iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** VIII **Month of publication:** Aug 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.54989

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue VIII Aug 2023- Available at www.ijraset.com*

# Unwavering Mathematical Equation Deciphering from the Paperwork Imagery

Dr. J. Sreerambabu[1], Mr. N. Santhosh[2], Mr. D. Rajkumar[3], Ms. V. N. Nevedha[4]
[1] *Head of the Department,* [2, 3] *Assistant Professor,* [3] *PG Scholar*

***Abstract:*** ***The paper proposes a new approach for identifying equations in document images, with an emphasis on handwritten equations. The proposed strategy makes use of Python-based machine learning algorithms, and for the user interface, HTML, CSS, and JS. The document image is transformed into a machine-readable format during the pre-processing stage utilizing optical character recognition (OCR). Then, image-improving methods like smoothing, thresholding, and morphological procedures are used. A deep neural network architecture is used to accurately identify math expressions in the pre-processed photos. Convolutional and recurrent layers make up this layout, which was developed utilizing machine learning techniques.***
***Keywords: Math formula recognition, Handwritten equations, Machine learning, Optical Character Recognition (OCR), Deep neural network, Image enhancement.***

## I. INTRODUCTION

For many years, the field of OCR has been the focus of intense study, and substantial developments have resulted in the creation of marketable, commercial OCR software [1]. But issues still exist, especially when working with smartphone-captured document photos that have erratic lighting, poor contrast, several types of noise, and complex layout structures. This is specifically true for mathematical formulas, where structural analysis is very difficult to do [2]. Despite the fact that Math Formula Recognition (MFR) has been successful on high-quality document images [12], [13], [15] and online handwritten systems, the recognition of math formulas in degraded Chinese document images is still quite difficult because of the effects of uneven illumination, low contrast, and various noise. They are dedicated to helping learners identify similar homework or exercises to practice their weak areas as a start-up company serving education via the web. In particular, the ability to recognize mathematical equations in deteriorated Chinese document images is crucial to our services [3]. The Convolutional Neural Networks (CNNs) to handle the difficulties of detecting arithmetic formulae in damaged document images in order to meet the growing demand for effective and accurate math formula detection. With the use of an easy-to-use graphical user interface (GUI), our system seeks to provide students with a dependable and user-friendly solution that will allow them to upload degraded photographs of arithmetic formulas and receive recognized formulas and solutions [4]. Both the combination of the Ostu based Niblack Binarization method put forth [5] and the unique background estimation strategy for degraded document image binarization introduced [3] offer important insights into how to handle damaged pictures successfully. The use of Niblack binarization on document images provides insightful suggestions for low-cost and noise-tolerant stochastic architecture [6]. These investigations serve as the basis for our study on Chinese document images that contain math equations. Using Flask and OCR with CNNs, we offer an in-depth description of our proposed end-to-end OCR architecture in this publication for the recognition of arithmetic formulas in Chinese document images [7]. The pre-process and character-based recognition module and the formula analysis module are the two main modules which make up the entire system. The first module employs OCR-based binarization techniques to pre-process images, locate characters, recognize them using a trained classifier, and analyse layout and text lines for accurate recognition [8]. The formula analysis module addresses challenges in MFR: analysing formula structure mixed with Chinese text and recognizing adhesive formula elements [9]. The development of a user-friendly GUI using Flask Webapp will ensure an intuitive and interactive experience for users, enhancing the overall usability of the system [11].

## II. RESEARCH GAPS

Mathematical formula recognition (MFR) from document images is a challenging but crucial issue in the sector of document image analysis [5]. Due to improvements in optical character recognition (OCR) and deep learning techniques, several existing solutions have been established to deal with this problem [1] [2]. The review evaluates the distinctive methods, approaches, and algorithms of these systems along with their pros and cons [9]. These systems' potential uses, performance metrics, and training and evaluation datasets are all described [10]. The information gathered from the present study will direct following research as well as improve the development of improved and reliable mathematical formula recognition systems [6].

