REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1S (2024)
http://www.veterinaria.org
Article Received: 11-03-2024 Revised:12-04-2024 Accepted: 15-05-2024

Deep Learning Approach For Handwritten Gujarati Script Recognition Using Convolutional Neural Network

Riddhi Kundal[1*], Dr. Bhagwati Parekh[2]

[1*]Bhakt Kavi Narsinh Mehta University - Gujarat - India riddhi.joshi931@gmail.com
[2]Bhakt Kavi Narsinh Mehta University - Gujarat - India parekhbhagwati@gmail.com

Abstract: - Handwritten character recognition, particularly in languages like Gujarati, poses both opportunities and challenges for technological advancement and cultural preservation. This paper explores the significance of handwritten character recognition in Indian languages, focusing on Gujarati script, which features a complex character set and intricate writing styles. The use of Convolutional Neural Networks (CNNs) in recognizing handwritten characters is discussed, highlighting their ability to extract hierarchical representations from input images and surpass traditional recognition techniques. The architecture and training process of CNNs for handwritten character recognition are detailed, emphasizing their effectiveness in capturing spatial dependencies and structural information in handwritten characters. Literature surveys further demonstrate the growing interest in CNNs for various natural language processing tasks, including sentiment analysis and text classification. The potential for CNN-based techniques to develop unified models for multilingual text processing and improve script identification accuracy is highlighted, suggesting promising directions for future research in artificial intelligence and language technologies.

Keywords: Convolutional Neural Networks (CNNs), Natural language processing (NLP), Recurrent neural networks (RNNs).

Introduction:
One of the trickiest and most significant problems in pattern matching, especially for Indian languages, is handwritten character identification. This will facilitate the development of numerous useful and practical apps for interacting with various digital gadgets. These programs can be used for anything from government applications to early schooling. These kinds of programs will primarily aid those who are illiterate in English and senior folks converse with digital gadgets. Due to their vast and intricate character sets, Indian languages present difficulties in handwritten character identification.One of the finest options may be handwritten character recognition.In the field of combination of patterns and machine learning, detection of handwritten characters is receiving significant interest. Based on the method of data capture, the identification of handwritten characters can be divided into both on- and offline categories. An image of a handwritten character is scanned in offline character recognition, after which features are retrieved and the character is identified using a feature set. Using a classifier and feature set, online character recognition uses a stylus or a cursor to detect pixel values, which are then translated into text.

Script in Gujarati:
One of India's official languages and one of the most widely spoken Indo-Aryan languages is Gujarati. The employment of diacritical marks, a wide and intricate character set, conjunctive and half characters, and other features give the Gujarati language its unique intricacy. Certain Gujarati characters need to be written with numerous strokes in order to be read. The alphabet used in Gujarati has 47 letters. Three different types of characters are used in the Gujarati language: dependent vowel symbols, thirteen different vowels, and thirty-four separate consonants [8].In this digital age, handwritten script detection is very important, especially for languages like Gujarati. Even if technology has made it easier to go from analog to digital communication, a large amount of historical and cultural knowledge is still kept in handwritten papers. Recognizing handwritten writing in Gujarati script opens us new possibilities for historical study, cultural preservation, and technological advancement in addition to protecting linguistic legacy.
Gujarati script's cultural relevance is one of the main factors contributing to its significance. Gujarati, an Indo-Aryan language spoken mostly in the Indian state of Gujarat, has a centuries-old literary heritage. This cultural legacy is captured in handwritten manuscripts, correspondence, and historical texts. Gujarati speakers' feeling of identification and belonging is enhanced by the preservation and interpretation of these handwritten relics, which also allows academics to explore the subtleties of Gujarati literature, history, and culture.Furthermore, the recognition of handwritten Gujarati script is essential for historical study and archive preservation. Gujarati script is used to write a lot of historical papers and records, from literary works to administrative paperwork. Researchers, historians, and academics can more easily access these handwritten papers and evaluate, interpret, and share important historical information thanks to the capacity to accurately transcribe and digitize them. Through the use of recognition technology, we are able to preserve handwritten documents and guarantee that the historical and cultural legacy they contain will be available to future generations.Moreover, recognition of handwritten Gujarati script is important for modern language processing and communication. Personal letters, notes, and signatures are examples of situations where handwritten communication is still used, despite the widespread use of digital platforms and keyboards. Effective identification of handwritten Gujarati script facilitates the

