



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume:** 10    **Issue:** IX    **Month of publication:** September 2022

**DOI:** <https://doi.org/10.22214/ijraset.2022.46874>

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

Call:  08813907089

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

# Study on the OCR of the Devanagari script using CNN

Prof. Aparna. S. Shirkande<sup>1</sup>, Sakshi. S. Sawant<sup>2</sup>, Neha. V. Shinde<sup>3</sup>, Sharanya. S. Rao<sup>4</sup>

<sup>1</sup>Assistant Professor, <sup>2,3,4</sup>Student, Department of Electronics and Telecommunications Engineering, S. B. Patil College Of Engineering, Vangali, Indapur (M.H)

**Abstract:** Optical Character Recognition (OCR) is the electronic conversion of images of typed, handwritten or printed text into machine encoded text, whether from scanned document, a photo of a document, a scene photo. The Optical Character Recognition is emerging as a useful technology for data entry, digital data storage and as an aid to visually impaired people. The Devanagari script is composed of 47 primary characters including 14 vowels and 33 consonants. It is fourth most widely adopted writing system in the world being used for 120 languages. Convolution Neural Network (CNN) are a specialized type of artificial neural networks that use a mathematical operation called convolution in place of general matrix multiplication in at least one of their layers. They are specifically designed to process pixel data and are used in image recognition and processing. In this paper, the character recognition of printed Devanagari script using CNN is proposed.

**Keywords:** OCR, Devanagari, CNN, Feature Extraction, Preprocessing, Classification.

## I. INTRODUCTION

The printed letters or documents are easy to read for human beings, and this is happened because of use of OCR which is Optical Character Recognition technique. OCR is the recognition of printed text using computer. It is the mechanical or electronic conversion of pictures of printed content into machine-encoded content. OCR is an area of research in artificial intelligence, computer vision and pattern identification. Character recognition is a process which allows computer to recognize written or printed characters such as numbers or letters and to change them into a form that the computer can use. Character recognition having some processes like of detection, segmentation[10], and identification characters from the image. Now a days Printed Character Recognition is challenging, the deep neural networks (CNN) is capable to solving computer sight problem such as detection of object, classification, identification, and so on. CNN is one of the types of deep neural network, it automatically detecting the features from the given image. Pattern identification is an eternal area of specialization in the context of artificial intelligence, and OCR is the one of the dynamic applications of an image classification[3].

## II. DEVANAGARI SCRIPT

Devanagari Script is widely used in India, as it is used to write national Language of India, which is Hindi. Along with Hindi, Devanagari Script is also used in Marathi, Maithili, Bhojpuri, Kokani, and Nepali language. Devanagari is also called as Nagari, it is a left-right abugida, this language is based on the ancient Brahmi script, used in the Indian subcontinent. The language was developed in ancient India from the 1<sup>st</sup> to 4<sup>th</sup> century CE and was in regular use by the 7<sup>th</sup> century CE[8].

The characters of Devanagari script basically consist of 36 consonants (Vyanjan) and 13 vowels (Swar). It has particular composition rules for joining consonants, vowels, and modifiers. Modifier symbols' set is called as Matras. The combination of two constants or a constant and a vowel are used to make a compound character[1]. The line which is accessible in the upper side of a character in the script is classified "Shirorekha". In view of this Shirorekha each character is separated into three particular parts. The segment in the upper side of Shiro Rekha is called upper modifiers, in the center segment the character is accessible and in the last part lower modifiers are accessible[2]. Writing system of Devanagari script is mixture of characters, numerals and syllabary. It follows phonetic principle where many characters follow mixture of vowels and consonants as well as writing is also accordingly to the sound of characters. So Devanagari script is also called as phonetic script[3].

अ	क	आ	का	अ	ब	औ	कौ	इ	कि
ई	की	उ	कु	उ	व	ए	के	ए	के
ऐ	कै	ओ	को	ओ	कौ	ओ	कौ	ऋ	ऋ
ॠ	ॠ	ऌ	ॡ	ऌ	ॡ	अ	क	अः	कः
क	क	ख	ग	घ	ङ	च	छ	ज	झ
झ	ञ	ट	ठ	ड	ढ	ण	त	थ	द
ध	न	प	फ	ब	भ	म	य	र	ल

Fig 1. Devanagari Script.

