

## OPTICAL CHARACTER RECOGNITION (OCR) FOR SCIENTIFIC EQUATION

Sagar Tete\*<sup>1</sup>, Lakhan Ghayade\*<sup>2</sup>, Archit Datarkar\*<sup>3</sup>, Kunal Jaiswal\*<sup>4</sup>, Yogesh Dahake\*<sup>5</sup>

\*<sup>1</sup>Professor, Department Of Computer Science And Engineering, Priyadarshini JL College Of Engineering, Nagpur, Maharashtra, India.

\*<sup>2,3,4,5</sup>UG Students, Department Of Computer Science And Engineering, Priyadarshini JL College Of Engineering, Nagpur, Maharashtra, India.

### ABSTRACT

During composing a theory or documentation or tasks on math point, science or physical science and like those subjects, journalists tackles with the circumstance where they can't compose the entire COMPLEX condition without anyone else utilizing the standard applications like Microsoft Word or Presentation or some other application and they additionally can't duplicate the condition containing picture straightforwardly to their documentation. So the examination and any exploration paper or articles are in picture design or checked archive design so imagine a scenario in which we can make the application or program such that it will take that picture or filtered report and convert that picture to some organize like it will be not difficult to duplicate that condition and there is compelling reason need to compose that condition without anyone else. There is sure innovation called Latex which is articulated «Lahtech» or «Lay-tech» (to rhyme with «blech» or «Bertolt Brecht»), is a record readiness framework for top notch typesetting. It is most frequently utilized for medium-to-enormous specialized or logical reports however it tends to be utilized for practically any type of distributing. Thus, If we convert that condition which is in the picture or filtered reports into plastic code then in the wake of gathering that plastic code we will get the ideal condition and composed or client assignment will be not difficult to utilize that plastic code or incorporated duplicate and pastable condition any place they need.

**Keywords:** Latex Code, Optical Character Recognition, Scientific Equations, Writing Thesis, Deep Learning, Machine Learning.

### I. INTRODUCTION

The GUI will permit you to take a cut of your current PC/Laptop screen where the condition is available and the cut will be taken as information (IMAGE). Subsequent to taking the information our program will deal with the picture as made sense of beneath in technique and give the clipped condition in editable structure. Tool compartment: PyQt5 for intuitive Graphical User Interface. The objective of this task is to make a learning-based framework that takes a picture of a numerical equation and returns comparing LaTeX code. The framework that snaps a picture of a printed condition and delivers a Latex code portrayal. The interaction utilizes versatile thresholding with mean sifting, morphological edge smoothing, and Hough Transform for Image binarization. The beneath picture will explain the idea of our task of LaTeX code age. Assuming anybody need a condition, which is in design like of can reorder it anyplace in postulation and so forth they simply need to pick choice of plastic code or condition.

### II. METHODOLOGY

Deep neural network model with an encoder-decoder architecture also known as seq2seq model. The ViT encoder is a convolutional neural network (CNN) that transforms images into a group of feature maps. The encoder for such applications is usually a convolutional neural network (CNN) which encodes the input images as abstract feature representations, and the decoder is usually a recurrent neural network (RNN) that represents a language model to translate the encoder output into a sequence of tokens drawn from a vocabulary. The decoder is a stacked bidirectional Transformer model integrated with the soft attention mechanism, which works as a language model to translate the encoder output into a sequence of Latex tokens. Translating math formula images to Latex sequences is a joint field of image processing and text processing. The sequence-to-sequence model (seq2seq), also called the encoder decoder architecture has been successfully applied to intersect these two fields. Before inserting array of image to the encoder the will be first normalize

using albumentation the library of python, Albumentation is a computer vision tool that boosts the performance of deep neural network.

### III. WORKING

The GUI is made using pyqt5, in below image using snip [Alt+S] user can take snip of present screen and image will be taken as input to the model and area having red boundary is webpage where the rendered math equation will be shown as input image after image gone through encoder and decoder. The blue boundary area is text box where is the latex equation will appear.

