International Journal of Latest Trends in Engineering and Technology
Vol.(11)Issue(4), pp.035-039
DOI: http://dx.doi.org/10.21172/1.114.07
e-ISSN:2278-621X

ANALYSIS OF KNN AND SVM CLASSIFIER FOR IMAGE CLASSIFICATION

Mr. Vishal B. Padole[1], Dr. V. R. Mankar[2]

Abstract: Content based image retrieval refers to search for an image from the huge database to retrieve an image with the required shape, text or size. Content-based image retrieval are judged on the parameters such as speed and efficiency, and a modified approach based on a composite color image histogram processing is introduced. This paper introduces a brief study and analysis of Multiclass SVM classifier and comparison of it with KNN classifiers, in application to retrieve a correct output image from database in reference to query image. The proposed approach is time efficient and provides good results. By using a pool of such classifiers we can make a system comprises of multiple CBIR systems providing complementary outputs of each other when yhese outputs are fused together we can have a best match for query image .
Keywords: SVM, KNN, CBIR, classifiers

1. INTRODUCTION

With the increase in the multimedia and large avability of images and videos, information system containing image retrieval function is very important. Content based image retrieval refers to search for an image from the huge database to receive an image with the required shape, text or size. Though there is large research is going on CBIR still there is not a single universal approach which defines the exact match for query image. A major direction in CBIR research emphasis is matching similar objects based on shape color and texture.
As there is large avaibility of images in medicine, military, science CBIR is used to retrieve the proper image. Insisted of matching the whole content of an image one should match the selected features of an image with the features of an image present in the database. There are two ways to extract the features of an image that is at low level like local edges and the high level such as color histogram or we can use both features for analysis. In CBIR query image is compared with the already existing sample image from the available database. The images from the database having highest score of matching is displayed as a result for user.

2. FEATURE REPRESENTATION:

In order to do the process of image matching and retrieval in short period of time , image dimensionality reduction which only represents the interesting part of an image is known as feature extraction. Feature extraction is widely used in computer vision problems for object recognition and detection, CBIR, Feature detection and texture classification. Some commonly used methods for feature extraction are

- Speeded up robust features (SURF)
- Histogram of Oriented Gradient (HOG)
- Local Binary Patterns (LBP)
- Haar wavelets

2.1 Color Histogram.

If two images are belonging to the same renowned class then we can say that they are similar images. This proves that similar images belonging to renowned classes have close probabilities. In this paper, a multi-class classification model (Multiclass SVM) is used to find the probabilistic match between object and database, and comparison of Multiclass SVM with KNN classifier. Multiclass SVM enhances the process of finding close match between query image and database (class) image. This is the key advantage of this system.

2.2 Classifiers Used

There are number of classifiers available to generate and reduce feature vector which will helps during preparation of model for retrieval system. The selected classifiers are Back Propagation Neural Networks (BPNN), Radial Basis Function (RBF) Neural Networks, Support Vector Machine (SVM) and Self-Organizing Maps (SOM), K-nearest neighbor

[1] Research Scholar at Sant Gadge Baba Amtavati university, Department of Electronics and Telecommunication Engineering Amravati, India
[2] Principal, Government Polytechnic, Washim, India

(KNN) This paper focuses mainly on the application of Support vector machine and KNN classifier for image classification process. This section presents details regarding Multiclass SVM and KNN.

2.3 K Nearest Neighbours algorithm

KNN algorithm simply uses the stored information of all existing cases and yields new results based on similarity measures. K-nearest neighbor is falls under the category of supervised learning algorithm in which the result of new image query is classified on the basis of maximum of k-nearest neighbor category. The ultimate aim of KNN algorithm is to categories a new image based on the training examples. With respect to the given query this algorithm finds K number of samples nearest to the input image.[5]. Classification of the images is done by taking into consideration the maximum votes while classifying the k number of objects with the facility of breaking ties randomly. This algorithm use nearby samples categorization as the probable guess value of the new input image[10].

K Nearest neighbours (kNN) is to use database in which the data points are separated into several separate classes to predict the classification of new points.

Input is consisting of k nearest perception examples in the feature space.

Output is class membership.

Maximum number of votes by neighbours will classify the object.

2.4 Multiclass SVM

While categorizing a particular point or object, there are N different classes to which the given object or point can be placed. Therefore it is required to construct a function which can effectively predict the class to which the given point or objects belongs.

