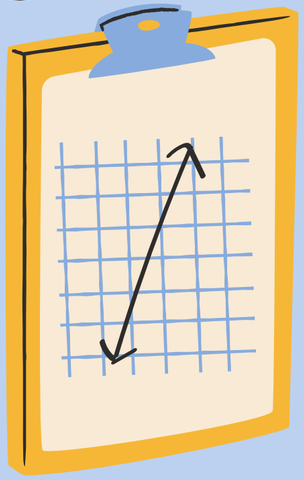
FACE RECOGNITION USING EIGENFACES

PROJECT OBJECTIVE

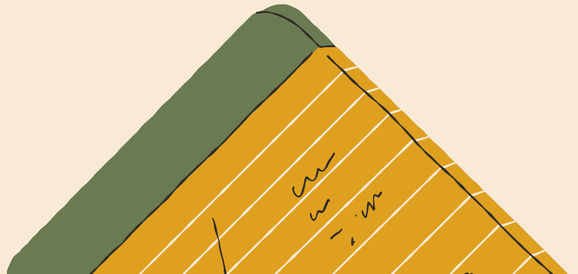


- **WHAT? -** using eigenfaces algorithm and PCA for dimensionality reduction to identify faces**.**
- **WHY eigenfaces? -** they are efficient in identifying different angles of same face and also retains essential features of a face because of PCA.
- **HOW? -** we will use PCA which speeds up the calculation while retaining the essential features and we will use KNN as it is more efficient in classification problems with small datasets.

WHY FACE RECOGNITION?

security and surveillance - Face recognition is widely used by police and security agencies to identify suspects from surveillance footage or to scan crowds for individuals on watchlists.

Attendance and access control - Face recognition systems are increasingly used in offices and schools to track attendance, replacing traditional time clocks or ID scans.



DATASET OVERVIEW(AT&T)

- Number of Subjects: 40 unique individuals.
- Number of Images: 400 images in total, with 10 images per person.
- Image Resolution: Each image is grayscale with a resolution of 92 x 112 pixels.

- **variability in images** - angles, facial expressions, accessories, lighting

- **Dataset format** - grayscale with simple background.

- Due to limited images of persons and variations it gives a semi realistic setting of an actual application.

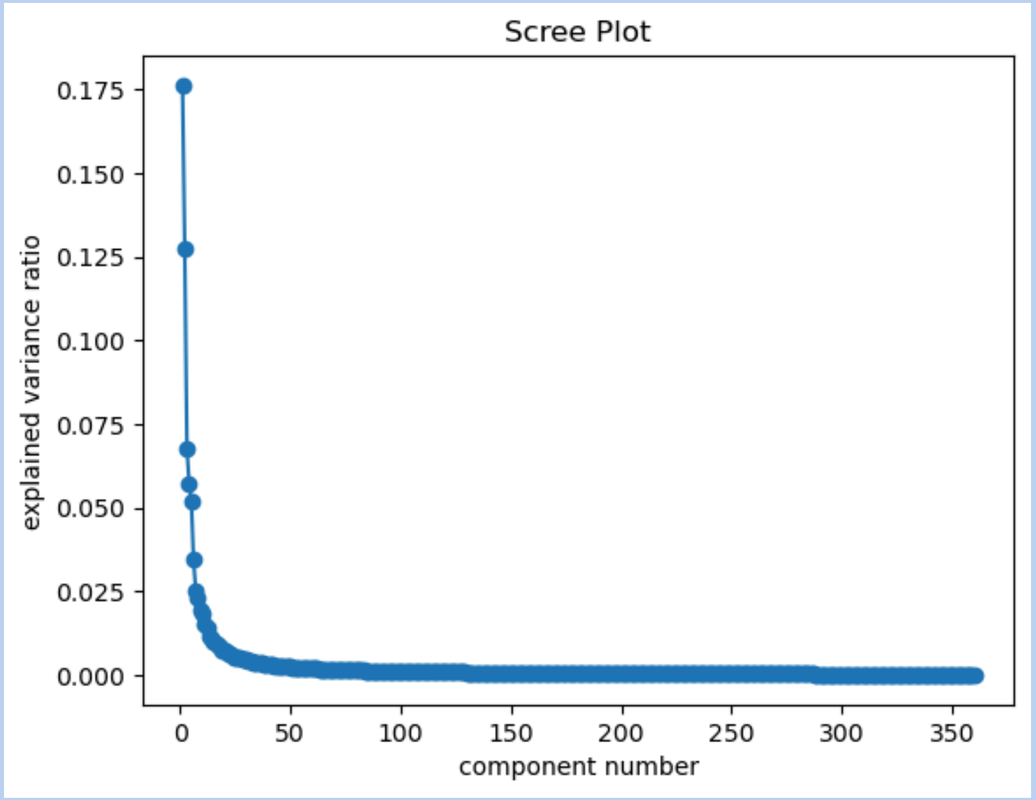
PRINCIPLE COMPONENT ANALYSIS - WHAT? AND WHY?

