**Gujarati Text Recognition**

**Prepared by**

Dhvanit Aghara (16IT001)

Keya Bhanvadia (16IT007)

**Under the supervision of**

Prof. Sandip Patel

A Report Submitted to

Charotar University of Science and Technology

for Partial Fulfillment of the Requirements for the

Degree of Bachelor of Technology

in Information Technology

IT350 Software Group Project-III (6th sem)

**Submitted at**



**DEPARTMENT OF INFORMATION TECHNOLOGY**

**Chandubhai S. Patel Institute of Technology**

**At: Changa, Dist: Anand – 388421**

**April 2019**



**CERTIFICATE**

This is to certify that the report entitled “**Gujarati Text Recognition**” is a bonafied work carried out by **Dhvanit Aghara (16IT001), Keya Bhanvadia (16IT007)** under the guidance and supervision of **Prof. Sandip Patel** for the subject **Software Group Project-III(IT350)** of **6th** Semester of Bachelor of Technology in **Information Technology** at Faculty of Technology & Engineering – CHARUSAT, Gujarat.

To the best of my knowledge and belief, this work embodies the work of candidate themselves, have duly been completed, and fulfils the requirement of the ordinance relating to the B.Tech. Degree of the University and is up to the standard in respect of content, presentation and language for being referred to the examiner.

|  |
| --- |
| Under supervision of,  Prof. Sandip Patel  Assistant Professor  Dept. of Information Technology  CSPIT, Changa, Gujarat. |
| Dr. Prof. Parth Shah  Head & Associate Professor  Department of Information Technology  CSPIT, Changa, Gujarat. |

**Chandubhai S Patel Institute of Technology**

At: Changa, Ta. Petlad, Dist. Anand, PIN: 388 421. Gujarat

**ACKNOWLEDGEMENT**

Working on this project “Gujarati Text Recognition” was a source of immense knowledge for us. We are really grateful to Parth Shah, HOD - Department of Information Technology, Dean - Faculty of Technology & Engineering for providing us with an opportunity to undertake this project in this university. We are highly thankful to Prof. Sandip Patel for their active support, valuable time and advice during the study and in completing the assignment of preparing the said project within the time provided.

The successful completion of this project would not have been possible without the dedicated support from all our mentors and friends.

**ABSTRACT**

This project ‘Gujarati Text Recognition’ is used to detecting text from image and correctly

recognize it. Model is trained to work for provided image as well as live image (through

webcam) in python language. It can read on almost all types of image formats including JPEG,

PNG, BMP. The solution developed for this project consists of three major parts: image pre-

processing, character recognition (using deep learning based CNN) and character

segmentation.

Model can recognize a word of more than one characters in English but can recognize only

single character in Gujarati. Also, made comparison between the accuracy of model working

on English dataset and Gujarati dataset.

