Efficient Image Retrieval Using Convolutional Neural Networks and Dimensionality Reduction Techniques: A Case Study on CIFAR-10

Rushi Patel[1], Krishi Patel[2], Het Patel[3], Prithamshu Singh[4], Roma Mevadawala[5]

[1, 2, 3, 4]Student, [5]Assistant Professor
Department of Information Technology, Indus institute of technology and engineering

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 swiftlylocating 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 navigate the labyrinth of high-dimensional data, distilling crucial insights while mitigating computational overhead.

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 navigate the labyrinth of high-dimensional data, distilling crucial insights while mitigating computational overhead.

The piece de resistance 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 Neighbours (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.

Keywords: Convolution Neural Network (CNN), Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), ResNet-50 model, Support Vector Machine (SVM)

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

- **Feature Extraction with CNNs:** Convolutional Neural Networks (CNNs) were employed to extract high-level features from images as shown in Figure

Specifically, a pre-trained ResNet-50 model was utilized to capture meaningful representations of images with the help of equation as shown in Figure 2.

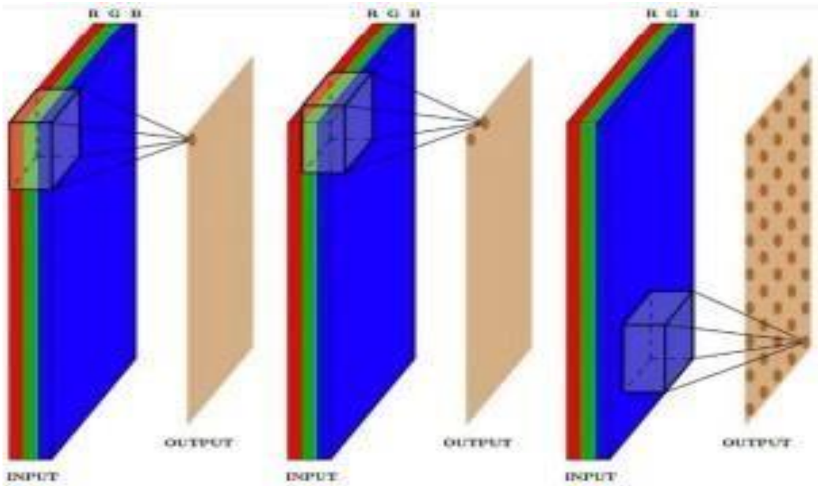


Figure 1: Example Classes

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

Figure 2: Example Classes

- **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.
- **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.

- **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 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:[5]

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 training accuracy achieved with this approach was 82.5%. 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 accuracy obtained was 82.26%. 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 optimise its performance.

3) Support Vector Machine (SVM):[67]

I. SVM Classifier:

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 classification accuracy of 84.5%. 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 functionand tuning its parameters to optimise the SVM's performance, as the choice of kernel can significantly impact the decision boundary.

II. Grid SearchCV for Hyperparameter Tuning:

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 accuracy of 86.69% with a polynomial kernel of degree 2. It helped in optimising the SVM's hyperparameters and improving its generalisation 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.

III. 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 capturecomplex decision boundaries.
Findings: Achieved the highest accuracy of 87.19% with the SVM classifier using the RBF kernel. The RBF kernel's flexibility allowed the model to capture intricate relationships in thedata and achieve superior performance.
Problem Faced: The main challenge was in tuning the SVM's hyperparameters, such as the regularisation 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 as shown in Figure 3 and preprocessed the images by splitting them into their respective red, green, and blue channels. The preprocessed data was then pickled for further processing.
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.

Figure 3: Example Classes

5 Anand Mishra-CNN Reference Code

6 Visualising and Understanding Convolutional Networks 7 Scikit-learn: Machine Learning in Python

3. **Dimensionality Reduction:** To reduce the dimensionality of the extracted features, we applied Linear Discriminant Analysis (LDA). LDA was utilised to transform the feature vectors into lower dimensional spaces while preserving relevant information.
4. **Classifier Training:** We trained various classifiers as shown in Table 1 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.