©IJRASET: All Rights are Reserved 

Accurate mathematical formula recognition from document images has drawn plenty of study as document digitization becomes more prevalent [14]. OCR, image processing, and deep learning methods are all used in combination by current mathematical formula recognition (MFR) systems [2]. Math formulas can be recognized in document images using existing MFR systems that use pre-processing techniques, CNN-based character recognition, feature extraction, and structural analysis [6], [12], [9]. Prospective research must concentrate on improving blended techniques, examining new structural analysis algorithms, and improving domain-specific difficulties with mathematical formula detection in document images [15].

This development will help to improve the MFR systems accuracy and efficiency, which will improve the fields of document analysis of images and recognition.

## III. PROPOSED METHODOLOGY

The objective of this research is to make a flexible framework that can recognize fine expressions with accuracy, handle different handwriting styles, analyse expression structure and symbols effective. The suggested method uses an iterative algorithm that takes use of similarities between structure analysis and symbol bracketing. The system can provide soft interpretations with confidence values rather than rigid outputs by employing machine learning approaches, which enhances symbol identification and structure recognition.

Decoupling the structure recognition process from the symbol identification process employing restrictions between them is the key challenge. Because of this decoupling, even when symbols are missing or statements are written in different ways, the system can recognize finer phrases with more precision.

Researchers prefer to improve an accurate and dependable optical character recognition (OCR) system for complicated encounters by using this method. Such a system may be necessary for the precise transfer and storage of complex mathematical formulas and precisely specified content.

The technique suggested includes machine learning methods with a sequential approach to fulfil demand for a trusted OCR system. The system may use the generated confidence scores to provide soft interpretation for both image and structure recognition. The system may accommodate altering the writing style and accurately evaluate the fine expression's component pieces by combining the symbol bracketing and structural analysis limitations.

The end result is a modernized OCR system which offers greater efficiency and dependability for proper expression identification and comprehension, enabling flawless saving and interchange of complex mathematical formulas as well as related documents.

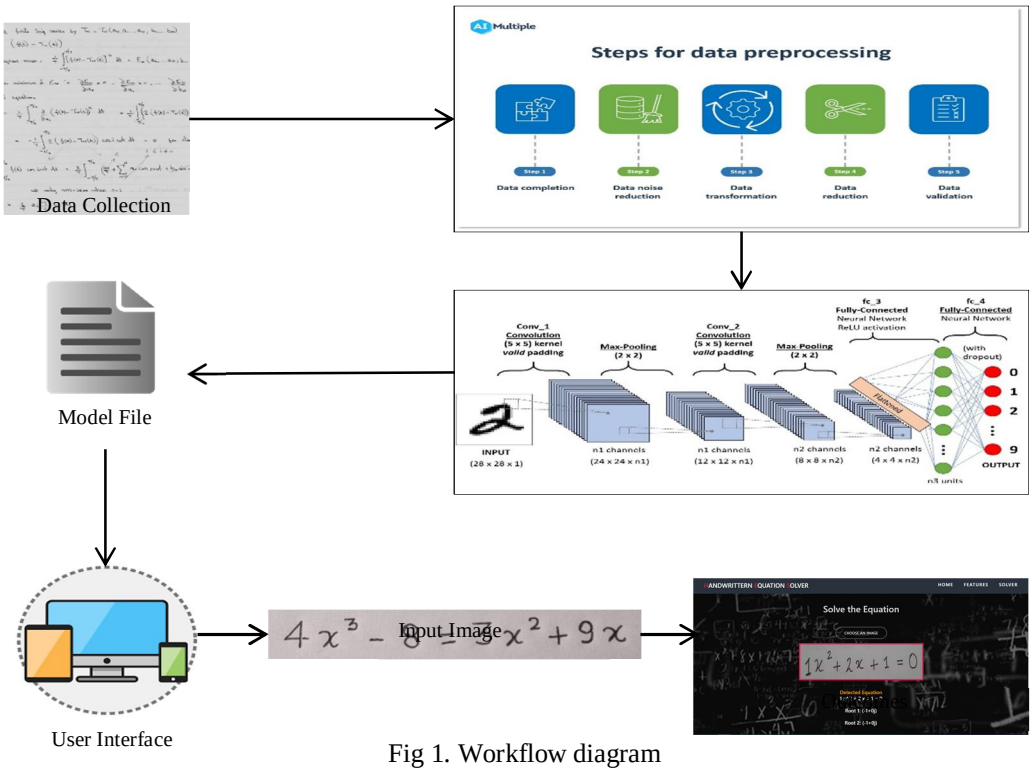



Fig 1. Workflow diagram

*A. Data collection*

The task is required for collecting enough quantities of document images containing equations to enable for the system's development. This dataset can be built via techniques like data augmentation for simulating various kinds of degradation, or it may exist from several kinds of resources, such as freely available databases. By system development and examination, this strategy ensures the dataset's inclusion of a diversity of representative samples. The collection of a sufficient amount of document images holding equations is required for the system to be implemented successfully.