TABLE OF CONTENTS

Acknowledgement 3

Abstract 4

Chapter 1 Introduction 7

1.1 Purpose 7

1.2 Scope 7

1.3 Objective 7

1.4 Applications 8

Chapter 2 System Analysis 9

2.1 User Characteristics (Roles of users who is dealing with the system) 9

2.2 Tools & Technology (Hardware/ Software Requirements- Minimum

requirements to run the system) 9

Chapter 3 Problems & Solutions 10

3.1 Problems Faced 10

3.2 Solutions 10

Chapter 4 System Design 11

4.1 Flow of System (Graphical representation of project) 11

4.2 Major Functionality (Data Flow Diagram) ……………………………………12

Chapter 5 Implementation 13

5.1 Implementation Environment (Single vs Multi user, GUI vs Non GUI) 13

5.2 Snapshots of Project 16

5.3 Accuracy Comparison 18

Chapter 6 Constraints and Future Enhancement 19

Chapter 7 Conclusion (Learning Outcome-In your own words) 20

References 21

**LIST OF FIGURES**

Fig 4.1.1 Flow of System 11

Fig 4.1.2 Data Flow 12

Fig 5.1.1 CNN 15

Fig 5.2.1 Predict English Word 16

Fig 5.2.2 Predict English Word-2 16

Fig 5.2.3 Predict Gujarati Character 16

Fig 5.2.4 Predict Gujarati Character-2 17

Fig 5.2.5 Loading Character Dataset 17

Fig 5.2.6 Training Model 18

Fig 5.2.7 Text Detection from Webcam 18

Fig 5.3.1 Accuracy Comparison 18

**Chapter 1**

**INTRODUCTION**

* 1. **Project Overview**

To develop character recognition software system to perform document image analysis which transforms documents in paper format to electronic format. For this process there are various techniques in the world. Among all those techniques we have chosen Optical Character Recognition as main fundamental technique to recognize characters. OCR works for various languages. We have worked on two languages English and Gujarati.

So, our project is based on the OCR technique to correctly detect the characters. Characters can be detected form provided image or from live image. A comparison between the accuracy of model working on Gujarati and English dataset is shown.

* 1. **Purpose**

The main purpose of this project is to detect and recognize the characters properly and thereby reducing the manual errors. The goal of Optical Character Recognition (OCR) is to classify optical patterns (often contained in a digital image) corresponding to alphanumeric or other characters.

* 1. **Scope**

Using optical character recognition system, we can detect and recognize text from provided image as well as from live image through webcam. Further detection and recognition can be done for different languages. Also can detect and recognize the text from paragraph having multiple languages.

* 1. **Applications**
* Banking : To process checks and passbooks without any human interaction
* Healthcare : Extract information of patient from form
* Automatic number plate recognition
* In airports, for passport recognition and information extraction
* More quickly make textual versions of printed documents, e.g. book scanning
* Data entry for business documents, e.g. invoice, bank statement and receipt

**Chapter 2**

**SYSTEM ANALYSIS**

**2.1) User Characteristics**

User must be familiar with concepts of python as technically the platform used is spyder.

**2.2) Tools and Technology**

* Programming language: Python
* IDE: Spyder
* Convolutional Neural Networks
* OpenCV (library for image processing)
* TensorFlow (python library for implementing deep learning)
* Keras (high level neural network API in python)

**Chapter 3**

**PROBLEMS & SOLUTIONS**

**3.1) Problems Faced:**

* The problem with manually written gujarati text becomes very inconvenient or can cause human errors.
* Writing the text from some hard paper like newspapers or magazine is very time consuming and is difficult to re-correct the mistakes.
* The handwritten text is separated into lines, lines into words and words into characters. Incorrect segmentation of line, word, or character decreases the recognition accuracy. Segmentation of handwritten script and Gujarati script in particular is a difficult task due to the curvature shapes of characters and varying writing style of different writers.
* Specially, character segmentation method needs to be improved which can handle joint character in gujarati language.

**3.2) Solutions:**

* Approaching a digital way, reduces the manual errors and ease the flow. We just need to capture the image and through the OCR engine, text is detected and recognition from that captured image.
* Time consumption is reduced as there is no need of manual writing and in fraction amount of time, text is recognized without any error.
* Need to improvise the Gujarati dataset for obtaining good accuracy for curvature shapes of characters.

**Chapter 4**

**SYSTEM DESIGN**

**4.1) Flow of system:**

Image Preprocessing

(e.g. thresholding , bluring)

Text Detection

Character Segmentation

Character Recognition

Output Text

Figure 4.1.1 | Flow of System

**4.2) Major Functionality (Data Flow Diagram)**

Training set

(Images)

Testing set

(Images)

Predictive model

New image for text prediction

Output

Feature Vectors

Feature Vectors

Feature Vector

Figure 4.1.2 | Data Flow

**Chapter 5**

**IMPLEMENTATION PLANNING**

**5.1) Implementation Environment**

**Four main steps are implemented to compute the output from the image. They are:**

* Image Preprocessing: In this stage input image will go through the various image preprocessing techniques like thresholding, blurring, noise removal etc. This stage is

important because it removes unnecessary information from input image.

