Grenze International Journal of Engineering and Technology, June Issue

Digital Recognition of English Handwritten Text

Vasanth Nayak[1], Shravan G Amin[2], Shreesha Shetty[3] and Shwetha Prabhu[4]
[1-4]Canara Engineering College, Information Science and Engineering, Mangalore, India
Email: vasanthnayak01@gmail.com, shravangamin311@gmail.com, shettyshreesha58@gmail.com, shweprab03@gmail.com

***Abstract*— In the modern world, data is widely recognized as one of the most important commodities on Earth. Therefore, it is essential that the data be digitally transformed. Handwritten text is used in many different businesses for operational purposes. For instance, handwritten tax returns, doctor's prescriptions, resumes, and financial and legal papers. This emphasizes how important it is to transform handwritten material because handwriting differs between people and is easily misread. After manually entering word crop photos into the application as a document, the text has to be retrieved from the image. One of the modifications is the field of artificial intelligence and machine learning known as Deep Learning. The suggested approach uses convolutional neural networks, specifically ANN to classify handwritten characters and then trains a model that can recognize them. This method allows for a more accurate and efficient conversion of handwritten text into digital format, saving time and effort. Additionally, advancements in this technology have led to improved accuracy rates in recognizing various handwriting styles and fonts.**

***Index Terms*— Deep Learning, Text Recognition, CNN- Convolutional Neural Network, IAM handwritten dataset, Image Processing**

I. INTRODUCTION

One of the most studied and researched pattern classification problems is handwritten text recognition. After years and years of extensive research, the problem of converting handwritten text to digital format has gradually advanced from the recognition of separated characters to complex cursive scripts. Handwritten text must be converted into a format that is machine readable by the handwriting recognition system. This can be used for writing text that is directly captured on an online recognition system, such as a digitizing device, or for offline images (scanned or based on cameras). When compared to printed texts, handwritten texts are recognized as being much more difficult and demanding to recognize. This is primarily because handwritten texts, especially those written in cursive, require the writer to make specific choices when writing characters and connecting them. This might differ depending on the person. The handwritten word recognition system processes the entire word contained in a scanned image. Its goal is to convert every word from an image format to a digital format that can be edited and searched. This system has been the focus of many researchers whose interests lie in the recognition of handwritten text. The Deep Learning system has demonstrated its ability to function as an intelligent system for these kinds of applications. It has been observed that the Deep-Learning system model integrates Transfer Learning, which requires a sizable data collection. While deep learning algorithms have shown promise in terms of accuracy, these models or systems rely heavily on big datasets with well labeled training data. It has hampered numerous study initiatives because of the requirement for a sizable data corpus. Even with the prevalence of numerous technologies in today's society, many people still choose to generate physical documents and papers that

Grenze ID: 01.GIJET.10.2.179

are necessary using paper and pen. It is a difficult task to save and retrieve data from traditional documents. The field of handwritten text recognition is in demand and has room to grow quickly in the current globalization period. When a computer is able to convert handwritten text into a digital version, it is known as handwritten recognition. Additionally, this recognition technique can be used with PDAs and tablet PCs. One benefit of this recognition system is that fewer workers are required for data retrieval and sorting. The preservation of historical scripts and documents, such as family archives and personal diaries, which may have been altered or tainted by mishaps or other circumstances, is a significant benefit of this approach. The main problem with this approach is that most people write words in different ways, and accuracy is another big worry.

II. LITERATURE SURVEY

In Hassan, Shahbaz, et.al [1], by using Urdu as a base or by treating it as a study, the system suggests an effective method for perceiving handwritten text in cursive format (these results can further be extended to different other scripts for cursive writing also). This system uses an analytical technique to work on implicit character segmentation, extracting and segmenting information through the usage of convolutional neural networks. The Long Short Term Memory (LSTM) network is a bi-directional network that does this. A dataset including 6000 odd handwritten lines of text with good character recognition rates is used to validate the system or model.

In Korichi, Aicha, et.al [2], The system uses CNN-based feature extraction techniques to assess the effectiveness and interpretation of Arabic handwriting extraction. The open AHDB benchmark database served as the basis for the trial outcome.

In Ashiquzzam, Arm, and Abdul Kawsar Tussah [3], Unlike other recognition methods that take Arabic numerals into account, the suggested system uses a revolutionary and new algorithm based on deep learning networking and a suitable layer of regularization. This helps to gain substantially higher accuracy. With an accuracy percentage of 97.4, the suggested method is thought to have the best accuracy for the dataset used in the experiment.

