Handwritten Equation Solver Using Convolutional Neural Network

Shuham Aggarwal
Department of Computer Science and Engineering
National Institute of Technology
Delhi , India
181210052@nitdelhi.ac.in

***Abstract*— Handwritten equation solving is a tricky job in the area of image processing. Among all the problem handwritten mathematical expression, recognition has been one of the complicated issue in the area of computer vision research. Segmentation and classification of specific character has been making the task more difficult. In this report ,a group of handwritten quadratic equation as well as a single quadratic equation are considered , the purpose being to recognize and make a solution for those equation For classification of specific character we apply Convolutional Neural Network. Each of the correct detection, character string operation is used for the solution of the equation. Finally , the stimulations and the results are shown below in the result section.**

Keywords-Data Preprocessing; Convolutional neural Network; Projection analysis; Quadratics; Connected component .

I. INTRODUCTION

Due to the swift improvement in computer technology and internet technology, most of the documents, books and literatures in the area of computer science as well as others are increasingly becoming digitalized. Mathematics is broadly used in almost all areas of science, such as physics, engineering, medicine, economics, etc. Digital document analysis and understanding is the major research concern today. For the recognition of English characters and numbers in electronic books OCR (optical character recognition) can attain higher recognition exactness. Handwritten mathematical expression recognition is still a most challenging job to do in the area of computer vision. Due to the two dimensional nesting assembly and different sizes, the correction rate of symbol segmentation and recognition still cannot achieved its actual requirements. The primary task for the recognition of mathematical expression is to segment the character and then classify those character. An operation that seeks to decompose an image of a sequence of characters into sub images of individual symbols is Character segmentation . Histogram grounded projection analysis is a well- known method and is commonly used in the segmentation of line and character from the image and also we can practice it at some detection stage. Convolutional neural network (CNN) is one of the mostly used classification model in computer vision area. In the last few years, deep Convolutional Neural Network(CNN) leaning has proved the outstanding performance in the field of image classification, machine learning and pattern recognition. Above all existing model, CNN is one of the most popular models and has been providing the state-of-the-art recognition accuracy on object recognition , segmentation , human activity analysis ,image super resolution , object detection , scene understanding , tracking , and image captioning . For the task of image classification CNN is outperforms above all the previous classification method . CNN extract feature from the image by a series of operations. In this research we mainly focus on the recognition of quadratic equation and after successful detection of quadratics we apply character string operation for the solution of those equation. The main contribution of recognition of the mathematical equation in this paper are summarized here. In this research we consider a single or a series of handwritten quadratic equation in the form $ay^2+by+c=0$ with any combination of it's part. Our job is to recognize the handwritten quadratics form the image and for each successful detection finding the solution of that equation
. In the case of segmenting each quadratics from the original image the square term is sometimes slightly placed out of the other parts of the equations horizontal projection curve. We solve this problem by compact horizontal projection. Another difficulty exists for applying connected component in the case of mathematical symbol such as for '=', which is a combination of two connected component. We solve this problem by using combined connected component analysis method. Feature extraction of each individual character is most difficult part for any handwritten due to it's different shape and structure. To solve this problem we apply convolutional neural network which doesn't required any predefined feature for classification of specific character. We use smallest number of hidden layer in order to reduce the training time with tolerable error rate. In section II literature review of the related work is described and in section three represents the proposed method of this work. The following section describe the experimental results and finally describe the summary of this work.

II. LITERATURE REVIEW

Several work has been done so far for handwritten character recognition. A scheme for mathematical expression recognition (MER) based on SVM and projection histogram for simple ME's have proposed in . which is a part of

