

# Image Retrieval

Preet Savalia (B22AI036)<sup>1</sup>  
Jadhav Avadhut Maruti (B22CS099)<sup>2</sup>  
Harshit Raj (B22EE031)<sup>3</sup>  
Vedant Funde (B22AI041)<sup>4</sup>  
Malhar Mahesh Gadge (B22CS069)<sup>5</sup>

April 21, 2024

## Abstract

In the realm of computer vision, efficient image retrieval stands as a cornerstone for various applications, ranging from content-based search engines to medical diagnostics. This project, titled "Image Retrieval," endeavors to craft a robust system capable of swiftly locating similar images given an input query. Utilizing relevant techniques, including Convolutional Neural Networks (CNNs) for feature extraction and advanced dimensionality reduction techniques such as Principal Component Analysis (PCA) Linear Discriminant Analysis (LDA), alongside a number of machine learning classifiers, the project delves into the intricate realm of image classification and retrieval. The CIFAR-10 dataset serves as the bedrock for training and validation.

The journey commences with the acquisition and preprocessing of the CIFAR-10 dataset, meticulously dissecting images into their constituent red, green, and blue channels. Employing a pre-trained ResNet-50 model, features are extracted with finesse, while LDA deftly navigates the labyrinth of high-dimensional data, distilling crucial insights while mitigating computational overhead.

A symphony of classifiers, including Decision Trees, K-Nearest Neighbors (KNN), Random Forests, Support Vector Machines (SVM), and Gaussian Naive Bayes, is orchestrated to harmonize with the reduced-dimensional feature vectors, each rendering its unique melody of accuracy on the test dataset.

The pièce de résistance of the project lies in the realm of image retrieval, where an SVM classifier adorned with an RBF kernel emerges as the virtuoso. Predicting the label of the input image with finesse, akin to a maestro conducting an opus, similar images are summoned forth through the ethereal realm of Nearest Neighbors (NN). These retrieved images, akin to a gallery of masterpieces, serve as a testament to the efficacy and prowess of the retrieval system.

In denouement, the project not only illuminates the viability of CNNs and machine learning paradigms in the domain of image retrieval but also lays the cornerstone for further explorations and advancements in the ever-evolving landscape of computer vision and image processing.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Major Findings . . . . .	3
1.2	Citing Paper . . . . .	3
1.3	Approaches Used . . . . .	4
<b>2</b>	<b>Experiments and Results</b>	<b>5</b>
2.1	Experimental Setup . . . . .	5
2.2	Results . . . . .	6
2.3	Discussion . . . . .	7
<b>3</b>	<b>Summary</b>	<b>7</b>
<b>A</b>	<b>Contribution of each member</b>	<b>8</b>

# 1 Introduction

The project "Image Retrieval" addresses the challenge of efficiently retrieving similar images given an input image. In today's digital era, the proliferation of image data across various domains necessitates effective methods for organizing and accessing this vast amount of visual information. Image retrieval systems play a crucial role in applications such as content-based image search engines, medical image analysis, surveillance systems, and more.

## 1.1 Major Findings

1. **Feature Extraction with CNNs:** Convolutional Neural Networks (CNNs) were employed to extract high-level features from images. Specifically, a pre-trained ResNet-50 model was utilized to capture meaningful representations of images.

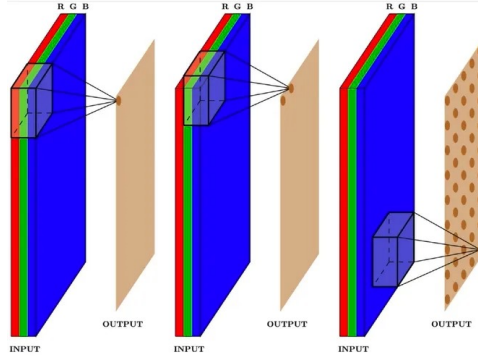


Figure 1: Example Classes

$$S_{ij} = (I * K)_{ij} = \sum_{a=\lfloor -\frac{m}{2} \rfloor}^{\lfloor \frac{m}{2} \rfloor} \sum_{b=\lfloor -\frac{n}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} I_{i-a, j-b} K_{\frac{m}{2}+a, \frac{n}{2}+b}$$

