HANDWRITTEN WORD DETECTION BASED ON MACHINE LEARNING APPROACH

BY

MD. SABBIR HASAN

ID: 183-15-2244

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Mushfiqur Rahman
Sr. Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Mohammad Jahangir Alam
Sr. Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

FEBRUARY 2023

APPROVAL

This Project titled "**HANDWRITTEN WORD DETECTION BASED ON MACHINE LEARNING APPROACH**", submitted by **Md. Sabbir Hasan, ID: 183-15-2244** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on February 2023.

BOARD OF EXAMINERS

6/2/23

Dr. Touhid Bhuiyan — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Nazmun Nessa Moon — **Internal Examiner**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Raja Tariqul Hasan Tusher — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Ahmed Wasif Reza — **External Examiner**
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Mushfiqur Rahman, Sr. Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Mushfiqur Rahman
Sr. Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

02.02.23

Mohammad Jahangir Alam
Sr. Lecturer
Department of CSE
Daffodil International University

Submitted by:

Sabbir

Md. Sabbir Hasan
ID: 183-15-2244
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First I express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

I really grateful and wish our profound our indebtedness to **Mushfiqur Rahman**, **Sr. Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of Machine Learning to carry out this project. His endless patience, scholarly guidance, continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

I would like to express our heartiest gratitude to **Professor Dr. Touhid Bhuiyan** Head of the Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

I would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, I must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

This thesis presents "Handwritten word detection based on machine learning approach". Handwritten focus is turning into a necessary problem in a number of years however it turns into an assignment to get suitable overall performance due to the alignment and many of them are similar. The mapping is represented as a deep convolutional neural network (CNN) that takes the low-resolution image as the input and outputs the high resolution one. The agenda for this field is to enable machines to view the world as humans do, perceive it in a similar manner, and even use the knowledge for a multitude of tasks such as Image & Video recognition, Image Analysis & Classification, Media Recreation, Recommendation Systems, Natural Language Processing, etc. Many research works have been done on that topic but most of are only capable of recognizing digits only. Here, try to develop a model called CRNN model (Convolutional Recurrent Neural Network) for detection, recognize and acknowledge handwritten. This system will be capable of various images and next automatically recognize which letter is in the image. In our model, I get 93.06% accuracy for detection character.

Keywords — Convolutional Neural Network (CNN), Letter Recognition, Machine Learning, Deep Learning.

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	i
Declaration	ii
Acknowledgements	iii
Abstract	iv
CHAPTER	
CHAPTER 1: INTRODUCTION	**1-2**
1.1 Introduction	1
1.2 Motivation	1
1.3 Research Questions	1
1.4 Aim and Expected Output	2
1.5 Organization of the report	2
CHAPTER 2: BACKGROUND	**3-9**
2.1 Preliminaries/Terminologies	3
2.1.1 Convolutional Neural Network (CNN)	4
2.1.2 CNN Variants	4
2.2 Related Works	4
2.3 Challenges	9
2.4 Research Summary	9
2.5 Scope of the Problem	9

CHAPTER 3: RESEARCH METHODOLOGY **10-21**

3.1 Introduction 10

3.2 Research Subject and Instrumentation 11

3.3 Data Collection 11

3.4 Image Preprocessing 12

3.5 Gaussian Blur 13

3.6 Edge Detection 13

3.7 Dilation and Erosion 14

3.8 Image Calibation 15

3.9 Preprocessing and preparing the images for training 15

3.10 CNN (Convolutional Neural Network) 17

3.11 DenseNet 121 18

3.12 InceptionV3 18

3.13 Xception 19

3.14 MobileNet 19

3.15 VGG16 19

3.16 ResNet 20

3.17 Descriptive Analysis 20

3.18 Implementation Requirements 21

CHAPTER 4: EXPERIMENT RESULTS AND DISCUSSION **22-25**

4.1 Introduction 22

4.2 Experimental Results 22

4.3 Model Accuracy 22

CHAPTER 5: IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY **26-27**

5.1 Impact on Society 26

5.2 Impact on Environment 26

5.3 Ethical Aspects 27

5.4 Sustainability Plan 27

CHAPTER 6: SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH **28**

6.1 Summary of the Study 28

6.2 Conclusions 28

6.3 Implication for Further Study 28

REFERENCES **29-30**

LIST OF FIGURES

FIGURES	**PAGE NO**
Fig. 3.1: Complete workflow	9
Fig. 3.2: Block diagram of a methodology	10
Fig. 3.3: View Data	14
Fig. 3.4: Handwritten Words	14
Fig. 3.5: Character Detection	15
Fig. 3.6: CNN	15
Fig. 3.7: DenseNet121	16
Fig. 3.8: InceptionV3	16
Fig. 3.9: Xception	17
Fig. 3.10: Testing Accuracy For a different Architecture	18
Fig. 4.1: Training and Validation Loss with Accuracy	20
Fig. 4.2: Confusion Martix of CRNN	20

LIST OF TABLES

TABLE	PAGE NO
Table 2.1: Performance Comparison	7
Table 4.1: Experimental Result Of Train Model	19
Table 4.2: Result Of Confusion Matrix CRNN	22

CHAPTER 1
INTRODUCTION

1.1 Introduction

The automatic handwriting character emphasis has a wide range of instructional and commercial purposes. The first step in handwritten character recognition is to cope with a wide range of handwriting styles with the help of different authors. Furthermore, some complex handwritten scripts feature separate word-writing patterns. They are cursive in certain circumstances, and characters are occasionally connected with each other. Many scholars in the field of NLP have already noticed this attempt [1-3].

Handwritten character attention is more difficult than printed character attention for the following reasons. PyTorch consists of detecting textual content from paper and changing the images in such a layout that a pc can manipulate and do quite several computational operations. And Tensorflow detects the handwritten character and words. There has been a variety of preprocessing steps that were once used to put together the statistics for attention.

1.2 Motivation

In this current era, technological know-how is our most reliable option. So, it is very vital to enrich our pc imagination and prescient via the English language. In the area of computer vision, the upward push of high-performance computing (HPC) structures flips the problem of photo focus very hard. So now there exists an extensive variety of algorithms and architecture to control such problems. Many models are created to remedy these issues using Deep CNN.

1.3 Research Questions

In this thesis, we use convolutional networks, a recent development in machine learning to remedy some of these problems. In particular, we simplify the system pipeline by using a single end-to-end trainable model that learns the relevant preprocessing and image features from raw data. We focus on improving the learning mechanism, both the model and algorithm, rather than on designing novel image features or improving specific subsystems.

Some of the questions which will be answered in this work are:

- Are convolutional networks sufficiently powerful to perform handwritten lesion segmentation as an end-to-end task? Is scarce data an obstacle to learning?
- Is deep learning feasible with the resources we have? Is our data and computational
- power sufficient? Is there any advantage to using GPU acceleration?
- Can we simplify the pipeline for handwritten detection? Can preprocessing be replaced
- Are convolutional networks a good option for future research?

1.4 Aim and Expected Output

These days most of the science is learned on the IoT due to the fact the future has a massive use of lots of devices. This science was once developed and will be developed to make people greater no longer extra skilled, fast, and environment friendly in their work. These days most software corporation is working in this quarter to enhance accuracy to reap the maximum plausible outcome. The purpose is to convey the theories, algorithms, and sketch patterns out of labs, scratch paper, and into people's day-to-day lives to make their existence easy. This project will be a simple word detection system. it will detect words from images and we can copy those words. No need to type.

1.5 Organization of the report

This thesis is organized as follows: Chapter 02 we will discuss the review of the research papers like the related work and the studies behind it. In Chapter 03 We will discuss the methodology of our project which we used before. In chapter 04 We are focusing on experimental results and discussion of the system. Chapter 05 is the conclusion and the future scope of the research-based project is described.

CHAPTER 2
BACKGROUND

2.1 Preliminaries

Here, DCNN model include VGG network for Bangla character detection U. Meier et al. [2] (2011). Transcribed person acknowledgment from a characteristic picture has an enormous arrangement of troubles. Bangla's manually written characters are made out of extremely complex shapes and strokes. The new improvement of the profound learning approach has solid capacities to remove undeniable level elements from a part of a picture. This paper will show a clever methodology that incorporates a multi-facet convolutional neural organization followed by a commencement module and a completely associated neural organization. The proposed design is utilized to fabricate a framework that can perceive Bangla characters from various essayists with changed penmanship styles. Dissimilar to past carefully assembled include extraction techniques, this CNN-based methodology learned additional summed up and exact elements from a huge scope preparing dataset W. Song et al. [3] (2011). Programmed manually written Bangla character acknowledgment (HBCR) is a moving issue in PC vision because of various varieties in the composing styles of a singular Bangla character and the presence of likenesses in shapes among various characters. Considering the intricacy of the issue, we want to foster an advanced convolutional neural organization (CNN) for exact acknowledgment, yet tragically, as of now, not many Bangla manually written datasets contain an enormous number of picture tests for each character appropriate for preparing profound learningbased strategies. In this paper, we present aibangla, another benchmark picture information base for secluded written by hand Bangla characters with point-by-point utilization and an exhibition gauge I.-J. Kim et al. [4] (2015). Bangla is one of the most communicated in dialects from one side of the planet to the other, regardless, endeavors on Bangla written by hand character acknowledgment are not sufficient. The utilization of Deep Convolutional Neural Network (DCNN) based classifiers has turned into a victory over the condition of craftsmanship AI strategies. Utilizing the DCNN model to arrange Bangla's secluded essential characters can give better results given its capacity to distinguish many concealed elements from a picture M. M. Rahman et al .[5] (2015).