offline handwritten expression recognition. This paper basically focus on numerous technique used for feature extraction and recognition. An effective and robust system for recognition of printed and handwritten mathematical characters have been proposed by Zanibbi (2002). Which evaluates an expression by using three successive passes i.e., the layout pass, lexical pass and operator tree. This tree manipulation which used in each pass can be represented by tree transformation. Bage .(2013) has proposed recognition system for offline handwritten mathematical symbol. For feature extraction, shape of the character is considered. The proposed system is based on relative study of feature extraction methods. Recognition has been carried out via SVM. Ramteke et.al. (2012) have transformed scientific and engineering documents into electronic form by bearing in mind of printed mathematical symbols and expressions. Features are computed using centroid of the image here and for segmentation bounding box technique is used. Classification has been carried out using neural network approach which recognition rate of 90%. Ahmad-Montaser Awal et al(2010) discussed some issues related to the problem of online mathematical expression recognition. In the development of this system, they have adapted object oriented methodology to describe the data abstraction for the hierarchical structure of the mathematical expression is given in the form of the expression tree. Pradeep et. al. (2010) have proposed diagonal feature extraction technique for the handwritten character by using feed-forward neural network algorithm. This technique uses diagonal, horizontal and vertical features for the classification purpose. A few paper are available online for recognition of mathematical expression using convolutional neural network(CNN).In recent years, Convolutional Neural Network (CNN) has made a series of revolution research results in the fields of image classification and detection. The past few years a branch of artificial neural networks called deep learning has shown great potential in solving classification problems. The field of deep learning started gaining popularity in 2012 when Alex et al. demonstrated their deep network architecture named AlexNet for image classification which outperformed.

III. Methodology Proposed

In our proposed method at first noise from the original input image is removed by applying binarization to it. After that we use compact horizontal projection for segmenting each line of equation from the input image. Then we consider each part of the segmented image as a full image for further process. For each line of equation image we then find specific characters in the form of connected component. Each segmented character is then providing as input to the convolutional neural network model for classification of the character. The resulting character that is the output of CNN is then used for making a character string which is similar to the originalequation.

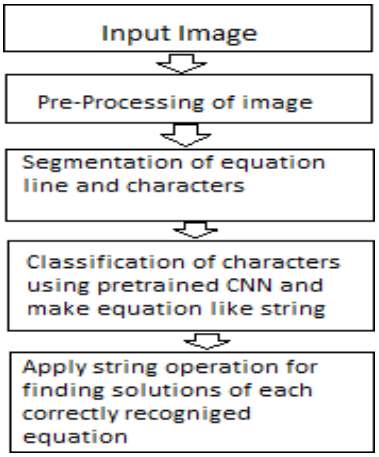


Figure 1. Workflow diagram of proposed method

For each correct detection we finally find the solution of the quadratics. In Figure 1 workflow diagram of our proposed method is shown.

A. Dataset preparation

Preparation of the dataset is the fundamental concern for this work. Characters such as English digit, alphabet and mathematical symbol all can be well defined by their edges. For this reason we first prepare dataset as the most precedence given to its edges that is enlighten the edges. We prepare some dataset by ourself and we also used some modified version of NIST dataset which is similar as the popular MINIST dataset for digit. For each category we use 2000 data item for the training of the network. And most of the case our network training we gained more than 98.5% training accuracy. Image size we used in dataset is 32×32 gray level image.

B. Pre-Processing

Pre-processing of the input image is the procedure which encompasses changes and modifications to the image to make it fit for recognition. The following techniques may be used for image enhancement.

1) Conversion of RGB to Gray-Scale

First of all this coloured image is transformed into a typical gray-scale image and is represented through a single matrix because the detection of characters on a coloured image is more challenging than on a gray-scale image. IF the gray bitmap Y and color bitmap is R,G and B then the formula we used is:

$$Y = 0.299R + 0.587G + 0.114B \quad (1)$$

2) Binarization

Binarization is the procedure of choosing a threshold value is for adaptation of pixel values into 0's and 1's. In This research for horizontal projection calculation 1's represent the black pixels while 0's characterize the white pixels. The threshold choice of binarization can approve two ways: overall threshold and partial threshold. Otsu's method is an overall threshold method, basing on image statistical characteristics. This

method can make the computer select automatically threshold.

