Procedure International Journal of Science and Technology

(International Open Access, Peer-reviewed & Refereed Journal)

(Multidisciplinary, Monthly, Multilanguage)

ISSN : 2584-2617 (Online)

Volume- 1, Issue- 2, January 2024

Website- www.pijst.com

Image Formation: Mathematical Expression of the Principle

Dr. Ankit Maurya

Assistant Professor, Department of Mathematics, S.L.B.S. Degree College, Gonda, U.P., Email id- drakslbsdc@gmail.com

Abstract-

The procedure for transferring written or printed text from an image file or scanned document into a machine-readable format for data processing applications like editing is known as optical character recognition, or OCR. Although handwriting has been a means of communication for millennia, the advent of contemporary computers has made it simpler. Although most individuals are used to inputting words on a keyboard, mathematical expressions and formulae require additional add-ons to be loaded in the word processor. The procedure might take a long time and be tiresome. Consequently, this study suggests an alternate approach that turns handwritten mathematical formulas into text that can be read by computers. Segmentation is accomplished using horizontal and vertical projection, while character recognition is accomplished by convolutional neural networks, which improve recognition accuracy. The suggested technique was able to extract and recognize each character written in addition to segmenting out handwritten mathematical equations from lined paper. It was then possible to correctly convert the handwritten equation to a digital one.

Key-words: Image formation, equations, mathematical expression, Segmentation, optical character recognition.

Introduction-

The mathematical expression for the principle of image formation can be abstracted to describe the general process across various imaging systems,

encompassing optics, photography, and computer vision. Given that handwritten mathematical expression recognition has been the subject of in-depth research for over 60 years, translating mathematical equations into textual format is not a recent development[1]. Since it necessitates mathematical computation, it is extensively employed in the fields of science and engineering. Additionally, there is a need to digitize historical mathematical formulas that were handwritten on paper. The segmentation of the mathematical equation is a barrier in this area since handwritten characters might include a large number of different symbols, some of which may have varying sizes. When utilizing optical character recognition (OCR), there's a potential that the image is blurry or that the character has been overlapped [2].

LaTeX is a typesetting system that is widely used for the creation of documents containing mathematical and scientific content. While it's not a coding tool in the traditional sense, it involves writing markup code to define the structure and formatting of a document. Learning LaTeX would be beneficial for most academics who must constantly write down mathematical equations, but it might not be something that students or other non-researchers are interested in learning. Handwritten mathematical equations (HME) are divided into two categories: online and offline. Since the stroke is captured and sent to the recognition system when the equation is typed on a tablet, the field of study and accuracy of character identification in online handwritten mathematical equations is greater. Handwritten mathematical equations offline are classified in the second category. Before the picture is transferred to the identification step in this approach, it is either scanned or captured using a camera. Since handwritten character pictures can be deformed and there might be differences in the size and scale of symbols, offline HME is recognized to be more demanding and difficult to recognize than online HME. As a result, less study has been done on this approach [3]. Offline HME has more advantages even if online HME is simpler and has greater recognition accuracy. On addition, mathematical equations written on non-digitized papers can also be converted using an offline HME. [4]. on addition to requiring less work on documents, this technique may be used to convert paper documents into electronic forms and, if necessary, transport them across devices [5].

The technique of breaking down a mathematical equation into distinct character equations is known as segmentation. Since every person's handwriting is unique, the segmentation step presents the biggest obstacle for offline identification, along with issues with overlapping letters and scale and size variations [2]. When the mathematical equation is handwritten clearly and with appropriate spacing between each word, a greater segmentation rate may be attained. The Onyshchak [6] created distinct characteristics, such as the equal sign "=" in place of two negative signs "-," to recognize characters with different contours and combine them. This document lacks a distinguishing subscript feature. The outcome demonstrates that the output could distinguish between the characters "i" and "j." Moreover, Hossain et al. [7] segment quadratic equations using a technique known as compact horizontal projection, in which the projection is limited to the y-axis. Researchers Drsouza et al. [8] used recursive projection profiling in their study. It uses a similar methodology to the earlier studies, but it now includes vertical projection. This helps to divide up each character. This article will use CNN for the character

categorization using MATLAB's neural network toolbox as CNN has the best success rate [9]. With the use of pre-defined layers like AlexNet, ResNet, or VGGNet, MATLAB offers a support package for transfer learning that would yield excellent accuracy [10]. Transfer learning is not intended for use with mathematical equations; rather, it is meant to be flexible. The network requires processing power and has layers that are not required for categorizing mathematical problems