2.1.1 Convolutional Neural Network (CNN)

LeCun et al. [7] (1998) used CNN to digit concentration tasks for the first time. CNN and its variants are constantly being modified to a wide range of tasks C. Dong et al . [8,9] (2016,2012). CNN is intended to mimic human visual processing, and it has amazingly tuned structures to analyze 2D pictures.

2.1.2 CNN Variants

In terms of CNN structure, it can be discovered that there are several important and urgent aspects that are employed to build an environmentally friendly DCNN design. The convolution layer, pooling layer, completely related layer, and Softmax layer are examples of these features. This community's superior structure comprises of a stack of convolutional layers and max-pooling layers watched by a totally connected and Softmax layer at the end. LeNet et al. [7] (1998), AlexNet et al . [9] (2012), VGG Net, All-Conv, and NiN are notable instances of such networks. There are other alternatives and superior architectures proposed, including GoogleNet with inception layers C. Szegedy et al . [11][20] (2015,2022), ResNet, FractalNet, and DenseNet.

2.2 Related Works

Bangla Handwritten person acknowledgment obtained extensive consideration in many examination regions, for example, PC vision and picture handling for its amazing applications. In such a manner, a Bangla transcribed person acknowledgment technique is proposed in this paper. The critical difficulties of this work are line division, word division, and character division. For that in this paper, the distance-based division (DBS) technique is utilized to fragment the sentence, word, and character independently. To execute the DBS technique proficiently, at first, the info Bangla report is pre-handled to resize and dispense with the commotion. From that point forward, the versatile thresholding strategy is used to eliminate the shadows from the info picture. The proposed DBS technique is applied to the handled picture to section the sentences, words, and characters. To portion words from a sentence, right off the bat fragment the lines from the archive. Because of these lines, the words in each line are sectioned. At long last, characters are sectioned from individual words. These fragmented characters are utilized as ROI to extricate the highlights and ship off SVM to the group. To assess the exhibition of the proposed strategy original copies of Rabindranath Tagore and various individuals'

Bangla written by hand archives are thought [1].

Transcribed digit acknowledgment is an average picture order issue. Convolutional neural organizations, otherwise called ConvNets, are amazing order models for such assignments. As various dialects have various styles and states of their numeral digits, the precision paces of the models shift from one another and from one language to another. Notwithstanding, unaided pre-preparing in such circumstances has shown further developed exactness for arrangement assignments, however, no such work has been found for Bangla digit acknowledgment. This paper presents the utilization of solo pre-preparing utilizing auto encoder with profound ConvNet to perceive manually written Bangla digits, i.e., 0-9. This paper presents the execution of solo pre-preparing, utilizing an autoencoder, as a pre-cursor to directed preparing for Bangla written by hand digit acknowledgment. Future investigations can investigate whether pre-preparing on small datasets . Because of its intricate letter forms and numerous practical applications, handwritten character recognition has recently become a hard and exciting area. A lot of research has already been done and is being done on English letter sets and numbers recognition. In any case, Bangla, although being the fifth most spoken language on the earth, has not undergone that kind of investigation. Furthermore, the numerous complicated states of the Bangla letter put the recognition to the test. In this study, we offer a directed example strategy for highlighting Bangla numeric characters that achieves excellent exactness of recognition. We utilize Local Directional Pattern (LDP) and Gradient Directional Pattern (GDP) to include extraction and afterward two notable AI calculations to arrange the numeric person [2].

Handwritten character identification is a difficult problem since it attempts to identify the right class of user-independent handwritten characters. This problem becomes significantly more difficult for thoroughly adapted, morphologically sophisticated, and perhaps juxtaposition characters carrying language, such as Bengali. As a result, as compared to other languages, the advancements in Bengali person recognition have been fundamentally smaller. We present a convolutional deep model to perceive Bengali transcribed characters in this study. We first developed a valuable arrangement of parts by utilizing bits and nearby open fields, and then we employed thickly connected layers for the separation job. On the BanglaLekha-Isolated dataset, we tested our approach [3].

Recognition of Handwritten Character has been one of the potential topics of research due to its applicability in a variety of sectors; nonetheless, it looks to be difficult research. In this study, we focus on unconnected manually written person recognition in the native language (Bengali) by first identifying individual letters. The primary approaches for unconnected manually written person acknowledgement may be classified into two types: divisional and all-encompassing. We used division-based hand written word recognition in our technique, and neural organizations were used to differentiate individual characters. We fabricated an application that can be valuable in making on-the-web duplicates of actual Bengali books and scripts. This element will be helpful in reclamation or safeguarding purposes and will work with making on-the-web forms of actual books which will make them more open to the majority because of simple accessibility and brought down cost. This task will likewise help in text-todiscourse interpretations of Bengali texts and thusly will be valuable as a movement help application to unfamiliar sightseers recognizes street signs, names of spots and structures, and so on A profound CNN is extremely encouraging for transcribed person acknowledgment. The states of Bengali person set probability is making future difficulties. With the appearance of a best-in-class framework more particular elements can be removed by building networks having a greater limit. The new best-in-class structure (GPU) can be used to set up the Deep CNN model with immense datasets having a wide scope of shapes [4].

There are a few words in Bangla written by hand character acknowledgment. Here another technique is proposed to perceive the person from nonstop Bangla written by hand character. This issue frequently occurs in transcribed texts like a successive person shows up on another person. About Bangla characters, division turns out to be considerably more troublesome. To fabricate a viable OCR arrangement of Bangla written by hand text, acknowledgment of characters is significant as the division of characters. Here the principal reason for existing is to make a framework, which takes nonstop Bangla written by hand text pictures as info and afterward fragments the information texts [5].

A conventional CNN architecture was utilized in this study to recognize online handwritten isolated Bangla characters. A detailed examination of the effects of utilizing various component types, pooling methods, and enactment capacities in CNN engineering has been carried out. An aggregate of 10000 person tests were used in this work, with 30% of the cases chosen as test sets and the remaining 70% used to develop the recognition model. The method achieved 99.40% acknowledgement precision on the test dataset. The outcome is superior than the recently suggested properly prepared highlights used for the recognition of online manually typed Bangla characters [6].

The testing's overall performance is in contrast to various current methods. The effects are introduced in Table 1 The experimental outcomes exhibit that the current DCNN models.

Table 2.1: Performance Comparison

SL No	**Author Name**	**Deep Learning Architecture**	**Accuracy(%)**
1	D. C. Cireşan, "U. Meier, L. M. Gambardella, and J. Schmidhuber	Recurrent Neural Network	97.56
2	U. Meier, D. C. Ciresan, L. M. Gambardella, and J. Schmidhuber	All Convolutional Network(All-Conv)	94.31
3	W. Song, S. Uchida, and M. Liwicki	Residual Network(ResNet)	96.73
4	I.-J. Kim and X. Xie	Convolutional Neural Networks	97.33

5	M. M. Rahman, M. Akhand, S. Islam, P. C. Shill, and M. H. Rahman	Convolutional Neural Networks	96.87
6	D. Ciregan and U. Meier	All Convolutional Network(All-Conv)	97.92
7	Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner	Recurrent Neural Network	95.76
8	C. Dong, C. C. Loy, K. He, and X. Tang	Convolutional Neural Networks	97.56
9	A. Krizhevsky, I. Sutskever, and G. E. Hinton	Convolutional Neural Networks	94.31
10	M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda	Convolutional Neural Networks	96.73
11	C. Szegedy, W. Liu, Y. Jia	Deep Learning Network	97.33
12	O. Russakovsky, J. Deng, H. Su	FractalNet	97.87
13	V. Nair and G. E. Hinton	All Convolutional Network(All-Conv)	97.92
14	M. Lin, Q. Chen, and S. Yan	Network in Network(NiN)	96.73
15	S. Ioffe and C. Szegedy	Deep Learning Network	95.76

2.3 Challenges

Collect the data from a distinctive character has continually been a huge challenge. First of the difficulties is to making a shape for gathering handwritten data. Making a training dataset without a photo is no longer possible. Most problems that show up that have an effect on the system when it is running. In this current COVID-19 situation we are failing to collect more datasets. We collect data from Google (CMATERdb dataset). We failed to add special characters. If we are able to add special characters, we must get better accuracy.

