Grenze International Journal of Engineering and Technology, June Issue

Handwritten Language Detection using Deep Learning

Sakshi D. Mahamuni[1], Sumant D. Gawade[2], Prithviraj P. Pednekar[3], Karan H. Vaghela[4] and Sachin Sambhaji Patil[5]
[1-4]Student, D Y Patil University School of Engineering and Technology, Ambi, Pune
Email: sakshi.mahamuni2802@gmail.com, yashgawade282002@gmail.com, prithvirajpednekar013@gmail.com, karanvaghela134@gmail.com
[5]Assistant Professor, D Y Patil University School of Engineering and Technology, Ambi, Pune
Email: sachin.patil@dypatiluniversitypune.edu.in

***Abstract*— Handwritten language detection plays a crucial role in various applications, such as document analysis, translation services, and forensic document examination. This research focuses on the development and evaluation of machine learning models for the automated identification of the language in handwritten texts. The proposed approach leverages advanced techniques in image processing and natural language processing to extract relevant features from handwritten documents. The first step involves preprocessing the handwritten images to enhance clarity and reduce noise. Subsequently, a combination of traditional and deep learning-based feature extraction methods is employed to capture the distinctive characteristics of different languages. The extracted features are then fed into a machine learning classifier, such as a support vector machine or a neural network, for training and validation. To assess the performance of the proposed system, a diverse dataset comprising handwritten samples from multiple languages is used. The evaluation metrics include accuracy, precision, recall, and F1 score. The experimental results demonstrate the effectiveness of the developed model in accurately identifying the language of handwritten text across various scripts and writing styles. Furthermore, the research explores the impact of dataset size, variability, and the transferability of the trained model to unseen data. The findings contribute to the advancement of handwritten language detection systems, paving the way for improved accuracy and applicability in real-world scenarios. This research has implications for document digitization, multilingual information retrieval, and linguistic analysis in forensic investigations.**

***Index Terms*— Handwritten Recognition, Deep Learning, Convolutional Neural Networks (CNN), Machine learning model, Image Processing, Pre-processing, Accuracy, Feature Extraction, Handwritten Language Detection.**

I. INTRODUCTION

Hand write recognition is a new technology that will be useful in this 21st century. It can act as base functionality for the birth of new necessaries. Operation of hand write recognition would be, processing of large set of paper document like answer scripts. With the help of hand-write recognition and AI, the answer answer Hand write recognition is a new technology that will be useful in this 21st century. It can act as base functionality for the birth of new necessaries. Operation of hand write recognition would be, processing of large set of paper document like answer scripts. With the help of hand-write recognition and AI, the answer scripts can be set without mortal

Grenze ID: 01.GIJET.10.2.222_1

involvement. Handwrite recognition constitutes a focal point that requires resolution. It represents one of the variants of Optical Character Recognition (OCR). OCR is identification of manual, which may be put out or hand-written. The task of developing an automatic handwriting recognition system presents a clear challenge. Among the various methodologies considered, the approach of using artificial neural networks is widely regarded as the most effective for creating systematic models that can accurately recognize handwritten characters. By emulating the cognitive processes of the human brain, expert systems play a significant role in achieving this objective. Moreover, reading handwritten text is a timeconsuming and tiresome process, especially when faced with the task of analyzing numerous reports written by different individuals. In this regard, a neural layered network proves to be the most suitable architectural structure as it possesses the ability to derive meaning from complex data and identify patterns that are difficult to detect using other traditional human techniques. The primary objective of this research paper is to develop a visual representation capable of efficiently analyzing images containing written numerals, symbols, and words. This goal will be accomplished by employing the concept of Convolution Neural Networks. The following sections will provide an overview of the associated task, the conceptual framework, the design process, the experimental results, and a brief conclusion. [1]

