Multi-font Optical Character Recognition Using Deep Learning

P.K. Sandhya Balakrishnan, L. Pavithira

Abstract--- Deep learning (DL) is a new area of research in classification, in which the aim is moving us closer towards the objective of artificial intelligent. However, the results of a DBN are often highly based on settings in particular the combination of runtime parameter values. In this paper, Simulated Annealing (SA) is proposed to increase the results of Convolution Neural Network (CNN), as an alternate method for traditional CNN. Then the proposed SA-CNN classifier is implemented using RETAS OCR dataset and provided the improved recognition accuracy than the CNN classifier.

Keywords--- Deep Learning, SA, CNN, OCR.

I. INTRODUCTION

Optical character recognition (OCR) is a procedure by which specially designed software and it is used to convert scanned images of text into electronic text, therefore the digitized data can be searched, indexed and recovered. OCR engines are created and enhanced for several real-world applications such as extracting data from business documents, checks, passports, invoices, bank statements, insurance documents, license plates and more[1]. These applications needs processing dataset with the purpose of involves hundreds of thousands scanned documents or images in order to train and optimize the classifiers. Processing the training data set is commonly performed by humans in order to give correct data that can be used by the engine to study and apply, making it "smarter" over time.

The common OCR system includes of three major steps such as image acquisition and preprocessing, feature extraction and classification [1]. The first step cleans up and enhances the quality of the image by noise removal, binarization,color adjustment and text segmentation. Second step, Featureextraction is introduced to extract and capture data from theacquired text image which can be used for classification. Final phase, the part of the segmented text in thedocument image is mapped towards the equal textual image.

In recent years, conventional algorithms in the area of OCR study have been roughlychanged with Convolutional Neural Networks (CNN). Oquab*et al.* introduced a new CNN towardsstudy image representations on a huge annotation dataset be able tosufficiently transfer this data toother visual identification tasks with a restricted amount of data [2]. Yejun Tang *et al.* proposed towardsinsert an adaptation layerin CNN by transfer learning, which obtainsresults improvement in past Chinese character detectiontasks[3]. Motivation from this work, we introduceda deep neuralnetwork with transfer learning used for broken English characterrecognition.

In common, methods of deep learning can be divided into deep discriminative models and generative models [4]. Some of the discriminative models are Deep Neural Networks (DNNs), Recurrent Neural Network (RNNs), and Convolutional Neural Networks (CNNs). On the other side, generative models, for example, are Restricted Boltzmann Machine (RBMs), Deep Belief Networks (DBNs),regularized autoencoders, and Deep Boltzmann Machines (DBMs). All of those methods, CNNs are the major focus of this paper. CNNs is a deep supervised-learning though this model is typically capable to train and test of samples, flexible to construct, and appropriate for end-to-end learning of complex system [4].

Though CNNs has good reputation for handling a variation of learning task, it is not easy to train [5]. Recently optimization techniques are used for training DL used layered wise pre-training [6]. Some of the methods are used for training of these CNNs methods are Stochastic Gradient Descent, Conjugate gradient, Hessian-free Optimization and Krylov Subspace Descent. In fact, a huge popular of optimization methods are generally heuristic and metaheuristic [7]. These optimization methods have been applied in various areas of science and engineering application. However, the research of meta-heuristic algorithms to optimize DL is hardly conducted. Because of that in this worked, we used metaheuristic algorithm asanother approach for optimal performance of DL.

In this paper, simulated annealing (SA) algorithm is proposed to optimize the Convolutional neural network. SA is a robust technique, which givesimproved solutions to single and multipleobjectiveoptimization problems by means of a considerabledecrease in computation time. The origin of this algorithm is Metropolis Algorithm, in statistical mechanics [8].

This paper is organized as follows. Section 2 gives the related works for OCR, Deep Learning and Simulated Annealing. Section 3 gives description of proposed methodology. Section 4 presents simulation result. Finally, Section 5 concludes this paper.