What is PCA? - PCA is a statistical method that reduces the number of variables in a dataset by finding new, uncorrelated variables (principal components).

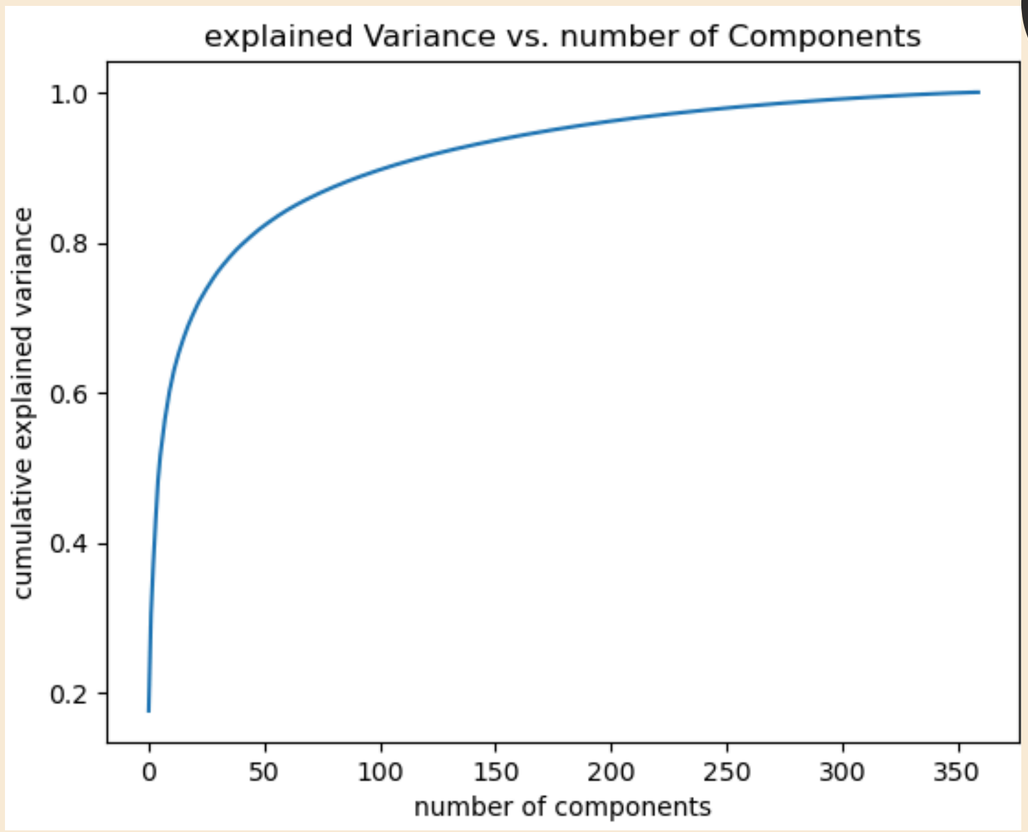
Why PCA in face recognition? - Face images are large in terms of pixels, making them high-dimensional data that requires simplification.PCA identifies the most essential components (eigenfaces) that represent unique features, like edges and contours, in faces reducing overall noise.

APPLYING PCA TO THE DATASET

- we split the available dataset of 400 images to 70%(280) for training and 30%(120) for testing.
- each image is of size 92x112 pixels, Hence the no. of total features would be 280x92x112 = 2885120, so it is a computationally intensive task to process them.
- so we will use PCA to reduce the dimensionality without losing the most essential features, we will use a scree plot to find the optimal no. of components to be used.
- Below plots will give us an idea on how PCA works.