II. Proposed System

A. Deep Learning

Deep learning, a subset of machine learning, harnesses the potential of neural network models to tackle complex problems. The "deep" in deep learning alludes to the multitude of layers that transform the data. It serves as the epitome of the fusion between machine learning and artificial neural networks (ANN). The learning process itself can be supervised, unsupervised, or semi-supervised. Prominent deep learning architectures include Deep Neural Networks (DNN), Deep Belief Networks (DBN), Recurrent Neural Networks (RNN), and Convolutional Neural Networks (CNN). These architectures have made significant strides in domains such as speech recognition, natural language processing (NLP), and audio recognition. The ability of deep learning algorithms to handle unsupervised learning tasks is particularly beneficial considering the abundance of unlabeled data compared to labeled data. Through deep learning, neural network models inspired by the human brain's functioning are utilized to address complex problems. These models, known as artificial neural networks, have multiple hidden layers that allow for the thorough analysis of data. Deep learning encompasses a range of learning types, including supervised, unsupervised, and semisupervised. In the field of machine learning, deep learning algorithms are particularly effective when employed in unsupervised learning tasks. [2]

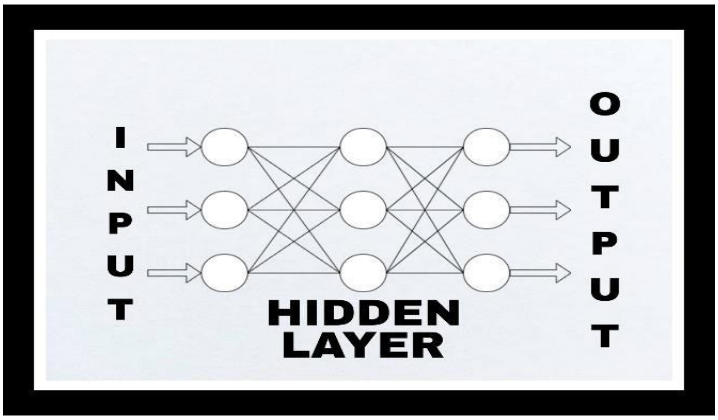


Fig:1. Hidden Layer diagram for Input and Output

B. Convolutional Neural Network(CNN)

Convolution Neural Network is a Deep Learning architecture which is mainly used for processing structured grid data (such as images). It uses convolutional operations to input data. Pooling Layers are used to reducing the resolution. Lower layers capture local patterns, while higher layers combine these patterns to represent more complex structures. 2D CNNs for images, 3D CNNs are used for processing video data. Convolutional layers use shared weights for the filters across different spatial locations. This parameter sharing significantly reduces the number of parameters in the model, making it more efficient and capable of learning translationinvariant features. Each neuron in a convolutional layer is connected to a local region in the input, known as its receptive field.

Activation functions, such as ReLU (Rectified Linear Unit), are applied to the output of convolutional layers, introducing non-linearity and enabling the network to learn complex relationships and representations. Following convolutional and pooling layers, fully connected layers process the global information captured in the higher-level features. These layers make the final decision based on the learned representations. CNNs are often used in transfer learning scenarios, where a pre-trained model on a large dataset (e.g., ImageNet) is fine-tuned for a specific task. The lower layers retain knowledge about general features, while the higher layers adapt to task-specific features. Convolutional Neural Networks have proven highly effective in computer vision tasks, including image classification, object detection, image segmentation, and facial recognition. Their architecture is tailored to exploit the spatial relationships present in grid-like data, making them a fundamental tool in image processing and analysis. [2]

C. Past Work

Handwriting text recognition (HTR) is a specific instance of optical character recognition (OCR). While both topics have been extensively studied, the focus of this paper is on HTR. In a comprehensive survey, HTR approaches are categorized into online and offline methods based on the type of data they utilize. Online methods have access to the pen location as the text is being written, enabling them to differentiate intersecting strokes. Conversely, offline methods only have access to the final text image, which may contain background noise or clutter. Although online methods have the advantage of better data quality, they require additional equipment like a touchscreen to capture pen stroke data, making it challenging to gather large quantities of online data in natural settings. Additionally, online methods are unsuitable for analyzing historical manuscripts and markings that are entirely offline. Therefore, this manuscript focuses solely on offline methods and excludes online approaches.[3]The field of handwriting text generation (HTG) is relatively young but promising. Graves's method was among the first to synthesize online data based on a recurrent neural network. Ji et al. further expanded on this concept by adopting the generative adversarial network (GAN) paradigm and introducing a discriminator. Deep Writing improved upon Graves's method by separating style generation from content. Haines et al. proposed a handwriting generation method that specifically focuses on emulating a particular author. However, this approach necessitates a time-consuming character-level annotation process for each new data sample[3].

