Content Based Image Retrieval using SVM Model

Alen Rujis (alenrujis@iitb.ac.in)
Aaryan Bagani (203110002@iitb.ac.in)
Avdhesh Kumar Sharma (203110001@iitb.ac.in)
Indian Institute of Technology Bombay

May 2021

1. Abstract

In CBIR, given a query image, feature vector is extracted and compared with the feature vectors of all the images in the database and the images are ranked. This project aims to reduce these comparisons by predicting the class of the query image and compare its feature vector with features of the all images in the top two classes and rank them. Automatic feature selection is used for classification of images.

Keywords- Visual semantic relationship, content-based, database, feature vector.

2. Introduction

Technological advancements have increased the usage of cameras, smartphones, and Internet. The shared and stored multimedia data are growing, and retrieving similar images from an archive is a research problem. The fundamental need of any image retrieval model is to search and arrange the images that are in a visual semantic relationship with the query given by the user. Most of the search engines on the Internet retrieve the images on the basis of text-based approaches that require captions as input. The user submits a query by entering some text or keywords that are matched with the keywords that are placed in the archive. The output is generated on the basis of matching in keywords, and this process can retrieve the images that are not relevant. There are two disadvantages with this approach. The first is that a considerable level of human labour is required for manual annotation. The second is the annotation inaccuracy due to the subjectivity of human perception.

Content-based image retrieval (CBIR) is a framework that can overcome the above-mentioned problems as it is based on the visual analysis of the query image and retrieve similar images from large databases. "Content-based" means that the search analyses the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. The term "content" in this context might refer to colours, shapes, textures, or any other information that can be derived from the image itself. CBIR is desirable as it doesn't rely purely on metadata.

To provide a query image as an input is the main requirement of CBIR and it matches the visual contents of query image with the images that are placed in the archive, and closeness in the visual similarity in terms of image feature vector provides a base to and images with similar contents.

3. Methodology

Figure 1 shows the framework for classification and retrieval. This contains two processes: Training and Testing.

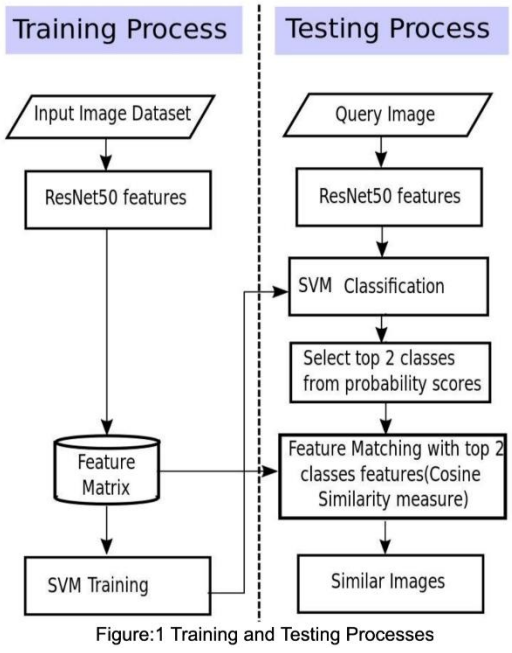

Figure 1 Training and Testing Processes

In the training process, features are extracted from the images in the database using ResNet50

pretrained model. ResNet50 is a convolutional neural network which is 50 layers deep and is trained on more than a million images from the ImageNet database. The pretrained network can classify images into 1000 object categories. As a result, the network has learned rich feature representations for a wide range of images.

Weights of ResNet50 pre-trained model are used as feature extractor and are frozen and are not updated during the training. These features are used to train a SVM 2 model. To implement transfer learning, the last predicting layer of the ResNet50 model is removed and replaced with fully connected layers to perform the classification with respect to the dataset chosen.

In the testing process, for a given query image, features are extracted from the ResNet50 pretrained model and these features are given to the SVM model to predict the class probabilities. The top two classes with highest probability are selected and the features of all the images from these two classes are compared with the features of the query image. Images of these two classes are ranked using cosine similarity measure and top ten images are presented as output.

The performed retrieval matches only feature vectors of images from two classes instead of matching all the images in the database. This reduces the time and increases precision for retrieval task since classification is done prior to retrieval. In this project Corel-5k dataset is used. There are 5,000 images from diverse contents such as sunset, beach, flower, building, car, horses, mountains, fish, food, door, etc. It contains 50 categories and each category contains 100 images.

4. Result and Conclusion

Classification and retrieval are evaluated on Corel-5k dataset. Table I shows accuracy of two classifiers on Corel-5k dataset.

Dataset	SVM
Coral-5k (50 classes)	98%

- Figure 2 and 3 shows confusion matrix for Corel-5k dataset using SVM From confusion matrix, we can infer that both SVM gave comparable results.

- Figure 4 and 5 shows Precision recall for Corel-5k dataset using SVM classifiers.

- Figure 6 gives Ranked Retrieval on test elephant image using SVM classifier in Corel-5k dataset.

