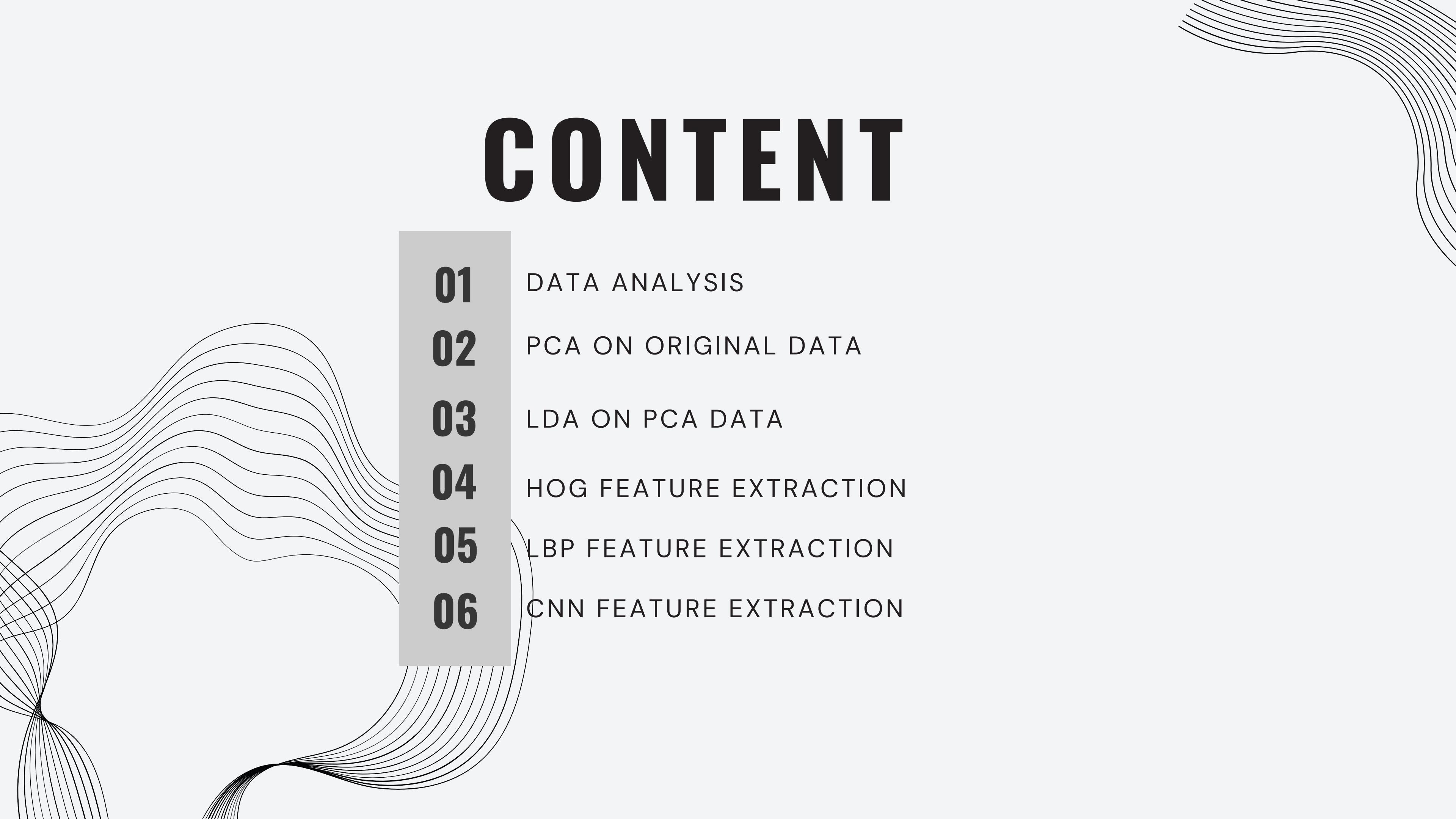


FACE IDENTIFICATION

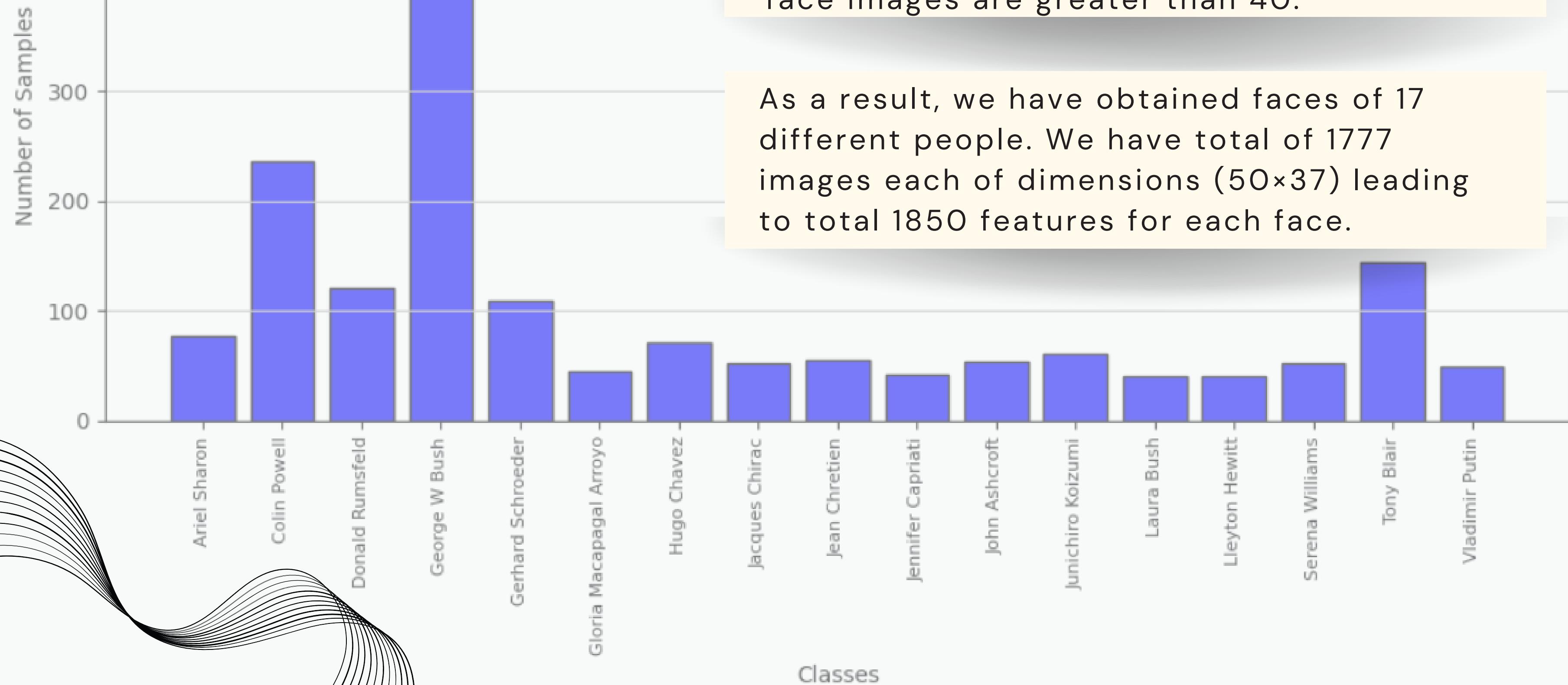
GROUP ID 10

CONTENT

- 
- 01** DATA ANALYSIS
 - 02** PCA ON ORIGINAL DATA
 - 03** LDA ON PCA DATA
 - 04** HOG FEATURE EXTRACTION
 - 05** LBP FEATURE EXTRACTION
 - 06** CNN FEATURE EXTRACTION