C. Segmentation

Identifying the objects or other significant information in digital images segmentation is mostly used in image processing and computer vision application. Which is the process of dividing an image into multiple parts[19]. The segmentation involves two main steps in our proposed method.

a) Equation line Segmentation

Equation segmentation is the split-up of the different lines of characters existing in the image. Each line is well-defined by a minimum vertical gap between the characters existing on a line and on the line overhead and below it. This gap can be used for the detection and separation of not the same lines of characters. For the solution of square term in quadratics as mentioned above we use compact horizontal projection. Combining of two curve into one that is representing one equation line with two curve is done by comparing with a threshold value that is obtained by averaging the maximum value of each horizontal projection vector. If any curve's maximum value is less than the threshold this curve is then considered as a part of the next projection curve.

b) Character Segmentation

Character segmentation is a procedure that look for to decompose an image of a series of characters into sub images of individual symbols[19]. In our proposed method we use connected component analysis method for the segmentation of specific character from the image. Problem arises at the point of extracting of math symbol '=' which is a combination of two connected component. If two consecutive component has the minimum height and are in a same horizontal direction we can consider this two into a single component. If an outside rectangle of two symbols region with connected regions is set to $X(x_1, y_1, x_2, y_2)$ and $Y(x_3, y_3, x_4, y_4)$. Now if this two symbol X and Y has minimum height and belong to a whole horizontal direction then the combination of X and Y produce a third component Z that is our desired. The position of Z can be obtained as

Where the even x and y variable represent the right and top position of the rectangle respectively, odd x and y variable represent respectively the left and bottom position of the rectangle.

D. Classification model

In this research convolutional neural network is used as a classification model , 30×30 gray scale image is used as input of the CNN input layer for the training section and also for the testing section. 5×5 filter is used at the convolutional layer after the convolution of the input image with the filter for each input image a 28×28 feature vector is produced. It is concord to apply a nonlinear layer (or activation layer) instantly after conv layer,. The persistence of this layer is to introduce nonlinearity to a system. To performs down-sampling by separating the input into rectangular pooling regions, and figuring the maximum of each region a max pooling layer is used. Pooling layers perform down-sampling operations. Calculation of the *output size, O* of a pooling layer with *input size ,I* Pooling *filter size,F Padding ,P* and *stide ,S* has done by

$$O = (I - F + 2\times P) / S + 1 \qquad (2)$$

In our proposed method we use 2×2 pool size and the output of the pooling layer is 14×14. Random selection of input elements to zero with a given probability is done in a dropout layer. It is a simple way to prevent the overfitting in the neural network. Overfitting is a stern difficulty in such networks. Large networks are also sluggish to use, making it more challenging to contract with overfitting by combining the predictions of many different large neural nets at the test time. The technique for addressing this problem is Dropout . The core idea is to randomly drop units (along with their connections) from the neural network during training . In our proposed method we use the probability 0.1 at the dropout layer during the training. After the convolutional and down-sampling layers one or more fully connected layers is used. The layer in which the neurons associate to all the neurons in the previous layer is fully connected layer. Fully connected layer combines all the features learned by the previous layers through the image to recognize the larger patterns. To classify the images the last fully connected layer combines the features. Hence, the Output parameter in the last fully

connected layer is equal to the number of classes in the target data. In our work, the output size is 14, corresponding to the 14 classes. An activation function softmax which normalizes the output of the fully connected layer is also used in the convolutional neural network model. The softmax layer output comprises of positive numbers that sum to one, in the nest that can be used as the classification probabilities at the classification layer. At the final layer ,classification layer uses the probabilities given by the softmax activation function to find the input image classes and find the loss by comparing it with the pre assigned ground truth classes.