These images can be developed by methodologies like data augmentation or collected from a variety of resources, like open-source databases, in order to demonstrate various kinds of degradation. The set of results is enlarged cover numerous kinds of degradation which could be noticed in everyday situations through data augmentation. The unique dataset allows the system to be properly planned and assessed, which improves the system's dependability and reliability.

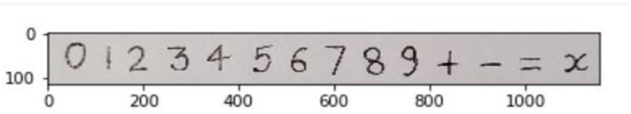



Fig 2 Dataset sample image

*B. Pre-processing*

Pre-processing is a crucial stage in the recognition of images which assures wider quality of image and enables accurate processing. A few examples of image improvement techniques that enhance input images by adjusting intensity levels and expanding the variety of results that enhance contrast and brightness include histogram equalization, contrasting stretches, and gamma correction. The process for segmenting of images divides the image into segments depending on appearance by applying these approaches, like thresholding, recognizing edges, & region development.

With the support of these techniques, it is possible to recognize foreground objects, evaluate limits, and categorize similar pixels into sections for further study. By reducing salt-and-pepper disturbances, a high-frequency disturbances, and also by specifically eliminating distortion from unique frequencies, noise reduction approaches like median elimination, Gaussian elimination, and wavelet denoising develop the reliability of the following analysis. By increasing their capacity for different uses, pre-processing techniques are used to change image recognition systems.

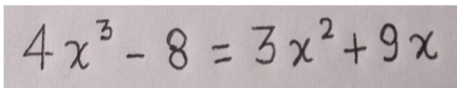



Fig 3(a). Input Image

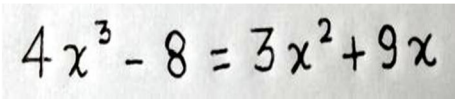



Fig 3(b). After Pre-processing

*C. Character Segmentation*

The first significant step in assessing the comprehension of equations is recognizing and identification of these assertions. Identifying equations in OCR data, using machine learning techniques to recognize the symbols, and developing mathematical equations are the three stages of this approach. The OCR output is initially examined to look for and recognize mathematical equations using techniques for object detection or image segmentation methods. In the second stage, machine learning algorithms includes SVMs and CNNs recognize the particular symbols that makes up the equations by analysing attributes includes shape, size, and orientation. The identified symbols are then arranged according to the fundamental concepts underlying the mathematical notation for the purpose for them to form the mathematical expressions. Reversible Polish Notation (RPN) as well as a tree-based notation can be utilized by these abilities. Both tree-based notation and RPN can be employed properly in these tasks, according to machine learning techniques. The machine-readable identification and recognition of equations depends on these three processes, which comprise symbol identification, detection, and expression development.

Fig 4 Dataset sample image

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue VIII Aug 2023- Available at www.ijraset.com*

## IV. CLASSIFICATION AND RECOGNITION

This approach depends on a structure known as a CNN (Convolutional Neural Network) that was developed clearly for classifying problems. Leveraging numerous layers, it accurately extracts essential characteristics from source images. The layout is made up of three convolutional stages, each containing 32 different filters of size 3x3. These layers will be followed by three max-pooling layers used to select the attributes preserving the most essential information. For the purpose of processing and analysing picture data, the 64x64 pixels & RGB channel source size is suitable for the CNN structure. Convolutional layers and max-pooling work together to extract significant features and patterns from the images. Because of its modular approach, the model is able to comprehend the fundamental layout of the data, which makes it perfect for image analytics.