In Rani, N. Shobha, et.al [4], with the aid of the Transfer Algorithm from the Devanagari system, which recognizes characters written in longhand, the suggested system offers a fresh and innovative method of identifying handwritten Kannada characters. The goal is to take into use the knowledge of a sizable data corpus for the training data that is taken into account from the Devanagari recognition model for Kannada longhand character recognition, which has a relatively small data corpus. The Deep Learning architecture to VGG19 NET facilitates the transmission of knowledge.

In Majid, Nishatul, and Elisa H. Barney Smith [5], the offline handwritten recognition performance of two devices is compared in the proposed system. These two devices—one is a mobile camera, the other is a scanner—are used for different types of image acquisition. Because it is easy to use and portable, the scanner produces distortion-free, high-quality images similar to those from a mobile phone camera. This system's or model's goal is to extract high-quality images from a scanner, which will influence offline handwritten recognition.

In Nikitha A., et.al [6], the suggested system is an implementation of the HTR system's deep learning framework. First, data is gathered in the system in question in order to train the handwritten messages. Subsequently, features are collected, extracted, and retrieved from these text datasets, and Deep Learning is used for training. To increase accuracy, the method use the recognition of words as opposed to characters.

In Bhagyasree P. V., et.al [7], the suggested method makes advantage of cutting-edge deep learning techniques, such as Directed Acyclic Graph—a.k.a. Convolutional Neural Network—which help recognize longhand characters. The characters are identified using both traditional and Deep-Learning methods. Deep Learning Method Convolutional-Neural-Networking provides more accuracy compared to other traditional methods.

In Remaida, Ahmed, et.al [8], The system's goal is to present a research that compares and analyzes various strategies and tactics using deep learning techniques. Moreover, Deep Learning applications are included in its scope, which may help to set the stage for further research.

In Hayashi, Taihei, et.al [9], The suggested system gathers the probability distribution of qualities that correspond to the character structure through a method, and then uses that distribution to create images of various character handwritings. By using strokes and learning from the character image data, the suggested model creates a statistical version of the character model, which is made up of probability distribution.

In Hasan, Md Mehedi, et.al [10], the suggested approach uses an adapted version of the ResNet-50 identification system and takes into account 171 handwritten characters in Bengali. The accuracy of the system is 96.15 percent, somewhat better than the previous result number of 90.33 percent.

In Chaurasia, Saumya and Suneeta Agarwal [11], To identify the longhand digits in any of the four regional languages—Bangla, Oriya, Devanagari, or Telugu—the system builds a model. It starts by normalizing the input

image, which only includes one digit. CNN is then performed to extract the features, then SVM is utilized in place of a classifier.
In Yu, Weibo, et.al [12], CNN and Softmax models are introduced in the suggested system. The role of activation is investigated in the model. The Tanh hyperbolic tangent function is employed to improvise on the rate of convergence throughout the identification process, and the ReLU function chooses to operate in lieu of the traditional sigmoid function. The use of dropout regularisation helps to prevent overfitting.
In Cheekati, et.al [13], The suggested model or system is predicated on a method of identifying Telugu longhand characters both offline and online using ResNet, a framework built on learning residually. Here's one deep neural network concept where data training can be useful. ResNet facilitates the construction of deep networks where the vanishing gradient problem—which may arise in deep CNNs—is addressed.
In Zhang, Huijie, et.al [14], The suggested strategy outlines a useful and accommodating HWR technique that should only be applied for identifying longhand digits and a small, restricted set of characters. Three steps involve the use of five different CNN classifier varieties: the Digit recognition step, the String-type classifier step, and the String recognition step.

