# Usage of Image Analytics for Text Recognition

Mehwish Shabbir, Abdul Razzaq
MSCS Student, Dept: Computing & Technology
Riphah International University, Lahore
Mehwish.shabbir22@gmail.com, abdulrazzaq1510@gmail.com

Abstract—The purpose of text recognition is to identify the content of the printed version of the report with the appropriate settings (such as .docx). The text recognition procedure involves several steps such as preparation, segmentation, highlighting, layout and post-processing. Preprocessing refers to performing basic steps on the input image, such as binarization to switch from a fine scale image to a binary image, and reduction of Clammer to expel strong signs from the image. A stage of segmentation where the image is fragmented line by line and each character in a sectioned line is segmented. Future extracts represent the character's character. The characterization involves the database and performs the mapping. Currently, he is doing important work at work and school.

**Index Terms— Classific Feature Extraction, OCR, Histogram, Segmentation, Recognition**

## 1. Introduction

OCR techniques are used text recognition, commonly [2] [5] [14], transforms an image of typed text into binary text, or a character image into a schematic. Includes a computer system designed to convert to standard encoding for representing text.OCR started as a research field of artificial intelligence [24] and computational vision [26]. Text recognition is a formal task that requires writing big data, such as post offices, banks, universities, etc., which is used in real-world applications that collect information from written text images.

## 2. Literature Review

The scanned image is checked for noise, sky, slate, etc. The picture can jump to the left or right or to Gaussian noise. Here the image is first converted to greyscale and then to binary images. So you get the right picture for the next action. After pre-processing, the noise-free image is transferred to the distribution phase, where it is merged into individual characters. I check the banner images for harmless surroundings. When an interline area is recognized, the image is divided into a series of paragraphs in the internal area. Paragraph lines are searched for horizontal intersections of the background. The image histogram [13] is used to determine the width of the horizon. Then the line at the intersection is scanned vertically. Here the histogram [13] is used to determine the width of the words. The word is then divided into letters according to the width of the letters. The Segmentation phase leads to feature extraction [2] [5] [14] to take glyphs into account in individual images and to extract features. A character has height and width.

The classification is based on the functions extracted in the previous step. This corresponds to the glyph for each character. These properties are analysed using a number of principles and identified with different classes. This classification has been generalized to work with the same fonts. In the event of conflicts, different distance metrics are selected as ranked candidates depending on the rank and character width. Classification rules for other letters are also written. This method is common because it is in the form of letters and requires no training. With a new symbol for this classification block [10] the properties are extracted, the properties compared according to rules, the letters identified and labelled.

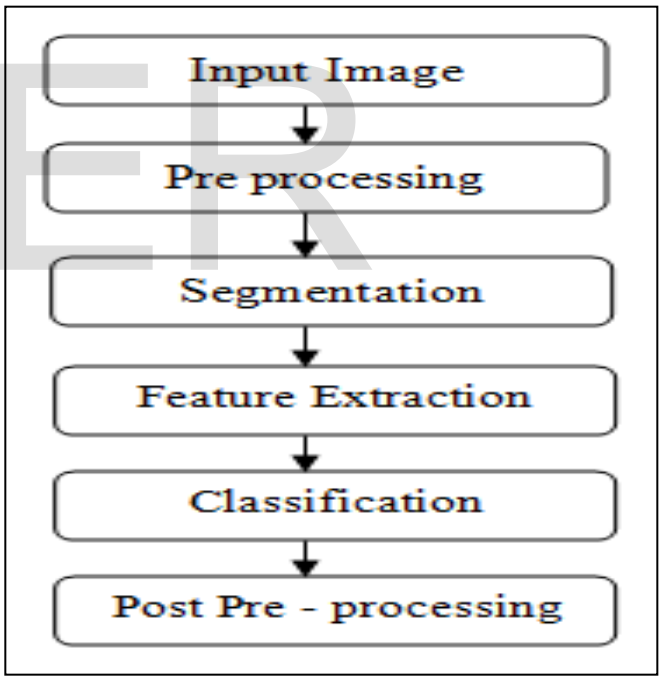


**Fig1. Text Extraction Process**

## 3. Algorithm

1. Begin
2. Input image
3. Conversion of image into greyscale
4. Greyscale to Binary image
5. Pre – processing (noise removal, skew correction)
6. Load Database
7. Segmentation of the text from image