In conclusion, typing down mathematical equations on a keyboard is a laborious process that takes longer to input each character of the equation than writing it down on paper if you are unfamiliar with LaTeX or MathML. Thus, the purpose of this study is to use CNN as the classification algorithm that is intended to be appropriate for classifying HME in order to translate mathematical equations into text based on segmentation utilizing Horizontal Projection Cutting (HPC) and Vertical Projection Cutting (VPC). The objective is to achieve high classification accuracy and improved segmentation accuracy with few CNN layers and training data, including the ability to separate overlapping characters.

Methodology-

The research project's general plan is depicted in Figure 1. This project aims to translate mathematical equations written by hand on paper into a text-based format. Therefore, having a handwritten image of a mathematical equation that contains integrals or derivatives is the first step. Following that, image processing will be applied to the photo. This stage is crucial because effective segmentation requires a black-and-white, noise-free picture. Following the pre-processing phase, segmentation will be applied to the picture. This procedure turns the picture into a distinct character. The goal is to create an algorithm that can determine if a segmented picture has a single character or whether it needs to be further combined into two smaller characters. For instance, the equals ("=") character is made up of two minus ("-") characters. It is necessary for the segmented algorithm to identify superscripts like x2. Furthermore, employing convolution, pooling, fully connected, and Softmax layers, Convolution Neural Networks (CNNs) classify mathematical formulae by identifying the key characteristics of each symbol. An appropriate quantity of training and test datasets should be made available to the network in order to train it and achieve good optical character recognition. Test datasets will be employed to assess the CNN architecture's accuracy. The software will transform the symbol into a text-based format after the system has recognized the character. This project is quantitative in nature since the number of right answers will be used to determine the study's correctness. MATLAB R2021b is the platform of choice for this research since it has deep learning and image processing toolboxes that would help with the technique.