अ	आ	इ	ई	उ	ऊ	ए	ऐ	ओ	औ
1	2	3	4	5	6	7	8	9	10
अं	अः								
11	12								

(a)

क	ख	ग	घ	ङ	च	छ	ज	झ	ञ
1	2	3	4	5	6	7	8	9	10
ट	ठ	ड	ढ	ण	त	थ	द	ध	न
11	12	13	14	15	16	17	18	19	20
प	फ	ब	भ	म	य	र	ल	व	श
21	22	23	24	25	26	27	28	29	30
ष	स	ह	क्ष	त्र	ज्ञ				
31	32	33	34	35	36				

(b)

Fig 2.(a) vowels of Devanagari (b) Consonants of Devanagari[3].

### III. TECHNIQUES

#### A. OCR

OCR (Optical Character Recognition) is a technology. It is a latest emerging technology in field of image processing. OCR technology is basically the process of converting printed handwritten, typed, etc. text in the image into the machine readable text. OCR system is hence the combination of hardware and software component to convert the printed documents into the machine/computer readable text. Hardware can be scanner or camera for reading text and software used for purpose of processing such as artificial intelligence.

Applications of OCR

- Conversion of printed text into machine editable text.
- Indexing print material for search engines.
- Aid for visually-impaired people to read document.
- Storing of the historic information.
- Translation of words.

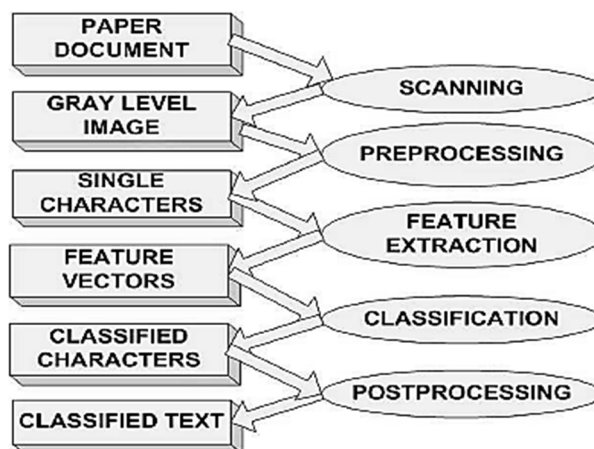


Fig 3. Basic steps of OCR[7]

#### 1) Basic steps of OCR

- Scanning:** Scanning process is taking an input image into the system for processing and recognition. Scanner is the device that takes the image and converts it into binary image. It differentiates the text from the background.

- b) *Preprocessing*: It is the step that improves the quality of the input image so that words are easily identified and can be further processed. Preprocessing is important steps as accuracy of OCR system depends on how well preprocessing is done. Preprocessing for printed text includes:
  - *Binarization*: Binarization is assigning all the pixel value present in the image either to value 0 or 1[5]. This is done by finding the threshold. Methods used Local Maxima and Minima, Otsu’s binarization, Adaptive thresholding.
  - *Skew Correction*: Skew correction step is to properly align the text with the horizontal line which might have misaligned due to improper scanning. Techniques used for skew Correction are Projection Profile method, Hough transform.
  - *Noise Removal*: Noise Removal is process of removing unwanted high intensity pixel present in an image which distorts the important features of the image.
- c) *Feature Extraction*: In this step of OCR, the character is represented as a feature vector[13]. The preprocessing stages gives single character input to the feature extraction stage and feature vectors are generated . Feature is basically information about important content of image. The use of convolution neural network done due to its effectiveness in feature extraction.[4].
- d) *Classification*: Classification is categorizing or grouping feature vectors. Two methods of Classification is Supervised and Unsupervised . It serves the basis for many problems of the computer vision.
- e) *Post processing*: This step makes the data clean by removing the induced noises due to the faults in OCR system. In this process identification and corrections are performed.

**B. CNN**

Convolution Neural Network is a type of artificial neural network that is designed to process on the pixel data of an image. Convolution was been used for image blurring and sharpening now it is also used for feature extraction. It uses the concept of acquiring the feature of the feature of the image as we human acquire that is in the layer format. Hence making the system closely related to human intelligence. Convolutional Neural Networks are a type of multi-layer neural network that is meant to extract visual patterns from pixels of the image.

The term 'Convolution' indicates the mathematical function of convolution which is a special type of linear operation wherein two functions are multiplied to produce a third function . In simple terms, two images which can be represented as matrices are multiplied to give an output that is used to extract features from the image. CNN is complex network because it has convolutional layers. CNN’s most important thing are convolutional layers. CNN forms spatial transformer layer[6].

