JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Handwriting Detection System Using Image Processing.

Dr.Santosh Singh[1], Sherlyn kevin[2], Sanket Govind Sonawane[3], Shiba Ibrahim shaikh[4]

[1]HOD, Department of IT, Thakur College of Science and Commerce, Thakur Village, Kandivali(East),Mumbai, Maharashtra, India

[2]Assisatnt Proffessor, Department of IT, Thakur College of Science and Commerce, Thakur Village, Kandivali(East),Mumbai, Maharashtra, India

[3,4]PG Student, Department of IT, Thakur College of Science and Commerce, Thakur Village, Kandivali(East),Mumbai, Maharashtra, India

***Abstract*—** Because handwriting recognition is used so frequently in document analysis, biometrics, and other forensics fields, it has earned a lot of notoriety. Therefore, this project entails through investigation and the development of an advanced image processing system for handwriting detection. Therefore, the suggested method begins with the preprocessing steps, which comprise picture acquisition, noise reduction, and image enhancement, to produce better-quality photos. Subsequently, diverse methodologies such as texture analysis, contouring function, and intensity pixel histograms are utilized to distinguish specific attributes of handwritten designs from others. After that, the features are further categorized into different kinds of handwritings using classification techniques like decision trees and convolutional neural networks (CNN), among others. Scalability problems are then addressed, if applicable.

***Keywords*— TensorFlow, Image processing, Convolutional Neural Network, Handwritten Recognition.**

I. INTRODUCTION

In order to create sophisticated handwriting identification systems, this research investigates the use of machine learning algorithms and image processing techniques. The main objective is to develop flawless systems that can reliably recognize and categorize written text, hence boosting efficacy and efficiency in areas like document forgery detection and signature verification. These are also used in many other domains, such as automated form processing, forensic analysis, biometric authentication, and document authentication.

During forensic investigations, handwriting identification is crucial for confirming the criminal charges. Furthermore, handwriting detection systems are especially crucial for this reason since handwritten signatures serve as unique identifiers in the field of biometrics. Specifically, it comprises a thorough literature assessment of techniques for handwriting detection. utilized, difficulties encountered, and most recent developments in the topic at hand. Additionally, it suggests a novel approach to handwriting detection by fusing machine learning and image processing; this will include strengthening the system by addressing challenges related to handwritten document analysis

II.LITERATURE REVIEW

Systems for recognising handwriting are interesting because they can handle a wide variety of scripts and languages. The most recent approaches and advancements in offline handwritten recognition for many languages are reviewed in this review paper, with a focus on word and character recognition methods.

Malayalam Character Recognition: George and Gafoor (2014) described an offline artificial neural network (ANN) based Malayalam character recognition method. Three tiers of hidden units with a log sigmoid activation function that is trained by backpropagation make up their network. With the use of contourlet transform features, the system's accuracy rate was 97.3%.

Marathi Character Recognition: As an offline recognition system, Kale and Deshmukhy (2014) proposed the use of Basic Marathi Script, which is developed from Devanagari. They outperformed previous techniques by 0.37% thanks to their strategy, which used Zernike moment feature descriptors and SVM/K-NN classifiers to produce reliable findings.

Afroge and Ahmed (2016) introduced English alphanumeric character recognition using their work English Alphanumeric Character Recognition: An ANN-based Technique. Binary matrices for feature extraction, sequential training, and classification by them were employed, along with preprocessing images. A stunning 99% accuracy in arithmetic, 96% in lowercase, 97% in uppercase, and 93% in alphanumeric characters was attained by the system.

Character Recognition in Tifinagh: Convolutional Neural Networks (CNNs) and Deep Belief Networks (DBNs) are two types of deep learning networks that Sadouk and Gadi (2017) studied for Tifinagh character recognition. CNNs outperformed DBNs in accuracy rate, reaching up to 98.25%, while DBNs were faster and had a more straightforward design.

Bangla Digit Recognition: CNNs with Gaussian and Gabor filters are among the DL techniques for Bangla digit recognition that Alom et al. (2017) suggested. CNNs helped their system perform well in terms of accuracy.

English Handwritten Character Recognition: Using feedforward backpropagation neural networks, Kharkar and Mali (2017) created an offline recognition method. It was more accurate, dependable, and effective than previous methods.

Recognition of Latin Characters: Firmani and Merialdo (2017) presented a deep convolutional network (DCNN) classifier for Latin character optical character recognition (OCR). After training on a vast corpus of documents, the DCNN demonstrated an astounding 96% accuracy rate.

Arabic Text Recognition: Ahmad and Naz (2020) introduced MDLSTM networks, which are multi-dimensional long short term memory networks as DL method, with the aim of recognizing handwritten Arabic text. Their methodology outperformed previous methods in terms of accuracy by utilizing preprocessing and data augmentation technique

English Handwritten Word Recognition: Gurg et al. (2020) created an effective model that combines CNNs and recurrent neural networks (RNNs) to recognize English handwritten words. Their preprocessing techniques, which included data augmentation and contrast normalization, greatly increased the accuracy of the IAM dataset.

This review showcases diverse methods and advancements in handwriting recognition systems across a range of scripts, languages, and applications. Researchers have pushed the limits of the accuracy and efficiency of handwritten recognition systems, using everything from cutting-edge DL algorithms to traditional ANN-based techniques.

III.METHODOLOGY

An image processing methodology-based handwriting detecting system goes through several steps, among other things. A general blueprint for developing such systems is as follows:

Gathering Information: Assemble files with handwritten images. Having a wide variety of handwriting styles, sizes, and variances in the collection would be fantastic. Two datasets have been created from the dataset.

