Modern Education Society's College of Engineering Pune-01

Department of Computer Engineering

Name of Student:	Class:
Semester/Year:	Roll No:
Date of Performance:	Date of Submission:
Examined By:	Assignment No: 1

Laboratory Practice – V (Deep Learning)

ASSIGNMENT NO: 02

AIM: **Classification using Deep neural network**

Multiclass classification using Deep **Neural** Networks: Example: Use the OCR letter recognition.

OBJECTIVES: Students should be able to design and develop the OCR letter recognition system using multiclass classification using Deep Neural Network.

PREREQUISITE:

1. Basic of programming language

2. Concept of Neural Network

3. Concept of Multiclass Classification

THEORY:

Optical Character Recognition is a method of extracting text from image. Its main purpose is to make editable document from existing paper or image files. Optical character recognition task involves identifying simple edge detection technique and matching them with predefined patterns. It is a compartment of image recognition and is extensively used as a form of data entry with the input being some sort of printed document or data record like statements from bank, invoices, resume, business card and passport.

INTRODUCTION:

Nowadays obtaining information and altering the content in pictures which are present within background images are time consumable. Optical Character Recognition (OCR) idea emerges to take care of this issue. OCR works primarily utilizing the AI based computation and it is significant in developing and exploring in man-made brainpower. Optical character acknowledgment permits to change over the characters in printed archives, computerized picture and examined records

with the word group. The conventional method for entering the information of the printed reports, checked records and picture into the PC is through console, which is inefficient when there is a huge volume of information. OCR is utilized to enter the information from those records electronically, without the intercession of people.

We live in times when any organization or company to scale and to stay relevant has to change how they look at technology and adapt to the changing landscapes swiftly.

We already know how Google has digitized books. Or how Google earth is using NLP (or NER) to identify addresses. Or how it is possible to read text in digital documents like invoices, legal paperwork, etc.

But how does it work exactly?

In this era of digitization, storing, editing, indexing and finding information in a digital document is much easier than spending hours scrolling through the printed/handwritten/typed documents. And moreover searching something in a large non-digital document is not just time-consuming but also, it is likely for us to miss the information while scrolling the document manually. Lucky for us, computers are getting better every day at doing the tasks humans thought only they could do, often performing better than us as well. Extracting text from images has found numerous applications.

Some of the applications are Passport recognition, automatic number plate recognition, converting handwritten texts to digital text, converting typed text to digital text, etc.

Before going through how we need to understand the challenges we face in OCR problem.

Many OCR implementations were available even before the boom of deep learning in 2012. While it was popularly believed that OCR was a solved problem, OCR is still a challenging problem especially when text images are taken in an unconstrained environment .for e.g complex backgrounds, noise, lightning, different font, and geometrical distortions in the image.

It is in such situations that the machine learning OCR (or machine learning image processing) tools shine. Challenges in the OCR problem arises mostly due to the attribute of the OCR tasks at hand. We can generally divide these tasks into two categories:

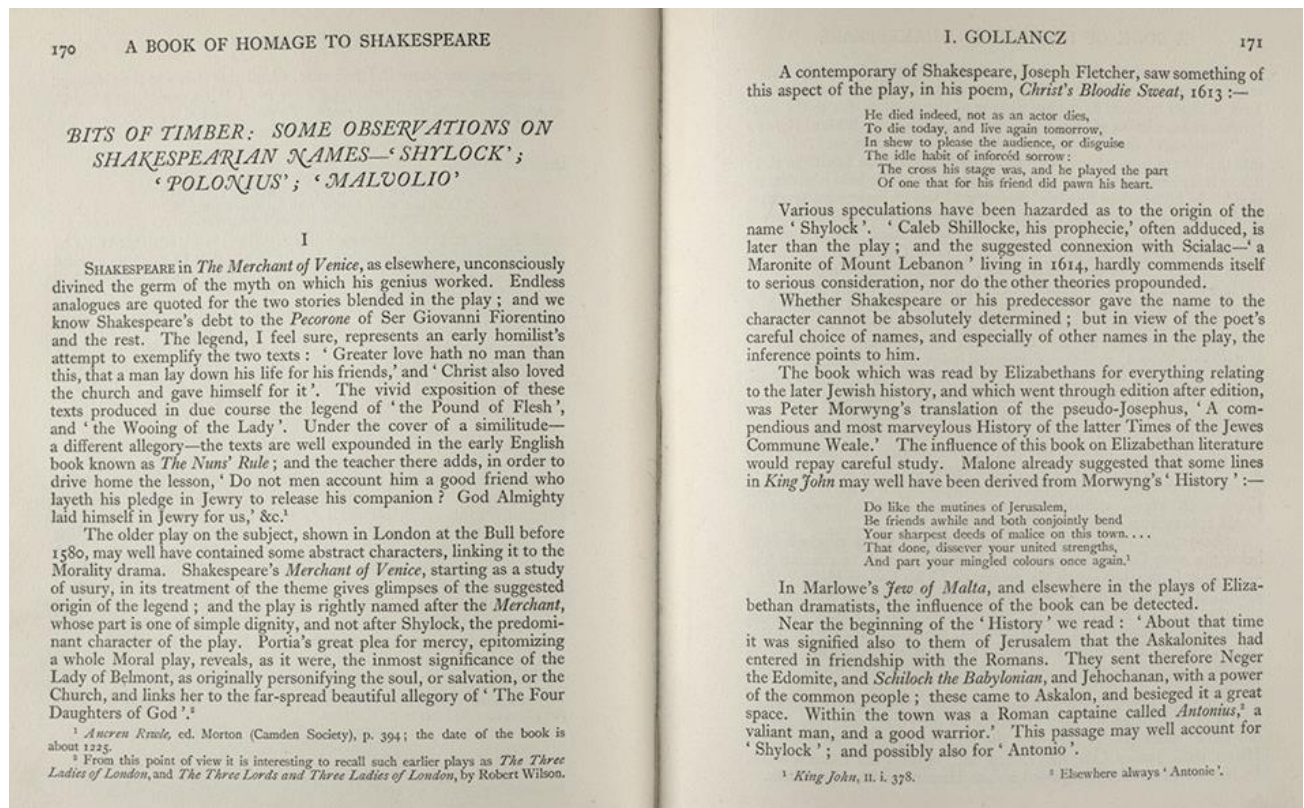**Structured Text**- Text in a typed document. In a standard background, proper row, standard font and mostly dense.

Figure 1: Structured Text: Dense, readable standard fonts; Image source: https://pixabay.com

Unstructured Text- Text at random places in a natural scene. Sparse text, no proper row structure, complex background, at random place in the image and no standard font.

Figure 2: Unstructured Texts: Handwritten, Multiple fonts and sparse; Image source: https://pixabay.com

Any Typical machine learning OCR pipeline follows the following steps shown in Figure 3:
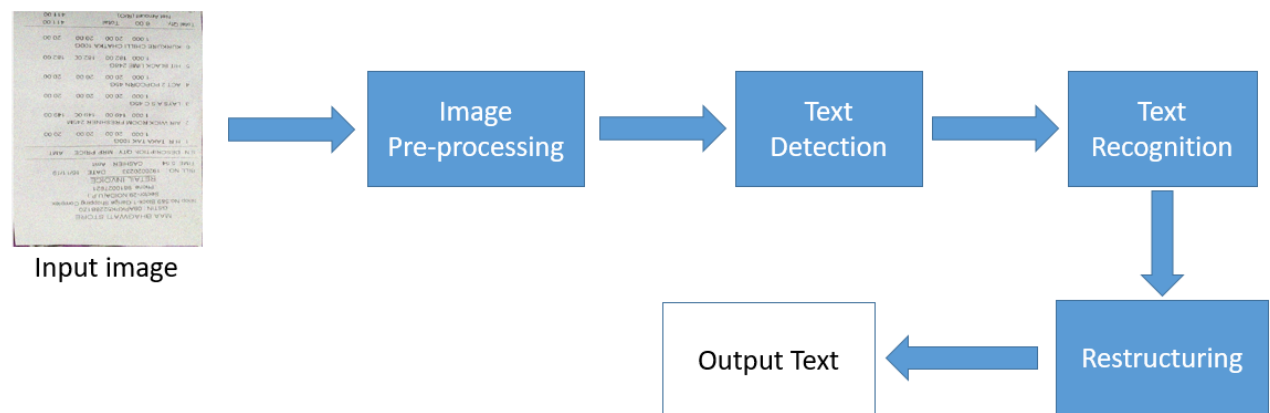

