2018 International Conference on Information, Communication, Engineering and Technology (ICICET)
Zeal College of Engineering and Research, Narhe, Pune, India. Aug 29-31, 2018

Offline Handwritten Mathematical Expression Recognition using Convolutional Neural Network

Lyzandra D'souza
Computer Engineering Department
Goa college of Engineering
Goa, India
lyzandra_06@yahoo.com

Maruska Mascarenhas
Computer Engineering Department
Goa college of Engineering
Goa, India
maruskha@gec.ac.in

***Abstract*— Offline Handwritten Mathematical Expression recognition is recognizing the Handwritten Mathematical Expression(HME) that are written on paper , paper is scanned and image is passed to the recognition system. This has caught the interest of most of the researchers and a lot of them are working on this topic and have used various classifiers. In past, Convolutional Neural Network also called CNN has gained high performance in recognizing the patterns. We propose an idea to recognize offline HME using CNN for classification. The steps involved are data collection of handwritten mathematical expression and symbols. Next, preprocessing steps are performed on the collected data. Segmentation of HME into individual symbols is done. To train CNN classifier, Handwritten Mathematical symbols (HMS) are used. The individual segmented symbols are then sent to the CNN classifiers to recognize which class they belong.**

Keywords*— *Offline Handwritten Mathematical Expression Recognition, Convolutional Neural Network, Isolated Handwritten Mathematical Symbols

I. Introduction

Mathematics is widely used in engineering, science and many other fields. With the growing demand to convert already existing handwritten data to text that is machine readable, many research scholars are working on it. Handwritten Mathematical Expression (HME) recognition is of two kind offline and online. Offline is when an image of handwritten data is taken and passed to the recognition system. Online is when handwritten data is written on tablet or mobile phone and this is recorded as strokes that is then passed to the recognition system.

In the past years, computer vision has done great improvement and it is currently also working on further improvement. The performance achieved by Convolutional Neural Network (CNN) is very high in terms of Handwritten Mathematical Symbol (HMS) recognition with 87.72% [1]. Though such high accuracy has been achieved, yet there are some difficulties faced like recognizing characters (1 ! |) are recognized as any one and not the correct one. This is due to variation in the writers' style of writing. Some more difficulties faced are twisted semantics [2][3] and large set of HMS [4][5].

II. Related studies

Many research scholars have worked on recognizing offline mathematical expressions using various classifications, feature extraction, segmentation and preprocessing methods. Manisha Bharambe[6] recognized logical expressions using the Support Vector Machine(SVM) classifier with 93.8% accuracy. Padmapriya and Karpagavalli[4] recognized simple algebraic expression with SVM and Multi-Layer Perceptron (MLP) as classifier.

The preprocessing methods involve were binarization, noise reduction. There are various methods for noise removal. One of them is median filtering [7] a nonlinear filter. Similarly there are various methods for thresholding. Adaptive thresholding is the most likely that can be used here as every image will have a different threshold value to binarize. The Sauvola Method[8] gives suitable threshold value.

The segmentation methods involved bounding box, connected component [4], both these methods were only for isolated symbols and order of symbol in the expression was lost. Hence the proposed segmentation algorithm [10] maintains the order of symbols using list like tree structure. Segmenting of online HME is simpler as while the writer writes the expression, strokes are recorded and then for segmentation grouping of strokes is carried out. The segmentation rate obtained for online HME is higher in comparison to offline HME.

The feature extraction methods used by Manisha Bharambe[6] are normalized chain code, moment invariant features, density features, histograms while Padmapriya and Karpagavalli [4] used zoning, skeletonization, directional features [4][6]. These features were then passed to the classifiers to recognize the symbols.

The classification methods used for offline HME are MLP, SVM, Hidden Markov Model (HMM), CNN [7][9]. CNN uses features that are extracted at the layers like in convolutional, pooling while MLP, SVM and HMM uses the feature extracted from feature extraction stage.

Most of them use LeNet-CNN which has four layers (two convolutional and two pooling) while SpNet has six layers (three convolutional and three pooling) while LeNet-5 has only one extra layer (pooling) than LeNet. Syrine Ben Driss[9] performs comparative study for LeNet, LeNet5 and SpNet and shows that SpNet has a high classification rate and learning rate than LeNet.

III. Proposed system

A. Data collection

A dataset of HME is collected. The dataset is of HME written by different authors. A dataset of symbols is collected that include alphabets, numbers, Greek letters,

978-1-5386-5510-8/18/$31.00 ©2018 IEEE

special characters like (… / + - * | , < > etc.). 83 different kinds of HMS are collected. The dataset collected is stored as image with .jpeg extension.