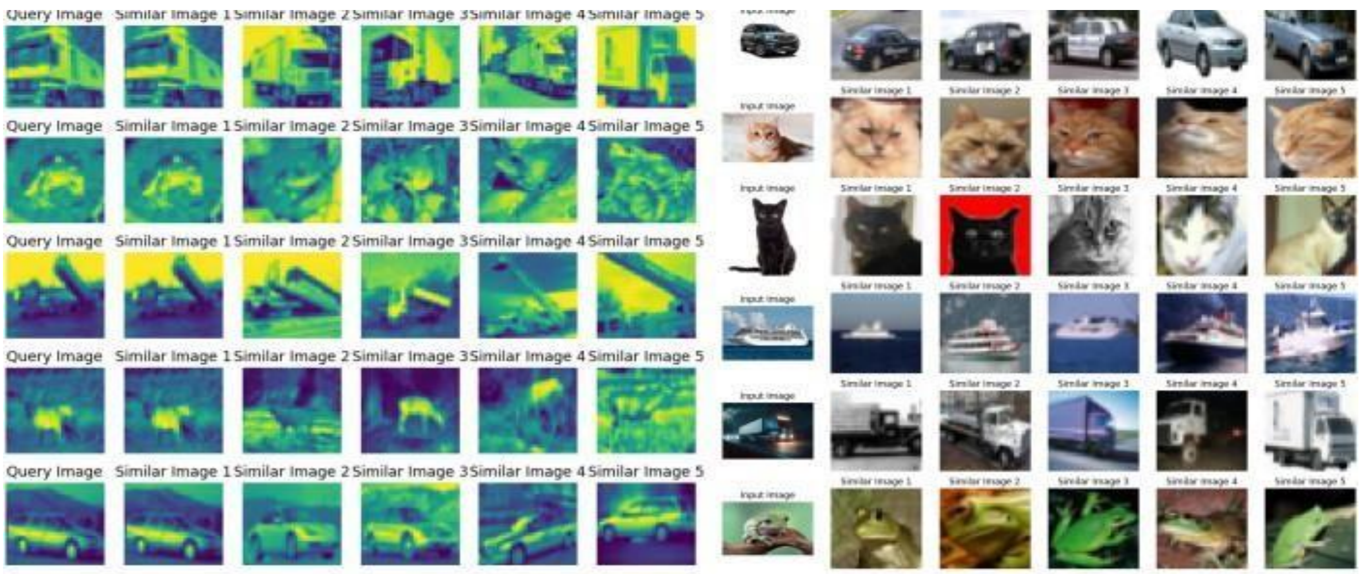
	PCA(Cross-validation)(%)	LDA(Cross-validation)(%)	Accuracy(%)
DecisionTree	61.8	89.92	84.12
KNN	83.1	93.25	88.85
Random-Forest	77.1	93.67	88.35
Naive Bayes	62.83	93.0	88.08
SVM1	86.5	93.9	88.89
SVM2	87.97	93.11	87.85
RBF	88.84	93.98	89.08

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 to a given input image.

2.2 Results

The results of the experiments as shown in figure 5 were evaluated based on the accuracy of the classifiers and the effectiveness of the image retrieval model.



(a) Feature-extracted train images (b) Internet Downloaded Images(test images)

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 cansubsequently be improved by utilising other batches as well.
2. **Feature Extraction:** We have firstly utilised 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 utilised 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 utilised Nearest neighbour for calculating accuracy, training accuracy would always 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 judged to be similar.

3. Summary

The project ”Image Retrieval” aimed to develop a robust system for efficiently locating similarimages given an input query.

The project ”Image Retrieval” aimed to develop a robust system for efficiently locating similarimages given an input query.

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 distil 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.

Conclusion

The research paper is successful at portraying the possibilities of the Convolutional Neural Networks (CNNs) as well as the other dimensionality reduction methods namely the Principal Component Analysis (PCA) and Linear Discriminant Analysis in developing a robust and efficient image retrieval system. So, if valiant attempts can jumpstart the problems the same dataset has, here the authors use CIFAR-10 dataset as a benchmark. In this case, the Support Vector Machine (SVM) with RBF kernel is highlighted as the most appropriate equipment in doing image classification and retrieval. The paper emphasises that the process of feature extraction using CNN such as ResNet-50 is important in obtaining high level and meaningful representation of images. Moreover, PCA and LDA have also assisted in reducing the computation load while retaining the crucial information because of Less propagated error. The system's performance, however, is improved with the use of SVM with the RBF kernelso that there is reasonable accuracy in correlation and retrieval of images and images with similar content. Ultimately, the picture retrieval system which has been developed validatesits real world application such that it successfully produces the top five images which aremost similar to the images presented as queries. The results obtained from this study formsthe basis for further development works in the domain of computer vision, focused on the enhancement of feature extraction techniques and seeking alternative approaches.

References:

1) Aigrain, P., Zhang, H., and Petkovic, D. 1996. Content-based representation and retrieval of visual media: A review of the state-of-the-art. Multimed. Tools Appl. 3, 3, 179--202.
2) M¨uller H, Michoux N, Bandon D, Geissbuhler A. A review of content-based image retrieval systems in medical applications – clinical benefits and future directions. International Journal of Medical Informatics 2004;(73):1– 23.
3) Sun Y, Zhang H, Zhang L, Li M. MyPhotos A System for Home Photo Management and Processing. In: ACM Multimedia Confernce. Juan-les-Pins, France; 2002. p. 81–82.
4) Markkula M, Sormunen E. Searching for Photos - Journalists' Practices in Pictorial IR. In: Electronic Workshopsin Computing – Challenge of Image Retrieval. Newcastle, UK; 1998. p. 1–13.
5) Armitage L H, Enser P G. Analysis of user need in image archives. Journal of Information Science 1997;23(4):287– 299.
6) Shirahatti N V, Barnard K. Evaluating Image Retrieval. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 05). vol. 1. San Diego, CA, USA: IEEE; 2005. p. 955–961.
7) Deselaers T, Weyand T, Keysers D, Macherey W, Ney H. FIRE in ImageCLEF 2005: Combining Contentbased Image Retrieval with Textual Information Retrieval. In: Accessing Multilingual

Information Repositories, 6th Workshop of the Cross-Language Evaluation Forum, CLEF 2005. vol. 4022 of Lecture Notes in Computer Science. Vienna, Austria; 2006. p. 652–661.