III. METHODOLOGY

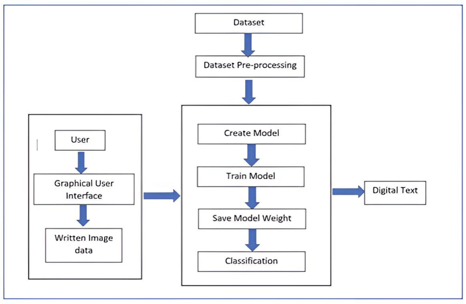

Figure 1: Architecture Design

In the fields of computer vision and artificial intelligence, handwritten text recognition (HTR) has grown in importance. Handwritten text is increasingly being converted into digital format thanks to HTR models, which are essential for handling data efficiently and the growing digitization of documents. The architecture design and workflow of a handwritten text recognition model, which consists of four separate modules- Pre-processing, Model Building, System Training, and Classification.
The accompanying figure shows the architecture design for the handwritten text recognition model. Every module has a distinct function within the overall workflow, helping to successfully convert handwritten text into digital data.
Initially we have the pre-processing module in which the process of recognizing handwritten text starts with the collection of handwritten text images. The quality, size, and orientation of these pictures could differ. The first stage in standardizing and getting these images ready for additional analysis is the pre-processing module. The pre-processing module carries out a number of tasks after obtaining the dataset photographs to guarantee consistency and best input for next phases. A basic step is to resize the pictures to the appropriate sizes. By ensuring that every image has a uniform size, this phase makes processing and model training more efficient. Furthermore, the pre-processing module could include functions like binarization, contrast enhancement, and noise reduction. By enhancing the image quality, these processes hope to facilitate the model's ability to extract pertinent characteristics in later phases.
Secondly, there is a model building module in which by using the pre-processed photos as a starting point, this module builds the machine learning model, which is the heart of the handwritten text recognition system. This lesson uses a method based on machine learning to create a model that can read and comprehend handwritten text. Choosing a suitable architecture, setting its parameters, and figuring out how many layers to include depending on the task's complexity and the processing power at hand are all part of the model-building process. In handwritten word identification challenges, a number of machine learning architectures, including recurrent neural networks (RNNs) and convolutional neural networks (CNNs), have demonstrated encouraging results. The model building module sets the network's initial weights and biases after the architecture is established, preparing the system for further training. The architecture of the model is carefully optimized to strike a balance between computational economy and performance.

Thirdly, we have the system training module in which a large amount of labeled data and processing power are needed to train a handwritten text recognition model. In order to properly train the system, the system training module uses the built-in model and the dataset photos. The model gradually gets better at reading handwritten writing as it gains recognition of patterns and characteristics in the input photos during training. In order to make sure the model generalizes successfully to new data, the training phase is essential for adjusting the model's weights and biases. Regularization and data augmentation are two further techniques that can be used to improve the resilience of the model and avoid overfitting.

Finally, we have the classification module in which the ultimate goal of the architecture design is to extract digital text data from handwritten photographs. This is accomplished by placing the classification module at the end. The model is prepared to identify text elements in fresh photos after it has been trained. The classification module uses the trained model to locate and classify text regions in a handwritten image. After segmenting and processing these regions, individual words are extracted. The output of the categorization module usually consists of digitally recognized text that can be stored, processed further, or integrated into other applications. The total functionality and performance of the handwritten text recognition system are directly impacted by the precision and effectiveness of this module.

The data flow path through any process or system is depicted in a data flow diagram (DFD). Both technical and non-technical users can benefit from data flow diagrams because they can convey concepts that are hard to understand in words visually. When levels and stages are included, a data flow diagram can become increasingly accurate. DFD Level 0 is also known as contextual diagrams. It provides a fundamental overview of the entire system or process that is being simulated or evaluated.

In DFD Level 1, the Context Level Diagram is separated more thoroughly. You will draw attention to the main tasks that the system completes when you dissect the elevated function of the context diagram into its component parts.

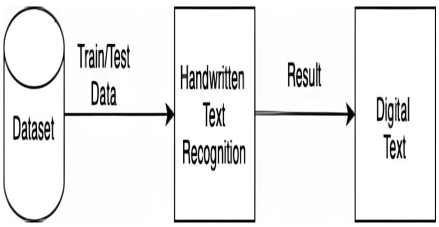

Figure 2: Dataflow Diagram Level 0

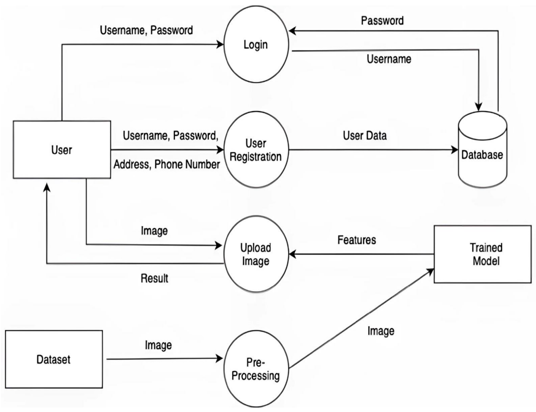

Figure 3: Dataflow Diagram Level 1

IV. Results And Discussion

This discussion focuses on the model's training performance using the handwritten words. While handwritten character recognition has been the subject of numerous studies in the past, the handwritten English language dataset has received less attention. Consequently, we concentrated on the English language dataset for this study, which is based on the English language's line characters. Preprocessing was completed first, and then the CNN architecture was updated and used to train the model.