Segmentation

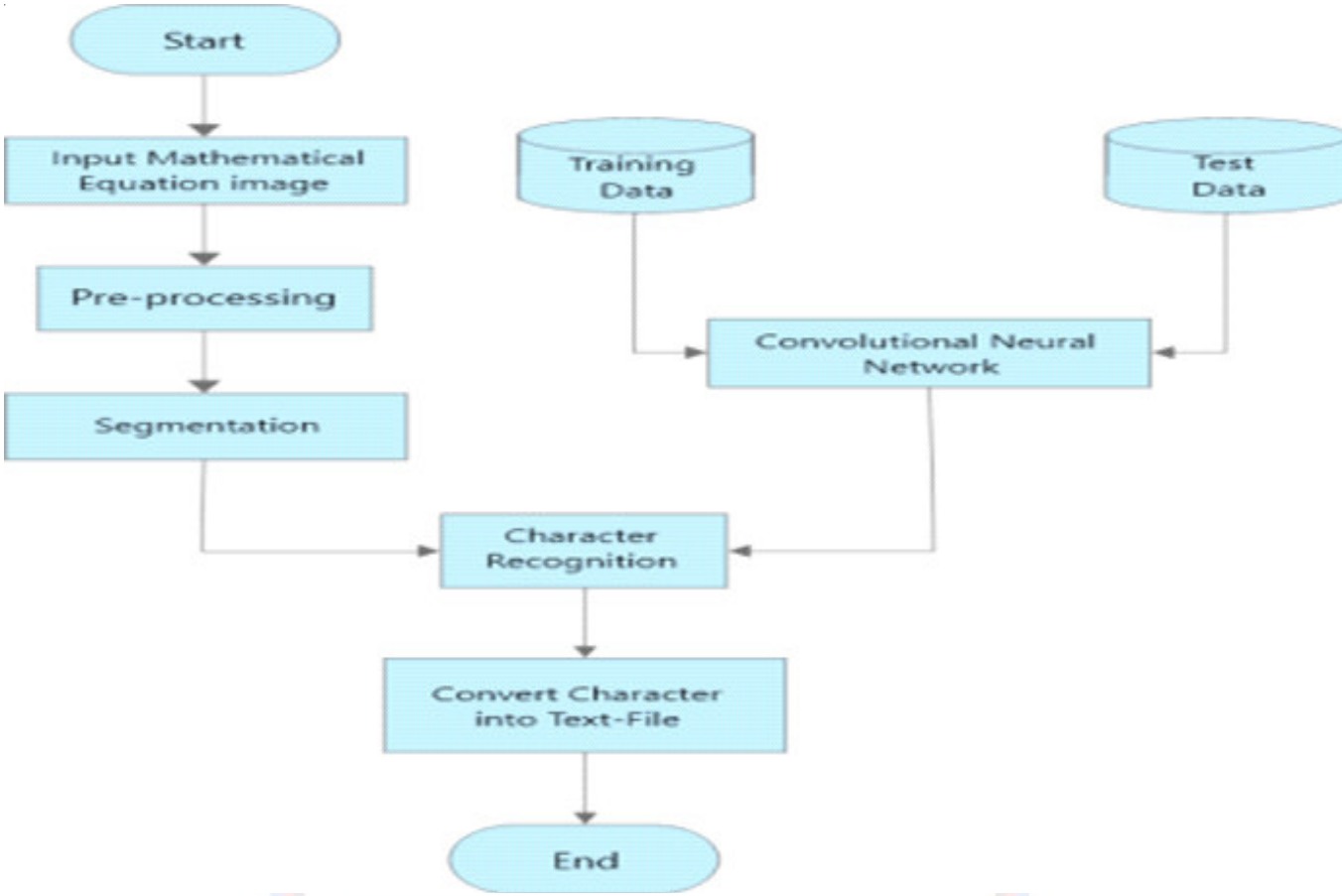


Figure 1. Overall Project

The segmentation approach must guarantee that the mathematical problem is segmented accurately in order to get good results. Because they demand less processing and yield promising results, the segmentation in this work will be done using the Horizontal Projection Cutting (HPC) and Vertical Projection Cutting (VPC) methods. When there are overlapping characters in the mathematical equation, the projection cutting approach has a drawback. A method has been created to circumvent the constraint. The segmentation flowchart is displayed in Figure 2. To get rid of lines and noise, go to the pre-processing step. The image will be filtered, then binaries and converted to gray scale. To make the segmentation process easier, the outcome will simply display the equation on a white backdrop. To find the equation's gaps, a horizontal projection of the binaries picture will be made. The current cell array photos' horizontal crop is where VPC is performed. At this point, a few of the characters are prepared for filing and CNN classification. A linked component will determine whether the cropped picture has one stroke in order to indicate that it is ready to be stored. This demonstrates that the reduced images only 1 stroke and was not capable of being divided further. However, some letters, like "=," contain two strokes and shouldn't be further divided to create "-." An picture is "=" and should be saved to a file rather than being further segmented if it includes more than one linked component after cropping and two peaks that

fall under the threshold value after applying HPC. By ensuring that there are no 0 values on the HPC, the algorithm will determine if cropped photos that fall outside of the parameters have overlapping characters. The label matrix of the linked components is used to isolate the character when there are no 0 values on the projection.