Distribution of Classes

DATA ANALYSIS

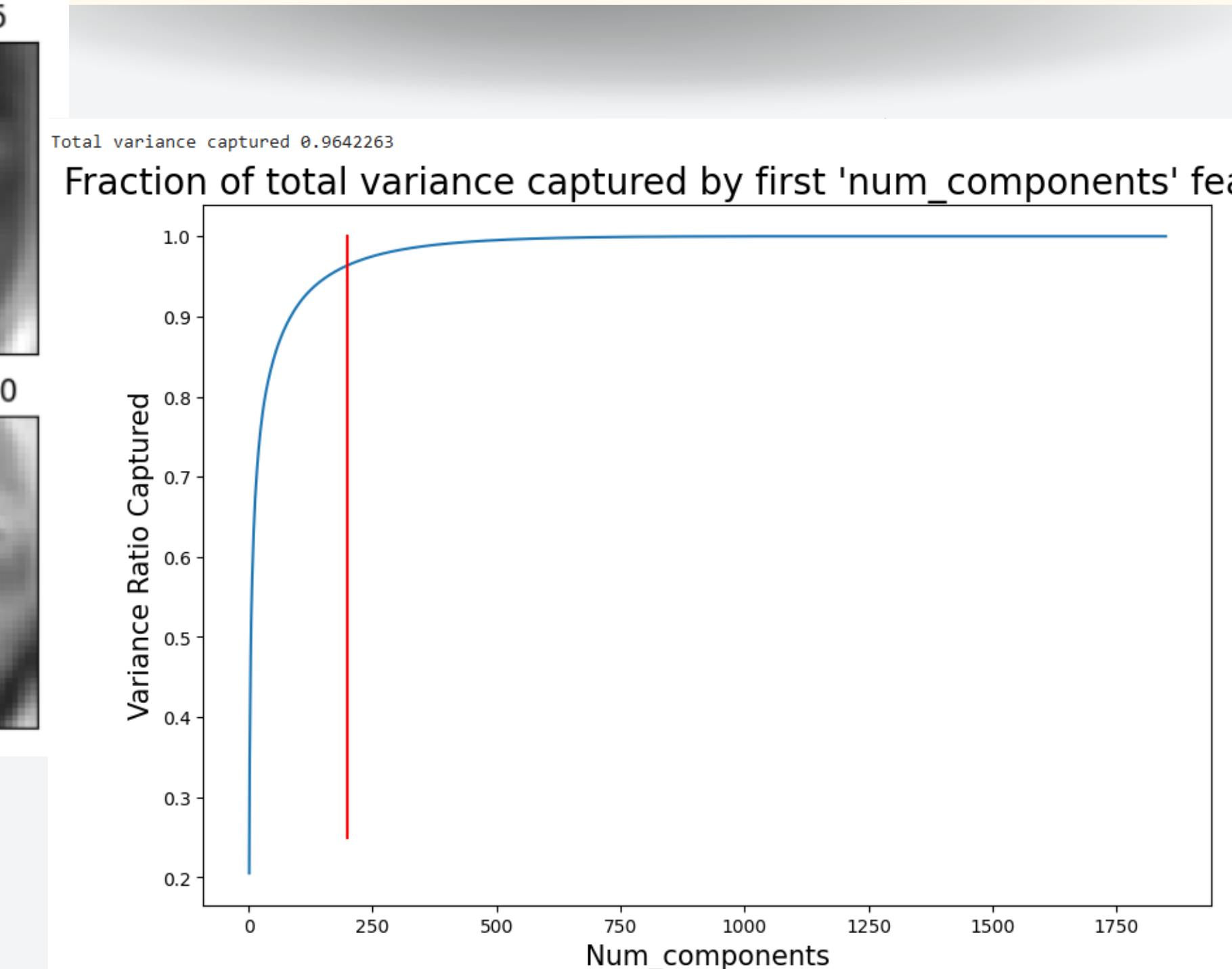
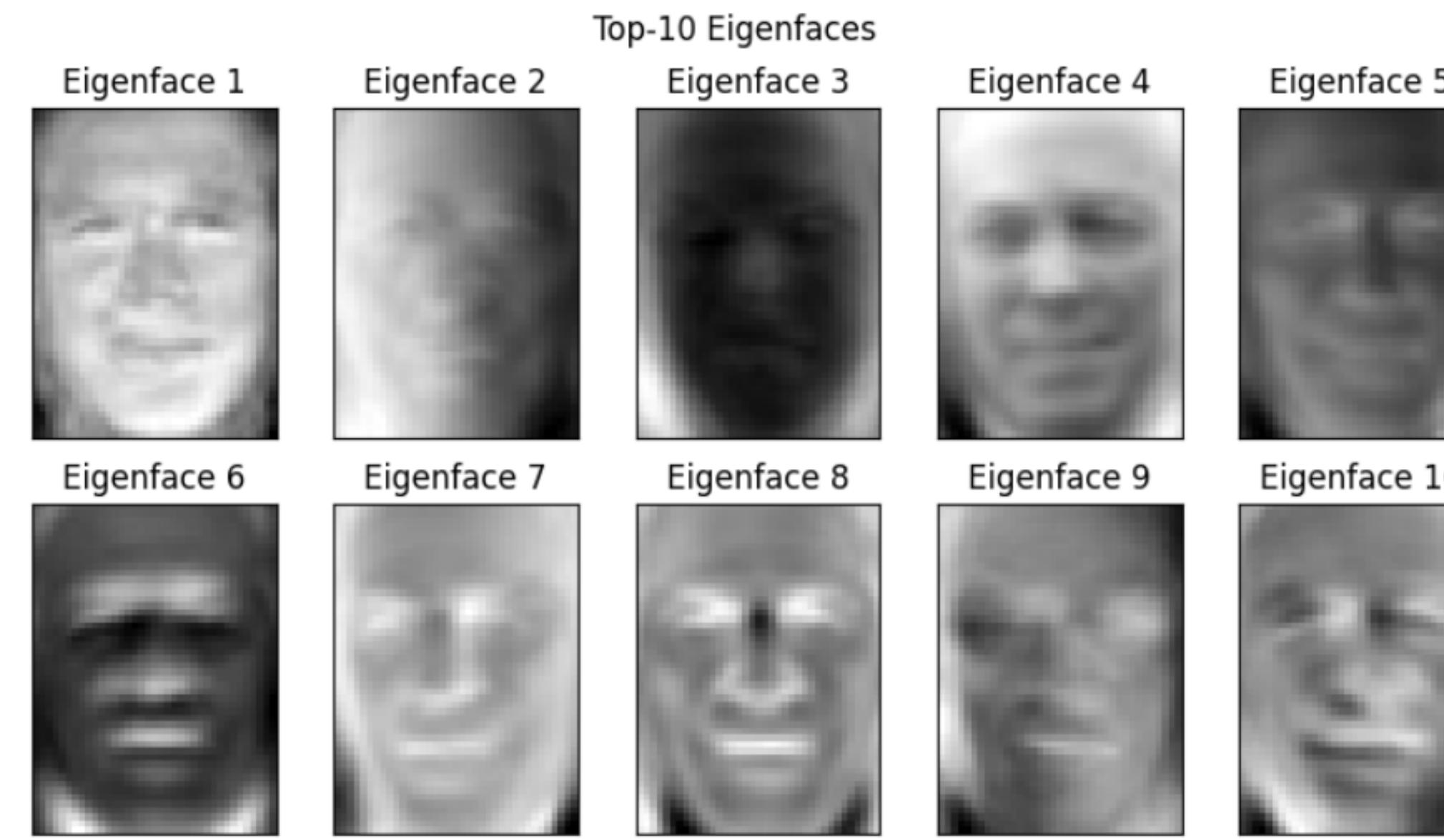


We have fetched the LFW dataset using the sklearn fetch_lfw_people() function. We have selected only those person whose sample face images are greater than 40.