2.4 Research Summary

In this report, We are working to find a way to recognize the handwritten consonant letter. There are many differences in the handwriting of different people. Many times it becomes very difficult to understand their handwriting. Due to the diversity of human handwriting, it is often difficult to understand what is written. I am implementing this because I need a model that can easily recognize handwriting.

2.5 Scope of the Problems

- To use this thesis people need to give a handwritten picture as input and a consonant letter can be found as output. Remember to give a consonant as input.
- In this project I am working on a method to find the consonant. Remember that you cannot identify the consonant from any paragraph, only the consonant can tell if the input is consonant.
- This project of mine will only apply to consonant recognition.it cannot work with any other consonant.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

Create model for character and handwritten recognition, there needs to be a model that can precisely apprehend the character so that it can pick out the character. The methodology consists of the processed dataset then the building model and then the train. The technique will be applied to a similar model of character recognition. Figure 1 displays the complete workflow of methodology.

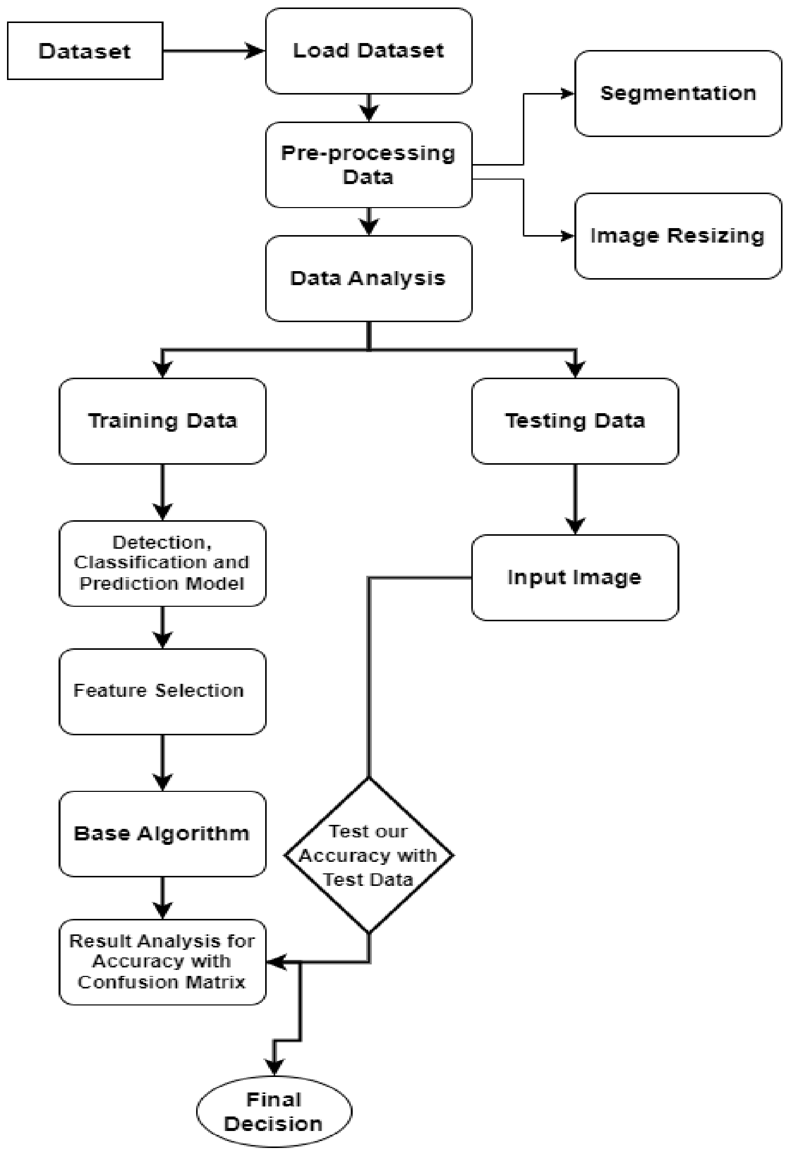

Fig. 3.1: Complete workflow

Due to the variation in the type of handwriting of different people Handwriting recognition has become difficult. We have tried to work with some deep neural networks in this project to create a new model such as a sequential CNN, various architecture of CNN. After we work through it, we contrast the aftereffect of the calculation and the outcome. This software has tried to create an image that will take input images and then identify which algorithms work best with its consonant input image.

3.2 Research Subject and Instrumentation

Handwriting identification is challenging due to the wide range of handwriting styles among different people. In this thesis, I attempted to create a model that delivers a higher recognition result. For this, we developed a CNN and various different CNN designs, as well as compared the outcomes. Figure 2 depicts the many steps of the whole procedure.

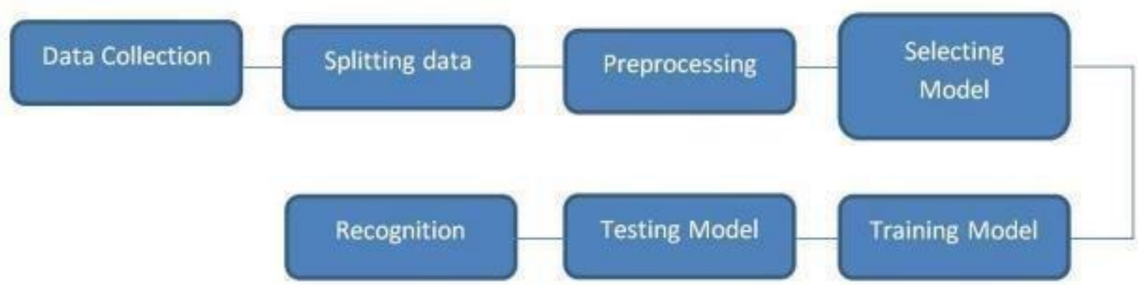

Fig. 3.2: Block diagram of a methodology

3.3 Data Collection

The records used to assist the findings of this to learn about are accessible [20]. This dataset contains almost 400 thousand handwritten names gathered through charitable programs. Image processing technologies are used in character recognition to translate characters on scanned documents into digital forms. It usually works well in machine-printed typefaces. However, due of the wide range of human writing styles, it remains challenging for robots to distinguish handwritten characters. There are a total of 206,799 first names and 207,024 surnames.

3.4 Image Preprocessing

Image scaling in the digital image occurs at some certain points whether this is in Bayer demosaicing or in photo enlargement. It happens at any time when we need to resize our image from the one-pixel grid to another grid. Conversion of an image is required if we want to increase or decrease the total number of transmissions in an image. Although the same dimension is made, the result may differ depending on the algorithm. In this project, Photos have been resized due to the number of reasons but one of them is very important to our project. All camera has its own solution, so when the system is created for some camera details it will not go correctly with any other camera depending on the same features.

So it is necessary to make a regular solution to the application and then make an image of conversion. The original image we have captured in a mobile device is with resolution 5520*4140. If the original resolution is processed to detect the object width it outdoes the screen resolution as the testing device screen resolution is 1366*768. For this the image is resized using a python code to a resolution of 600*450. This one is the only one and major preprocessing used in our project.

People analyze colors using wavelength-sensitive cells known as cones. There are three different types of cones, each with a different sensitivity of magnetic radiation of a wide range of waves. Different cones are sensitive to different lights. One is sensitive to green light, one is to red and one is to blue light. When it flashes a combination of three colors (red, green, blue) and promotes three types of cones, it is also possible to make almost any of the colors to be seen by us. This is why the features of colors are always kept as three separate image matrices; One keeps as red (R) in each pixel, another one in green pixel (G) and the other one as Blue (B). This model is known as the RGB color model. However, we do not consider the amount of effuse in the grayscale image, rather than we emit the same amount in each channel of a grayscale image. Here little light gives dark pixels and much light is perceived as bright pixels. During the conversion of an RGB image into a Grayscale image, we need to consider the RGB parameters for each of the pixel and make as a single value output reflecting the brightness of that respective pixel. (R + B + C) / 3. Since the light is understood to be governed by the green section, a different, human-oriented method, the way to make it is to consider a weighted average of the notes, for example, 0.2R + 0.39G + 0.47B.

