KATHFORD INTERNATIONAL COLLEGE OF

ENGINEERING AND MANAGEMENT

Balkumari, Lalitpur

A

Minor Project mid Term Report

On

"NUMBER RECOGNITION OF RESULT PAPER USING MACHINE LEARNING ALGORITHM"

[Subject Code: EX654]

Project Members

Bharat Karki (003/BEX/2074)

Niranjan Tamang (007/BEX/2074)

Sagar Shrestha (011/BEX/2074)

DEPARTMENT OF COMPUTER AND ELECTRONICS &

COMMUNICATION ENGINEERING

LALITPUR, NEPAL

FEBRUARY, 2021

KATHFORD INTERNATIONAL COLLEGE OF ENGINEERING AND MANAGEMENT

Balkumari, Lalitpur

(Affiliated to Tribhuvan University)

"NUMBER RECOGNITION OF RESULT PAPER USING MACHINE LEARNING ALGORITHM"

A

PROJECT REPORT

SUBMITTED TO THE DEPARTMENT OF COMPUTER AND ELECTRONICS & COMMUNICATION ENGINEERING IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF BACHELOR IN ELECTRONICS & COMMUNICATION ENGINEERING

Submitted by:

Bharat Karki (2074/BEX/03)

Niranjan Tamang (2074/BEX/07)

Sagar Shrestha (2074/BEX/11)

Submitted to:

DEPARTMENT OF COMPUTER AND ELECTRONICS & COMMUNICATION ENGINEERING

FEBRUARY, 2021

COPYRIGHT

No part of this report can be copied without the permission of the authors. However, the authors have agreed that the Library, Department of Computer and Electronics & Communication, Kathford International College of Engineering and Management may take this project freely available for inspection. Moreover, authors have agreed that the permission for extensive copying of this project for scholarly purpose may be granted by the Professor(s), who supervised the project. It is understood that recognition will be given to the authors and to the department in any use of the parts of this project. Copying of publication or other use of this project for financial gain without approval of the department and author's written permission is prohibited.

Request for permission to copy or to make any use of the material in this project in whole or part should be addressed to:

Head of the Department
Department of Computer and Electronics & Communication Engineering Kathford International College of Engineering and Management

Balkumari, Lalitpur,
Nepal

TRIBHUWAN UNIVERSITY
KATHFORD INTERNATIONAL COLLEGE OF ENGINEERING AND MANAGEMENT
DEPARTMENT OF COMPUTER AND ELECTRONICS & COMMUNICATION ENGINEERING

Letter of Approval

The undersigned certify that they have read and recommended to the Department of Computer and Electronics & Communication Engineering for acceptance, a project entitled **"NUMBER RECOGNITION OF RESULT PAPER USING MACHINE LEARNING ALGORITHM"**, submitted by **Bharat Karki, Niranjan Tamang** and **Sagar Shrestha** in partial fulfillment of the requirement for the minor project of **"Bachelor in Electronics and Communication Engineering"**.

External Examiner:

……

Project Coordinator:

…….

Er. Saban Kumar K.C.

Lecturer

Department of Computer and Electronics Engineering

Kathford International College of Engineering and Management

FEBRUARY, 2021

Balkumari, Lalitpur-8, Nepal
Tel.: +977-1-5201241/5201911
Fax: +977-1-5201899
G.P.O.Box: 23483, Kathmandu
E-mail: info@kathford.edu.np
Web: www.kathford.edu.np

DEPARTMENTAL ACCEPTANCE

The project entitled "**NUMBER RECOGNITION OF RESULT PAPER USING MACHINE LEARNING ALGORITHM**", submitted by ***Bharat Karki, Niranjan Tamang*** *and* ***Sagar Shrestha*** in partial fulfillment of the requirement for the minor project in **"*Bachelor in Electronics and Communication Engineering*"** has been accepted as a bonfire record of work carried out by them in the department.

…….

Er. Prabin Kumar Jha
Deputy Head of the Department
Department of Computer and Electronics & Communication Engineering Kathford International College of Engineering and Management

Balkumari, Lalitpur
Nepal

FEBRUARY, 2021

ACKNOWLEDGEMENT

We would like to acknowledge the guidance and constructive suggestions provided by the teachers and individuals in course of the preparation of this report. We would like to express our sincere thanks to **Er. Saban Kumar K.C**, Project Coordinator for his assistance and help during the process.

We are also obliged to all the individuals and groups who have directly or indirectly assisted in the preparation of this report, especially **Er. Bikal Adhikari** and **Er. Kiran Bagale** for providing the comments and assistance in preparing the report. We extend our gratitude towards DHOD of Department of Electronics and Communication of Kathford International College of Engineering and Management **Er. Prabin Kumar Jha** for providing us the opportunity and the required environment, support and help to do this project.

We would like to convey our thanks to all teaching and non-teaching staff members of Department of Computer and Electronics & Communication Department of Kathford Engineering College of Engineering and Management for their invaluable support and help throughout the project period.

Also, we would like to express our deep sense of gratitude to our seniors **Mr. Saman Ratna Buddhacharya** and **Mr. Santosh Adhikari**, our friend **Mr. Abiral Bhandari, all** our beloved friends and family members who tried to encourage us while doing this project with their valuable suggestions to approach our goals. The entire team is thankful to all the peoples who have either directly or indirectly helped for the successful completion of the project.

Lastly, we would humbly welcome any suggestions and guidance for the improvement of this report.

Thank You.

