International Journal of Mathematical, Engineering and Management Sciences
Vol. 7, No. 5, 643-655, 2022
https://doi.org/10.33889/IJMEMS.2022.7.5.042

CNN-Based Optical Character Recognition for Isolated Printed Gujarati Characters and Handwritten Numerals

Sanket B. Suthar
Department of Information Technology,
Chandubhai S. Patel Institute of Technology (CSPIT), Faculty of Technology & Engineering (FTE),
Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India.
Corresponding author: sanketsuthar.it@charusat.ac.in

Amit R. Thakkar
Department of Computer Science and Engineering,
Chandubhai S. Patel Institute of Technology (CSPIT), Faculty of Technology & Engineering (FTE),
Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India.
E-mail: amitthakkar.it@charusat.ac.in

(Received on March 30, 2022; Accepted on June 28, 2022)

Abstract
Optical character recognition (OCR) technologies have made significant progress in the field of language recognition. Gujarati is a more difficult language to recognize compared to other languages because of curves, close loops, the inclusion of modifiers, and the presence of joint characters. So great effort has been laid into the literature for Gujarati OCR. Recently deep learning-based CNN models are applied to develop OCR for different languages but Convolutional Neural Networks (CNN) models are not yet giving a satisfactory performance to recognize Gujarati characters. So, this paper proposes a revolutionary Gujarati printed characters and numerals recognition CNN models. CNN-PGC (CNN for - Printed Gujarati Character) and CNN-HGC (CNN for - Handwritten Gujarati Character) are two optimally configured Convolutional Neural Networks (CNNs) presented in this research for printed Gujarati base characters and handwritten numbers, respectively. Concerning particular performance indicators, the suggested work's performance is evaluated and proven against that of other traditional models and with the latest baseline methods. Experimental analysis has been carried out on well-segmented newly generated Gujarati base characters and numerals dataset which includes 36 consonants, 13 vowels, and 10 handwritten numerals. Variation in the database is also taken into consideration during experiments like size, skew, noise blue, etc. Even in the presence of printing irregularities, writing irregularities, and degradations the proposed method achieves a 98.08% recognition rate for print characters and a 95.24 % recognition rate for handwritten numerals which is better than other existing models.

Keywords- Optical character recognition (OCR), Convolution neural network (CNN), Recognition, Gujarati character and symbols, Handwritten numerals, Printed character classification.

1. Introduction

Over the last five decades, OCR achieved many milestones and grown from a dream to reality. OCR development for Indic languages is a very challenging job because of the variety of shapes of characters, Complex character sets, and presence of modifiers. OCR has proven to be one of the most effective applications of artificial intelligence and pattern recognition technologies. In today's era scanned or camera-captured images of documents in a huge set of Indic scripts are available. OCR makes it 100% searchable. Variety of devices used to capture the Images like mobile phones, tablets, DSLR cameras, etc. Data entry from specific original paper data sources, such as passport documents, invoices, bank statements, receipts, business cards, mail, or any number of printed records, is typically done using OCR. Gujarati Text recognition will help to recognize sentences written in ancient books like Granth, Vedas, and historical documents and will convert them into digital form. Further, it also can be converted into speech. That digitized documents can be translated into English, Hindi, and many more languages too. The suggested method enables individuals who are blind or visually impaired to scan printed material and have it read aloud or saved to a computer file. This research will help to develop a portable scanner specifically for the Gujarati language, which helps to scan any type of document and convert them to digital form. Another application where Gujratai OCR plays an important role is scene text recognition from real-time images. OCR is a research field in pattern recognition, artificial intelligence, and computer vision. From multiple sources, it is possible to collect data for Indic languages and such sources of documents are printed books, newspapers, Granths, Vedas, religious books, and other historical documents. In various domains, OCR plays a very important role. Some examples and applications are education, banking, defense, healthcare, finance, government agencies, etc. The interest of this research work is specifically for Gujarati printed and handwritten characters and numerals. In today's decade, lots of research work has been done and many applications are developed for commercial OCRs in printed as well as handwritten "Roman", "Japanese", "Korean", "Chinese", "English", "Arabic" and other European and Asian scripts (Jeong et al., 2003; Mehta et al., 2016; Das and Mohanty, 2020). But very few works are found in this field for structural features-based methods for the feature extraction and classification of "Printed Gujarati Characters, printed numerals and handwritten numerals". A structural feature extraction method is introduced and achieved noticeable accuracy (Goswami and Mitra, 2013).

This research attempts to propose a suitable method for Gujarati Characters and numerals and focuses on how deep learning will increase the efficiency specifically in classification and recognition and feature extraction. Various architectures are already developed by different authors based on the different languages. The autoencoder, convolutional neural network (LeCun, 2015), deep belief network, and constrained Boltzmann machine are four deep learning architectures presented (Liu et al., 2017). Optical Character Recognition (OCR) is a strong tool for retrieval tasks, but it requires a well-designed image mining system. The image content can be expressed in several ways, to beat this issue a new technique called Content-Based Image Retrieval for image mining is being widely used (Yakin et al., 2021). Recently Tesseract an open-source optical character recognition (OCR) platform is used to extract text from images and documents without a text layer and outputs the document into a new searchable text file, PDF, or most other popular formats. It can read characters from ancient books with tiny font sizes and nearly incomprehensible content. Tamil, Malayalam, Oriya, Gujarati, Kannada, and other Indic scripts are supported by Tesseract OCR. Another good application developed by Google is Google Lens. It can get details or take actions on the photos, objects around people, and image searches with Google Lens. Google lens can translate text in real-time from over 100 languages For Indic scripts, specifically in Gujarati Tesseract and Google Lens both fail to provide good accuracy because of two major reasons. The first reason is that the Gujarati script contains modifiers, special characters, joint characters, and half letters. Another reason is available Gujarati handwritten character dataset does not contain a sufficient number of samples of each Gujarati character with a variety of characteristics.