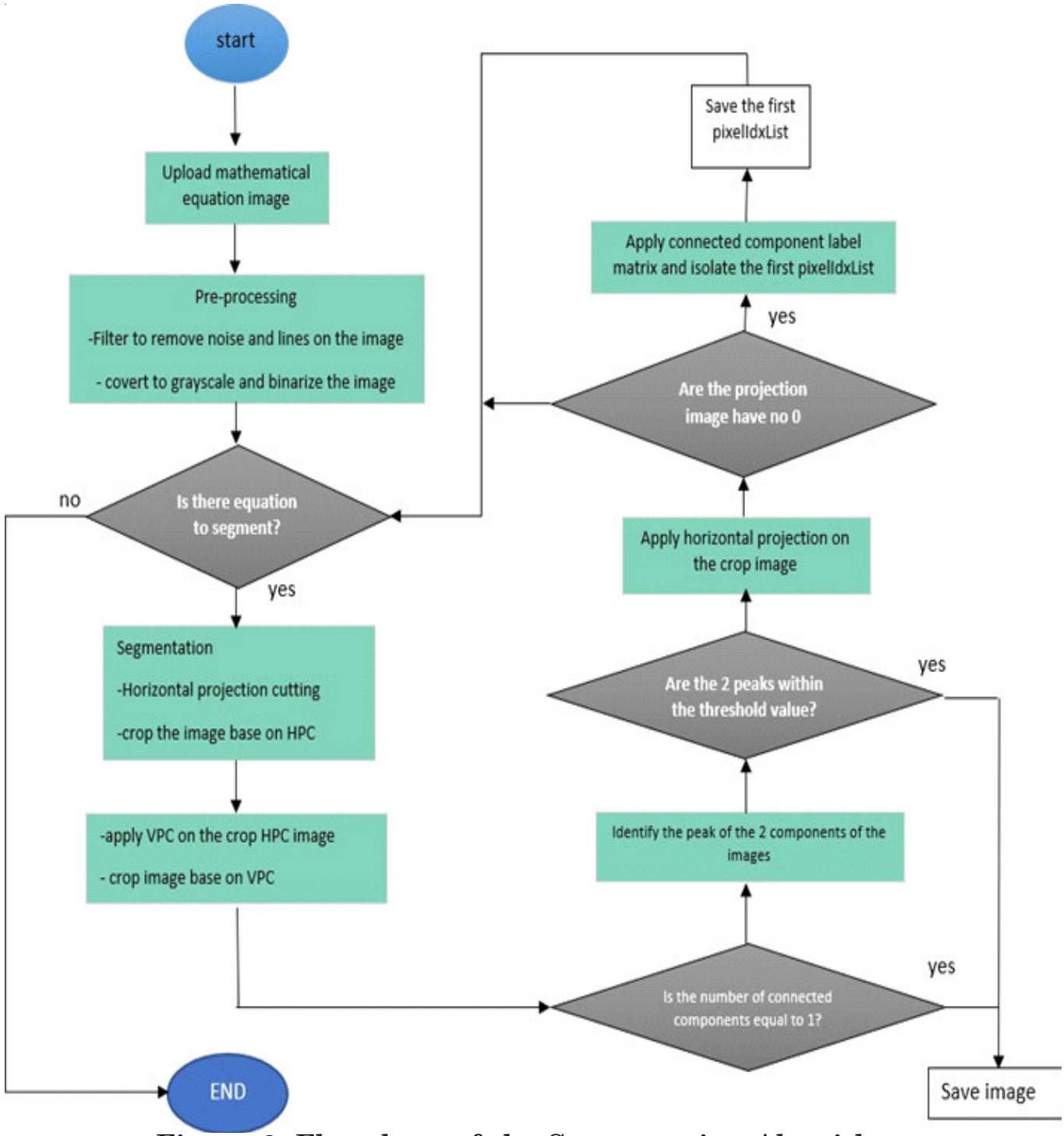


Figure 2. Flowchart of the Segmentation Algorithm

Classification-

Constructing CNN is a difficult undertaking. As a result, in order to conserve computing resources, this project would adopt the high accuracy architecture of L. Drsouza et al. [8]. This would reduce the number of layers required. A 45x45 jpeg picture serves as the CNN network's input image. Both the pooling layer and the convolution layer have a 5x5 filter size. SGDM, with a maximum of 30 epochs and an initial learning rate of 0.0001, is the learning option that is employed. The network is trained using the CROHME dataset. In this collection, 10,000 distinct photos are included. Hours will pass until all 10,000 photos are trained. As a result, 1000 photos representing each of the digits 0 through 9 are extracted and stored in a separate file. 40% of the photos were utilized for accuracy testing, while 60% of the images were used to train the network.

