MAT JOURNALS

Journal of Communication Engineering and its Innovations
Volume 5 Issue 1

Content Based Image Retrieval Using Edge Based Feature Extraction In Deep Learning Algorithm

***Siva Krishna Sajarao*[1,*], *M Pradeep*[2]**
[1]*PG Student, Department of Electronics and Communication Engineering, A1 Global Institute of Engineering & Technology, Markapur, Andhra Pradesh, India*
[2]*Assistant Professor, Department of Electronics and Communication Engineering, A1 Global Institute of Engineering & Technology, Markapur, Andhra Pradesh, India*
***Email:** koteshproject@gmail.com*
DOI: http://doi.org/10.5281/zenodo.2606766

Abstract

Content-based image retrieval (CBIR) using edge-based feature extraction in deep learning algorithm has been implemented in this paper. CBIR is widely used in securities, entertainment, medicine and many military applications. the retrieval system process is based on the decomposition of the input image and different levels of filtering analysis such as edge, enhancement and different scaling factors of the training images. Deep learning analysis is majorly concentrated on the different convolution layers. The convolutional neural network layer has been implemented on this process. The convolutional neural network is having one of the layers is convolutional. It is used to extract the feature based on the different kernels to form the feature extraction map. This paper is implemented on the COREL database. As compared to existing methods this paper gives better results in terms of recognition efficiency.

Keywords: *convolution neural network and distance measure techniqes and recognition performance measures*

INTRODUCTION

Picture recovery is an examination zone of Information Retrieval [1] of extraordinary scientific enthusiasm since 1970s. Prior investigations incorporate manual comment of pictures utilizing catchphrases and seeking by content [2]. Content Based Image Retrieval (CBIR), [3], has been proposed in 1990s, so as to defeat the difficulties of content based picture recovery, getting from the manual explanation of pictures, that depends on the abstract human recognition, and the time and work necessities of comment. CBIR alludes to the way toward acquiring pictures that are significant to a question picture from an extensive accumulation dependent on their visual substance [4]. Given the component portrayals of the pictures to be sought and the question picture, the yield of the CBIR system incorporates an inquiry in the element space, so as to recover a positioned set of pictures as far as likeness (for example cosine likeness) to the question portrayal. A key issue related with CBIR is to extricate important data from crude information so as to take out the alleged semantic-hole [5]. The semantic-hole alludes to the distinction between the low-level portrayals of pictures and their more elevated amount ideas. While prior works center around crude highlights that depict the picture substance, for example, shading, surface, and shape, various later works have been explained on the heading of finding semantically more extravagant picture portrayals. Among the best are those that utilization the Fisher Vector descriptors [6], Vector of Locally Aggregated Descriptors (VLAD) [7,8] or com-bine sack of-words models [9] with nearby descriptors, for example, Scale-Invariant Feature Transform (SIFT) [10]. A few late investigations present Deep Learning calculations [11] against the shallow previously mentioned ways to deal with a wide scope of PC vision errands, including picture recovery [12– 15]. The fundamental purposes for their prosperity are the accessibility of extensive commented on datasets, and the GPUs computational power and af-foldability. Profound Convolutional Neural Networks (CNN), [16,17], are viewed as the more efficient Deep Learning design for visual data examination. CNNs contain various convolutional and subsampling layers with non-direct neural initiations, trailed by completely associated layers. That is, the information picture is acquainted with the neural system as a three-dimensional tensor with measurements (i.e., width and tallness) equivalent to the components of the picture and profundity equivalent to the quantity of shading channels (generally three in RGB pictures). Three dimensional filters are found out and ap-employed in each layer where convolution is performed and the yield is passed to the neurons of the following layer for non-straight change utilizing proper initiation capacities. After various convolution layers and sub sampling the structure of the profound engineering changes to completely associated layers and single dimensional signs. These initiations are typically utilized as profound portrayals for classification, grouping or recovery.

