OPTICAL CHARACTER RECOGNITION USING IMPROVED ROOTSIFT FEATURES

Sandhya Balakrishnan P K[1], Dr.L.Pavithira[2]

[1]*Research Scholar, Department of Computer Science,*
CMS College of Science and Commerce, Coimbatore.
[2]*Associate Professor, Department of Computer Science,*
CMS College of Science and Commerce, Coimbatore.

Abstract:

Nowadays, the biggest challenge in the field of image processing is to recognize documents both in printed and handwritten format. Character recognition is one of the most widely used biometric traits for authentication of person as well as document. Optical Character Recognition is a process that can convert text, present in digital image, to editable text. It allows a machine to recognize characters through optical mechanisms. A Neural network is designed to model the way in which the brain performs a particular task or function of interest. Each image character is comprised of 30×20 pixels. We have applied improved RootSIFT feature extraction technique for calculating the features. Features extracted from characters are directions of pixels with respect to their neighboring pixels. These inputs are given to a back propagation neural network with hidden layer and output layer. Finally, we evaluated the proposed approach using RETAS OCR evaluation dataset and obtained the better recognition accuracy which demonstrate the effectiveness of the proposed approach.

Keywords: *OCR; Character Recognition; RootSIFT; BPN Network; English Alphanumeric Characters; Image Acquisition*.

I. INTRODUCTION:

Optical Character Recognition (OCR) is the process of converting the scanned documents, pdf into computer readable and editable format, i.e., ASCII Characters. OCR is the process of classification of optical patterns contained in a digital image corresponding to alphanumeric or other characters [1]. In spite of the title being the Optical Character Recognition, the work of digital image processing also has to be done. So it is essential to have the background knowledge for both image processing as well as for character recognition.

OCR is the process of reading characters in images. It consists of two main tasks[2]: segmentation of the individual characters, and the classification of the segmental characters. Concretely, an OCR system typically consists of the following processing steps: (1) gray-level scanning at an appropriate resolution, typically 300-1000 dots per inch. (2) preprocessing: (a) binarization (two-level thresholding), using a global or a locally adaptive method; (b) segmentation to isolate individual characters; (3) feature extraction; (4) recognition using one or more classifiers; (5) contextual verification or postprocessing.

Character segmentation is a main preprocessing step. In character segmentation, many methods have been developed, e.g. thresholding, extraction of connected components, subpixel-precise thresholding[3], and morphological operations[4], etc.. Selection of a feature extraction method is probably the single most important factor in achieving high recognition performance. Nowadays, a large number of feature extraction methods have been reported in the literatures[5][6].

Local invariant feature based methods extracts the stable local invariant features for matching. Compared with the aforementioned methods, these methods are invariant to rotation, scale and translation, which makes them appropriate for contact-free vein recognition. Pierre *et al.* [7] used SIFT [8] for feature extraction and matching after preprocessing and binarization on digital images. However, because the binarization will result in the loss of useful feature information, some researchers [9]–[10] have directly extracted SIFT and SURF [11] features from the digital image after noise removal and illumination normalization, thereby improving the correct matching rate.

In this paper, modified RootSIFT algorithm is implemented for feature extraction and back propagation neural network is implemented. The rest of the paper is organized as follows. Section II provides an overview of the recent state-of-the-art feature detection and description algorithms proposed in literature. In Section III, the modified RootSIFT algorithm is studied along with back propagation Neural network classification algorithm. Finally, Section IV concludes the paper.

II. RELATED WORKS

The traditional approach for image retrieval uses bag of words (BoW) model. The BoW approach proposed in [16] was motivated by techniques used in text retrieval. The BoW model uses the powerful “Scale Invariant Feature Transform” (SIFT) descriptors [17] and represents an image by counting the number of descriptors that is assigned to each of the different visual words [18].

Perronnin et al. [19] proposed the use of fisher vectors for representing the images. The Fisher vectors can be represented compactly and requires no labeled training data, but computing them is costly. J´egou et al. [20] proposed the construction of vector of locally aggregated descriptors(VLAD). These descriptors represent images compactly and they can be computed efficiently. J´egou et al. further improved the VLAD representation in [21] by using residual normalization and local coordinate system.

A. F. Mollah et al. [10] presented an OCR system for camera based handheld devices. M. F. Kader et. al. [11] proposed a neural network based size and color invariant character recognition system using feed forward neural network. Prerna Kakkar et. Al. [13] presented a novel approach for recognition of English characters based on artificial neural network. Kauleshwar Prasad et al. [18] presented a handwriting recognition system to recognize English characters using Matlab’s neural network toolbox. Anita Pal et. al. [15] presented a handwritten English character recognition system using neural network.

