Handwritten Mathematical Symbol Recognition using Convolutional Neural Network

Girish S. Katkar[1], Pooja C. Ukey[2], Ajay S. Ramteke[3]
[1]*Assistant Professor, Department of Computer Science, Taywade College, Mahadula, Koradi*
[2]*Research Scholar, Taywade College, Mahadula, Koradi, Dist.-Nagpur*
[3]*Assistant Professor, Department of Computer Science, Taywade College, Mahadula, Koradi*

Abstract: **In our day-to-day life handwritten symbols are seen everywhere. We do involve automating data entry, letters, from writing cheques to writing notes manually. Symbol recognition is a process of detecting and recognizing symbols from input images and converting them into machine editable form. The technique by which a computer system can recognize symbols and other text written by hand is called a handwriting recognition system. Handwritten mathematical symbol recognition used widely for performing automating data entry, and processing handwritten applications. It is challenging for a computerized system to carry out certain types of tasks is not easy. Handwritten style of an individual person is different and varies from time to time. While understanding handwritten symbols, there are many challenges which we have to deal with. The degradation of a document or image over time can impair the ability to accurately interpret the symbols. However, we can find a solution for this using deep learning. This paper presents a model that accurately interprets handwritten characters using a dataset employed to train the model. The primary aim of this paper is to establish effective and dependable methods for recognizing handwritten symbols, and the accuracy of handwritten symbol recognition using a sequential model is 99%.**

Index Terms: Deep learning, Feature Extraction, Handwritten Mathematical Symbol recognition, Neural Networks.

I. INTRODUCTION

In recent years, handwriting recognition is one of the most captivating and challenging study areas in the subject of image processing and sample recognition. Many research works have been specialized in techniques and new strategies that reduce the processing time at the same time as providing better accuracy. In the neural network, most of those structures apply tool learning mechanisms. After the reputation device is very

beneficial like documentation analysis, Mail processing and license plate recognition, signature verification, postal addresses. Handwritten recognition may be very useful in the correct global sense. Numerous methods have been used in each online method and have been used in both online and offline handwriting recognition fields like statistical methods, structural strategies. A few recognition machines pick out strokes and others the recognition of individuals. So, the handwriting recognition tool works as a verbal communication between humans and machines. A Convolutional Neural Network (CNN) is a special type of deep learning algorithm. It is designed for tasks that make necessary for object recognition, image classification, detection, and segmentation. A sequential model in a Convolutional Neural Network (CNN) is a conventional Convolutional Neural Network where each layer only connects to the layer before and after it. This model is created by stacking layers together. Each layer having one input and one output. The Sequential model is available in Keras.

II. EXISTING SYSTEM

• Different persons have different handwriting styles but the symbols are handwritten it becomes difficult for the evaluator to recognise because of variance in size, translation.
• There exists a system which can scan symbols and there even exists a system where we can write our handwritten symbols directly but there doesn't exist a system which can interpret handwritten symbols.

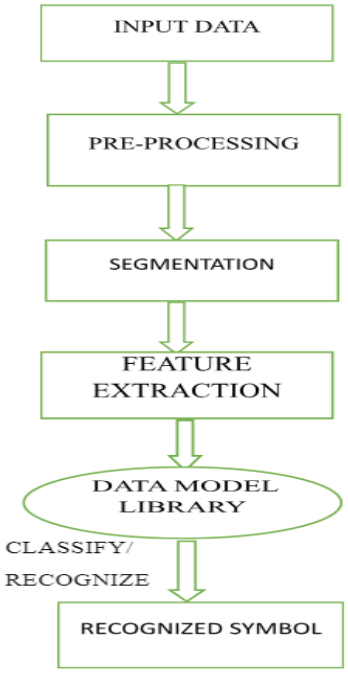


Figure1: PROCESS OF HANDWRITTEN MATHEMATICAL SYMBOL RECOGNITION

III. PROPOSED SYSTEM

• Handwritten Symbols recognition system is the work and training to recognize the symbols which is the solution for the existing problem.

• Recognizing Symbols from different sources like Document analysis, Postal office automation, National ID number recognition systems, Bank automation.