8) Yavlinski A, Pickering M J, Heesch D, R¨uger S. A comparative Study of Evidence Combination Strategies. In: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004). vol. 3. Montreal, Canada; 2004. p. 1040–1043.
9) Heesch D, R¨uger S. Performance boosting with three mouse clicks - Relevance feedback for CBIR. In: European Conference on Information Retrieval Research. No. 2633 in LNCS. Pisa, Italy: Springer Verlag; 2003. p. 363–376.
10) M¨uller H, M¨uller W, Marchand-Maillet S, Squire D M. Strategies for positive and negative Relevance Feedback in Image Retrieval. In: International Conference on Pattern Recognition. vol. 1 of Computer Vision and Image Analysis. Barcelona, Spain; 2000. p. 1043–1046.
11) Wang J Z, Li J, Wiederhold G. SIMPLIcity: SemanticsSensitive Integrated Matching for Picture LIbraries. IEEE Transactions on Pattern Analysis and Machine Intelligence 2001;23(9):947–963.
12) Lehmann T M, G¨uld M-O, Deselaers T, Keysers D, Schubert H, Spitzer K, et al. Automatic Categorization of Medical Images for Content-based Retrieval and Data Mining. Computerized Medical Imaging and Graphics 2005;29(2):143–155.
13) Deselaers T, Keysers D, Ney H. Features for Image Retrieval – A Quantitative Comparison. In: DAGM 2004, Pattern Recognition, 26th DAGM Symposium. vol. 3175 of Lecture Notes in Computer Science. T¨ubingen, Germany; 2004. p. 228–236.
14) Bloehdorn S, Petridis K, Saathoff C, Simou N, Tzouvaras V, Avrithis Y, et al. Semantic Annotation of Images and Videos for Multimedia Analysis. In: Europoean Semantic Web Conference (ESWC 05). Heraklian, Greece; 2005. .
15) Di Sciascio E, Donini F M, Mongiello M. Structured Knowledge Representation for Image Retrieval. Journal of Artificial Intelligence Research 2002;16:209–257.
16) Meghini C, Sebastiani F, Straccia U. A Model of Multimedia Information Retrieval. Journal of the ACM 2001;48(5):909–970.
17) MacArthur S D, Brodley C E, Shyu C-R. Relevance Feedback Decision Trees in Content-based Image Retrieval. In: Content-based Access of Image and Video Libraries. Hilton Head Island, SC, USA: IEEE; 2000. p. 68–72.
18) Smeulders A W M, Worring M, Santini S, Gupta A, Jain R. Content-Based Image Retrieval at the End of the Early Years. IEEE Transactions on Pattern Analysis and Machine Intelligence 2000;22(12):1349–1380.
19) Forsyth D A, Ponce J. Computer Vision: A Modern Approach. Prentice Hall; 2002. p. 599–619.
20) Rui Y, Huang T, Chang S. Image Retrieval: Current Techniques, Promising Directions and Open Issues. Journal of Visual Communication and Image Representation 1999;10(4):39–62.
21) Datta R, Li J, Wang J Z. Content-based Image Retrieval – Approaches and Trends of the New Age. In: ACM Intl. Workshop on Multimedia Information Retrieval, ACM Multimedia. Singapore; 2005.
22) Lew M S, Sebe N, Djeraba C, Jain R. Content-based Multimedia Information Retrieval: State of the Art and Challenges. ACM Transactions on Multimedia Computing, Communications and Applications 2006;2(1):1–19.
23) de Vries A P, Westerveld T. A comparison of continuous vs. discrete image models for probabilistic image and video retrieval. In: Proc. International Conference on Image Processing. Singapore; 2004. p. 2387–2390
24) Faloutsos C, Barber R, Flickner M, Hafner J, Niblack W, Petkovic D, et al. Efficient and Effective Querying by Image Content. Journal of Intelligent Information Systems 1994;3(3/4):231–262.

25) Pentland A, Picard R, Sclaroff S. Photobook: ContentBased Manipulation of Image Databases. International Journal of Computer Vision 1996;18(3):233–254.
26) Carson C, Belongie S, Greenspan H, Malik J. Blobworld: Image Segmentation Using Expectation-Maximization and its Application to Image Querying. IEEE Transactions on Pattern Analysis and Machine Intelligence 2002;24(8):1026–1038.
27) Siggelkow S, Schael M, Burkhardt H. SIMBA — Search IMages By Appearance. In: DAGM 2001, Pattern Recognition, 23rd DAGM Symposium. vol. 2191 of Lecture Notes in Computer Science. Munich, Germany: Springer Verlag; 2001. p. 9–17.
28) Clough P, M¨uller H, Sanderson M. The CLEF Cross Language Image Retrieval Track (ImageCLEF) 2004. In: Fifth Workshop of the Cross–Language Evaluation Forum (CLEF 2004). vol. 3491 of LNCS; 2005. p. 597–613.