1.1 Working Principles OCR System

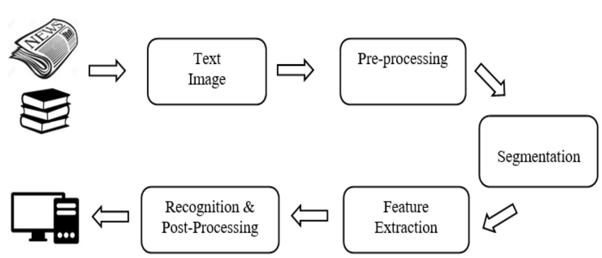

Figure 1. Components of OCR.

The components of OCR are shown in Figure 1. The very First step is the Acquisition of the documents from libraries, museums, laser printed books, inject printed books, typewritten books, newspapers, Granths, Vedas, etc. and scanned that all documents in some good quality scanner with resolutions 300x300 dpi which can give good quality images so it is easy for users to segment words and characters in future. Good quality images with 300x300 also help us to apply various pre-processing steps which are necessary before feature extraction and classification method. The second component is pre-processing and it has a very important role in OCR. Scanned images of documents are usually displayed in either RGB color or grayscale. The documents that are scanned are very challenging to process because they contain printing noise. Skew correction is also equally important before applying feature extraction methods to recognize characters efficiently. This step helps to segment the character images in the next part. The third phase is segmentation, which includes issues such as line segmentation, word segmentation, and character segmentation, and varies by script. It depends on curves, shapes, and modifiers used in the script. Segmentation for the Indic languages compared to western languages is more difficult due to complicated shapes, curves, and circles (Mehta et al., 2016; Liao et al., 2018). Image-based features, statistical features, transform-based features, structural features, and other features are all included in the feature extraction step of the OCR component. In the Bangla script, a lot of work has gone into structural feature extraction. (Fink et al., 2010). The last component is post-processing which helps OCR to replace similar-looking and most frequently appeared characters with wrong recognized characters. Some other post-processing tasks are grouping and context information.

1.2 The Different Areas of Character Recognition

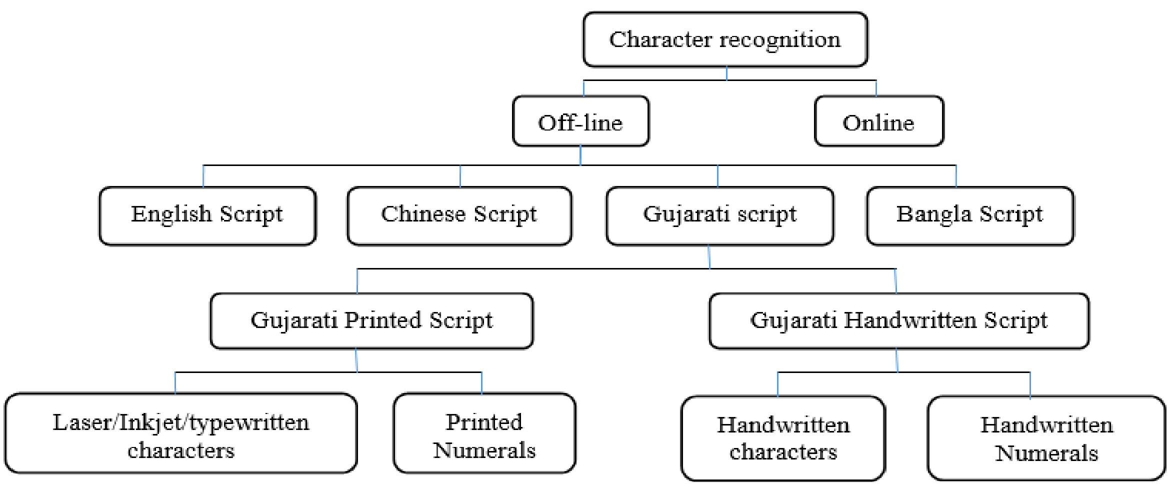

Figure 2. Character recognition areas.

This research work is focused on off-line Gujarati character recognition in which thousands of documents are collected from a variety of resources like the Printed books, Newspapers, Granths, Vedas, Historical documents, Government documents, etc. This research focused on only Gujarati printed script and Handwritten numerals. In the future, this research will be expanded to include the recognition of offline-handwritten numerals in Gujarati and Sanskrit, two significant Indian scripts. Few research works have been found on offline handwritten Chinese characters and for North Indian script symbols recognition in the years 2017 and 2016 respectively (Macwan et al., 2016; Zhang et al., 2017; Wu et al., 2017; Roy et al., 2017; Ukil et al., 2021).