The subsequent feature maps are shaped into a 1D vector complying with the processes of convolution and max-pooling layers. Then this vector is sent through two highly dense layers containing a total of 128 pixels each using the activation function of ReLU. These deep layers process the retrieved characteristics after that, aiding the model's comprehension of the complex relationships between them. Dropout layers, which randomly deactivate certain portions of the units during training, are inserted after the two initial dense layers to reduce over fitting and improve generalization on untrained input. It is projected that 50% and 30% of the pupils will leave school. Using 14 units, the output layer—the final layer in the model accompanies the required number of classes to finish the classification assignment. The probability values for every category are calculated using the softmax activating function, and they represent the degree of certainty with which the model can categorize an input image. The model leverages the optimizer developed by Adam, an established neural network optimization techniques, during training. To improve the model's performance, the categories cross- Entropy function of loss is utilized as the target. This combination allows the model to efficiently modify its weights and biases, hence minimizing the discrepancy between expected and actual category labels.

Table I
Model architect summary.

| Layer (Type) | Output Shape | Param# |
|---|---|---|
| Conv2D-1 | ( 62, 62, 32) | 896 |
| Conv2D-2 | (29, 29, 32) | 9248 |
| MaxPool2d-1 | (31, 31, 32) | 0 |
| Dropout-1 | (128) | 0 |
| Conv2D-3 | (12, 12, 32) | 9248 |
| MaxPool2d-2 | (14, 14, 32) | 0 |
| Dropout-2 | (128) | 0 |
| Flatten-1 | (1152) | 0 |
| MaxPool2d-3 | (6, 6, 32) | 0 |
| Dense-1 | (128) | 147584 |
| Dense-2 | (128) | 16512 |
| Dense-3 | (14) | 1806 |
| Total params: 186,294 | | |
| Trainable params: 186,294 | | |
| Non-trainable params: 0 | | |

According to this kind of approach, an optimizer by the term of RMSprop is applied to enhance results reliability. The size of weight changes during training are managed by the model by setting the learning rate to 0.001. A self-adjusting learning rate reduction method is also utilised. The model can evaluate validation accuracy and automatically lower the learning rate if necessary by establishing a minimum learning rate of 0.00001. The model may fine-tune its parameters to improve its overall efficiency according to this progressive lowering of the learning rate that eventually converges to an ultimate learning rate value of 1e-05.

The CNN architecture employed in this situation, then, adheres to a standard procedure for dealing with picture categorization problems. Layers using convolution and pooling are used to gather traits, followed by dense layers for categorizing the data and layers of dropouts to reduce overfitting. The performance and consistency of the model are enhanced by using the optimization tool for the RMSprop algorithm when coupled with an adaptive learning rate reduction methodology. Its accuracy for validation was 97.91%, whereas its accuracy during training across thirty epochs was 97.74%. Using test datasets, the model's feasibility and reliability in generating predictions were further evaluated. Figure 5 depicts the evolution of train-validation accuracy as well as loss.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue VIII Aug 2023- Available at www.ijraset.com*

