Handwritten Math Problem Solver Using Convolutional Neural Network

Mohini Lokhande, Kirti Murudkar, Arbaaz Nadaf, Niyati Shah

Department of Computer Engineering Trinity College of Engineering and Research, Pune

Abstract: The main objective of this project to develop an Website which calculates the presents a Handwritten Equation Solver trained by handwritten digits and mathematical symbol using Convolutional Neural Network with some image processing techniques to achieve a decent accuracy of 98.46%.

Keywords: Machine learning, Python, Handwritten Equation, Simplification, Preprocessing, Segmentation, Implementation, Recognition, CNN, Polynomial expressions, Image processing.

I. INTRODUCTION

Polynomials are used to graph the curves of roller coasters and bridges, in economics to do cost analyses, to find the motion of a particle under the influence of gravity, to forecast sales trends, to develop profit margins. Complexities involved in Human Computer Interaction (HCI) are reducing at an extraordinary rate and computers are gradually catching up with the nuances, irregularities and imprecisions of the real world to enable an easier interaction. Tremendous amount of research has gone into identifying and classifying some irregularities of the real world including handwriting recognition, natural language processing, face detection and so on.

II. LITERARTURE SURVEY

Many various papers can be found on handwritten character segmentation. Some scheme is also working on mathematical expressions recognition "MER". Like "Using SVM and projection histogram identification of ME". Some are for offline printed mathematical expressions and recognition mentioned by Zanibbi et. al. (2002), "Recognition of printed mathematical symbols", "Using SVM Mathematical symbols identification", "Recognition of online mathematical symbols using template matching distance", "Offline Handwritten Mathematical symbols recognition using character geometry". This all proposed method for recognition of symbols and segmentation using various actions. Recognition has been also done by CNN based models for mathematical symbols and character. Some discussion about concerned to the labyrinth of online mathematical expression recognition. Majority of those papers concentrated in the recognition scheme of the equations. In this approach, the main hub is on handwritten polynomial equation simplify. As there is not any proper work that can successfully handle the problem for handwritten images. The past work only focuses on the recognition level of the equation and not concentrated on multivariable identification in the equation to solve the problem. This work motivated to handle a multivariable segmentation and recognition efficiently and show the result is and every variable in the equation.

After preprocessing, segmentation, and recognition of the input image, generating a string equation from that image and simplify that expression is the main target.

III. EXISTING SYSTEM

SolveIt : An Application for Automated Recognition and Processing of Handwritten Mathematical Equations(2018).
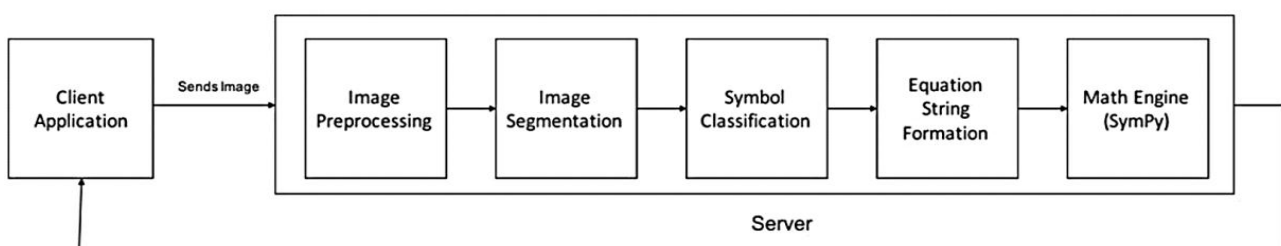