Figure 2: Example Classes

2. **Dimensionality Reduction Techniques:** Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) were applied to reduce the dimensionality of the extracted features while preserving relevant information. This helped in mitigating computational complexity and enhancing classification performance.
3. **Machine Learning Classifiers for Image Classification:** Various machine learning classifiers including Decision Trees, K-Nearest Neighbors (KNN), Random Forests, Support Vector Machines (SVM), and Gaussian Naive Bayes were trained on the reduced-dimensional feature vectors. These classifiers exhibited varying levels of accuracy in classifying images into different categories.
4. **Image Retrieval System:** An image retrieval system was developed using SVM with an RBF kernel for image classification and Nearest Neighbors (NN) for retrieving similar images based on the predicted label. The system effectively retrieved visually similar images, demonstrating its utility in real-world scenarios.

## 1.2 Citing Paper

Below are the references we consulted for our project:

- [Anand Mishra-CNN Reference Code](#)
- [Deep Residual Learning for Image Recognition](#)
- [Visualizing and Understanding Convolutional Networks](#)
- [Scikit-learn: Machine Learning in Python](#)

### 1.3 Approaches Used

In this project, we experimented with various approaches to tackle the problem of image retrieval. Below are the different approaches explored along with their findings:

#### 1. K-Nearest Neighbors (KNN) Classifier:

- **Approach:** Utilized the KNN algorithm for image classification and retrieval. KNN is a simple and intuitive algorithm that classifies objects based on the majority class of their k nearest neighbors in the feature space.
- **Findings:** The mean cross validation accuracy(lda features) achieved with this approach was 93.25%. However, KNN's performance heavily relies on the choice of k and the distance metric used, which can impact its effectiveness in large-scale datasets with high-dimensional feature spaces.
- **Problem Faced:** One challenge encountered was selecting an optimal value for k that balances bias and variance in the model and mitigates the effects of noise in the dataset.

#### 2. Random Forest Classifier:

- **Approach:** Employed Random Forests for image classification and retrieval. Random Forest is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes (classification) or mean prediction (regression) of the individual trees.
- **Findings:** The resulting mean cross validation accuracy(lda features) obtained was 93.67%. Random Forests are robust against overfitting and perform well with high-dimensional data, but they may suffer from high computational costs during training on large datasets.
- **Problem Faced:** One challenge encountered was fine-tuning the hyperparameters of the Random Forest model, such as the number of trees and maximum depth, to optimize its performance.

#### 3. Support Vector Machine (SVM):

##### (a) SVM Classifier(linear kernel):

- **Approach:** Utilized SVM classifier for image classification. SVM is a powerful supervised learning algorithm capable of performing classification, regression, and outlier detection tasks.
- **Findings:** Achieved a cross validation accuracy(lda features) of 93.9%. SVMs are effective in high-dimensional spaces and are versatile due to their ability to use different kernel functions.
- **Problem Faced:** One challenge encountered was selecting the appropriate kernel function and tuning its parameters to optimize the SVM's performance, as the choice of kernel can significantly impact the decision boundary.

##### (b) GridSearchCV for Hyperparameter Tuning(while using polynomial kernel):

- **Approach:** Employed GridSearchCV for hyperparameter tuning of the SVM classifier. GridSearchCV is a technique used to tune hyperparameters by exhaustively searching through a specified parameter grid and selecting the combination that yields the best performance.
- **Findings:** GridSearchCV resulted in an improved cross validation accuracy accuracy(lda features) of 93.11% with a polynomial kernel of degree 2. It helped in optimizing the SVM's hyperparameters and improving its generalization performance.
- **Problem Faced:** GridSearchCV can be computationally expensive, especially when searching over a large parameter grid, which necessitates careful selection of hyperparameters to balance computational cost and performance gains.