3.5 Gaussian Blur

In picture preparing, a Gaussian obscure (otherwise called Gaussian smoothing) is the consequence of obscuring a picture by a Gaussian capacity (named after mathematician 13 and researcher Carl Friedrich Gauss). It is a broadly utilized impact in designs programming, commonly to lessen picture clamor and decrease detail. The visual impact of this obscuring procedure is a smooth obscure looking like that of a survey the picture through a translucent screen, particularly not the same as the Bokeh impact delivered by an out-of-center focal point or the shadow of a question under normal enlightenment. Gaussian smoothing is likewise utilized as a pre-preparing stage in PC vision calculations keeping in mind the end goal to upgrade picture structures at various scales like scale space portrayal and scale space execution. Numerically, if we apply a Gaussian dim to a picture it becomes the same as involving the image with a Gaussian capacity. This is otherwise called a two-dimensional Weierstrass change.

3.6 Edge Detection

Edge detection incorporates an assortment of scientific techniques that go for distinguishing focuses in an advanced picture at which the picture splendor changes forcefully or, all the more formally, has discontinuities. The focuses at which picture brilliance changes strongly are commonly sorted out into an arrangement of bended line portions named edges. A similar issue of discovering discontinuities in one-dimensional signs is known as step identification and the issue of discovering signal discontinuities after some time is known as change location. Edge location is a principal device in picture handling, machine vision, and PC vision, especially in the regions of highlight discovery and highlight extraction.

In the perfect case, the consequence of applying an edge finder to a picture may prompt an arrangement of associated bends that demonstrate the limits of items, the limits of surface markings and additionally bends that relate to discontinuities in a surface introduction. In this manner, applying an edge identification calculation to a picture may essentially decrease the measure of information to be handled and may, along these lines, sift through data that might be viewed as less applicable, while saving the imperative basic properties 14 of a picture. On the off chance that the edge identification step is effective, the ensuing assignment of translating the data substance in the first picture may, along these lines, be significantly improved. In any case, it isn't generally conceivable to get such perfect edges from genuine pictures of direct multifaceted nature.

3.7 Dilation and Erosion

Dilation and Erosion Morphology is an expansive arrangement of picture handling activities that procedure pictures in light of shapes. Morphological tasks apply an organizing component to an information picture, making a yield picture of a similar size. In a morphological task, the estimation of every pixel in the yield picture depends on a correlation of the relating pixel in the information picture with its neighbors. By picking the size and state of the area, you can build a morphological task that is delicate to particular shapes in the info picture.

The most fundamental morphological tasks are widening and disintegration. Enlargement adds pixels to the limits of articles in a picture, while disintegration expels pixels on question limits. The quantity of pixels included or expelled from the articles in a picture relies upon the size and state of the organizing component used to process the picture. In the morphological expansion and disintegration tasks, the condition of any given pixel in the yield picture is controlled by applying an administer to the relating pixel and its neighbors in the information picture. The govern used to process the pixels characterizes the activity as an enlargement or a disintegration. This table records the principles for both widening and disintegration.

3.8 Image Calibation

ACCD cluster is mechanically very steady; the pixels have a settled geometric relationship. Every pixel inside the exhibit, notwithstanding, has special light affectability attributes. Since these qualities influence camera execution, they should be expelled through adjustment. The procedure by which a CCD picture is aligned is known as fat fielding or shading rectification. Geometric camera alignment, additionally alluded to as camera re-sectioning, gauges the parameters of a focal point and picture sensor of a picture or camcorder. You can utilize these parameters to revise for focal point twisting, measure the extent of a protest in world units, or decide the area of the camera in the scene. These errands are utilized as a part of uses, for example, machine vision to distinguish and measure objects. They are likewise utilized as a part of apply autonomy, for route frameworks, and 3-D scene remaking.

3.9 Preprocessing and preparing the images for training

To begin, we divided my consonant letter dataset into a test set and a train set in order to test and prepare the model. We divided the dataset into 75% train set and 25% test set for successive CNN. Furthermore, the separation size for other engineers was 80% in the train set and 20% in the testset.

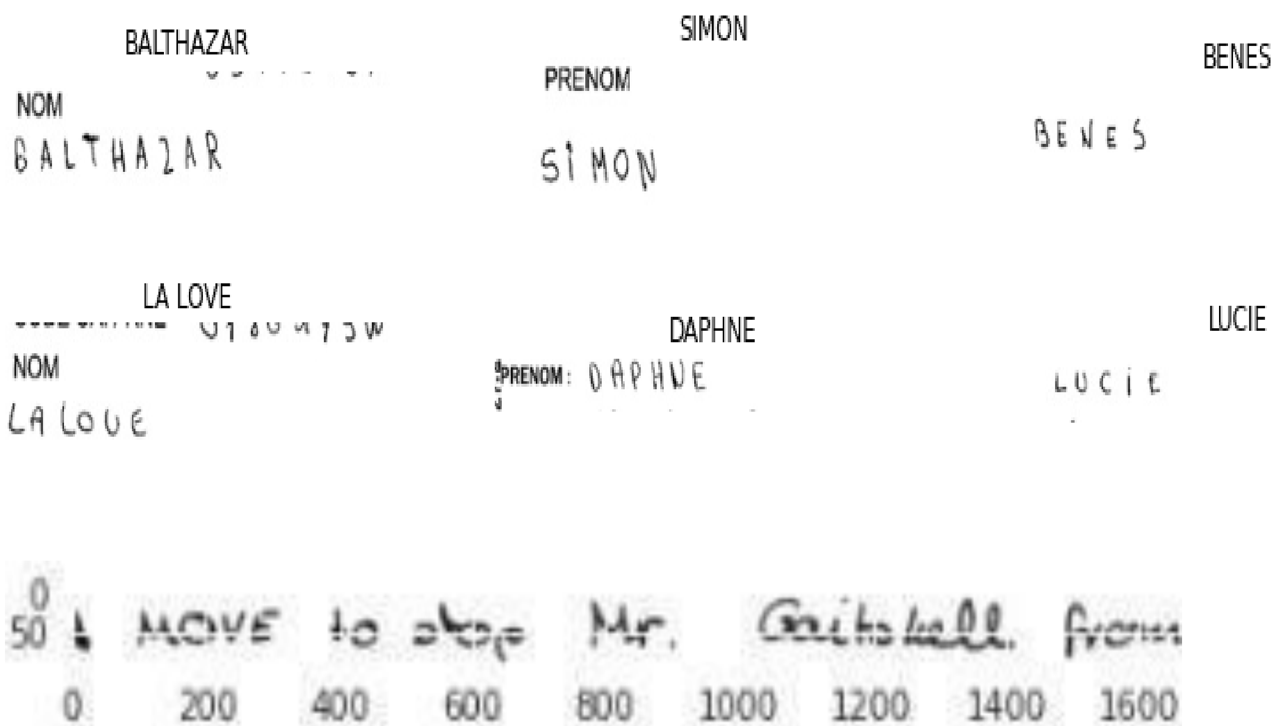

Fig. 3.3: View Data Handwritten

words look like below:

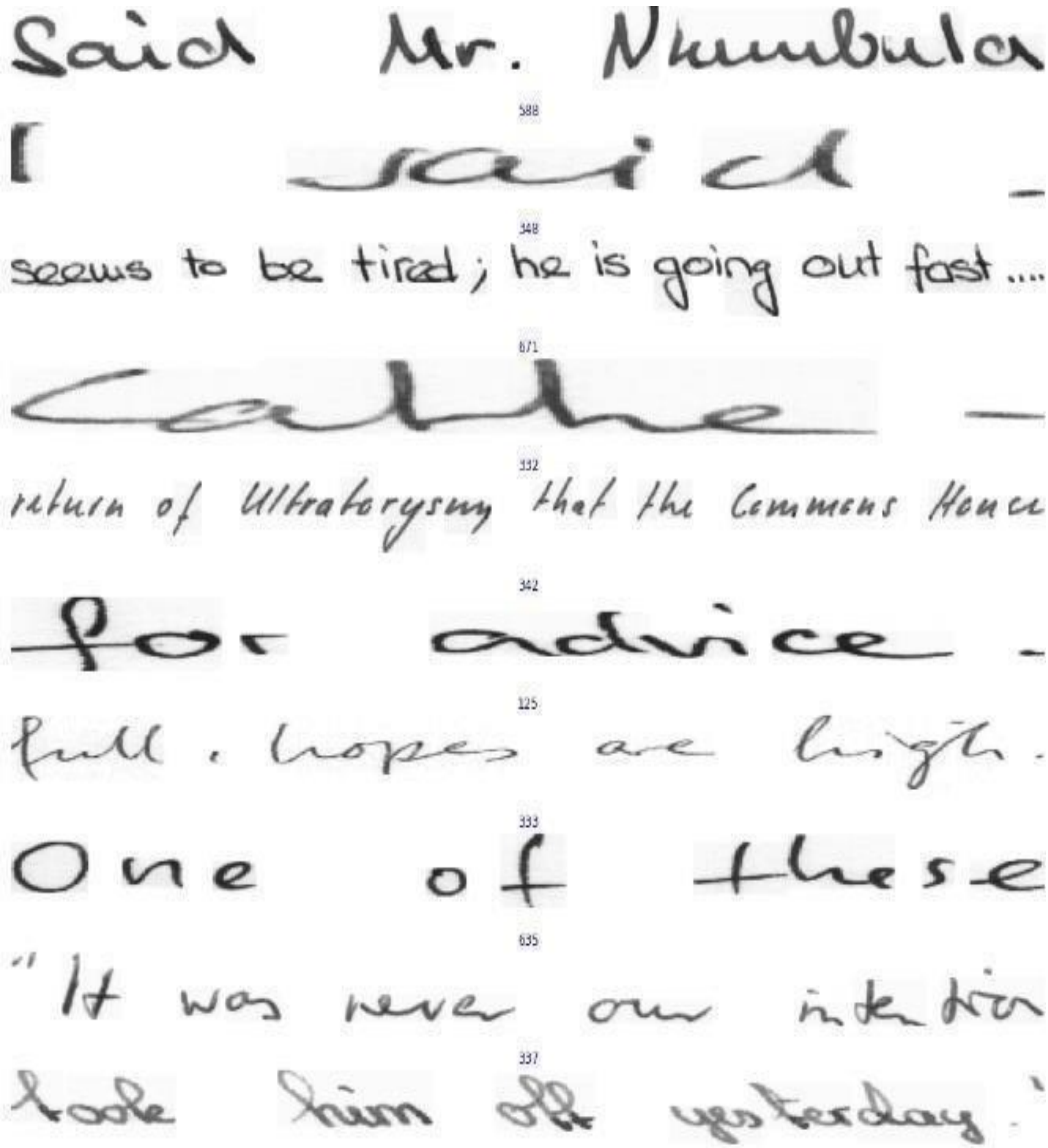

Fig. 3.4: Handwritten Words

The photos are loaded in grayscale and resized to 256 width and 64 height. If the width and height are more than 256 and 64, respectively, they are chopped. If they are smaller, white pixels are used to pad the picture. Finally, the picture is rotated clockwise to get the desired image form (x, y). After that, the picture is normalized to the range [0, 1].

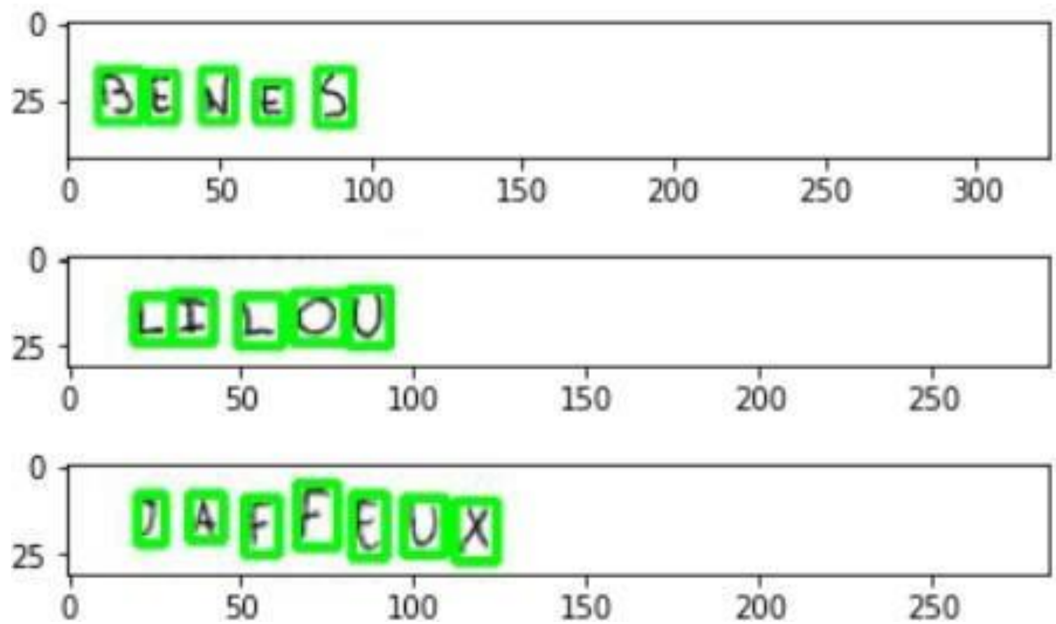

Fig 3.5: Character Detection

3.10 CNN (Convolutional Neural Network)

CNN is a type of network architecture for deep learning which is specifically involved in image detection, pixel data processing. The array data is used for tasks. Neural networks have CNN network architecture for detection and detection.

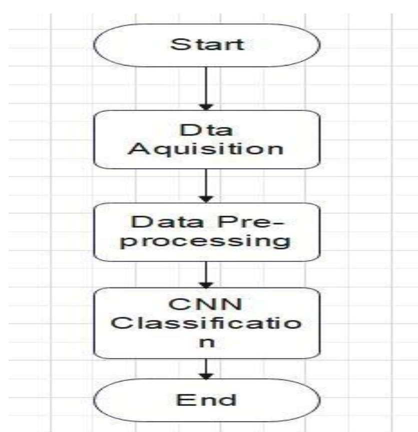

Fig 3.6: CNN

3.11 DenseNet 121

The number 121 corresponds to the number of your shots with the teachable weight. For, example convolutional neuron networks have layers and are fully connected to this layer. And additional layers include primarily 7x7conventional layers and three transitional layers and are fully connected to this layer.

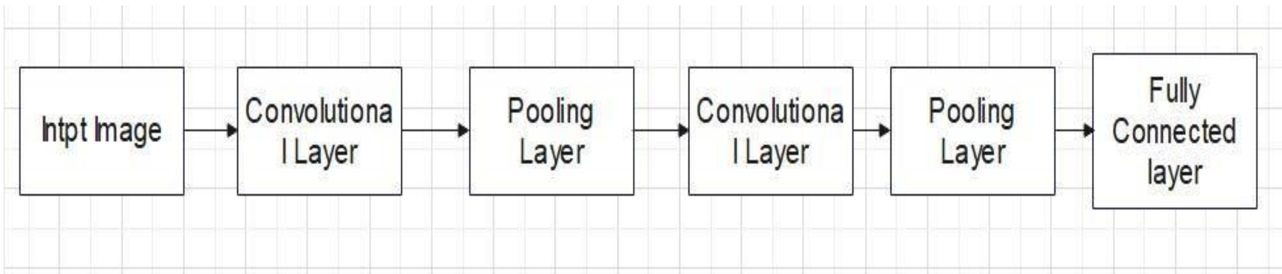

Fig 3.7: DenseNet121

3.12 InceptionV3

Inception V3 is a convolutional neural network that is composed of maximum pooling and convolutional with it. It is fully incorporated into the network that you cannot know its structure by heart. To know its structure we need to understand Keras and manage the relationship.

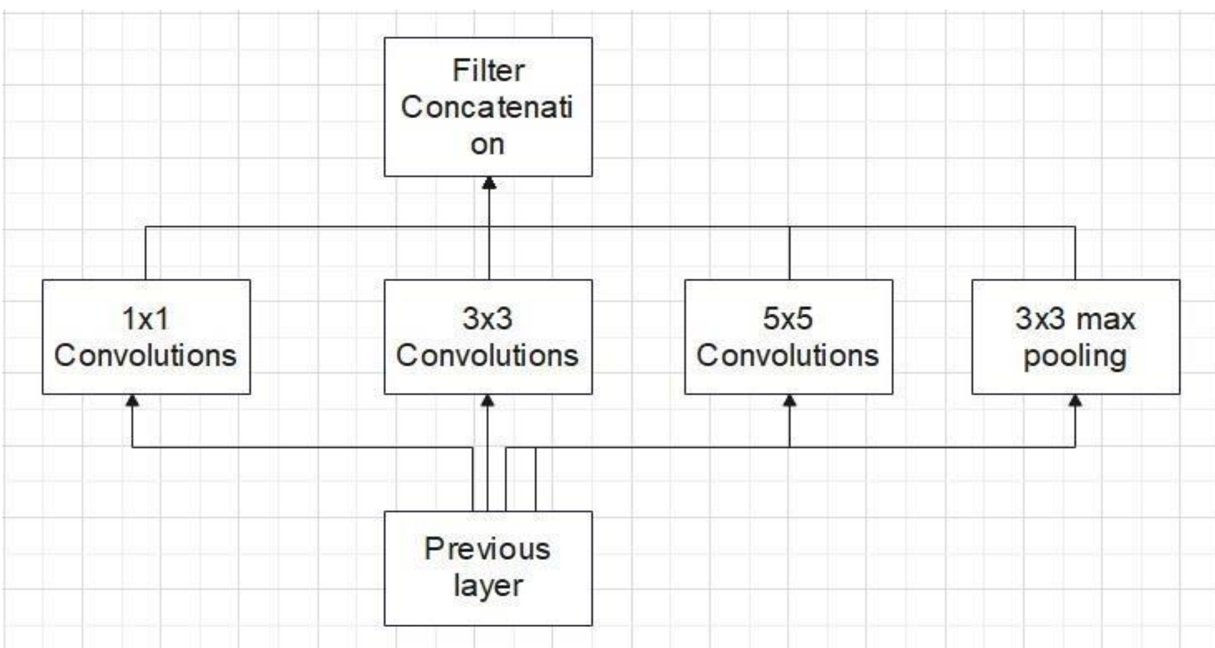

Fig 3.8: InceptionV3

3.13 Xception

The exception is the 71st layer of a convolutional neural network, which is much deeper. ImageNet can store more than one million pre-trained image networks from the database. Networks can classify objects into 1000's of categories. Like mouse, keyboard, pencil, pen, many types of pets etc.