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1S (2024)
http://www.veterinaria.org
Article Received: 11-03-2024 Revised:12-04-2024 Accepted: 15-05-2024

easy incorporation of handwritten input into digital systems, improving accessibility and user experiences for Gujarati speakers across a range of industries, including business, education, and administration.
The creation of handwritten Gujarati script recognition systems offers particular potential and problems from a technology perspective. Gujarati characters are complicated, and handwriting styles differ, hence sophisticated pattern recognition methods are needed. Machine learning advances, especially in deep learning techniques such as convolutional neural networks (CNNs), have demonstrated potential in handwritten script recognition. Researchers provide advances in handwritten Gujarati script recognition, which not only benefits language technology but also has wider applications in artificial intelligence, text mining, and document analysis.

Deep learning Approach for handwritten character recognition:
Deep learning has been a potent tool in recent years for handling difficult pattern recognition problems, such as handwritten character identification. An overview of deep learning methods, in particular Convolutional Neural Networks (CNNs), and how they are used to recognize Gujarati characters written by hand is given in this essay.In the machine learning subfield of deep learning, multiple-layer neural networks are trained to learn hierarchical data representations. Deep learning algorithms have the ability to automatically learn useful features from raw data, which makes them particularly ideal for problems with high-dimensional input spaces, such as image identification. This is in contrast to typical machine learning approaches, which need handcrafted feature extraction.
A subclass of deep neural networks called convolutional neural networks (CNNs) is especially made for handling structured, grid-like data, like photographs. Convolutional, pooling, and fully linked layers are among the layers that make up a CNN. While the pooling layers downsample the feature maps to lessen the computational complexity of the network, the convolutional layers apply convolutional filters to the input image to capture local spatial patterns. In order to provide predictions, the fully linked layers finally combine the acquired characteristics.Because Gujarati script is complex and writing styles vary, it might be difficult to recognize handwritten characters. Support Vector Machines (SVMs) and Hidden Markov Models (HMMs) are two examples of machine learning algorithms that are frequently used in traditional methods for handwritten character recognition.
These techniques might, however, find it difficult to generalize effectively across various handwriting styles and character form variances.In a variety of image recognition applications, including handwritten character identification, CNNs have demonstrated impressive performance. The spatial dependencies and structural information included in handwritten characters can be efficiently captured by CNNs by utilizing the hierarchical features acquired through convolutional filters. CNNs may be trained to distinguish between various characters and handwriting style variants in the Gujarati character recognition environment, which will yield more accurate recognition results.There are various phases involved in training CNNs for handwritten Gujarati character recognition, including gathering data, preprocessing it, designing the model, training it, and evaluating it.
The CNN model is trained using annotated datasets of handwritten Gujarati characters, with preprocessing methods like augmentation and normalization employed to improve the caliber and variety of the training data. With suitable selections of convolutional filter sizes, pooling techniques, and activation functions, the CNN architecture is made to fit the distinctive qualities of Gujarati characters. Stochastic gradient descent (SGD) and Adam are two optimization methods that are used during training to minimize the loss function and update the network parameters. Ultimately, the accuracy and generalization capacity of the trained CNN model are evaluated by analyzing its performance on an independent test dataset.