Input image

Figure 3: OCR Flow

Preprocessing

Remove the noise from the image

Remove the complex background from the image

Text Detection

Text detection techniques required to detect the text in the image and create and bounding box around the portion of the image having text. Standard objection detection techniques will also work here.

Text Recognition

Once we have detected the bounding boxes having the text, the next step is to recognize text. There are several techniques for recognizing the text.

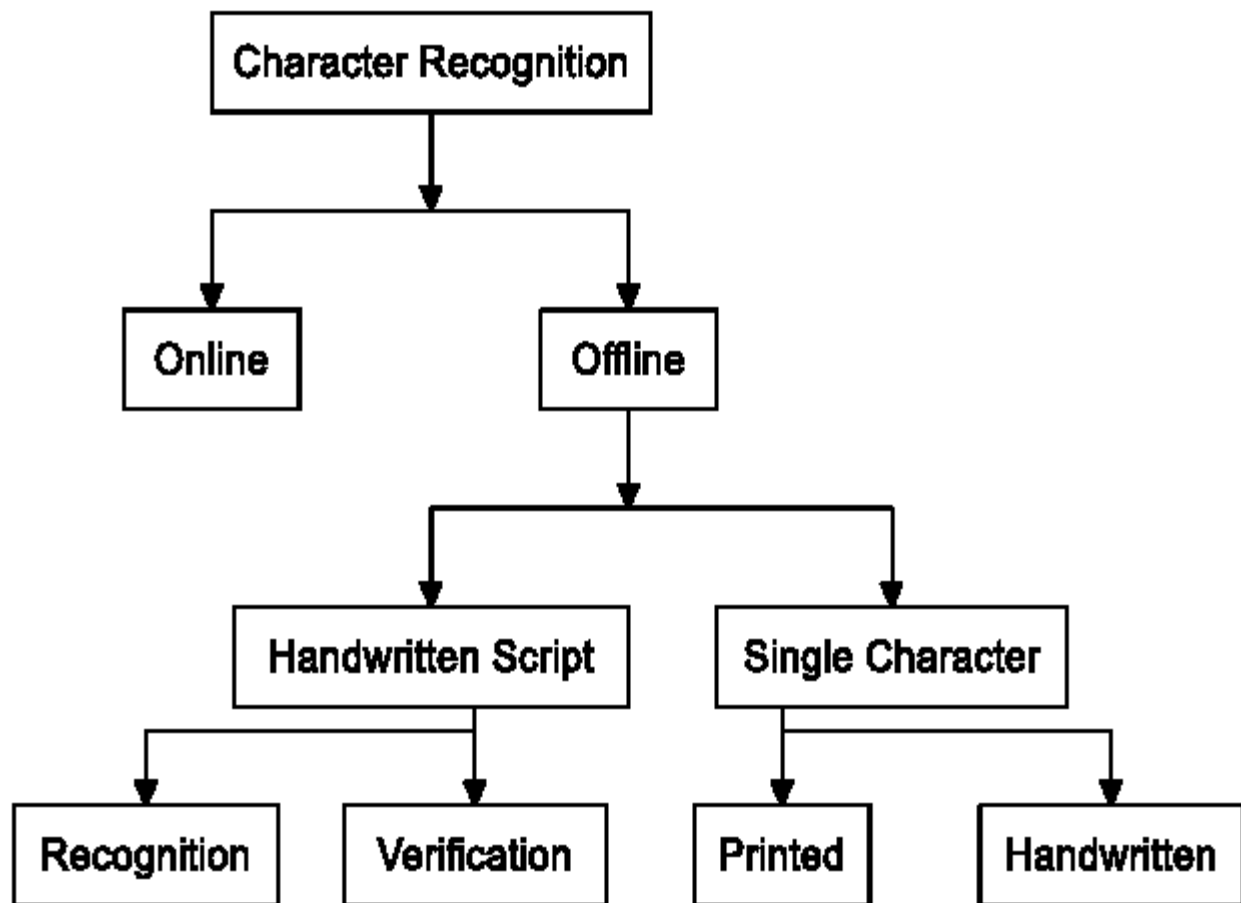Types of OCR are shown in the Figure 4.