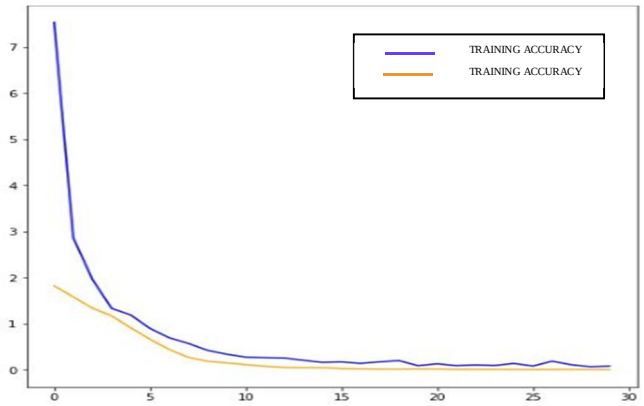



Fig 5(a). Cross Entropy Loss Graph

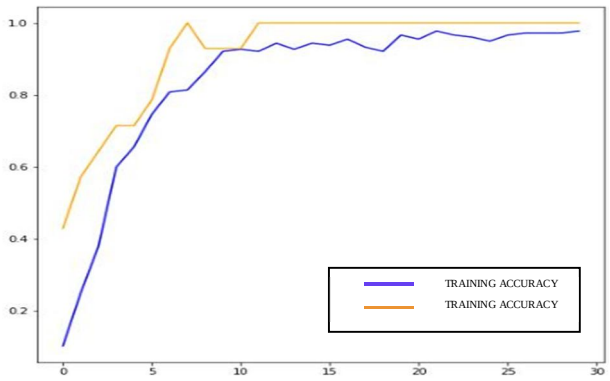



Fig 5(b). Classification Accuracy Graph

## V. SOLUTION OF MATHEMATICAL EXPRESSIONS

The primary image is divided into discrete signs and mathematical symbols, and the predictions for each are listed one after another. The created strings are then analysed to see if they represent trinomials, binomials, or monomials. Utilizing the Python Cmath module, that we store every set of characters in a different list according to word count in order to facilitate calculation of strings and equation solving.

TYPE 1

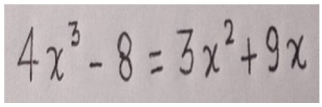
| INPUT IMAGE | GENERATED LIST |
|---|---|
|  | $4x^3-8=3x^2+9x$ |

Table II (a) Sample Images

Let's look at a Type 1 sample from the table equation "$4x^3 - 8 = 3x^2 + 9x$" to demonstrate how to solve equations. Finding the value or values of x that satisfies this equation is the objective. As we examine the equation, we first take note of the coefficients & degrees of x.

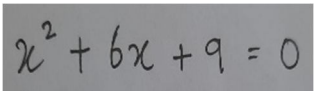
There are a total of four terms in this instance: one with a coefficient of 4 and a degree of 3 (4x3), another with a constant coefficient of -8 and a degree of 0 (-8), one with a coefficient of 3 and a degree of 2 (3x2), and one with a coefficient of 9 and a degree of 1 (9x).Using the Python the problems may be solved.

The process changes the outcome to [4, 13, 3, 9] after eliminating the correct symbol along with the upper-right section. It first determines the polynomial's degree, which in this case is 3, and then uses the coefficient's signs in line with basic regulations to produce [4, -13, 3, -9].

The equations 4x^3, 3x^2, and 9x represent the estimated coefficient of correlation values with correlation degree and level of significance as inputs.

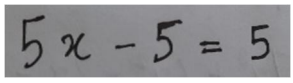
The solution to these equations is [0, 3, 9]. An issue with the root of the problem was identified, involving an integer of -8 in a formula. The code searches for the equation's roots using a cubic formula solver. The formula used in this instance has just one squared root, x = -2.0, it is discovered. We can determine that the formula "4x3 - 8 = 3x2 + 9x" includes one real root at x = -2.0 by modifying the initial code to the attached equation.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

| INPUT IMAGE | GENERATED LIST |
|---|---|
|  | $x^2+6x+9=0$ |

TYPE 2
Table II (b) Sample Images

Considering a Type 2 instance. From the table, we observe that the equation is x2 + 6x + 9 = 0. We may resolve this equation by employing the supplied algorithm. Since the polynomial is a quadratic equation, we first establish that its degree is 2. The equation "x2 + 6x + 9" can be output with the print-Eq function. The changed equation is therefore "x2 + 6x + 9" once the equal sign has been removed. We are not necessary to adjust the array because there are no substantial modifications to this equation. The equation's coefficients are then discovered. Term 9 of this coefficient is a constant, and the following coefficients are [1, 6, 9]. The equation has two recurring roots that are equivalent to -3.0, according to the quadratic solver function. Therefore, x2 + 6 x + 9 = 0 has two solutions: -3.0 and -3.0.We successfully established that the square roots in the quadratic formula x2 + 6x + 9 = 0 were both exactly -3.0 by using the code provided to the equation.

| INPUT IMAGE | GENERATED LIST |
|---|---|
|  | 5 x - 5 = 5 |

TYPE 3
Table II (b) Sample Images

Considering a Type 3 instance from the table, offers a methodical method for examining and resolving polynomial problems of many different degrees and kinds. It dissects a formula, such as "5x - 5 = 5," to expose its coefficients and constants. In this illustration, a constant term is -5, while the corresponding value of x is 5. The code recognizes the equality symbol and modifies the symbols of the equations by breaking the equation down into a set of tokens. The code determines the variables' degrees and establishes the coefficients for linear formulas like "5x - 5 = 5." In this instance, the value of the constant is -5, and the corresponding coefficient of x had a degree of 1. The code locates the equation's root, corresponding to x = 2 in our example, by using the proper solver. The algorithm offers an effective and precise solution to polynomial problems through its systematic methodology.