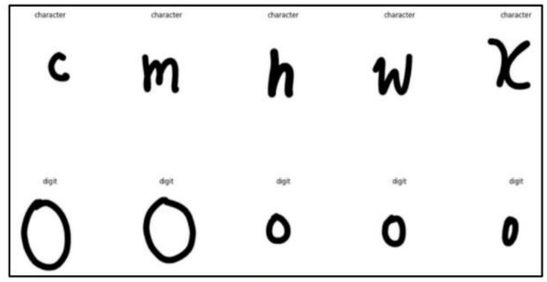

Fig 1: Classification of Dataset

Preprocessing: This step improves the quality of the photos and gets them ready for additional analysis. This could include reducing noise, scaling, normalizing, and binarizing—a process that turns images into black and white. Extraction of significant features from previously processed images is referred to as feature extraction. Among the instances are stroke width, Characteristics or descriptors that capture the distinctive qualities of handwriting, such as histograms of oriented gradients (HOG).
The CNN method is used on the dataset in the diagram below.

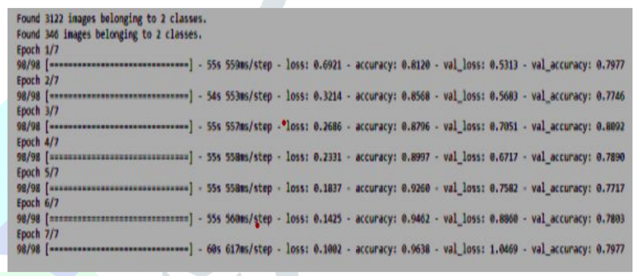

Fig 2: Applying CNN

Model Training: Use the features that were gathered to train a deep learning or machine learning model. For handwriting recognition, one can use convolutional neural networks (CNNs), recurrent neural networks (RNNs), or both together, such as convolutional-recurrent neural networks.

Testing and Validation: After using a different validation dataset to validate the trained model, modify its parameters. Evaluate a different testing dataset's accuracy, precision, recall, and related metrics before using it for model analysis.
Post-processing: To improve the final output, apply post-processing methods like as mistake correction, text smoothing, or improved character segmentation.

Integration & Deployment: Create a graphical user interface (GUI) for the model or incorporate it into an already-existing software system to create an intuitive application that allows the input of handwritten images to produce text.
Iterative Improvement: Based on user input and fresh advancements in machine learning and image processing techniques, the system's functionality and design are continuously improved. This could entail gathering additional data, streamlining preprocessing workflows, or changing the model's architecture.
The development process of a handwriting recognition system must include scalability, efficiency, and resilience in order to guarantee that the system will function well in real-world

scenarios and be able to process a wide range of input image types.

IV.RESULTS

For example, there are multiple phases involved in developing an image processing-based handwriting recognition system. Let's go over some of the fundamental procedures needed to create one of these systems.

Capture handwritten text images using various methods like scanning old documents or taking photos with a camera. Prior to processing, enhance text quality by cleaning up graphics. This involves steps such as adjusting brightness, contrast, and sharpness, as well as removing noise and artifacts to improve readability and overall appearance.

Resizing: Resize image resolution or denoise pictures by applying techniques like Gaussian blurring or median filtering to remove unwanted noise.

Binarization: Convert the image to binary format, using a foreground text and a uniform background. Techniques like thresholding can be applied here.

Skew correction: Adjust any rotation or skewness in the picture.

Text Detection: Locate in an image any places with handwritten text. Here are several techniques that can be applied.

Edge detection: Techniques like the Canny edge detector can be used to locate edges in an image. Using contour detection, one can locate contours in an image that could represent text-containing areas. Using connected component analysis, determine which connected areas in the image are most likely related to the text.

VII. CONCLUSION

In conclusion, developing an image processing system for handwriting identification has shown to be an effective method for automating the reading and evaluation of handwritten content. This has been accomplished by precisely recognizing handwritten words and characters using a variety of image processing methods, including feature extraction, edge detection, and machine learning frameworks. This method has several advantages, such as better accessibility options for those with impairments, less time spent on transcription tasks, and higher efficiency when handling handwritten materials. Furthermore, the system's versatility in handling many languages, handwriting styles, and document formats makes it useful to a variety of areas and industries.
In summary, Overall, the handwriting identification technique used in this study shows how image processing technology can drastically alter how handwritten documents are handled and used in a variety of applications. However, if research and development efforts are not stopped, this field should continue to improve in the future, leading to more advanced and trustworthy systems available for handwriting recognition.

VIII. REFERENCES

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. International Journal of Computer Vision, 115(3), 211-252.

Seales, W. B., Parker, C. S., Segal, M., Tov, E., Shor, P., & Porath, Y. (2016). From damage to discovery via virtual unwrapping: Reading the scroll from en-gedi. Science Advances, 2(9), e1601247.

C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 60.

imard, P. Y., Steinkraus, D., & Platt, J. C. (2003). Best practices for convolutional neural networks applied to visual document analysis. In ICDAR, 2003.

Souibgui, M. A., Bensalah, A., Chen, J., Fornés, A., & Waldispühl, M. (2023). A user perspective on htr methods for the automatic transcription of rare scripts: The case of codex runicus. J. Comput. Cult. Herit., 15(4).

Sternberg, S. R. (1986). Grayscale morphology. Computer Vision, Graphics, and Image Processing, 35(3), 333-355.

Tensmeyer, C., & Martinez, T. (2020). Historical document image binarization: A review. SN Computer Science, 1(3), 173.

Vats, E., Hast, A., & Singh, P. (2017). Automatic document image binarization using bayesian optimization. In HIP, page 89-94.

Wilkinson, T., & Brun, A. (2016). Semantic and verbatim word spotting using deep neural networks. In 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), pages 307-312.

Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 2242-2251.