2. Background and Literature Survey

Gujarati is a widely spoken West Indian language that hails from Gujarat and belongs to the Indo-Aryan language family (Mehta et al., 2016). India is a multilingual country and more than 29 scripts are there for communication more than 46.1 million people in India Speak Gujarati. It is generally written from left to right and top to down without shirorekha on top of words like other languages Devnagri and Hindi. A total of 34 Consonants 13 Vowels 15 Gujarati modifiers are present in Gujarati Script. The presence of these Modifiers and Vowels in Gujarati makes the script very difficult to recognize compared to other western languages like English. Building Optical Character Recognition (OCR) technology for Gujarati text is more difficult than developing OCR technology for western languages due to the complicated character set, the presence of modifiers, joint characters, and other difficulties depicted in Figures 3, 4 and 5.

Figure 3. Curves & Circles (Closed Loops)

Figure 4. Character with Modifiers

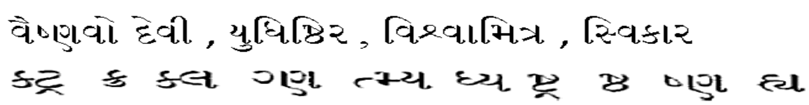

Figure 5. Joint characters.

From the literature survey, it is come to know the use of various deep learning techniques in OCR that reduces the number of steps performed in traditional methods and can classify characters with high accuracy. The following section provides a brief overview of some important work reported by authors using many traditional methods like self-organizing map, K Nearest Neighbor, Artificial Neural Network, Structural Feature-based Method, etc. and also Arabic, English, Cursive Urdu, Roman, Handwritten Chinese, Bangla handwritten, other Indic languages, are recognized using various deep learning approaches. Artificial intelligence (AI)-based machine learning and deep learning approaches have been introduced in recent years to recognize the characters of the Gujarati language efficiently. Some of the previous approaches based on machine learning are provided by authors recently (Jain and Arolkar, 2018). The SVM method is introduced to recognize both written characters as well as vowel recognition. But this method shows some inaccurate outcomes in recognizing the letters (Khan et al., 2022). Another approach is based on a multi-label and multi-class classification-based OCR approach for training the Gujarati characters (Narang et al., 2021). However, this method needs more time to train the word and a high loss in the outcome.

The first reference is to classify Gujarati printed characters, which dates back to 2011. Authors achieved the results by integrating the Self-Organizing Map (SOM) with the k-Nearest Neighbor Classifier (K-NN), an elegant classification technique was developed and used for a subset of printed Gujarati characters (Goswami et al., 2011), with an average accuracy of 82.36 % for a newly generated test dataset. The nearest neighbor classifier was applied to a dataset of 4500 samples from 45 different classes by the authors with k-fold cross-validation to recognize Gujarati characters (37 characters and 8 numerals), overall Recognition rate reported was 93% (Gohel et al., 2015). Goswami et al. (2011) developed a new structural feature extraction method for the categorization of printed Gujarati text in 2011, which he tested on a 4000 printed character symbol data set. The accuracy reported was low specifically for the character which contains a greater number of components (Goswami and Mitra, 2013). Another work reported on the Arabic language dated back to 2006 when Al-Jarrah et al. (2006) first use neural networks. The existence of dots and hamzas adds to the system's intricacy, as do multiple shapes for the same character in Arabic. The author presents a unique approach that takes advantage of the features of the Arabic language; it uses the information in the mainline of text to generate tokens that describe the characters. They achieved 87% accuracy across different fonts and font sizes. (Al-Jarrah et al., 2006). Later, Sahu and Kubde (2013) applied feedforward backpropagation neural networks for English handwritten characters. The authors conclude that accurate recognition is directly dependent on the nature and quality of the information to be read. TextBoxes++ is an end-to-end trainable fast scene text detector for SIW-13 (English and other Latin-based scripts specifically for Scene text detection and recognition) that detects arbitrary-oriented scene text with good accuracy and efficiency in a single network forward pass, according to Minghui Liao and colleagues. (Liao et al., 2018). The author used multidimensional recurrent neural networks for Cursive Urdu-Nastaliq script recognition. Accuracy claimed with the highest character recognition was 96.40% (Naz et al., 2016). ANN and Nearest Neighbor a new approach was applied for Roman script and English language from scanned Images (Mehta et al., 2016). The authors examine neural network language models (NNLMs) and hybrid NNLMs for Chinese handwritten text in depth (Wu et al., 2017). On the CASIA-HWDB test set, the character-level accurate rate (AR) and correct rate (CR) are 95.88 percent and 95.95 percent, respectively (Wu et al., 2017).

Francis and Sreenath (2020) published a paper in 2020 focusing on recognizing possible text objects from a scene using the Least-Square Support Vector Machine Technique. The results were assessed using ICDAR 2015 scene pictures, MSRA500, and SVT datasets, as well as other methodologies, yielding promising results. Roy et al. (2017) used SVM and DCNN to recognize Isolated Bangla Compound characters. Recognition accuracy reported using this approach was 90.33%. Thaker (2017) stated in his thesis that he used the transfer learning approach on Inception-V3 to categorize Gujarati characters using a tiny character image dataset to train neural networks and compare the results with existing solutions. Also, for feature extraction on word-level handwritten Indic script, (Ukil et al., 2021) used CNN and a multilayer 2D discrete Haar wavelet transform. Using a multi-layer perceptron (MLP), the authors were able to attain a maximum script detection rate of 94.73%. It is obvious from the conversation that a greater part of the work done for optical character recognition in previous and recent years used a machine learning approach for achieving a good accuracy rate. Post-processing also performed a major role after successful character recognition that reducing the errors in the results (Ahamed et al., 2021). Dwivedi et al. (2020) developed a Sanskrit-specific OCR system by applying LSTM to their model to overcome the problem of long-length word recognition due to characters' complexities. A word error rate of 15.97% and a character error rate of 3.71 percent have been attained.