## 4. Related Work

Developments and advances in a variety of different methods for extracting text attributes from images and videos ,including page distribution, text color extraction [2], text addressing of video images [3]. Add and index content-based images or videos. However, extensive research, designing a versatile series system is not easy. Indeed, there are different sources of different possibilities when extracting text. Images created with textured backgrounds, less contrasting or complex images, or variations in font size, style, color, orientation, and requirements. This change makes it very difficult to fix the problem automatically. I can divide text detection methods into three types. The first method comprises component methods. This assumes that the characters in the text are the same color and meet certain restrictions regarding the size, shape, and local space. However, if the colors are like the background of the text, these methods do not work. The second comprises structural methods that should have a specific structure in the text area. These methods are less sensitive to the background color, but cannot distinguish between text and background text. The third comprises of border-based methods. I find the text area assuming that the background and lines of the object are smaller than the text area. Conversely, such an approach is not very effective in recognizing large font size text. We compare methods based on Support Vector Machine (SVM) [1] and Multiplexing Presetter (MLP) [1]. Four independent functions: distance map function, local grayscale derivation function, basic function and DCT coefficient function for text authentication in relation to permanent gradient variations. They found that SVM gave better recognition results than MLP. Several solution-based text recognition methods are commonly used to recognize text at different scales.
Video → Image Frame Extraction → Segmentation → Classification → Text Extraction → Character Recognition

## 5. Text Extraction

OCR aims to categorize visual patterns in alphabetical or other order. The OCR process includes numerous steps such as segmentation, characteristic drawing out, and classification. In principle, standard OCR software [2] [7] can identify text in video frmaes. However, a closer look at the characteristics of the candidate role area of a fragmented frame or icon [13] shows that most OCR software is very difficult to recognize from the text. The images in the document are original from a nature map since they mainly contain graphics and text with images. The resolution of images captured on portable devices is so low that it is more difficult to apply the complete layout structure of a document (logical or physical) and even more difficult to apply to standard OCR systems. For this reason, I have proposed it to render captured images of documents at low resolution and shallow depth. For the original stored electronic document, it is easy to extract the same signature. Converts PDF or PowerPoint files into relatively high resolution signed images. Finally, the order is compared to the signatures of all the original electronic documents to compare the signatures of the captured documents.

a. *Pre- Training Phase*
Symptoms in WSD come from a variety of sources, and distributors always understand them regardless of the relationship between them. The hidden part of RBM is the combination of features that provide a model for this relationship. Breaking down RBM allows you to study the complex relationships of knowledge sources. The learned layer will be used as a guide to insert another hidden layer. This step is not protected and does not require name tags.

## 6. Text Extraction Process

Extracting text and recognizing it includes five steps: text detection, text detection, text tracking, truncation or binary [6], and character recognition. Figure 2 shows the architecture of the text export process.

1. *Finding Phase*
2. *Localization Phase*
3. *Tracking Phase*
4. *Binarization Phase*
5. *Recognition Phase*

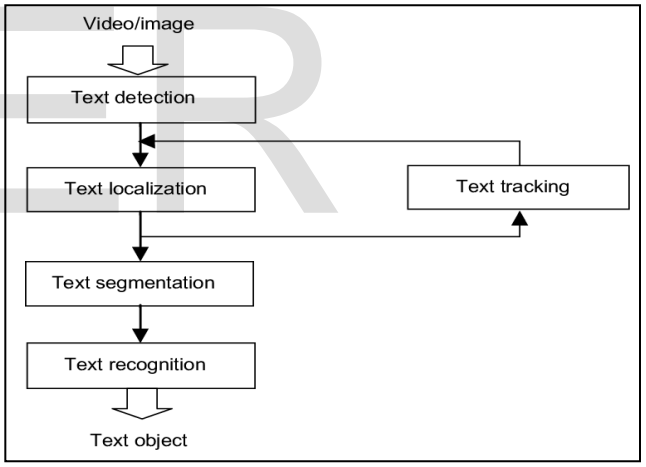


**Fig2. Text Extraction Architecure**

## 7. Text Extraction Application

There are many applications for extracting text from images. With the rapid growth of multimedia data, the need to understand content has increased. The following describes some text extraction applications.