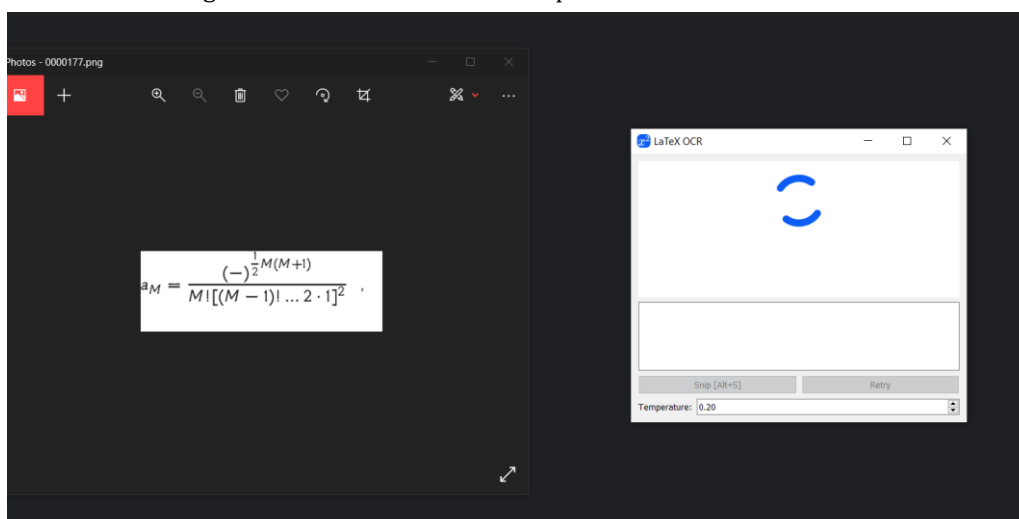
User can also write or paste the latex equation (with or without \$ sign) to get respected math equation which will be rendered with MathJax.



**Figure 1**

After taking the snip firstly the program will normalize the image, then numpy will convert it into array having values between 0 and 1 according to density of particular pixel.

Encoder will take array to predict the specific characters and also will take data of center of mass of that character then decoder will generate mathematical mark-up



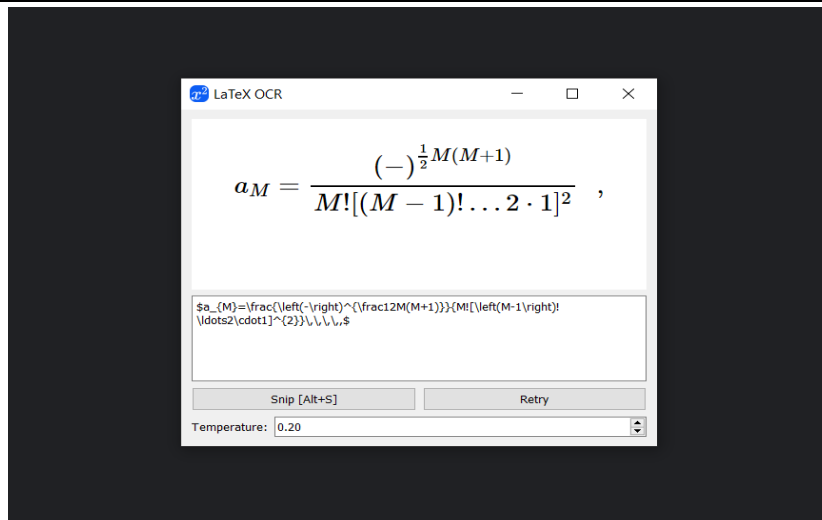


Figure 2: Capturing and Processing math equation.

#### IV. DISCUSSION

This model focuses on generating latex equations from input images which contains mathematical or scientific equations, helps adding (writing) math equations in official documents, thesis, or books. Deep neural network model with an encoder-decoder architecture also known as seq2seq model, visual transformer (vit) as encoder which is CNN (convolutional neural network) without the last classification layer and the decoder is transformer decoder.

#### V. CONCLUSION

The application will allow you to take snip of your present screen or upload image and takes that image as input which will contain math equation and after processing that image you will get latex code of that equation as output.