Scree plot: A scree plot is a graphical representation used to visualize the eigenvalues or explained variance of principal components, typically in factor analysis or principal component analysis (PCA), to determine the number of components to retain..

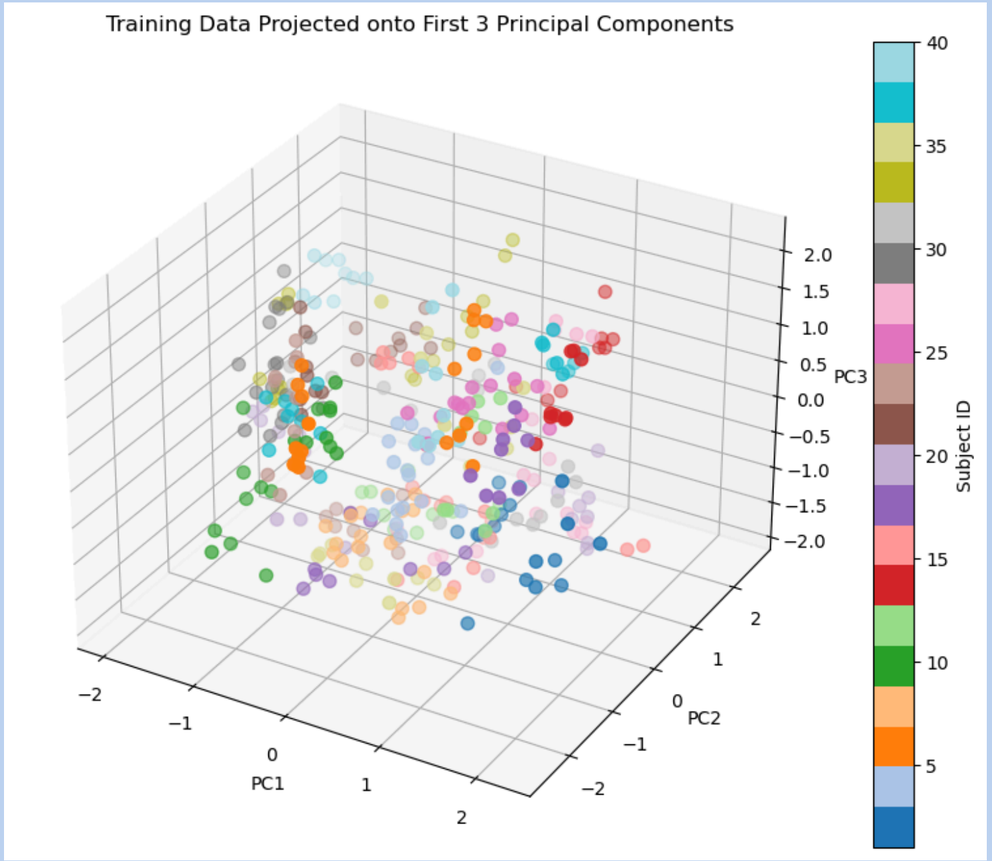


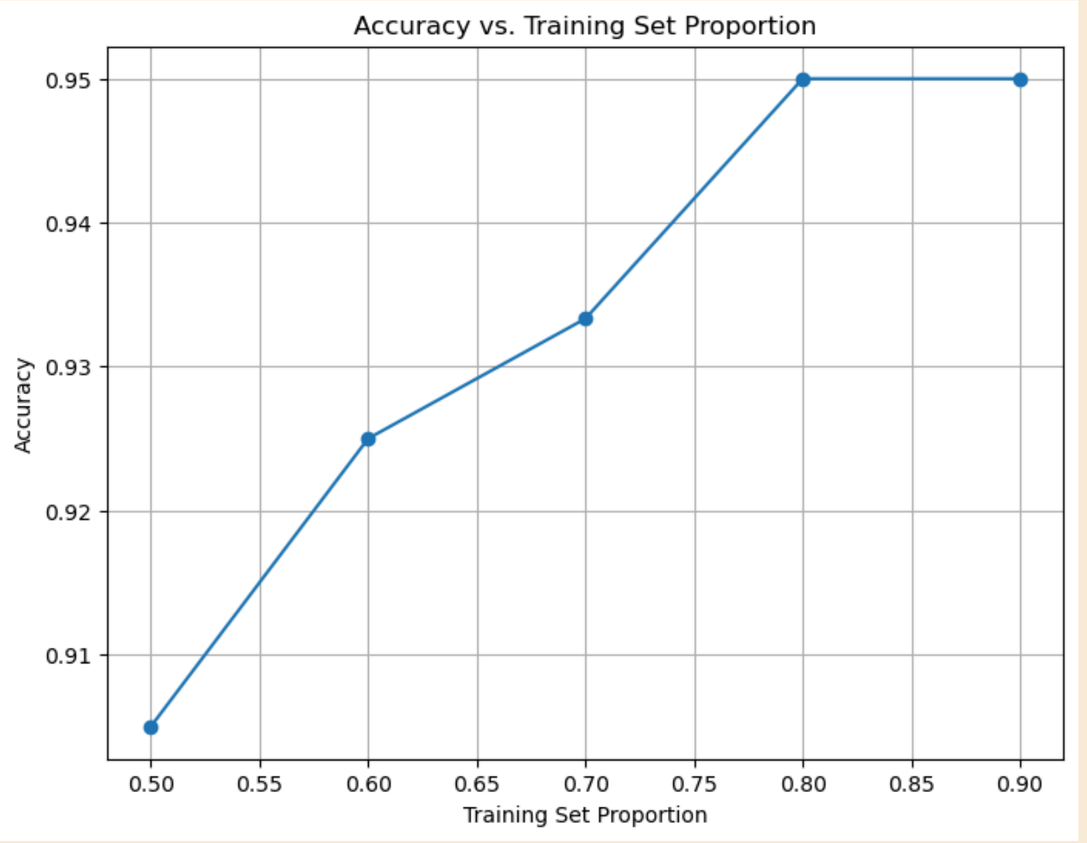
Explained variance: Explained variance is the proportion of a dataset's total variance captured by each principal component in PCA, indicating how much information each component holds.

K-NEAREST NEIGHBOURS(KNN) - WHAT? AND WHY?

What is KNN? - KNN (K-Nearest Neighbors) is a supervised machine learning algorithm that classifies a data point based on the majority class of its 'k' closest neighbors in the feature space.

Why KNN? How? - KNN is used in face recognition with eigenfaces because it is simple, effective in low-dimensional spaces, and can classify new faces without needing retraining, KNN is used in