Improved OCR Using Histogram of Oriented Gradients And Deep Belief Network Learning

SANDHYA BALAKRISHNAN P K

Research Scholar,

Department of Computer Science,

CMS College of Science and

Commerce, Coimbatore.

Dr.L.PAVITHIRA

Associate Professor

Department of Computer Science,

CMS College of Science and

Commerce, Coimbatore.

***ABSTRACT*: Optical character recognition is one of the key research topic in the past few years. It has been demonstrated that high acknowledgment rates can be accomplished in explicit application situations utilizing some standard and all-around contemplated techniques, for instance, support vector machine (SVM), etc. However, the main aim is the reliable and speedy recognition in the fairly controlled environments, the major issues related to the system's adaptivity are not fully investigated in the literature. Mainly, the recognition rate of the OCR systems will then be affected by the image noise and the difference between the training databases. In this paper, we explore the likelihood of learning a features for structuring OCR. We have proposed a deep belief network for the acknowledgment of content. Histogram of arranged inclinations is utilized as the features. Better recognition accuracy is achieved when experimented with a dataset consisting of more than 10,000 samples.**

Keywords: OCR, HOG, Deep belief networks.

I. INTRODUCTION:

OCR (optical character recognition) is the acknowledgment of printed or composed content characters by a PC. This includes image scanning of the content character-by-character, investigation of the scanned picture, and afterward interpretation of the character picture into character codes, for example, ASCII, generally utilized in information processing. In OCR processing, the examined in picture or bitmap is dissected for light and dim regions so as to distinguish each alphabetic letter or numeric digit. At the point when a character is identified, it is changed over into an ASCII code. Exclusively designed processor chips made for OCR are utilized to accelerate the recognition procedure.

The Deep Belief Network (DBN) [1] and Deep Boltzmann Machine (DBM) [2] are two famous profound probabilistic generative models that give cutting edge results in numerous issues. These models contain more than two layers of hidden factors, and use an undirected graphical model called the Restricted Boltzmann Machine (RBM) [3] as the building square. An important property of the RBM is that angle appraises on the model parameters rush to ascertain, and stochastic inclination plummet gives moderately productive deduction. Be that as it may, assessing the likelihood of an information point under RBM is nontrivial because of the computationally recalcitrant parcel work, which must be evaluated, for instance utilizing an annealed significance sampling algorithm [4].

The paper is divided into five sections thus: (1) Introduction, (2) Pre-processing, (3) segmentation and feature extraction, (4) template matching and (5) conclusion.

II PRE PROCESSING

Preprocessing is the first step in the development of any OCR system which prepares images for the subsequent phases. Depending upon the application and the type of input images, preprocessing may involve binarization [6], noise removal and, skew and slant detection and correction [7]. In our study, we intend to work on contemporary images which are not likely to suffer from noise or degradations. The preprocessing in our case, therefore, comprises binarization of image to segment text from the background. In our implementation, we have employed the well-known Otsu‘s global thresholding to binarize the text image.

II. SEGMENTATION AND FEATURE EXTRACTION

In any OCR system, character segmentation and feature extraction are essential stages after the image is being acquired. They constitute two major phases upon which largely OCR performance depends on as shown in the figure 1.

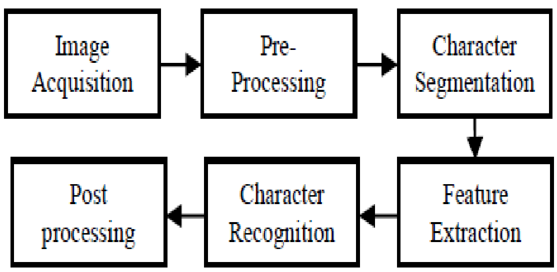


Figure 1: Overview of OCR.

A. SEGMENTATION

Segmentation processes, including following processes:

a. Line segmentation

Line segmentation is the process in which from the image, we extract only lines or differentiate the lines. Horizontal projection of a scanned image is used to extract the horizontal text lines of an image. This will have an isolated pinnacles and valleys for the lines that are very much isolated and are not tiled, which fill in as the separators of the content lines. Those valleys are effectively identified and used to find the area of limits between the lines.

b. Word segmentation

Word Segmentation is a procedure of isolating a sentence into its segment words. Word part is the way toward parsing linked content to gather where word breaks exist. By utilizing vertical projection profile, one can get section totals. By searching for minima in horizontal projection profile of the page, we can isolate the lines and after that different words by taking a gander at minima in the vertical projection profile of a solitary line. By utilizing the valleys in the vertical projection of a line picture, one can extricate words from a line and furthermore separating individual characters from the word [5].

c. Character segmentation

In character segmentation, we extricate just characters from word. Character segmentation is a troublesome advance of OCR frameworks as it separates important locales for investigation. This progression decays the pictures into classifiable units called character [5].

B. Feature extraction

Histogram of oriented gradients algorithm is executed for extraction of features. Histogram of oriented gradients (HOG) is a component descriptor used to distinguish features in digital image processing and computer vision technology. The HOG descriptor system includes events of inclination introduction in restricted segments of a picture – discovery window, or region of interest (ROI).

HOG descriptor algorithm steps are:

1. Split the images in to small isolated regions called cells, and HOG is calculated for all the cells splitted from the image.

