Handwriting Recognition By Logistic Regression And Neuralnetworks

Prof. Doke Kanchan Kiran[1], Abhinas Mishra[2], Suraj Desai[3], Simran Nagpurkar[4]
[1, 2, 3, 4] Dept of Computer Engineering
[1, 2, 3, 4] Bharati Vidyapeeth College of Engineering, Navi Mumbai

Abstract- *English characters are foremost wide adopted writing systems within the world. Previous analysis had main focus on recognizing handwritten English numbers but to teach a machine recognition of English characters is a challenge and a fascinating task. This paper aims to review existing strategies for the handwritten character recognition using machine learning algorithms and implement it using the logistical regression as a discriminative model together with multilayer perceptron for recognizing English characters. The application provides a solution for handwriting recognition supported through camera frames or an image file, learning new characters, and learning through user's feedback. Multilayer perceptron is chosen as the recognition model which is a feedforward ANN, particularly thanks to its high performance on nonlinearly divisible issues. It's additionally proved powerful in OCR and ICR systems that might be seen as an additional extension of this work. And when evaluating the perceptron's performance and configuring its parameters within the Octave programming language, it is implemented as an Android application using the identical perceptron design, learning parameters and optimization algorithms. The application is tested on a set of training data having digits with the flexibility to find out alphabetical or completely different characters.*

Keywords- Character recognition, Multilayer Perceptron, Image processing, OCR.

I. INTRODUCTION

The key goal of this project is to recognise handwritten characters which is provided by the user as input through touch and image/camera in an offline manner. Also, the application provides means of adding and learning a new character and learning interactively from user's feedback.

Handwritten character identification is a subject of research in NLP, computer vision, artificial intelligence and pattern recognition. A computer performing handwriting recognition is alleged to be ready to acquire and distinguish characters in paper archives, pictures and different sources and transform them into machine-encrypted form. Its purpose is found in OCR and more advanced ICRsystems.Themajority of these frameworks these days execute ML components, for example, neural networks.

Machine learning is a domain of artificial intelligence inspired by psychology and biology that deals with learning from a set of data and can be applied to solve wide spectrum of problems. In supervised learning the model is provided with cases of data to a problem case and also solution precise that solves the problem for each case. When learning is complete, the model is able not only to provide answers to the data it has learned on, but also, to yet unseen data with high precision.

In machine learning neural networks are used as learning models. The learning process is simulated in the NN that occurs in an animal or human neural system. Being a powerful learning model, they useful in automation of tasks where the decision of a human being takes too long, or is imprecise. A neural network can be extremely quick at conveying results and may detect relation between observed instances of data that human can't see.

II. IDENTIFY,RESEARCHANDCOLLECT IDEA

Handwritten character recognition is currently used extensively in OCR and ICR systems. Some of the existing systems are listed below.

2.1. Full-featured Document OCR and ICR Systems

ABBYY FineReader [12] by the ABBYY company is a piece of software with worldwide recognition that deals with OCR and ICR systems, as well as applied linguistics. The company has also developed a business card reader mobile application that uses a 12 smartphone's camera for text recognition to import contact information. The application is available for the Android platform. Tesseract-OCR was an OCR engine which was developed in the year 1985 and 1995 at HP Labs. It is claimed that this engine is the most precise open source OCR engine existing, supporting a wide choice of image formats and over 60 languages. It is open source software and free, licensed under Apache License 2.0. Google Goggles is an image recognition Android and iOS application, featuring searching based on images taken by compatible devices and using character recognition for some use cases.

2.2.Input methods

Microsoft has been supporting a tablet handwriting-based input method since the release of Windows XP Tablet PC Edition [13]. This allows users of devices with this platform to write text using a digitizing tablet, a touch screen, or a mouse, which is converted into text that can be used in most applications running on the platform.

Google Translate [14], a machine translation Android application from Google, features handwriting recognition as an input method, as well as translating directly from the camera. This closely resembles a possible extension of the work in the future.

III. METHODOLOGY

The algorithms used in character recognition can be dissented up into three categories: feature extraction, image pre-processingand classification. They are normally used in sequence-image pre-processing helps make feature extraction a smoother process, while feature extraction is necessary for correct classification.