* Text Detection: Text is detected from image is done by using OpenCV’s EAST text

detector which give detected text.

* Character Segmentation: After text is detected, now we need to segment each and every

character from detected text. The segmentation is done with OpenCV’s contour

detection.

* Character Recognition: Each segmented character will now pass to the our trained model which will predict the character. The model is trained using convolutional neural networks.

Model for English text recognition is trained on 4627 images and validated on 1157 images.

Model for Gujarati text recognition is trained on 4360 images and validated on 1090 images.

* Output text: After proper recognition, we need to combine each and every predicted

character and make it in the text form.

* Our model is working on supervised machine learning algorithm. We have used

Convolution Neural Network to build our model.

**Convolution Neural Network:**

To implement OCR to correctly recognize the character , we need to use neural network, i.e. Convolutional Neural networks (subpart of deep learning).

A convolutional neural network (CNN) is a specific type of artificial neural network that uses perceptrons, a machine learning unit algorithm, for supervised learning, to analyse data. CNNs are applied to image processing, natural language processing and other kinds of cognitive tasks.

**First Layer: Convolutional layer**

* The first layer in a CNN is always a Convolutional Layer. Convolution is the first layer to extract features from an input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data. The convolution layer comprises of a set of independent filters. Each filter is independently convolved with the image.
* After each conv. layer, it is convenient to apply a nonlinear layer (or activation layer) immediately i.e. ReLU layer for this model.ReLU stands for Rectified Linear Unit for a non-linear operation.
* Why ReLU is important : ReLU’s purpose is to introduce non-linearity in our ConvNet. Since, the real world data would want our ConvNet to learn would be non-negative linear values.

**Second Layer: Pooling Layer**

* Pooling Layer is also called as down sampling layer.Pooling layers section would reduce the number of parameters when the images are too large. Spatial pooling also called downsampling which reduces the dimensionality of each map but retains the important information.
* Pooling layer serves two main purposes. The first is that the amount of parameters or weights is reduced by 75%, thus lessening the computation cost. The second is that it will control overfitting.
* A symptom of overfitting is having a model that gets accuracy 100% or 99% on the training set, but very less on the test data.

**Third Layer: Fully Connected Layer**

* The layer we call as FC layer, we flattened our matrix into vector and feed it into a fully connected layer like neural network. Fully Connected Layer (FC) basically takes an input volume (whatever the output is of the convNet or ReLU or pool layer preceding it) and outputs an N dimensional vector where N is the number of classes that the program has to choose from.
* Here we are using SoftMax approach to represent output.