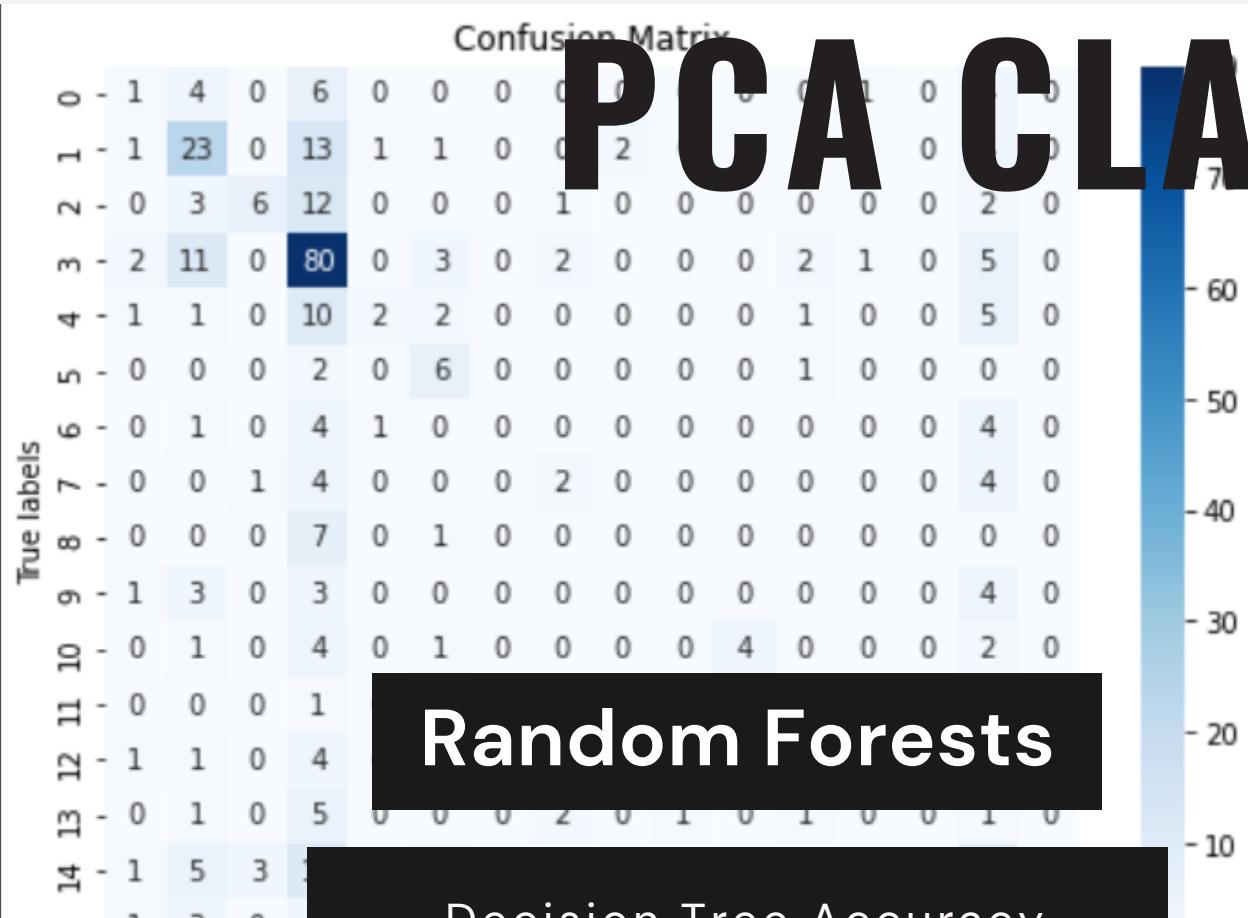
As a result, we have obtained faces of 17 different people. We have total of 1777 images each of dimensions (50×37) leading to total 1850 features for each face.

PRINCIPAL COMPONENT ANALYSIS

Applying principal component analysis, we have reduced 1850 features of the training data to 200 features still capturing 96.4% of the total variance.

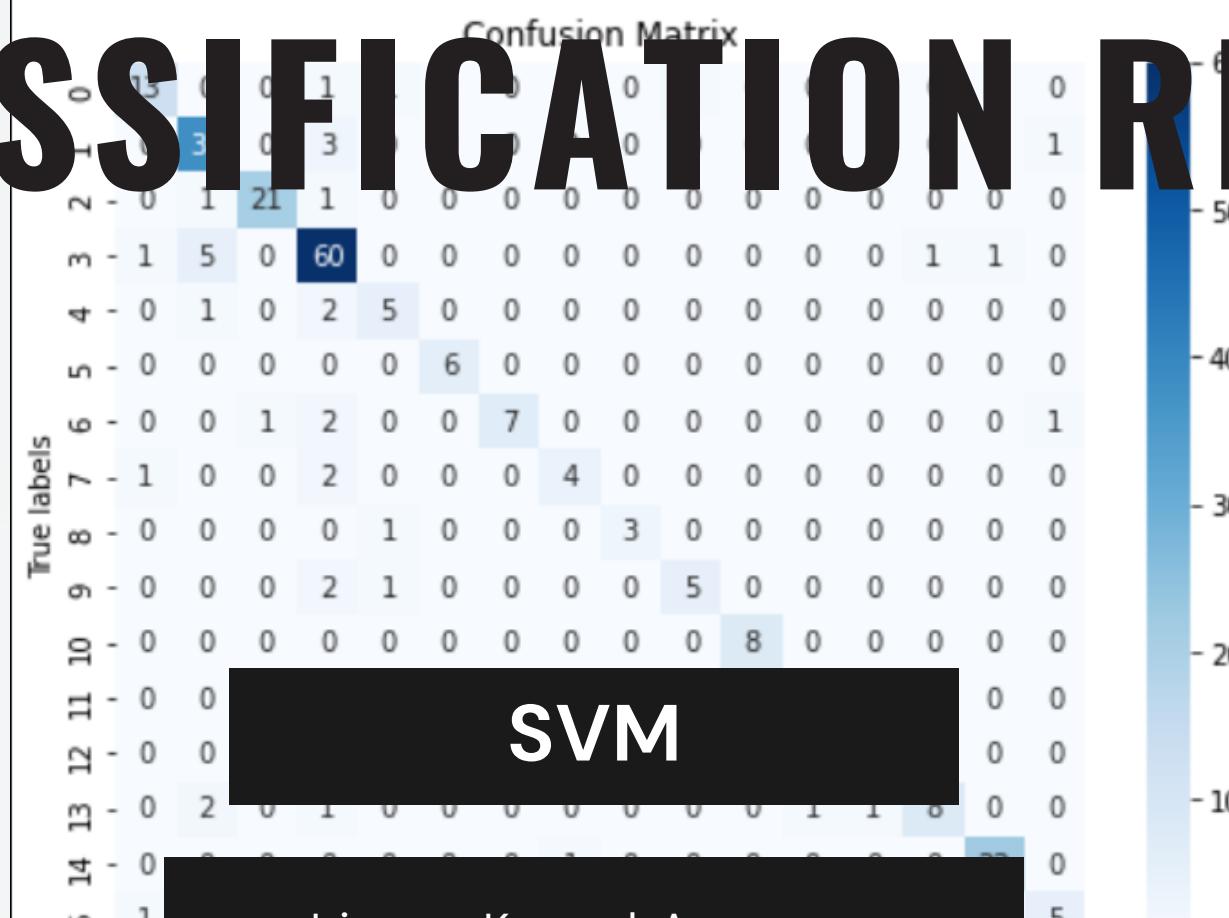
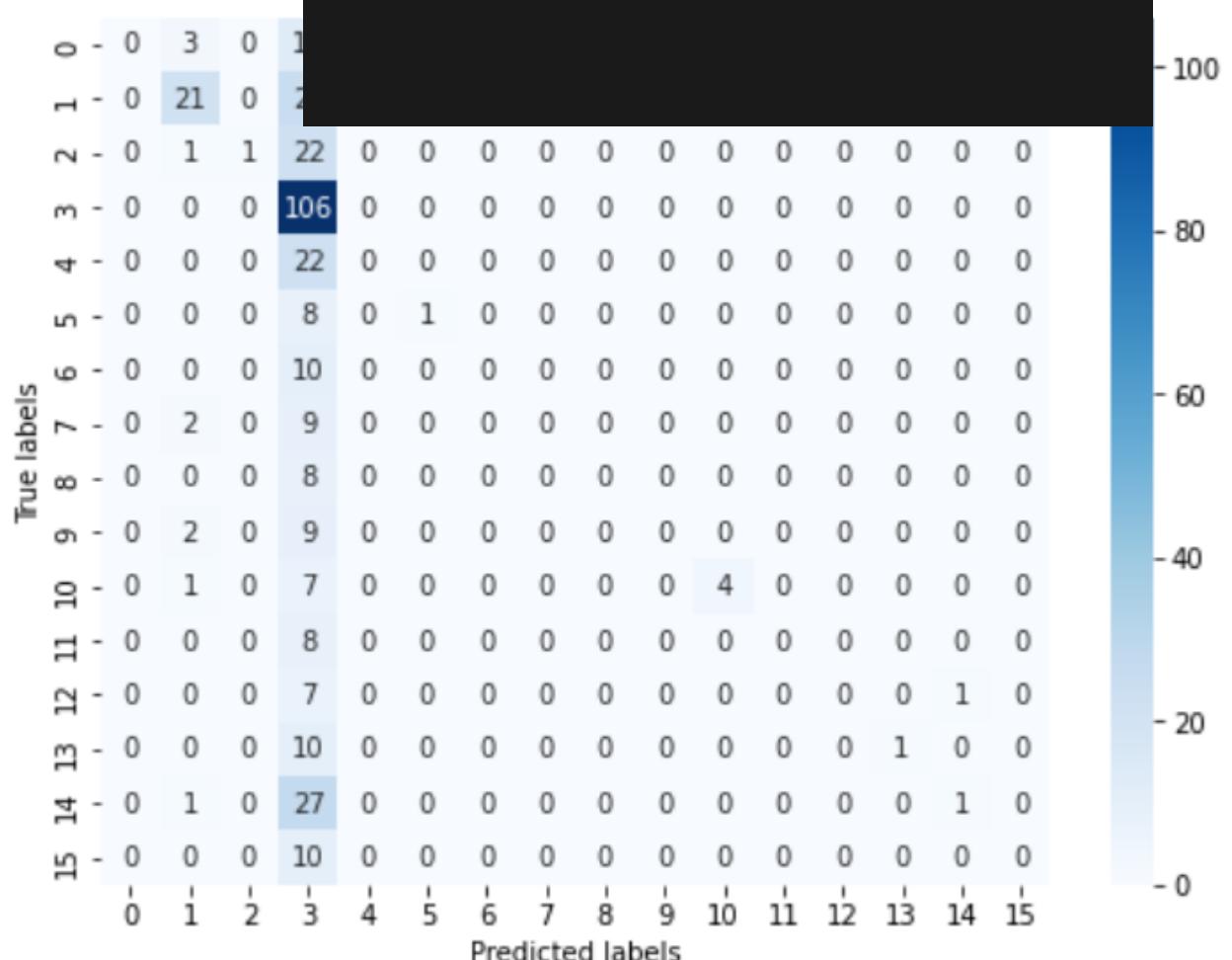