#### ACKNOWLEDGEMENTS

In The progress of any work relies upon endeavors of numerous people. We might want to take this opportunity to offer our profound thanks to the people who expanded their help and have directed to finish the venture work. We wish to communicate our genuine and most profound appreciation to our aide Er. Sagar Tete for his significant and exceptional direction. We might likewise want to express gratitude toward her for the consistent wellspring of help, motivation and support in the fruitful completion of task. It has been our honor and joy to work under his master direction. We like to express gratitude toward Dr. V.P. Balpande (Head of Department) for giving us the fundamental data about point. We might again want to express gratitude toward Dr.A.M. Shende , Principal of our school, for giving us the important assistance and offices we really wanted. We express our gratitude to all the staff individuals from CSE office who have straightforwardly or in a roundabout way expanded their caring co-activity in the finishing of our task report.

Place: Nagpur

Lakhan Ghayade

Yogesh Dahake

Kunal Jaiswal

Archit Datarkar

#### VI. REFERENCES

- [1] Ganesh History Of OCR: - <https://history-computer.com/optical-character-recognition-history-of-optical-character-recognition-ocr/>
- [2] Hough Algorithm: - <https://youtu.be/5zAT6yTHvP0>
- [3] Combining python with Latex: - <https://www.youtube.com/watch?v=LufMPezEN6Y>
- [4] What is OCR: - <https://www.youtube.com/watch?v=EQo8nXynOTw>
- [5] Preprocessing the image: - <https://www.youtube.com/watch?v=ADV-AjAXHdc>

- [6] Text recognition in OCR: - <https://youtu.be/oesLvefGEzs>
- [7] Anderson, R. H. 1967. Syntax-directed recognition of handprinted two-dimensional mathematics. In Symposium on Interactive Systems for Experimental Applied Mathematics: Proceedings of the Association for Computing Machinery Inc. Symposium, 436–459. ACM.
- [8] Belaid, A., and Haton, J.-P. 1984. A syntactic approach for handwritten mathematical formula recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*(1):105–111.
- [9] Cireşan, D. C.; Meier, U.; Gambardella, L. M.; and Schmidhuber, J. 2010. Deep, big, simple neural nets for handwritten digit recognition. *Neural computation* 22(12):3207–3220.
- [10] Gehrke, J.; Ginsparg, P.; and Kleinberg, J. 2003. Overview of the 2003 kdd cup. *ACM SIGKDD Explorations Newsletter* 5(2):149–151.
- [11] Hochreiter, S., and Schmidhuber, J. 1997. Long short-term memory. *Neural computation* 9(8):1735–1780.
- [12] Ioffe, S., and Szegedy, C. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *Proceedings of The 32nd International Conference on Machine Learning*, 448–456.
- [13] Jaderberg, M.; Simonyan, K.; Vedaldi, A.; and Zisserman, A. 2016. Reading text in the wild with convolutional neural networks. *International Journal of Computer Vision*116(1):1–20.
- [14] Karpathy, A., and Fei-Fei, L. 2015. Deep visual-semantic alignments for generating image descriptions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3128–3137.
- [15] Karpathy, A.; Johnson, J.; and Li, F.-F. 2015. Visualizing and understanding recurrent networks. arXiv preprint arXiv:1506.02078
- [16] Miller, E. G., and Viola, P. A. 1998. Ambiguity and constraint in mathematical expression recognition. In *AAAI/IAAI*,784–791
- [17] Mouchere, H.; Viard-Gaudin, C.; Kim, D. H.; Kim, J. H.; and Garain, U. 2012. Icfhr 2012 competition on recognition of on-line mathematical expressions (crohme 2012). In *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on*, 811–816. IEEE.
- [18] Mouchere, H.; Viard-Gaudin, C.; Zanibbi, R.; Garain, U.; Kim, D. H.; and Kim, J. H. 2013. Icdar 2013 crohme: Third international competition on recognition of online handwritten mathematical expressions. In *2013 12<sup>th</sup> International Conference on Document Analysis and Recognition*, 1428–1432. IEEE.
- [19] Papineni, K.; Roukos, S.; Ward, T.; and Zhu, W.-J. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics*, 311–318. Association for Computational Linguistics.
- [20] Suzuki, M.; Tamari, F.; Fukuda, R.; Uchida, S.; and Kanahori, T. 2003. Infty: an integrated ocr system for mathematical documents. In *Proceedings of the 2003 ACM symposium on Document engineering*, 95–104. ACM.