Fig 1: Pipeline style architecture of the application

IV. PROPOSED METHODOLOGY

In this activity to identify the handwritten equation and to simplify it, we have used many different steps right from taking the input image to generating the output.
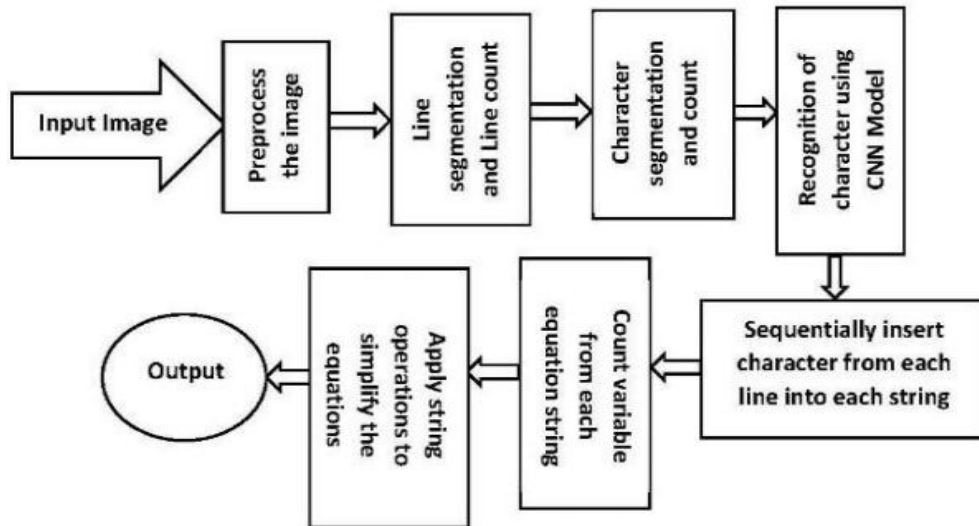

Fig 2: Workflow Model

4 A. Pre processing

In pre processing the input image is transformed to make it of suitable quality for recognition purpose. At first convert the image into grayscale image, because in original images, it becomes difficult to identify the characters. Then the image is converted into binary image. That is all pixels are converted to 0-255. This reduces the computational time and eliminates the unnecessary pixels, hence making the image more compact.
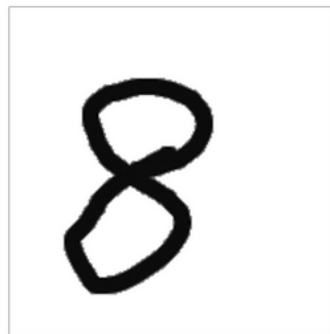

Fig 3.a Input Image   Fig 3.b After pre processing

4 B. Line Segmentation

If there are multiple lines in the input image, each line needs to be separated. By applying line segmentation, each line can separated from one another.
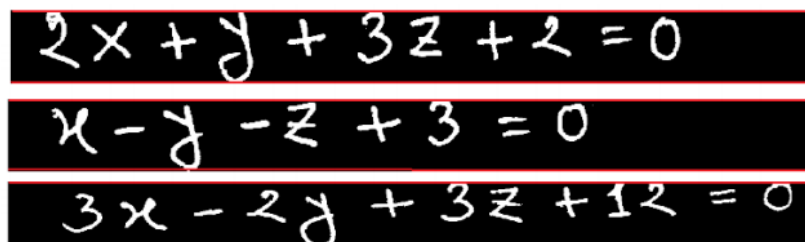

Fig 4 Segmented lines

4 C. Character Segmentation

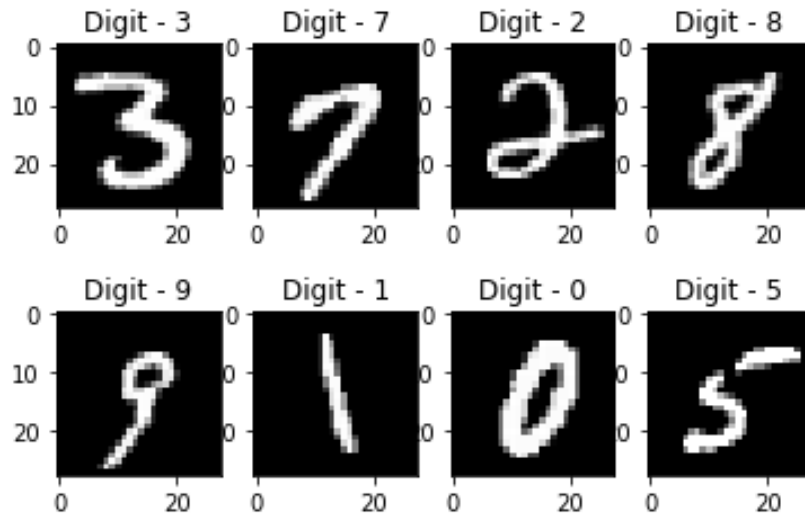Character segmentation is nothing but separating each character from the equation.

Fig 5 Character segmentation

4 D Dataset preparation

For training, this proposed model used MNIST for handwritten digits. Which include 0-9 numerals and 3 mathematical symbols and operator. For better computational purposes all images converted into the 28×28-pixel size and kept in Comma-separated values (CSV) format.

Fig 6 MNIST dataset

V. TECHNOLOGIES USED

A] Frontend Application: The front end webpage is created using HTML ,CSS and JavaScript with VisualStudio used as the IDE.