Support Vector Machines are primarily designed for binary classification that is for only two classes possibility but in most of the cases single data point can belongs to a several classes to solve this constraint multiclass SVM is the solution.

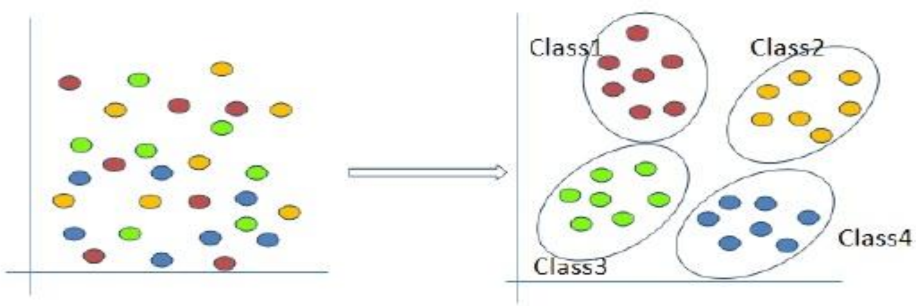


Figure: 1.1 An example illustrating a multiclass SVM

How to convert SVM into multiclass SVM is still a enlightening research issue. A typical method to construct multiclass classifier is by combining several binary classifiers[4].

Conventional Methods for Multiclass Problems are as follows,

One-vs-rest approaches

Pairwise approaches

To overcome the drawbacks of these conventional systems recent development for Multiclass Problems are as follows,

Simultaneous Classification

Various loss functions

While designing Multiclass SVM one has to consider following constraints

(i) Decompose the multiclass classification problem into multiple binary classification problems.

(ii) Use the majority voting principle (a combined decision from the committee) to predict the label. Some simple and effective approaches are,

One-vs-rest (one-vs-all) approaches

Pairwise (one-vs-one, all-vs-all) approaches

One-vs-rest (one-vs-all) approaches

Amongs all the classifiers built for real valued binary classifiers, is to train N number of binary classifiers . Each classifier is trained such that it will place the example in single class out of all remaining classes. When one has to classify new querry all classifiers (N) are run and classifier whose output gains maximum votes will be chosen. This scheme is referred as One Vs All approach (OVA)[9].

Out of all the multiclass Classifiers this one is the simplest multiclass classifier; commonly used in SVMs; also known as the one-vs-all (OVA) approach

Solve K different binary problems: classify class k" versus the rest classes" for

k = 1; _ _ _ ;K.

Assign a test sample to the class giving the largest fk (x) (most positive) value, where fk (x) is the solution from the kth problem

2.5 Properties:

Implementation is very simple. perform well in practice Not optimal (asymptotically): the decision rule is not Fisher consistent if there is no dominating class (i.e. arg max pk (x) < 1

Pairwise (one-vs-one, all-vs-all) approaches

The major problem associated with the OVA classification approach is multiclass classification. This scheme converts multiclass problem into series of two class problem, each pair of classes is assigned with a individual problem. This scheme is more efficient than OVA classification. The basic idea behind this type of classification is to transform n-class problem into n(n-1)/2 binary problems. This classification scheme is elaborated as follows.

This type of classification is also known as all-vs-all (AVA) approach

(i) Solve (L/2) different binary problems: classify class "l" versus m class "m" for all m ≠ l. Each classifier is called hij .

(ii) For prediction at a point, each classifer is queried once and issues a vote. The class with the maximum number of (weighted) votes is the winner.

2.6 Properties:

As this process deals with minute binary problems it is very efficient.

If L is big, there are too many problems to solve. If L = 10, we need to train 45 binary classifiers.

Its implementation is very simple.

It performs comparatively faster in practice.

