

**KATHFORD INTERNATIONAL COLLEGE OF  
ENGINEERING AND MANAGEMENT**

Balkumari, Lalitpur



A

Minor Project Mid Term Report

On

**“CHARACTERS AND NUMBER RECOGNITION OF RESULT PAPER  
USING SUPPORT VECTOR MACHINE 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**

**DECEMBER 2020**

## **Abstract**

In this world advancing with innovation, everybody likes to spare time and lessen dreary work. Also, 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 algorithm for text and number conversion. The aim of this project is to create numbers and Character Recognition that converts English letters and numbers into digital format and make 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 indirectly helps in the student's academics. Unlike taking a photo, the symbol number, marks, subject name will be stored in database which will be then used to make the mark sheet. The main motive for working on this project is to create a value, save time and reduce repetitive work. Talking about the procedure the text and numbers is scanned which is further noise reduced, skew corrected, gray scaled and other filtering process that makes the scanning process fast and error free. A photo of mark sheet will be captured with the help of pi Camera and raspberry pi microcontroller. Then the scanned text and number is converted to digital format. So, the final converted format (i.e. Text (subject), number (symbol number, marks etc.) can be saved in the database and later the data stored in the database will be used to make mark sheet. This application allows academics section of Nepal to spend less amount of time to make the marks sheet instead of taking huge amount of time in making mark sheet. In terms of market use it can be used by all the academics section such as SLC board, NEB, University, colleges and many more.

**Keywords:** Number and Character Recognition, Pi Camera, Digitalization, Machine Learning algorithms, Optical Character Recognition

## Table of Contents

Abstract.....	i
Table of Contents.....	ii
List of Figures.....	iii
List of Abbreviations.....	iv
1 Introduction.....	1
1.1 Background.....	1
1.2 Problem Statement.....	1
1.3 Objectives.....	2
1.3.1 Main Objective.....	2
1.3.2 Specific Objectives.....	2
2 Literature Review.....	3
3 Project Methodology.....	5
3.1 Block Diagram.....	5
3.2 Flow Chart.....	7
4 Work Completed.....	8
4.1 Working with Images.....	8
4.2 Building a Model.....	11
5 Work in Progress.....	12
6 Conclusion.....	13
References.....	14

## List of Figures

Figure 3.1 Block Diagram.....	5
Figure 3.2 Flow Chart.....	7
Figure 4.1 Skew Image.....	8
Figure 4.2 Bitwise-Not and Thresholding of Image.....	9
Figure 4.3 Skew Corrected Image.....	9
Figure 4.4 Original Image.....	10
Figure 4.5 Image After Masking.....	10
Figure 4.6 Convolutional Neural Network Model.....	11
Figure 4.7 Neural Network Model Parameters.....	11

## **List of Abbreviations**

JPEG Joint Photographic Experts Group

ML Machine Learning

OCR Optical Character Recognition

PNG Portable Network Graphics

CNN Convolution Neural Network

# **1 Introduction**

## **1.1 Background**

OCR is a conversion of handwritten or any text into machine encoded text. It is a technology that enables us to convert different types of documents, such as scanned paper documents, PDF files or images captured by a digital camera or phone into editable and searchable data. It can recognize text in images and convert it into editable text by going through a simplified process. This conversion can be electronic or mechanical. 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 very useful since it saves time without the need of retyping the document. It can perform the actions in a few minutes.

The modern English alphabet is a Latin alphabet consisting of 26 letters, each having an upper and lower-case form. A character is any mark or symbol that can appear in writing. A letter is a character that is part of an alphabet. Basically, a character that represents a sound in the language and that can be combined with other characters to form words. Number is a mathematical object used to count, measure, and label. Number can be used to calculate the marks in exam paper.

The main goal of this project is to convert mark sheets paper which consists of handwritten marks, symbol number, subject into text and number in a digital format which can be stored in database and can be used to create result paper.

## **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 characters and number recognition of result paper using machine learning.

#### **1.3.2 Specific Objectives**

- i. To capture the alphanumeric characters using pi camera
- ii. To implement Support Vector Machine (SVM) algorithm to recognize characters and numbers on the raspberry pi.
- iii. To train the system using various data sets.
- iv. To store the detected data (symbol number and marks) into database and create the mark sheet.

## 2 Literature Review

Rachit Adhvaryu [1] described about the OCR. According to him OCR is the process of classification of optical pattern 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 technique used for speech recognition, optical mark reading, online character recognition and offline character recognition etc.