II. RELATED WORKS

Character recognition is procedure of identifying and recognizing characters from input image and converts them intoASCII or other equal machine editable form [9], [10], [11].OCR system is most appropriate for the applications such as data entry for businessdocuments, automatic number plate recognition,multi choice examination; approximatelyevery one kind of form processing system.

Manuscript received June 10, 2019.

P.K. SandhyaBalakrishnan*, Department of Computer Science, CMS College of Science and Commerce, Coimbatore, T.N, India.

Dr.L. Pavithira, Associate Professor Department of Computer Science, CMS College of Science and Commerce, Coimbatore, T.N, India.

Generally, traditional methods in the definition step are depending on either on Template Description (TD) or on Feature Description (FD) [12]. In TD Methods [12] [13], characters are labelled depending on their pixel data. In spite of their simple make use of these methods are not robust on noisy characters.

In addition, computation is loaded with redundant pixel description. Less complex, the FD methods [14] [15] [16] carryout the description of characters depending on some specific Features.

In the other hand, the decision making is performed related to a matrix or extracted feature matching. As mentioned above, pixel comparison in matrix matching is very susceptible towards noisy characters [12] [13]. This step is able to be carryout using classification methods such as RNA, SVM [14] [15] [16].

They give improved results; on the other hand they still complex comparing to matching procedure. Certainly, simple feature matching creates a good trade-off among accuracy and computation complexity [16].

In the recent work, deep learning becomes anextremelyrelevant method [17], particularlyfollowing the winning of AlphaGooverhuman world champion. On image classification deeplearning has obtained good results, and remainsas a very competitive classifier for this task. On the other hand, related to mainly learning algorithms, deep learning is susceptible to parameter settings.

III. PROPOSED METHODOLOGY

Convolution Neural Network (CNN) is a variation of Multi-Layer Perceptron (MLP) [12] classifier. The structural design of new CNN is typically organized in two major layers: convolution layers andsubsampling layers are shown in Fig.1 [12].

In this CNN technique, a classifier trained on a multilevel architecture which includes of certain stages. Every stage has input and output in the caller of feature map, which are set in array.

Every feature map on the output denotes a certain feature extracted at every one locations on the input. Where each step includes of three layers: a layer of the *filter bank*, a layer of *nonlinearity*,and a layer of *feature pooling* [16].

Theoretically, a feature map is established by convolving *k*-th input image with a linier filter. Ranges of biasterm are included and then apply it to the non-linear function. If *k*-th feature map at a known layer as h^k, whose filters are computed by the weights W^k and bias b_k, then the feature map h^k is achieved as follow:

$$h_{ij}^{k} = \tanh((W^{k} * x)_{ij} + b_{k}$$

The design of proposed CNN technique uses: 6c – 2s – 12c – 2s. The architecture of this CNN classifier is illustrated in Fig.1. The kernel size of 1st and 2nd convolution layer is 5x5. Conversely, the scale of 1st and 2nd subsampling layer is 2 [12].

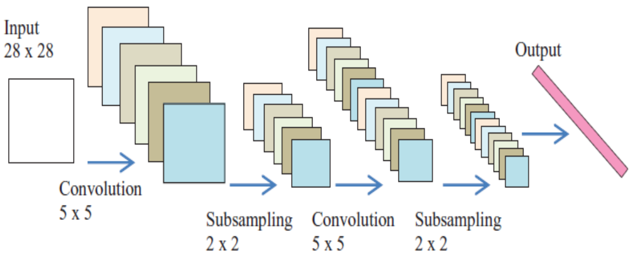

Fig.1: Proposed Architecture

The objective of using SA to train CNN is to meet definite accurateness constraint and to make the estimate error and network complexity indicators to the lower side. These can be done by computing standard error on the training set and fitness function of vector solution. The fitness functions used in this work are described as follow:

$$f = \frac{1}{2}\sqrt{\frac{\sum_{i=1}^{R}(o-y)^2}{R}}$$

Where *o* is denoted as the desired output, *y* is denoted as the actual output, and *R* is denoted as the number of training samples. In caseof termination criterion, the following two conditions need to be satisfy. First is the iteration reaches the particular maximumiteration, and second is the fitness function is less than anexacting constant. At this moment, the estimated error and minimum network complexity indicators have obtained the mainly optimal condition.