3. RESULT AND ANALYSIS:

In this chapter comparative performance of KNN classifier and Multiclass SVM is analyzed. The dataset consist of three categories Beach, Bus and Dinosaurs all having Size of 256*256 pixel. For training purpose we have considered 100 images of each category and for testing we have considered 50 images of each category. Category and image description is shown in Table 3.1

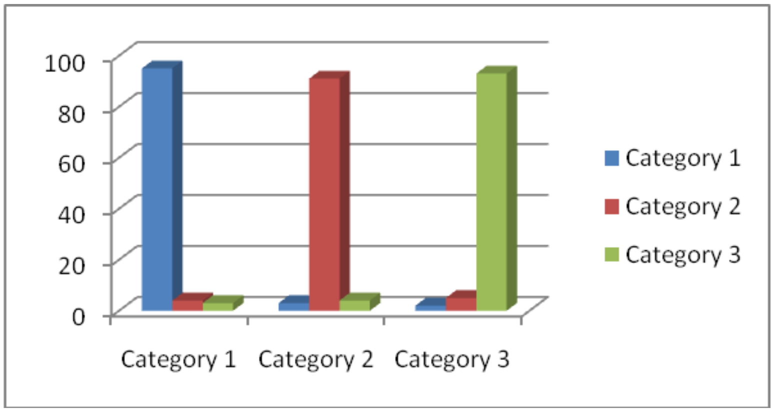
Table 3.1 Database Image Category description Table

Sr. no	Category	Image Description	No. of Images for Training	No. of images for Testing
1	Category-1	Beach	100	50
2	Category-2	Bus	100	50
3	Category-3	Dinosaurs	100	50

The above table 3.1, contains three categories and the image description for each category .It also shows the number of images considered for training and testing.

Table.3.2 Performance matrix for Multiclass SVM classifier for Training images (100)

		Classification Result		
		Category 1	Category 2	Category 3
Input Image	Category 1	95	3	2
	Category 2	4	91	5
	Category 3	3	4	93

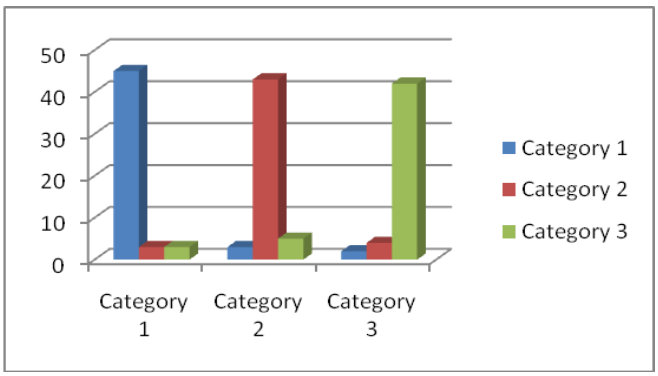


The above table 3.2 Depicts the performance of SVM classifier against 100 images belonging to each category as a training images. Beach, Bus and Dinosaurs belong to category 1, category 2 and category 3 respectively. The table 3.2 shows the number of images classified correctly for each category.

From table 3.2 it is observed that out of 100 images of Beach the Multiclass SVM has correctly classified 95 images in category 1 and wrongly classified 3 images in category 2 and 2 images in category 3. For category 2, the Multiclass SVM has classified 91 images correctly in category 2 and wrongly classified 4 images in category 1 and 5 images in category 3. Similarly for category 3, 93 images were correctly classified and 3 images were classified in category 1 and 4 images in category 2.This is overall performance of Multiclass SVM Classifier.

Table.3.3 Performance matrix for Multiclass SVM classifier for Testing Images (50)

		Classification Result		
		Category 1	Category 2	Category 3
Input Image	Category 1	45	3	2
	Category 2	3	43	4
	Category 3	3	5	42



The above table 3.3 Depicts the performance of SVM classifier against 50 images belonging to each category as a testing images.

Table.3.4 Performance matrix for KNN classifier for Testing Images (50)
For K=1

		Classification Result		
		Category 1	Category 2	Category 3
Input Image	Category 1	40	4	6
	Category 2	5	38	7
	Category 3	9	4	37