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

Computer vision is the science and technology of machines that machine is capable to extract the information from the information such as numbers, text and many more to solve some task. Alphanumeric recognition system 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 gray scale) so that the image pixel is either 1 or 0. But, in 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 scale in the documents to give an importance and also the character could also appear in different orientation 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 purpose.

Alphanumeric recognition results vary a bit from different types of classifiers. Classifier like Projection distance (PD), Subspace method (SM), Linear discriminant

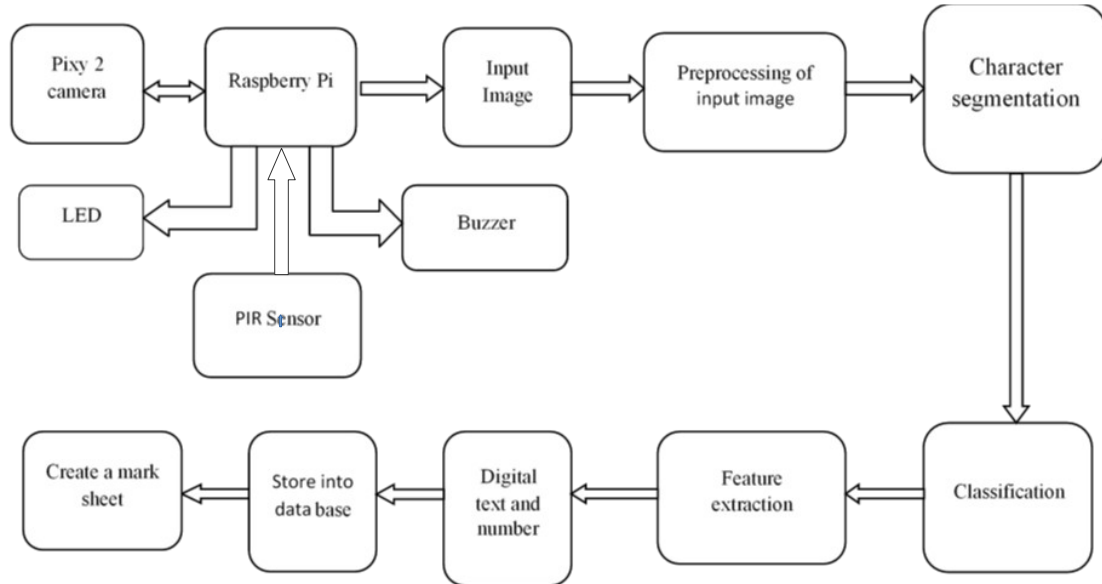


function (LDF), Support vector machines (SVM), Modified quadratic discriminant function (MQDF), Mirror image learning (MIL), Euclidean distance (ED), Nearest neighbor, k-Nearest neighbor (k-NN), Modified Projection distance (MPD), Compound projection distance (CPD), and Compound modified quadratic discriminant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

Sometimes photo we click may be blurring artifacts, varying illuminations, changing background due to which some of recognition system may not work properly, so this above problem can be efficiently solve 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 new image is formed by grouping the connecting rays.

### 3 Project Methodology

#### 3.1 Block Diagram



**Figure 3.1 Block Diagram**

The design of characters and number recognition of result paper using Raspberry pi as micro-controller is show in above figure 1. Raspberry pi is micro-controller which is specially designed for implementing image processing and deep learning model. It operates on Linux operating system. Linux operating system is open-source software. PIR sensor is chosen to detect the motion, if there is motion then it sends signal to micro-controller that there is motion and not to capture the photo using the pi camera. If, not then it captures the images and save it in memory of micro-controller. And if there is no motion for about 1 minutes then, makes sound with the help of buzzer and lights up LED. The image captured by the micro-controller containing paper is further processed by micro-controller to detect numbers. Some preprocessing of image such as skew correction, conversion of BGR image into the grayscale and further into the binary image is done. In preprocessing steps, we also do color-detection to detect the red part of the image. For all the preprocessing steps, OpenCV-Python libraries are used. OpenCV-Python is a library of Python bindings designed to solve computer vision problems.

For building our neural networks model, keras libraries are used. Keras is an open-source software library that provides a Python interface for artificial neural networks. Convolution neural networks(CNN) are built because it's performs best on problems

such as image classification. Using keras, mist dataset is downloaded and used to train as well as to test the model. At the output layer we use SoftMax activation function to classify the images. Both the convolution operation and maxpooling operation is done to extract the useful information from the image. For convolution operation, filter of size 3 by 3 is used. After testing the model, individual image is feed into model to predict the number contain in image using the learned parameters. After layer 2, each pixel in image is flattened and dropout is done to prevent from overfitting. The neurons in third layer is feed to output layer which uses the SoftMax activation function.