PCA CLASSIFICATION RESULTS



Decision Tree Accuracy
= 40%

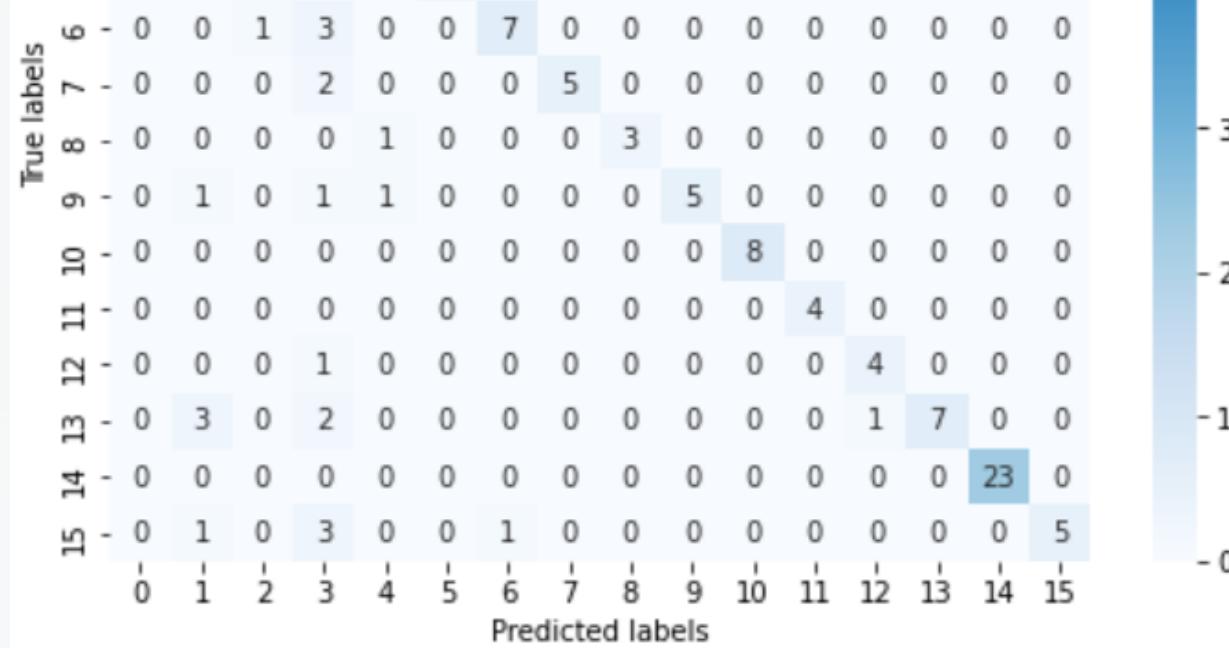
Random Forests Accuracy
= 40%



Linear Kernel Accuracy
= 83.2%

Polynomial Kernel Accuracy
= 65.2%

RBF Kernel Accuracy
= 84.7%



PCA+LDA CLASSIFICATION RESULTS

Confusion Matrix																
True labels	0	1	3	0	2	2	0	0	0	3	1	1	2	0	0	1
	0	12	1	1	4	0	1	0	0	2	0	1	1	0	0	0
	1	4	12	1	1	4	0	1	0	0	2	0	1	1	0	0
	2	2	2	0	5	0	0	4	0	0	3	2	0	0	1	0
	3	8	2	4	45	8	0	9	3	4	1	0	6	0	0	9
	4	1	1	1	1	1	7	0	4	3	0	1	3	0	0	0
	5	0	0	0	1	0	0	0	3	4	0	0	0	1	0	0
	6	1	1	1	1	1	0	0	2	0	1	0	0	0	1	2
	7	0	1	1	1	1	0	1	4	0	0	0	1	0	0	1
	8	1	0	0	0	1	0	0	0	2	1	0	0	1	2	0
	9	0	0	0	0	3	0	1	0	0	1	2	1	0	1	0
	10	1	0	0	3	2	0	1	0	0	0	4	1	0	0	0
	11	0	1	1	0	0	0	0	4	1	0	0	0	0	0	0
	12	0	1	0	2	0	0	0	0	4	1	0	0	0	0	0
	13	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	15	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	

Random Forests

Decision Tree Accuracy
= 58%

Random Forests Accuracy
= 79%

Confusion Matrix																
True labels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	0	4	1	1	18	2	1	3	7	64	6	1	1	2	0	3
	1	5	18	2	0	0	0	0	4	2	0	0	2	0	0	0
	2	1	3	7	0	0	0	0	1	0	0	0	0	0	0	0
	3	5	3	4	64	6	1	1	1	2	0	3	1	2	5	5
	4	0	3	1	2	8	0	0	0	4	2	0	0	2	0	0
	5	0	0	1	3	0	4	0	0	0	1	0	0	0	0	0
	6	1	1	1	1	1	1	2	0	0	0	1	0	1	0	0
	7	0	1	1	1	1	2	0	0	1	0	0	0	2	1	0
	8	0	0	1	1	1	0	1	0	3	0	0	0	0	1	0
	9	0	0	0	1	1	0	1	0	5	1	0	0	1	0	1
	10	1	1	0	3	0	0	0	0	1	4	0	0	1	1	0
	11	0	2	0	0	1	0	0	0	4	0	0	0	0	0	0
	12	1	2	0	3	0	0	1	0	0	0	0	0	0	0	0
	13	2	3	1	0	0	1	0	0	0	0	0	2	0	0	0
	14	1	2	2	5	0	0	1	0	2	2	0	0	1	1	1
	15	2	2	0	3	1	0	2	0	0	0	0	0	0	0	0
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	

Confusion Matrix																
True labels	0	2	1	0	3	1	0	0	2	1	0	0	0	1	0	0
	0	2	1	0	3	1	0	0	2	1	0	0	0	1	0	0
	1	0	2	1	0	3	1	0	0	2	1	0	0	0	1	0
	2	0	1	0	5	0	0	4	0	0	0	0	0	1	0	0
	3	8	2	4	45	8	0	9	3	4	1	0	6	0	0	9
	4	1	1	1	1	1	7	0	4	3	0	1	3	0	0	0
	5	0	0	0	1	0	0	0	3	4	0	0	0	1	0	0
	6	1	1	1	1	1	0	0	2	0	1	0	0	0	1	2
	7	0	1	1	1	1	0	1	4	0	0	0	1	0	0	1
	8	1	0	0	0	1	0	0	0	2	1	0	0	1	2	0
	9	0	0	0	0	3	0	1	0	0	0	0	0	1	0	0
	10	1	0	0	3	2	0	1	0	0	0	0	0	0	1	0
	11	0	1	0	1	0	0	0	0	4	0	0	0	0	0	0
	12	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
	13	2	0	0	3	0	0	1	0	0	0	0	0	0	0	0
	14	1	2	2	5	0	0	1	0	2	2	0	0	0	1	1
	15	2	2	0	3	1	0	2	0	0	0	0	0	0	0	0
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	

SVM

Linear Kernel Accuracy
= 30%
Polynomial Kernel Accuracy
= 74%
RBF Kernel Accuracy
= 76%

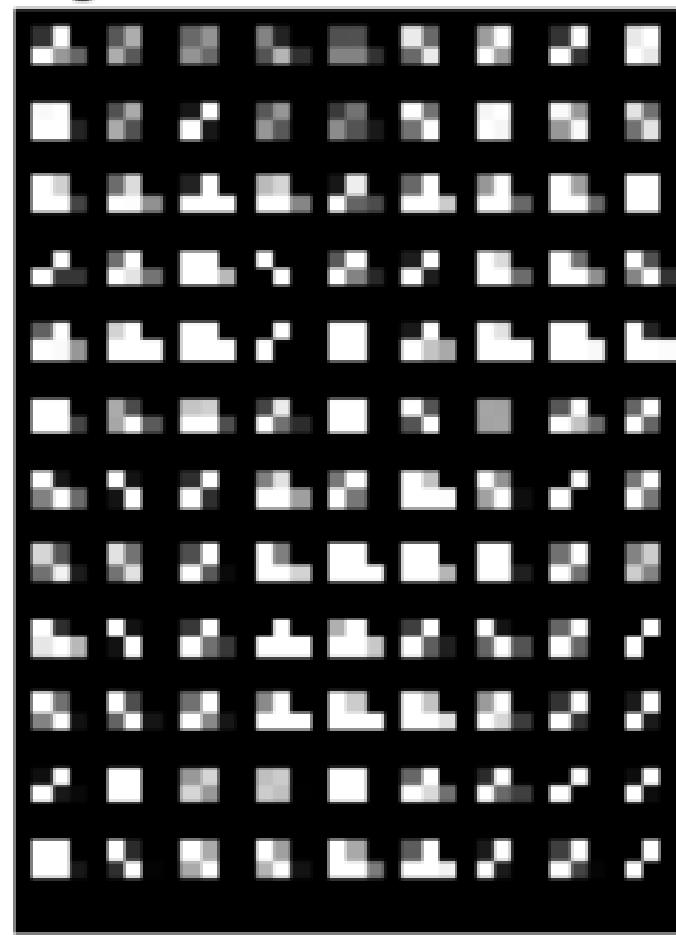
|<th colspan="1
| |

HISTOGRAM OF ORIENTED GRAPHS

Resized image



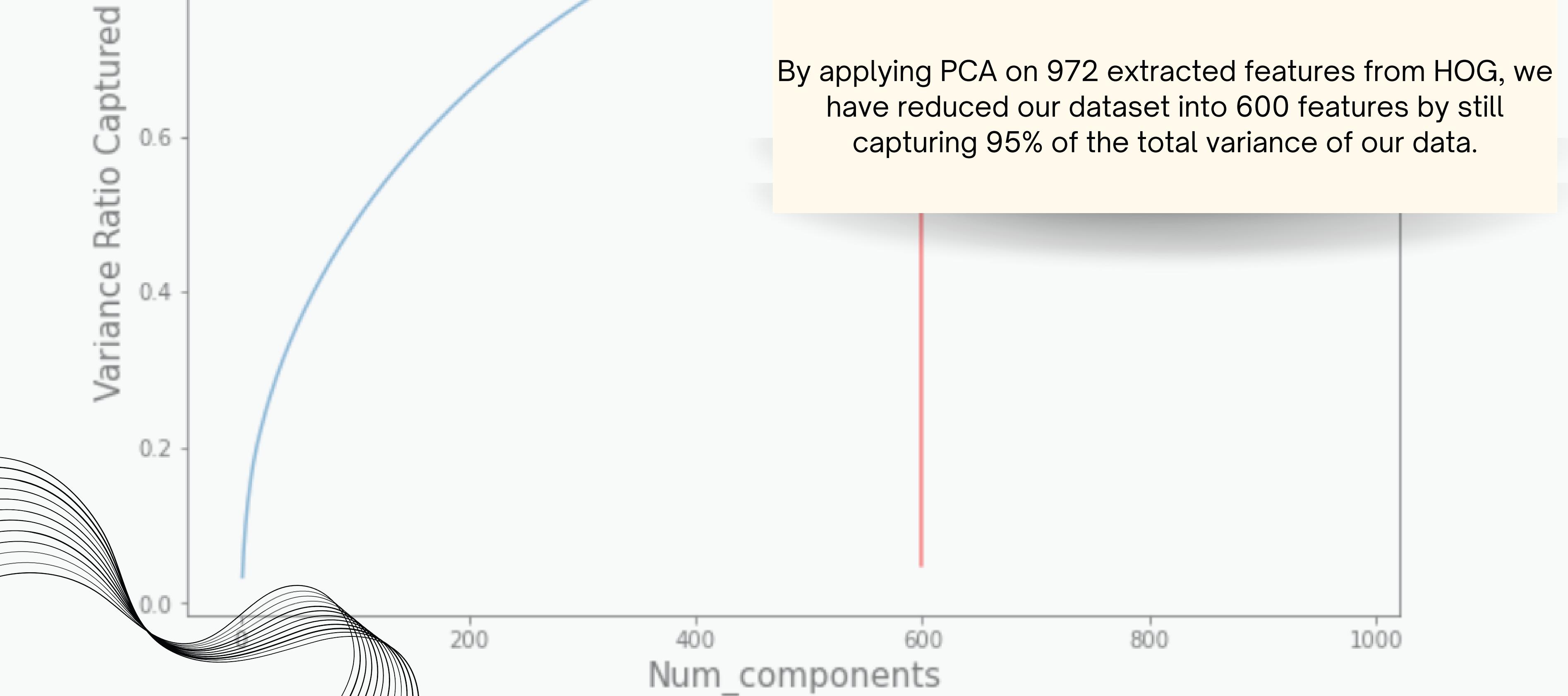
Histogram of Oriented Gradients



HOG extracts local gradient orientation information from an image by dividing it into cells, computing histograms of gradient orientations within each cell, and then normalizing these histograms.

Here each image is divided into cells of (4×4) pixels each and computed 9 HOG orientations for each cell, leading to total of 972 features. Then we have applied techniques like PCA and LDA to further reduce the extracted features for more efficient model training.

