



# Extraction of Text from Images using Machine Learning

Dr.B Sujatha | K.Sri Lasya | N.NagaSuneetha Devi | A.Gnana Surya Bhavani Prasad | P.Dileep Surya Manikanta

Department of Computer Ccience and Engineering, Godavari Institute of Engineering and Technology(A), JNTUK, Kakinada.

## To Cite this Article

Dr.B Sujatha, K.Sri Lasya, N.NagaSuneetha Devi, A.Gnana Surya Bhavani Prasad and P.Dileep Surya Manikanta. Extraction of Text from Images using Machine Learning. International Journal for Modern Trends in Science and Technology 2022, 8(S03), pp. 01-05. <https://doi.org/10.46501/IJMTST08S0301>

## Article Info

Received: 26 April 2022; Accepted: 24 May 2022; Published: 30 May 2022.

## ABSTRACT

*This article focuses primarily on text extraction from images using machine learning. Extracting and recognizing text from images is an important step in building an efficient indexing and search system for multimedia databases. Our main goal is to use machine learning to build an unlimited image indexing and search system. The block identified as text is then provided as an input for OCR (Optical Character Recognition). The OCR output in the form of ASCII characters that form the word is stored in the database as keywords for later retrieval. Many extraction techniques have been developed to search for relevant information. Therefore, successful implementation of text extraction from images in your organization requires identifying business goals and analyzing data accessible from both open source and private datasets. In addition, you need to determine if additional security measures are needed to check for accuracy errors in the OCR mechanism. The result of extracting text from an image provides an accuracy of over 95, depending on the image or dataset.*

**Keywords**—Text Extraction from images, Text Recognition, OCR technique, CNN Algorithm.

## INTRODUCTION

Extraction of text from an image is a technique which uses machine learning algorithms to extract text directly from an image without the intervention of humans. Text extraction from images generally comes to mind when thinking about how to teach artificial intelligence algorithms to read. The first step here is to teach an algorithm (text recognition) to display the text. The next step is to process it and convert it to another format (such as a text file). In this project, we'll take a closer look at these two phases of the text extraction process. We also use a variety of machine learning techniques and algorithms to get successful results from this project. Text recognition in natural images plays an important role in artificial intelligence, augmented reality, and other innovations. Helps to denoise the image and identify the text. Still, the variability of imaging conditions such as

lighting, specular reflection, agitation, darkness, proximity of blocks on the content, and the size, introduction, text style, and style of the content itself. Therefore, good text recognition calculations need to take positive action against such fluctuations. Text recognition plays an important role in our lives because it is used in visual applications. We are currently facing issues such as background complexity, different text orientations, background complexity, scene text diversity, and interfering factors. To overcome these difficulties, kmeans clustering and related methods are used, and in some cases valley and ridge techniques are also used. This is the result of recognizing the text and improving its accuracy. [9] Here we implement various methods of text recognition and classification. The main goal of the project is to recognize images that recognize text using files loaded using deep neural networks. Then, in order

to convert the image to text, we need to extract the text from pytesseract for this purpose. By doing so, we can recognize real-time video text primarily using deep neural networks.[10] Here you can see the text from the video. In this project, we'll take a closer look at these two phases of the text extraction process. We also use a variety of machine learning techniques and algorithms to get successful results from this project.

Manual text recognition is a very important innovation today. Before properly implementing this innovation, we relied on the fact that writing a typeface could end up with its own mistakes. It is difficult to keep this data accurate and collect it effectively. Difficult painting is required to maintain proper association of data. The Recurrent Neural Community is applied to find character relevance. Today we have an efficiently achievable OCR in English. Formal text English OCR is also available, but handwritten content material OCR is rare. Moreover, what is achievable is not very accurate. We plan to use a recurring neural community to create OCR that highlights the popularity of handwritten text. The proposed version runs using the conda used by the tensor float framework. The motivation for the Recurrent Neural Community is to improve accuracy.

## EXISTING MODEL

- Non-Automated Segmentation
- This method is not fully automatic because it involves few parameters that need manual specification.
- ImageThreshold

Today, textbooks are one of the most important careers in this digital word knowledge. Text can be thought of as the structure of symbols that we are accustomed to communicating, communicating, and way of life. It is one of the most important creativities of mankind. Text plays an important role in human existence. Texts can be found everywhere in our daily lives. Therefore, it is important to extract data from the text that exists in our daily life scenarios. Working with text that exists in natural scenarios has two main tasks: text identification and text review.