E. Solution of Quadratics

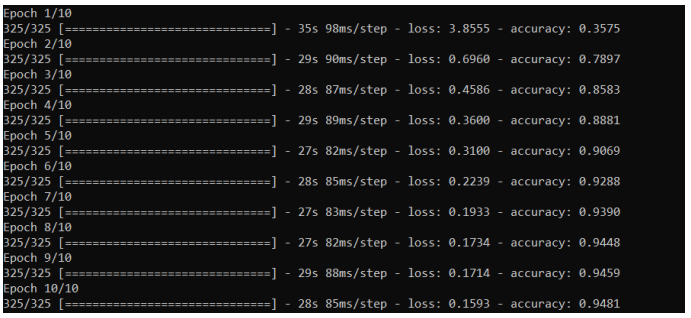
After successful detection of the quadratics we organized the character string in the form of *'+ay^2+by+c=0'*. Now for solving this quadratics our job is to find the value of *a, b, c*. where *a,b,c* can be a single digit number or a multiple digit number. To find the value of a, b, c we apply a character string operation algorithm as *Algorithm 2*. After finding the parameter of the quadratics we apply the formula for the solution of the quadratic equation.

IV. RESULT

After training of the network we use more than 10000 test image for the recognition of the quadratics. we use image with one quadratics ,and more than one quadratics. we consider each of the quadratics character are separating with each other in the test image. For applying the traditional algorithm we face some problem and we modify those part in our system for getting correct solution. Firstly we train the network after completing the training we apply each test image for recognition. For each of the quadratics recognition process in the case of correct recognition we will provide a command for finding a solution otherwise providing a command for next quadratic recognition. For the segmentation of each line of quadratics we use horizontal projection at each row of the experimental image.

Most complex part is the classification part generally several character are distinct in shape. But for the different handwritten strategy some the character creates ambiguity to each other categories. Here we use convolutional neural network for the solution of critical classification problem.

After that the noise free image , expected outcome is also shown. For running the model , the below mentioned image is there which shows the accuracy of the model and the impact the model had after each epoch.

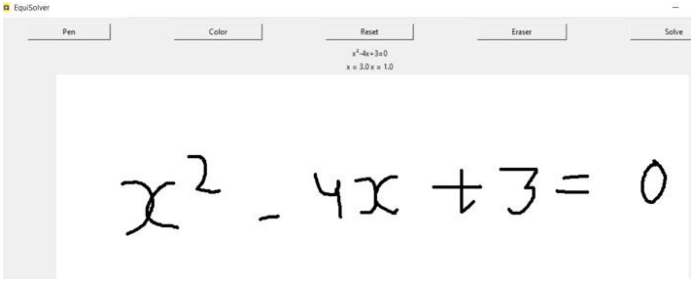


The accuracy now comes out to be 0.9481. The other thing it shows that accuracy improves after each epoch. This seems to be much better than the traditional CNN method which we have been using ever since. Traditional method with higher hidden layerthough sometimes provide high accuracy rate for the classification task but requires lots of training time as shown in the Figure 5. For the case of 100 training data elapsed time for training is several times larger than our proposed network layer. Handwritten character are sometime even becomes difficult for human to recognize due to several reason ,such as different people writes character in different shapes and style, different writing tools are used by people that makes the character sometimes more complex, another most important reason is writing style of single character by a man is most often different from a group of words or specially in mathematical equation. For classification part the actual performance, P can be measured by

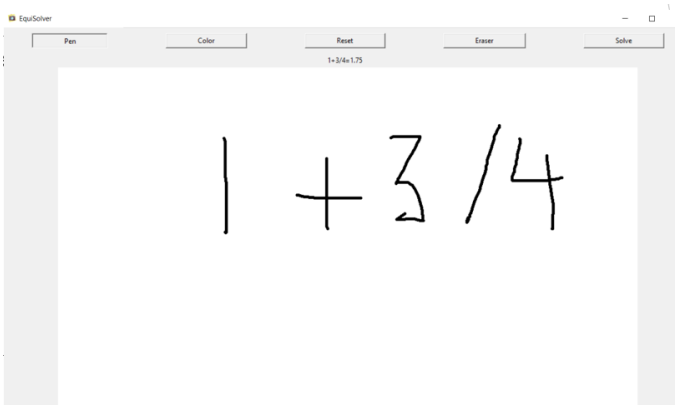
$$P = \frac{Y}{X-m} \times 100\%$$

Where X is the number of image character given as input for classification and Y is actual number for correct classification and m is the number of character those are conflict shape with other character that even can't classified by human.