ABSTRACT

With the development of technology, we would like to save time. Another thing to notice is that the storage of proper information is crucial. So, to create an efficient way of managing text and numbers, we have used OCR using machine learning algorithms for text and number conversion. The aim of this project is to create a numbers Recognition system that converts numbers into digital format and makes a mark sheet. The main focus of this project is to decrease the delay time that was required for making the mark sheet of student's exam results which directly or indirectly helps in the student's academics. Unlike taking a photo, the symbol number, marks, the subject name will be stored in a database which will be then used to make the mark sheet in portable document format (pdf). The main motive for working on this project is to create save time and reduce repetitive work. With the help of a web application, the user can upload the photo containing the numbers. The image that is fed by the user undergoes image processing techniques and is fed into the Convolution Neural Network model. Afterward, produce the numbers and the scanned number is saved into the database. Later, the saved mark is fetched from the database is used to generate the mark sheet of the students in pdf and stored in a specific directory saved with the filename of the id of the student. Furthermore, the final converted format is displayed on the User Interface (UI) of the web application and it can be seen by the user as well. This application allows the academics section of Nepal to spend less time in making the Mark sheet instead of taking a huge amount of time in making the mark sheet. This application will move the academic section of Nepal towards the digital world. In terms of market use, it can be used by all the academic sections such as the SLC board, NEB, universities, colleges, and other academic institutions.

Keywords: Number Recognition, Digitalization, Machine Learning algorithms, Optical Character Recognition

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

JPEG Joint Photographic Experts Group

ML Machine Learning

OCR Optical Character Recognition

PNG Portable Network Graphics

SVM Support Vector Machine

1 INTRODUCTION

1.1 Background

OCR is a technology that recognizes the text and number contain in the digital image. OCR can be used to convert a physical paper document or an image into an accessible electronic version with text. OCR program will recognize the text and convert the document to an editable text file. It is used in data entry for paper documents like passport, business cards, receipt and so on which can be stored compactly in a digital format that can later be edited which helps in document management. This technology is innovative since it saves time without the need to retype the document.

A digit is a single symbol used to make numerals. 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9 are the ten digits which we use in everyday numerals. For example, the numeral 153 is made up of 3 digits ("1", "5" and "3"). The main goal of this project is to convert mark sheets paper which consists of handwritten marks, symbol numbers in a digital format that can be stored in a database and can be used to create result paper in pdf. For the OCR program to work well we need to perform two things. They are:

1.1.1 Image Preprocessing

Image preprocessing is a crucial part of the OCR program. After performing image preprocessing, it can be fed to the CNN model and produce a more accurate prediction of digits. Image preprocessing tasks involve various image processing techniques such as binarization, gray scaling, line segmentation, thresholding, resizing character segmentation, local and non-local filtering, skew correction, color detection, dilation, and erosion and so on. Since handwritten numerals are tightly packed, it leads to difficulty in character segmentation. This problem can be resolved using the erosion function to thin the lines that help to pull apart packed characters. Since our train data is a binary image we use a technique called binarization. Before that, since most of the image is color we perform the gray scaling of the image.

1.1 .2 Character Recognition

A preprocessed image is fed as input to the CNN model and produces a more accurate prediction of digits. CNN model is a multilayer neural network model. Our CNN model takes the input image of size 28*28. The input image to the CNN model is a binary image. As the name of Convolution Neural Network consists of the word

Convolution, it performs the convolution operation of the image with a filter. In our model, we use the filter of size 3*3. Furthermore, in this model, we use the max-pooling with the filter of size 2*2. After passing through the series of layers, it is capable of predicting whether the image contains 0 or any other digits.

1.2 Problem Statement

Optical Character Recognition would be very much appropriate for the characters and number recognition of result paper with much faster and efficient outcome. As compared to the traditional methods of recording alphanumeric character, typing is done manually. Handwritten alphanumeric characters may be difficult to be recognized by recorder and multiple times checking of numbers is required which requires more time. So, OCR is more favorable for scanning marks and symbol numbers and storing it in database in mark sheet format which is more efficient and accurate compared to tradition methods. Besides that, there are also some problems seen in the OCR projects done earlier. OCR is still in the field of research area. Sometimes there may be delay in academic mark sheets of national level examinations which may result in many issues and problems such as students may not be able to get admission in school and colleges on time and there will be gap in their academic progression which will hamper students' academic study and career.

1.3 Objectives

1.3.1 Main Objective

- To design and develop a system for number recognition of result papers using machine learning and generate the mark sheet of students in pdf format.

1.3.2 Specific Objectives

- To capture the numeric characters.
- To implement a Convolution Neural network to recognize numerals in images.

- To train the system using the Mnist dataset.
- To store the detected data (symbol number and marks) into the database and create the mark sheet.

1.4 Scopes and Application

These are the fields where this project can be used as listed below:

I. This project can be extensively used in Academics.
II. It can play a significant role in making Academics modernized and easy.

2 Literature Review

Rachis Adhvaryu [1] described the OCR. According to him, OCR is the process of classification of optical patterns contained in a digital image corresponding to alphanumeric characters. OCR Technology allows us to convert scanned documents, pdf files, and images from digital camera to editable and readable form. It is also described as the different techniques used for speech recognition, optical mark reading, online character recognition, and offline character recognition, etc.

M. Jordan [2] described in this book the concept of recognition system and machine learning helpful for the different types of modeling systems. Richard O. Duda [3] described how to implement the modeling system on different types of design cycles to help design algorithms.

Computer vision is the science and technology of machines that machines are capable of extracting information from information such as numbers, text, and many more to solve some task. Alphanumeric recognition systems can recognize almost all the letters and numbers correctly and their respective locations as they appear in the image. In general, alphanumeric recognition is classified into two types as off-line and online handwriting recognition methods. In the off-line recognition takes the raster image from the scanner (scanned images), digital camera. The image is binarized based on, for instance, color pattern (color or grayscale) so that the image pixel is either 1 or 0. But, in the case of on-line character recognition, the current information is presented to the system, and recognition is carried out at the same time. Basically, it accepts the string of (x, y) coordinate pairs from an electronic pen touching a pressure-sensitive digital tablet. Adaptive Template matching and Feature Extraction using curvelet transform can be used to recognize the alphanumeric character [4].

Text characters are often in different scales in the documents to give an important and also the character could also appear in different orientations other than the usual horizontal and vertical direction. Partha [5] used the convex Hull-based approach for multi-oriented character recognition from the graphical documents. A Support Vector Machine (SVM) classifier has been used for recognition purposes.