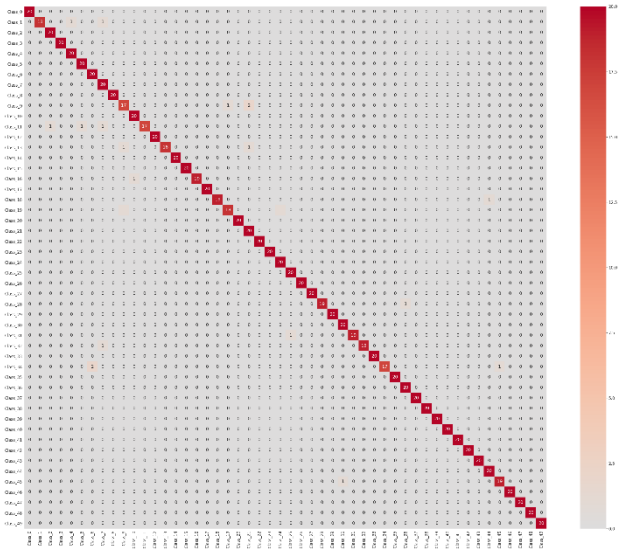

Figure 2Confusion Matrix for Corel-5k dataset using SVM classifier

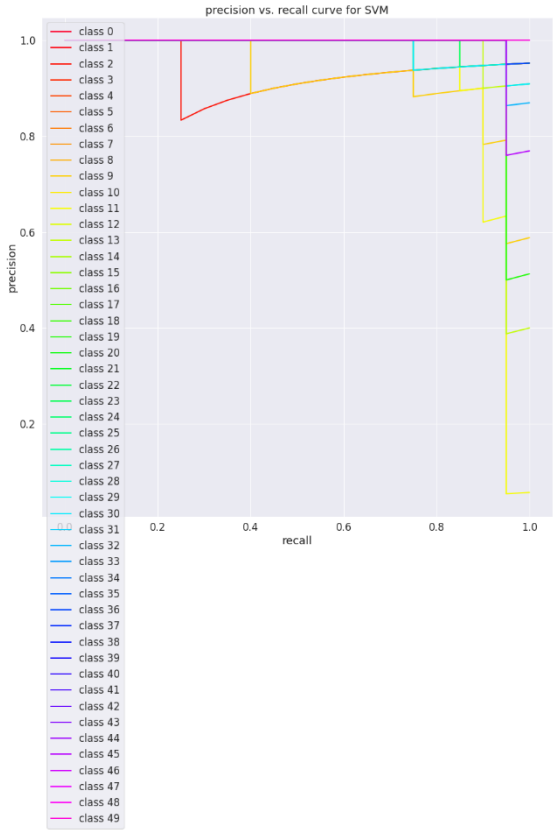

Figure 3 Precision recall graph for Corel-5k dataset using SVM classifier

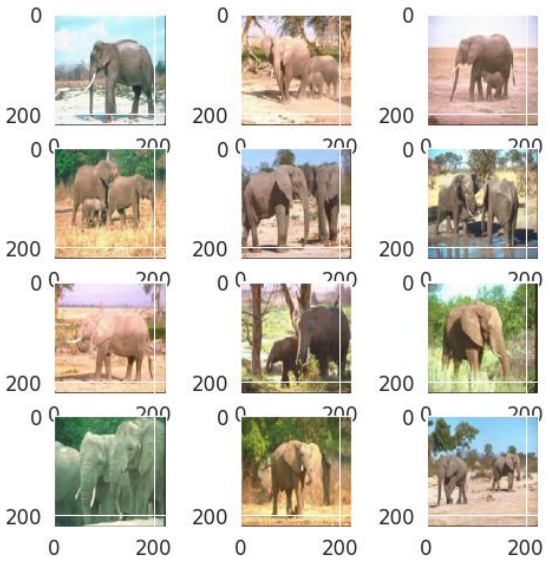

Figure 4 Image Retrieval on elephant image using SVM classifier

In this work we demonstrated that the number of matching are reduced by predicting the class of the query image and compare its feature vector with features of the all images in the top two classes. ResNet50 pretrained model is used for feature extraction. We have evaluated this framework on Corel-5k dataset and achieved good accuracy. The problem about this approach is that we first need labelled data to train the model. The labelling task can be costly and time consuming. Alternative approach for image retrieval task is to use an unsupervised deep learning algorithm.

References

1. Latif, Afshan, Aqsa Rasheed, Umer Sajid, Jameel Ahmed, Nouman Ali, Naeem Iqbal Ratyal, Bushra Zafar, Saadat Hanif Dar, Muhammad Sajid, and Tehmina Khalil. "Content-based image retrieval and feature extraction: a comprehensive review." Mathematical Problems in Engineering 2019 (2019).

2. Liu, Ying, Dengsheng Zhang, Guojun Lu, and Wei-Ying Ma. "A survey of content-based image retrieval with high-level semantics." Pattern recognition 40, no. 1 (2007): 262-282.

3. Guang-Hai Liu, Jing-Yu Yang, etc,. Content-based image retrieval using computational visual attention model, Pattern Recognition, 48(8) (2015) 2554-2566.

4. Babenko, Artem, Anton Slesarev, Alexandr Chigorin, and Victor Lempitsky. "Neural codes for image retrieval." In European conference on computer vision, pp. 584-599. Springer, Cham, 2014.