3.1. Image Pre-Processing

Image pre-processing is importantfor accurate correct character prediction recognition pipeline. These methods typically include

- Noise reduction using filtering
- Image Segmentation using thresholding
- Cropping and Scaling

Digital capture and conversion of an image often introduces noise which makes it problematical to decide what is actually a part of the object of interest and what is not. Considering the problem of character recognition, one must reduce noise as much as achievable, while preserving the character strokes, since they are important for correct classification. There are many ways of achieving this. Local processing is one of them.

Local Pre-processing makes use of a small area around the pixel from input image called mask to produce output value. This is called filtering. For this task one must use convolutional masks that scan an image, ideally reducing all unwanted noise. Masks are square matrices with elements representing weights of the surrounding area pixels that determine the light intensity value of the pixel at hand.

$$h = 1/9 \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Typical Average Value

$$h = 1/10 \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Centroid highlight

In order to preserve character strokes in image a non-linear operation median filtering has been used. In median filtering the pixel's value are substituted with the median intensity of surrounding pixels. The example is shown in the figure below.

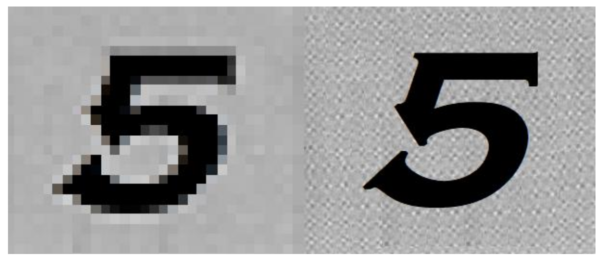

Figure 1. Median Filtering

The process of splitting an image into fragments with strong relation with objects or the real-world properties the image represents is called as image segmentation. Probably the simplest image segmentation method is thresholding. Thresholding is the extraction of the foreground, which is a character in this case, from the rather monotonic background. Any image which is gray-scale, this is equal to binarization of its real-valued data. An image threshold is selected, according to the input image which is converted. Given an input pixel *f(i, j)* and threshold *T*, the output pixel *g(i, j)* is calculated as follows:

$$g(i,j) = \begin{cases} 1, & if\ f(i,j) \geq T \\ 0, & if\ f(i,j) < T \end{cases}$$

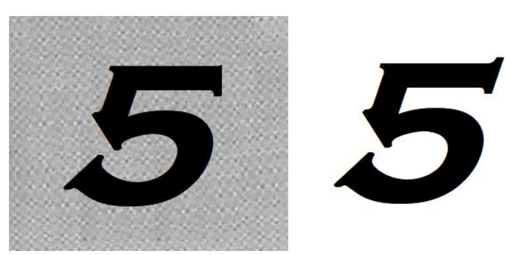

Figure 1. Image Thresholding

Finally, in both touch and image-based recognition cropping and scaling of the images to a small fixed size in the application has been used.

3.2. Feature Extraction

Features of input data are the quantifiable properties of observations, which one uses to classify these occasions of data. The process of feature extraction is to settle on relevant features that separate the occasions well and are autonomous of one another.According to [3], High recognition performance can be achieved by selecting the most applicable of a feature extraction method. There are several methods for feature extraction from character images, each having invariance properties, different characteristics and re-constructability of characters. [3] states that in order to answer to the question of which method is best suited for a given situation, an experimental evaluation must be performed.Themethods explained in [3] are projection histograms, deformable templates,zoning, graph description, template matching geometric moment invariants, unitary image transforms,spline curve approximation,Zernike moments, contour profiles and Fourier descriptors. To describe the way feature extraction is sometimes done in handwriting recognition, briefly explained one of them.

Projection histograms were presented in 1956 out of an OCR framework by Glauberman and are utilized in division of characters, words, and content lines, or to detect if a scanned text page is rotated [3].Then collect the horizontal and vertical projections of an image by setting each vertical and horizontal "bin" value to the count of pixels in respective rows and columns (Figure 3), where neighboring bins can be merged to make the features scale independent. The projection is, however, variant to rotation and difference in writing style.

The two histograms are then compared and a dissimilarity measure is obtained, which can be consumed as a feature vector.

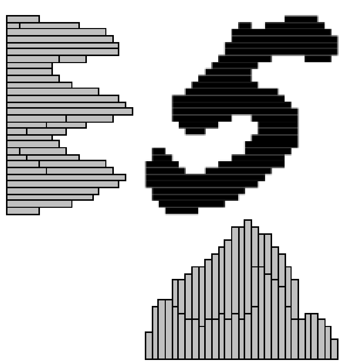

Figure 3:
Horizontal and verticalprojection histograms[3].