3. Data Collection

Gujarati characters were collected from a variety of sources, including digitized articles, periodicals, newspapers, books, etc. Gujarati Handwritten Numerals samples were collected from 300 people of various ages, professional backgrounds, and genders. For laser and computer printed, for handwritten numbers, 100 samples of each character are collected, and 300 samples of each number are gathered (Vyas and Goswami, 2015). The basic characters of the Gujarati language are shown in Figure 6.

ક	ખ	ગ	ઘ	ચ	છ	જ	ઝ	ટ	ઠ	ડ	ઢ
ka	kha	ga	gha	Cha	chha	ja	za	ta	tha	da	dha
ણ	ત	થ	દ	ધ	ન	પ	ફ	બ	ભ	મ	ચ
aNa	ta	tha	da	dha	na	pa	fa	ba	bha	ma	Ya
ર	લ	વ	સ	શ	ષ	હ	ળ	ક્ષ	જ્ઞ		
ra	la	va	sa	sha	shha	ha	ala	ksha	gna		

Figure 6. Gujarati consonants.

3.1 Data Processing

Before the data can be converted into a numerical format, it needs to be pre-processed to remove noise and resize it into a uniform matrix. To denoise the data, an algorithm similar to the one used for line detection was used. To further smooth out the data, the grayscale image was binarized. Binarization is like a floor and ceiling functiox1n being applied to every pixel in the image. Pixels with values lower than the threshold are turned white and help remove unnecessary values from the image. For efficient working of the Machine Learning algorithm, all the images in the dataset are resized into 28x28.

Table 1. Dataset collection.

Factors	Number of Image Symbols
Machine Printed Text (Books)	5693
Machine Printed Text (Newspaper)	5000
Laser Printed Text	4825
Gujarati Handwritten Numerals	3000

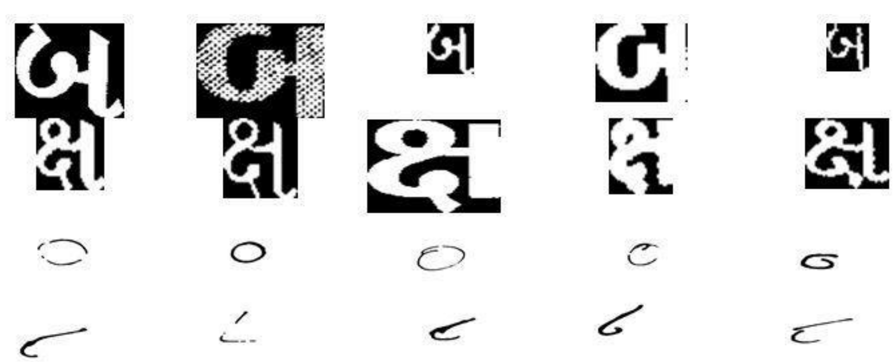

Figure 7. Sample dataset images for printed Gujarati characters and handwritten numerals.

A set of 42 Gujarati character classes are there in printed Gujarati characters and a set of 10 Gujarati numeral classes are there in handwritten numerals. Sample isolated characters from the dataset taken for the experiments are shown in Figure 8. Characters were collected from machine-printed textbooks, newspapers, laser-printed textbooks, etc. Number of symbols as per the factors mentioned in Table 1.

4. Proposed Methodology with Experiments and Results

Four major steps are handled by the deep learning technique in this approach which is shown in Figure 8. Many things need to be taken care of, especially for character recognition from the data collection process to final recognition. Data collection, data pre-processing, data segmentation, and data recognition are all part of the architecture described below. The gathering of old Gujarati manuscripts from various age groups people around the state is the first step in this project. The proposed model is processed in three stages namely, pre-processing, feature extraction and recognition. For both the character and numeral recognition, the proposed model is processed separately and at the final stage, the features from both are fused and forwarded to the recognition model for recognition.

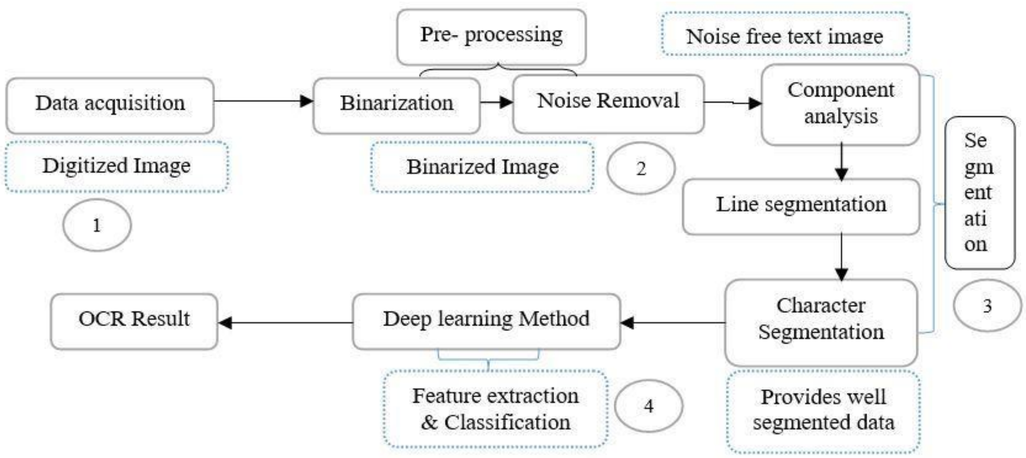

Figure 8. Proposed architecture step.