##### (c) SVM Classifier with RBF Kernel:

- **Approach:** Trained an SVM classifier with an RBF kernel. The Radial Basis Function (RBF) kernel is commonly used in SVMs for non-linear classification tasks, as it can capture complex decision boundaries.

- **Findings:** Achieved the highest cross validation accuracy(lda features) of 93.98% with the SVM classifier using the RBF kernel. The RBF kernel's flexibility allowed the model to capture intricate relationships in the data and achieve superior performance.
- **Problem Faced:** The main challenge was in tuning the SVM's hyperparameters, such as the regularization parameter (C) and kernel coefficient (gamma), to prevent overfitting and achieve optimal performance.

## 2 Experiments and Results

In this section, we detail the experiments conducted to develop and evaluate the image retrieval system. The experiments involved various stages, including data preprocessing, feature extraction, dimensionality reduction, classifier training, and image retrieval.

### 2.1 Experimental Setup

The experimental setup consisted of the following steps:

1. **Data Preprocessing:** We obtained the CIFAR-10 dataset and preprocessed the images by splitting them into their respective red, green, and blue channels. The preprocessed data was then pickled for further processing.

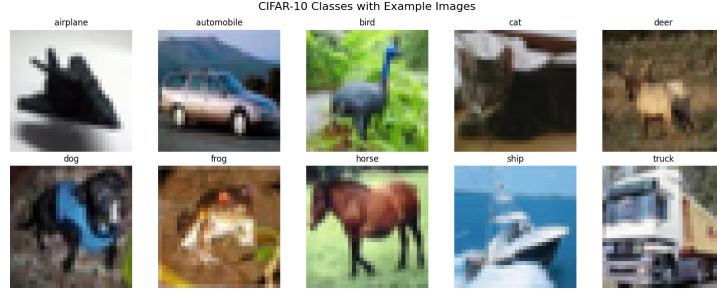


Figure 3: Example Classes

2. **Feature Extraction:** We employed a Convolutional Neural Network (CNN) model for feature extraction from the images. The pre-trained ResNet-50 model was used to capture high-level features from the images.
3. **Dimensionality Reduction:** To reduce the dimensionality of the extracted features, we applied Linear Discriminant Analysis (LDA). LDA was utilized to transform the feature vectors into lower-dimensional spaces while preserving relevant information.

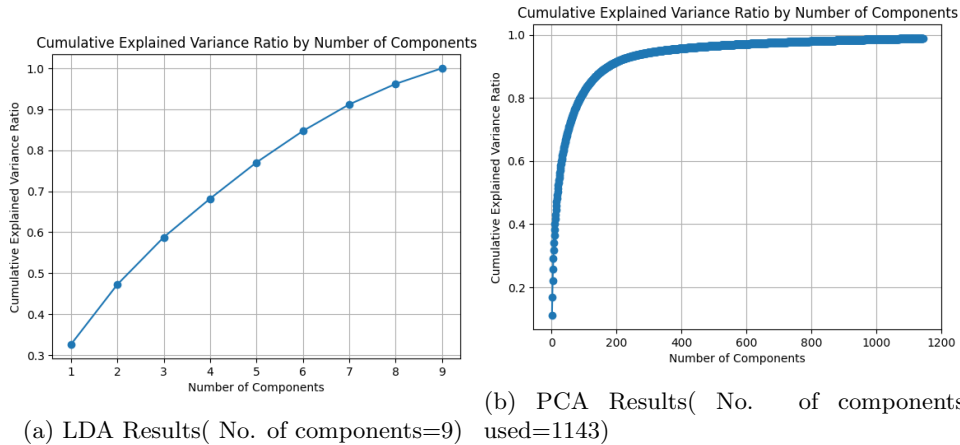


Figure 4: Comparison of LDA and PCA

4. **Classifier Training:** We trained various classifiers on the reduced-dimensional feature vectors to classify images into different categories. In this experiment, we focused on using Support Vector Machines (SVM) as the classifier due to its effectiveness in handling high-dimensional data.