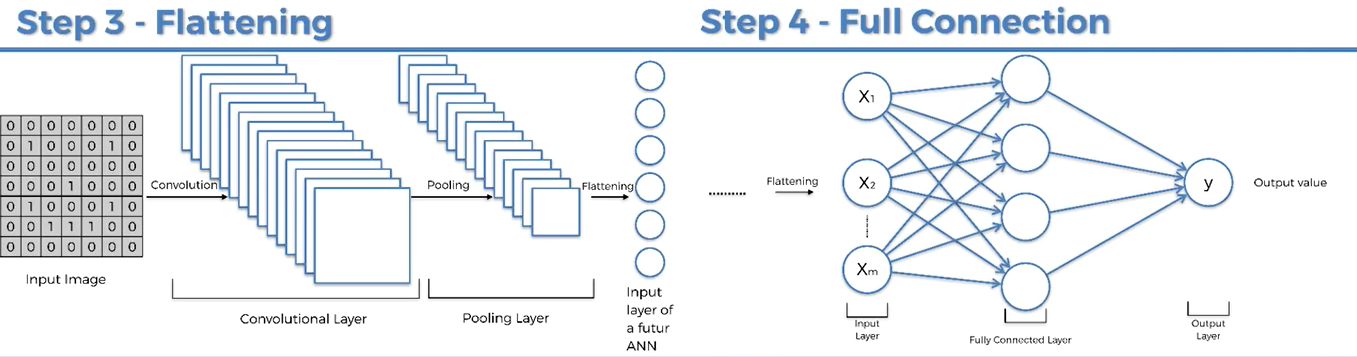


Figure 5.1.1 | CNN

**5.2) Snapshots of project**

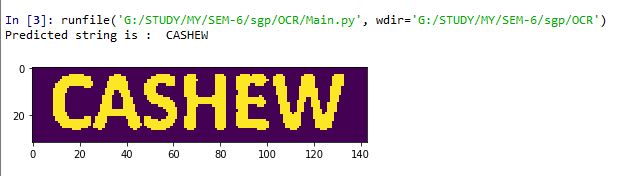
****

Figure 5.2.1 | Predict English Word

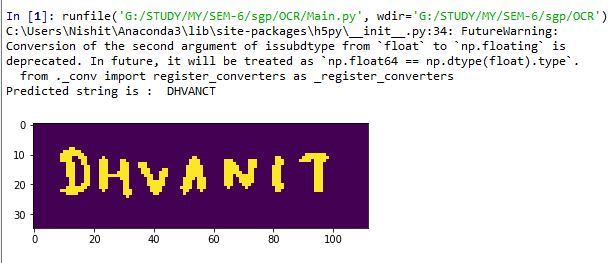
****

Figure 5.2.2 | Predict English Word 2

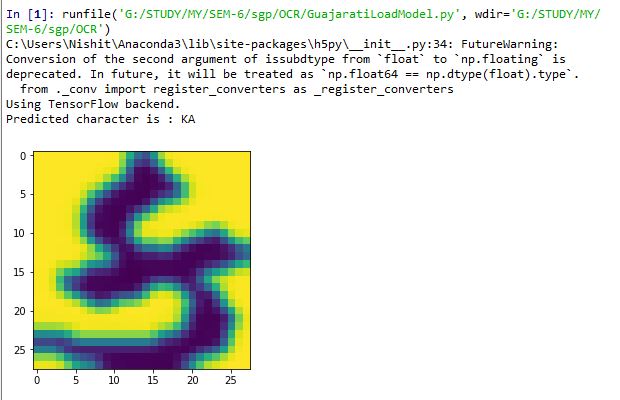
****

Figure 5.2.3 | Predict Gujarati Character

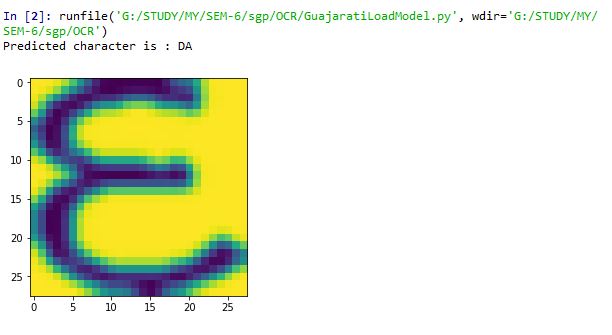
****

Figure 5.2.4 | Predict Gujarati Character-2

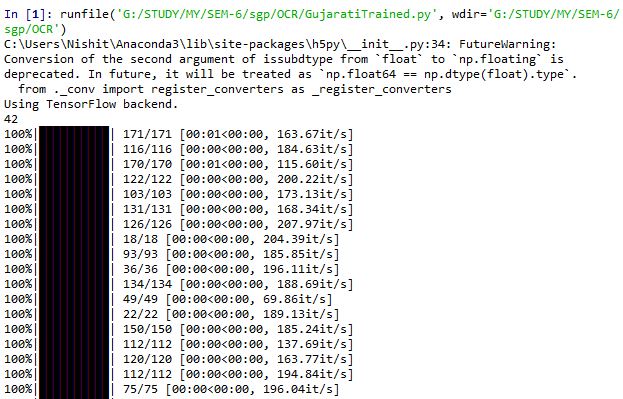
****

Figure 5.2.5 | Loading Character dataset

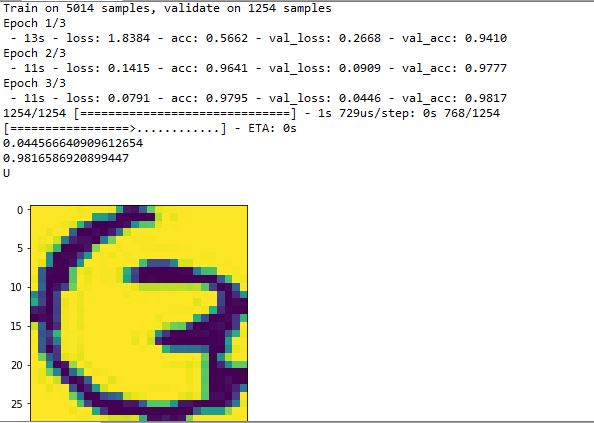
****

Figure 5.2.6 | Model Training

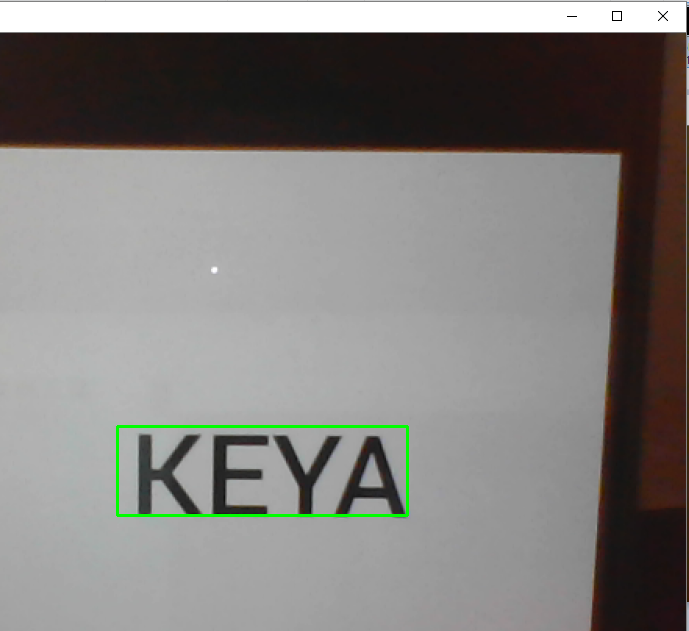
****

Figure 5.2.7 | Text Detection from Webcam

**5.3) Accuracy Comparison**

|  |  |  |
| --- | --- | --- |
|  | English | Gujarati |
| In characters | More than 90% | Around 80% |
| In words | Around 90% | Less than 50% |

Figure 5.3.1 | Accuracy Comparison

**Chapter 6**

**CONSTRAINTS AND FUTURE ENHANCEMENT**

**Constraints:**

* Not properly detecting the characters from live images.
* For noisy images, the model will fail to predict the right output.

**Future enhancement:**

* A proper GUI application can be made for better user experience in android and iOS.
* Increasing the efficiency and segmentation algorithm for live image.
* For noisy images, better noise removal techniques can be used.

**Chapter 7**

**CONCLUSION**

Working on this project has helped us to understand various concepts regarding python programming language and image processing. We also learnt algorithms and techniques of machine learning and deep learning to implement our project. Also, became aware about various high level API’s like TensorFlow, Keras and Opencv.

It was essential for project to understand the requirements in terms of functional requirements, non-functional requirements, platforms, technicality, timelines, information availability. The project was completed within the given time period.

No major problems were encountered and no major risk was involved during development of the project. All the system error has been removed successfully. The program runs successfully.

**References:**

* <https://pythonprogramming.net/>
* <https://www.coursera.org/learn/convolutional-neural-networks/>
* <https://neuralnetworksanddeeplearning.com/chap1.html>
* <https://stackoverflow.com/>
* <https://github.com/dhaval1212/Optical-Character-Recognition>