Alphanumeric recognition results vary a bit from different types of classifiers. Classifier like Projection distance (PD), Subspace method (SM), Linear discriminant function (LDF), Support vector machines (SVM), Modified quadratic discriminant

function (MQDF), Mirror image learning (MIL), Euclidean distance (ED), Nearest neighbor, k-Nearest neighbor (k-NN), Modified Projection distance (MPD), Compound projection distance (CPD), and Compound modified quadratic discriminant function (CMQDF) are there [6]. For recognition purposes, four sets of features are there (two are from binary and two are from gray-scale images). Most of the time we use the binary dataset, but our data set is greyscale, to get the feature on the binary images we convert the gray-scale image into binary using the Otsu method.

Sometimes the photo we click may be blurring artifacts, varying illuminations, changing background due to which some of the recognition systems may not work properly, so this problem can be efficiently solved by using stroke width transform [7]. At first, pre-processing of the image is done to remove blurring artifacts. Then the significant edges in the image are detected and a new image is formed by grouping the connecting rays.

The term Deep Learning or Deep Neural Network refers to Artificial Neural Networks (ANN) with multilayers. Over the last few decades, it has been considered to be one of the most powerful tools and has become very popular in the literature as it is able to handle a huge amount of data. The interest in having deeper hidden layers has recently begun to surpass classical methods performance in different fields; especially in pattern recognition. One of the most popular deep neural networks is the Convolutional Neural Network (CNN) [8].

Reading text from natural images is a hard computer vision task. Deep convolutional neural networks can be used to recognize numbers in natural scene images. Convolution Neural Network (CNN) requires fixed dimensional input while number images contain unknown amount of digits [9].

3 Project Methodology

3.1 Block Diagram

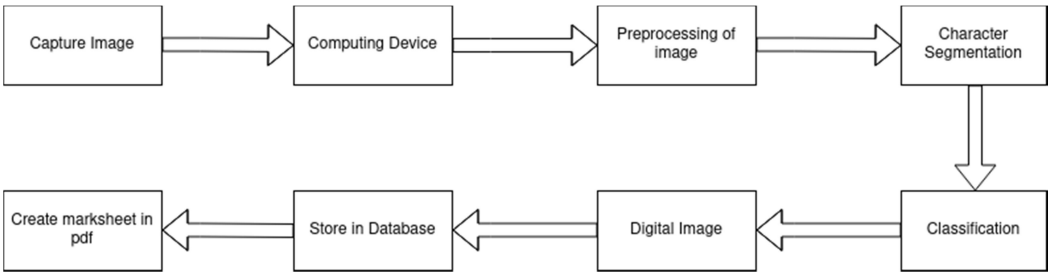


Figure 3-1: Block diagram

The design of number recognition of result paper using Computing Device is show in above figure 1. Our Computing device operates on Linux operating system. Linux operating system is open-source software. Camera is used to capture the image. Some preprocessing of image such as skew correction, conversion of BGR image into the grayscale and further into the binary image is done. For all the preprocessing steps, OpenCV-Python libraries are used.

3.1.1 Input Image Text

The input image was formed of pixels. The image was a multidimensional array, the image was converted to 2D array called grayscale image and a 3D array if the image was colored. Each row and columns form pixels.

3.1.2 Skew correction

In this process, the position was corrected and letters were aligned properly so that it will be easier for the character classification. First, the image was projected horizontally to get a histogram of pixels along with the height of the image. The image was rotated at various angles and the difference between the peaks was calculated. The angle at which the maximum difference between peaks was found and the skewness of the image was corrected through that angle.

3.1.3 Gray Scaling

Gray Scaling is used to save a lot of computational power by reducing a lot of information that is not required. Before getting the binary image, we perform the gray scaling of the image.

3.1.4 Thresholding

In Binarization, a colored image was converted into binary image in which each pixel is represented by 1bit only. For this as first, the threshold value was determined but normally, threshold =150. If the current pixel was greater than the threshold value then its value was set to 255 i.e. white pixels else 0 i.e. black pixel.

3.1.5 Line segmentation

Line segmentation was done using the global horizontal projection method. Using the histogram, we found the rows containing no white pixels. In this way we can segment the line containing the numerals.

3.1.6 Character Segmentation

Character Segmentation was done by using the column histogram. When going through the column histogram, the column where there were no values, they were taken as the space between the Characters. This way the characters were distinguished from each other. Finally, the characters were fed to the CNN model.

3.1.7 Character Classification

Segmented character was then fed into CNN model. A simple CNN model with four layer neural network was used. The model was trained using about 34,826 parameters.

3.1.8 Digital Text

The recognized characters were sent to the view of the web application and displayed to the Users.

3.1.9 Database and pdf

The recognized number was marks that were saved into the SQLite database. Afterward, data of specific ID was obtained. The obtained data was used to create the mark sheet in pdf format.

3.1.10 Artificial neural network

An artificial neural network (ANN) is the piece of a computing system designed to simulate the way the human brain analyzes and processes information. An ANN is based on a collection of connected units or nodes called artificial .An artificial neuron that receives a signal then processes it and can signal neurons connected to it. The "signal" at a connection is a real number, and the output of each neuron is computed by some non-linear function of the sum of its inputs. The connections are called *edges*. Neurons and edges typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection.

3.1.11 Forward propagation

As the name suggests, the input data is fed in the forward direction through the network. Each hidden layer accepts the input data, processes it as per the activation function and passes to the successive layer. In order to generate some output, the input data should be fed in the forward direction only. The data should not flow in reverse direction during output generation otherwise it would form a cycle and the output could never be generated. Such network configurations are known as feed-forward network. The feed-forward network helps in forward propagation.