A convolutional neural network is made up of number of layers, such as convolution layers, pooling layers, and fully connected layers. Convolution Neural Network makes use of different filters to detect the features present throughout the image. Convolution is used in Convolution neural network because it gives out the probabilities and hence it becomes easier to predict the output image or text.

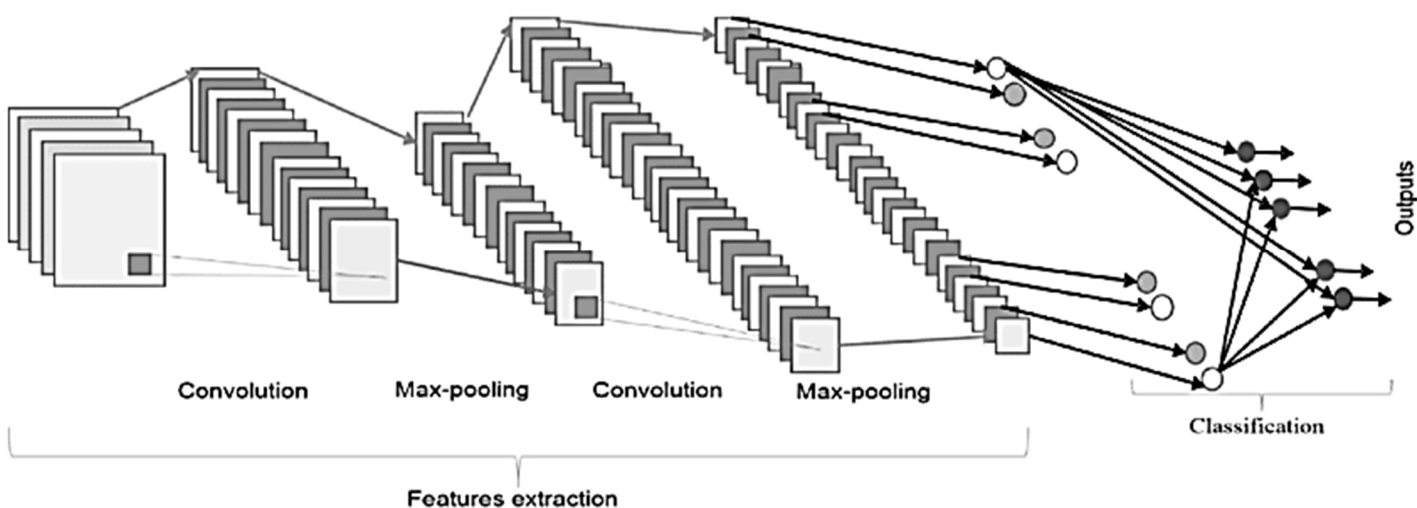


Fig 3. Architecture of CNN[9].

- 1) *Convolutional Layer*: A convolutional layer is the main building block of a CNN. It contains a set of filters (or kernels), parameters of which are to be learned throughout the training. The size of the filters is usually smaller than the actual image. Each filter convolves with the image and creates an activation map. For convolution the filter slid across the height and width of the image and the dot product between every element of the filter and the input is calculated at every spatial position[11]. Until the complete image is scanned, the kernel makes horizontal and vertical adjustments dependent on the stride rate. The kernel is less in size than a picture, but it has more depth, the kernel height and width will be modest spatially.
- 2) *Pooling Layer*: This layer is in charge of reducing dimensionality. It aids in reducing the amount of computing power required to process the data.

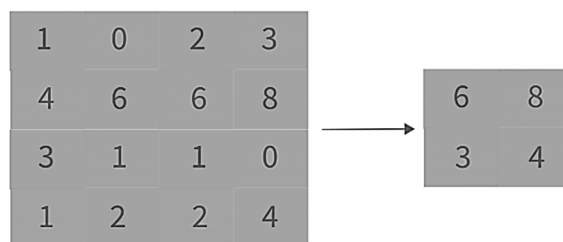


Fig 4. Pooling layer[12].

As the filter moves across the input, it selects the pixel with the maximum value to send to the output array in max pooling. As an aside, this approach tends to be used more often compared to average pooling. As the filter moves across the input, it calculates the average value within receptive field to send to the output array in average pooling.