Over the most recent couple of years, profound CNNs have been set up as a standout amongst the most encouraging roads of research in the PC vision region because of their extraordinary execution in a progression of vision acknowledgment assignments, for example, picture classification [18, 19], face acknowledgment [20], digit acknowledgment [21, 22], present estimation [23], and item and person on foot discovery [24, 25]. It has likewise been shown that highlights separated from the enactment of a CNN prepared in a completely managed manner on a substantial, fixed set of item acknowledgment errands can be re-purposed to novel conventional acknowledgment undertakings, [26]. Persuaded by these outcomes, profound CNNs presented in the striking examination territory of CBIR. The essential methodology of applying profound CNNs in the recovery space is to extricate the element portrayal from a pretrained show by encouraging pictures in the information layer of the model and taking actuation esteems drawn either from the completely associated layers [27– 30] which are intended to catch abnormal state semantic data, or from the convolutional layers misusing the spatial data of these layers, utilizing either entirety pooling procedures [31,32] or max-pooling [33]. Momentum inquire about likewise incorporates show retraining approaches, which are increasingly pertinent to our work, while different investigations center around the mix of the CNN descriptors with customary descriptors like the VLAD portrayal. The current related works are talked about in the accompanying area.

Discriminant and saving projection techniques are clarified in area II, a propelled approach for the face recovery framework has been proposed talked about in segment III, exploratory outcomes and discourse are portrayed based diverse databases in segment IV, and the end has been examined in segment V

CONVOLUTIONAL NEURAL NETWORK

A significant CNN is retrained with comparability learning target work, considering triplets of material and irrelevant models procured from the totally related layers of the pretrained illustrate, in [29]. A related system has moreover been proposed in the face affirmation errand which, using a triplet-based setback work, achieves stand out execution, [42], while an appropriate idea starting late adequately introduced in the cross-secluded recuperation zone [43]. These approaches are using triplet test acknowledging which is difficult to be completed in extensive scale, and ordinarily powerful learning is used in order to pick essential triplets that can without a doubt add to learning [42].

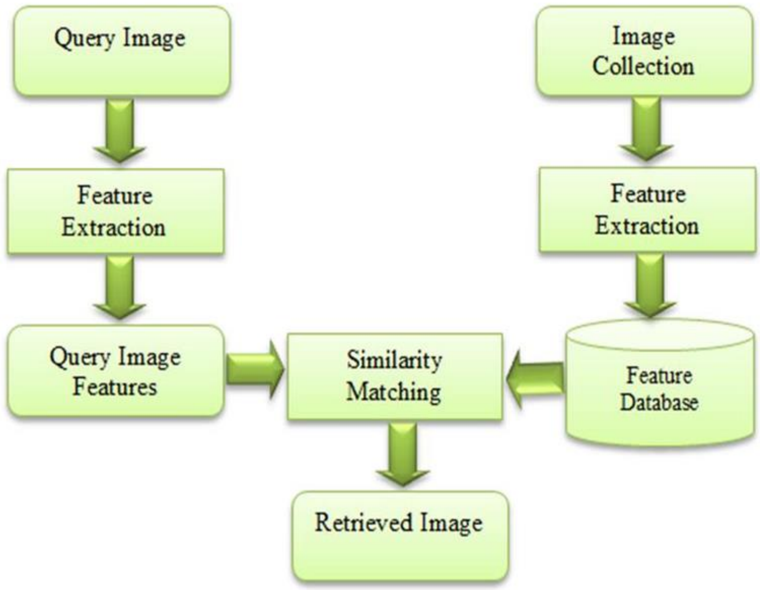

Figure 1: *Basic block diagram of content-based image retrieval*

In our approach we widen these theories by considering diverse imperative and different arbitrary precedents in the planning framework for every readiness test. Moreover, we help the readiness speed by defining depiction centers for the arrangement tests and backslide on the hid sanctum layers, as opposed to defining logically complex hardship works that need three precedents for every planning venture. That is, our procedure uses single precedent getting ready considering speedy and flowed learning. Moreover, the proposed methodology is in like manner prepared to manhandle the geometric structure of the data using unsupervised adapting, similarly as to abuse the customer's analysis using congruity input. Finally, since our inside is to make low-dimensional descriptors, which improve both the recuperation time and the memory necessities, we apply our system on convolutional layers using max-pooling procedures, rather than the past methodologies which utilize the totally related layers.