The process of extracting the body of text from an image is known as text recognition. There are many applications

for scene text recognition in different disciplines. Therefore, Street View text recognition and recognition has been increasingly studied in recent years. This paper focuses on how to recognize text in natural scene images as a first step. Extracting text from an image can be a tedious process. Most people just re-enter text / data from the image. However, this is time consuming and inefficient when processing many images in the real world. Text extraction plays an important role in finding important and valuable information. Text extraction includes recognizing, searching, tracking, binarizing, extracting, extending, and recognizing text from a particular image. These text characters are difficult to see and recognize due to their different sizes, fonts, styles, alignments, alignments, contrasts, and textured backgrounds of complex colors. With the increased growth of various multimedia documents and the growing need for information, identification, indexing, and search, much research has been done on text extraction in images. Various techniques have been developed to extract text from images. The proposed methods were based on morphological operators, wavelet transform, artificial neural network, skeletonization operation, edge detection algorithm, histogram technique etc. All these techniques have their benefits and restrictions. Thus, the accuracy of the results obtained from these existing methods are less when compared with the system which is proposed.

## PROPOSED SYSTEM

- The proposed system uses image preprocessing. This is the term for manipulating images at the lowest level of abstraction.

These operations do not increase the image information content, but decrease it if entropy is an information measure.

Applications:

Automatic sign recognition technique

Computer vision techniques

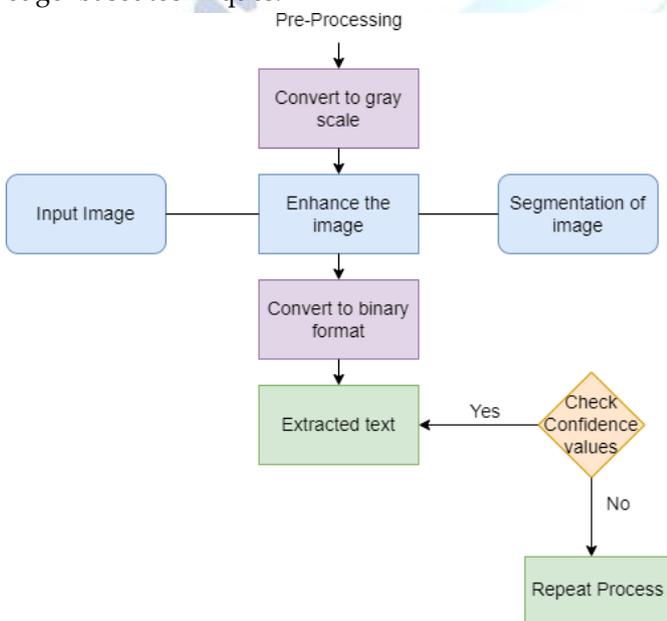
*Region Based Technique*

Region-based techniques use the color or grayscale properties of an area of text and how they differ from the corresponding properties of the background. When it comes to image representation, region-based image representation provides image simplification in that it

reduces the number of typical elements. In this representation, the objects in the scene are retrieved by merging the area into the initial partition.

*Edge Based Technique*

Edge-based text extraction algorithms are a common method of text extraction. Quickly and effectively find and extract text from documents and images. Edges are considered a very important part of the perceptual information content of a document image and represent significant variations in strength, depth discontinuities, surface orientations, changes in material properties, and so on. Vertical edges are detected using smoothing filters and combined with text clusters for text extraction using edge-based techniques.



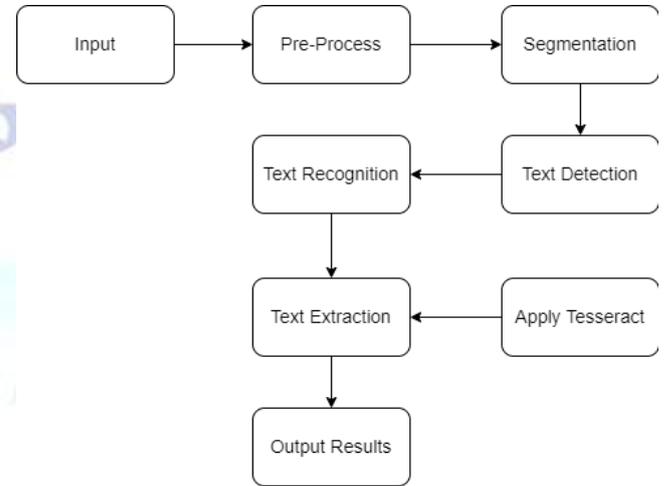
Block Diagram of Proposed System

*Texture Based Technique*

The texture-based method uses the observation that the text in the image has various texture properties that distinguish it from the background to determine if a pixel or block of pixels belongs to the text. Text feature extraction basically relies on image preprocessing techniques and is typically performed using energy measurements or non-linear operators after linearly transforming or filtering texture images. You can use this texture-based text extraction to build a robust license plate localization system.

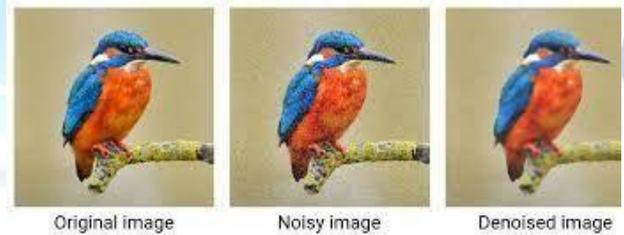
*Morphological Based Technique*

Mathematical morphology is a topology and geometry-based approach to image analysis. It provides powerful tools for extracting geometric structures and expressing shapes in many applications.