Subsequently, the trained model is taken into account and employed to forecast the classes of the test set of the target data, which was only provided by the current data source.

The graph shows the training accuracy that our model has attained.

After obtaining the dataset, the training is completed. The model is trained on multiple data lines that are taken into consideration. We employ new target data that is not included in the training dataset once the model has been trained. The handwritten image with the words "will have" is fed as an input to the model.

For the provided data, a prediction has been made. The handwritten image is sent into the model as an input, and the model makes word predictions. The result that appears when the submit button is pressed to examine the handwritten input image's predicted outcome is shown in the image below.

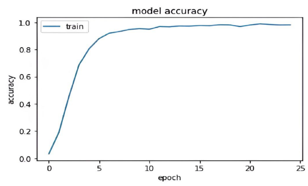

Figure 4: Graph showing training accuracy

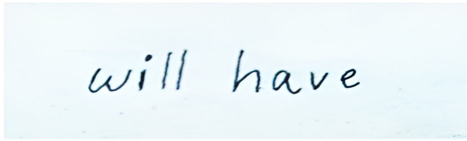

Figure 5: Handwritten image given as the input

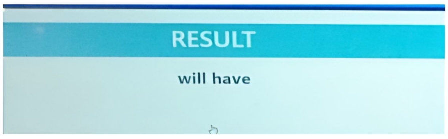

Figure 6: The result displaying the recognized text

Based on the training data, which consisted of handwritten text lines, predictions are made in this manner. 92% of the model is achieved with precision.

V. Conclusion And Future Scope

Significant progress has been made in recent years in deep learning-based handwritten text recognition. This robust technology can be used for a variety of tasks, such as digitizing financial records, historical documents, automatic data entry, and handwritten user interfaces.

Lastly, machine learning with deep learning has completely changed the way handwritten writing is recognized, offering up to date accuracy and creating new opportunities for a variety of uses. However, the quality of the data, the architectural decisions made, and the particular requirements of the task all play a major role in how well these systems perform. To make the models more functional in real-world use cases, researchers have been working hard, addressing issues, and strengthening the models' robustness.

References

[1] Cursive Handwritten Text Recognition using Bi-Directional LSTMs: A case study on Urdu Handwriting published with 978-1-7289-2020 IEEE

[2] Arabic handwriting recognition: Between handcrafted methods and deep learning techniques published with 977-1-7269-2020 IEEE

[3] Handwritten Arabic Numeral Recognition using Deep Learning Neural Networks published with 9 78-1-5090-6004-7/17/$31.00 c 2017 IEEE

[4] Deep Learning Network Architecture based Kannada Handwritten Character Recognition IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2

[5] Performance Comparison of Scanner and Camera-Acquired Data for Bangla Offline Handwriting Recognition 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)

[6] Handwritten Text Recognition using Deep Learning 2020 5th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT-2020), November 12th & 13th 2020

[7] A Proposed Framework for Recognition of Handwritten Cursive English Characters using DAG-CNN

[8] Handwriting Personality Recognition with Machine Learning: A Comparative Study published with 9 78-1-7281-6921-7/20/$31.00 ©2020 IEEE

[9] A Study of Data Augmentation for Handwritten Character Recognition Using Deep Learning 2018 16th International Conference on Frontiers in Handwriting Recognition

[10] Bengali Handwritten Isolated Compound Characters Recognition by Applying Transfer Learned Deep Convolutional Neural Network 2020 23rd International Conference on Computer and Information Technology (ICCIT), 19-21 December, 2020

[11] Recognition of Handwritten Numerals of various Indian Regional Languages using Deep Learning 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering

[12] Handwritten Digital Recognition Optimization Method based on Deep Learning published with 78-1-7281-7687-1/20/$31.00 2020 IEEE

[13] Telugu handwritten character recognition using deep residual learning IEEE Xplore Part Number:CFP20OSV-ART; ISBN: 978-1-7281-5464-0

[14] A user-adaptive deep machine learning method for handwritten digit recognition 1st IEEE International Conference on Knowledge Innovation and Invention 2018