iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IX **Month of publication:** September 2024

DOI: https://doi.org/10.22214/ijraset.2024.64315

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 12 Issue IX Sep 2024- Available at www.ijraset.com

An Overview of Handwritten Text Recognition Using Generative AI: Cutting-Edge Approaches, Key Challenges, and Future Directions

Roshan R P[1], Swetha K[2], Vimala Mathew[3]
[1, 2, 3]*Information Technology Group, National Institute of Electronics and Information Technology, Calicut, Kerala, India*

Abstract: ***Recognizing handwritten text involves several stages, including data collection, data preparation, feature extraction, and categorization. It is essential for tasks like document processing, robotic automation, and historical document analysis, given the challenges of variations in size and shape. The survey focuses on enhancing the model's capabilities through data augmentation. Additionally, it delves into the intricate realm of recognizing handwritten numbers, presenting relevant work and emphasizing the use of various methods such as neural networks, k-nearest neighbor (KNN), and Support Vector Machine (SVM) for classifying handwritten numbers. The main goal of this research is to develop a technique using image processing methods to detect handwritten text within an image, addressing research gaps, challenges, and prospects in the field of text recognition.***

Keywords: Generative AI, Artificial Intelligence, Deep Learning, Handwriting Recognition, Text Recognition.

I. INTRODUCTION

Computer vision is a pivotal domain within artificial intelligence (AI) that empowers intelligent software systems to extract meaningful information from various forms of graphical data, including text, personal photographs, fingerprints, and health-related information. Despite advancements in this field, the recognition of handwritten text (HWT) remains a significant challenge due to the variability in handwriting styles among individuals and the complexities introduced by different languages [1]. Handwritten text can be processed through two primary methods: offline, which involves extracting text from images and converting it into digital formats such as ASCII code, and online, which captures text as it is being written using specialized devices.

Numerous methods have been developed for handwritten text recognition (HWTR), ranging from traditional approaches that segment each character to modern techniques that utilize machine learning algorithms to detect and recognize entire segments of text [2]. Handwritten text recognition provides a practical solution for various applications, including automated document processing, enhancing human-robot interaction, and enriching historical document analysis.

Convolutional Neural Networks (CNNs) have transformed the landscape of HWTR by effectively learning and extracting complex features from image data. Optical Character Recognition (OCR)plays a crucial role in this process by converting images of handwritten or printed text into machine-readable formats [3]. However, recognizing handwritten text poses greater difficulties compared to machine-printed text due to the non-uniformity in handwriting styles.

In OCR systems, datasets are categorized into offline and online types [4]. Offline datasets are preloaded into classifiers for training, while online datasets record the coordinates of pen movements during writing to facilitate real-time prediction. The challenges associated with handwriting recognition include variations in individual handwriting styles, inconsistencies in writing quality over time, and difficulties in collecting high-quality datasets [5].

This survey paper aims to explore the current state-of-the-art techniques in handwritten recognition using generative AI methodologies. By examining both traditional and modern approaches, we will provide insights into their effectiveness and applicability across different contexts, ultimately contributing to advancements in the field of handwritten text recognition [6].

In the upcoming parts, Section 2 will provide a literature review, summarizing key research and developments in handwritten text recognition using Generative AI. Section 3 will explore the different approaches utilized for handwritten recognition, including traditional methods and state-of-the-art AI techniques. Section 4 will offer a detailed analysis of the pros and cons of the methods outlined in Section 3, evaluating their strengths, limitations, and application areas. Finally, the survey will conclude by summarizing the findings and proposing future research directions in the last section.

©IJRASET: All Rights are Reserved

II. LITERATURE REVIEW

Over the past two decades, significant advancements in Artificial Intelligence (AI) have greatly enhanced the identification and recognition of handwritten text. Research in Handwritten Text Recognition (HWTR) utilizing TensorFlow has primarily focused on transcribing scanned documents that contain handwritten text [7]. A key dataset often used for training in this domain is the IAM dataset, which comprises a large collection of handwritten English text samples. To process these images effectively, researchers have turned to powerful neural network architectures that integrate multiple techniques, including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Connectionist Temporal Classification (CTC) [8]. These models have demonstrated impressive results, with some achieving accuracy levels exceeding 90%, showcasing the power of deep learning in HWTR.

A comprehensive evaluation of various methods employed in offline handwritten text and character recognition has highlighted the critical components of the process. These stages include preprocessing, image acquisition, segmentation, classification, feature extraction, and the final recognition phase [9]. Each of these steps plays a pivotal role in the overall accuracy of the text recognition system. The accuracy of the output depends greatly on the type and quality of the input data. For instance, a study titled "Handwritten Text Recognition and Digital Text Conversion" [10] explored the effectiveness of neural networks in recognizing segmented words from processed images. This research utilized the IAM dataset and achieved an accuracy rate of 75% for the words that were included in the training phase, emphasizing the importance of high-quality training data for reliable recognition.

Further reviews have delved into a range of techniques used in handwritten text recognition. Key methods include Line and Word Segmentation, Incremental and Semi-Incremental Recognition Methods, Part-Based Methods, as well as Slope and Slant Correction Methods. Each of these methods offers distinct advantages and comes with its own set of challenges, particularly when it comes to detecting, extracting, and recognizing handwritten text across various document formats [11]. The performance of each method often depends on the type of handwriting being processed and the specific features of the text.

One notable review examined part-based handwritten character recognition (PBHW) and experimentally evaluated three different methods [12]. The class distance method emerged as the most accurate, while the single voting method performed the worst. This analysis also underscored the role of distinct character styles in achieving high accuracy, particularly in the recognition of the English language. By contrast, whole word recognition was found to be less accurate due to the wide variation in individual writing styles. The complexity of recognizing Arabic handwriting presents additional challenges compared to English, primarily due to the language's intricate structure and the lack of clear letter separation [13]. To address these challenges, a study proposed a deep learning solution that leveraged morphological gradients and a Multilayer Neuron Network, resulting in 100% accuracy under certain conditions, depending on the quality of the training database used [14].

In conclusion, the research highlights significant advancements in handwritten text recognition through AI, illustrating how different methodologies and neural network architectures are contributing to improved accuracy across various languages and writing styles. Despite these advancements, challenges remain, particularly in recognizing more complex scripts, but the future of HWTR looks promising as AI continues to evolve.

III.DISCUSSION AND RESULT

The methods and approaches delineated in the preceding sections are crucial for text recognition, encompassing both printed and handwritten text. However, specific methodologies may exhibit expedited and more efficient outcomes in particular scenarios, while others may demonstrate a slower and less effective performance in diverse contexts. The ensuing table furnishes a comprehensive overview of the principal disparities, advantages, and limitations of all the discussed techniques and strategies.

TABLE I
REVIEWED APPROACHES FOR FINDING AND RECOGNIZING HANDWRITTEN CONTENT

S.NO	Approaches	Description
I.	Connectionist Temporal Classification (CTC)	The approach involves combining recognized words into lines and blocks [15].
II.	Large Language Models (LLM) Optical Character Recognition (OCR)	It is a process of digitizing, categorizing, and summarizing documents using AI, which represents a revolutionary approach in mechanical and electronic methods that aid in converting visual text representations into a digital format [16].

III.	Free Form Reader	Each row is dedicated to a specific question, and each column is dedicated to a single answer. If the system detects an "Error" cell or more than one answer marked, it will raise an alarm, prompting a human operator to review and input the correct answer manually [17].
IV.	Conditional Generative Adversarial Networks (cGANs)	Using the Frechet Inception Distance (FID) metric to determine the best hyperparameters, specifically for variable-length samples such as handwritten text images [18].
V.	Controllable generative model	Training a controllable generative model to generate missing data involves optimizing data synthesis and strategically integrating synthetic and authentic data to train recognition models, ensuring that both content and style are effectively encompassed.
VI.	Handwritten Text Recognition (HTR)	• Pre-Processing • Segmentation • Feature Extraction • Classification • Post Processing [19]
VII.	LSTM (Long Short-Term Memory) and a Connectionist Temporal Classification (CTC)	The system comprises a line and word segmentation module, along with a neural recognition model. The methods include Word Segmentation, Line Segmentation, Segmentation Methods, and Spellchecking [20].
VIII.	Generative Adversarial Networks (GAN)	The first set of new data is sourced from the same statistical distribution as the training set, while the second set aims to distinguish data generated by the first set from the data in the original training set [21].
IX.	BiLSTM-CTC architecture with generated synthetic handwritten words	The process entails the creation of two distinct types of extensive and varied handwritten word datasets: • overlapped • non-overlapped. • Subsequently, the synthetic images are input into the CNN model to extract the features [22].
X.	Off-Line Automatic Assessment System (OFLAAS)	• The Gaussian Grid (GGF) • Modified Direction Feature Extraction (MSF) Techniques [23].
XI.	Capsule Networks (CapsNets)	• EMNIST balanced • EMNIST letters • EMNIST digits [24]
XII.	Two-layer CNN architecture	• Automated Student Assessment Prize (ASAP) dataset • Input image is processed to gradually extract various features using convolution layers, pooling layers, and fully connected layers [25].

IV. CHALLENGES

The generative network's inability to correct grammatical errors that are deemed valid according to the displayed text in Connectionist Temporal Classification (CTC) is a significant issue [15]. Despite the effectiveness of Optical Character Recognition (OCR), instances of misinterpreted characters can arise. The final stage, encompassing error detection, correction, and data extraction, presents notable challenges. Inaccuracies in OCR have the potential to propagate errors in AI-generated corrections and analyses, particularly in contexts involving sensitive data identification and document authentication, where inaccuracies can carry substantial legal and privacy implications.

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

Moreover, while AI demonstrates proficiency in extracting specific information and analysing sentiments, its comprehension is confined to explicit text, often overlooking subtleties, sarcasm, or implied meanings crucial for fully grasping a document's context [16]. When the system identifies a marked "Error" cell or two or more marked answers, it triggers an alarm, necessitating human operator intervention to inspect and manually input the correct answer in the Free Form Reader [17]. In the Handwriting Recognition System, post-processing enhancements should incorporate spellchecker awareness of the system's operation on handwritten and recognized data [20]. Addressing overconfidence and overfitting issues necessitates fine-tuning the complex structure of the Generative Adversarial Networks (GAN) [18]. The OCR domain presents myriad challenges, encompassing variations in writing and style, poor source quality, and more. While the performance on printed documents is commendable, it exhibits deficiencies in handling overlapping handwritten words using the skew detection method [23]. Segmentation methods further necessitate a reliable ground truth or labelled data, which poses inherent difficulties [23]. Offline recognition techniques are not only costly but also at times unreliable, potentially leading to the loss of critical information [20],[24]. Although Convolutional Neural Networks (CNNs) excel in understanding low-level and high-level image features, they sacrifice valuable information at pooling layers [26]. The resource demands of conditional Generative Adversarial Networks (cGANs) surpass those of standard GANs due to the added conditioning information. Designing effective cGAN architectures is a complex endeavour, as flawed designs can yield unstable or inaccurate outputs [22]. Ensuring that the training data is diverse and representative of the desired outputs, along with having sufficient labelled data to train the model on specific control parameters, is paramount [27],[28]. The integration of language models to comprehend the context and semantics of text is essential yet challenging in Controllable generative model [29],[30].

V. CONCLUSION

The integration of generative AI with handwriting recognition technology presents a promising opportunity to enhance the precision and efficiency of HWR systems. By facilitating the creation of high-quality training data and refining error correction mechanisms, generative AI not only mitigates the inherent challenges posed by handwriting variability but also lays the groundwork for more dependable applications across a wide spectrum of industries. As ongoing research in this domain continues to progress, we can anticipate further advancements that will fine-tune these technologies, rendering handwritten input more accessible and usable in digital formats. This continuous development is poised to drive the widespread adoption of HWR solutions, reshaping the processing and utilization of handwritten documents across various sectors.

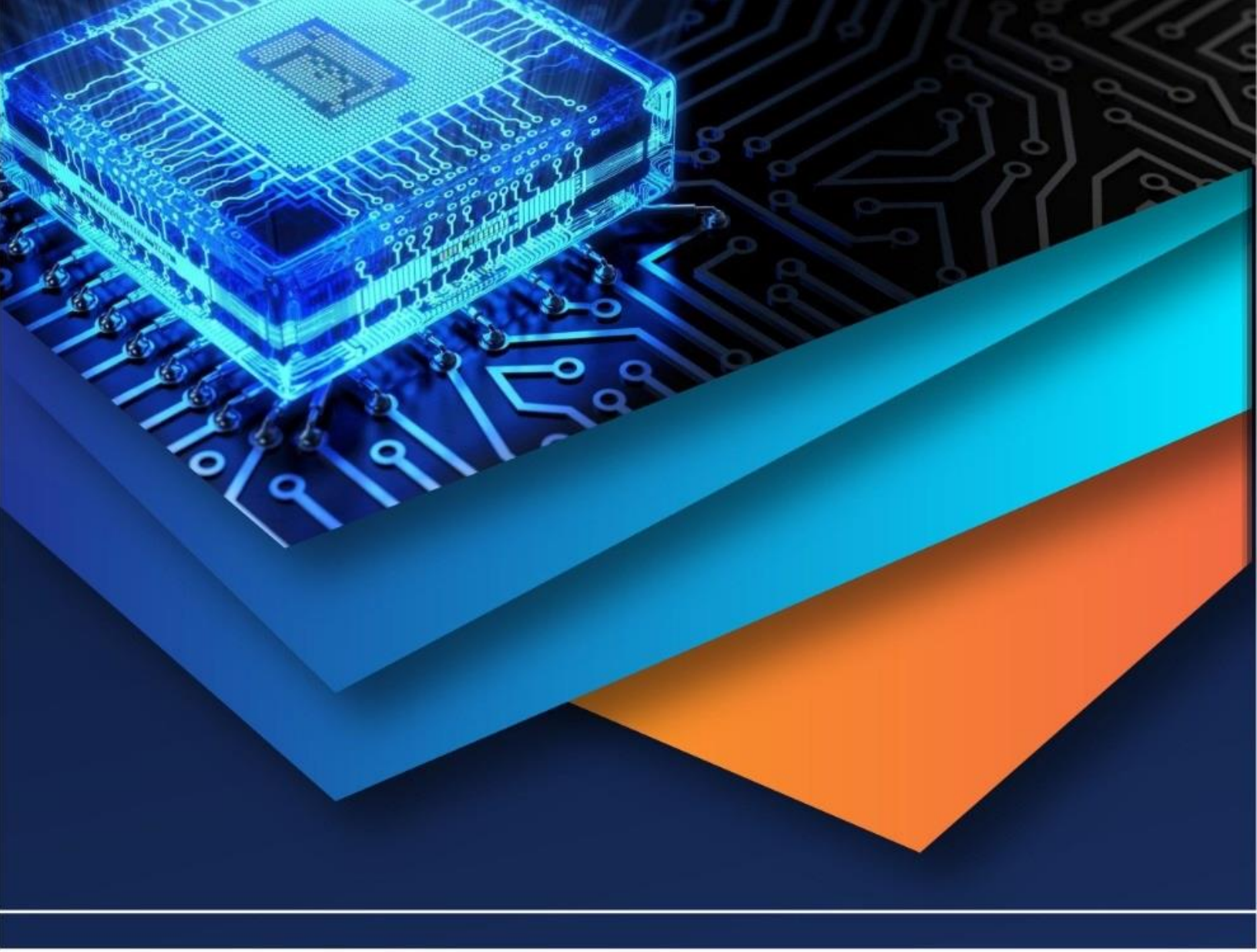
REFERENCES

[1] Mbida, M., & Ezzati, A. (2022). Artificial intelligence auscultation system for physiological diseases. International Journal on Technical and Physical Problems of Engineering (IJTPE), (49), 97-103.
[2] Fanany, M. I. (2017, May). Handwriting recognition on form document using convolutional neural network and support vector machines (CNN-SVM). In 2017 5th international conference on information and communication technology (ICoIC7) (pp. 1-6). IEEE.
[3] Dhande, P., & Kharat, R. (2017, May). Recognition of cursive English handwritten characters. In 2017 International Conference on Trends in Electronics and Informatics (ICEI) (pp. 199-203). IEEE.
[4] Khamparia, A., & Singh, K. M. (2019). A systematic review on deep learning architectures and applications. Expert Systems, 36(3), e12400.
[5] Soni, S., & Bhushan, B. (2019, July). Use of Machine Learning algorithms for designing efficient cyber security solutions. In 2019 2nd international conference on intelligent computing, instrumentation and control technologies (ICICICT) (Vol. 1, pp. 1496-1501). IEEE.
[6] Manchanda, C., Rathi, R., & Sharma, N. (2019, October). Traffic density investigation & road accident analysis in India using deep learning. In 2019 international conference on computing, communication, and intelligent systems (ICCCIS) (pp. 501-506). IEEE.
[7] Manchala, S. Y., Kinthali, J., Kotha, K., Kumar, J. J. K. S., & Jayalaxmi, J. (2020). Handwritten text recognition using deep learning with Tensorflow. International Journal of Engineering and Technical Research, 9(5).
[8] Odeh, A., Odeh, M., Odeh, H., & Odeh, N. (2022). Hand-written text recognition methods: Review study.
[9] Sahu, V. L., & Kubde, B. (2013). Offline handwritten character recognition techniques using neural network: a review. International journal of science and Research (IJSR), 2(1), 87-94.
[10] Reddy, M. B. R., Nandini, J., & Sathwik, P. S. Y. (2019). Handwritten text recognition and digital text conversion. International Journal of Trend in Research and Development, 3(3), 1826-1827.
[11] Rosyda, S. S., & Purboyo, T. W. (2018). A review of various handwriting recognition methods. International Journal of Applied Engineering Research, 13(2), 1155-1164.
[12] Song, W., Uchida, S., & Liwicki, M. (2011, September). Comparative study of part-based handwritten character recognition methods. In 2011 International Conference on Document Analysis and Recognition (pp. 814-818). IEEE.
[13] Patel, M., & Thakkar, S. P. (2015). Handwritten character recognition in english: a survey. International Journal of Advanced Research in Computer and Communication Engineering, 4(2), 345-350.
[14] El Atillah, M., & El Fazazy, K. (2020). Recognition of Intrusive Alphabets to the Arabic Language Using a Deep Morphological Gradient. Rev. d'Intelligence Artif., 34(3), 277-284.

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

[15] Yakovchuk, O., & Vasin, M. (2023). Increasing the accuracy of handwriting text recognition in medical prescriptions with generative artificial intelligence. Technology audit and production reserves, 4(2/72), 18-21.
[16] Abdelaziz, T. A. I., & Fazil, U. (2023). Applications of integration of AI-based Optical Character Recognition (OCR) and Generative AI in Document Understanding and Processing. Applied Research in Artificial Intelligence and Cloud Computing, 6(11), 1-16.
[17] Supic, M., Brkic, K., Hrkac, T., Mihajlović, Ž., & Kalafatić, Z. (2014, May). Automatic recognition of handwritten corrections for multiple-choice exam answer sheets. In 2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (pp. 1136-1141). IEEE.
[18] Kang, L., Riba, P., Rusinol, M., Fornes, A., & Villegas, M. (2021). Content and style aware generation of text-line images for handwriting recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(12), 8846-8860.
[19] Chang, J. H. R., Bresler, M., Chherawala, Y., Delaye, A., Deselaers, T., Dixon, R., & Tuzel, O. (2022, May). Data incubation—synthesizing missing data for handwriting recognition. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 4188-4192). IEEE.
[20] Tayal, P., Shetty, P. P., Pratheeksha, P., Harshita, R., & Sreenath, M. V. (2023, November). Evaluation of Handwritten Descriptive Responses Using Machine Learning-A Survey. In 2023 2nd International Conference on Futuristic Technologies (INCOFT) (pp. 1-5). IEEE.
[21] Gold, C., & Zesch, T. (2020, September). Exploring the impact of handwriting recognition on the automated scoring of handwritten student answers. In 2020 17th international conference on frontiers in handwriting recognition (ICFHR) (pp. 252-257). IEEE.
[22] Du, Y., & Yau, C. J. (2021, September). Handwriting Image Recognition Based on a GAN Model. In 2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE) (pp. 487-491). IEEE.
[23] Akter, M. S., Shahriar, H., Cuzzocrea, A., Ahmed, N., & Leung, C. (2022, December). Handwritten word recognition using deep learning approach: A novel way of generating handwritten words. In 2022 IEEE International Conference on Big Data (Big Data) (pp. 5414-5423). IEEE.
[24] Suwanwiwat, H., Nguyen, V., & Blumenstein, M. (2012, December). Off-line restricted-set handwritten word recognition for student identification in a short answer question automated assessment system. In 2012 12th International Conference on Hybrid Intelligent Systems (HIS) (pp. 167-172). IEEE.
[25] Suwanwiwat, H., Blumenstein, M., & Pal, U. (2015, July). Short answer question examination using an automatic off-line handwriting recognition system and a novel combined feature. In 2015 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.
[26] Jayasundara, V., Jayasekara, S., Jayasekara, H., Rajasegaran, J., Seneviratne, S., & Rodrigo, R. (2019, January). Textcaps: Handwritten character recognition with very small datasets. In 2019 IEEE winter conference on applications of computer vision (WACV) (pp. 254-262). IEEE.
[27] Rahaman, M. A., & Mahmud, H. (2022). Automated evaluation of handwritten answer script using deep learning approach. Transactions on Machine Learning and Artificial Intelligence, 10(4).
[28] Sethi, R., & Kaushik, I. (2020, April). Hand written digit recognition using machine learning. In 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT) (pp. 49-54). IEEE.
[29] Sivaanandh, M., Surya, S., & Priyanka, G. (2018, July). Hand written Indian numeral character recognition using deep learning approaches. In 2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE) (pp. 1301-1304). IEEE.
[30] Padmasree, P., & Maheswari, R. A Novel Technique for Image Compression in Hand Written Recognition using Back Propagation in Neural Network. vol, 4, 763-768.

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



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)