

## Smart OCR for Document Digitization

### 2. Introduction

#### 2.1 Overview

- .)Building text extraction code
- .)Building HTML code
- .)Building Python code
- .)Running the App

#### 2.2 Purpose

With the advent of OCR techniques, much time has been saved by automatically extracting the text out of a digital image of any invoice or a document. Currently, this is where most organizations that use OCR for any form of automation are Digital copies of invoices or documents are obtained by scanning or taking pictures. The text is extracted from these documents is entered into a template-based data entry software.

### 3.Results

#### 3.1

Smart OCR

#### Smart OCR :

Optical character recognition or optical character reader is the electronic or mechanical conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo or from subtitle text superimposed on an image.

Upload PDF Here

Choose...



## A Simple PDF File

This is a small demonstration .pdf file -

just for use in the Virtual Mechanics tutorials. More text. And more text. And more text. And more text. And more text.

And more text. And more text. And more text. And more text. And more text. And more text. Boring, zzzzz. And more text. And more text. And more text. And more text. And more text. And more text. And more text.

And more text. And more text. And more text. And more text. And more text. And more text. And more text. Even more. Continued on page 2 ...

### 4.Applications

OCR can be used for: Data entry for business documents, e.g. Cheque, passport, invoice, bank statement and receipt.

### 5.Conclusion

Hence, optical character recognition for recognizing text inside images.

### 6.Future Scope

OCR may be used in humanoid robots, scanned text, information extraction in software.