2. Discretize each cell into angular bins according to the gradient orientation.

3. Each cell's pixel contributes weighted inclination to its relating angular bin.

4. Groups of contiguous cells are considered as spatial areas called squares. The gathering of cells into a square is the reason for gathering and standardization of histograms.

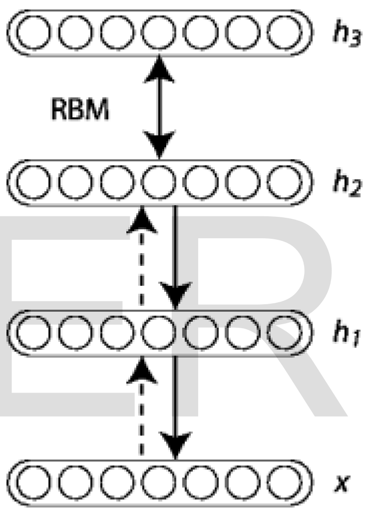
5. Normalized gathering of histograms speaks to the square histogram. The arrangement of these square histograms speaks to the descriptor.

III. Template Matching

In this module, deep belief network algorithm is executed for template matching. DBNs are graphical models which figure out how to detect a deep hierarchical portrayal of the training data. They show the joint dissemination between observed vector x and the hidden layers ℓ as pursues:

$$P(x, h^1, \ldots, h^\ell) = \left(\prod_{k=0}^{\ell-2} P(h^k|h^{k+1})\right) P(h^{\ell-1}, h^\ell)$$

Where $x = h^0$, $P(h^{k-1}|h^k)$ is a restrictive conveyance for the visible units molded on the hidden units of the RBM at level k, $P(h^{\ell-1}, h^\ell)$ and is the noticeable concealed joint dissemination in the top level RBM. This is shown in the figure beneath.



The standard of layer-wise unsupervised training can be connected to DBNs with RBMs as the building obstructs for each layer [8]. The procedure is as per the following:

1. Train the initial layer as a RBM that models the input $x = h^{(0)}$ as its visible layer.

2. Utilize that first layer to acquire a portrayal of the input that will be utilized as information for the second layer. Two normal arrangements exist. This representation can be picked just like the mean enactments $p(h^{(1)} = 1|h^{(0)})$ or tests of $p(h^{(1)}|h^{(0)})$.

3. Train the second layer as a RBM, taking the changed information (tests or mean initiations) as preparing models (for the obvious layer of that RBM).

4. Repeat (2 and 3) for the ideal number of layers, each time spreading upward either tests or mean values.

5. Calibrate every one of the parameters of this profound architecture as for an intermediary for the DBN log- likelihood, or regarding an administered training standard.

We center around adjusting through supervised gradient descent. In particular, we utilize a logistic regression classifier to order the input x dependent on the yield of the last hidden layer $h^{(l)}$ of the DBN. Adjusting is then performed by means of directed slope drop of the negative log- likelihood cost function. Since the supervised gradient is just non-invalid for the weights and hidden layer inclinations of each layer (for example invalid for the noticeable predispositions of each RBM), this methodology is equal to instating the parameters of a profound MLP with the weights and hidden layer inclinations got with the unsupervised training procedure.

IV.EXPERIMENTAL RESULTS

The performance of the proposed method is found by experimenting on the natural image samples. 400 samples for each character are collected from different images and with different background. The accuracy is calculated as the average of recognition rates for the five test subsets. We have studied the effect of training sample set size on the recognition accuracy and found that using 80% of the data for training results in improvement of accuracy.

V.CONCLUSION:

A new approach for OCR was proposed using modified template matching. The deep belief network is implemented to learn the HOG descriptors features. The deep learning network algorithm contributed to increase in recognition accuracy. The algorithm will have impressive results with the high training database.

VI.REFERENCES

[1]. G. Hinton, S. Osindero, and Y.-W. Teh. A fast learning algorithm for deep belief nets. Neural computation, 2006.

[2]. R. Salakhutdinov and G. E. Hinton. Deep Boltzmann machines. AISTATS, 2009.

[3]. G. E. Hinton. Training products of experts by minimizing contrastive divergence. Neural computation, 2002.

[4]. R. Salakhutdinov and I. Murray. On the quantitative analysis of deep belief networks. ICML, 2008.

[5]. A. Rehman, D. Mohamad, and G. Sulong "Implicit Vs Explicit based script segmentation and recognition: A performance comparison on benchmark database," International Journal Open Problems Compt. Math., vol. 2, pp. 352-364, 2009.

[6]. M. Michael and N. Papamarkos, "An adaptive layer-based local binarization technique for degraded documents," International Journal of Pattern Recognition and Artificial Intelligence, vol. 24, no. 245, 2010.

[7]. P. Sanjoy, P. Bhowmick, S. Sural, and J. Mukhopadhyay, "Skew correction of document

images by rank analysis in farey sequence," International Journal of Pattern Recognition and Artificial Intelligence, vol. 27, no. 1353004, October 2013.

[8]. Y. Bengio, P. Lamblin, D. Popovici and H. Larochelle, Greedy Layer-Wise Training of Deep Networks, in Advances in Neural Information Processing Systems 19 (NIPS'06), pages 153-160, MIT Press 2007.