Object Recognition (P-22)

Pranav Firake
ppfirake
ppfirake@ncsu.edu

Saloni Desai
sndesai
sndesai@ncsu.edu

Harshal Gurjar
hkgurjar
hkgurjar@ncsu.edu

FNU Manushri
manush
manush@ncsu.edu

Introduction

- To extract features and apply classifier to recognize objects in images.
- We aimed on using PCA for image compression and comparing object recognition results of KNN, clustering and CNN

Motivation

- Vision is one of the most essential human senses and it is important role in human perception about surrounding environment.
- It is not easy for the blind to organize daily simple activities and is especially difficult for them to distinguish between different items just by the sense of touch.
- Aim of the project is to provide accessibility for visually challenged people

Dataset

CALTECH 256 Dataset

- 256 object categories
- Total 30608 pictures
- Average of 119 pictures in each category
- Dimensions vary for each image







Principal Component Analysis (PCA)

- Uses **orthogonal** transformation to convert a high-dimensional dataset into a low dimensional space.
- Approach : Flatten the image,applied PCA on it and reconstructed by stacking the pca applied images for each R, G, B.
- We have used 75 principal components.
- It captures 98.6% of the variance

Why:

- For dimensionality reduction
- Faster image processing

K-Means Clustering

Approach :

- Resized the image to fixed dimension.
- Flatten the images.
- Used 70% of these flatten images as a training set to train the K-Means Classifier.

Why:

- Extracts local features of an image and groups them into appearance clusters.
- Captures a large variability of the local image structure.

K-Nearest Neighbors

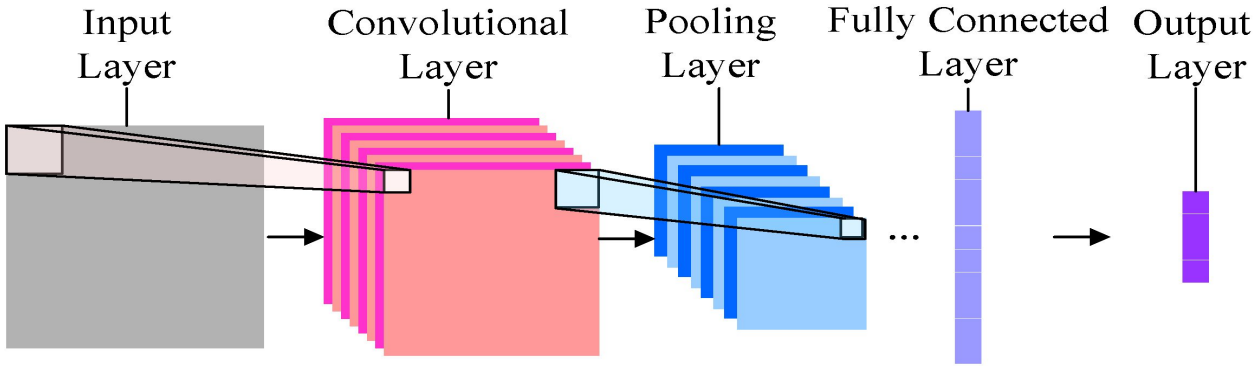
- Inputs used :
  - Raw pixels – Images are resized into a fixed width and length (32x32) and the pixel intensities have been flattened into a single list of numbers.
  - Color Histogram – Color histogram is created for each image to characterise their color distribution.
- We tried KNN with 10 different values of k and analysed results.

Why:

- Permits recognition based on color, texture and shape
- Good results when we have versatile and large number of training images

Convolutional Neural Networks (CNN)

- Convolutional Neural Networks are made up of neurons that have learnable weights and biases.
- We have used 64*64 dimension images as a feed to input layer
- 128 epochs
- Adam Optimizer



Why:

- Better accuracy
- Can learn the high level as well as low level features of an image and hence we don’t need to do feature engineering