At each neuron in a hidden or output layer, the processing happens in two steps:

a. **Pre activation:** it is a weighted sum of inputs i.e. the linear transformation of weights w.r.t to inputs available. Based on this aggregated sum and activation function the neuron makes a decision whether to pass this information further or not.

b. **Activation:** the calculated weighted sum of inputs is passed to the activation function. An activation function is a mathematical function which adds non-linearity to the network. There are four commonly used and popular activation functions — sigmoid, hyperbolic tangent (tanh), ReLU and Softmax.

c. **Cost function**: A cost function, $E(X,\theta)$, which defines the error between the desired output and the calculated output of the neural network on input for a set of input-output pairs $X(xi,yi) \in X$ and a particular value of the parameters θ. Cost function is also sometimes called as error function.

d. **Back-propagation:** It is the essence of neural net training. It is the practice of fine-tuning the weights of a neural net based on the error rate (i.e. loss) obtained in the previous epoch (i.e. iteration). Proper tuning of the weights ensures lower error rates, making the model reliable by increasing its generalization. In fitting a neural network, backpropagation computes the gradient of the loss function with respect to the weights of the network

3.1.12 Building the model

We have used the Convolution neural network model to detect the number from the image shown in figure 8. We use keras to build the CNN model. We used the layer 4 neural networks for building our model. We used various optimization techniques such as rms prop, momentum-based optimizer. But among them, we found that using Adam optimization is best. We use the kernel also called filter of the size of 3 by 3 in convolution operation, max pooling of filter size 2 by 2, relu activation and softmax activation function. The number of filters used in the first layer is 32 and in the second layer we use 64 filters.

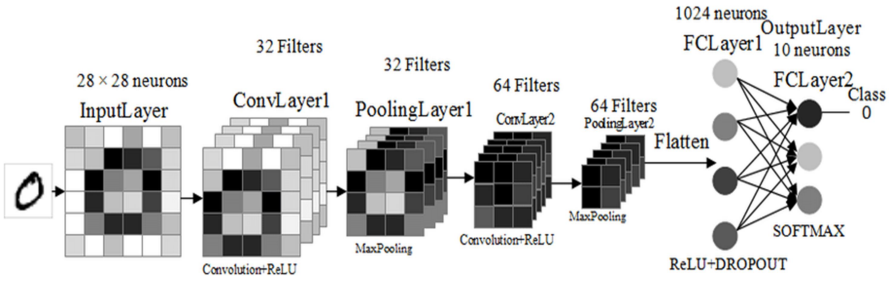


Figure 3-2: Building the Model

At the output layer, we use the SoftMax activation function for classification purposes. Also, convolution operation and max polling were done for feature extraction and from the fully connected layer to the output layer, which we used for classification purposes. By using our model, we get an accuracy of about 98%.

We use the mnist dataset to train the neural network model. We split the dataset into two-part train set which contains about 90% of the dataset and the test set contains 10% of the dataset respectively. The numbers of the parameters in each layer are given below:

Input Layer	784
Layer 1	320
Layer 2	18496
Layer 3	16010
Layer 4	10

Table 1: Number of Parameters in CNN model

3.2 Flow Chart

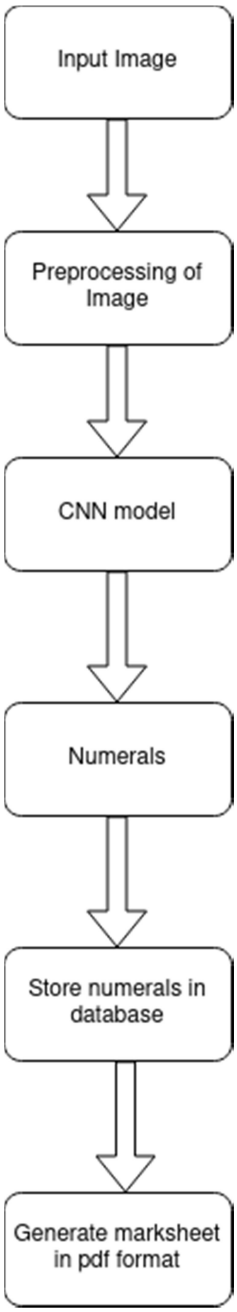


Figure 3-3: Flow Chart

3.3 ER diagram

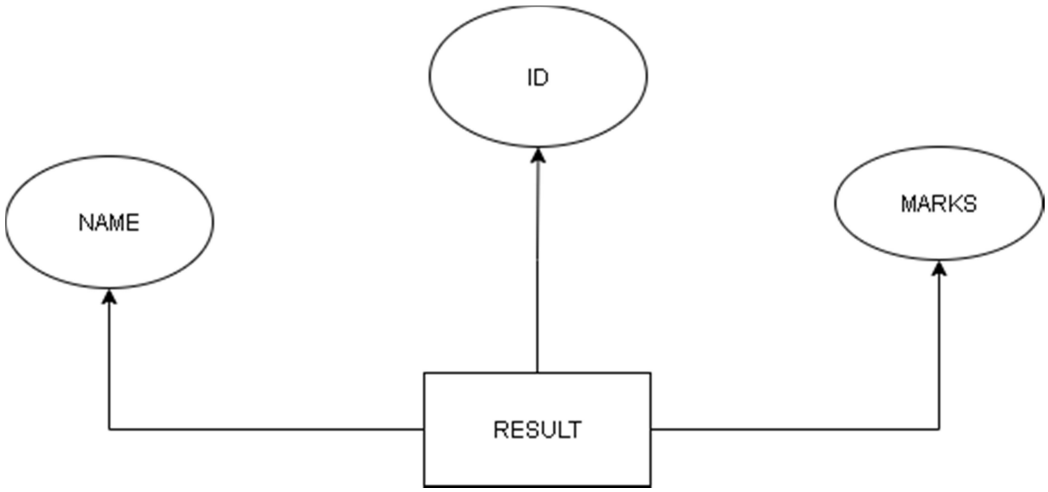


Figure 3-4: ER diagram

ER diagram:

We have used the relational database which consists of table named result. In the result table, ID attribute is used as primary key. Name and marks attributes are used as Secondary Key. With the help of Primary key we can get the information of Specific Students.