• In this project we use a deep learning algorithm. We used CNN algorithm to clarify images of handwritten mathematical symbols.

• The uniqueness and variety in the handwriting of different individuals also influence the formation and appearance of the alphabets.

• The aim of a handwriting recognition system is to convert handwritten symbols into machine understandable format using CNN with accuracy.

IV. TRAINING AND RECOFGNITION

After feature extraction, the next step is classification. The classification process is to make decisions for new input symbols. The stage of categorizing symbols is recognized and assigned labels. Good performance depends on the classification and feature extraction selection. When the input image is used as the input sample for a handwritten mathematical symbols recognition system, all significant features will be recovered and input into a trained classifier. Preparation of the input feature is compared with the save patterns to find a matching categorization of the input image. It will be complete with the help of the classifier. The correct label training data is needed to analyze the testing samples. Then, train the convolutional neural network model in Python. Perform training with weight decay and exit on each fully connected layer. The recognition process is done using three layers of CNN. CNN uses filters on the pixels of any image to learn detailed patterns compared to global patterns with a traditional neural network. To create CNN, we have to define:

1. A convolutional Layer: Applying the number of filters to the feature map. After convolution, we need to use a feed activation function to add non-linearity to the network.

2. Pooling Layer: The next step after the Convention is to down sampling the maximum facility. The objective of pooling is to reduce the mobility of the feature map to prevent overfitting and improve the computation speed. Max pooling is a traditional technique and splits feature maps into subfields and only holds maximum values.

3. Fully connected Layers: All neurons from the past layers are connected with the other next layers. The CNN has classified the label according to the features from convolutional layers and make less with any pooling layer. This proposed address employs convolutional layer, max pooling layer and dropout layer, dense layer and flatten layer. The image is firstly passed into a convolutional layer for two times which apply certain extracted features on the filter. The max pooling layer was used to reduce the size of the given image for the further transfer into a dropout layer which prevents overfitting.

Output Recognition: We use Sequential Model to read images and after undergoing all the processes the final image is used to make predictions. The output screen contains an image and its predicted value. After the execution screen is displayed which contains the result.

V. EXPERIMENTAL WORK

The dataset used in this research consists of images of handwritten mathematical symbols. It includes symbols such as plus, minus, integral, square root, and others, covering a wide range of mathematical notations. The dataset is pre-labelled and split into training and testing sets.

Preprocessing

To improve the model's performance, the images are pre-processed:

- Normalization: Pixel values are scaled to the range [0, 1] to standardize the input.
- Resizing: All images are resized to a consistent shape, e.g., 28x28 pixels.
- Augmentation: Techniques such as rotation, translation, and scaling are applied to increase dataset diversity.

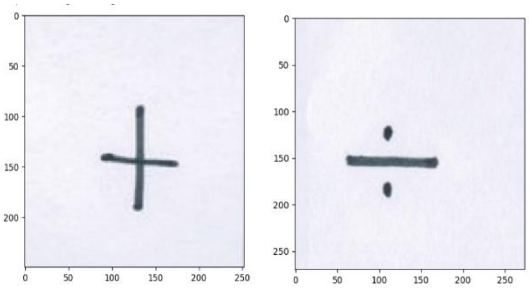


Figure 2- INPUT MATHEMATICAL SYMBOLS

VII. MODEL ARCHITECTURE.

Sequential Model

A Sequential model is working for its simplicity and ease of implementation. The model architecture includes:

- Input Layer: Accepts the pre-processed 28x28 grayscale image.
- Convolutional Layers: Spatial features extract using filters.
- Pooling Layers: Spatial dimensions and computational complexity are reduced.
- Flatten Layer: The 2D feature map converts into a 1D vector.
- Fully Connected Layers: the symbols based on the extracted features are classified.
- Output Layer: Uses a SoftMax activation function to output probabilities for each symbol class.

