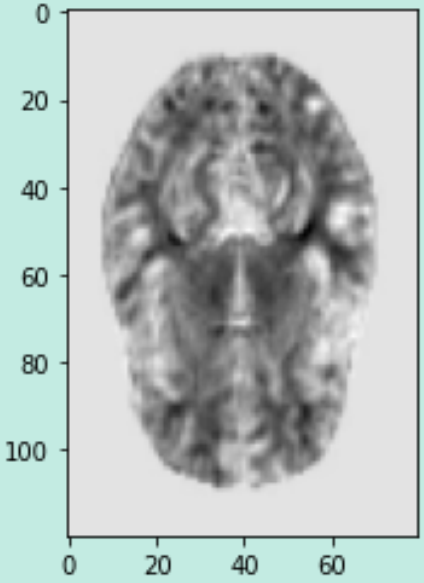


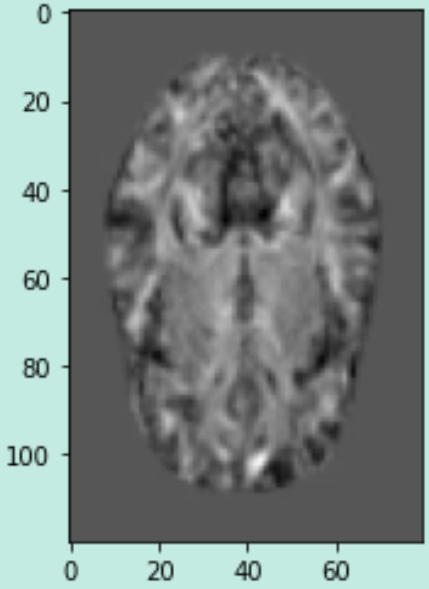
PROCESS - PCA

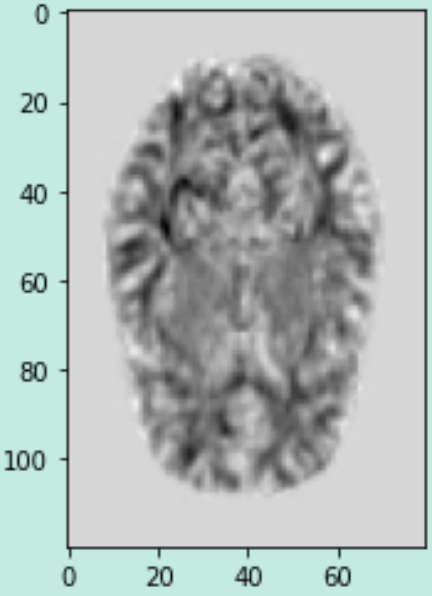

0
20
40
60
80
100
0
20
40
60

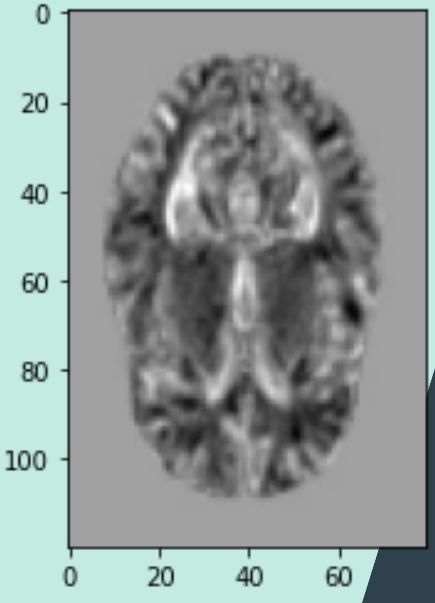

0
20
40
60
80
100
0
20
40
60


20
40
60
80
100
0
20
40
60


0
20
40
60
80
100
0
20
40
60

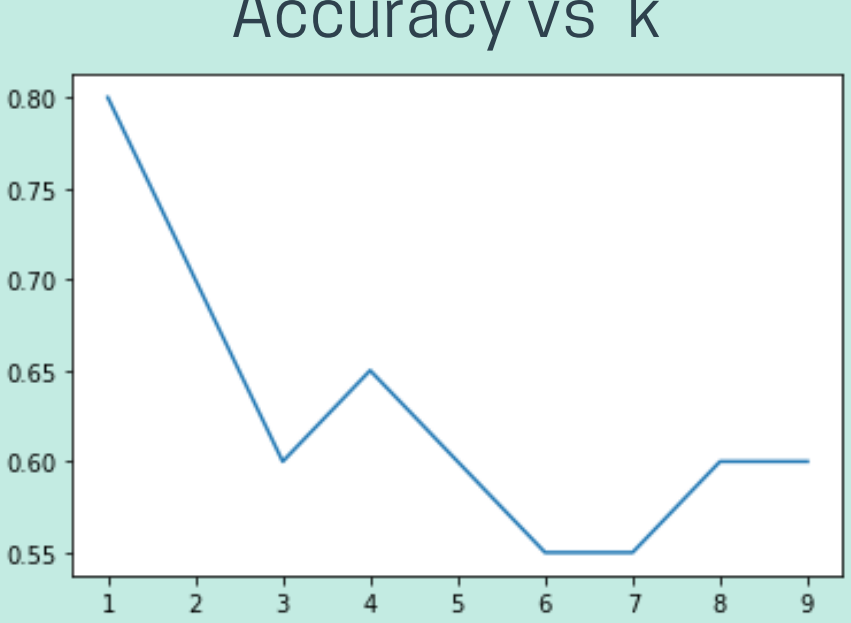

0
20
40
60
80
100
0
20
40
60


0
20
40
60
80
100
0
20
40
60


0
20
40
60
80
100
0
20
40
60

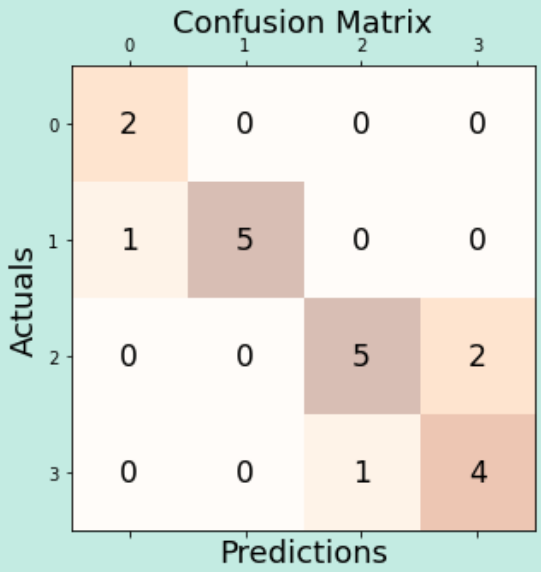
PROCESS - CLASSIFICATION

	Predicted	Label
0	MildDemented	ModerateDemented
1	VeryMildDemented	VeryMildDemented
2	ModerateDemented	ModerateDemented
3	ModerateDemented	ModerateDemented
4	ModerateDemented	ModerateDemented
5	VeryMildDemented	NonDemented
6	VeryMildDemented	VeryMildDemented
7	NonDemented	VeryMildDemented
8	NonDemented	NonDemented
9	ModerateDemented	ModerateDemented
10	NonDemented	NonDemented
11	MildDemented	MildDemented
12	VeryMildDemented	VeryMildDemented
13	ModerateDemented	ModerateDemented
14	VeryMildDemented	VeryMildDemented
15	VeryMildDemented	NonDemented
16	MildDemented	MildDemented
17	NonDemented	NonDemented
18	NonDemented	NonDemented
19	NonDemented	NonDemented



RESULTS

```
[17] print("Accuracy: ", knn.score(x_test, y_test))

     Accuracy:  0.8

[14] print('Precision: ', precision_score(y_test, df_test['Predicted'], average='macro'))

     Precision:  0.7916666666666666

[15] print('Recall: ', recall_score(y_test, df_test['Predicted'], average='macro'))

     Recall:  0.836904761904762
```



Our Observations

Custom Dataset - Character images

- Accuracy was low
- Extensive preprocessing required
- No uniformity
- Lesser number of images and more number of classes

Open Source Dataset Kaggle Alzheimer's

- Better accuracy
- No extensive preprocessing required
- High Uniformity
- Lesser number of classes and more number of images

References

- Sarvesh Debey, Kaggle's Alzheimer's Dataset, Kaggle
- Alan Bi, "Detecto — An object detection library for PyTorch", Medium
- Face Detection using Haar Cascades, OpenCV tutorials
- Yash, "Face Recognition with FaceNet", Tech Musings - Medium
- Face Recognition Using Eigenfaces (PCA Algorithm), GeeksforGeeks