a. *Video and picture Retrieval*:
Content-based image and video restoration have been the focus of many researchers in recent years. The text displayed in the image shows the essence of the actual content of the image and shows the human perception of the content. It is an essential tool for indexing and retrieving multimedia content [3]. This tool can provide much better results than other recovery methods based on shape, texture, or color [8]. Text

embedded in videos and images tells the human identity of the content. Therefore, it is ideal for indexing and searching multimedia data.

b. *Multimedia summarization:*
With the exponential growth of multimedia data, more information is available. Because of this excellent information, there is the problem of a sizeable amount of information. The text summary provides a solution to the problem. The text in the video clip probably provides useful information about its content. It displays text statistics in videos that contain information to help you interpret and automatically create content summaries. I have found unique ways to solve this problem. Sports video summaries and news summaries are popular apps for visual information summaries.

c. *Retrieval of the web pages:*
Web image text extraction methods can really improve the indexing and retrieval of web pages. The main indexing terms are integrated into the title image or banner. Most sites use images instead of text to display the titles of web pages. Therefore, in order to accurately index and retrieve web pages, you need to understand the text of the images. This improves indexing and allows for more efficient and more precise searches [10]. You can also filter the image by offensive words by extracting text from the web image. It also helps convert web pages to voice. The above application is not the only example of a text extraction method. There are many other applications such as voice coding of the shutter, intelligent transport systems, image marking, robotic vision, and scene analysis.

## 8. CONCLUSION

In this article, we suggest an algorithm to solve the offline character identification problem. We have provided contributions in the form of images. I trained the algorithm on the training data that was initially in the database. We have preprocessed, segmented, and traced the line. This article presents a quick survey of applications in various fields with experience in some fields. The proposed method is effective for extracting many bimodal images, including opacity and lighting. This article will serve as a good bibliographic investigation for researchers starting work on optical character recognition.

## REFERENCES

[1] Ramayana. R. ET, "Nuclear Technology for Identifying English Fonts Using Support Vector Machines".
[2] [Even lines, "optical character recognition",
[3] And then. , "Extract text using artificial neural networks.
[4] Matisse "Rejection Strategy for Recognizing Offline Handwritten Text"
[5] P. Santhanam, "A Survey of Different Methods of Extracting Text from Images",
[6] Chen, Kim Share, "A Survey on Text Recognition and Identification of Photos and Videos."
[7] Suwang Yin and Hong-Wei Hao, "Find powerful text on landscape icons.
[8] For BS, "End-to-end scene text recognition".
[9] Yongkang Song, "Identification of landscape text in multi-channel connection component segments",
[10] Muhammad Hanif, "Text detection and location in complex landscape images.
[11] Teofilo e Bodla "Identifying Features in Natural Maps",
[12] Jerry Mattas, "Scene Location and Real-Time Text Recognition,"
[13] Yi Xiong, "Extracting text from landscape images.
[14] [Wang, Bei Gao, "Scene Text Detection Using the Most Stable, Super-Regional Constructed Graphic Model."
[15] JML. Azul, "Neural Network-Based Systems Manual Fingerprint OCR Application".
[16] R. Gass Wheeler, M. Camvar, S. Bluja, "What's going on with the capture?
[17] B. Joanna, "Building Institutional Resources at the University of Pablo.
[18] Scooby, S. Kumar, "Automatic music player based on OMR".
[19] S. Shrihari, "Identifying Online and Offline Handwriting: A Comprehensive Survey.
[20] TSS Chen, "Automatic License Plate Identification".
[21] J. Maui, "Survey on Image Separation".
[22] MS Callus, "Artificial Neural Networks for Early Detection of Biological Brain Tumors",
[23] [twenty four] /. Price, "Simple and effective Horstick for global modifications in continuous space",
[24] I.H. Chang, "Hypertext Assisted Video Indexing and Content Recovery".
[25] SK Paul, "Outline of Image Segmentation Techniques",