As shown in the Figure 8, the first step is the Collection of data from various sources. In this step, the main focus is to collect base characters from Gujarati script and Handwritten Numerals. To improve the quality, before the extraction and recognition stage the stage of pre-processing is required. The second step is Binarization, noise cleaning, Thinning, Skewing, etc. A noise-free text image at this point is obtained. In the pre-processing stage, filtering, character extraction, and morphological operations are performed to improve recognition accuracy. The next step is Segmentation, which includes component analysis, Line segmentation, and character segmentation. After applying step 3, well-segmented isolated character and numeral images are obtained. To extract lines and words from the document image, the histogram Projection Method is applied. The lines can be separated by selecting the row corresponding to the lower peak in the histogram. For the word segmentation lower peaks should be selected so it spans through a certain width (threshold). Further, the character level segmentation can be achieved by leveraging the small gap between the characters, i.e., by projecting the image vertically, it is possible to segment the characters. The feature extraction stage is the final stage, and it is used to quickly extract hybrid features such as zoning features, diagonal features, centroid features, and peak extent features. The work will be done in Python, and the performance measurements will be assessed and compared to existing approaches to determine the efficacy of the model.

4.1 Execution Details

The Keras and Tensor Flow frameworks are used to run the CNN – PGC, and CNN – HGC on a GPU. All of the tests were run on an NVIDIA GPU with a single GP 100 accelerator card and 32 GB of RAM. Performance analysis will be carried out with the existing state-of-the-art techniques.

4.2 CNN – PGC for Printed Gujarati Characters

CNN – PGC is introduced to recognize printed Gujarati characters. Well-segmented base characters are taken into consideration. The total number of character images used to train models and to test is 15518 containing different variations like Noisy character images, Skewed character images, Images with variations in the size of characters, broken character images, Blurred images, etc. Architecture for CNN – PGC is shown in Table 2 and Table 3 shows experiments and results with different image size variations is displayed.

Table 2. CNN architecture for printed characters (CNN – PGC).

Layers	Image Size	Kernel Size	No. of Filters	Stride	Activation
	28x28	-	-	-	-
Convolution	24x24	5x5	32	1	Relu
Convolution	12x12	3x3	64	1	Relu
Convolution	10x10	2x2	128	2	Relu
max pooling	5x5	3x3	64	1	Relu
Drop out (0.1)	5x5	-	64	-	-
Fully Connected	-	128 Units	-	-	-
Fully Connected	-	256 Units	-	-	-
Output		-	-	-	softmax

Table 3. Results obtained using CNN – PGC architecture.

	1	2	3	4	5	6
Image Size	28x28	32x32	32x32	32x32	50x50	50x50
Batch	100	500	200	500	100	500
Epoch	500	100	200	200	200	100
Loss	0.11	0.12	0.18	0.14	0.13	0.0924
Accuracy	0.9756	0.9757	0.9758	0.9808	0.9754	0.9792

The recognition accuracy also depends on the size of the image file fed into the network. So, the experiment has been conducted for various image sizes mentioned in Table 3. It is observed from the experiment that the highest accuracy reported is approximately 98% for the image size of 32x32. The low image size will not able to extract the minor details like thin edges of character and the high image size will not able to capture details effectively, especially for characters with curves and corners. So, in all subsequent experiments, the image size is fixed as 32 x 32.

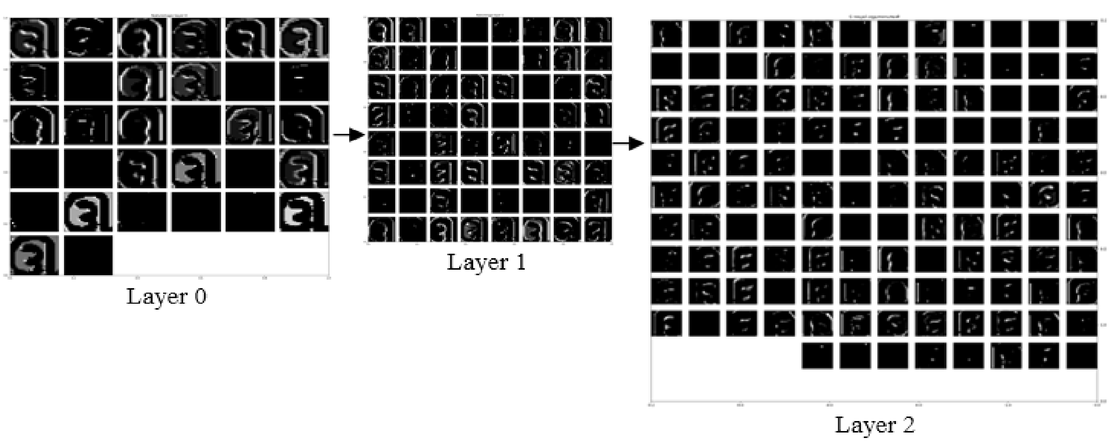

Figure 9. Visualization for intermediate layers in CNN.