Number of Parameters in each layer of the convolution neural networks:

input layer	784 unit
layer 1	320
layer 2	18496
layer 3	16010
layer 4	10

### 3.2 Flow Chart

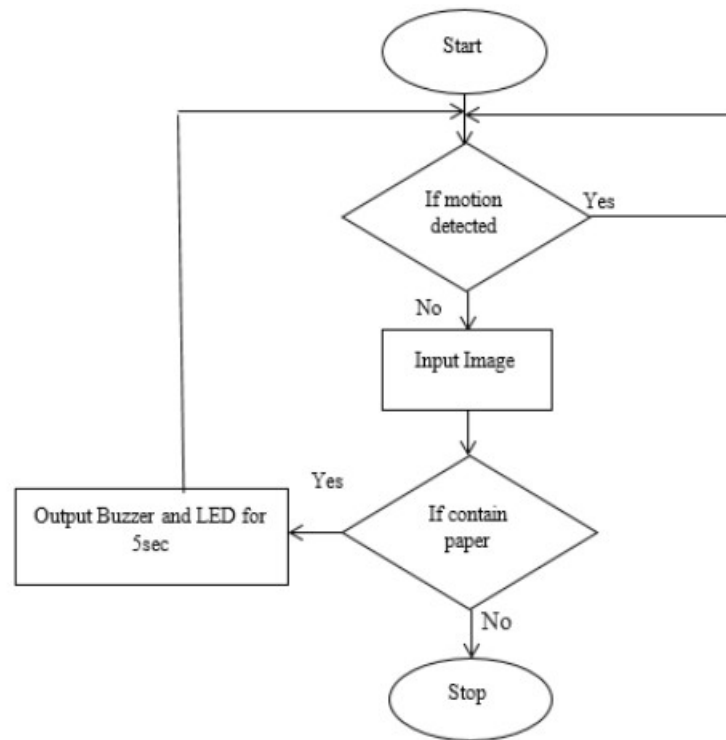


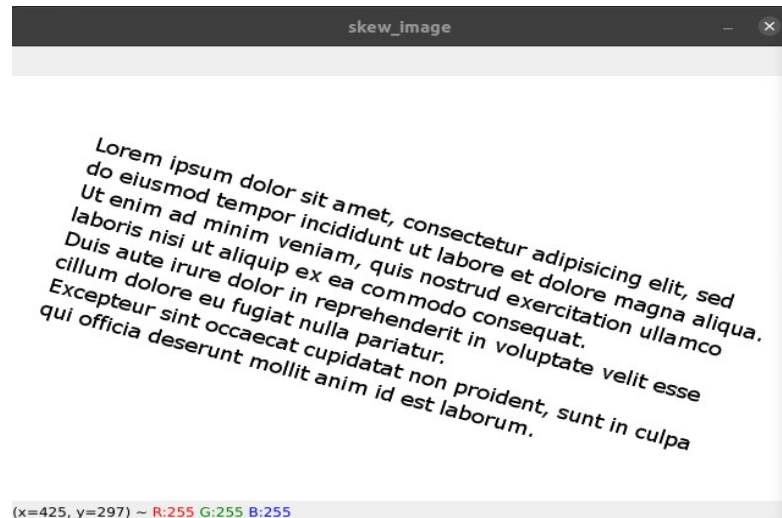
Figure 3.2 Flow Chart

## 4 Work Completed

**Image Preprocessing:** The best type of image that works best for the OCR model is a binary image. In the image preprocessing step, we do some conversion of the image from RGB to gray-scale image and use thresholding to binarize the image and do some skew correction step before feeding into our model. It will help in significant improvement in result in recognizing the digits. By keeping in mind that our model will tend to work best for a binary image, but our camera captures the RGB image, so we convert it gray-scale image. We use a different library for working with images such as NumPy, open-cv, and many more.