B. *Preprocessing*

Once the HME image is collected then preprocessing has to be done on images so that the image is clear for further processing. Grayscale image is obtained. The preprocessing methods used are:

- Noise Reduction: The HME image may contain some noise, to remove this noise we use Median filtering [7]. In median filtering, the center value is replaced with median value of neighborhood in the window.
- Binarization: Binarization helps to differentiate foreground as 1 and background as 0. Adaptive threshold method is used in obtaining a binary image having pixel intensity values as 0 or 1. To find the adaptive threshold Savoula Method [8] is used. In this method the threshold is obtained where window size is size of image.
- Thinning: Thinning is used to reduce the thickness of foreground to approximately 1 pixel. Using thinning algorithm, the skeletal image of HME is obtained. Thinning is also performed on HMS.

The preprocessing methods like binarization and thinning are done on the HMS dataset also.

C. *Segmentation*

The segmentation of the HME into individual math symbols is done using:

- Segmentation using recursive projection profiling algorithm [10]:

This algorithm segments the HME into individual symbols. Then merging together of the small segments as one symbol like (=, i, !) is done.

To segment the expression with vertical projection called Vertical Projection Profile Cutting (VPPC) and horizontal projection called Horizontal Projection Profile Cutting (HPPC), a binary image is necessary. Vertical projection is parallel to y-axis while horizontal to x-axis. The image is segmented where either VPPC or HPPC, sum of pixels along the projection line is zero [11].

We first have to apply VPPC to the image and obtain horizontal partitions. Once horizontal partitions are obtained then HPPC must be applied to each partitioned block from VPPC to obtain vertical partitions. These steps above are performed in recursion until further partitioning is not possible. The result of this is passed to next segmentation algorithm [11].

- Segmentation using connected component labeling [10]:

The above projection segmentation algorithm has a disadvantage when symbols like square root (√) or any other symbol that is difficult to segment appears.

Connected component (CC) is applied on HME binarized image. Here it checks the neighborhood and if any pixel in neighborhood has same pixel intensity then it is one component.

Once the HME is segmented to HMS than the tree structure is obtained with the links as shown in Fig. 1.

This tree structure keeps a track of the HME order. It is useful in obtaining the final Latex output in correct order. Once the expression is segmented and the tree structure is obtained than the HMS is passed to the classifier.

The above two algorithm works best only for isolated symbols and not for connected or joined symbols.

D. *Classification*

The classifier used is CNN [12] with 83 different classes. The proposed CNN is called SpNet-CNN. The SpNet-CNN shown in figure (Fig.2.) comprises of the following layers input, convolutional, max pooling, convolutional, max pooling, convolutional, max pooling, fully connected and softmax.

In convolution layer, the input image is applied to a filter of size 3x3. All the three convolution layers have same filter size. The activation function after Convolutional layer used is Rectified Linear Unit (ReLu). In ReLu, the negative values are replaced with zero while positive remain the same.

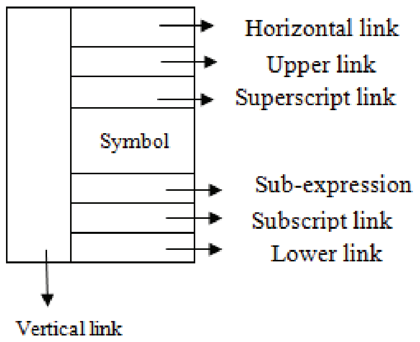

Fig 2: Tree structure Node for Handwritten Mathematical Symbol

In pooling layer, output size of the image from pooling is half the size of input image to the layer when subsampling size of 2x2 is applied. Convolutional layers & pooling layers extract features. These extracted features are passed to fully connected layer. In fully connected layer, every input neuron connects every output neuron. The activation function applied is Softmax.