PRINCIPAL COMPONENT ANALYSIS ON 'HOG'



HOG+PCA CLASSIFICATION RESULTS

Confusion Matrix																
True labels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	1	3	0	2	2	0	1	1	1	1	1	1	1	1	1	1
1	4	12	1	1	4	0	1	0	2	7	0	1	0	1	0	1
2	2	2	0	5	0	0	4	0	0	0	3	2	2	0	0	0
3	8	2	4	45	8	0	9	3	4	1	1	0	6	0	6	9
4	1	1	1	1	1	7	0	4	3	0	1	3	0	0	0	0
5	0	0	0	1	0	0	3	4	0	0	0	1	0	0	0	0
6	1	1	1	1	1	0	0	2	0	1	0	0	0	1	2	0
7	0	1	1	1	1	0	1	4	0	0	0	1	0	0	1	0
8	1	0	0	0	1	0	0	0	2	1	0	0	1	2	0	0
9	0	0	0	0	3	0	1	0	0	1	2	1	0	1	0	2
10	1	0	0	3	2	0	1	0	0	0	4	1	0	0	0	0
11	0	1	1	0	0	2	0	1	0	0	0	0	0	0	0	0
12	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0
13	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Random Forests

Decision Tree Accuracy
= 58%
Random Forests Accuracy
= 79%

Confusion Matrix																
True labels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	4	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1	5	18	2	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	3	7	0	0	0	0	0	0	0	0	0	0	0	0	0
3	5	3	4	64	6	1	1	1	2	0	3	1	2	5	5	3
4	0	3	1	2	8	0	0	0	0	4	2	0	0	2	0	0
5	0	0	1	3	0	4	0	0	0	0	1	0	0	0	0	0
6	1	1	1	1	1	1	2	0	0	0	0	1	0	1	0	0
7	0	1	1	1	2	0	0	2	0	0	1	0	0	0	2	1
8	0	0	1	1	1	0	1	0	3	0	0	0	0	1	0	0
9	0	0	0	1	1	0	1	0	0	5	1	0	0	1	0	1
10	1	1	0	3	0	0	0	0	0	1	4	0	0	1	1	0
11	0	2	0	0	1	0	0	0	0	4	0	0	0	0	0	0
12	1	2	0	3	0	0	1	0	0	0	0	0	0	0	0	0
13	2	2	3	1	0	0	1	0	0	0	0	0	0	2	0	0
14	1	1	2	2	5	0	0	1	0	2	2	0	0	1	1	1
15	2	2	0	3	1	0	2	0	0	0	0	0	0	0	0	0
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	0

Confusion Matrix																
True labels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	1	1	3	1	3	1	1	0	1	0	0	0	1	0	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	2	0	3	8	2	1	0	0	0	1	1	0	0	0	0	0
4	0	6	3	47	2	1	3	2	2	0	2	2	2	3	3	0
5	0	0	0	0	0	6	0	0	0	0	0	1	0	0	0	0
6	-1	0	3	1	0	0	0	0	0	0	0	0	0	1	2	0
7	0	2	1	1	1	0	0	2	0	0	1	0	0	0	0	0
8	0	0	0	2	0	1	0	2	0	0	0	0	0	0	0	0
9	0	1	0	2	0	1	0	0	0	0	4	0	0	0	0	0
10	0	1	0	0	2	0	0	0	0	6	0	0	0	0	0	0
11	0	1	0	0	0	0	0	0	0	1	3	0	0	0	0	0
12	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
13	0	1	0	0	0	0	0	0	0	0	0	0	3	0	0	0
14	0	2	2	3	0	1	0	0	2	0	0	0	0	9	0	0
15	0	2	1	1	0	0	1	0	0	0	0	0	1	0	0	0

SVM

Linear Kernel Accuracy
= 30%
Polynomial Kernel Accuracy
= 74%
RBF Kernel Accuracy
= 76%

Confusion Matrix																
True labels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

<tbl_r cells="16" ix="5" max

LOCAL BINARY PATTERNS

It characterizes the local texture patterns of an image by comparing the intensity of a pixel with its neighbors.

