

Employability Analysis of Person

Farzana Alam^{#1}, Mrunali Dandgavhal^{#2}, Ishani Naik^{#3}, Bharati Pujari^{#4}, Geeta S. Navale^{#5}

[#]Computer Engineering Department, Savitribai Phule Pune University

¹alamfarzana786@gmail.com

²mrunali231999@gmail.com

³ishani.naik5@gmail.com

⁴pujaribharati123@gmail.com

⁵gsnavale_sits@sinhgad.edu

Abstract

Handwriting analysis or Graphology both are the scientific scheme of identifying, recognizing and distinguishing behaviour and patterns related with the handwriting. In today's world it has become easier to train deep neural networks using Convolutional Neural Networks (CNN) because of availability of huge amount of data and various algorithmic innovations which are taking place. After pre-processing of images using OpenCV (Open Source Computer Vision Library) and CNN, then can check for numerous trails present in the sample using structural graphic elements for analysis and predict the personality trait present in the handwriting sample. The features extracted are compared with the SVM trained results stored in the knowledge base using Support Vector Machine (SVM) classifier. World is constantly developing particularly in the Information Technology (IT) industry, recruiting countless number of new employee's each year. These recruiting methods are often typically burdening for the Human Resource (HR) department and this is where a handwriting analysis can be used to provide a heads up regarding the character of an individual, as an individual could lie in an interview, however the brain writing can be tough to manipulate. Therefore, the authors can say that prediction through handwriting analysis can be a non-biased parameter which can be included as a pre-test in any recruitment process thereby analysing the Employability skills in a person. This paper focuses on the related work carried out by different researchers.

Keywords— Handwriting analysis, CNN, SVM Classifier, Feature-Extraction, Data Analysis

I. INTRODUCTION

Handwriting is one of the unique attributes of human beings which provide insights into the physical, emotional and mental condition of a person. Every person has a distinctive style of handwriting, and even if the writing is done using one's foot or mouth, it will depict the same unique features as those delineated by hand. Handwriting analysis or Graphology both used for identifying, recognizing and distinguishing behaviour over the movements and patterns related with the handwriting are the scientific scheme of doing so. Thus, graphology identifies these unique features and strokes with their corresponding personality trait. The strokes and patterns which define and distinguish one sample of handwriting from the rest are associated with a personality trait. Each of the traits is linked to a particular brain pattern. A very distinct neuromuscular movement is associated with a particular brain pattern and bound to occur unconsciously during writing. And thus, handwriting is popularly termed as 'brain-writing' also. Handwriting process is analysed through structural graphic elements in which the order to derive information about the writer's personality.

II. MOTIVATION

The motivation of this topic came when the placement activities at Sinhgad Institutes started. As per the observation the ratio of opportunities when compared to the number of students, were less. In order to select students, it was noticed that companies are taking pre-tests to evaluate students and thereby eliminating the students on the basis of various behavioural parameters. In those times, the idea of handwriting analysis came where it is thought that an individual could lie in an interview but brain writing could be tough to manipulate.

III. RELATED WORK

In this section, the recent works are reviewed in order to extract the concepts that are required in the proposed system. Reviewing of the papers from other authors also provides acknowledgement of the future work scopes which can be implemented in the proposed system.

Rohan Vaidya et.al[1] have described a system which is an android application. They have made use of OpenCV for performing Image processing and have used Tensorflow for training the Neural Network. The system is developed using python programming language. The system comprises of two parts: Firstly, an Android application- The android is used. application helps the user to click a picture of text and picture is passed on to a python script running on a server which further processes this image to extract the relevant information. Secondly, a server – This server is a computer which is capable of executing a python script. Use of the publicly available NIST (National Institute of Standards and Technology) Dataset is done which contains samples of handwritten characters from thousands of writers. The neural network that is used is Convolutional Neural Network. CNN's are State-of-Art neural networks which have huge applications in field of Computer Vision. The neural network model was trained using Tensorflow which is an open source library used for Machine learning applications. OpenCV was used to perform various image processing operations like segmentation, thresholding and Morphological Operations. OpenCV which is used for Image processing, is an open source library. Presently the system can only recognize letters and digits. There are many possible applications of this system. Some of the applications are Cheques processing in Banks, helps in desktop publishing, text recognition from business cards, helping the blind in recognition of handwritten texts on letters.

The new method for estimating the slant of word in hand written text was proposed by Rustam Latypov et. al [2]. The main objective of the research is to provide a tool to observe small changes of slant height in the texts during work. Feature extraction is the most important step in handwriting analysis; on the other end a list of the features to be fetched from a manuscript was then developed. Most of these features are qualitative: baseline form, size of margins, height of letters, slant of word. The problem of observing of the human characteristics is very crucial now, but it seems to be rather difficult to reveal any kind of behavioural changes without a direct contact with that person. The slant of words is a very important feature of a handwritten text. Mostly, authors are dealing with a qualitative determination of the slant: controlled left, extreme right, extreme left, controlled right and vertical. The precision of the measure depends on the parameters of the procedure. An evaluation is found for the average slant of words, that turned out to be close to an angle of 75 degrees. Although, this paper recognizes the slant height of single letter but it has no provision for recognition of the slant and height of the whole word which when considered can improve the recognition from an input image. In addition, our proposed system will take yet more parameters to enhance the prediction.

