

**Conversion of Hand Written Text into**

**Computer-Understandable Text**

*Piyush Mahajan*

Roll no. 7 (Reg no. 11803539)

*Raghavendra Pratap Singh*

Roll no. 8 (Reg no. 11803550)

*Nakulesh Gosh*

Roll no. 9 (Reg no. 11803553)

**Department of Intelligent Systems**

**School of computer Science Enginnering**

**Lovely Proffesional University, Jalandhar**

April-2020

**Contents**

*Topics Page no.*

1. Introduction …................................. 3
2. Steps for the conversion of texts …................................. 4
3. How it works …................................. 5
4. Important points for conversion …................................. 6

5. Need of conversion …................................. 7

**Introduction**

Handwriting Detection is a technique or ability of a Computer to receive and interpret intelligible handwritten input from source such as paper documents, touch screen, photo graphs etc.

Handwritten Text recognition is one of area pattern recognition.

The purpose ofpattern recognition is to categorizing or classification data or object of one of the classes or categories.

Handwriting recognition is defined as the task of transforming a language represented in its spatial form of graphical marks into its symbolic representation.

The goal of handwriting is to identify input characters or image correctly then analyzedto many automated process systems

The automatic recognition of handwritten text can be extremelyuseful in many applications where it is necessary to process large volumes of handwritten data, such as recognition of addresses and postcodes on envelopes, interpretation of amounts on bank checks, document analysis, and verification of signature

Therefore, computer is needed to be able to read document or data for ease of document processing.

**Steps for the conversion of texts**

To run this project we need two things

* Tesseract
* Pip install pytesseract

We need to import pytesseract as tess.

After this, we give the location of pytesseract.

From PIL (Python Image Library), import image.

By using variable image.open, we open the image.

We convert this image into string to get the converted version of image.

**How it works ?**

To convert the image into text we use OCR.

OCR stands for Optical Character Recognition.

It is a widespread technology to recognise text inside images, such as scanned documents and photos.

OCR technology is used to convert virtually any kind of images containing written text (typed, handwritten or printed) into machine-readable text data.

Probably the most well known use case for OCR is converting printed paper documents into machine-readable text documents.

Once a scanned paper document went through OCR processing, the text of the document can be edited with word processors like Microsoft Word or Google Docs.

**Important points to be considered for conversion**

* Text Density
* Structure of text
* Fonts
* Character type
* Artifacts
* Location

**Need Of Conversion**

Before OCR technology was available, the only option to digitise printed paper documents was to manually re-typing the text.

Not only was the massively time consuming, it also came with inaccuracy and typing errors.

OCR is often used as a “hidden” technology, powering many well known systems and services in our daily life.

Less known, but as important, use cases for OCR technology include data entry automation, indexing documents for search engines, automatic number plate recognition, as well as assisting blind and visually impaired persons.