face recognition with eigenfaces to classify faces by comparing the test image's reduced feature vector to those of stored images based on proximity.


Training Data Projected onto First 3 Principal Components
PC1
PC2
PC3
Subject ID


Accuracy vs. Training Set Proportion
Accuracy
Training Set Proportion

MODEL TRAINING AND TESTING

- **Training process :** The K-Nearest Neighbors (KNN) classifier was trained using the PCA-transformed training dataset. By reducing the dimensionality with PCA, the training data was projected onto a lower-dimensional subspace, capturing essential features while reducing computation complexity for KNN.

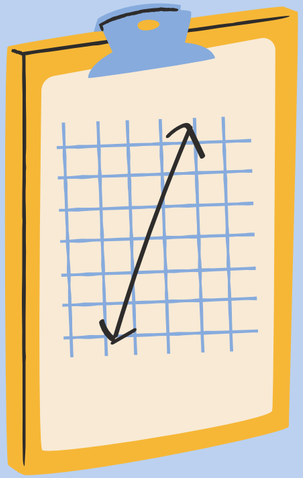
- **Testing approach :** The trained KNN classifier was evaluated on the PCA-transformed test dataset. This involved using the PCA-reduced feature representation of the test samples to predict the labels, providing a consistent basis for both training and testing phases.