CNN:
Another well-liked and frequently applied deep learning method is CNN [9]. It has been widely used in many applications, including computer vision, speech processing , and natural language processing , to mention a few. Its architecture is modeled after the neurons found in the brains of animals and humans, just like in conventional neural networks. In particular, it mimics the intricate network of cells that makes up the visual cortex in a cat's brain [10]. Three key benefits of CNN are parameter sharing, sparse interactions, and equivalent representations, as stated in [11]. Rather than using standard fully connected networks, the network makes use of local connections and shared weights to fully leverage the two-dimensional structure of an input data (such as an image signal).
Very few parameters are produced by this procedure, which speeds up and simplifies the training of the network. This function is comparable to that of the cells in the visual cortex. Small portions of a scene, as opposed to the entire scene, are what these cells respond to. Put another way, the cells extract spatially local correlation present in the data and act as local filters over the input.A natural language inquiry should be interpreted by an automatic question-and-answer system, which should then utilize logic to provide a suitable response. The renown FREEBASE dataset and other contemporary knowledge bases enable this area to grow and transcend the days of manually crafting features and rule sets for certain domains. A multicolumn CNN approach was developed by Dong et al. [12] that can examine a question from multiple perspectives, such as which context to select, the answer's underlying semantic meaning, and the answer's formation. They employ a multitasking strategy that simultaneously learns the affiliations and correlations of the low-level word semantics and scores the question-answer pairs. It is suggested to use a deeper learning architecture that is more universal and not language-specific.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1S (2024)
http://www.veterinaria.org
Article Received: 11-03-2024 Revised:12-04-2024 Accepted: 15-05-2024

CNN Architecture:
CNN architectures usually have the following essential elements:

- Convolutional Layers: Convolutional layers use learnable filters, or kernels, to extract local spatial patterns from input pictures, including forms, edges, and textures.
- Pooling Layers: Using a downsampled approach, pooling layers reduce the spatial dimensionality of feature maps derived from convolutional layers while maintaining significant features.
- Activation Functions: By introducing nonlinearity into the network, activation functions help the network discover intricate patterns and connections within the data.
- Fully Connected Layers: Using a sequence of weighted connections, fully connected layers combine the characteristics that were retrieved by convolutional and pooling layers and map them to output classes.

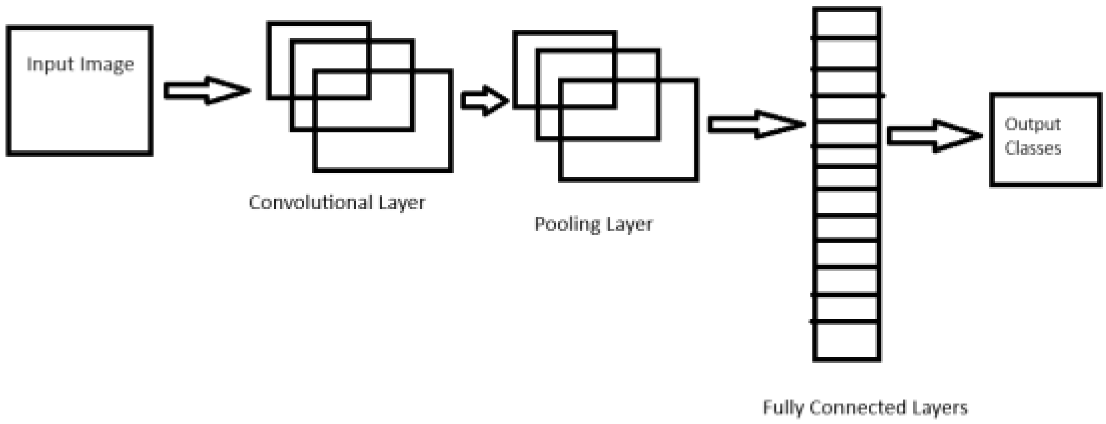

Fig. 1 Architecture of CNN

Input Layer: Typically, an image containing the text to be recognized serves as the CNN's input. Pixel values are arranged in a grid to depict this image.

Convolutional Layer: The fundamental component of a CNN is the convolutional layer. It is composed of a number of kernels, or learnable filters.Although each filter is compact in size, it covers the entire input volume depth.Convolution is the process of applying the filter to the input image by multiplying the filter values by the input values element-by-element and then summarizing the results. Features from the input image, such as edges, corners, and textures, are successfully extracted at various spatial positions using this technique.