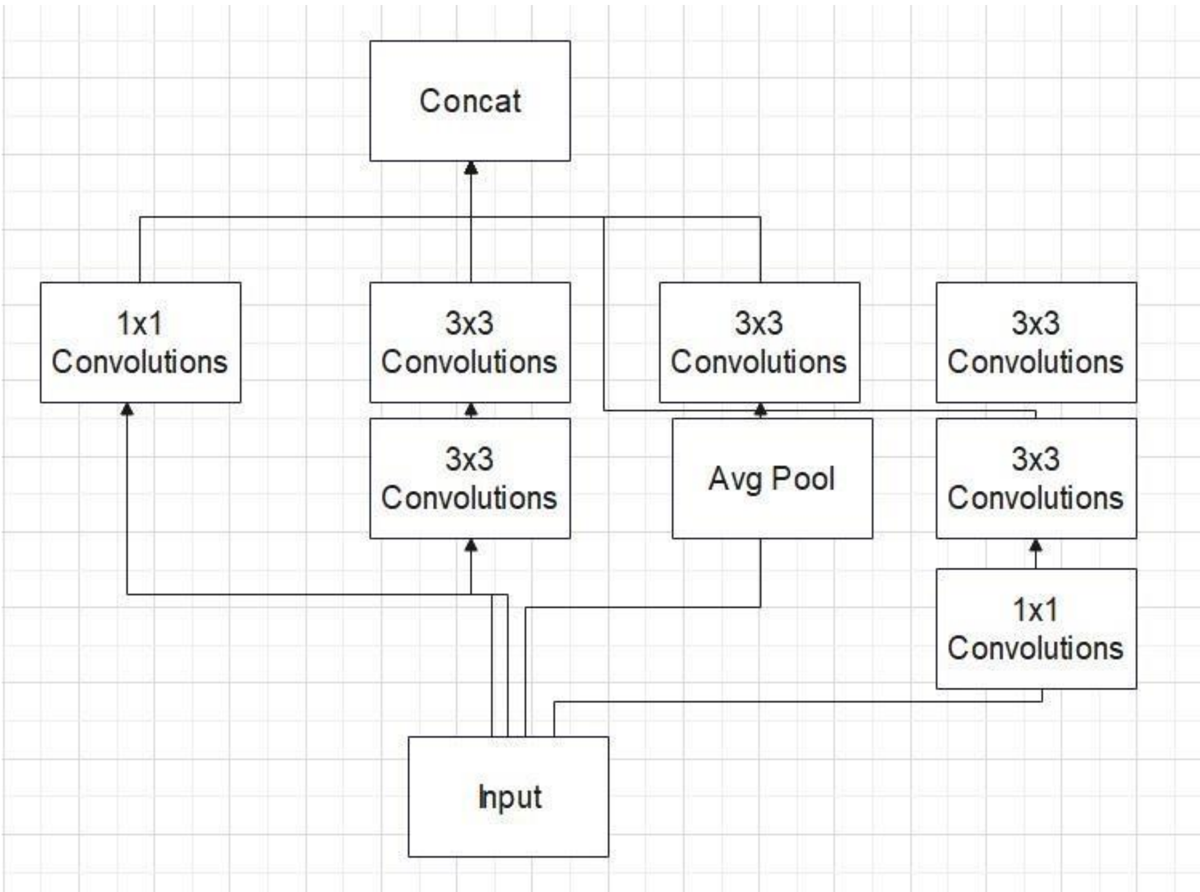

Fig 3.9: Xception

3.14 MobileNet

Mobile is a streamlined architecture that uses very deep convolutional neural networks to operate in a lightweight manner. And models are provided for mobile and embedded vision applications using only confusion.

3.15 VGG16

VGG16 is a 16th layer deep convolutional neural network. ImageNet can store more than one million pre-trained image networks from the database. Networks can classify objects into 1000's of categories. Like mouse, keyboard, pencil, pen, many types of pets etc.

3.16 ResNet

ResNet is an architecture that allows artificial neural networks to skip layers without affecting their model.

3.17 Descriptive Analysis

We used a confusion matrix for accurate visualization of CNN and other architectures we plot graph. For Xception 92.41% accuracy, for VGG16 96.04% accuracy, for ResNet50 96.37% accuracy, for MobileNet 93.07% accuracy, for Inception V3 93.07% accuracy, for DenseNet 121 91.75% accuracy and CNN accuracy is 87.30%.

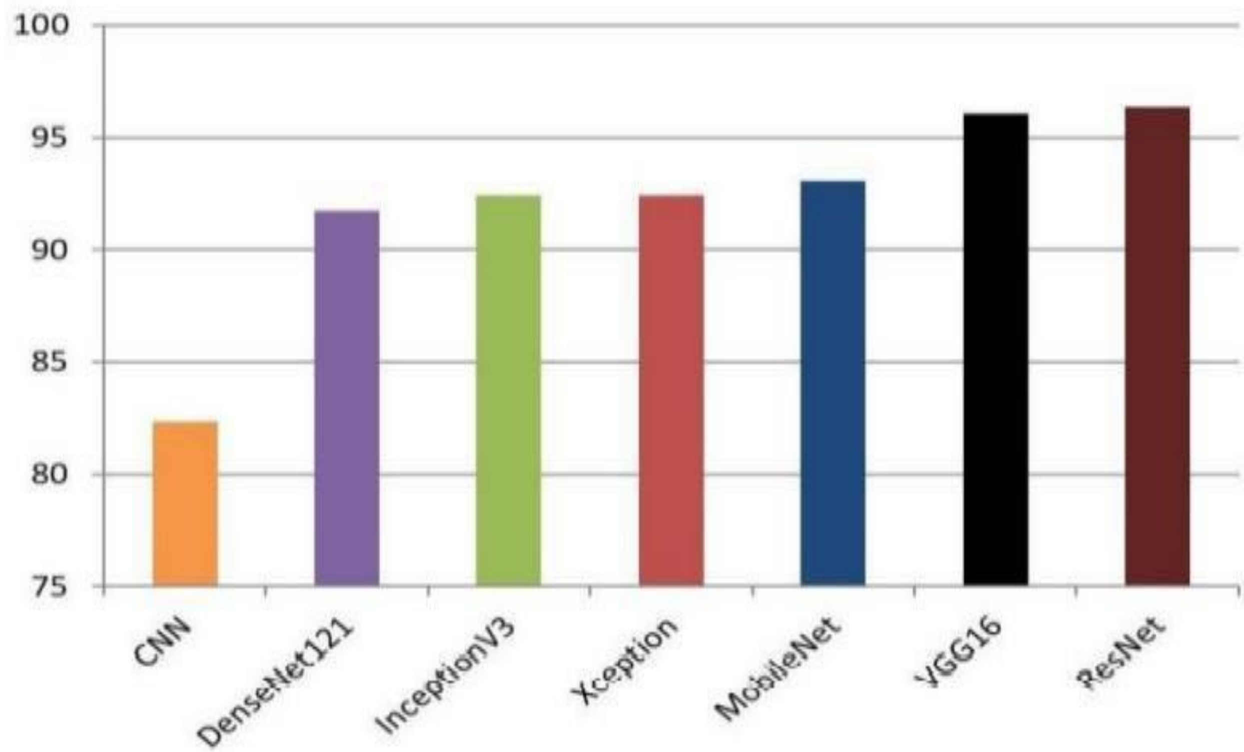

Fig 3.10: Testing Accuracy For a different Architecture

3.18 Implementation Requirements

We are researching Handwritten Letter recognition. The list of tools, devices, languages, and programs that we used in this thesis are given below:

Back end:

- RAM: 8GB
- Operating system: Windows 11
- Python
- CPU: Intel core i7
- Kaggle
- OpenCV
- Numpy
- Tensorflow
- Sklearn
- Panda

CHAPTER 4

EXPERIMENT RESULTS AND DISCUSSION

4.1 Introduction

In this section, we will portray the test that we made. In the wake of perusing this section, anybody can get information about the aftereffect of my venture's anticipated result and get information about the consequences of the proposed model shown on an acknowledgment of Bangla consonant letters dependent on every calculation which is carried out in my task.