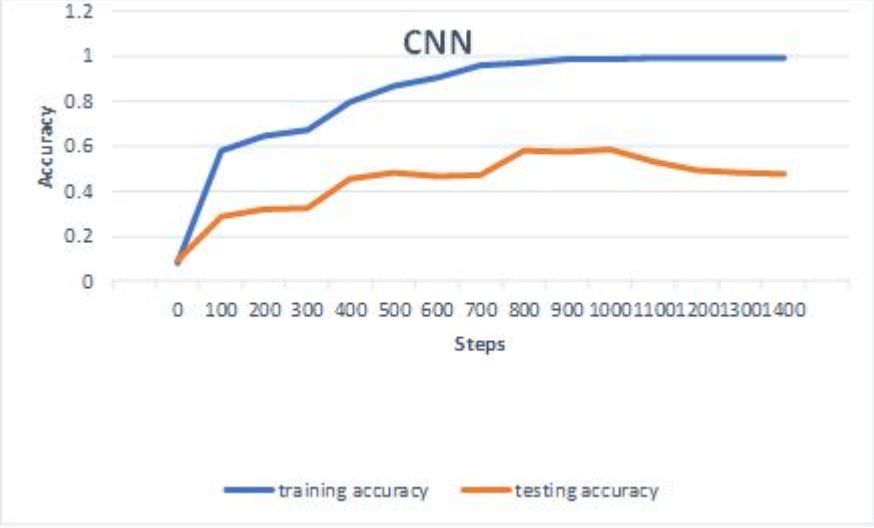
Results

KNN with and without PCA

Max Accuracy 38%

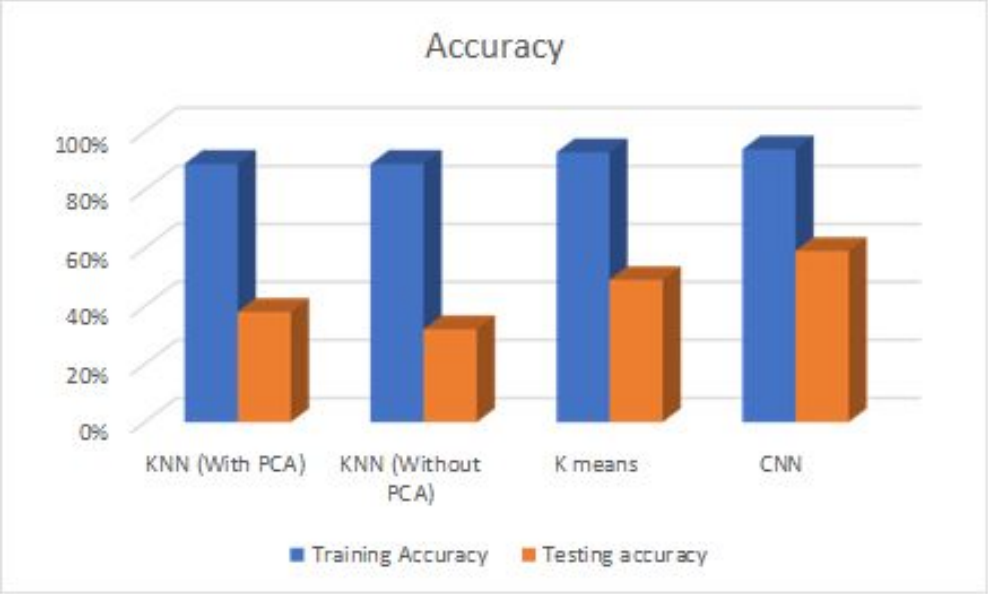
CNN

Max Accuracy 59%



Comparison

Method	Training Accuracy	Testing accuracy
KNN (With PCA)	89%	38%
KNN (Without PCA)	89%	32%
K means	93%	49%
CNN	94%	59%



Conclusion

- Unsupervised techniques like K-Means clustering and CNN tend to perform better than supervised techniques like KNN.
- CNN is more efficient as compared to KNN
- CNN model requires more training time as compared to KNN
- With PCA we get better accuracy with KNN
- K means gave better results than KNN

Applications of Object Recognition

- Accessibility for visually challenged people
- Object and face recognition over social media
- Computer vision
- Image Search
- Face recognition for security


Thank
you

REFERENCES

- *SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition*
- *Efficient clustering and matching for object class recognition*
- *Recurrent Convolutional Neural Network for Object Recognition*