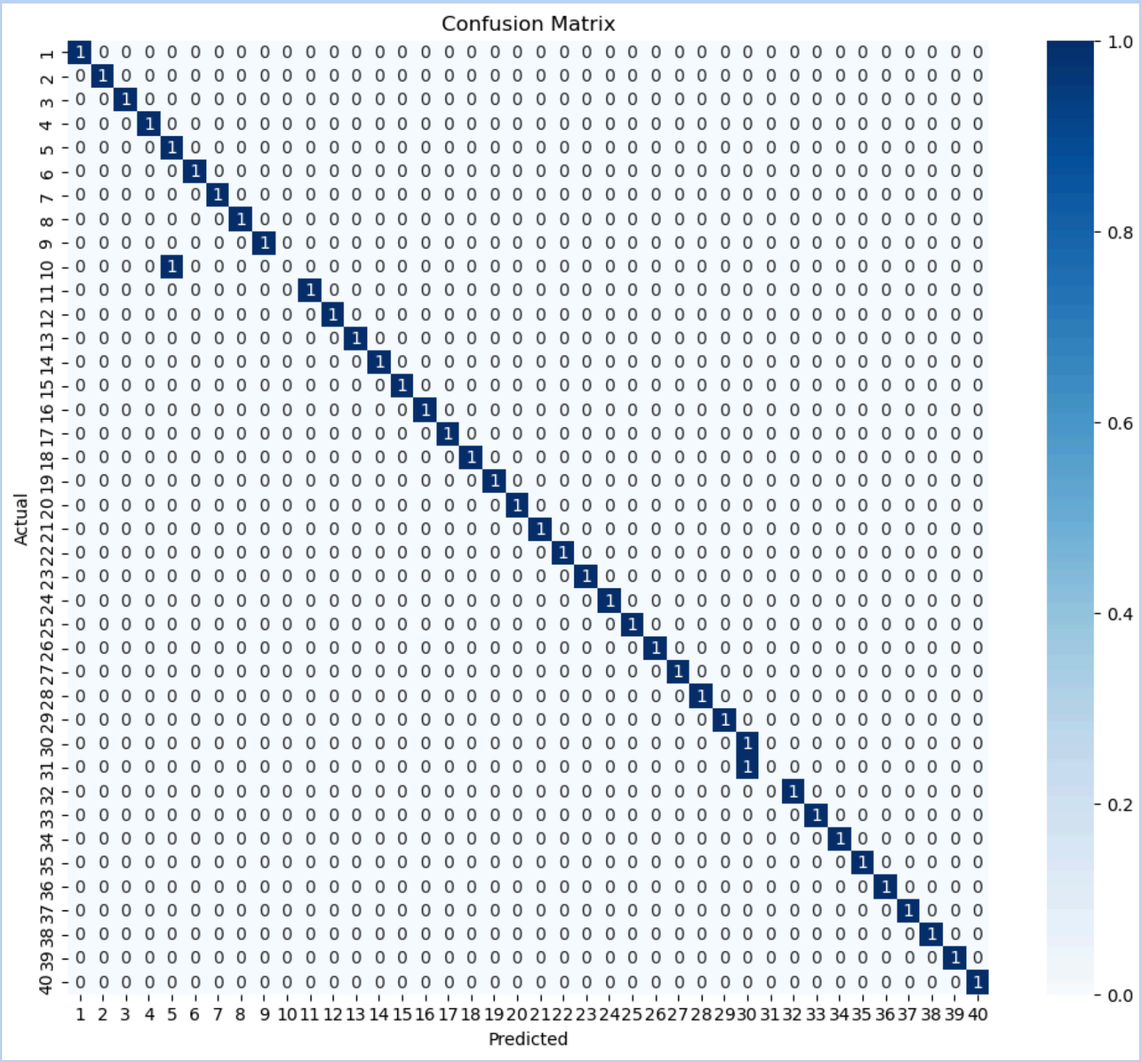
EVALUATION METRICS

Confusion matrix : A confusion matrix summarizes a model's performance, showing true and false predictions for each class category.

- **Accuracy:** The proportion of total predictions that are correct, calculated as the sum of true positives and true negatives divided by all predictions.
- **Precision:** The ratio of correctly predicted positive observations to the total predicted positives, measuring the accuracy of positive predictions.
- **Recall:** The ratio of correctly predicted positive observations to all actual positives, showing the model's ability to identify all relevant cases.