Fundamentally in this proposed CNN classifier, weights and the bias of the system are calculated, and then the results on the last layers are used to calculate the lost function, as the finest solution vector *x*0 to be optimized in SA. Choosing a new result vector from the neighbourhood of the latest solution is conducted by adding Δx randomly. It is updated based on the objective function $f(x_0+\Delta x)<f(x_0)$. When SA algorithm has obtained the best objectivefunction, the value of weights and the bias are updated for each and every layers of the system. Flowchart of thisproposed SA-CNN classifier is shown in Fig. 2.

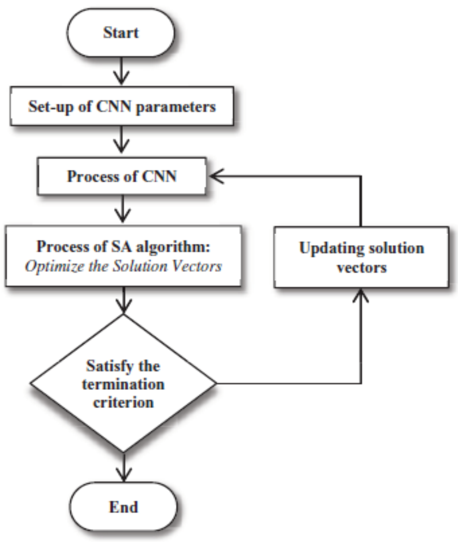

Fig.2: Flowchart of SAalgorithm for optimal CNN

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

IV. EXPERIMENTAL RESULTS

A number of scanned books in different languages are downloaded from [18] and their OCR accuracy is evaluated.According to the metadata, these books are recognized usingABBYY FineReader 8.0. The ground truth texts are obtainedfrom [19]. For each book, word and character recognition accuracies are estimated. The OCR accuracy metric is definedas follows:

$$Acc = \frac{m}{c}$$

where m is the total number of matching characters/words inthe alignment and c is the total number of characters/wordsin the ground truth. This metric accounts for the containmentof the ground truth text in the OCR output. The rationale behind this approach is to obtain a statistical evaluation ofthe OCR accuracy for the portion of the text for which wehave ground truth.

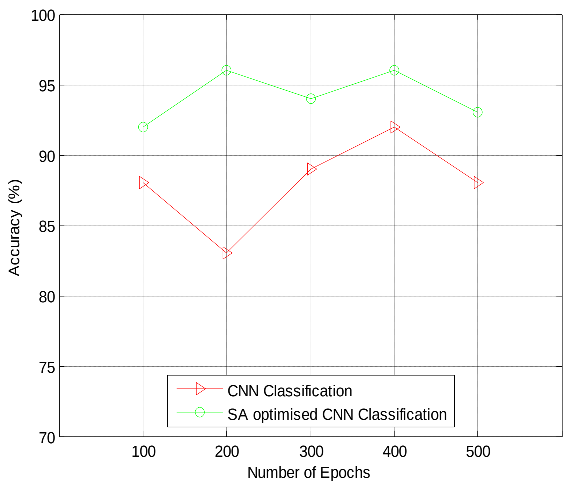

Fig.3: Simulation results of CNN and SA Optimised CNN

The primary concern of simulation in this paper is to improve the performance of original CNN byusing SA.Demonstration of the performance of implementation CNN and CNN by SA on RETAS dataset aregiven in Fig. 3. In general, the simulation results show that percentage accuracy of CNN by SA isbetter than the original CNN.