D. Offline Handwitten Text Recognition

Offline handwritten text recognition poses significant challenges due to the extensive length of its sentences, which can reach up to 90 characters. Additionally, the problem is compounded by various writing styles and the complexity of accurately recognizing individual characters. With these complexities in mind, it becomes an ideal testing ground for assessing the robustness and effectiveness of the Decoupled Attention Network. To ensure thorough evaluation, we also conducted experiments using two commonly used attentional decoders: Bahdanau's attention method (Bahdanau, Cho, and Bengio 2015) and Luong's attention method (Luong, Pham, and Manning 2015). These attentional decoders have been widely employed in text recognition studies (Shi et al. 2018; Cheng et al. 2018; Luo, Jin, and Sun 2019; Li et al. 2019). To maintain fairness in our comparisons, we replaced the CAM and decoupled text decoder with these methods.[4]

E. Dataset

TextBlob is a powerful library designed for working with textual data. It offers an intuitive API for various natural language processing tasks, including part-of-speech tagging, sentiment analysis, translation, and much more. With TextBlob, you can effortlessly perform these NLP operations without the need for extensive knowledge of complex algorithms.With TextBlob, you can train your own text classifiers for various tasks like detecting spam or categorizing topics. This flexibility allows for customized solutions to meet specific requirements. TextBlob offers features for word inflection and lemmatization, enabling the normalization of words and reducing them to their base or root forms. This linguistic processing assists in standardizing text and simplifying subsequent analyses. TextBlob also facilitates language translation. Through the translate method, you can effortlessly translate text from one language to another, expanding the reach and usability of your applications. Integrating with WordNet, a lexical database of the English language, TextBlob provides access to valuable information about word meanings, synonyms, antonyms, and more. This integration enhances the comprehensiveness and accuracy of your NLP tasks. One of TextBlob's greatest strengths lies in its user-friendly design. The API is intentionally crafted to be easy to use, making it an ideal choice for developers and learners alike. Additionally, TextBlob's flexible nature allows users to incorporate custom models, corpora, or other resources, expanding the library's functionality. TextBlob serves as a convenient tool for applications requiring quick and straightforward NLP functionality.

Whether you need efficient text analysis or robust prototyping of NLP tasks in Python, TextBlob delivers the necessary capabilities.

III. IMPLEMENTATION

A. Processing

As shown in Fig:2, Pre-processing and feature extraction are critical steps in the automatic recognition of handwritten digits from documented images[5]. The primary goal is to enhance the discriminative ability of the pixels or raw features derived from the input images. A significant amount of effort has been dedicated to enhancing the preprocessing[6]. One challenge encountered in the recognition process relates to the detection and correction of skew/slant in documented images, which poses difficulties for segmentation[7].

B. Letter Division

As shown in Fig:2, The recognition of handwritten characters involves various complexities, with division being the most daunting aspect[8]. The main cause of this challenge lies in the uncertain size, number, and gaps between characters[9]. An effective solution to overcome this hurdle is the implementation of a flawless character segmentation algorithm. With the inclusion of rules, such as individual character boxing and identifying gaps between strings, the segmentation of handwritten characters becomes significantly easier. By employing a basic algorithm, these string types can be segmented through connected component analysis, after the removal of any noise elements[10]. The primary objective of this research is to develop an ideal segmentation algorithm for accurate handwritten character recognition.

Structure of a Model

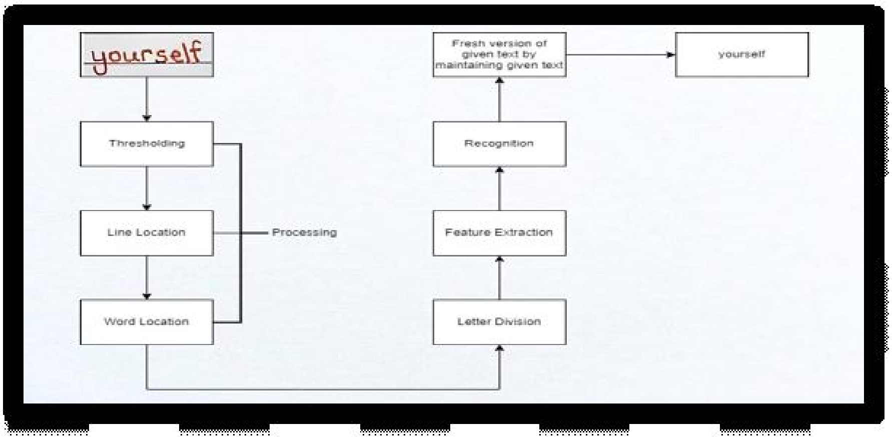


Fig:2. Process for analyzing handwritten text

Feature Extraction

As shown in Fig:2, Feature extraction involves the process of extracting the distinctive elements of characters from a given image sample[11]. There are basically two types of feature extraction:

1. Statistical feature extraction: In statistical feature extraction the feature vector is the combination of all the features extracted from each character and these association feature vector relative to positions of features in the character image matrix.
2. Structural feature extraction: In structural feature extraction extracts the morphological features of a character from the image matrix. It considers the edges, curvature, regions, etc. The functions that are used in feature extraction for indexing and labeling the dataset and it helps in the classification and recognition of handwritten characters.

Recognition

As shown in Fig:3, The handwritten character recognition system is based on the input image, pre-process the image, segment the image, extract the feature of the image and classify the character that is based on feature extracted module, and the proposed handwritten character recognition[12], [13]. In the following, we briefly describe feature generation method, classifiers and the technique of recognition the character.

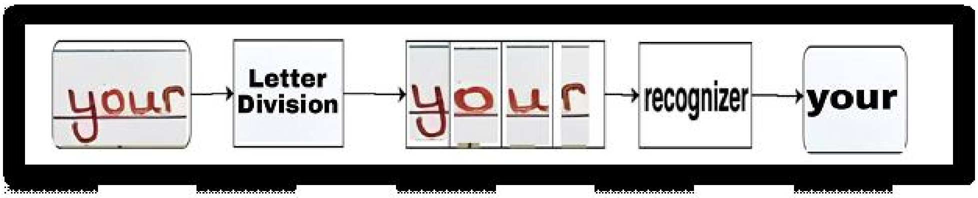


Fig: 3. Character Recognizer

IV. Result

Accuracy in English, Hindi, Marathi Text

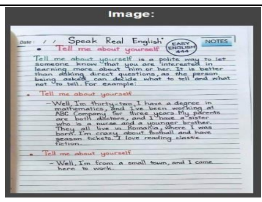
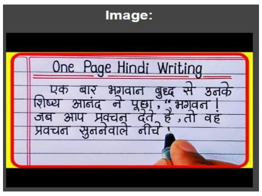
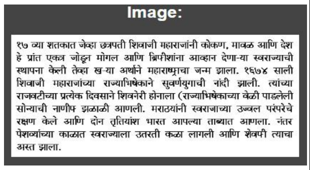
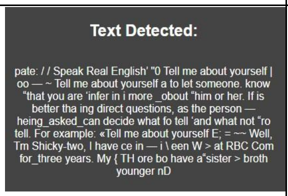
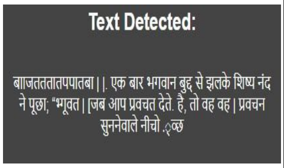
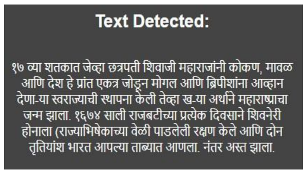
	English	Hindi	Marathi
Input Image			
Output Text	 Text Detected: pate: / / Speak Real English' "0 Tell me about yourself | oo — ~ Tell me about yourself a to let someone. know "that you are 'infer in i more _obout "him or her. If is better tha ing direct questions, as the person — heing_asked_can decide what fo tell 'and what not "ro tell. For example: «Tell me about yourself E; = ~~ Well, Tm Shicky-two, I have ce in — i \ een W > at RBC Com for_three years. My { TH ore bo have a"sister > broth younger nD	 Text Detected: बााजततततातपपातबा | |. एक बार भगवान बुद्द से झलके शिष्य नंद ने पूछा; "भूवत | [जब आप प्रवचत देते. है, तो वह वह | प्रवचन सुननेवाले नीचो .्वछ	 Text Detected: १७ व्या शतकात जेव्हा छत्रपती शिवाजी महाराजांनी कोकण, मावळ आणि देश हे प्रांत एकत्र जोडून मोगल आणि ब्रिपीशांना आव्हान देणा-या स्वराज्याची स्थापना केली तेव्हा ख-या अर्थाने महाराष्ट्राचा जन्म झाला. १६७४ साली राजबटीच्या प्रत्येक दिवसाने शिवनेरी होनाला (राज्याभिषेकाच्या वेळी पाडलेली रक्षण केले आणि दोन तृतियांश भारत आपल्या ताब्यात आणला. नंतर अस्त झाला.

Fig:4. Actual code Input and Output

Accuracy in English: 90%
Accuracy in Hindi:70-80%
Accuracy in Marathi:80%

V. Conclusion

In Conclusion handwritten language detection using machine learning is a valuable and versatile application with numerous real-world use cases. It involves the development of models and systems that can automatically identify and classify handwritten text into different languages or scripts. Here are some key points to summarize the significance and implications of this technology.

Multilingual Recognition: Handwriting language detection enables the automatic recognition of multiple languages or scripts within handwritten documents. This is particularly useful in diverse linguistic and cultural contexts.

Automated Data Entry: It facilitates the automation of data entry tasks, making it more efficient and reducing the risk of human error when dealing with handwritten text. **Historical Document Analysis:** Handwritten language detection can be applied to historical documents, allowing researchers and archivists to digitize and analyse valuable handwritten records from the past. **Cursive and Varied Handwriting:** The technology can handle cursive writing and different styles of handwriting, making it adaptable to a wide range of handwritten text.

Acknowledgment

We thank our guide, Sachin Sambhaji Patil, Assistant Professor, D Y Patil University for his remarkable supervision. We thank him for his relentless support, patience, and encouragement. This work would not be possible without his immense knowledge and exceptional guidance.

REFERENCES

[1] Singh, J., Kohli, S., Sharma, T., Goyal, M. and Singh, N., Handwritten Character Recognition using Artificial Neural Networks.

[2] Thangamariappan, P. and Pamila, J.C., 2020. Handwritten recognition by using machine learning approach. International Journal of Engineering Applied Sciences and Technology, 4(11), pp.2455-2143

[3] Fogel, S., Averbuch-Elor, H., Cohen, S., Mazor, S. and Litman, R., 2020. Scrabblegan: Semi-supervised varying length handwritten text generation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 4324-4333).

[4] Wang, T., Zhu, Y., Jin, L., Luo, C., Chen, X., Wu, Y., Wang, Q. and Cai, M., 2020, April. Decoupled attention network for text recognition. In Proceedings of the AAAI conference on artificial intelligence (Vol. 34, No. 07, pp. 12216-12224).

[5] Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R. and LeCun, Y., 2013. Overfeat: Integrated recognition, localization and detection using convolutional networks. arXiv preprint arXiv:1312.6229.

[6] Moghaddam, R.F. and Cheriet, M., 2012. AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization. Pattern Recognition, 45(6), pp.2419-2431.

[7] Bansal, K. and Kumar, R.G., 2012. Cleaning and Recognition of Numerals in a Handwritten Devnagari Document Consisting of Roll Numbers and Marks (Doctoral dissertation).

[8] Knerr, S., Personnaz, L. and Dreyfus, G., 1992. Handwritten digit recognition by neural networks with single-layer training. IEEE Transactions on neural networks, 3(6), pp.962-968.

[9] Ribas, F.C., Oliveira, L.S., Britto Jr, A.S. and Sabourin, R., 2013. Handwritten digit segmentation: a comparative study. international journal on document analysis and recognition (IJDAR), 16(2), pp.127-137.

[10] Elnagar, A. and Alhajj, R., 2003. Segmentation of connected handwritten numeral strings. Pattern Recognition, 36(3), pp.625-634.

[11] Peng, H., Long, F. and Ding, C., 2005. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on pattern analysis and machine intelligence, 27(8), pp.1226-1238.

[12] Lopes, G.S., da Silva, D.C., Rodrigues, A.W.O. and Reboucas Filho, P.P., 2016. Recognition of handwritten digits using the signature features and Optimum-Path Forest Classifier. IEEE Latin America Transactions, 14(5), pp.24552460.

[13] Liu, C.L., Nakashima, K., Sako, H. and Fujisawa, H., 2004. Handwritten digit recognition: investigation of normalization and feature extraction techniques. Pattern Recognition, 37(2), pp.265-279.