Activation Function: To add non-linearity to the network, an activation function such as ReLU (Rectified Linear Unit) is applied element-by-element following the convolution operation.ReLU preserves positive values while setting all negative values to zero.

Pooling Layer: The feature maps that are acquired from the convolutional layers are downsampled using the pooling layer.Max pooling and average pooling are common pooling methods that minimize the spatial dimensions of the input while preserving its most significant properties.By helping to reduce the size and manageability of the representations, pooling lowers the network's computational complexity.

Fully Connected Layers: The high-level features are compressed into a one-dimensional vector following a number of convolutional and pooling layers.The next step involves passing this vector through one or more fully connected (dense) layers, which use these features to conduct classification.Making predictions and recognizing intricate patterns in the features that were extracted are the responsibilities of the fully connected layers.

Output Layer: The output layer converts the result scores into probabilities by use of a softmax activation function.When it comes to text recognition, each class (such as an alphanumeric character) usually has a node in the output layer, and the network predicts the probability distribution over these classes.

Training: Stochastic gradient descent (SGD) and Adam optimization methods, as well as backpropagation, are used in CNN training.In order to reduce the discrepancy between the expected output and the ground truth labels, the network learns to modify the weights of its filters and fully connected layers during training.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1S (2024)
http://www.veterinaria.org
Article Received: 11-03-2024 Revised:12-04-2024 Accepted: 15-05-2024

CNNs may effectively identify text in fresh images by training on a dataset that includes text images and the labels that correlate to those images. The network gains the ability to recognize patterns and traits that distinguish various characters and words, allowing it to predict outcomes accurately even when no data is available.

Literature Survey:
A comparison of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for natural language processing (NLP) tasks is presented by Wenpeng Yin, Katharina Kann, Mo Yu, and Hinrich Schütze. It is likely that the study looks into how well these systems perform in a range of NLP tasks, including sentiment analysis, text classification, language modeling, and machine translation. It probably goes over the advantages and disadvantages of each architecture for processing text-style sequential data, emphasizing how well CNNs capture local patterns.The study may also explore CNNs' variations in architecture, training methods, and computing efficiency.[2]

The study by Amit Kumar Sharma, Sandeep Chaurasia, and Devesh Kumar Srivastava focuses on sentiment analysis of brief words using a Convolutional Neural Network (CNN) deep learning model. In particular, it uses Word2Vec embeddings that are optimized in the training phase. The study probably looks into how well CNNs can convey the emotion conveyed in brief textual bursts. Through the use of Word2Vec embeddings tailored to the sentiment analysis job, the model may be able to pick up rich representations of words and their related contexts in the provided short phrases.. [3]

In his discussion of CNN architectures tailored for textual data processing, N I Widiastuti highlights the technology's capacity to automatically learn hierarchical representations. The evaluation might emphasize how well CNNs perform in a range of natural language processing (NLP) applications, including named entity recognition, text classification, and sentiment analysis..[4]

Natural language processing (NLP) is one area in which Wei Wang and Jianxun Gang investigate the use of convolutional neural networks (CNNs). In NLP, CNNs have become more and more popular due to their capacity to automatically extract hierarchical feature representations from unprocessed input data. CNNs achieve state-of-the-art results in sentiment analysis, topic classification, and spam detection by identifying both local and global patterns in textual data. This makes CNNs excellent at text classification tasks. Additionally, by identifying sentiment-bearing characteristics in text inputs, CNNs have proven their efficacy in sentiment analysis. [5]
Regardless of the script being utilized, Durjoy Sen Maitra, Ujjwal Bhattacharya, and Swapan K. Parui showed how well CNNs capture complex patterns and features from handwritten characters. Through the utilization of CNNs' hierarchical learning capabilities, scientists have produced state-of-the-art outcomes in character recognition challenges, opening the door for the creation of unified models that can identify characters across a variety of scripts. Providing an overview of the approaches and results in CNN-based methods for handwritten character identification across several scripts is the goal of this review of the literature. It also highlights the potential for enhanced accuracy and generalization in multilingual text processing systems.[6]