	PCA(Cross-validation)(%)	LDA(Cross-validation)(%)	Accuracy(%)
DecisionTree	61.8	89.92	84.12
KNN	83.1	93.25	88.85
Random-Forest	77.1	93.67	88.35
Naive Bayes	62.83	93.0	88.08
SVM1	86.5	93.9	88.89
SVM2	87.97	93.11	87.85
RBF	88.84	93.98	89.08

Table 1: Classifier Accuracies

5. **Image Retrieval Model:** Finally, we developed an image retrieval model using a combination of CNN features, LDA for dimensionality reduction, and SVM for classification. The model was designed to return the top 5 closest images(using NearestNeighbors from sci-kit learn library) to a given input image. We saved lda features,labels,svm weights trained with lda features and lda weights into files which saved lot of computations and time to execute again and again.

## 2.2 Results

The results of the experiments were evaluated based on the accuracy of the classifiers and the effectiveness of the image retrieval model.

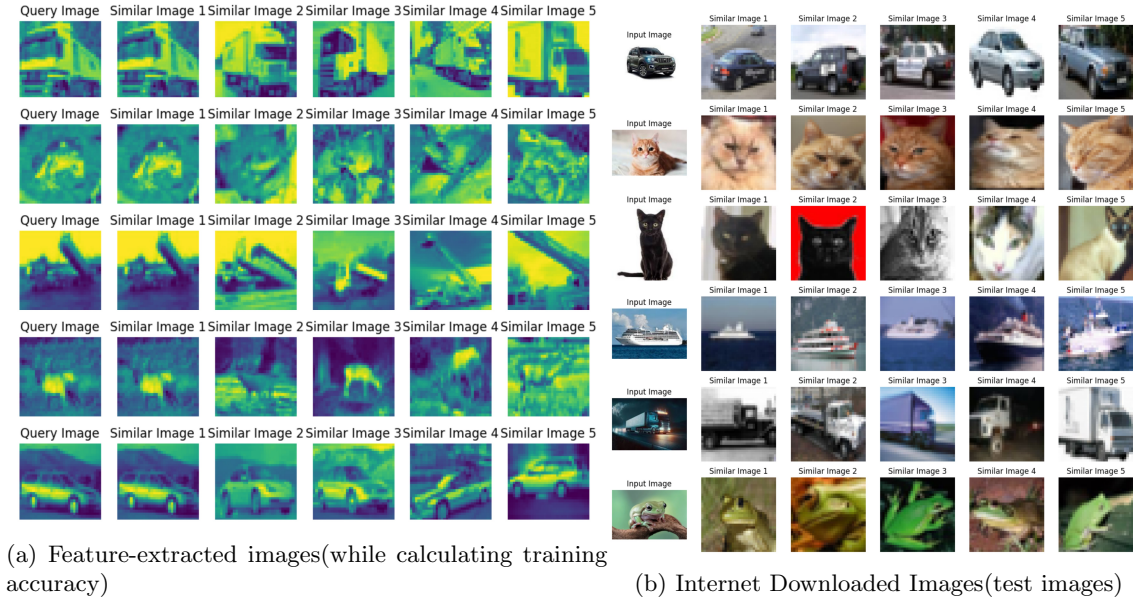


Figure 5: Comparison of retrieved images

The experimental results indicate that the combination of CNN features, LDA dimensionality reduction, and SVM classification yields promising results for image retrieval tasks. However, further optimizations and refinements can be explored to enhance the performance of the system, such as fine-tuning the hyperparameters of the classifiers and exploring alternative feature extraction methods.