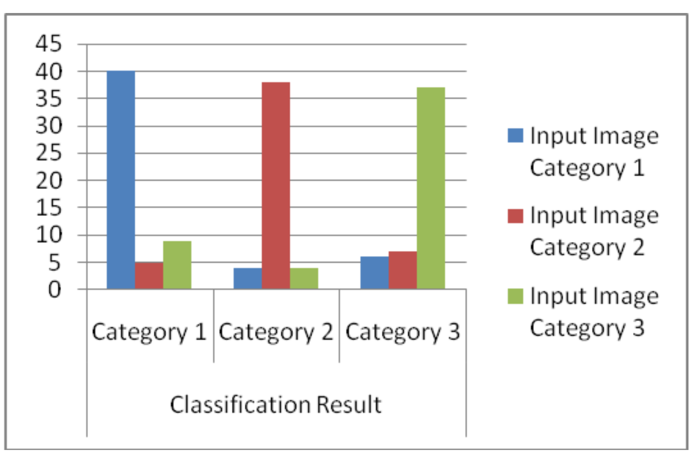
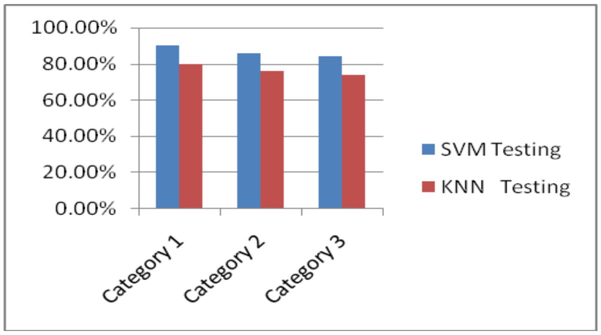


Table.3.5 Accuracy comparison of ANN and KNN classifier for Testing Images For 50 images

	ANN	KNN
	Testing	Testing
Category 1	90.00%	80.00%
Category 2	86.00%	76.00%
Category 3	84.00%	74.00%



The above table 3.5 shows the comparative study of SVM classifier and K-nearest neighbor classifier for 50 testing of images. From the analysis given in table 8.3 it can be concluded that categorization accuracy of SVM classifier is better as compared to KNN classifiers.

4. CONCLUSION:

This paper contributes the application of Multiclass SVM in comparison with the KNN classifier for detecting the correct image from the database. Multiclass SVM provides output which is having better matching percentage with the query image as compared with KNN classifiers.

However by taking into consideration the complex features present in the query image and computational limitation of feature extraction algorithms these methods cannot provides the effective results. In order to overcome the drawbacks and limitations of individual systems, individual systems need to be device and combined in such a way the fused system will be more robust than the individual systems.

5. REFERENCES

[1] Sandeep Singh1, Er.Rachna Rajput “Content Based Image Retrieval using SVM, NN and KNN IJARCCE Vol. 4, Issue 6, June

[2] Dharani, I. Laurence Aroquiaraj A Survey on Content Based Image Retrieval Proceedings of the 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, February 21-22

[3] Aster Wardhani and Tod Thomson Content Based Image Retrieval Using Category-Based Indexing2004 IEEE International Conference on Multimedia and Expo (ICME)

[4] Jitendra kumar, V. S. Chakravarthy “Designing an optimal classifier Ensemble for online character Recognition”

[5] R.Priya, Dr. Vasantha Kalyani David Optimized Content based Image Retrieval System based on Multiple Feature Fusion AlgorithmInternational Journal of Computer Applications (0975 – 8887) Volume 31– No.8, October 2011

[6] Chih-Wei Hsu and Chih-Jen Lin A Comparison of Methods for Multi-class Support Vector Machines (cjlin@csie.ntu.edu.tw)

[7] Deok-Hwan Kim et. Al Support Vector Machine Learning for Region-Based Image Retrieval with Relevance Feedback ETRI Journal, Volume 29, Number 5, October 2007

[8] L. Zhang, F. Lin, and B. Zhang, “Support Vector Machine Learning for Image Retrieval,” Proc. IEEE ICIP, 2001

[9] Dr.A.Kannan Content Based Image Retrieval Using Nearest Neighbour and Hybrid KNNSVM Methods to Diagnose MR Images International Journal of Innovative Research in Science, Engineering and Technology Volume 3, Special Issue 3, March 2014

[10] T. Dharani et al. Content Based Image Retrieval System using Feature Classification with Modified KNN Algorithm

[11] Khoa Duc Tran Content-Based Retrieval Using a Multi-Objective Genetic Algorithm 0-7803-8865-8/05/$20.00 ©2005 IEEE.

[12] Ryan Rifkin et al. In Defense of One-Vs-All Classifcation Journal of Machine Learning Research 5 (2004) 101-141

[13] Sang-Hyeun Park and Johannes F¨urnkranz TU Darmstadt, Knowledge Engineering Group, D-64289 Darmstadt, Germany Efficient Pairwise Classification