III. PROPOSED METHODOLOGY

In this paper, three main steps involved Optical Character Recognition. They are Preprocessing, Feature extraction and Classification. The process of enhancing the image, which should be used for further processing, is called preprocessing. Preprocessing is the major step in handwriting recognition system.

The scanned images not only have noises which are inbuilt within it, but also the noise may be during the scanning of that image. So the noises and the unwanted information should be removed from the image. Preprocessing is not the single step rather it contains sequences of steps. They are Binarization, Normalization, Sampling, Denoising and Thinning.

SIFT method was first proposed by Lowe in [12] and has been successfully applied in many fields due to its capability to extract very distinctive and scale invariant features from images [13]. SIFT is usually the preferred method of feature extraction in applications such as object retrieval and object detection. RootSIFT and SIFT follow the same principle for the extraction of the features with the only difference being that SIFT uses an Euclidean distance for similarity measurement while RootSIFT uses the Hellinger kernel. By using the Hellinger kernel instead of the Euclidean one, significant performance improvements can be obtained [14]. This is due to the fact that the Euclidean distance is much less efficient than the Hellinger kernel for comparison of histograms. Furthermore, the different scanning procedures of the documents can cause variations in the illumination for the training and query documents. Therefore, these challenges should be taken care of by the feature extraction method to ensure robustness against such variations in order to provide a reliable result. RootSIFT has proven to be an efficient method to address these geometric distortions as demonstrated in the field of character identification [15]. RootSIFT algorithm operates in four stages. First, an image is broken down into a Gaussian pyramid of octaves where the original image is then convolved with its corresponding octaves of the pyramid with difference of Gaussian (DoG) filters at different variances. In the next stage, referred to as key point localization, the stable key points are detected. Then the orientations, scales and locations of these key points are calculated. Finally, 128 dimension descriptors are generated to represent the image features. This is based on the histogram of oriented gradients (HoG).

Finally, Back propagation neural network algorithm is implemented to classify the features extracted from training and testing dataset to classify the characters. BPNN is layered feed-forward supervised network and the popular tool to solve wide range of problems like classification, recognition and control. The architecture of BPNN is a multilayer neural network with one layer of hidden units is shown in Figure 1.

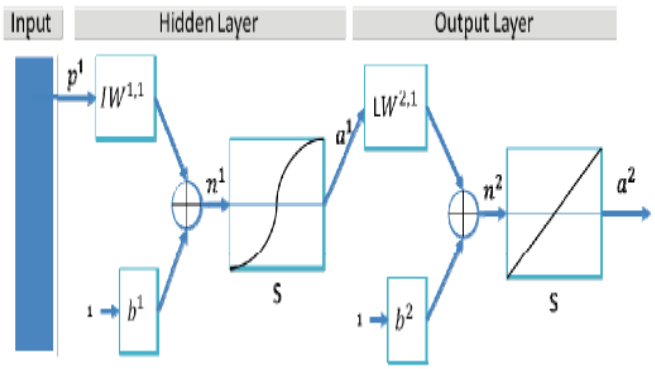

Fig. 1. Backpropagation neural network.

IV. EXPERIMENTAL RESULTS

This section depicts the experiments led to look at and assess the recognition of characters utilizing ABC improved the neural network algorithm. The experiments are directed utilizing MATLAB programming on a 64-bit Windows 7 Professional machine. The processor utilized is Intel Pentium, 2.20 GHz, and 4.00 GB.

Images of 62 characters of 500 examples are taken for preparing the ABC improved back propogation neural network algorithm. Each character contains 7-10% salt and pepper write noise.10 tests of each character are taken to make a test dataset where each character contains 7-10% salt and pepper compose commotion. Finally, the trained system is reproduced with the test informational collection which perceives the test characters. The system is recreated for both individual classes and joined classes. The system determination is given underneath table 1:

Input layer	1
Hidden layer	1
Output layer	1
No. of neuron in input layer	96
No. of neurons in hidden layer	79
No. of neurons in output layer	62

Table 1. Specification Of The Network

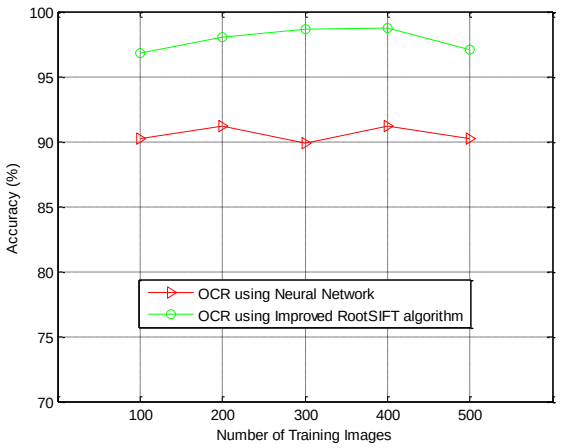

Figure 2: Accuracy comparison

The above figure shows that the modified RootSIFT with back propagation neural network algorithm has achieved a maximum accuracy of 97.3077% compared with the neural network algorithm, which has a 90.8022% accuracy.