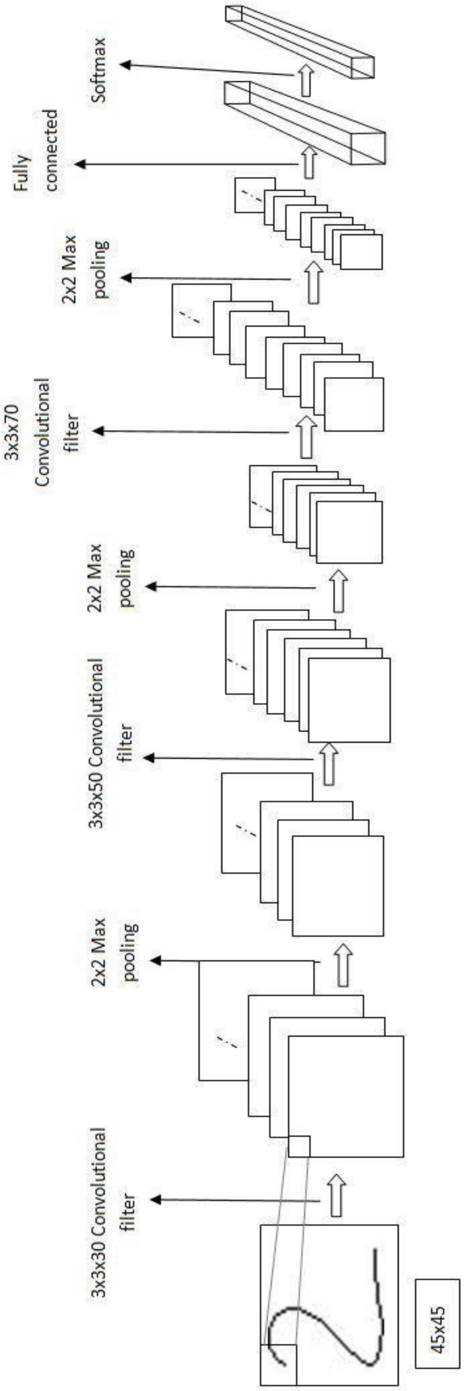

Fig 3: SpNet architecture for Handwritten Mathematical Symbol

- Training: The HMS dataset is used in training the symbols after preprocessing. Train the CNN classifier for 83 different classes.

  To train the classifier 100 epochs are used with learning rate 1.0. The values calculated for each class are stored that can be retrieved while testing.

- Testing: The segmented symbol of the HME is individually tested using the trained CNN classifier. The Softmax layer comes into use here. The Softmax layer works on the principle, as in

$$P(y=c|x) = \frac{e^{x^T w_c + b_c}}{\sum_j e^{x^T w_j + b_j} } \quad (1)$$

In (1), y is the class c to be predicted, x is the vector of the symbol to be recognized after passing the fully connected layer, w is the weights and b is the bias for j different classes[12].

IV. CONCLUSION AND FUTURE SCOPE

We have proposed a system that will recognize offline HME. The HME will be converted to latex. This system will work best for isolated symbols. In future, merged or connected or joined symbols will be segmented to give better recognition results.

REFERENCES

[1] Irwansyah Ramadhan, Bedy Purnama, Said Al Faraby, “Convolutional Neural Networks Applied to Handwritten Mathematical Symbols Classification” 4th International Conference on Information and Communication Technologies (ICoICT), 2015IEEE

[2] Sagar Shinde, Rajendra Waghulade, “Handwritten Mathematical Expression Recognition using Back Propagation Artificial Neural Network” Communication on Applied Electronics (CAE) , Foundation of Computer Science FCS, New York, USA Vol.4, No. 7, March 2016, pp. 1-6.

[3] Ahmad-Montaser Awal, Harold Mouchère and Christian Viard-Gaudin, “The problem of Handwritten Mathematical Expression Recognition” 2010 IEEE 12th International Conference on Frontiers in Handwriting Recognition, pp. 646-651

[4] R.Padmapriya, S. Karpagavalli, “Offline Handwritten Mathematical Expression Recognition” International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 1, January 2016, pp. 52-59.

[5] Yassine Chajri, Abdelkrim Maarir, Belaid Bouikhalene, “A comparative study of Handwritten Mathematical Symbols recognition” IEEE 13th International Conference Computer Graphics, Imaging and Visualization, 2016, pp. 448-451.

[6] Manisha Bharambe “Recognition of Offline Handwritten Mathematical Expressions” International Journal of Computer Applications, National conference on Digital Image and Signal Processing, DISP 2015, pp. 35-39

[7] Abdalla Mohamed Hambal, Dr. Zhijun Pei, Faustini Libent Ishabailu, “Image Noise Reduction and Filtering Techniques” International Journal of Science and Research (IJSR) Volume 6 Issue 3, March 2017 pp. 2033-2038

[8] Jyotsna, Shivani Chauhan, Ekta Sharma, Amit Doegar, “Binarization Techniques for Degraded Document Images - a review” 2016 5th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), 7-9 Sep. 2016, pp163-166.

[9] Syrine Ben Driss, M Soua, Rostom Kachouri, Mohamed Akil, “A comparison study between MLP and Convolutional Neural Network models for character recognition” SPIE Conference on Real-Time Image and Video Processing, Apr 2017, Anaheim, CA, United States.

[10] Xue-Dong Tian, Hai-Yan Li, Xin-Fu Li and Li-Ping Zhang, “Research on Symbol Recognition for Mathematical Expressions” 1st International Conference on Innovative Computing, Information and Control (ICICIC'06), 2006 IEEE, Sept. 2006

[11] Masayuki Okamoto, Bin Miao, “Recognition of Mathematical Expressions by Using the Layout Structures of Symbols” Technical Report, Institute of Information Engineering, Shinshu University, pp. 242-250.

[12] Moacir A. Ponti, Leonardo S. F. Ribeiro, Tiago S. Nazare, Tu Bui, John Collomosse, “Everything you wanted to know about Deep Learning for Computer Vision but were afraid to ask” 30th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials, 2017 IEEE, pp. 17-41.