Original image



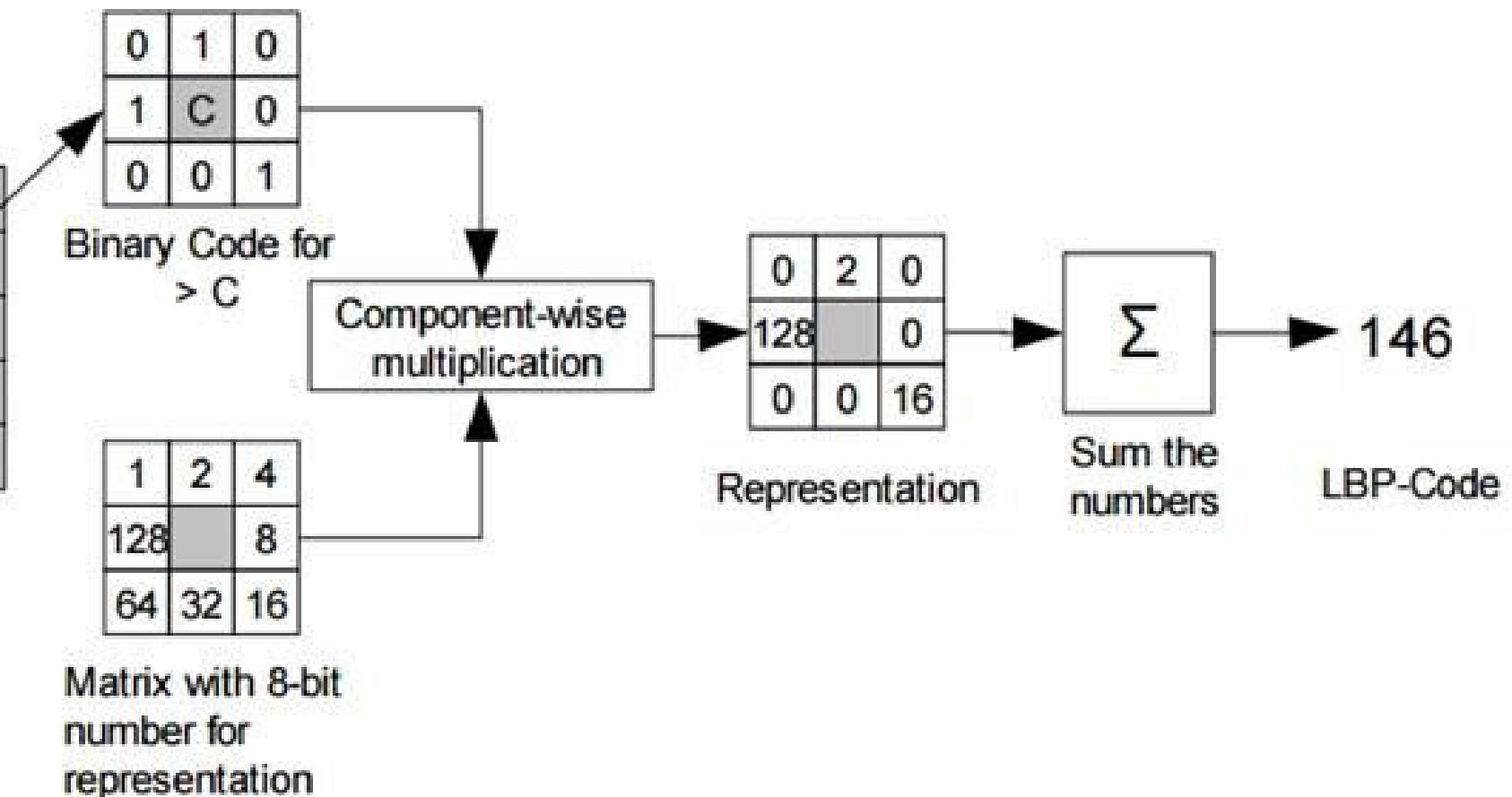
LBP Image



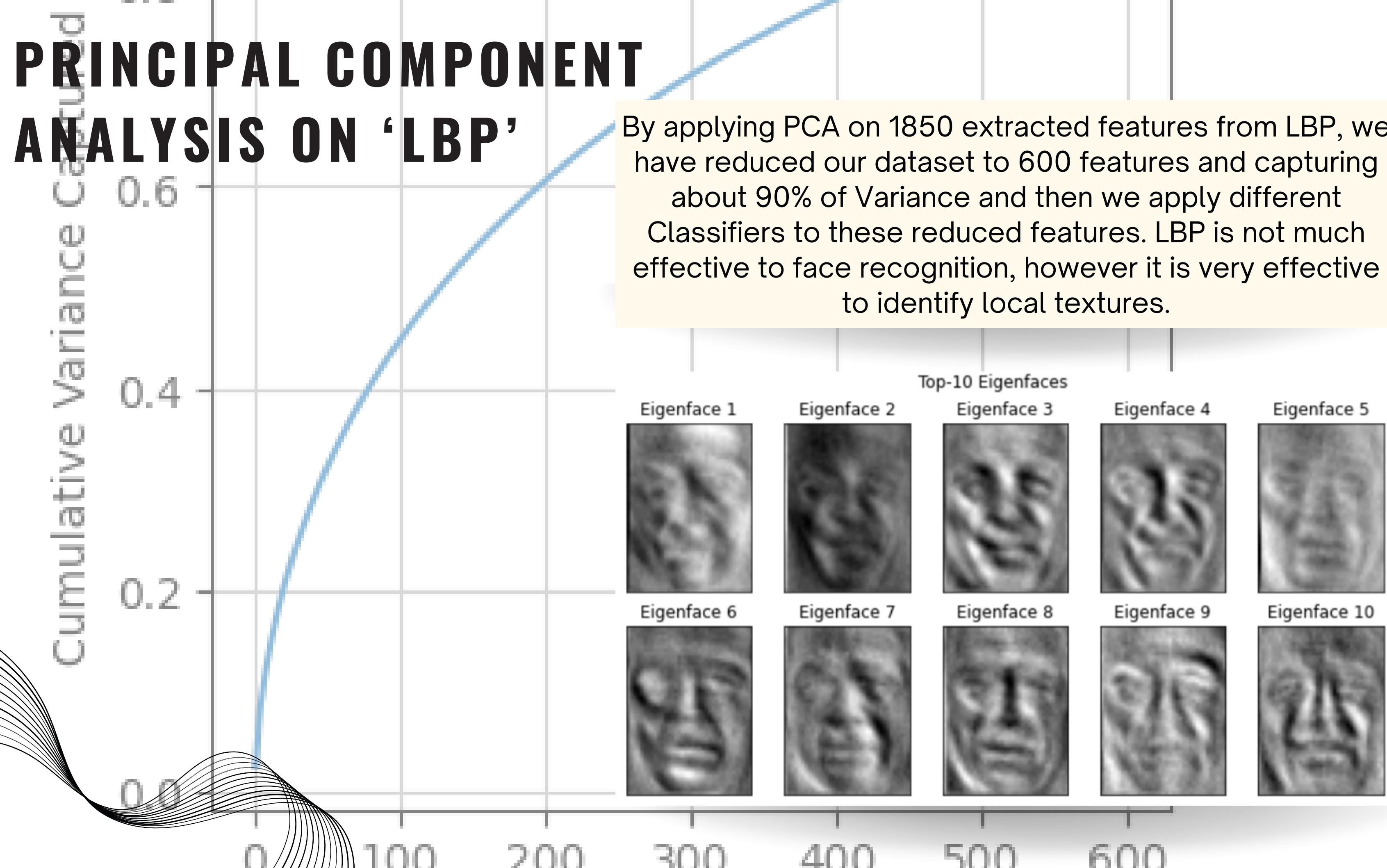
Neighbourhood
of a grey scale
image

3	7	2
8	4	1
2	3	5

Here we are Calculating new LBP pixel value By Taking its 8 neighbors into account. The difference between the central pixel and its adjacent 8 pixels is extracted as local texture feature representation.



PRINCIPAL COMPONENT ANALYSIS ON 'LBP'



LBP+PCA CLASSIFICATION RESULTS

Confusion Matrix																	
True labels	0	1	3	0	2	2	0	0	0	3	1	1	0	0	0	1	
	0	-1	3	0	2	2	0	0	0	3	1	1	0	0	0	1	
	1	-4	12	1	1	4	0	4	3	1	0	2	5	7	-4	-4	
	2	-2	2	0	5	0	0	4	0	0	3	2	2	0	-35	-35	
	3	-8	2	4	45	8	0	9	3	4	1	1	0	6	0	6	9
	4	-1	1	1	1	1	7	0	4	3	0	1	3	0	0	0	0
	5	-0	0	0	1	0	0	0	3	4	0	0	0	1	0	0	0
	6	-1	1	1	1	1	0	0	2	0	1	0	0	0	1	2	0
	7	-0	1	1	1	1	0	1	4	0	0	0	1	0	0	1	0
	8	-1	0	0	0	1	0	0	0	2	1	0	0	1	2	0	0
	9	-0	0	0	0	0	3	0	1	0	0	1	2	1	0	1	0
	10	-1	0	0	0	3	2	0	1	0	0	0	4	1	0	2	0
	11	-0	1	1	1	0	0	0	4	1	0	0	0	0	0	0	0
	12	-0	1	0	2	0	0	0	0	4	1	0	0	0	0	0	0
	13	-2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
	14	-0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	15	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Random Forests

Decision Tree Accuracy
= 33.7%
Random Forests Accuracy
= 36%

Confusion Matrix																
True labels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	0	-4	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	-5	18	2	1	1	1	1	1	1	1	1	1	1	1	1
	2	-1	3	7	1	1	1	1	1	1	1	1	1	1	1	1
	3	-5	3	4	64	6	1	1	1	2	0	3	1	2	5	5
	4	-0	3	1	2	8	0	0	0	4	2	0	0	2	0	0
	5	-0	0	1	3	0	4	0	0	0	0	1	0	0	0	0
	6	-1	1	1	1	1	1	1	1	2	0	0	0	1	0	0
	7	-0	1	1	1	1	1	1	1	1	2	0	0	0	2	1
	8	-0	1	1	1	1	1	1	1	1	2	0	0	0	2	1
	9	-0	0	1	1	1	1	0	1	0	3	0	0	0	1	0
	10	-0	0	0	1	1	0	0	0	5	1	0	0	1	0	1
	11	-1	1	0	3	0	0	0	0	1	4	0	0	1	1	0
	12	-0	2	0	0	1	0	0	0	4	0	0	0	0	0	0
	13	-1	2	0	3	0	0	1	0	0	0	0	0	0	0	0
	14	-2	2	3	1	0	0	1	0	0	0	0	0	2	0	0
	15	-1	1	2	2	5	0	0	1	0	2	2	0	0	1	1
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