V. CONCLUSION

The objective of our proposed method, to optimize Convolution Neural Networks (CNN) usingSimulated Annealing (SA), has been achieved for optical character recognition. Theclassification accuracy from the proposed method is lower than the original of CNN for variation of fonts. This proposed method could potentially be employed and tried for benchmark dataset RETAS.The proposed method proves the efficiency of character recognition.

REFERENCES

1. R. Anil, K. Manjusha, S. S. Kumar, and K. P. Soman,"Convolutional Neural Networks for the Recognition ofMalayalam Characters," in Proceedings of the 3rdInternational Conference on Frontiers of IntelligentComputing: Theory and Applications (FICTA) 2014, vol.328, Springer International Publishing, 2015, pp. 493–500.
2. M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning andTransferring Mid-level Image Representations UsingConvolutional Neural Networks," in Proceedings of the2014 Computer Vision and Pattern Recognition (CVPR),2014, pp. 1717–1724.
3. Y. Tang, L. Peng, Q. Xu, Y. Wang, and A. Furuhata,"CNN Based Transfer Learning for Historical ChineseCharacter Recognition," in Proceedings of the 12th IAPRWorkshop on Document Analysis Systems (DAS), 2016,pp. 25–29.
4. L. Deng and D. Yu, Deep Learning: Methods and Application, NOW Publisers 2014.
5. J.D. Lamos Sweeney, Deep Learning using Genetic Algorithms, master thesis, Dept. of Computer Science, RochesterInstitute of Technology, 2012.
6. G. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural network", Science, 313:504-507, 2006.
7. Xin-She Yang, Engineering Optimization: an introduction with metaheuristic applications, John Wiley & Sons, Hoboken,New Jersey, 2010.
8. El-GhazaliTalbi, Metaheuristics from Design to Implementation, John Wiley & Sons, Hoboken, New Jersey, 2009.
9. Kai Ding, Zhibin Liu, LianwenJin, Xinghua Zhu, "A Comparativestudy of GABOR feature and gradient feature for handwritten Chinese character recognition", International Conference on Wavelet Analysisand Pattern Recognition, pp. 1182-1186, Beijing, China, 2-4 Nov.2007
10. Pranob K Charles, V.Harish, M.Swathi, CH. Deepthi, "A Review onthe Various Techniques used for Optical CharacterRecognition",International Journal of Engineering Research andApplications, Vol. 2, Issue 1, pp. 659-662, Jan-Feb 2012
11. Om Prakash Sharma, M. K. Ghose, Krishna Bikram Shah, "AnImproved Zone Based Hybrid Feature Extraction Model forHandwritten Alphabets Recognition Using Euler Number",International Journal of Soft Computing and Engineering, Vol.2, Issue2, pp. 504-58, May 2012
12. J.R. Prasad, Template matching algorithm for Gujrati Character Recognition,IEEE 2nd International conference on Emerging Trends in Engineeringand Technology, ICETET (2009)
13. Z.You-qing, A Recognition Method of Car License Plate CharactersBased on Template Matching Using Modified Hausdorff Distance. IEEEInternational Conference on Computer, Mechatronics, control and electronicEngineering (CMDE) (2010)
14. F.Aghdasi, Automatic Licence Plate recognition system, IEEE AFRICON(2004)
15. J.Jiao, A configurable method for multi-style license plate recognition.Pattern Recognitionn, Vol: 42, No: 3, pages 358-369 (2009)
16. M.M Farhad, An efficient Optical Character Recognition Algorithm usingAritificial Neural Network by CurvativeProperties.IEEE Internationalconference on informatics, electronics ´e vision (2014)
17. Y. Bengio, A. Courville, and P. Vincent. Representation learning: Areview and new perspectives. IEEE Transactions on Pattern Analysisand Machine Intelligence, 35(8):1798–1828, Aug 2013.
18. "The Internet Archive," http://www.archive.org, 2011.
19. "Project Gutenberg," http://www.gutenberg.org, 2011.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication