



# **iJRASET**

International Journal For Research in  
Applied Science and Engineering Technology



---

# **INTERNATIONAL JOURNAL FOR RESEARCH**

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume:** 12    **Issue:** XII    **Month of publication:** December 2024

**DOI:** <https://doi.org/10.22214/ijraset.2024.65737>

**[www.ijraset.com](http://www.ijraset.com)**

**Call:** ☎ 08813907089

**E-mail ID:** [ijraset@gmail.com](mailto:ijraset@gmail.com)

# A Review on Automatic Subjective Answer Evaluation

Dr. Dipak D. Bage<sup>1</sup>, Tina H. Deore<sup>2</sup>, Samarth S. Abak<sup>3</sup>, Shruti D. Godse<sup>4</sup>, Vishakha P. Mandawade<sup>5</sup>, Dr. Aniruddha S. Rumale<sup>6</sup>

<sup>1</sup>Associate Professor, <sup>2,3,4,5</sup>Research Scholar, <sup>6</sup>Professor, Department of Information Technology  
Sandip Foundation's Sandip Institute of Technology and Research Centre, Nashik

**Abstract:** *Offline handwritten text recognition from images is a long-standing research topic for researchers trying to digitize a large number of hand-scanned documents. We propose a novel neural network architecture that integrates feature extraction, sequence learning, and transcription into a unified framework. The documentation process converts paper-based and handwritten text into electronic information. Extracting text lines, keywords, and images from such a complex document can be a difficult task. Identifying text lines from handwritten or printed document images is a crucial stage in the optical character recognition (OCR) system process. This paper addresses the challenge of evaluating handwritten subjective answers using deep learning techniques. Traditional manual evaluation methods are prone to human error, time-consuming, and subjective. Automating the evaluation process ensures objectivity and efficiency, particularly in educational settings. Deep learning models, like convolutional neural networks (CNNs), are used in the proposed system to read handwritten text. Machine learning (SVM or NB) or transformer architectures are used to understand and rate the answers' semantic relevance. Natural Language Processing (NLP) has created an opportunity for computers to learn about written text data and make important decisions based on the learned model. We designed the system to handle varying writing styles and assess answers based on predefined marking schemes.*

**Keywords:** *NLP, Question Answering, machine learning, sentiment, deep learning*

## I. INTRODUCTION

The evaluation of subjective answers in educational settings requires human intervention, which can lead to bias, inconsistencies, and delays. With the advent of deep learning, there is potential to automate this process, reducing human error and ensuring uniformity. Handwriting recognition, paired with semantic evaluation models, offers a means to not only digitize handwritten scripts but also automatically assess their content. This research focuses on developing a system that recognizes and evaluates handwritten subjective answers using deep learning approaches. Subjective answer evaluation requires a comprehensive understanding of the context, semantics, and the grading rubric. Traditional methods of evaluation rely heavily on human assessors, which introduces bias and inconsistency. In recent years, natural language processing (NLP) techniques have advanced significantly, and with the integration of deep learning, it is now possible to build models that automatically assess and grade subjective answers. This paper investigates the use of deep learning models to replicate human-like evaluation, ensuring objectivity and minimizing the time and effort needed for largescale assessments. Optical Character Recognition, or OCR, is the electronic translation of handwritten, typewritten, or printed text into machine-translated images. OCR is the machine replication of human reading and has been the subject of intensive search for more than three decades. The system has proposed an approach to OCR detection and classification that utilizes deep learning, similar to CNN. The system integrates various stages of OCR, such as optical scanning, location segmentation, preprocessing, feature extraction, and recognition postprocessing, throughout the entire execution process. Work on optical character recognition (OCR) was limited until a few decades ago to capturing uninhabited images with flatbed desktop scanners. The processing speed and memory size of former handheld devices were not yet sufficient to run desktop-based OCR algorithms, which are computationally expensive and require a large amount of memory. Scientists have conducted extensive research on phone OCR.

## II. LITERATURE SURVEY

In [1], the development of natural language processing (NLP) and optical character recognition (OCR) methodologies for the automated evaluation of subjective responses. This article evaluates several natural language processing methodologies on prominent datasets, including the SICK dataset, STS benchmark, and Microsoft Paraphrase Identification. They may assess optical character recognition methodologies using MNIST, EMNIST, IAM datasets, and others.

According to [2], the examination of the research uncovers diverse methodologies for assessing subjective response sheets. The system's benefit is that it uses a weighted average of the most precise approaches to get the optimum outcome. TESA is a methodical and dependable technique that facilitates assessors' responsibilities and delivers faster and more effective results. This technology generates a dependable, resilient, and evident rapid reaction time.

According to [3], a voice-over-guided system to teach visually impaired individuals how to compose multilingual letters. The technology constantly observes and records the learner's strokes, while a voice-over guide provides appropriate suggestions. It will also notify if the student executes an incorrect stroke or positions the stylus outside the permissible range. This method may effectively teach any alphabet and language, allowing visually challenged students to engage in writing. They have created a language-agnostic algorithm to assist visually challenged individuals in writing multilingual alphabets. In this paper, they have implemented a voice-over guiding system in the educational process, which removes the necessity for heavy or costly equipment installations. The system integrates machine learning algorithms to assess the progress of learners. They evaluate an effective and user-focused system through usability testing.

In [4], an advanced deep learning architecture that combines convolutional neural networks (CNN) and bidirectional long short-term memory (BiLSTM) to accurately find and grade handwritten responses, just like an expert grader would. The model is specifically designed to evaluate responses consisting of 40 words, 13 of which are lengthy. They constructed the model using several methodologies, which involved modifying parameters, deep layers, neuron count, activation functions, and bidirectional LSTM layers. They systematically adjusted each parameter several times and included or eliminated layers, LSTMs, or nodes to identify the most efficient and best model.

In [5], the system utilizes a personal computer, a portable scanner, and application software to automatically correct handwritten response sheets. The Convolutional Neural Network (CNN), a machine learning classifier, processes scanned pictures for handwritten character identification. They developed and trained two CNN models using 250 photos from students at Prince Mohammad Bin Fahd University. The suggested approach would ultimately provide the student's final score by juxtaposing each categorized response with the right answer.

According to [6], the first model employs deep convolutional neural networks (CNNs) for feature extraction and a fully connected multilayer perceptron (MLP) for word categorization. The second model, termed SimpleHTR, employs convolutional neural network (CNN) and recurrent neural network (RNN) layers to extract data from images. They also offered the Bluechet and Puchserver models for data comparison. Owing to the scarcity of accessible open datasets in Russian and Kazakh languages, they undertook the task of compiling data that included handwritten names of nations and towns derived from 42 distinct Cyrillic words, inscribed over 500 times in various handwriting styles.

In [7], a self-supervised, feature-based categorization problem that is capable of autonomously fine-tuning for each inquiry without explicit supervision. The use of information retrieval and extraction (IRE) and natural language processing (NLP) techniques, together with semantic analysis for self-evaluation in handwritten text, creates a set of useful character traits. They evaluated their methodology on three datasets derived from diverse fields, with assistance from students of varying age groups.

In [8], they discuss the needs, relevant research towards handwritten recognition, and how to process it. They outline the steps and stages used in the recognition of Kannada handwritten words. The main aim of proposed work is to identify Kannada handwritten answer written in answer booklets and to solve recognition problem by using machine learning algorithms. System provides a detailed concept on pre-processing, segmentation, and the classifier used to develop systematic OCR tool.

Kumar, Munish, et al. [9], discuss the necessary conditions, relevant studies on handwriting identification, and techniques for processing.

They outline the procedures and phases involved in identifying Kannada handwritten words. The primary goal of the proposed study is to recognize Kannada handwritten responses in answer booklets and address the identification challenge using machine learning methods. The system provides a comprehensive framework for pre-processing, segmentation, and classification that is used in the development of a systematic OCR tool.

Mukhopadhyay, Anirban, et al. [10], Information given by one form-based and two texture-based data characteristics are combined from handwritten text images using classifier mixture techniques for script recognition (word-level) purposes. Based on the confidence scores supplied by the Multi-Layer Perceptron (MLP) classifier, the word samples from the specified database are listed. For this pattern recognition problem, major classifier combination techniques such as majority voting, Borda count, sum rule, product rule, max rule, Dempster-Shafer (DS) combination rule and secondary classifiers are evaluated.

Summary Table

Author(s)	Title	Methodology	Algorithms	Limitations
Souibgui, Mohamed Ali, et al. [11]	Docentr: An End-to-End Document Image Enhancement Transformer	Proposes an end-to-end Transformer model specifically designed for enhancing document images; includes techniques for image denoising, enhancement, and text clarity improvement	Transformer, U-Net Architecture, Image Enhancement Techniques	May require extensive computational resources and training data; performance may vary based on the quality and type of input images
Shailash Acharya, Ashok Kumar Pant, Prashanna Kumar Gyalwali [12]	Deep Learning Based Large Scale Handwritten Devanagari Character Recognition	Dataset increment, Dropout layers, Stochastic gradient descent (SGD) with momentum, Local response normalization, ReLU activation	Model A: Deep Convolutional Neural Network (CNN), Model B: Shallow Convolutional Neural Network (CNN), Overlapping kernel scheme, Nonoverlapping kernel scheme	High visual similarity between some characters, leading to ambiguity, Variability in handwriting styles across individuals.
Ali, Amani Ali Ahmed et al. [13]	Intelligent Handwritten Recognition Using Hybrid CNN Architectures Based-	Utilizes a hybrid CNN architecture combined with an SVM	CNN (Convolutional Neural Network), SVM (Support Vector Machine),	May face challenges with varying handwriting styles and quality;
	SVM Classifier with Dropout	classifier for handwritten recognition; includes dropout techniques for regularization and model robustness	Dropout Regularization	requires careful tuning of dropout rates and model parameters
Teslya, Niko lay et al. [14]	Deep Learning for Handwriting Text Recognition: Existing Approaches and Challenges	A comprehensive review of deep learning techniques for handwriting text recognition; includes analysis of various models, architectures, and their performance	CNN (Convolutional Neural Network), RNN (Recurrent Neural Network), LSTM (Long Short-Term Memory), Transformer	Challenges with diverse handwriting styles, data variability, and model generalization; requires extensive datasets and computational resources
Alro bah, Nase	A Hybrid Deep Model for	Development of a hybrid	CNN (Convolutional Neural	Performance may be affected
em et al. [15]	Recognizing Arabic Handwritten Characters	deep learning model combining CNNs and RNNs to recognize Arabic handwritten characters; includes preprocessing, feature extraction, and classification	Network), RNN (Recurrent Neural Network), LSTM (Long Short-Term Memory)	by variations in handwriting styles and character shapes; requires extensive training data



### III. PROPOSED SYSTEM

The Proposed System illustrates a hybrid machine learning system that combines Convolutional Neural Networks (CNNs) with traditional Machine Learning (ML) classifiers as given in Figure (1) to enhance the accuracy and robustness of the classification tasks. The overall architecture involves both training and testing processes for a dataset, leveraging the strengths of both CNNs and ML algorithms. The system is composed of two main modules: the CNN and the ML Classifier, each playing specific roles in the pipeline. Here is a detailed breakdown of the architecture (Refer Fig.1):

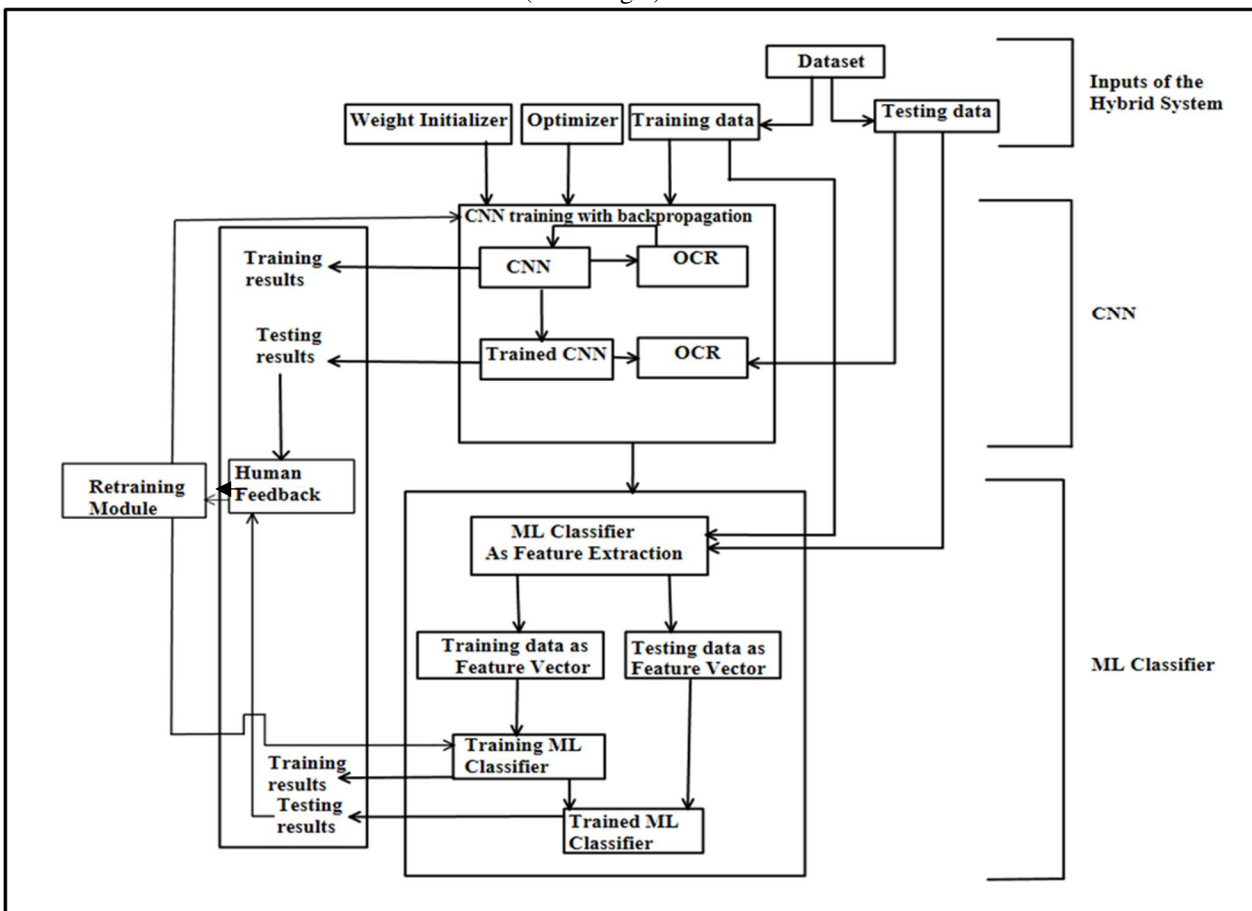


Fig.1. Proposed System Architecture

#### A. Working of the Proposed System

The system begins by accepting a dataset, which is split into two then backpropagating this error to adjust the weights, reducing parts: training data and testing data. The training data is used to the error in future predictions. This iterative process continues train both the CNN and the ML classifier, while the testing data until the CNN is well-trained to recognize patterns in the input is used to evaluate the model's performance. The first step in the data. After training, the CNN is used to perform OCR (Optical pipeline is the CNN block, where the Weight Initializer is responsible for initializing the CNN's weights. These weights are both training and testing data. The CNN acts as a feature parameters that the network adjusts as it learns. To optimize the extractor by identifying high-level patterns, such as shapes or learning process, an Optimizer (like Adam or SGD) is used to text characters in images. Features from the input. The CNN learns by calculating the error (or loss) between the predicted output and the actual labels and Character Recognition) or similar feature extraction tasks on minimize the loss function, guiding the weight adjustments an during training. Once the CNN has extracted these features, they are passed to ML Classifier. The ML classifier here plays a dual role: it serves Once initialized, the CNN undergoes training with both as a feature extraction tool and as the final classification backpropagation. This process involves feeding the training data model. The training data and testing data, now converted into the network, where the CNN applies multiple layers of feature vectors by the CNN, are fed into the ML classifier.

This convolutions, pooling, and activations to extract meaningful classifier could be any traditional model like Support Vector Machines (SVM), Random Forests, or Decision Trees. The classifier uses these feature vectors to learn the relationships between the extracted features and their corresponding labels during training, creating a model capable of making accurate predictions on new, unseen data.

The testing phase then begins, where the trained CNN and trained ML classifier are used together to predict the outcomes of the testing data. The CNN extracts features from the testing data, and the ML classifier uses these features to produce the final classification results. These results are evaluated to determine the system's performance.

The system also incorporates an essential retraining module that uses human feedback to improve its accuracy over time. After the testing phase, if the results are unsatisfactory or need refinement, a human user can provide feedback on the results. This feedback is then processed by the retraining module, which adjusts the system based on the insights gained. For example, if the CNN is not extracting the right features or if the ML classifier is misclassifying data, the retraining module can re-initiate the training process, fine-tuning the CNN and/or the ML classifier with the updated information.

#### IV. APPLICATIONS

The proposed system for automatic subjective answer evaluation has a wide range of applications, particularly in educational and assessment settings.

- 1) *Online Examination Systems:* The integration of deep learning techniques for handwritten recognition and semantic evaluation streamlines the assessment of subjective answers in online exams. Automated systems can read students' handwritten responses, interpret the semantic content, and grade based on accuracy and completeness. This eliminates the need for manual evaluation, reducing bias and increasing efficiency.
- 2) *Chatbot Exam Applications:* Chatbots can be equipped with automatic subjective answer evaluation models to help students practice subjective questions. By integrating NLP and deep learning, chatbots can provide real-time feedback on handwritten answers sub
- 3) *Subjective Evaluation Systems:* For educational institutions, these systems automate the grading of handwritten subjective responses, such as essays or short answers, using OCR and deep learning models. They analyse text for semantic meaning and compare responses against predefined grading rubrics, offering a consistent and objective evaluation process.
- 4) *Legal Document Analysis:* In legal sectors, automatic subjective answer evaluation can assist in digitizing and analysing handwritten legal documents, contracts, and case notes. Deep learning models can interpret complex language and extract meaningful insights, helping lawyers save time in reviewing documents.
- 5) *Handwritten Form Processing:* Used in industries that handle forms (e.g., healthcare, banking), OCR combined with deep learning can automate the extraction mitted via image input. This application can help users improve their answers by assessing content relevance and structure of handwritten information from forms and documents, reducing manual data entry errors and improving data processing speed.

#### V. CONCLUSION

We have studied five papers, thoroughly examining their methodologies, algorithms, and results. Each of these papers contributes valuable insights into the domain of handwriting recognition and subjective answer evaluation, but they also reveal certain limitations. From this comparative study, we identified gaps in handling diverse handwriting styles, semantic evaluation of subjective content, and robustness against variations in document quality. By analyzing these limitations, we identified opportunities to address these issues through alternative approaches and algorithms.

Our system integrates deep learning and NLP techniques to achieve a more consistent, objective, and efficient evaluation process, addressing challenges present in other papers.

Now, we are currently working on implementing this concept and very soon, we are aiming to publish a paper on our findings.

#### REFERENCES

- [1] Desai, Madhavi B., et al. "A Survey on Automatic end document image enhancement transformer." 2022 Subjective Answer Evaluation." Advances and 26th International Conference on Pattern Recognition Applications in Mathematical Sciences 20.11 (2021): 2749-2765.
- [2] Singh, Shreya, et al. "Tool for evaluating subjective scale handwritten Devanagari character recognition." answers using AI (TESA)." 2021 International Proceedings of the 9th International Conference on Conference on Communication information and Software, Knowledge, Information Management & Computing Technology (ICCICT). IEEE, 2021.



- [3] Islam, Muhammad Nazrul, et al. "A multilingual handwriting learning system for visually impaired people." IEEE Access (2024). architectures based-SVM classifier with dropout."
- [4] Rahaman, Md Afzalur, and Hasan Mahmud. "Automated evaluation of handwritten answer script using deep learning approach." Transactions on Machine Learning and Artificial Intelligence 10.4 learning for handwriting text recognition: existing (2022). approaches and challenges." 2022 31st Conference of
- [5] Shaikh, Eman, et al. "Automated grading for handwritten answer sheets using convolutional neural networks." 2019 2nd International conference on new model for recognizing Arabic handwritten characters." trends in computing sciences (ICTCS). IEEE, 2019. IEEE Access 9 (2021): 87058-87069.
- [6] Nurseitov, Daniyar, et al. "Classification of handwritten names of cities and handwritten text recognition using various deep learning models." arXiv preprint arXiv:2102.04816 (2021).
- [7] Rowtula, Vijay, Subba Reddy Oota, and C. V. Jawahar. "Towards automated evaluation of handwritten assessments." 2019 International Conference on Document Analysis and Recognition (ICDAR). IEEE, 2019.
- [8] Sampathkumar, S. "Kannada Handwritten Answer Recognition using Machine Learning Approach." (2022).
- [9] Kumar, Munish, et al. "Character and numeral recognition for non-Indic and Indic scripts: a survey." Artificial Intelligence Review 52.4 (2019): 2235-2261.
- [10] Mukhopadhyay, Anirban, et al. "A study of different classifier combination approaches for handwritten Indic Script Recognition." Journal of Imaging 4.2 (2018): 39.
- [11] Souibgui, Mohamed Ali, et al. "Docentr: An end-to-(ICPR). IEEE, 2022.
- [12] Pant, Ashok Kumar, et al. "Deep learning based large Applications (SKIMA). IEEE, 2015.
- [13] Ali, Amani Ali Ahmed, and Suresha Mallaiah. "Intelligent handwritten recognition using hybrid CNN Journal of King Saud University-Computer and Information Sciences 34.6 (2022): 3294-3300.
- [14] Teslya, Nikolay, and Samah Mohammed. "Deep Open Innovations Association (FRUCT). IEEE, 2022.
- [15] Alrobah, Naseem, and Saleh Albahli. "A hybrid deep





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)