Figure 4: Types of OCR

DATA SET DESCRIPTION: (dataset https://archive.ics.uci.edu/ml/datasets/letter+recognition)

1. Title: Letter Image Recognition Data

2. Source Information

  -- Creator: David J. Slate

   -- Odesta Corporation; 1890 Maple Ave; Suite 115; Evanston, IL 60201

  -- Donor: David J. Slate (dave@math.nwu.edu) (708) 491-3867

  -- Date: January, 1991

3. Past Usage:

  -- P. W. Frey and D. J. Slate (Machine Learning Vol 6 #2 March 91):

     "Letter Recognition Using Holland-style Adaptive Classifiers".   The   research   for   this article investigated the ability of several   variations of Holland-style adaptive classifier systems to learn to correctly guess the letter categories associated with vectors of 16 integer attributes extracted from raster scan images of the letters. The best accuracy obtained was a little over 80%. It would be interesting to see how well other methods do with the same data.

4. Relevant Information:

   The objective is to identify each of a large number of black-and-white rectangular pixel displays as one of the 26 capital letters in the English alphabet. The character images were based on 20 different fonts and each letter within these 20 fonts was randomly distorted to produce a file of

   20,000 unique stimuli. Each stimulus was converted into 16 primitive numerical attributes (statistical moments and edge counts) which were then scaled to fit into a range of integer values from 0 through 15. They typically train on the first 16000 items and then use the resulting model

   to predict the letter category for the remaining 4000.

5. Number of Instances: 20000

6. Number of Attributes: 17 (Letter category and 16 numeric features)

7. Attribute Information:

    1.   lettr   capital letter   (26 values from A to Z)
    2.   x-box   horizontal position of box   (integer)
    3.   y-box   vertical position of box   (integer)
    4.   width   width of box   (integer)
    5.   high   height of box   (integer)
    6.   onpix   total # on pixels   (integer)
    7.   x-bar   mean x of on pixels in box   (integer)
    8.   y-bar   mean y of on pixels in box   (integer)
    9.   x2bar   mean x variance   (integer)
    10.   y2bar   mean y variance   (integer)
    11.   xybar   mean x y correlation   (integer)
    12.   x2ybr   mean of x * x * y   (integer)
    13.   xy2br   mean of x * y * y   (integer)
    14.   x-ege   mean edge count left to right (integer)
    15.   xegvy   correlation of x-ege with y   (integer)
    16.   y-ege   mean edge count bottom to top   (integer)
    17.   yegvx   correlation of y-ege with x   (integer)

8. Missing Attribute Values: None

9. Class Distribution:

    789 A   766 B   736 C   805 D   768 E   775 F   773 G
    734 H   755 I   747 J   739 K   761 L   792 M   783 N
    753 O   803 P   783 Q   758 R   748 S   796 T   813 U
    764 V   752 W   787 X   786 Y   734 Z

PERFORMANCE ANALYSIS: - To evaluate the performance of the system, four metrics have been used to assess device efficiency, namely precision, accuracy, recall and F-measure.

CONCLUSION:

OCR letter recognition system (Multiclass classification) using Deep Neural Networks detection system is developed successfully.

QUESTIONS:

1. What is Multiclass Classification?

2. What is a Deep Neural Network?