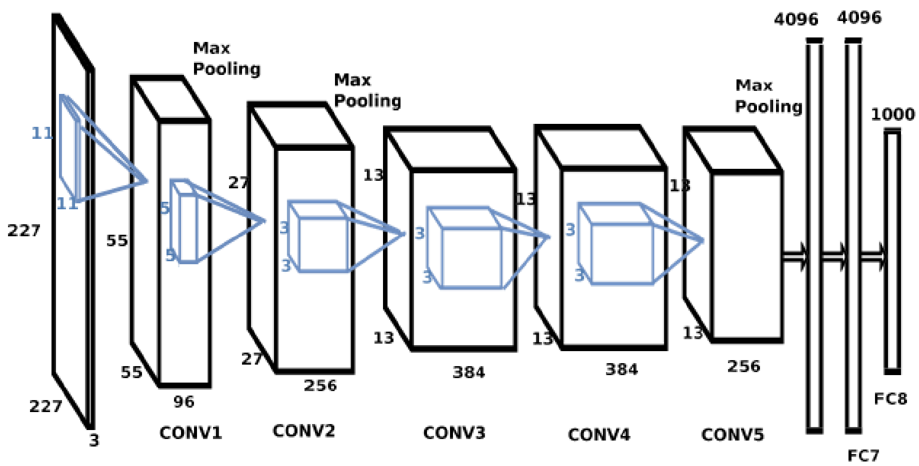

Figure 2: *Basic architecture of convolutional neural networks*

In the FU approach, we mean to enhance the essential recovery assumption that the applicable picture portrayals are nearer to the specific question portrayal in the component space. The proportion behind this methodology is attached to the bunch speculation which expresses that records in a similar group are probably going to fulfill a similar data need [49]. That is, we retrain the pretrained CNN demonstrate on the given dataset, going for amplifying the cosine similarity between each picture portrayal and its n closest representations, as far as cosine remove.

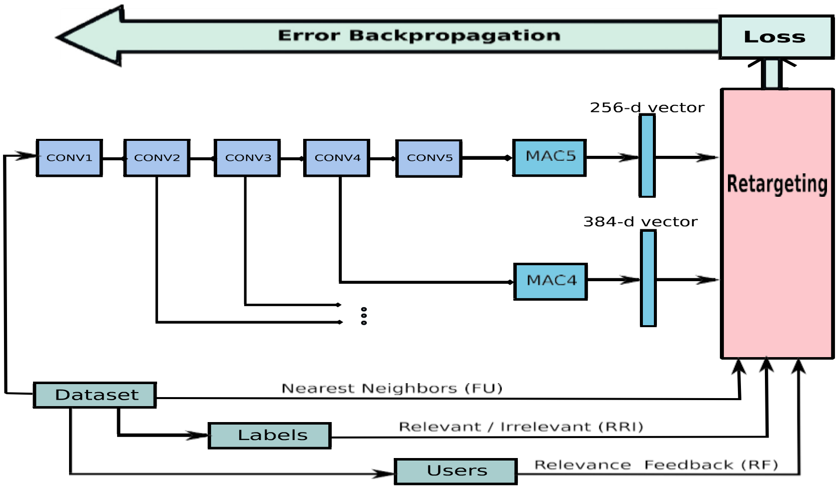

***Figure 3:** Detailed architecture of neural network*

PROPOSED METHOD

Content based image retrieval methodology has been implemented as per following steps.

1. Choose a color image $X_{m, n, p}$ and it can be converted into gray level image as based on below criteria to select the shape features. Where m, n and p are number of rows, columns and color(R, G, and B) dimensions.
   X = 0.3*R+ 0.591*G+0.1*B.
2. To exact the features based on different moments of shape representation.
   $$M_{i,j} = \sum_{m,n}^{\Re} (X - X_c)^i (Y - Y_c)^j$$
   Where X_c and Y_c are central moments
3. To calculate the normalized function for individual moments of the input image as can be applied for following formula.
4. Apply the different moments on given input database for extracting local features.
5. First apply the convolutional neural network on the database.
6. To extract the better coefficient, apply the moments coefficients on the contourlet transform sub bands.
7. Apply the feature extraction process on each 32-sub band and merge the all local features.
8. To test the query image, apply the same procedure for test image and extract the features.
9. Calculate the precision and recall method for getting reconditioned images.
10. Calculate the recognition rate and computational rates for all database images.

EXPERIMENATAL RESULTS

Experiment using YALE database

A recovered picture is viewed as a match if and just in the event that it is in indistinguishable class from the inquiry. This presumption is sensible since the 10 classes were picked with the goal that each delineates an unmistakable semantic subject. Each picture in the sub-database was tried as an inquiry and the recovery positions of all pictures were recorded.