## VI. PROTOTYPE EVALUATION

The performance of the recognition module greatly influences the accuracy of the overall result, and a tiny percentage of the test images segmented from the input image are improperly recognized. The resulting simplification will be inaccurate if even one symbol or number is erroneous. We analysed thousands of equations throughout the experiment and discovered that a handful of them were incorrectly identified, resulting in the creation of false string equations. Out of the thousand photos analysed, almost 97% were correctly identified, as shown in Figure 7.

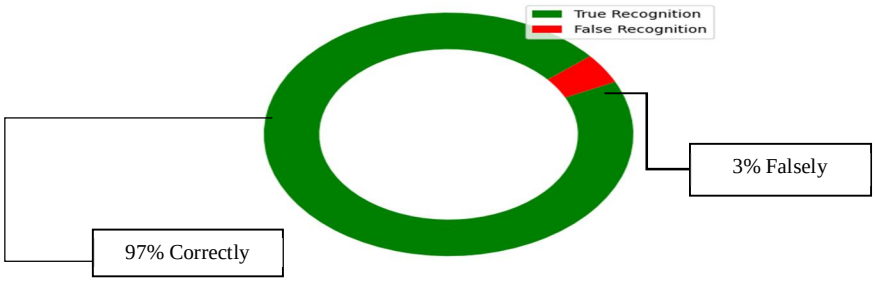



Fig 7. Visualization of the recognition rate

Table III (a) shows some correctly generated lists and correct output

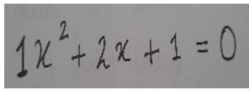
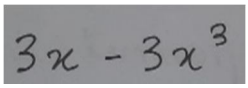
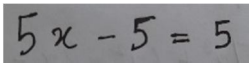
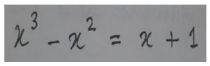
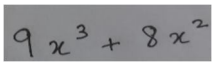
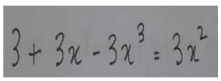
| NO | INPUT IMAGE | TRULY RECOGNIZED EQUATIONS IN LISTS | TRUE ANSWER |
|---|---|---|---|
| 1 |  | $1x^2+2x+1=0$ | [-1,-1] |
| 2 |  | $3x-3x^3$ | [-1,0,1] |
| 3 |  | 5x-5=5 | [2] |

Image evaluation and identification problems may occur for many kinds of explanations, including confusing or figuring source images as well as incorrectly labelled information in datasets. Because of these problems, the data derived from such images may be incorrect. Erroneous evaluation, that could be triggered by human error in image classification, restrictions on the usual number of input images, or confusion in the data, has an impact on the reliability of both the conclusions generated from the data and also on the methods used to make decisions. Applying restricted quality control techniques, continually enhancing data extraction procedures, & implementing novel algorithms that are able deal with noise or unclear records are essential in safeguarding data interpretation accuracy.

Table III (b) shows some wrongly generated lists and wrong output

| NO | INPUT IMAGE | THE WRONG GENERATED EQUATIONS IN LISTS |
|---|---|---|
| 1 |  | $x^3-x^2=x+2$ |
| 2 |  | $9x^3+8x$ |
| 3 |  | $3+3x-3x^2=32$ |

## VII. CONCLUSION AND FUTURE WORK

The main objective of this investigation is to establish a reliable and efficient method for identifying mathematical formulas from document images. To accomplish this, the system use Convolutional Neural Network (CNN), a Deep Learning strategy, to identify and evaluate equations. The system's display of the ability to handle polynomial problems up to a third degree utilizing CNNs increases its value for clients who require the both formulae recognition and polynomial issue solving. Because of the incorporation of CNNs, the system can successfully analyse and evaluate equations, providing users with quick and efficient responses. The implementation of a following formula the solver in the project demonstrates the portability and complexity of deep learning approaches. It explains how similar techniques as well as strategies may be employed for various kinds of mathematical problems, which means enhancing the project's possible applications. In addition, future advances in this field of study aim to include actual detection of arithmetic formulas using an imaging device. The system may be included to detect equations from streaming video or images acquired with an imaging device using methods such as computer vision, allowing for quick identification and assessment.