CONFUSION MATRIX

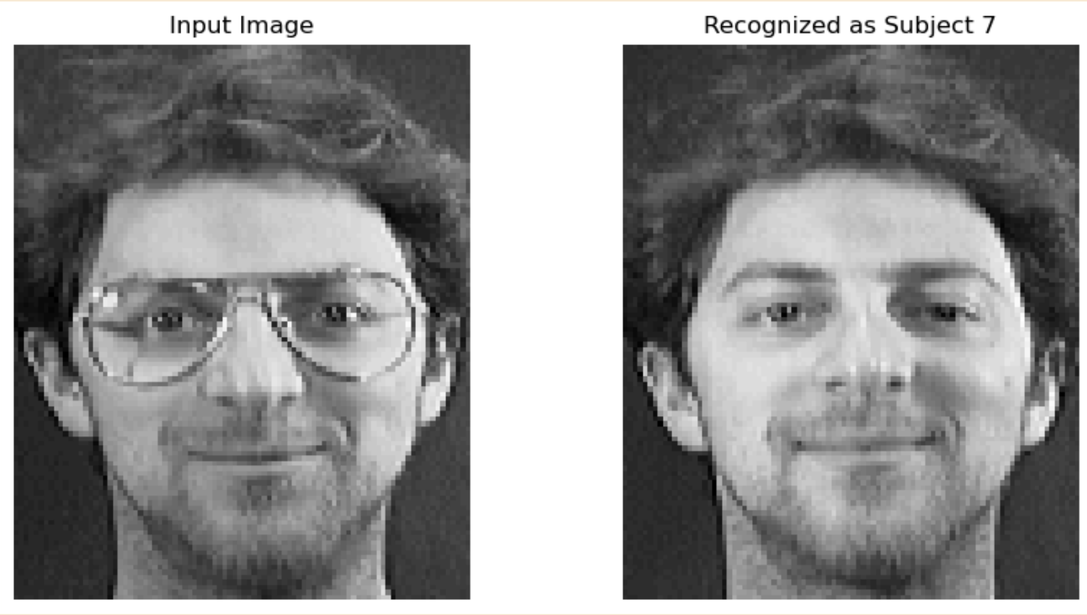




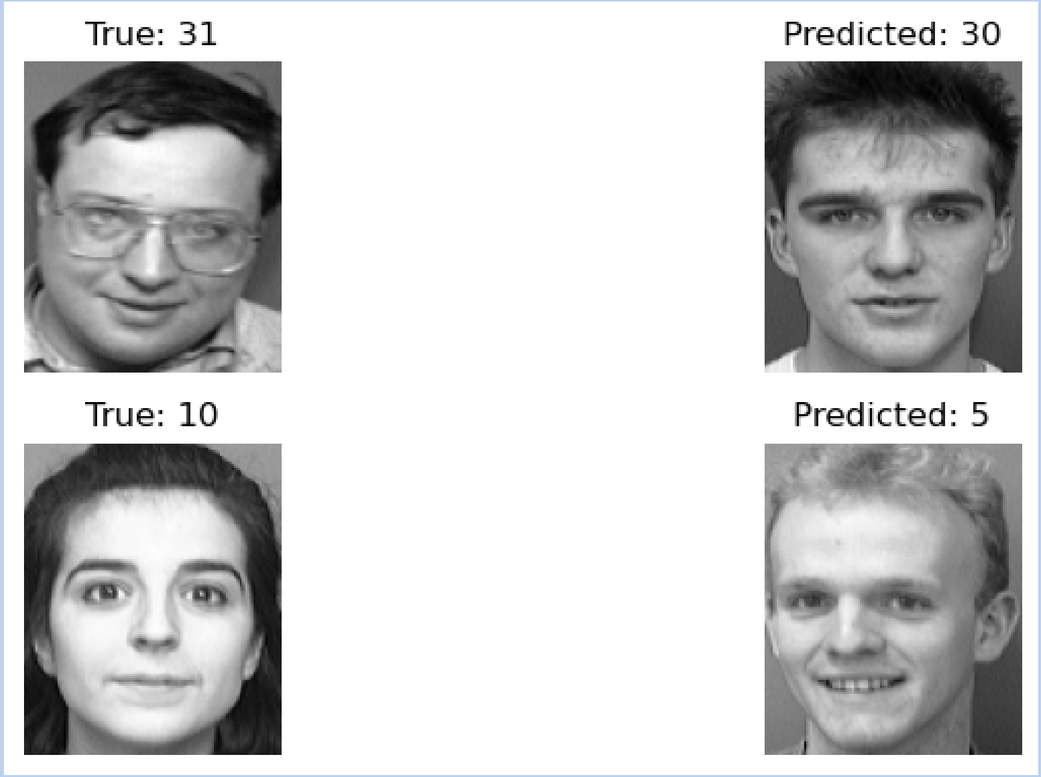
VISUALIZATION OF EIGENFACES

SAMPLE PREDICTION AND ERROR ANALYSIS

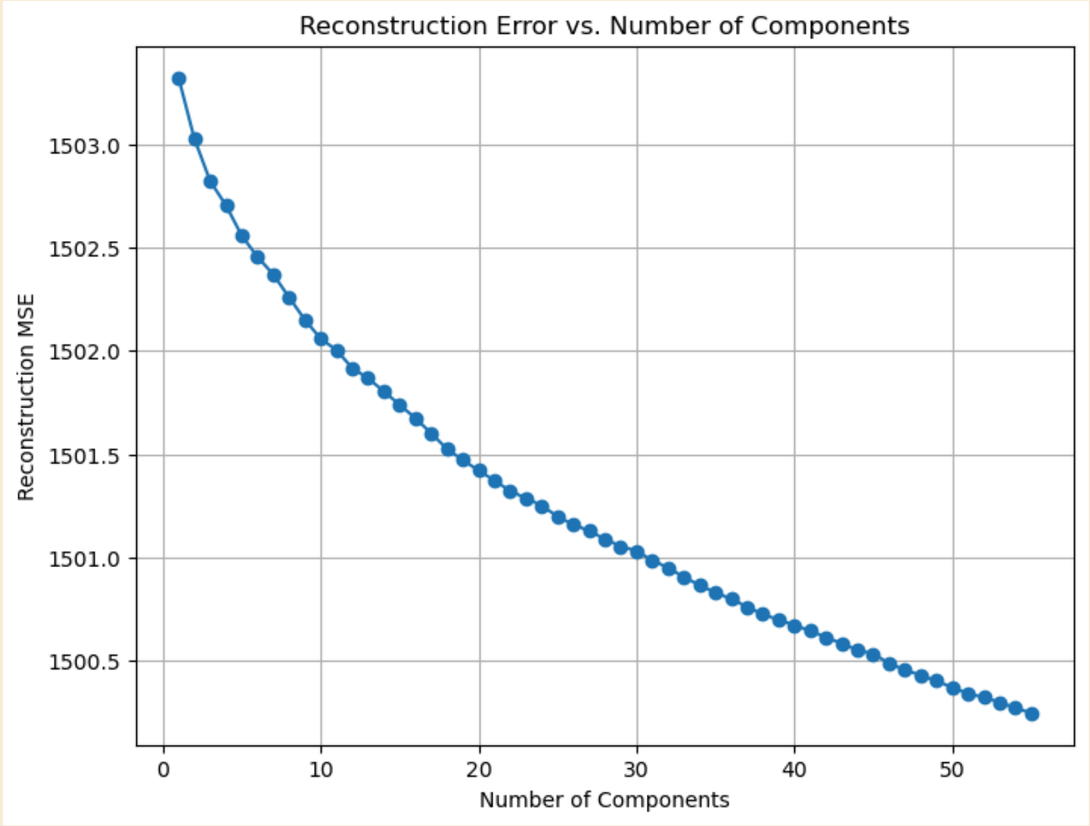
- Due to the limited size of our dataset and quality of images our model is proven to be 93% accurate.
- It even predicts the face without any accessories.
- Our model failed to classify class 10 i.e, person 10 and person 31.
- In our case the background lighting and very close i.e, similar features affected the model's accuracy.
- Therefore, even if eigenfaces algorithm proved to be useful on small datasets like AT&T it may not perform well on datasets containing visually similar images.



MISSCALSSIFIED SAMPLES



RMSE PLOT



CONCLUSION & FUTURE IMPROVEMENTS

- Instead of just recognizing gray scale images we can use three different matrices to try and classify colored images of faces.
- Despite the overall success of the system, we identified several limitations that affect recognition accuracy. One significant limitation is the system’s difficulty in distinguishing between faces with very similar features
- Implementing convolutional neural networks (CNNs) could enable the system to learn more complex and hierarchical features from facial images, improving its ability to distinguish between similar faces and handle variations in lighting, pose, and expression


THANKING YOU