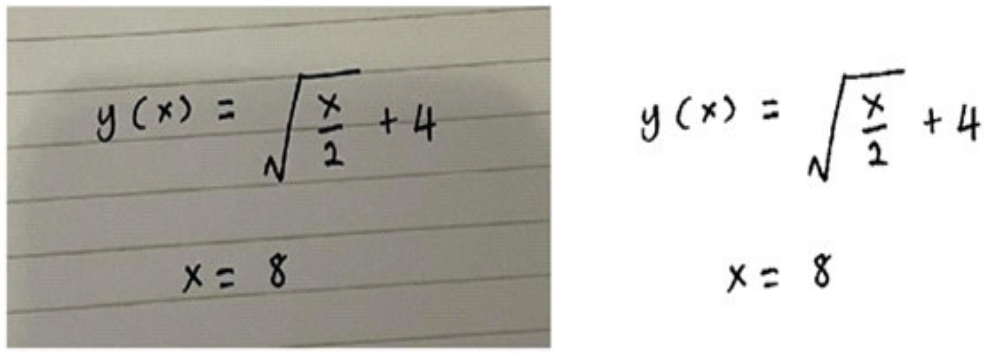
Figure 3. CNN Network Architecture

1.Results and discussion-

Outcome of image processing-

Figure 4 displays both the initial and pre-processed images. To get rid of any undesired lines and noise, a median filter and an adaptive filter were applied to the image. It is then binarized and turned to grayscale. Otsu's approach is MATLAB's default binarization method. Grayscale pictures are transformed into bi-level images using this global thresholding technique [12]. With a 0.7 gel pen, write the equation. Noise reduction is therefore unnecessary. Given that the image is composed of straight lines, skew correction is not applied. The picture underwent satisfactory pre-processing, and just the desired equation was binaryized.

The whole count of the horizontal projection's pixels is displayed in Figure 5(a). White pixels value of "1," whereas black pixels value of "0." A dark pixel has the equation printed on it. There is a gap between the two equations, as seen by the graph with all white pixels. Once all the holes have been identified, After then, the picture was horizontally cropped. The outcome of cutting with a horizontal projection is seen in Figure 5(b) picture.



(a) (b)

Figure4. Image binarization result (a) Original image (b) Pre-processed image

Classification Result-

The CROHME competition provides a benchmark dataset for evaluating and comparing the performance of OCR systems on handwritten mathematical expressions. The dataset contains a collection of online handwritten symbols, characters, and mathematical expressions, often captured through electronic devices like stylus-based tablets. The efficacy of the CNN training will be assessed by the test dataset outcome.

Table 1. CNN training result

Epoch	Iteration	Time Elapsed (hh : mm : ss)	Mini – batch Accuracy	Mini - batch Loss	Base Learning Rate
1	1	00:00:04	4.69%	38.1729	0.0001
2	50	00:00:42	95.31%	0.3484	0.0001
3	100	00:01:08	98.44%	0.4015	0.0001
4	150	00:01:28	100.00%	0.0067	0.0001
5	200	00:01:48	100.00%	0.0007	0.0001
6	250	00:02:06	100.00%	0.0007	0.0001
7	300	00:02:25	100.00%	0.0011	0.0001
8	350	00:02:43	100.00%	0.0009	0.0001
9	400	00:03:01	100.00%	0.0007	0.0001
10	450	00:03:18	100.00%	0.0002	0.0001
11	500	00:03:38	100.00%	0.0004	0.0001
12	550	00:03:56	100.00%	0.0001	0.0001
13	600	00:04:15	100.00%	0.0002	0.0001
14	650	00:04:34	100.00%	0.0002	0.0001
15	700	00:04:52	100.00%	0.0001	0.0001

The projected test data's confusion chart is displayed in Figure 9. Because it is simpler to recognize one stroke, the network appears to misclassify between numbers two and six, placing the greatest confidence in number one. The test's overall accuracy is 85%. The number that has been processed by segmentation is displayed in Figure 10. The categorization is quite certain that the two numbers provided are eight and four.