As mentioned, there is no method that is intrinsically perfect for a given task. Evaluation of such methods would take a lot of time. We set the focus on multilayer feedforward NN, which can be viewed as a combination of a feature extractor and a classifier [3], the latter of which will be explained further.In this application multilayer layered neural network model has been used, which will be more deeply described in the next chapter. This model is a directed graph having of atleast 3 layers of nodes.

The input layer is the first layer and output layer isthe last layer and there a number of hidden intermediate layers. Except the input layer the nodes of neural networks are also called neurons or units. Each node of a layer typically has a weighted relation to the nodes of the next layer.These hidden layers are important for feature extraction, as they create an internal abstraction of the data which is given into the network. The more intermediate layers there are in a network, the more abstract the extracted features will be.

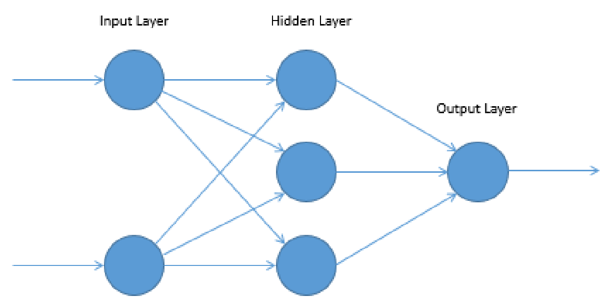

Figure 4:
Basic view at the multilayer perceptron architecture.

3.3. Classification

Classification is defined as the activity of assigning labels (categories, classes) to yet unseen observations (instances of data). In machine learning, this is done on the basis of training an algorithm on a training set of examples. Classification is a supervised learning problem, where a “teacher” links a label to each instance of data. Label is a discrete number that identifies the class a particular instance belongs to. It is usually represented as a nonnegative integer. There are many machine learning models that implement classification; these are known as classifiers. The aim of classifiers is to fit a decision boundary (Figure 4) in feature-space that separates the training examples, so that the class of a new observation instance can be correctly classified. In general, the decision boundary is a hyper-surface that separates an N dimensional space into two partitions, itself being N−1-dimensional.

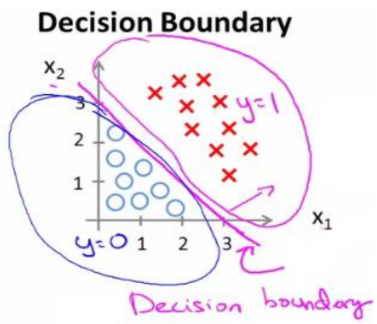

Figure 5:
Visualization of a decision boundary.

3.3.1 Logistic Regression

Logistic regression which is a simple linear classifier algorithm tries to find the decision boundary by iterating over the training examples, trying to fit parameters that describe the decision boundary hyper-surface equation. During this learning process, the algorithm computes a cost function (also called error function), which represents the error measure of its hypothesis (the output value, prediction). This value is used for penalization, which updates the parameters to better fit the decision boundary. The prime objective is to get the cost function to its minimum by converging to parameter values. It has been proved that logistic regression is always convex, therefore the minimization process can always converge to a minimum, thus finding the best fit of the decision boundary this algorithm can provide.

Until now binary classification has been discussed. To apply logistic regression to handwriting recognition, one needs more than 2 distinguishing classes, hence we need multi-class classification. This can be solved by using the one-vs-all approach.

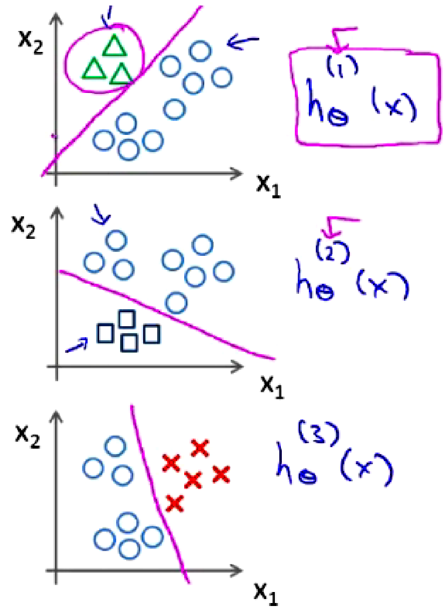

Figure 5:
Multi-class classification of three classes as it is split into three sub-problems.