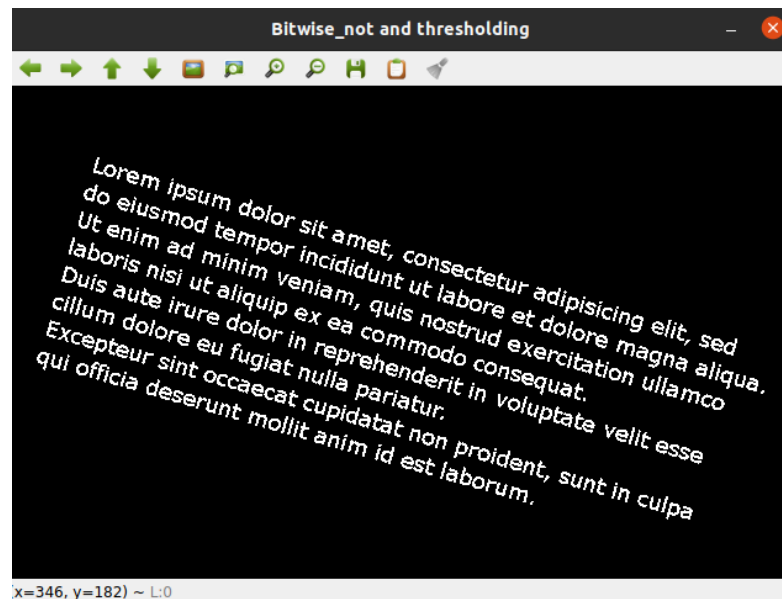
### 4.1 Working with Images

**Skew-correction:** Before feeding the picture into the model, the captured image might be skewed by some angle with a negative or positive slope. So, to convert the skew image into a deskew image, we first convert the BGR image into the gray-scale image, then by using the technique of thresholding we binarize the image. All the algorithm needs for us to convert image was already defined in the OpenCV libraries. So, by just calling a few function we were able to convert images into a binary image. Afterward, we pass that binary image into the function which gives the coordinates of edges in an image which helps us to determine by what angle it was rotated so that we can deskew the image. After finding by what angle it is to be rotated, we use affine transformation to deskew the image.



**Figure 4.3 Skew Image**

Figure3.3 shows the skew image which we must deskew the image. For this, we have applied bitwise operation and thresholding in the image and get the image shown below in figure3.4.



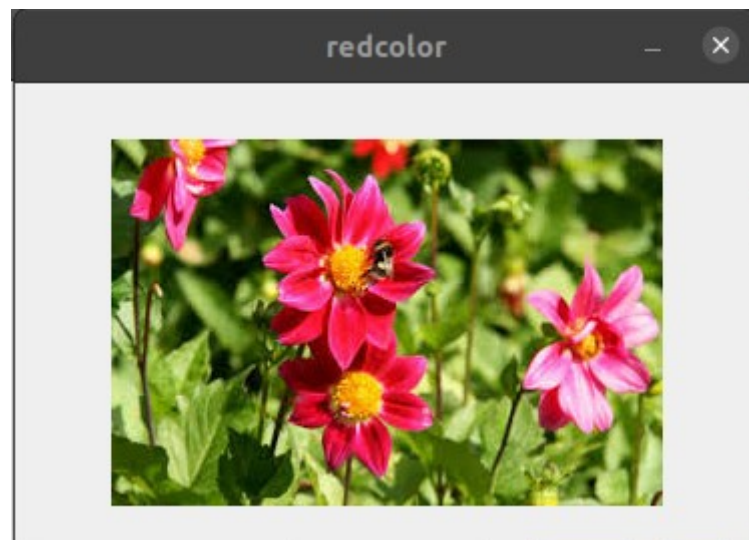
**Figure 4.4 Bitwise-Not and Thresholding of Image**

Afterward, we calculate the angle by which the image is skewed or rotated. The angle we calculate was used to rotate the original image. We rotated the image and got an image that was deskewed. The image that we got by rotating the original image by that angle is shown in figure3.4 which is a deskew image which is our first step in image preprocessing.



**Figure 4.5 Skew Corrected Image**

**Color-Detection:** In most mark sheets, marks are written with a red pen so to extract the digits written with a red pen we used a masking technique that only extract the red part of the image. So, we extract only the red part of the image and apply the gray-scaling in that image so that it can be feed into our model for the detection of the number and store it in excel format. The figure 6 below represents the original image from which we want to extract only the red part of the image using the masking. The figure 7, shows the result after performing the masking and converting it into the gray-scale image.