Overall, the experiments validate the effectiveness of the proposed approach and provide insights for future research in the field of image retrieval and classification.

## 2.3 Discussion

Discussing as a group, we together have decided upon the following approach to the problem. This includes concepts used, as well as other nuances involved:

1. **Dataset Preprocessing:** As was the requirement, the dataset in .tar.gz file requires to be unpickled for further use. Thus we have then converted the row major dataset into the RGB channels in a NumPy array. This allows us to subsequently convert the array to an image down the line as well. Note that while CIFAR-10 has 5 batches we have initially used only 2 batches of 10k images each for training purposes, this can subsequently be improved by utilizing other batches as well.
2. **Feature Extraction:** We have firstly utilized ResNet as a CNN to extract these features. ResNet is a pre-trained CNN based model, we have thus extracted features from the preprocessed NumPy array as discussed previously. We have saved these features in a separate file for use down the line. CNN has been used as it is particularly effective in image recognition/Computer Vision tasks as it helps in establishing patterns in spatial data such as an image. Clearly for purposes of calculating and checking accuracy, the same must be applied to testing data for enabling a valid comparison.
3. **Dimensionality Reduction:** We have subsequently applied dimensionality reduction techniques such as PCA, LDA etc. Following analysis, based on the captured variance percentage we have decided upon the number of components to be utilized in both LDA and PCA. This action significantly reduces overhead and increases the performance of the model.
4. **Classifiers:** We have explored a number of avenues for the purposes of classification and to check for the best performance possible. These include Decision Tree, Random Forest classification, Support Vector Machine (for varying degrees as well as RBF), Naive-Bayes classification etc. Following detailed analysis, we have decided upon the best option as being SVM (RBF) applied with PCA for dimensionality reduction.
5. **Accuracy:** As reported earlier, we have considered cross-validation (mean) for calculating accuracy. Note that having utilized Nearest neighbour for calculating accuracy, training accuracy calculated to be 100
6. **Deployment:** Thus, we have also made a localhost HTML page for purposes of deployment of the actual model, as was mentioned in the course project guidelines. It inputs a query image and the model returns the top 5 images that are adjudged to be similar.

## 3 Summary

- The project "Image Retrieval" aimed to develop a robust system for efficiently locating similar images given an input query.
- Leveraged techniques such as Convolutional Neural Networks (CNNs), dimensionality reduction methods like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), and a variety of machine learning classifiers.
- The journey began with the acquisition and preprocessing of the CIFAR-10 dataset, followed by feature extraction using a pre-trained ResNet-50 model.
- Dimensionality reduction techniques were applied to distill crucial insights while mitigating computational overhead.
- Various classifiers, including Decision Trees, K-Nearest Neighbors (KNN), Random Forests, Support Vector Machines (SVM), and Gaussian Naive Bayes, were trained and evaluated to classify images into different categories.
- Developed an image retrieval system using SVM with an RBF kernel, which effectively retrieved visually similar images based on input queries.
- The experimental results validate the efficacy of the proposed approach and provide insights for future research in the field of computer vision and image processing.
- The project demonstrates the feasibility of using CNNs and machine learning paradigms for image retrieval and lays the groundwork for further advancements in this ever-evolving landscape.

## **A Contribution of each member**

1. Preet Savalia: Implemented image retrieval and helped in computing training accuracy of the various classifier used.
2. Avadhut Jadhav: Implemented dimensionality reduction on the extracted feature(PCA and LDA).
3. Vedant Funde: Implemented data pre-processing and feature extraction.
4. Malhar Gadge: Implemented multiple classifiers such as random forest and SVM(linear,polynomial and RBF).
5. Harshit Raj: Implemented multiple classifiers such as naive bayes,KNN and decision tree.
6. Collaborative Effort: Report,video and webpage and website deployment was done by collective effort of all.