The principle of one versus all is to fragment the training dataset into a number of binary classification problems. Considering to classify handwritten digits, the problem degrades into 10 sub-problems, where individual digits are separated from the rest. Figure 6 shows the same for 3 classes of objects.The output of each sub-problem is a probability measure representing how likely a data instance fits to the class at hand. The overall class is chosen as the class as the sub-problem of which has the highest probability.

To correctly classify data that are not linearly separable, one can use a logistic regression algorithm while using a non-linear hypothesis to separate the classes. One then has to include high-order polynomials in the feature vector to approximate the complex decision boundary. This has an implicit problem of choosing the right polynomials, which can be an enormous task; moreover, this problem is very specific to the task being solved. To avoid this, all of the possible polynomials could be included, but in this way, the sum of features grows exponentially and it becomes computationally challenging to train on the data.

3.3.2Multilayer Perceptron

Multilayer perceptron's (MLPs) are artificial neural networks, learning models inspired by biology. As opposed to logistic regression, which is only a linear classifier on its own, the multilayer perceptron learning model, which already has been mentioned in terms of feature extraction, can also distinguish data that are not linearly separable.

As already outlined the architecture of an MLP, as seen in (Figure 4). In order to calculate the class prediction, one must perform feedforward propagation. Input data is provided to the input layer and propagated further, passing through weighted connections into hidden layers, using an activation function. Hence, the node's activation (output value at the node) is a function of the sum of weights of the connected nodes at a previous layer. Until the output layer is reached this process continues.

The learning algorithm, backpropagation, is different from the one in logistic regression. First, the cost function is measured on the output layer, propagating back to the connections between the input and the first intermediate layer afterwards, updating unit weights. MLPs can perform multi-class level classification as well, without any modifications. To simply set the output layer size to the number of classes one wants to recognize. After the hypothesis is calculated, pick the one with the maximum value. An activation function which is nonlinear is required for the network to separate nonlinearly separable data instances.

IV. RESULT AND DISCUSSION

The character recognition model is trained to recognize handwritten English alphabets. The alphabets are divided into 25 various segments, Neural Network architecture is proposed primarily in order to process 25 input bits. Following are parameters of network which is used for training:

Number of Intermediate layers (Hidden) = 2
Rate of learning= 0.05
Number of Unit in Intermediate layers (Hidden)= 25
Number of Unit in First layers (Input) = 25
Initial Wt. = Random [0,1]
Transfer Function Used for intermediate Layer 1= "Logsig"
Transfer Function Used for intermediate Layer 2= "Tansig"

4.1. Training

The training set consist of the alphabets binary code. It is not rational to input these shapes separately when generating training sets as the shape of character is dependent on the handwriting. Therefore, this was automatized so that the whole letter is input to the framework, the shape of the alphabet segment needed is extracted from this full letter. The examples for training set patterns can be seen below:

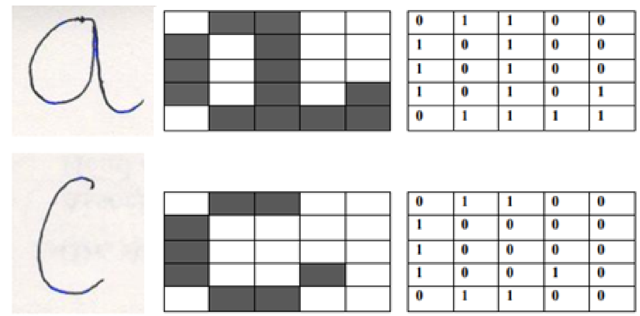

Figure 6: Training Set

4.2. Testing

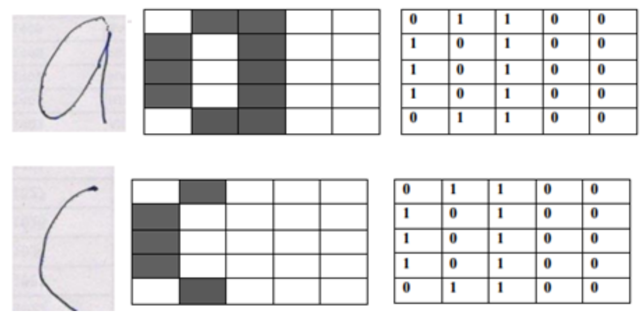

Figure 6: Testing Set

Results can be computed as:

Alphabet	No. of samples for training	No. of samples for testing	No. of epochs	% Recognition Accuracy
a	20	5	294	94.0
b	20	5	321	83.0
c	20	5	587	71.0
d	20	5	282	88.0
e	20	5	548	64.0
f	20	5	254	85.0
g	20	5	247	89.0
h	20	5	263	92.0
i	20	5	658	72.0
j	20	5	599	73.0
k	20	5	300	91.0
l	20	5	652	71.0
m	20	5	456	86.0
n	20	5	398	82.0
o	20	5	356	94.0
p	20	5	264	88.0
q	20	5	287	82.0
r	20	5	669	70.0
s	20	5	202	88.0
t	20	5	252	79.0
u	20	5	458	80.0
v	20	5	488	77.0
w	20	5	511	94.0
x	20	5	341	91.0
y	20	5	268	71.0
z	20	5	296	90.0

Table 1: Result Set

V. CONCLUSION

This paper aims on the machine learning techniques used in the project. At first, reviewed the approaches that are nowadays used in similar applications. There are different methods through which 'handwritten character recognition' is achieved. The method to give efficient and effective result both for feature extraction as well as recognition. Project focuses on implementation part which uses logistic regression algorithm. Finally, the results of the implementation of the learning algorithms have been contrasted. The logistic regression algorithm has been chosen due to its better efficiency and accuracy from another algorithm. The application performs handwritten character identification based on touch, image, and camera input.Moreover, new features can be added to improve the accuracy of recognition.This algorithm can be tried on large database of handwritten text.The work can be extended to work on degraded text or broken characters.This work further extended to the character recognition for other languages.

VI. ACKNOWLEDGMENT

Having endured the experience, there were many who helped us in our project and we very much like to thank everyone.

We are profoundly obliged to our HOD Dr. D. R. Ingle, for giving us this significant chance to do this task and we express our generous gratitude.

We also thank our guide Prof. Kanchan Doke and our coordinator Prof. Rahul Patil for kindly helping and advising during the work and also who gave this opportunity to work on this project on “HANDWRITING CHARACTER RECOGNITION” through which we came to know about so many new things.

It is incredible joy to recognize the assistance and recommendation, which we got from the Department of Computer Engineering. We wish to express our sincere thanks to everyone who helped us in finding information about report. Much good help and support has given on various events by our entire family.

REFERENCES

[1] Drawing and Recognizing Chinese Characters with Recurrent Neural Network-2018 IEEE Xu-Yao Zhang, Fei Yin, Yan-Ming Zhang, Cheng-Lin Liu, Fellow, IEEE, and YoshuaBengio.
Available At:
https://arxiv.org/pdf/1606.06539.pdf

[2] Deep Random Vector Functional Link Network for Handwritten Character Recognition-2016 IEEE Hubert Cecotti School of Computing and Intelligent Systems, Ulster University Londonderry, Northern Ireland, UK
Available At:
http://uir.ulster.ac.uk/35837/1/paper_ijcnn16_hubert.pdf

[3] Multi-Column Deep Neural Networks for Offline Handwritten Chinese Character Classification-2015 IEEE Dan Cire§anUeli Meier IstitutoDalleMolle di studisull'lntelligenzaArtificiale (IDSIA) SUPSI-USI

[4] Accurate Handwritten Digits Recognition using Convolutional Neural Networks-2016 IEEE Yan Yin, JunMin Wu, HuanXin Zheng University of Science and Technology of China ISuzhou Institute for Advanced Study Department of Computer Science and Technology

[5] Feature Set Evaluation for Offline Handwriting Recognition Systems: Application to the Recurrent Neural Network Model -IEEE 2016 Youssouf Chherawala, ParthaPratim Roy, and Mohamed Cheriet

[6] Machine Learning-Stanford University Andrew Ng
Available At:
https://www.coursera.org/learn/machine-learning

[7] EdX Deep Learning using TENSORFLOW
Available At:
https://www.edx.org/course/deep-learning-with-tensorflow

[8] Scikit learn for Datasets
Available At: http://scikit-learn.org

[9] Toward Data Science for articles
Available At: https://towardsdatascience.com

[10] Science Direct for journals and open access content.
Available At: https://www.sciencedirect.com

[11] Online Statistics Education for introductory statistics
Available At: http://onlinestatbook.com

[12] ABBYY FineReader
Available At: https://finereaderonline.com/en-us

[13] Windows XP Tablet PC Edition
Available At:
https://msdn.microsoft.com/en-us/library/ms812485.aspx

[14] Google Translate
Available At:
https://translate.google.com/#auto/en/A