B] Backend Application: Python in used in the backend to perform all other operations. Anaconda jyupiter is used as the IDE.

VI. EXPERIMENTAL RESULTS

The figure below shows the accuracy of the model developed which is 97%. That means that 97% of the times the image is recognized correctly.
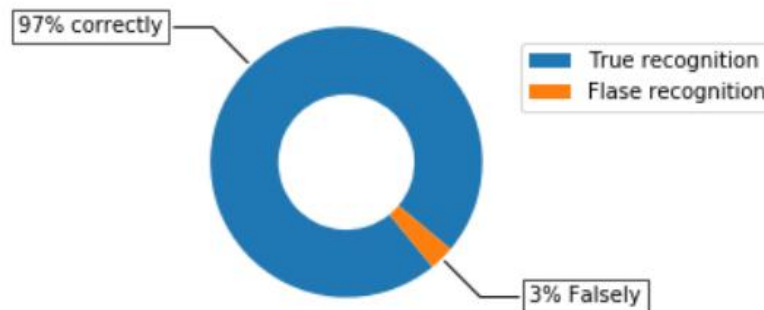

Fig 7 Recognition Rate

VII. CONCLUSION AND FUTURE SCOPE

To simplify the math, the main task done in the feature extraction from the image and recognition with the help of the CNN model. If the CNN model classifies correctly all of the segmented images then this will be generated the correct list of equations. Which will be better forthe simplification part. This is a successful representation of the state of the art. Any person can easily use this as the process is easy and does not require any advance knowledge of the technology.

In future days the main focus will be to try to raise the precision level and build a segmentation system that can successfully segment two connected digits, and also increase the performance level of the dataset.

ACKNOWELEDGEMENT

REFERENCES

[1] Fuad Hasan, Shifat Nayme Shuvo et al, "Bangla Continuous Handwriting Character and Digit Recognition Using CNN", 7th international Conference on Innovations in Computer Science & Engineering (ICICSE 2019), vol 103, springer, Singapore, pp. 555- 563.
[2] Jalal Uddin Mahmud, Mohammed Feroz Raihan and Chowdhury Mofizur Rahman, "A Complete OCR System for Continuous Bangla Characters", IEEE TENCON-2003: Proceedings of the Conference on Convergent Technologies for the Asia Pacific, 2003.
[3] Gharde, S. S., Baviskar, P. V., & Adhiya, K. P. (2013) Identification of Handwritten Simple Mathematical Equation Based on SVM and Projection Histogram. International Journal of Soft Computing and Engineering (IJSCE),3(2),425-429
[4] Francisco Álvaro and Joan Andreu Sánchez, "Comparing Several Techniques for Offline Recognition of Printed Mathematical Symbols", IEEE,ICPR 2010.
[5] Christopher Malon, Masakazu Suzuki, and Seiichi Uchida, "Support Vector Machines for Mathematical Symbol Recognition", Technical Report of IEICE.
[6] Fotini Simistira, Vassilis Katsouros and George Carayannis, "A Template Matching Distance for Recognition of On-Line Mathematical Symbols", Institute for Language and Speech Processing of Athena - Research and Innovation Center in ICKT, Athens, Greece.
[7] Dipak D. Bage et. al,"A new approach for recognizing offline handwritten mathematical symbols using character geometry", International Journal of Innovative Research in Science, Engineering and Technology Vol. 2, Issue 7, July 2013, ISSN: 2319-8753
[8] Catherine Lu Karanveer Mohan.' Recognition of Online Handwritten Mathematical Expressions Using Convolutional Neural Networks' cs231n project report stanford 2015.
[9] Sanjay S. Gharde, Vidya A. Nemade and K. P. Adhiya,"Evaluation of Feature Extraction and Classification Techniques on Special Symbols", IJSER, Volume 3, Issue 4, 2012.
[10] Ahmad-Montaser Awal , Harold Mouchère, Christian Viard- Gaudin.'Towards Handwritten Mathematical Expression Recognition' 2009 10th International Conference on Document Analysis and Recognition.
[11] G. Louloudisa *, B.Gatosb,1, I.Pratikakisb,1, C.Halatsisa (2009). Text line and word segmentation of handwritten documents.
[12] Richard G. Casey, Eric Lecolinet " A survey of methods and strategies in character segmentation'1996
[13] Shifat Nayme Shuvo,Fuad Hasan,Mohi Uddin Ahmed,Syed Akhter Hossain,Sheikh Abujar,"MathNET: Using CNN Bangla Handwritten Digit, Mathematical Symbols and Trigonometric.