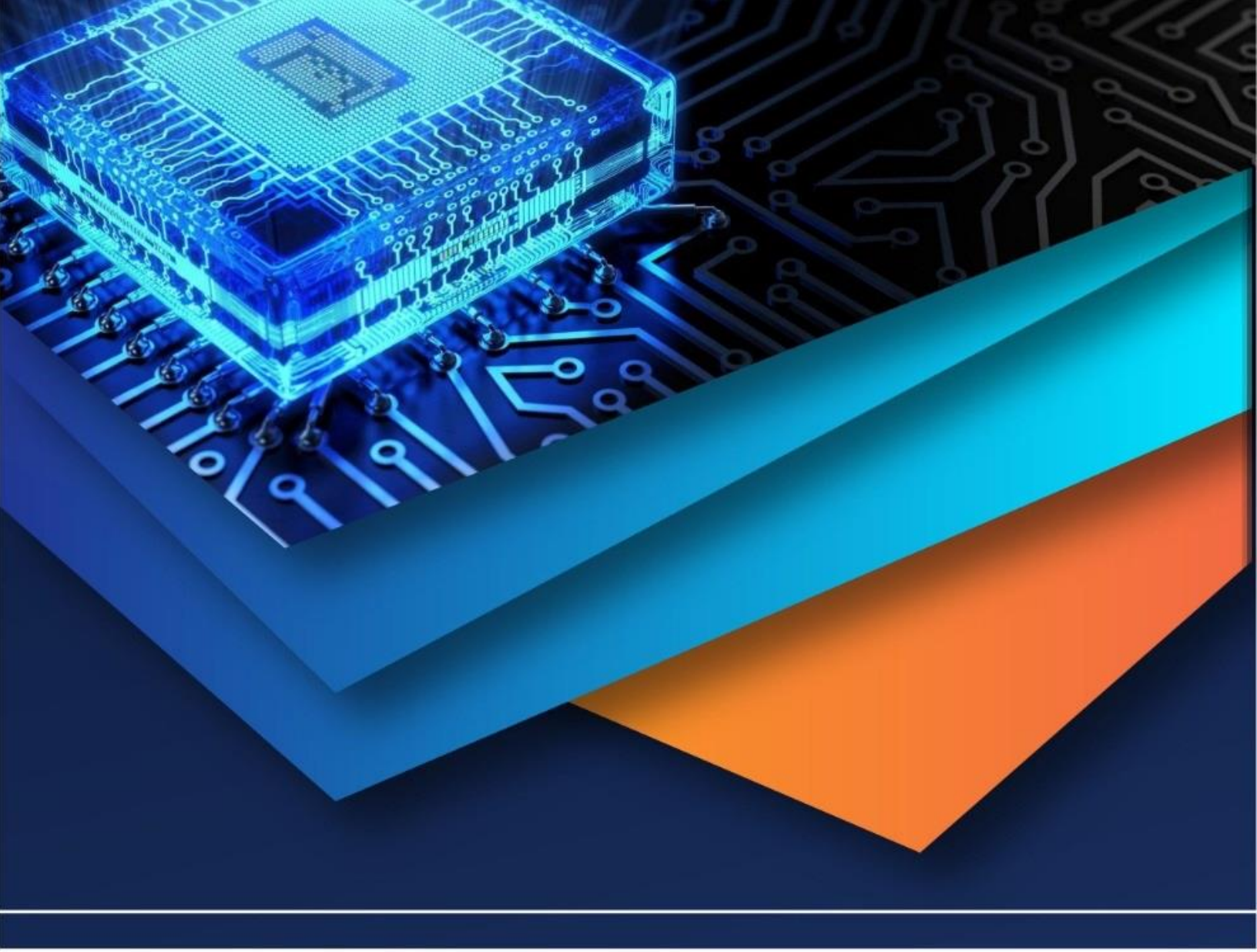
As the technique emerges, representations of equation-solving techniques may be necessary. By which includes all of these features, the software's equation-solving approach could grow simpler to figure out as well as comprehend for users. The method develops easier and more user-friendly by providing successive images, providing significant instructional tools for classification in addition to the ability to resolve problems. Such possible development show the program's durability and potential in mathematical evaluation and the debugging process

## VIII. ACKNOWLEDGMENT

I acknowledge our Head of the Department Mr Dr. J. Sreerambabu, M.E., Ph.D., PDF, FIE and our mentor Mr N. Santosh, MCA, who provided insight and expertise that greatly helped the research, for suggestions that greatly improved this manuscript. Special thanks, to our supervisor Mr M. Mohammed Riyaz, MCA, for the support in this research work.