4.2 Experimental Results

Instantiate the Custom Callback and train the model

Table 4.1: Experimental Result Of Train Model

Epoch	loss	Accuracy	Validation loss	Validation accuracy	Learning Rate	Next Learning Rate	Monitor	Improve in %	Duration
1 /40	9.536	26.317	7.25683	51.020	0.00100	0.00100	Accuracy	0.00	98.69
2 /40	5.932	61.117	4.53708	79.184	0.00100	0.00100	Accuracy	132.24	75.93
3 /40	3.907	76.867	2.78082	92.245	0.00100	0.00100	Accuracy	25.77	76.06
4 /40	2.685	83.867	1.98233	90.204	0.00100	0.00100	Accuracy	9.11	75.78
5 /40	1.923	87.783	1.40244	95.918	0.00100	0.00100	Accuracy	4.67	79.03

4.3 Model Accuracy

The training loss for exclusive CRNN models is proven in Figure 10. It is clear that the DenseNet suggests fine convergence in contrast with the different CRNN approaches. DenseNet once more suggests ultimate validation accuracy in contrast with different CRNN approaches. Now, plot training data, evaluate and save the model.

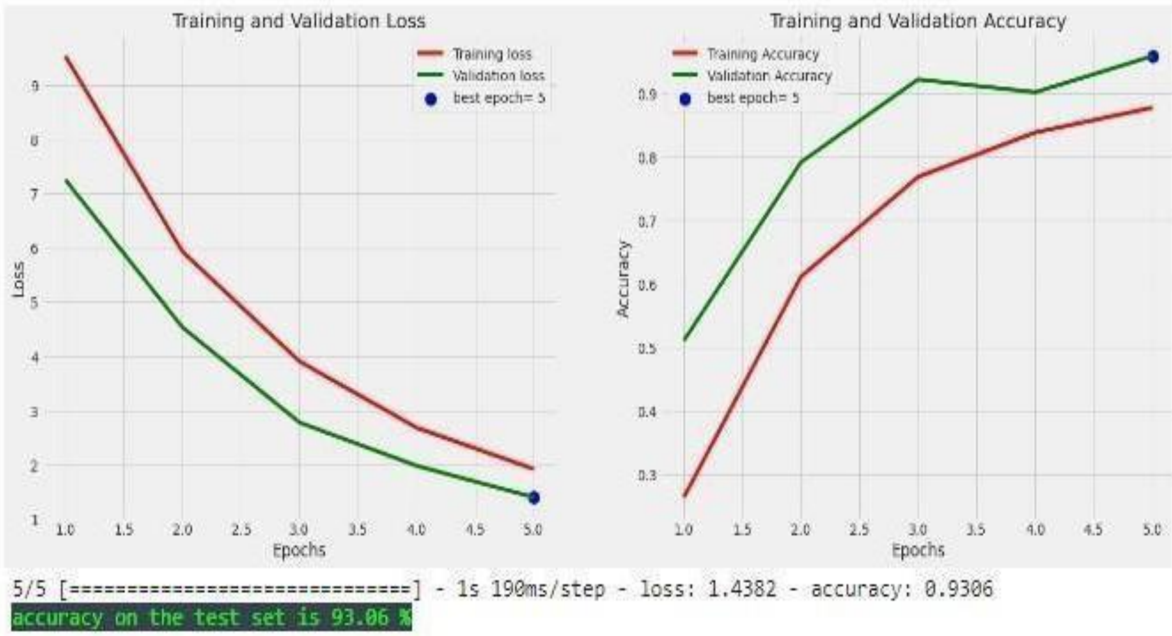

Fig. 4.1: Training and Validation Loss with Accuracy

For the sequential CRNN model, we get an accuracy rate of 93.06%.

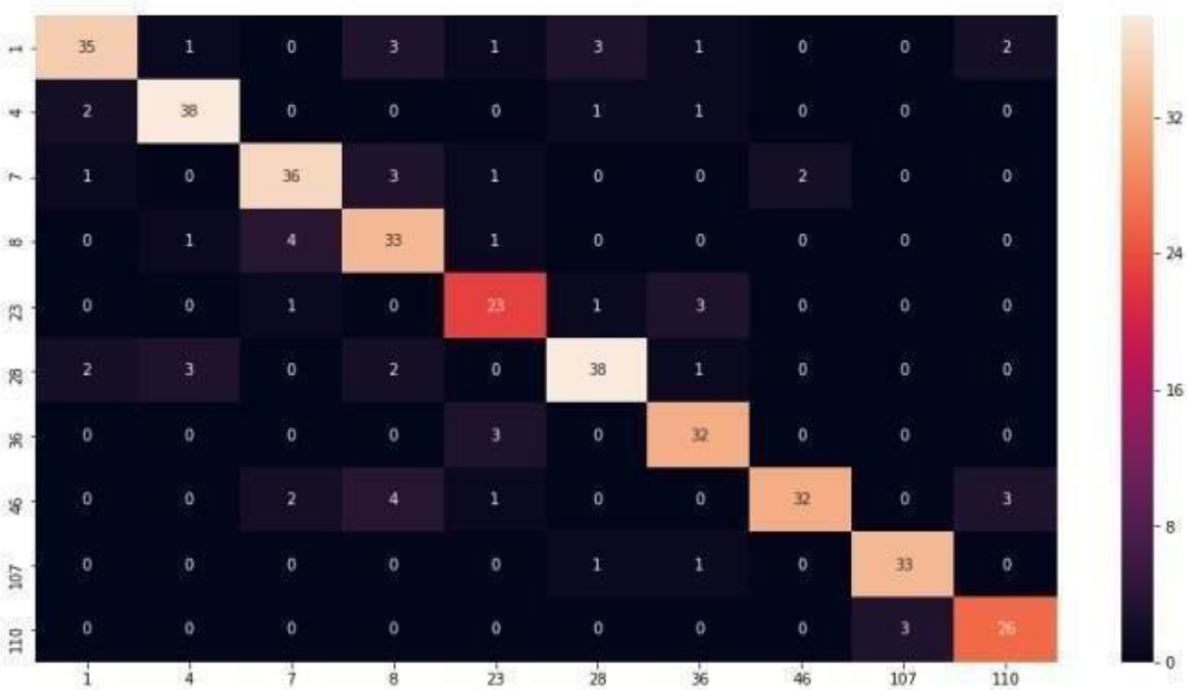

Fig 4.2: Confusion Martix of CRNN

Classification Result Report:

Table 4.2: Result Of Confusion Matrix CRNN

	Precision	Recall	F1- Score	Support
000	1.00	0.94	0.97	35
025	1.00	1.00	1.00	3
026	1.00	1.00	1.00	3
037	1.00	1.00	1.00	2
085	1.00	1.00	1.00	2
118	1.00	1.00	1.00	2
123	1.00	1.00	1.00	2
125	1.00	1.00	1.00	3
126	1.00	1.00	1.00	3
128	1.00	1.00	1.00	2
130	1.00	1.00	1.00	3
133	1.00	1.00	1.00	3
150	1.00	1.00	1.00	5
151	0.83	1.00	0.91	5
152	1.00	1.00	1.00	5
153	1.00	1.00	0.83	5
154	1.00	1.00	0.92	6
155	1.00	1.00	1.00	5
173	1.00	1.00	1.00	3
174	1.00	0.67	0.80	3
209	1.00	1.00	1.00	3
315	1.00	0.67	0.80	3
332	1.00	1.00	1.00	4
333	1.00	0.83	0.91	6
334	1.00	1.00	1.00	5
335	1.00	1.00	1.00	4
336	1.00	1.00	1.00	4
337	1.00	1.00	1.00	3
338	1.00	0.75	0.86	4
339	0.67	0.80	0.73	5
340	0.75	0.75	0.75	4
341	1.00	0.80	0.89	5
342	0.67	1.00	0.80	4
343	1.00	1.00	1.00	4
344	1.00	1.00	1.00	5
345	1.00	1.00	1.00	6
346	1.00	1.00	1.00	6
347	1.00	1.00	1.00	4
348	0.75	0.60	0.67	5
349	1.00	1.00	1.00	4
384	1.00	1.00	1.00	6
415	1.00	0.75	0.86	4
551	1.00	1.00	1.00	6
552	0.88	1.00	0.93	7
567	1.00	1.00	1.00	4

588	1.00	0.50	0.67	6
634	0.75	0.86	0.80	7
635	0.80	0.80	0.80	5
670	1.00	0.83	0.91	6
671	0.60	1.00	0.75	6
Accuracy			0.93	245
Macro Avg	0.95	0.93	0.93	245
Weighted Avg	0.94	0.93	0.93	245

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

Handwritten word detection It helps to convert textually. It can also be done as a format at the beginning of the online platform. It can also be stored, reviewed, understood and explained to others. It also works well in developing systems for detection in the case of neural networks. It also helps simulate what a person's brain does when they read. Also helps to match handwriting through machines. At last we can say it can be studied as a text book. When we purposefully change our handwriting, we change the attitudes that others may have about us. In this case, our achievements can motivate and affect risk-taking. And through us it can bring out the best result. Mainly because handwriting, thought and language can activate the brain for healing working memory and reflection. Handwriting strengthens our reading and language processes. Handwriting slows down the thought process, enabling the brain to think about how words can shape their spelling and writing. So we can say it has a great impact on our society. Thus the environment of our society is balanced through handwritten word detection.