V. CONCLUSION

In this paper, we have proposed an improved RootSIFT descriptor for feature extraction of english characters in the digital scanned images. After that the artificial neural network classification algorithm is used to classify the

english alphabets of different fonts. Our system produces very good result for individual classes but produces satisfactory result when the classes are combined together. In future, we can try the machine learning algorithm for better training of characters and can improve the recognition accuracy.

REFERENCES

[1]. Arindam Chaudhary, Krupa Mandaviya, Pratixa Badelia and Soumya K Ghosh, "Optical Character Recognition Systems for Different Languages with Soft Computing", Springer International Publishing AG, Picassoplatz 4, CH-4052 Basel, Switzerland,2017.

[2] Carsten Steger, Markus Ulrich, Christian Wiedemann. Machine Vision Algorithms and Applications[M]. Beijing: Tsinghua University Press, 2008

[3] Kashyap Paidimarri, Sharat Chandran. Ad-Hoc Multi-planar Projector Displays, in: Computer Vision, Graphics and Image Processing[M]. Heidelberg: Springer Berlin Press, 2006, 4338: 289-298

[4] Huizhu Luo, Rangachar Kasturi. Improved Directional Morphological Operations for Separation of Characters from Maps/graphics, in: Graphics Recognition Algorithms and Systems[M]. Heidelberg: Springer Berlin Press, 2006

[5] Michael Blumensteina, Xinyu Liu. An Investigation of the Modified Direction Feature for Cursive Character Recognition. Pattern Recognition[J], 2007, 40(2): 376-388

[6] X. Wang, X. Ding, C. Liu. Gabor Filters-based Feature Extraction for Character Recognition. Pattern Recognition[J], 2005, 38(3): 369-379

[7]. Ladoux PO, Rosenberger C, Dorizzi B (2009) Palm vein verification system based on sift matching. In: Advances in Biometrics, Springer. pp. 1290–1298.

[8]. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. International journal of computer vision 60: 91–110. View Article Google Scholar

[9]. Xiuyan L, Tiegen L, Shichao D, Jin H, Yunxin W (2011) Fast recognition of hand vein with surf descriptors. Chinese Journal of Scientific Instrument 32: 831–836.

[10]. Kim HG, Lee EJ, Yoon GJ, Yang SD, Lee EC, et al.. (2012) Illumination normalization for sift based finger vein authentication. In: Advances in Visual Computing, Springer. pp. 21–30.

[11]. Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). Computer vision and image understanding 110: 346–359.

[12] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International journal of computer vision, vol. 60, no. 2, pp. 91–110, 2004.

[13] D. Nister and H. Stewenius, "Scalable recognition with a vocabulary tree," in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol. 2. IEEE, 2006, pp. 2161–2168.

[14] R. Arandjelovi´c and A. Zisserman, "Three things everyone should know to improve object retrieval," in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012, pp. 2911–2918.

[15] S. Fiel and R. Sablatnig, "Writer retrieval and writer identification using local features," in Document Analysis Systems (DAS), 2012 10th IAPR International Workshop on. IEEE, 2012, pp. 145–149.

[16] Josef Sivic and Andrew Zisserman. Video google: A text retrieval approach to object matching in videos. In Ninth IEEE International Conference on Computer Vision, 2003. Proceedings., pages 1470–1477, 2003.

[17] David G Lowe. Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2):91–110, 2004.

[18] Bastian Leibe Kristen Grauman. Visual object recognition. In Thomas Dietterich Ronald Brachman, editor, Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan and Claypool Publishers, 2009.

[19] Florent Perronnin, Jorge S´anchez, and Thomas Mensink. Improving the fisher kernel for large-scale image classification. In Computer Vision–ECCV 2010, pages 143–156. Springer, 2010.

[20] Herv´e J´egou, Matthijs Douze, Cordelia Schmid, and Patrick P´erez. Aggregating local descriptors into a compact image representation. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 3304–3311. IEEE, 2010.

[21] Jonathan Delhumeau, Philippe-Henri Gosselin, Herv´e J´egou, and Patrick P´erez. Revisiting the vlad image representation. In Proceedings of the 21st ACM international conference on Multimedia, pages 653–656. ACM, 2013.