Compilation

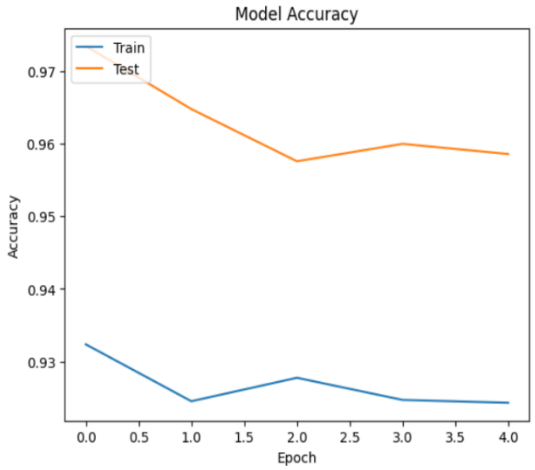
The model is compiled with the following configurations:

- Optimizer: Adam, chosen for its adaptive learning rate.
- Loss Function: For multi-class classification categorical_crossentropy is suitable.
- Metrics: Accuracy, for evaluating the model's performance.

VIII. EXPERIMENTAL RESULTS

Training

The model is trained using the training dataset over several epochs. The training process involves backpropagation and optimization, where the model iteratively improves its ability to classify symbols.



Evaluation

After training, the model's performance is evaluated on the testing dataset. The key metrics include accuracy, precision, recall, and F1-score. Classification report generated the accuracy 0.99 by deep layer.

Accuracy			0.99	10000
Macro avg	0.99	0.99	0.99	10000
Weighted avg	0.99	0.99	0.99	10000

Figure 3- CLASSIFICATION REPORT

This research demonstrates the effectiveness of a Sequential deep learning model in recognizing handwritten mathematical symbols. While the model achieves promising results, ongoing improvements and experimentation are necessary to enhance its robustness, particularly in dealing with ambiguous or poorly written symbols.

IX. CONCLUSION

Deep learning has become a dominant method in a variety of complex tasks such as object detection and image classification. We learned how the system recognizes handwritten mathematical symbols using Keras, sklearn, and Pyplot, NumPy, Matplotlib and this handwriting recognition system utilized basic image processing algorithms to segment symbols from an input image and then passed symbols through trained handwriting recognition model to recognize each symbol. This algorithm first reads the symbols, binarizes, and extracts features of handwritten mathematical symbols images, and sets up the database of handwritten mathematical symbols. A recognition model based on Deep learning algorithm was set up in matplot which is trained by sample images in the database. All the training samples have been studied; the learning effects of the model tested by test samples. With the proper pre-processing tasks and valid database will lead to the successful detection of handwritten mathematical symbols from the images with better efficiency. Handwriting mathematical symbols recognition based on deep learning algorithms provides an effective way to improve the overall recognition rate of data *and the accuracy by using a sequential model is 99%.*

REFERENCES

[1] Aggarwal, C. C. (2018). Neural Networks and Deep Learning: A Textbook. Basingstoke, England: Springer

[2] Monika*1, Monika Ingole2, Khemutai Tighare2, “Handwritten Character Recognition Using Neural Network”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Volume 7, Issue 4, July-August-2021.

[3] Chooi Shir Ley1, Aimi Syammi Binti Ab Ghafar1*, “Handwritten Character Recognition Using Convolutional Neural Network”, Progress in Engineering Application and Technology Vol. 2 No. 1 (2021) p. 593-611.

[4] S. Anandh Kishan, J. Clinton David, P.Sharon Femi, “Handwritten Character Recognition Using Cnn”, IJRAR September 2018, Volume 5, Issue 3.

[5] Thatikonda Somashekar, “A Survey on Handwritten Character Recognition using Deep Learning Technique”, Journal of University of Shanghai for Science and Technology, Volume 23, Issue 6, June – 2021.

[6] V.K. Govindan and A.P. Shivaprasad, “Character Recognition – A review,” Pattern Recognition, vol. 23, no. 7, pp. 671- 683, 1990

[7] Boukharouba, A., & Bennia, A. (2017). Novel feature extraction technique for the recognition of handwritten digits. Applied Computing and Informatics, 13(1), 19- 26.

[8] Balas, V. E., Roy, S. S., Sharma, D., & Samui,P. (2019). Handbook of Deep Learning Applications. Basingstoke, England: Springer.