Visualization of the intermediate layer is shown in Figure 9. The absolute first layer is ostensibly holding the full state of the character "UU", even though there are a few filters that are not actuated and are left blank. At that stage, most of the information extracted by the model and extracted maximum of the data are present in the underlying picture. As it goes further in the layers, the activations become progressively less interpretable. They start to encode higher-level features like fringes, corners, curves, and edges. Higher representation contains progressively fewer data about the character. As per our observation, the last layers are not fetching enough information and there is nothing more to learn at that stage.

Table 4. Comparative analysis of CNN – PGC with traditional models and other existing work for printed Gujarati characters.

Authors	Types of datasets	Classifier	Dataset	Accuracy
Classification of printed Gujarati characters (Goswami et al., 2011)	Character	SOM & K-NN	160 (40 different characters)	82.36%
Online character recognition (Gohel et al., 2015)	Character	K-NN	4500	93%
Similar looking Gujarati printed character recognition (Chaudhary et al., 2012)	Character	ANN	1500	96%
Structural feature-based classification of printed Gujarati characters (Goswami and Mitra, 2013)	Character	Structural Feature based Method	4000	85%
CNN – PGC	Character	CNN	15000	98.08%

Significant research has been conducted in recent years to produce an effective character recognition system. The majority of the work done by character recognition researchers in the literature was based on feature extraction and classification using traditional methods. Table 4 tabulates the recognition result summary and it can be used to compare and contrast the performance of different recognition methods. On a newly developed Gujarati dataset, recognition accuracy was recorded as follows: The first model was run with 500 epochs for a 28x28 size image and subsequently changed hyperparameters to tune the model like batch size, epochs, etc. The highest testing accuracy achieved using CNN – PGC model is 98.08 % with 500 batch sizes and 200 epochs. In training networks, the epoch is very essential. During the training of the proposed model, real-time data augmentation and dropout processes are used to reduce overfitting problems induced by a large number of trainable parameters. A comparison of the Gujarati printed characters dataset is shown in Table 4. On our dataset, the results obtained by the proposed Fine-tuned model are giving good accuracy.

4.3 CNN-HGC for Handwritten Gujarati Numerals

CNN – HGC is developed to recognize handwritten Gujarati characters and numerals. Well-segmented handwritten numerals only are taken into consideration to check the model's performance. The total number of images used in the training and testing of the model is 3000 (300 images for each numeral) which contains numerals with different variations like the intensity of pen, size variation, style of writing numeral, different age groups, and from -10 to +10 Skew levels. The architecture for CNN – HGC is shown in Table 5.

Experiments on a novel Gujarati dataset indicate that the proposed CNN-HGC model produces promising results. Our test results show that compared to traditional models like SVM, KNN, and FFNN, accuracy increased for the novel dataset by approximately 5 to 10% using the proposed CNN. The accuracy achieved for handwritten Gujarati numerals is 95.24%. The two problems established by traditional approaches have arguably been solved. Convolution neural networks can recognize any handwriting, in any style, from any alphabet. Also, the main reason behind achieving good accuracy is the large-sized dataset with multiple variations in the database, because the model must learn a great amount of variation in order to handle diverse handwriting styles, it must learn a large amount of variance.

Table 5. CNN architecture for handwritten Gujarati characters and numerals.

Layers	Image Size	Kernel Size	No. of Filters	Stride	Activation
	28x28	-	-	-	-
Convolution	26x26	5x5	32	1	Relu
Convolution	24x24	3x3	64	1	Relu
max pooling	-	2x2	-	2	Relu
Convolution	12x12	3x3	64	1	Relu
max pooling	-	2x2	-	2	Relu
Convolution	10x10	3x3	128	1	Relu
max pooling	-	2x2	-	2	Relu
Drop out (0.1)	-	-	-	-	-
Fully Connected	-	512 Units	-	-	-
Output	-	-	-	-	softmax

Table 6 shows a comparison with other comparable work on the same Gujarati dataset.

Table 6. Comparison of proposed model with traditional models for handwritten Gujarati numerals.

Authors	Types of datasets	Classifier	Dataset	Accuracy
Feature extraction based on stroke orientation estimation technique (Nagar and Mitra, 2015)	Numeral	SVM	12889	88.93%
Affine Invariant Moments Approach (Baheti and Kale, 2012)	Numeral	KNN, PCA, SVM, Gaussian distribution function	1600	SVM=92.28%, Gaussian=87.2%, KNN=90.04%, PCA=84.1
optical character reorganization through the neural network (Desai, 2010)	Numeral	Feed Forward Backpropagation neural network (FFNN)	3000	81.66%
Proposed CNN-HGC model	Numeral	CNN	3258	95.24%

4.4 CNN-HGC for Handwritten Gujarati Characters

Gujarati is a language for which very less work is present in literature, particularly for the handwritten script. To extend this work further for handwritten Gujarati characters, a new benchmark dataset of isolated handwritten character images is generated. Few recent research works have been analyzed and come to know that authors have prepared their own CNN architecture for the Indian languages like Bangla, Tamil, Devanagari, etc. Thus, to test the performance of the existing model for the Gujarati dataset, experiments have been carried out with existing models trained on similar work of Narang et al. (2021) and Huda et al. (2022) which are designed to recognize Devnagari and Bangla characters respectively. The dataset used to train all three models is the newly generated Gujarati handwritten characters and numerals. Among all existing work, the proposed CNN – HGC is compared with the CNN architecture of Narang et al. (2021) and Huda et al. (2022). The results are shown in Table 7.