4 Tools Requirement

4.1 Hardware Requirement

Following are the hardware requirement:

- Raspberry pi
- PIR sensor
- LED
- Pixy2 camera

4.1.1 Raspberry pi

The Raspberry Pi itself is an embedded computer, or also an SBC (single board computer). The Raspberry Pi is a low cost, credit-card sized computer that plugs into a computer monitor or TV, and uses a standard keyboard and mouse. Main difference between the general computer and raspberry pi is that general computer doesn't contain the gpio pins whereas the raspberry pi contains the gpio pin so that we can add the sensor and can be used to solve the particular problem.

Features and Technical Specifications

- Broadcom BCM2711, Quad core Cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz
- 2GB, 4GB or 8GB LPDDR4-3200 SDRAM (depending on model)
- Gigabit Ethernet
- 2 USB 3.0 ports; 2 USB 2.0 ports.
- Raspberry Pi standard 40 pin GPIO header (fully backwards compatible with previous boards)
- 2 × micro-HDMI ports (up to 4kp60 supported)
- 2-lane MIPI DSI display port
- 2-lane MIPI CSI camera port
- OpenGL ES 3.0 graphics
- Micro-SD card slot for loading operating system and data storage
- 5V DC via USB-C connector (minimum 3A*)
- 5V DC via GPIO header (minimum 3A*)

Figure 4-1: Raspberry pi

4.1.2 PIR sensor

A passive infrared sensor (PIR sensor) is an electronic sensor that measures infrared (IR) light radiating from objects in its field of view. They are most often used in PIR-based motion detectors. PIR sensors are commonly used in security alarms and automatic lighting applications. PIR sensors detect general movement, but do not give information on whom or what moved. For that purpose, an imaging IR sensor is required. PIR sensors are commonly called simply "PIR", or sometimes "PID", for "passive infrared detector". The term passive refers to the fact that PIR devices do not radiate energy for detection purposes. They work entirely by detecting infrared radiation (radiant heat) emitted by or reflected from objects.

Features and Technical Specifications:

Pin Number	Pin Name	Description
1	Vcc	Input voltage is +5V for typical applications. Can range from 4.5V- 12V

2	High/Low Output (Dout)	Digital pulse high (3.3V) when triggered (motion detected) digital low(0V) when idle(no motion detected
3	Ground	Connected to ground of circuit

Table 2: Dataset of PIR Sensor

Figure 4-2: PIR sensor

4.1.3 LED

In the simplest terms, a light-emitting diode (LED) is a semiconductor device that emits light when an electric current is passed through it. Light is produced when the particles that carry the current (known as electrons and holes) combine together within the semiconductor material. Since light is generated within the solid semiconductor material, LEDs are described as solid-state devices.

Features and Technical Specifications:

Forward Current (IF)	30mA
Reverse Voltage	5V
Operating Temperature	-30℃ to +85℃
Storage Temperature	-40℃ to +100℃

Table 3: Dataset of LED

Figure 4-3: LED

4.1.4 Pixy2 camera

Pixy2 camera is smaller, faster and more capable than the original Pixy. Like its predecessor, Pixy2 can learn to detect objects that you teach it, just by pressing a button. And frame rate of Pixy2 is 60 frames-per-second, so work can be fast too.

Features and Technical Specifications:

Processor	NXP LPC4330, 204 MHz, dual core
Lens field-of-view	60 degrees horizontal, 40 degrees vertical
Power consumption	140 mA

Power input	USB input (5V)
RAM	264K bytes
Flash	2M bytes

Table 4: Dataset of Pixy2 camera

Figure 4-4: Pixy2 camera

4.2 Software Requirement

Following are the software requirement:

- OpenCV
- Python
- VScode
- Fpdf
- SQlite3
- CSV
- Numpy
- Mnist Dataset
- HDF5

Descriptions are as follows:

4.2.1. OpenCV

OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in commercial products. OpenCV makes it easy for businesses to utilize and modify the code.

The library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms.

4.2.2 Python

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built-in data structures combined with dynamic typing and dynamic binding; make it very attractive for Rapid Application Development. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Nowadays, python is extensive used to implement the machine learning algorithm. It is also used in raspberry pi.

4.2.3 VScode

Visual Studio Code is a lightweight but powerful source code editor which runs on desktop and is available for Windows, macOS and Linux. It comes with built-in support for python.

4.2.4 Keras

Keras is a deep learning API written in Python, running on top of the machine learning platform TensorFlow. It was developed with a focus on enabling fast experimentation. Being able to go from idea to result as fast as possible is key to good research.

Keras contains numerous implementations of commonly used neural-network building blocks such as layers, objectives, activation functions, optimizers, and a host of tools to make working with image and text data easier to simplify the coding necessary for writing deep neural network code.

4.2.5 Fpdf

Fpdf is a library for PDF document generation under Python. Compared with other PDF libraries, PyFPDF is simple, small and versatile, with advanced capabilities and easy to learn, extend and maintain.

4.2.6 Sqlite3

SQLite is a C library that provides a lightweight disk-based database that doesn't require a separate server process and allows accessing the database using a nonstandard variant of the SQL query language. Some applications can use SQLite for internal data storage.

4.2.7 Numpy

Numpy is a library for the python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical function to operate on these arrays. It provides a powerful N-dimensional array object, sophisticated (broadcasting) functions, useful linear algebra, Fourier transform, and random number capabilities and much more.

4.2.8 Mnist database

The MNIST database (Modified National Institute of Standards and Technology database) is a large database of handwritten digits that is commonly used for training various image processing systems. The database is also widely used for training and testing in the field of machine learning. It was created by "re-mixing" the samples from NIST's original datasets.

4.2.9 HTML

Hypertext Markup Language (HTML) is the standard markup language for documents designed to be displayed in a web browser. It can be assisted by technologies such as Cascading Style Sheets (CSS) and scripting languages such as JavaScript.

Web browsers receive HTML documents from a web server or from local storage and render the documents into multimedia web pages. HTML describes the structure of a web page semantically and originally included cues for the appearance of the document.

4.2.10 CSS

Cascading Style Sheets (CSS) is a style sheet language used for describing the presentation of a document written in a markup language such as HTML. CSS is a cornerstone technology of the World Wide Web, alongside HTML and JavaScript.