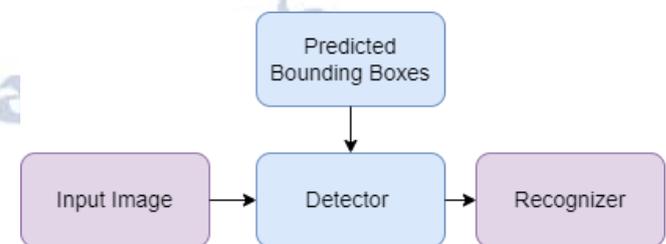
*Steps Involved*

- i) Pre-Processing
- ii) Text Detection
- iii) Text Recognition Pre-Processing



Remove the noise from the image. Remove the complex background from the image. Handle the different lightning conditions in the image.

*Text Detection*



The text recognition technique needed to recognize the text in an image and create a bounding box around the

portion of the image that contains the text. Again, standard objection detection techniques work.

The various text recognition algorithms are:

- . Sliding Window Technique
- . Single shot and Region Based Detectors(YOLO)
- . EAST(Efficient Accurate Scene Text Detector)

### Text Recognition

After recognizing the bounding box in the text, the next step is to recognize the text. There are various techniques for recognizing text. The next section describes some of the best techniques.

- . CRNN(Convolutional Recurrent Neural Network)
- . ML OCR with Tesseract

## RESULTS



Figure-1 : Image before text extraction

The above shown image is the original image which was taken as input to extract text from the image. We can also take any other image or datasets as input.



Figure-2 : Image after text extraction

Final output image after text extraction is shown in above figure-2. The bounding boxes have been detected

from each and every word in the image and the text which is extracted from the image have been displayed along with the confidence values in the shell.

```
IDLE Shell 3.9.8
File Edit Shell Debug Options Window Help
Python 3.9.8 (tags/v3.9.8:bb3fdef, Nov 5 2021, 20:48:33) [MSC v.1929 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
===== RESTART: E:\tesseract_text_tesseract.py =====
Confidence: 26
Text: a
Confidence: 96
Text: Apple
Confidence: 96
Text: Support
Confidence: 96
Text: 1-800-275-2273
>>>
```

Figure-3 : Confidence values

Finally, this method and algorithms which were used are helpful to extract text from horizontally textured images.

## CONCLUSION

With this proposed framework, various images are scanned and optically recognized. By using the CNN algorithm, the output text is printed with an accuracy of over 90. To further improve the finished exhibit, various other algorithms need to be designed to further investigate the problem and find a better placement.

In summary, the demand for text extraction from images is currently increasing. Many extraction techniques have been developed to search for relevant information. Therefore, successful implementation of text extraction from images in your organization requires identifying business goals and analyzing data accessible from both open source and private datasets. In addition, you need to determine if additional security measures are needed to check for accuracy errors in the OCR mechanism.

### Conflict of interest statement

Authors declare that they do not have any conflict of interest.

## REFERENCES

- [1] S. Sharma, D. R. Bhat, N. R. Hegde, Y. A. P. Visually Impaired People", International Journal of Engineering Research & Technology, 2018.
- [2] Text Reader", 2018 Fourth International Conference on Advances in Computing Communication & Automation (ICACCA), pp. 1-6, Oct. 2018.

- [3] D. Velmurugan, M. S. Sonam, S. Umamaheswari, S. Parthasarathy and K. R. Arun, "A Smart Reader for Visually Impaired People Using Raspberry PI", pp. 5, 2016.
- [4] H. M. Htun and T. Zin, "Text To Speech Conversion Using Different Speech Synthesis", vol. 4, no. 07, pp. 5, 2015.
- [5] D. Velmurugan, Srilakshmi, S. Umamaheswari, S. Parthasarathy and K.R. Arun, "Hardware Implementation of Smart Reader for Visually Impaired People Using Raspberry PI".
- [6] Recognition (ICDAR 2007), vol. 2, pp. 629-633, Sep. 2007.
- [7] S. Musale and V. Ghiye, "Smart reader for visually impaired", 2018 2nd International Conference on Inventive Systems and Control (ICISC), pp. 339-342, Jan. 2018.
- [8] S. R. Mache, M. R. Baheti and C. N. Mahender, "Review on Text-To-Speech Synthesizer", vol. 4, no. 8, pp.6.
- [9] Parvathi, D. S. L., Leelavathi, N., Ravikumar, J. M. S. V., & Sujatha, B. (2020, July). Emotion Analysis Using Deep Learning. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 593-598). IEEE.
- [10] Kumar, J. R., Sujatha, B., & Leelavathi, N. (2021, February). Automatic Vehicle Number Plate Recognition System Using Machine Learning. In IOP Conference Series: Materials Science and Engineering (Vol. 1074, No. 1, p. 012012). IOP Publishing."