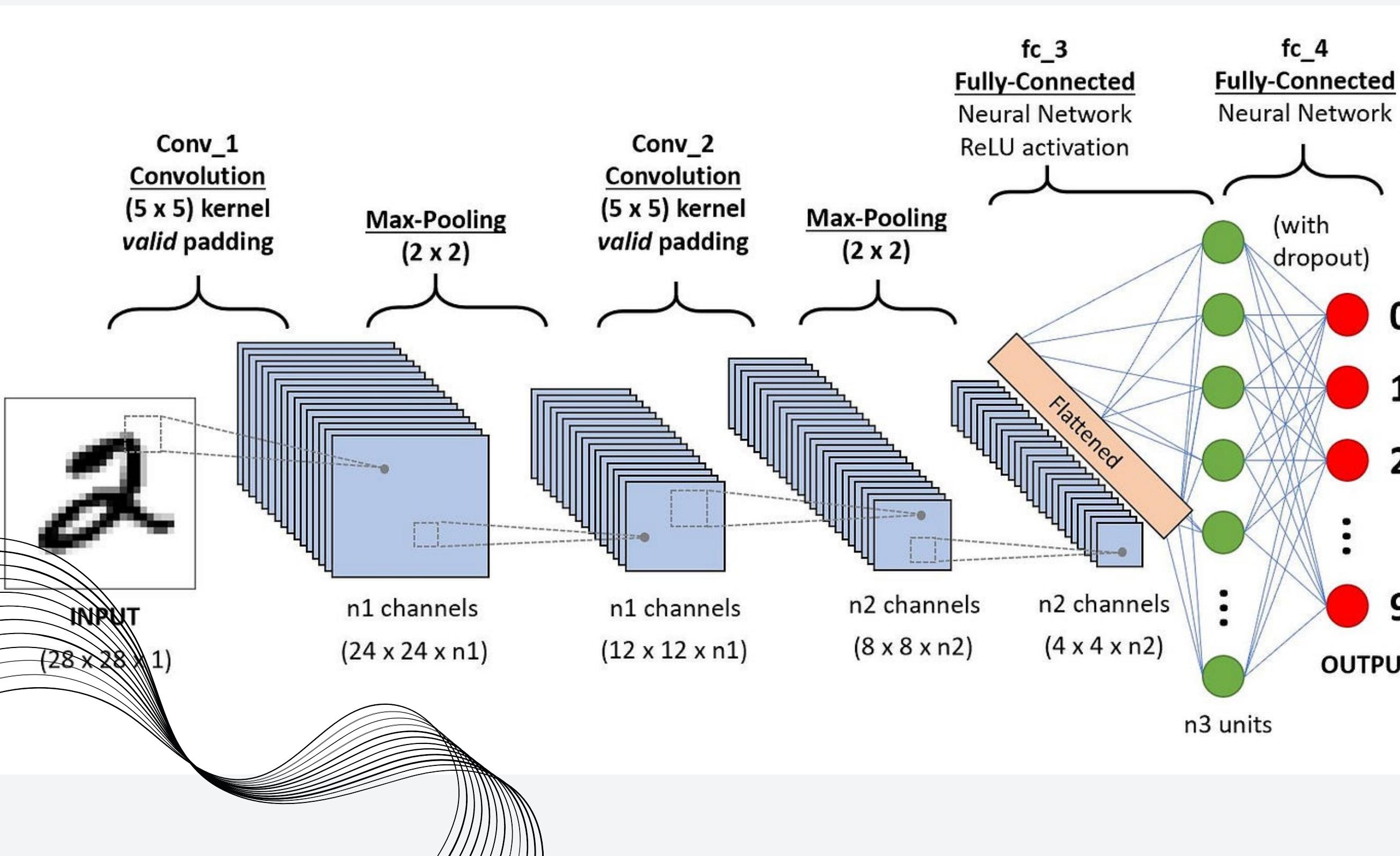
Confusion Matrix																
True labels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	0	-1	1	0	2	1	1	0	0	0	0	0	0	0	0	0
	1	-1	16	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	-2	0	3	8	2	1	0	0	0	1	1	0	0	0	0
	3	-3	0	6	3	47	2	1	3	2	2	0	2	2	3	3
	4	-1	1	1	1	1	8	0	0	0	1	0	0	2	0	0
	5	-0	0	0	0	0	6	0	0	0	0	1	0	0	0	0
	6	-1	0	3	1	0	0	0	0	0	0	0	0	1	2	0
	7	-0	2	1	1	1	0	0	2	0	0	1	0	0	0	0
	8	-0	0	0	2	0	1	0	2	0	0	0	0	1	0	0
	9	-0	1	1	0	2	0	0	0	4	0	0	0	0	0	0
	10	-0	1	0	0	2	0	0	0	6	0	0	0	0	0	0
	11	-0	1	0	0	0	0	0	1	3	0	0	0	0	0	0
	12	-1	0	0	4	0	0	0	0	0	0	0	0	1	0	0
	13	-1	0	0	4	0	0	0	0	0	0	0	1	0	0	0
	14	-1	2	2	3	0	1	0	0	2	0	0	0	0	9	0
	15	-0	2	1	1	0	0	1	0	0	0	0	0	1	0	0

SVM

Linear Kernel Accuracy
= 71.12%
Polynomial Kernel Accuracy
= 44.5%
RBF Kernel Accuracy
= 67%

Confusion Matrix																
True labels	0	<th														

CONVOLUTIONAL NEURAL NETWORKS



To apply CNN on the face data, we have first converted our face images to 60 batches each of batch size 28 samples.

Using CNN, we have extracted 2048 features. Then we have used classification techniques like SVM, Logistic Regression etc.

CNN CLASSIFICATION RESULTS

Confusion Matrix																
True labels	0	1	3	0	2	2	0	0	0	0	3	1	1	0	0	-45
	0	1	4	12	1	1	4	0	4	2	1	3	1	0	0	-44
	1	2	2	2	0	5	0	0	4	0	0	4	0	0	0	-43
	2	8	2	4	45	8	0	9	3	4	1	1	0	6	0	-35
	3	1	1	1	1	1	7	0	4	3	0	1	3	0	0	-30
	4	0	0	0	1	0	0	3	4	0	0	0	1	0	0	-25
	5	1	1	1	1	0	0	2	0	1	0	0	0	0	1	-20
	6	0	1	1	1	1	0	1	4	0	0	0	1	0	0	-15
	7	1	0	0	0	1	0	0	0	2	1	0	0	1	2	-10
	8	0	0	0	0	3	0	1	0	0	1	2	1	0	1	-5
	9	0	0	0	0	0	3	0	1	0	0	1	2	1	0	0
	10	1	0	0	0	3	2	0	1	0	0	0	4	1	0	0
	11	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0
	12	0	1	0	0	0	3	2	0	1	0	0	0	0	0	0
	13	2	2	2	0	1	0	0	0	4	1	0	0	0	0	0
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	15	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

SVM

Linear Kernel Accuracy
= 85.5%

Polynomial Kernel Accuracy
= 86%

RBF Kernel Accuracy
= 86%

Confusion Matrix																
True labels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	2	0	1	0	3	8	2	1	0	0	1	1	0	0	0	-40
	3	0	6	3	47	2	1	3	2	2	0	2	2	2	3	-35
	4	1	1	1	1	8	0	0	0	1	0	0	1	0	2	-30
	5	0	0	0	0	0	6	0	0	0	0	1	0	0	0	-25
	6	-1	0	3	1	0	0	0	0	0	0	0	0	1	2	-20
	7	0	2	1	1	1	0	0	2	0	0	0	1	0	0	-15
	8	0	0	0	2	0	1	0	2	0	0	0	1	0	0	-10
	9	0	1	1	0	2	0	0	0	0	4	0	0	0	0	-5
	10	0	1	0	0	2	0	0	0	0	0	6	0	0	0	0
	11	0	1	1	0	0	0	0	0	0	0	1	3	0	0	0
	12	1	0	0	3	0	0	0	1	0	0	0	0	1	0	0
	13	0	1	1	0	1	0	1	0	0	0	0	3	0	0	0
	14	2	3	1	3	0	1	0	0	2	0	0	0	0	9	0
	15	0	1	1	1	1	0	0	1	0	0	0	0	1	0	0

Confusion Matrix																
True labels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	0	2	2	0	3	0	0	0	1	2	0	1	0	0	0	1
	1	0	15	4	6	1	0	0	2	1	2	2	0	1	0	1
	2	0	4	8	2	0	0	0	0	1	1	0	0	1	1	0
	3	1	3	2	53	2	1	2	2	3	0	2	1	3	0	4
	4	1	2	1	2	8	0	0	0	0	0	0	1	0	1	0
	5	0	0	0	0	0	6	0	0	0	0	1	0	0	0	0
	6	1	1	3	1	0	0	0	0	0	0	0	0	0	0	2
	7	0	1	3	1	1	0	1	1	0	0	0	0	0	0	0
	8	0	0	2	0	2	0	1	0	0	0	0	0	1	0	0
	9	0	2	1	0	3	0	0	0	2	0	0	0	0	0	0
	10	0	1	0	2	0	0	0	0	6	0	0	0	0	0	0
	11	0	2	0	0	1	0	0	0	0	1	3	0	0	0	0
	12	1	0	0	4	0	0	0	0	0	0	1	0	0	0	0
	13	0	1	1	0	1	0	1	0	0	0	3	0	0	0	0
	14	1	2	2	2	3	0	1	0	0	2	0	0	0	9	0
	15	0	2	1	1	1	0	0	1	0	0	0	0	1	0	0

Confusion Matrix																
True labels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	2	0														

FINAL CLASSIFIER SELECTION



FEATURE EXTRACTION

HOG
**Histogram of Oriented
Gradients**



FEATURE REDUCTION

PCA
**Principal
Component
Analysis**



CLASSIFIER

RBF SVM
**Support Vector
Machine**



ACCURACY

88.38 %



FEATURE
EXTRACTION

CNN
Convolutional Neural
Networks



CLASSIFIER

Multilayer
Perceptron



ACCURACY

89.52 %