The other thing we below is how the equisolver is actually working and how after the input it shows the output.

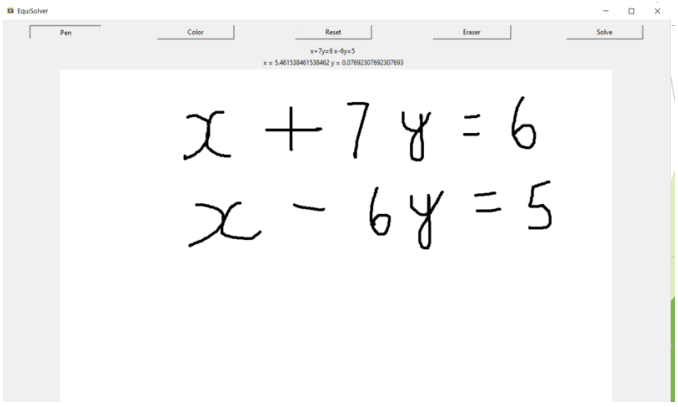


As we can see here the equation is easily solved and the result is showing just above the equation given as an input. Other snapshots of the simulation results are also mentioned below.



This is the result which we have achieved by using very basic and simple mathematic . An interesting thing to note here is that fractions are being represented as decimals in the given output. As we have seen in the above equation , ¾ is being shown as 0.75

Another interesting thing which is to be seen is how it solves the 2 variable equation. Lets see that in the snapshot provided below:



As we can see , the equation is being solved , but also the solution is coming in the recurring form , it is giving us the recurring solution up till it can actually provide us with a solution.

The most important thing here to note is that we have used the polynomial equation of both quadratic and linear form so that It is very flexible for the user.

V.CONCLUSION

In this paper we mainly focused on recognizing handwritten mathematical quadratics. we consider a single quadratics and also series of quadratics for the recognition. Projection analysis specially compact horizontal projection is used for the segmentation of each line of quadratics. Connected component which has very high success rate is used for character segmentation [19]. Improved version of connected component is used for the symbol like '=' detection which is a single symbol combined with two distinct connected component .Feature extraction the most complicated part of classification. Moreover with some predefined feature about handwritten it is difficult to recognize handwritten. Convolutional Neural Network the most powerful classification model is used in the classification part[24]. Once successful recognition of the quadratics in any combination we further process the detected equation for finding the solution of the quadratics. In this part we apply an string operation algorithm for finding the value of a, b, c of each quadratics in the form of ay 2+by+c=0. Finally we obtained a state of art performance in the detection phases and also in the solution phases. In future we will try to improve the accuracy and try to make the system workable for multiple mathematical formula simultaneously. Through the use of convolutional neural network we will try to make a vision based mathematical formula dictionary in future.

VI . FUTURE SCOPE

Deploying the project on the internet so that it can be used by the general public.

Being able to solve more complicated mathematical equations such as integration and differentiation of the given input.

If possible have the facility of taking the input from the voice modulation so that it would be useful for people having certain disabilities.

Integration of the above working model in the educational sector.

The whole idea of the project is to create an Alexa like model using different tensorflow techniques so that we would be able to have an equation solver machine which would be able to solve many different equation on its own , just on the command of the voice modulation provided by the user .

VII. REFERENCES

[1] Widad Jakjoud , Azzeddine Lazrek 'Segmentation method of offline mathematical symbols', International Conference on Multimedia Computing and Systems (ICMCS), 2011

[2] Zouaoui Abderaouf 'license plate character segmentation based on horizontal projection and connected component analysis' 2014 world symposium on computer applications & research (wscar).

[3] Catherine Lu Karanveer Mohan.' Recognition of Online Handwritten Mathematical Expressions Using Convolutional Neural Networks' cs231n project report stanford 2015.

[4]https://www.researchgate.net/publication/252012021_Handwritten_mathematical_formula_recognition_using_a_statistical_approach

[5]https://www.researchgate.net/publication/326710549_Recognition_and_Solution_for_Handwritten_Equation_Using_Convolutional_Neural_Network