Narang et al. (2021) performed CNN-based Devanagari ancient manuscripts recognition which contains 5484 characters arranged in 33 classes. In the experimental scenario, the overall accuracy attained was about 93.73% but on the newly generated Gujarati dataset, the accuracy recorded is 56.96%. Huda et al.

(2022) presented a deep CNN model. In the experiments, the model exhibits an average accuracy of 96.42% in recognizing BanglaLekha augmented. The model also achieved 98.92% accuracy on the NumtaDB dataset. It contains 4,74,792 character images. Using the same model but for the newly generated Gujarati dataset, the accuracy recorded is 47.56%. The proposed CNN-HGC achieves 66.76% accuracy which is 10% and 19.2% more compared to the existing CNN model of Narang et al. (2021) and Huda et al. (2022).

Table 7. Comparison of the proposed model with latest baseline models for handwritten Gujarati characters.

	Dataset	Image size	Batch size	Epoch	Loss	Accuracy
CNN (Narang et al., 2021)	Handwritten Gujarati characters & numerals	28x28	64	100	3.76	56.96%
CNN (Huda et al., 2022)		28x28	64	100	2.32	47.56%
Proposed CNN - HGC		28x28	64	100	1.24	66.76%

5. Conclusions

The variety of letter shapes, complicated character sets, and the inclusion of modifiers make OCR development for Indian languages extremely difficult. Available techniques developed in the past few years did not always produce correct results for complex languages like Gujarati. Two modified CNN models are presented in this paper for character identification on Gujarati printed base characters and handwritten numbers. Results obtained using deep learning methods are compared with traditional and some baseline methods used by various authors on similar kinds of datasets like Devanagari and Bangla. On average, the proposed CNN - PGC model achieves an 18.29% improvement in accuracy compared to other existing algorithms. The second proposed CNN - HGC model achieves an average 7.60% and 14.6% improvement in accuracy compared to other existing algorithms for handwritten Gujarati numerals and characters respectively. In spite of the large and diversified dataset used, the proposed CNN architecture achieves promising results in terms of overall accuracy compared to other traditional and baseline algorithms available in the literature. More work is needed in the future to generate a complete benchmark dataset for Gujarati handwritten characters including modifiers. In the future hybrid, CNN models should be developed to perform state-of-the-art OCR, specifically for the Gujarati script.

Conflict of Interest
There are no conflicts to declare.

Acknowledgments
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors would like to thank the editor and anonymous reviewers for their comments that help improve the quality of this work.

References

Ahamed, I., Jahan, M., Tasnim, Z., Karim, T., Reza, S.S., & Hossain, D.A. (2021). Spell corrector for Bangla language using Norvig's algorithm and Jaro-Winkler distance. *Bulletin of Electrical Engineering and Informatics*, *10*(4), 1997-2005.

Al-Jarrah, O., Al-Kiswany, S., Al-Gharaibeh, B., Fraiwan, M., & Khasawneh, H. (2006). A new algorithm for arabic optical character recognition. *WSEAS Transactions on Information Science and Applications*, *3*(4), 832-845.

Baheti, M.J., & Kale, K.V. (2012). Gujarati numeral recognition: affine invariant moments approach. *Soft Computing*, 140-146.

Chaudhary, M., Shikkenawis, G., Mitra, S.K., & Goswami, M. (2012, December). Similar looking Gujarati printed character recognition using locality preserving projection and artificial neural networks. In *2012 Third International Conference on Emerging Applications of Information Technology* (pp. 153-156). IEEE. Kolkata, India.

Das, A., & Mohanty, M.N. (2020). *Use of deep neural network for optical character recognition*. In Advancements in Computer Vision Applications in Intelligent Systems and Multimedia Technologies (pp. 219-254). IGI Global.

Desai, A.A. (2010). Gujarati handwritten numeral optical character reorganization through neural network. *Pattern Recognition*, *43*(7), 2582-2589.

Dwivedi, A., Saluja, R., & Sarvadevabhatla, R.K. (2020). An OCR for classical Indic documents containing arbitrarily long words. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 560-561).

Fink, G.A., Vajda, S., Bhattacharya, U., Parui, S.K., & Chaudhuri, B.B. (2010, November). Online Bangla word recognition using sub-stroke level features and hidden Markov models. In *2010 12th International Conference on Frontiers in Handwriting Recognition* (pp. 393-398). IEEE. Kolkata, India.

Francis, L.M., & Sreenath, N. (2020). TEDLESS–Text detection using least-square SVM from natural scene. *Journal of King Saud University-Computer and Information Sciences*, *32*(3), 287-299.

Gohel, C.C., Goswami, M.M., & Prajapati, V.K. (2015, December). On-line handwritten Gujarati character recognition using low level stroke. In *2015 Third International Conference on Image Information Processing (ICIIP)* (pp. 130-134). IEEE. Waknaghat, India.