Three measurements were figured for each question: 1. the accuracy inside the initial 100 recovered pictures (p), 2. the mean position of all the coordinated pictures (r), and 3. The standard deviation of the positions of coordinated pictures (σ). The review inside the initial 100 recovered pictures is indistinguishable to the accuracy in this uncommon case.

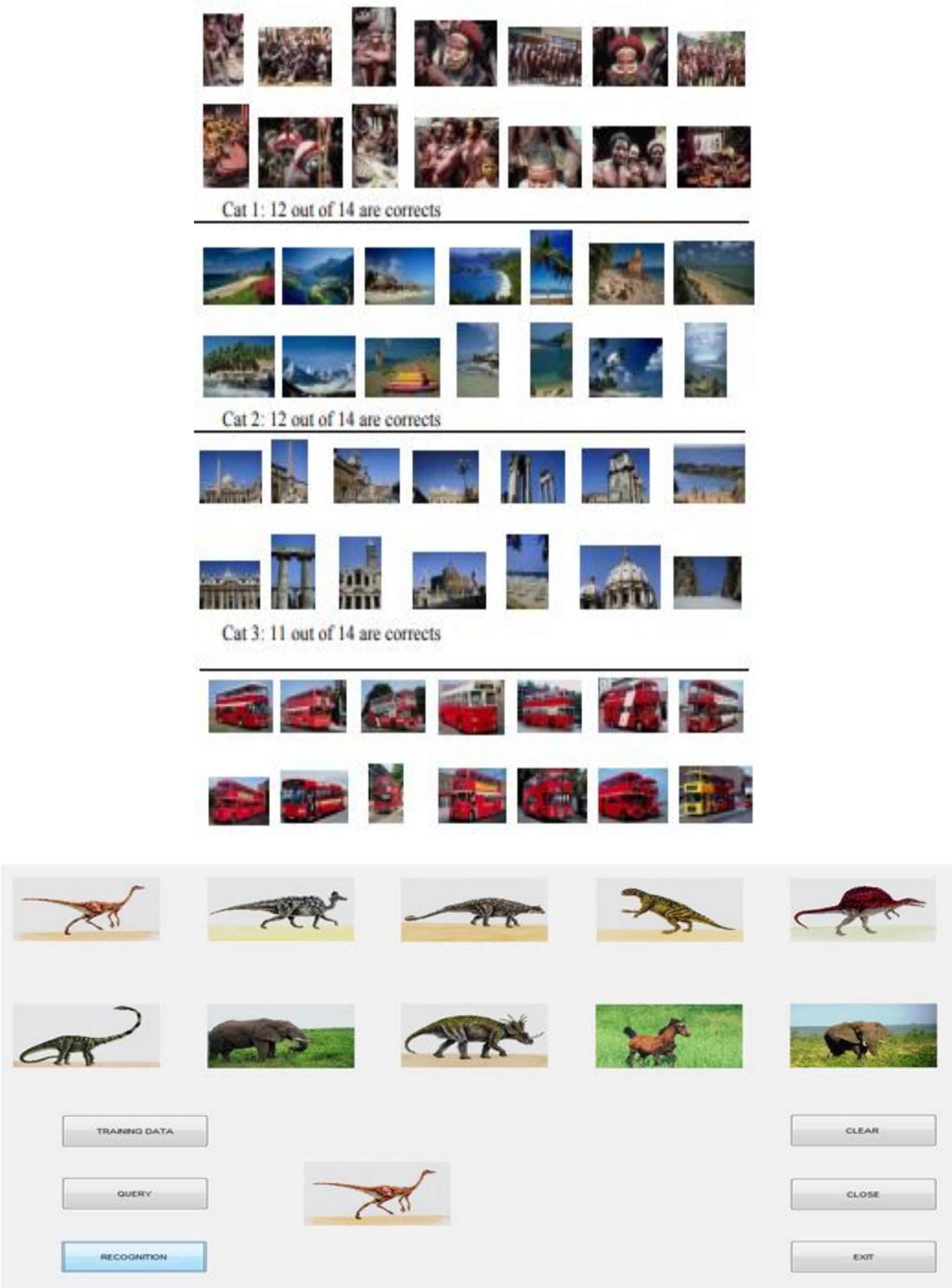

Figure 4: *Retrieval system for COREL Database*

Figure 5: *COREL Database*

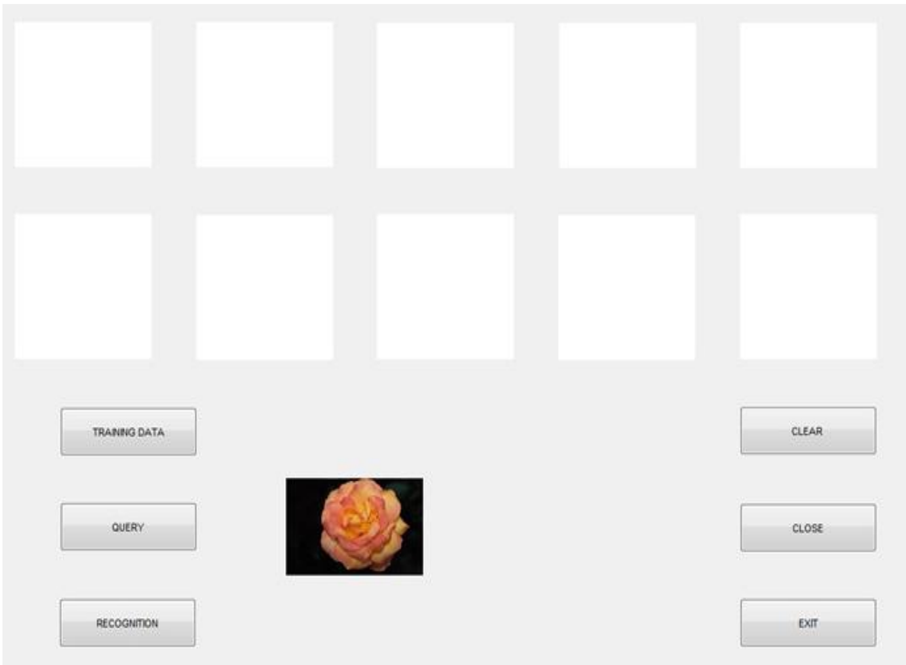

Figure 6: *Test image of the retrieval images*

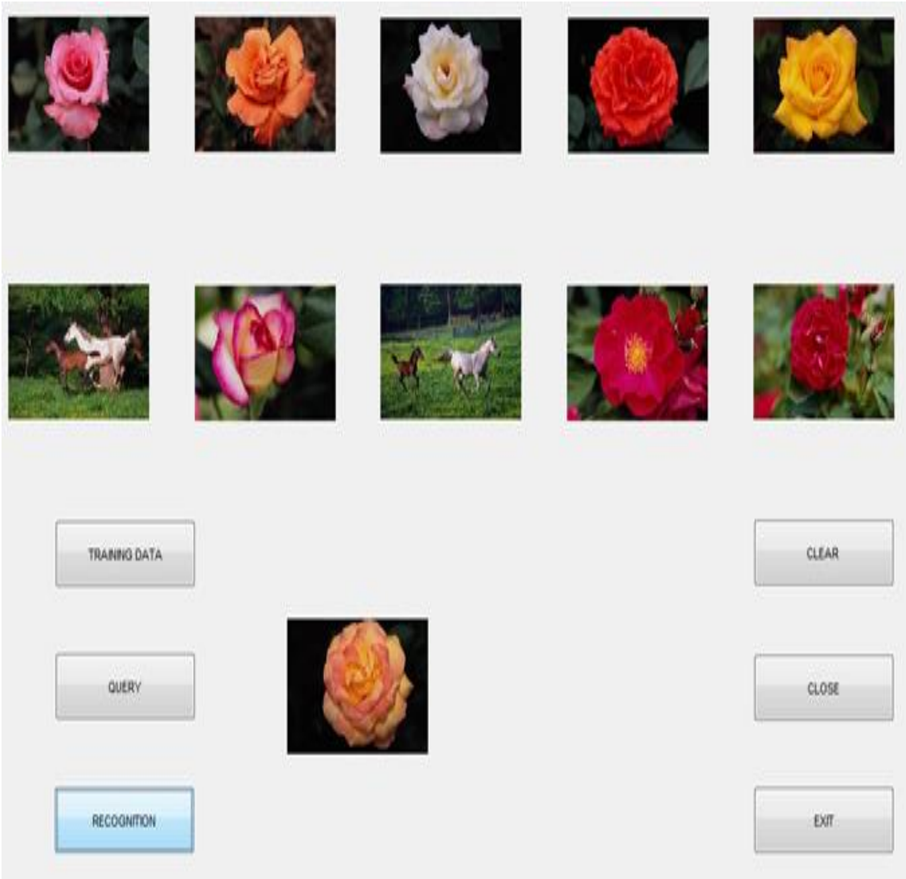

Figure 7: *Retrieval images*

Table 1: *Tabular form of content based image retrieval*

	Recognized images				
Methods	***1***	***3***	***5***	***7***	***10***
K-NN Algorithm	*100*	*70*	*65*	*63*	*62*
Neural network	*100*	*75*	*72*	*69*	*66*
LDA	*100*	*82*	*81*	*79*	*75*
Convolutional Neural Network	***100***	***98.5***	***97***	***88***	***82***

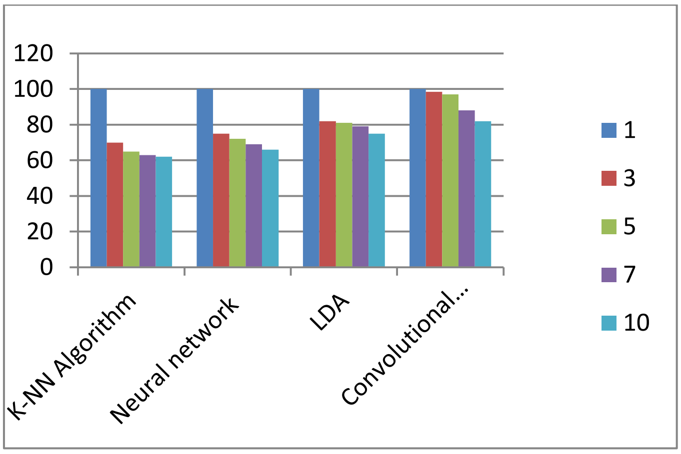

Figure 8: *Comparison Graph for Retrieval System*

CONCLUSION