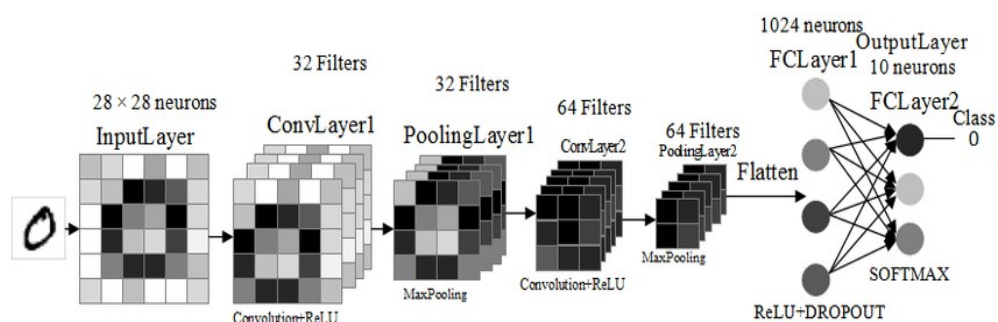
**Figure 4.6 Original Image**



**Figure 4.7 Image After Masking**

## 4.2 Building a Model

We have used Convolution neural network model to detect the number from the image shown in figure 8. We used Keras libraries which are built on top of TensorFlow libraries for making our neural network 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, I found myself using Adam optimization is best. We use the kernel also called filter of the size of 3 by 3, max pooling of 2 by 2, and relu activation function.

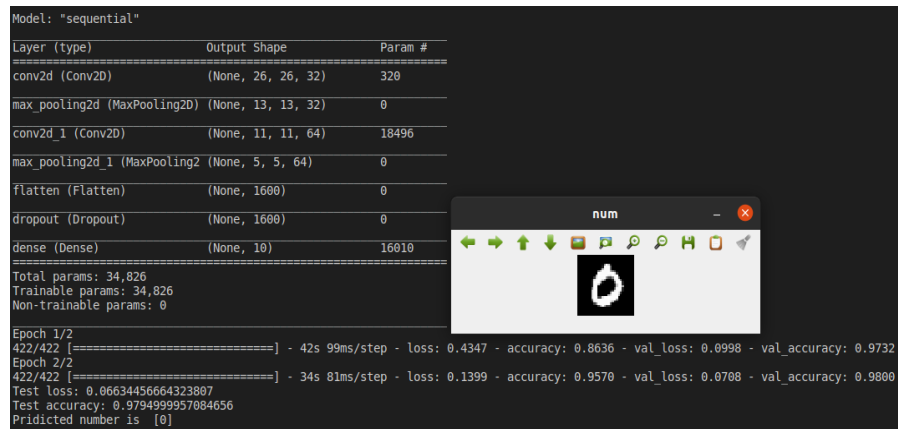


**Figure 4.8 Convolutional Neural Network Model**

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 three-part trainset which contains about 90% of the dataset and the test and dev set contains 5% of the dataset respectively. All the information about our model can be get from the figure 9 below.



**Figure 4.9 Neural Network Model Parameters**

## **5 Work in Progress**

- Image segmentation
- Saving the recognized image into excel file
- Saving the learned parameter from model.
- Finalizing the code in python file by importing another python file
- Writing the docstring of different functions

## **6 Conclusion**

- Images preprocessing has been done successfully. Skew of correction of the skew image has been done successfully.
- Extraction of red part of the image has been done successfully using the masking of image.
- 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 accuracy of about 98 %.

## References

- [1] R. Adhvaryu, R. Parikh and K. Vora, "Design and Implementation of OCR to identify English Characters and Numbers," *International Journal of Scientific Research in Science, Engineering and Technology*, vol. 4, no. 2, pp. 57-62, 2018.
- [2] M. Jordan, J. Kleinberg and B. Scholkopf, *PATTERN RECOGNITION AND MACHINE LEARNING*, New York, 2006.
- [3] R. O.Duda, P. E.Hart and D. G.Stork, "Pattern Classification".
- [4] K. S.Machhale and P. Zode, "Implementation of Number Recognition using Adaptive Template Matching and Feature Extraction Method," 2012.
- [5] U. P. J. L. F. K. P. P. Roy, "Convex Hull based Approach for Multi-Oriented Character Recognition from Graphical Documents," IEEE, 2008.
- [6] T. W. and F. K. U. Pal, "Comparative Study of Devnagari Handwritten Character Recognition using Different Feature and Classifiers," 2009.
- [7] J. C. A. P. J. K.V, "Automatic Number Plate Recognition system," 2013.
- [8] S. ALBAWI and T. A. MOHAMMED, "Understanding of a Convolutional Neural Network," *ICET*, 2017.