Goswami, M., & Mitra, S. (2013, December). Structural feature-based classification of printed Gujarati characters. In *International Conference on Pattern Recognition and Machine Intelligence* (pp. 82-87). Springer, Berlin, Heidelberg.

Goswami, M.M., Prajapati, H.B., & Dabhi, V.K. (2011, November). Classification of printed Gujarati characters using SOM based k-Nearest Neighbor Classifier. In *2011 International Conference on Image Information Processing* (pp. 1-5). IEEE. Shimla, India.

Huda, H., Fahad, M.A.I., Islam, M., & Das, A.K. (2022, January). Bangla handwritten character and digit recognition using deep convolutional neural network on augmented dataset and its applications. In *2022 16th International Conference on Ubiquitous Information Management and Communication* (IMCOM) (pp. 1-7). IEEE. Seoul, Korea.

Jain, A.A., & Arolkar, H.A.A. (2018). Survey of Gujarati handwritten character recognition techniques. *International Journal for Research in Applied Science & Engineering Technology* (IJRASET), *6*(9), 461-465.

Jeong, C.B., Kwag, H.K., Kim, S.H., Kim, J.S., & Park, S.C. (2003, December). Identification of font styles and typefaces in printed Korean documents. In *International Conference on Asian Digital Libraries* (pp. 666-669). Springer, Berlin, Heidelberg.

Khan, M.M., Uddin, M.S., Parvez, M.Z., & Nahar, L. (2022). A squeeze and excitation ResNeXt-based deep learning model for Bangla handwritten compound character recognition. *Journal of King Saud University - Computer and Information Sciences*, *34*(6), 3356–3364. https://doi.org/10.1016/j.jksuci.2021.01.021

LeCun, Y. (2015). "LeNet-5, convolutional neural networks." URL: http://yann. lecun. com/exdb/lenet 20, no. 5 14.

Liao, M., Shi, B., & Bai, X. (2018). Textboxes++: A single-shot oriented scene text detector. *IEEE Transactions on Image Processing*, *27*(8), 3676-3690.

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F.E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*, 11-26.

Macwan, J.J., Goswami, M.M., & Vyas, A.N. (2016, March). A survey on offline handwritten north Indian script symbol recognition. In *2016 International Conference on Electrical, Electronics, and Optimization Techniques* (ICEEOT) (pp. 2747-2752). IEEE. Chennai, India.

Mehta, H., Singla, S., & Mahajan, A. (2016, May). Optical character recognition (OCR) system for roman script & English language using artificial neural network (ANN) classifier. In *2016 International Conference on Research Advances in Integrated Navigation Systems* (RAINS) (pp. 1-5). IEEE. Bangalore, India.

Nagar, R., & Mitra, S.K. (2015, January). Feature extraction based on stroke orientation estimation technique for handwritten numeral. In 2015 *Eighth International Conference on Advances in Pattern Recognition* (ICAPR) (pp. 1-6). IEEE. Kolkata, India.

Narang, S.R., Kumar, M., & Jindal, M.K. (2021). DeepNetDevanagari: a deep learning model for Devanagari ancient character recognition. *Multimedia Tools and Applications*, *80*(13), 20671-20686.

Naz, S., Umar, A.I., Ahmad, R., Ahmed, S.B., Shirazi, S.H., Siddiqi, I., & Razzak, M.I. (2016). Offline cursive Urdu-Nastaliq script recognition using multidimensional recurrent neural networks. *Neurocomputing*, *177*, 228-241.

Roy, S., Das, N., Kundu, M., & Nasipuri, M. (2017). Handwritten isolated Bangla compound character recognition: A new benchmark using a novel deep learning approach. *Pattern Recognition Letters*, *90*, 15-21.

Sahu, V.L., & Kubde, B. (2013). Offline handwritten character recognition techniques using neural network: a review. *International Journal of Science and Research*, *2*(1), 87-94.

Thaker, H.S. (2017). *Robot Vision Application: Using Transfer Learning on Inception-V3 to Classify Gujarati Characters* (Doctoral dissertation, Instytut Techniki Lotniczej i Mechaniki Stosowanej).

Ukil, S., Ghosh, S., Obaidullah, S.M., Santosh, K.C., Roy, K., & Das, N. (2021). Deep learning for word-level handwritten Indic script identification. In *International Conference on Recent Trends in Image Processing and Pattern Recognition* (pp. 499-510). Springer, Singapore.

Vyas, A.N., & Goswami, M.M. (2015, August). Classification of handwritten Gujarati numerals. In 2015 *International Conference on Advances in Computing, Communications and Informatics* (*ICACCI*) (pp. 1231-1237). IEEE. Kochi, India.

Wu, Y.C., Yin, F., & Liu, C.L. (2017). Improving handwritten Chinese text recognition using neural network language models and convolutional neural network shape models. *Pattern Recognition*, *65*, 251-264.

Yakin, S., Hasanuddin, T., & Kurniati, N. (2021). Application of content based image retrieval in digital image search system. *Bulletin of Electrical Engineering and Informatics*, *10*(2), 1122-1128.

Zhang, X.Y., Bengio, Y., & Liu, C.L. (2017). Online and offline handwritten chinese character recognition: A comprehensive study and new benchmark. *Pattern Recognition*, *61*, 348-360.