By combining data from text recognition systems, Yupeng Cao, Jing Li, Qiufeng Wang, Kaizhu Huang, and Rui Zhang investigated unique techniques to improve script identification. Through the utilization of text recognition model output, researchers hope to increase the precision and resilience of script identification methods. More contextual signals and linguistic elements are made available by integrating text recognition information, which makes it possible to identify scripts in multilingual publications more precisely. This overview of the literature highlights the potential for improvements in multilingual text processing and document analysis systems by combining the results of several studies examining the complementary link between text recognition and script identification.[7]

Conclusion:
In conclusion, there are potential and challenges associated with handwritten character recognition, especially in languages like Gujarati. Because of its capacity to capture intricate patterns and features, Convolutional Neural Networks (CNNs) have become highly successful tools in this field, surpassing the capabilities of conventional techniques. Convolutional, pooling, and fully connected layers are three features that CNN architectures use to their advantage when deriving hierarchical representations from input images so that character recognition is correct. With further research, CNN-based techniques could lead to unified models that can identify characters from different scripts, improving multilingual text processing systems. Combining CNNs with text recognition models has the potential to increase script identification accuracy and develop artificial intelligence and language technologies.

References:
1. Hassan, Abdalraouf, and Ausif Mahmood. "Deep learning for sentence classification." *2017 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*. IEEE, 2017.
2. Yin, Wenpeng, et al. "Comparative study of CNN and RNN for natural language processing." *arXiv preprint*

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1S (2024)
http://www.veterinaria.org
Article Received: 11-03-2024 Revised:12-04-2024 Accepted: 15-05-2024

arXiv:1702.01923 (2017).
3. Sharma, Amit Kumar, Sandeep Chaurasia, and Devesh Kumar Srivastava. "Sentimental short sentences classification by using CNN deep learning model with fine tuned Word2Vec." *Procedia Computer Science* 167 (2020): 1139-1147.
4. Widiastuti, Nelly Indriani. "Convolution neural network for text mining and natural language processing." *IOP Conference Series: Materials Science and Engineering*. Vol. 662. No. 5. IOP Publishing, 2019.
5. Wang, Wei, and Jianxun Gang. "Application of convolutional neural network in natural language processing." *2018 international conference on information Systems and computer aided education (ICISCAE)*. IEEE, 2018.
6. Maitra, Durjoy Sen, Ujjwal Bhattacharya, and Swapan K. Parui. "CNN based common approach to handwritten character recognition of multiple scripts." *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*. IEEE, 2015.
7. Cao, Yupeng, et al. "Improving Script Identification by Integrating Text Recognition Information." *Aust. J. Intell. Inf. Process. Syst.* 16.3 (2019): 67-75.
8. Magare, S. S., Gedam, Y. K., Randhave, D. S., & Deshmukh, R. R. (2014). Character Recognition of Gujarati and Devanagari Script: A Review. International Journal of Engineering Research & Technology (IJERT), 3(01).
9. Pareek, J., Singhania, D., Kumari, R.R. and Purohit, S., 2020. Gujarati handwritten character recognition from text images. Procedia Computer Science, 171, pp.514-523.
10. Patel, C., & Desai, A. (2013). Gujarati handwritten character recognition using hybrid method based on binary tree-classifier and k-nearest neighbour. International Journal of Engineering Research & Technology (IJERT), 2(6), 2337-2345.
11. Naik, V. A., & Desai, A. A. (2017, July). Online handwritten Gujarati character recognition using SVM, MLP, and K-NN. In 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.
12. Macwan, S. J., & Vyas, A. N. (2015, August). Classification of offline Gujarati handwritten characters. In 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 1535-1541). IEEE.