Content based image retrieval using edge-based feature extraction in deep learning algorithms has been successfully implemented in this paper. Different neural networks are implemented to take the feature extraction for retrieval system. The retrieval system for top matching images 1,3,5,7, and 10 images are 100%, 98.5%, 97%, 88%, 82 respectively. This method has given good computational time 0.4sec, 0.5sec, 0.55sec, 0.78sec, 0.45sec for proposed method.

ACKNOWLEDGMENT

Authors would like to thank A1 Global Institute of Engineering & Technology, Markapur, management for providing facilities to finish this work.

REFERENCES

1. J. T. Favata, G. Srikantan, and S. N. Srihari, "Handprinted character/digit recognition using a multiple feature/resolution philosophy," in *IWFHR-IV*, pp. 57–66, Dec 1994.
2. A. Marzal and E. Vidal, "Computation of normalized edit dis- tance and applications," *IEEE Transactions on Pattern Analy- sis and Machine Intelligence*, vol. 15, pp. 926–932, Sep 1993.
3. P. A. V. Hall and G. R. Dowling, "Approximate string match- ing," *ACM Computing Surveys*, vol. 12, pp. 381–402, Dec 1980.
4. CEDAR, "Intelligent character recognition." http://www.cedar.buffalo.edu/ICR, 1997.
5. R. A. Wagner and M. J. Fischer, "The string-to-string correc- tion problem," *Journal of the ACM*, vol. 6, pp. 168–178, Jan 1974.
6. G. M. Landau and U. Vishkin, "Introducing efficient paral- lelism into approximate string matching and a new serial al- gorithm," in *Proceedings of the 18th Annual ACM Symposium on Theory of Computing*, pp. 220–230, 1986.
7. R. R. Bradford and R. B. Bradford, *Introduction to Handwrit- ing Examination and Identification*. Nelson-Hall Publishers: Chicago, 1992.
8. G. Kim and V. Govindaraju, "Efflcient chain code based image manipulation for handwritten word recognition," in *the SPIE sym. on electronic imaging science and technology (Document Recognition III)*, vol. 2660, pp. 262–272, Feb 1996.
9. CEDAR, "http://www.cedar.buffalo.edu/databases/cdrom1/."
10. J. J. Hull, "A database for handwriten text recognition re- search," *Transactions on Pattern Analysis and Machine Intel- ligence*, pp. 550–554, May 1994.
11. R. O. Duda and P. E. Hart, *Pattern classification and scene analysis*. New York, Wiley, 1st ed., 1973.