5.2 Impact on Environment

When a person writes about the impact on the environment, it is not immediately possible to implement it. When we send any postal address of a letter, the type of text we write while writing the letter, while writing the bank check like date, amount, name of the customer etc. And when we fill the form of any school college or university it affects. The process of writing and typing slows down our memory and concentration refers to our ability to remember and write or create something specific on paper. It is important to study handwriting to discover the characteristics of a person through handwriting analysis. It will tell you about the people in your life, their characteristics, and is a powerful means of knowing their behavior. You can get to know someone better or analyze yourself by collecting their handwriting samples or taking a look at your writing yourself.

5.3 Ethical Aspects

The problem with handwriting is that writing is good and bad in many ways. It plays an important role for a student if the writing is good then it will be understood and read and if the writing is bad then if the writing is not catchable or read then it has a positive effect. Also sometimes the texts look the same in this case for the computer to catch that text. That's why we write the texts through convolutional neural network through the computer using hand detection.

5.4 Sustainability Plan

Stimulates brain function. Children and adults need time as well as brain and exercise. Increases learning disabilities. Standardly it acts as an evaluation and critical part. Boosts confidence of adults and children. Practicing it makes a good writer. Helps to increase comprehension and understanding quickly. We will use handwritten word detection to highlight the natural variations in writing. It plays a huge role in writing patterns and image processing using it. Texts can easily convey the features that exist.

CHAPTER 6

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH

6.1 Summary of the Study

Handwritten focus is turning into a necessary problem in a number of years however it turns into an assignment to get suitable overall performance due to the alignment and many of them are similar. The mapping is represented as a deep convolutional neural network (CNN) that takes the low-resolution image as the input and outputs the high resolution one.

6.2 Conclusions

In this thesis, we looked at the exhibition of a few famous profound CNN. We investigated the overall performance of countless famous deep CNN for handwritten personality recognition. The experimental consequences indicated that DenseNet is a fine performer in classifying alphabets. Specifically, we performed an attention rate for the handwritten alphabet is more than 90% using CNN. To the excellent of my knowledge, these are the satisfactory awareness consequences on the dataset.

6.3 Implication for Further Study

In the future, some fusion-based Densly connected neural network models, such as IRCNN, will be explored and developed for handwritten character recognition. In the future, we will try to include voice features in it. When someone uploads a document it will also convert into text plus it will automatically read the document. In fact, it will help those people who cannot read & write.

REFERENCE

[1] D. C. Cireşan, “U. Meier, L. M. Gambardella, and J. Schmidhuber, Deep, big, simple neural nets for handwritten digit recognition, Neural Computation, vol. 22, no. 12, pp. 3207–3220, 2010.

[2] U. Meier, D. C. Ciresan, L. M. Gambardella, and J. Schmidhuber, Better digit recognition with a committee of simple neural nets, in Proceedings of International Conference on Document Analysis and Recognition (ICDAR), pp. 1250–1254, Beijing, China, September 2011.

[3] W. Song, S. Uchida, and M. Liwicki, Comparative study of part-based handwritten character recognition methods, in Proceedings of International Conference on Document Analysis and Recognition (ICDAR), pp. 814–818, Beijing, China, September 2011.

[4] I.-J. Kim and X. Xie, Handwritten Hangul recognition using deep convolutional neural networks, International Journal on Document Analysis and Recognition, vol. 18, no. 1, pp. 1–13, 2015.

[5] M. M. Rahman, M. Akhand, S. Islam, P. C. Shill, and M. H. Rahman, Bangla handwritten character recognition using convolutional neural network, International Journal of Image, Graphics and Signal Processing, vol. 7, no. 8, pp. 42–49, 2015.

[6] D. Ciregan and U. Meier, Multi-column deep neural networks for offline handwritten Chinese character classification, in Proceedings of 2015 International Joint Conference on Neural Networks (IJCNN), pp. 1–6, IEEE, Killarney, Ireland, July 2015.

[7] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[8] C. Dong, C. C. Loy, K. He, and X. Tang, Image super-resolution using deep convolutional networks, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 295–307, 2016.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet classification with deep convolutional neural networks, in Proceedings of Advances in Neural Information Processing Systems, pp. 1097–1105, Lake Tahoe, NV, USA, December 2012.

[10] M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, Subject independent facial expression recognition with robust face detection using a convolutional neural network, Neural Networks, vol. 16, no. 5-6, pp. 555–559, 2003.

[11] C. Szegedy, W. Liu, Y. Jia et al., Going deeper with convolutions, in Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9, Boston, MA, USA, June 2015.

[12] O. Russakovsky, J. Deng, H. Su et al., Imagenet large scale visual recognition challenge, International Journal of Computer Vision, vol. 115, no. 3, pp. 211–252, 2015.

[13] V. Nair and G. E. Hinton, Rectified linear units improve restricted Boltzmann machines, in Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 807–814, Haifa, Israel, June 2010.

[14] M. Lin, Q. Chen, and S. Yan, Network in network, Cornel University, <<https://arxiv.org/abs/1312.4400>>, March 10, 2021 6:30 PM GMT.

[15] S. Ioffe and C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, in Proceedings of International Conference on Machine Learning, pp. 448–456, Lille, France, July 2015.

[16] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 770–778, Las Vegas, NV, USA, June-July 2016.

[17] G. Larsson, M. Maire, and G. Shakhnarovich, Fractalnet: ultra-deep neural network without residuals, <<https://arxiv.org/abs/1605.07648>>, March 11, 2021 8:00 PM GMT.

[18] G. Huang, Z. Liu, K. Q. Weinberger, and L. van der Maaten, Densely connected convolutional networks, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, no. 2, p. 3, Honolulu, HI, USA, July 2017.

[19] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, Learning hierarchical features for scene labeling, IEEE Transactions on Pattern Analysis and Machine Intelligence,

[20] Dataset Handwriting Recognition https://www.kaggle.com/datasets/landlord/handwriting-recognition, access 2022."

3e

ORIGINALITY REPORT

28%	23%	16%	20%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	5%
2	Submitted to Daffodil International University Student Paper	4%
3	Submitted to Multimedia University Student Paper	1%
4	Md. Shahriar Islam, Eimdadul Haque Moon, Md. Ashikujjaman Shaikat, Mohammad Jahangir Alam. "A Novel Approach to Detect Face Mask using CNN", 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), 2020 Publication	1%
5	Saja Salim Mohammed, Jamal Mustafa Al-Tuwaijari. "Skin Disease Classification System Based on Machine Learning Technique: A Survey", IOP Conference Series: Materials Science and Engineering, 2021 Publication	1%

6	Mingyang Li, Zequn Wang. "LSTM-augmented Deep Networks for Time-Variant Reliability Assessment of Dynamic Systems", Reliability Engineering & System Safety, 2021 Publication	1 %
7	www.researchgate.net Internet Source	1 %
8	coek.info Internet Source	1 %
9	isrpgroup.org Internet Source	1 %
10	americaspg.com Internet Source	1 %
11	s3.amazonaws.com Internet Source	1 %
12	www.ijesi.org Internet Source	1 %
13	www.journal-aquaticscience.com Internet Source	1 %
14	Rayan Shaik, Sai Krishna Bodhapati, Abhiram Uddandam, Lokesh Krupal, Joydeep Sengupta. "A Deep Learning Model that Diagnosis Skin Diseases and Recommends Medication", 2022 1st International Conference on the Paradigm Shifts in Communication, Embedded Systems,	1 %