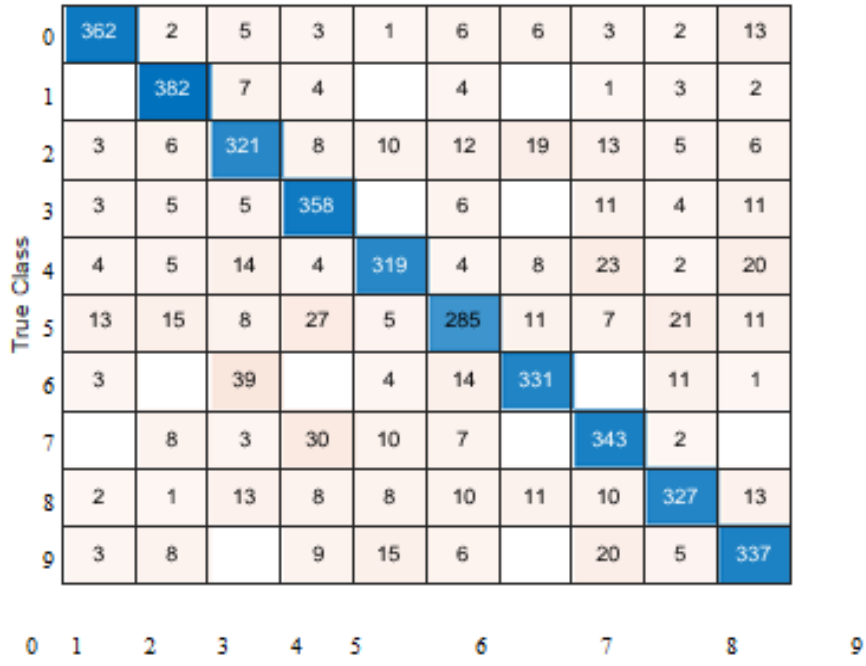


Figure 9. Confusion chart of data

Conclusion-

In summary, on lined paper, the project can preprocess mathematical equations written by hand, which it uses as the input picture, and then turn that image into a binary image that only contains the equations. The mathematical expression of the principle of image formation depends on the specific imaging system or process being considered. However, I'll provide a general overview of the mathematical principles involved in image formation, particularly in the context of optical imaging. The segmented equation is stored and scaled to 45 by 45 pixels so that CNN can classify it. The CNN used 600 pictures for training and 400 for testing, resulting in an 85% test accuracy for numerals 0 through 9. To improve CNN's accuracy, more testing has to be done. The project's relevance lies in its ability to generate mathematical formulae on a computer using optical character recognition (OCR) instead of requiring tedious hand typing of the equation. The project will address the problem of greater test accuracy and imaginary number segmentation in the future.

References-

1. Anderson, R.H., 1967, August. Syntax-directed recognition of hand-printed two-dimensional mathematics. In Symposium on interactive systems for

experimental applied mathematics: Proceedings of the Association for Computing Machinery Inc. Symposium (pp. 436-459).
2. D'souza, L. and Mascarenhas, M., 2018, August. Offline handwritten mathematical expression recognition using convolutional neural network. In 2018 International Conference on Information, Communication, Engineering and Technology (ICICET) (pp. 1-3). IEEE.
3. Veres, O., Rishnyak, I. and Rishniak, H., 2019. Application of Methods of Machine Learning for the Recognition of Mathematical Expressions. In COLINS (pp. 378-389).
4. Chan, C., 2020. Stroke extraction for offline handwritten mathematical expression recognition. IEEE Access, 8, pp.61565-61575.
5. Dai, J., Sun, Y., Su, G., Ye, S. and Sun, Y., 2019, January. Recognizing offline handwritten mathematical expressions efficiently. In Proceedings of the 10th international conference on E-Education, E-Business, E-Management and E-Learning (pp. 198-204).
6. Onyshchak, O., 2019. "A Tutorial on Optical Character Recognition in Mathematical Domain Image Recommendation for Wikipedia Articles View project A Tutorial on Optical Character Recognition in Mathematical Domain." doi: 10.13140.
7. Hossain, M.B., Naznin, F., Joarder, Y.A., Islam, M.Z. and Uddin, M.J., 2018, June. Recognition and solution for handwritten equation using convolutional neural network. In 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR) (pp. 250-255). IEEE.
8. D'souza, L. and Mascarenhas, M., 2018, August. Offline handwritten mathematical expression recognition using convolutional neural network. In 2018 International Conference on Information, Communication, Engineering and Technology (ICICET) (pp. 1-3). IEEE.
9. Le, A.D., 2020. Recognizing handwritten mathematical expressions via paired dual loss attention network and printed mathematical expressions. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 566-567).
10. Ramadhan, I., Purnama, B. and Al Faraby, S., 2016, May. Convolutional neural networks applied to handwritten mathematical symbols classification. In 2016 4th international conference on information and communication technology (ICoICT) (pp. 1-4). IEEE.
11. Debnath, B., Anika, A., Abrar, M.A., Chowdhury, T., Chakraborty, R., Khan, A.I., Fattah, S.A. and Shahnaz, C., 2018, October. Automatic Handwritten words on Touchscreen to Text file converter. In TENCON 2018-2018 IEEE Region 10 Conference (pp. 0219-0223). IEEE.