CSS is designed to enable the separation of presentation and content, including layout, colors, and fonts. This separation can improve content accessibility, provide more flexibility and control in the specification of presentation characteristics, and enable multiple web pages to share formatting by specifying the relevant CSS in a separate.

4.2.11 Flask

Flask is a micro web framework written in Python. It is classified as a micro framework because it does not require particular tools or libraries. However, Flask supports extensions that can add application features as if they were implemented in Flask itself.

5 RESULTS AND DISCUSSION

We used CNN model for predicting the digits contained in the image which contains the total number of the parameter 34,826. Our CNN model is multilayer neural network. For the first layer, we perform the convolution operation and after that we perform the maxpooling operation of filter size 2*2. In second layer we perform the same operation with same filter size as in first case. Difference in the first and second layer is that number of filter used in first layer is 32 whereas the number of filter used in second is 64. And we get accuracy about 98.5% and test loss of 1.5%. Before feeding image into the CNN model, we have done preprocessing of the input image. Preprocessing step include the gray scaling of the image, thresholding of image. First we performed gray scaling of the image and then we performed the thresholding .Furthermore, skew correction was done to align the letters properly. After performing skew correction, we performed the line segmentation to determine the where character is presented in row. After that we performed the character segmentation.

Segmented character was feed into the CNN model; CNN model takes input image of size 28*28. So, the segmented character is resized into image size of 28*28. So, the resized image of size 28*28 was fed into the CNN model, CNN model predicted what numerals were contained in the image. After predicting numerals, we store the numerals in the database. Stored numerals were fetched to create the mark sheet in pdf of specific id of student.

KATHFORD INTERNATIONAL COLLEGE OF ENGINEERING AND MANAGEMENT

Name :SAMAN

Bachelor's in Engineering

Subjects	Marks_obtained
Math	12.0

Figure 5-1: Mark sheet in pdf format

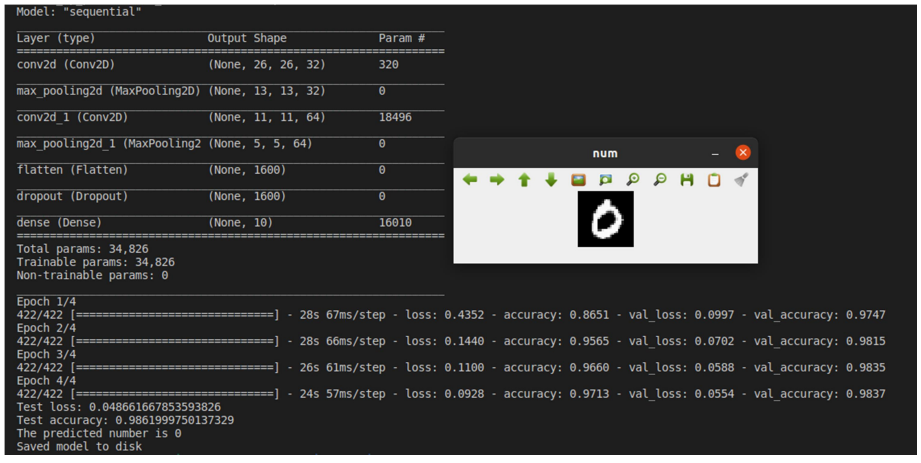


Figure 5-2: Predicting the number with CNN model

6 CONCLUSION

Building convolution neural networks has been done successfully. Training the neural networks using Mnist datasets and testing of model has been done successfully with high of about 98 %. The project has implemented successfully and numerals was recognized and displayed in separate text box in the text format. Artificial neural network can be widely used in various tasks such as image recognition, object detection and many more. From this we got an opportunity to learn about artificial neural network, various image processing technique, relational database, SQLite, keras framework for developing the CNN model efficiently and faster, raspberry pi, to train the CNN model and many more.

7 LIMITATION AND FUTURE ENHANCEMENT

7.1 Limitations

The system has the following limitations:

- Fully automatic system is not yet ready.
- Application can only recognized the numerals.
- Only web based application is there.
- Only able to recognize numerals written in blank white paper.
- ER model don’t consist the whole information of the students.

7.2 Future Enhancements

In future, we might add several features listed below to improve our system.

- Able to recognize the numerals in real result paper.
- Increase in Character segmentation capability.
- Fully automatic system for use.
- Reconstruction of Database, so that it contain the full information of the students.

8. REFERENCES

[1] R. P. a. K. V. R. Adhvaryu, "Design and Implementation of OCR to Identify English Characters and Numbers," *International Journal of Scientific Research in Science, Engineering and Technology,* vol. 4, pp. 57-62, 2018.

[2] J. a. B. S. M.Jordan, PATTERN RECOGNITION AND MACHINE LEARNING, New York, 2006.

[3] E. a. D. G. S. Richard O. Duda, Pattern Classification.

[4] K. S. a. P. Zode, "Implementation of Number Recognition using Adaptive Template Matching and Feature Extraction Method," 2012.

[5] U. P. J. L. a. F. K. P. P. Roy, "Convex hull based approach for multi-oriented character recognition from graphical documents," in *IEEE*, USA, 2008.

[6] T. W. a. F. U.Pal, "Comparative Study of Devnagari Handwritten Character Recognition using different Feature and Classifiers," in *IEEE*, Spain, 2009.

[7] D. S. a. A. P. C. I. Patel, "Automatic Number Plate Recognition System," 2013.

[8] T. A. M. S. Albawi, "Understanding of a Convolutional Neural Network," in *ICET*, Antalya, Turkey, 2017.