## REFERENCES

[1] MATHEW, J. "Recovery of badly degraded Document images using Binarization."
[2] Du, Z., & He, C. "Nonlinear diffusion equation with selective source for binarization of degraded document images." Applied Mathematical Modelling, 99, 243-259, 2021.
[3] Jindal, H., Kumar, M., Tomar, A., & Malik, A. "Degraded Document Image Binarization using Novel Background Estimation Technique." In 2021 6th International Conference for Convergence in Technology (I2CT), pp. 1-8, IEEE, April 2021.
[4] Xiong, W., Zhou, L., Yue, L., Li, L., & Wang, S. "An enhanced binarization framework for degraded historical document images." EURASIP Journal on Image and Video Processing, 2021(1), 1-24.
[5] Ranjitha, P., & Shreelakshmi, T. D. "A Hybrid Ostu based Niblack Binarization for Degraded Image Documents." In 2021 2nd International Conference for Emerging Technology (INCET), pp. 1-7, IEEE, May 2021.
[6] Mitra, S., Santosh, K. C., & Naskar, M. K. "Niblack binarization on document images: Area efficient, low cost, and noise tolerant stochastic architecture." International Journal of Pattern Recognition and Artificial Intelligence, 35(04), 2154013, 2021.
[7] BJ, B. N., VA, N. A., & Akhil, A. "A Novel Binarization Method to Remove Verdigris from Ancient Metal Image." In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 884-888, IEEE, May 2021.
[8] Lü, Y., Lin, H., Wu, P., & Chen, Y. "Feature compensation based on independent noise estimation for robust speech recognition." EURASIP Journal on Audio, Speech, and Music Processing, 2021(1), 1-9.
[9] Bera, S. K., Ghosh, S., Bhowmik, S., Sarkar, R., & Nasipuri, M. "A nonparametric binarization method based on ensemble of clustering algorithms." Multimedia Tools and Applications, 80(5), 7653-7673, 2021.
[10] Mechi, O., Mehri, M., Ingold, R., & Amara, N. E. B. "A two-step framework for text line segmentation in historical Arabic and Latin document images." International Journal on Document Analysis and Recognition (IJDAR), 1-22, 2021.
[11] Wu, Y., Gao, Z., Feng, Y., Cui, Q., Du, C., Yu, C., ... & Sun, J. "Harnessing selective and durable electrosynthesis of H2O2 over dual-defective yolk-shell carbon 57 nanosphere toward on-site pollutant degradation." Applied Catalysis B: Environmental, 298, 120572, and 2021.
[12] Sharan, S. P., Aitha, S., Kumar, A., Trivedi, A., Augustine, A., & Sarvadevabhatla, R. K. "Palmira: a deep deformable network for instance segmentation of dense and uneven layouts in handwritten manuscripts." In International Conference on Document Analysis and Recognition, pp. 477-491, Springer, Cham, September 2021.
[13] Khan, M. A., Zhang, Y. D., Alhusseni, M., Kadry, S., Wang, S. H., Saba, T., & Iqbal, T. "A fused heterogeneous deep neural network and robust feature selection framework for human actions recognition." Arabian Journal for Science and Engineering, 1-16, 2021.
[14] Xiong, W., Jia, X., Yang, D., Ai, M., Li, L., & Wang, S. "DP-LinkNet: A convolutional network for historical document image binarization." KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS, 15(5), 1778-1797, 2021.
[15] Mansouri-Benssassi, E., & Ye, J. "Generalisation and robustness investigation for facial and speech emotion recognition using bio-inspired spiking neural networks." Soft Computing, 25(3), 1717-1730, 2021.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)