- 3) *Fully Connected Layer (FC)*: The fully connected layer (FC) works with a flattened input, which means that each input is coupled to every neuron. After that, the flattened vector is sent via a few additional FC layers, where the mathematical functional operations are normally performed. The classification procedure gets started at this point. FC layers are frequently found near the end of CNN architectures if they are present.
- 4) *Activation Functions*: Finally, one of the most important parameters of the CNN model is the activation function. They are used to learn and approximate any kind of continuous and complex relationship between variables of the network. In simple It adds non-linearity to the network. There are several commonly used activation functions such as the ReLU, Softmax, tanH and the Sigmoid functions.

### C. Advantages

- 1) High accuracy in image recognition problems.
- 2) Automatically detects the important features without any human supervision.
- 3) Weight sharing.

## IV. CONCLUSION

Character recognition is currently a trending topic in the image processing. OCR is emerging as a powerful technology for character recognition and data entry. Devanagari script is used in many Indian and Nepalese languages, hence digital storage of this script is must. In available literature mostly handwritten recognition is done, very few literature includes printed text as well. In this paper we have presented the Optical Character Recognition of printed Devanagari script using CNN algorithm is proposed. A Devanagari character dataset is used for training and testing. We use several pre-processing and post-processing techniques for OCR. CNN algorithm yield better results when compared with other machine learning algorithm. This technology makes printed Devanagari script recognition easier.

## V. ACKNOWLEDGEMENT

We would like to thank to S. B. Patil College Of Engineering, HOD of Electronics and Telecommunications department Prof. V. U. Bansude sir, project coordinator Prof. A. S. Shirkande mam, and to department's teaching staff for their continuous guidance, and moral support. Thanks for all the help and cooperation .



## REFERENCES

- [1] Pooja Sharma, "A Review on Devanagari Character Recognition", 2018 IJRAR August, Volume 5, Issue 3, ISSN-2349-5138.
- [2] Anupama Thakur, Amrit Kaur "Devanagari Handwritten Character Recognition Using Neural Network" , International Journal Of Scientific And Technology Research Volume 8, Issue 10, October 2019, ISSN-2277-8616.
- [3] Shalaka Prasad Deore, Albert Pravin " Devanagari Handwritten Character Recognition Using fine-Tuned Deep Convolutional Neural Network on trivial dataset", Indian Academy of Sciences.
- [4] Yash Gurav , Priyanka Bhagat, Rajeshri Jadhav, " Devanagari Handwritten Character Recognition Using Convolutional Neural Networks", International Conference on Electrical , Communication and Computer Engineering , June 2022.
- [5] Prasad Chavan, Suyog Sankpal, Akshay Sonawane, Shahid Shaikh, Prof. Anup Raut, "Handwritten Devanagari Optical Character Recognition", International Journal of Innovative Research in Computer Science and Technology, ISSL-2347-5552, Volume 2, Issue 2, March 2022.
- [6] Kartik Datta, Praveen Krishnan, Minesh Mathew and C.V.Jawahar, "Offline Handwriting Recognition on Devanagari Using new Benchmark Dataset".
- [7] Farjana Yeasmin Omee, Shiam Shabbir Himel, Md. Abu Naser Bikas," A Complete Workflow for Development of Bangla OCR" International Journal of Computer Applications (0975 – 8887) Volume 21– No.9, May 2011.
- [8] <https://en.m.wikipedia.org/wiki/Devanagari>
- [9] Md Zahangir Alom , Tarek M. Taha , Chris Yakopcic , Stefan Westberg , Paheding Sidike , Mst Shamima Nasrin, Mahmudul Hasan, Brian C. Van Essen , Abdul A. S. Awwal and Vijayan K. Asari , "A State-of-the-Art Survey on Deep Learning Theory and Architectures".
- [10] Aparna Shirkande, Snehal Sabale, Dr S. T. Shirkande, "Different techniques of grey level transformation for image enhancement", International Journal For Research in Engineering Applications and Management, ISSN:2454-9150, Vol-08, Issue-01, April 2022.
- [11] <https://www.sciencedirect.com/topics/engineering/convolutional-layer>
- [12] <https://www.interviewbit.com/blog/cnn-architecture/>
- [13] Mamta Nayak, Ajit Nayak," Odia Running Text Recognition Using Moment-Based Feature Extraction and Mean Distance Classification Technique".



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)