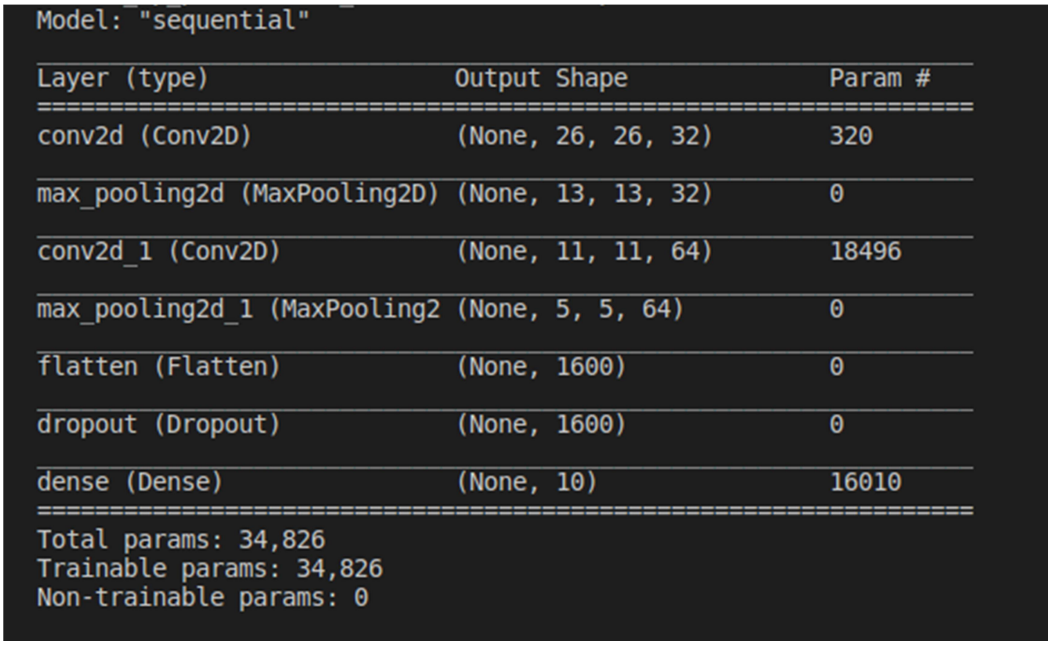
[9] J. L. D. T. a. G. L. Q. Guo, "Reading Numbers in Natural Scene Images with Convolutional Neural Networks," in *IEEE*, Hubei,China, 2014.

APPENDICES

APPENDICES A1 (Snapshots)

Model Summary:

For the model, CNN was used. The summary of the model is given below:



```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 26, 26, 32)        320
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 13, 13, 32)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 11, 11, 64)        18496
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 5, 5, 64)          0
_________________________________________________________________
flatten (Flatten)            (None, 1600)              0
_________________________________________________________________
dropout (Dropout)            (None, 1600)              0
_________________________________________________________________
dense (Dense)                (None, 10)                16010
=================================================================
Total params: 34,826
Trainable params: 34,826
Non-trainable params: 0
```

Figure 0-1: Model Summary

Model Training:

No. of Epochs: 4

Loss = Categorical Cross-Entropy

Batch Size = 128

Optimizer= ADAM Optimization

Test loss: 4.8%

Test accuracy: 98.4%o

Validation Split: 0.1

Max Pulling 2D Filter Size: 2x2

Convolution Filter Size: 3x3

Dropout: 0.5

Activation Function: Relu, Softmax

```
Epoch 1/4
422/422 [==============================] - 53s 125ms/step - loss: 0.4210 - accuracy: 0.8670 - val_loss: 0.0952 - val_acc
uracy: 0.9740
Epoch 2/4
422/422 [==============================] - 50s 118ms/step - loss: 0.1364 - accuracy: 0.9585 - val_loss: 0.0635 - val_acc
uracy: 0.9830
Epoch 3/4
422/422 [==============================] - 50s 118ms/step - loss: 0.1051 - accuracy: 0.9682 - val_loss: 0.0575 - val_acc
uracy: 0.9853
Epoch 4/4
422/422 [==============================] - 57s 135ms/step - loss: 0.0919 - accuracy: 0.9718 - val_loss: 0.0504 - val_acc
uracy: 0.9865
Test loss: 0.04857610538601875
Test accuracy: 0.984000027179718
The predicted number is 0
Saved model to disk
```

Figure 0-2: Model Training

Input Image and binarized image:

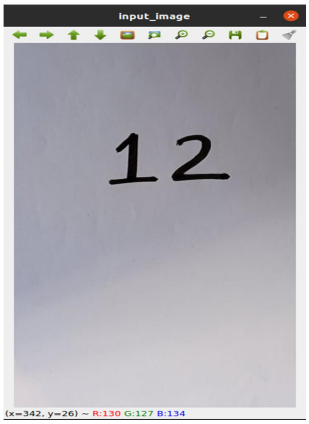


Figure 0-3: Input Image

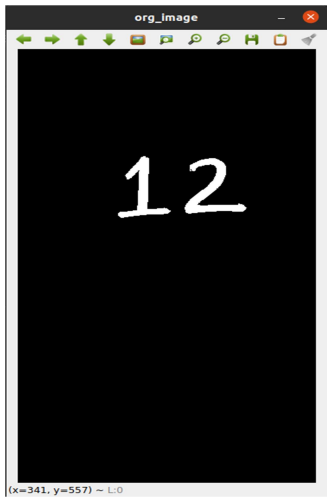


Figure 0-4: Binarized Image

Segmented Image:

Figure 0-5: Segmented Image

Resized Image:

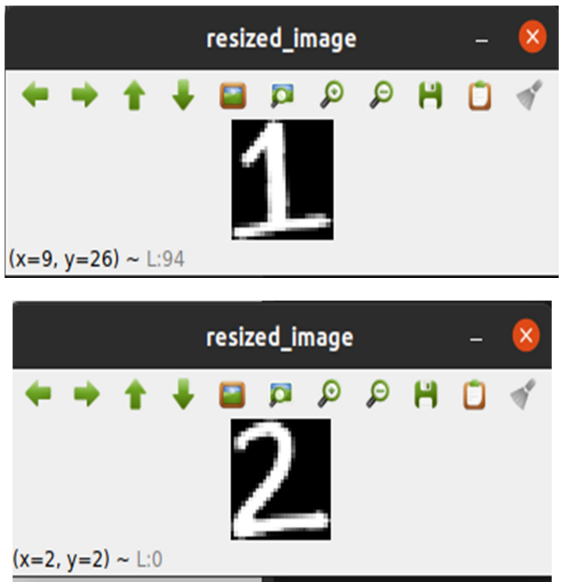
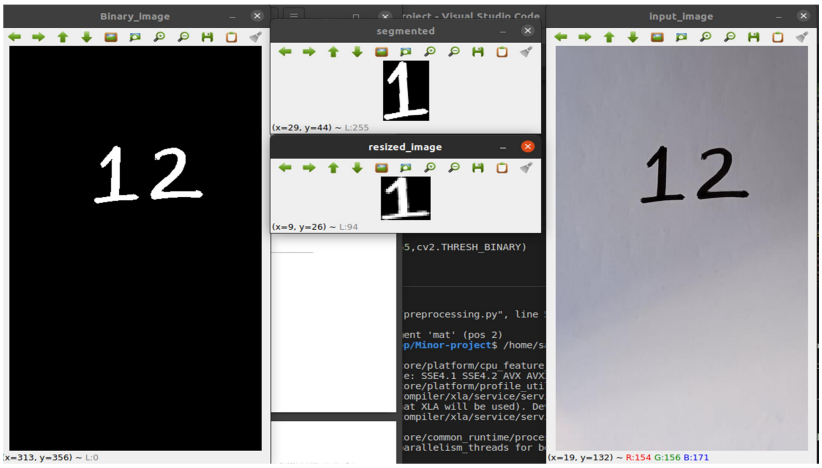


Figure 0-6: Resized Image

Combined Image:



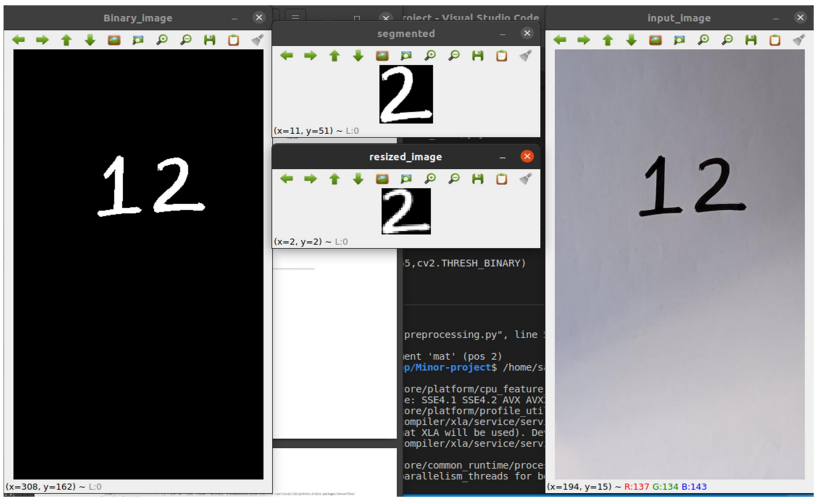


Figure 0-7: Combined Image

Database Browser:

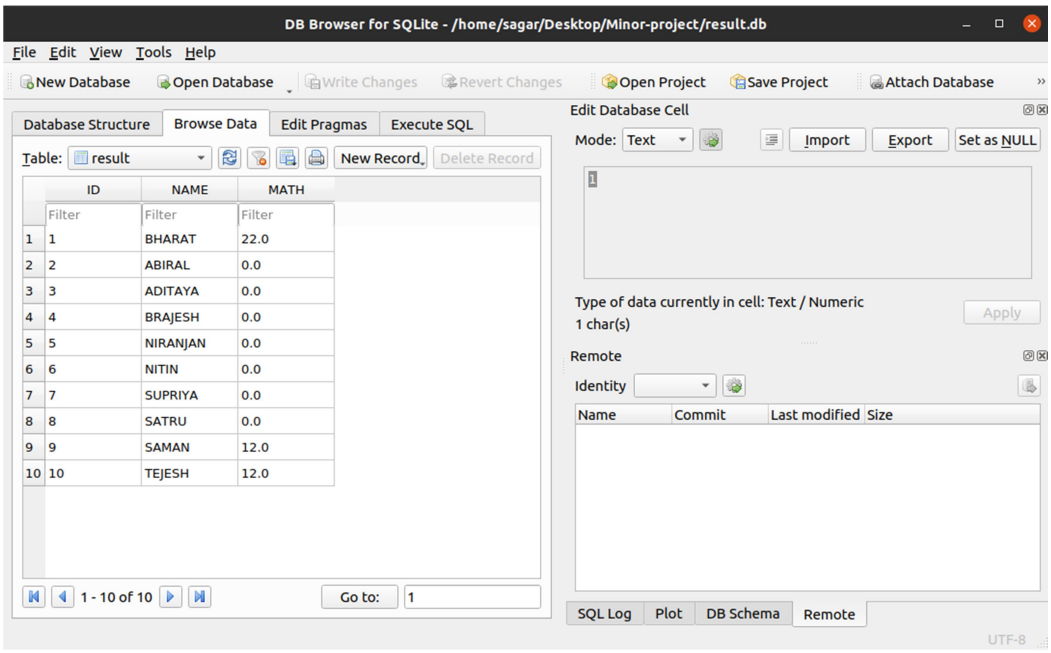


Figure 0-8: DB Browser

APPENDIX A2

1. Cost Estimation:

S.N.	Part of list	Quantity	Price (Rs.)
1.	Raspberry pi with case(4GB)	1	12000
2.	Pi camera (8 megapixel)	2	7000
3.	Male to Male wire	15	150
4.	Female to Female wire	15	250
5.	SD Card (16GB)	2	1800
6.	Universal micro USB power supply	1	150
7.	PIR sensor	2	500
8.	Male to female wire	15	150
9.	Buzzer	2	100
10.	LED	5	25
11.	Ethernet cable	1	300
12.	Card Reader	1	120
13.	Glue Gun	1	300
14.	Glue stick	3	60
15.	HDMI cable	3	600
16.	Miscellaneous		2000
Total			25355

Table 0-1: Cost Estimation