Syeda Asra et.al [3] had implemented the fusion of two domains, psychology and computer science engineering. The problem is recognized in the field of psychology and realized in software

engineering. The proposed work is subdivided into two phases called testing and training. Hand written images of the word are taken as the input image. A pre-processing procedure is done by image colour conversion and image resize. A termed interested region is used for segmenting the middle loop letter. The segmented middle loop is then passed to feature extraction methods like FCC and zoning for extracting features. The training part is done by previously trained samples of hand written images and stored in the knowledge base. These features are then comparing with knowledge based stored data using SVM Classifier. They used Deterministic Finite Automation (DFA), Federal Communications Commission (FCC), SVM training and zoning feature algorithm and got highest accuracy of 86.66obtained predicts the nature of person depending upon hand writing style. The proposed system mainly focuses on writing style of letter O to define behaviour of person. However, the author has used only one approach which is SVM classifier to obtain the results. So, the system is implemented using hybridized approach which will be using SVM and CNN as classifiers along with the mathematical calculation used in the paper to reach higher accuracy. Also, we will consider writing style of whole word with the features like slant, slope, height, pen pressure.

The concept and the process of implementation of Graphology using ANN was described by Anupam Varshney et.al[4]. They have appropriately defined Graphology as the scientific scheme of identifying, recognizing and distinguish behaviour over the movements and patterns inform with the handwriting. In Graphology, process of handwriting is analysed through graphically structured elements in order to derive information about the scripiter's personality. This paper basically checks the performance of a person on the basis of various parameters that is words from the baseline, pen pressure that a person applies, the slant of word. There are two parameters "f" and "I" letter which represents the input of artificial neural network which is used to predict activities of the writer. Some of the terms which are used in the paper are Zones, Connections, Slant, Spacing, Margin, Letter size, A large middle zone, Small middle zone, Upper zone extension, Speed, Clarity, Pressure. We are going to implement the concept of graphology to detect the behavior of a person keeping in mind its leverage in analysis of handwriting. Also, system will use CNN for image classification to provide better recognition of characters in the image and system to view 3D models. Also, the idea of this system is very similar to the project, so it is chosen as a reference.

VanCuong Kieu et.al[5] had described about a study that focuses on the complexity measure of an object which is based on the reported analysis of various details that are limited by contours of object. There may be holes or convexity evolution along the contour line. This study leads to a novel measure of the complexity in topology, measure of vacuity that quantifies the relation between objects and vaccum-ness or space. Based on the vacuity measure, proposal to define a novel shape descriptor and the associated dissimilarity measure is made. They can be applied in analysis of handwritten character and in recognition of objects in general. The experiments are then performed on a handwritten English character dataset and the object shapes dataset.

Anamika Sen et.al [6] implemented a system which analyses features and gives a detailed analysis employing the principles of Graphology. Extraction of six different English handwriting features are proposed and then further implemented. Processing of image was used for feature extraction using MATLAB.

In order to identify the character parameters in training stage, Behnam Fallah et.al [7] suggested application of The Minnesota Multiphase Personality Inventory, and for identification of one's personality from his/her handwriting, a neural network (MLP) to perform classification and a hidden

Markov model was used. MLP was used to identify the properties which are not similar to that handwriting and a hidden Markov model was used to classify those properties which are similar to the target handwriting. In this stage, in order to select the most relatable image to the image of the input context, the image of input context after fetching its property will be compared with every existing pattern in saved database. At the end, the test output regarding the personality of the image of the target handwritten text which is recommended by the system will be provided as the parameters of output character.

A method has been proposed by Champa H N et.al [8] to predict the personality of a person from the pen pressure, the baseline, the letter ‘t’, the slant of the writing and the lower loop of letter ‘y’ as found in an individual’s handwriting. These parameters are taken as inputs to a Rule-Base which provide output as the personality trait of the writer.

U V Marti et.al [9] presented a system for writer identification. From handwritten text, twelve sophisticated features are extracted which are used to recognize personality, based on their handwriting. The features extracted mainly correspond to visible characteristic traits of the writing, for example, the slant, the width and the height of the three important writing zones. Additionally, features based on the fractal behaviour of the text input are used which are correlated with the handwriting’s legibility.

Aim of Nuno Miguel Pessoa Saiitos et.al [10] is to describe a neural-network based system, inspired by the optical structure of human cortex area and its application to pattern recognition. Author adopted an unsupervised learning algorithm that is applied to each presented pattern. Insensitivity to spatial shift variance is the system's most relevant feature. The system's behaviour is based on the identification and decomposition of the relevant input image characteristics, that will be used) to achieve classification.

Table I summarizes highlights and observations of the related work discussed above.

TABLE I

Reference No	Highlights	Observations
[1]	Image segmentation based Handwritten character recognition system. and OpenCV for performing Image processing is used.	Helping Hand in Desktop publishing. Recognition of text from business card. It recognizes characteristics in letters and not in words.
[2]	Input images are compared with the knowledge base using SVM classifier. The hand writing style of cursive O is taken into consideration, and thus it defines the character traits of a person.	Used to observe a person, holistically and provide ways for improvising the personality traits or human behaviour using handwriting. System uses only single approach as classifier and thereby recognizing the characters
[3]	Main goal is to present a tool to observe small changes of slant in the text during work. Slant of word is found out using entropy value and projection of all the points in the text.	Provides the technique to determine the slant of text which depends on the speed of writing. Slant of a letter and not word is considered.
[4]	Handwriting analysis or Graphology as scientific	Complete study of Graphology can be

	scheme of identifying, recognizing and distinguishes behaviour over the movements and patterns inform with the handwriting.	used to develop a system. System makes use of Artificial Neural Networks (ANN) which cannot recognize characters in an image.
[5]	Study on the complexity measure of an object. To define novel shape descriptor and its associated measure of dissimilarity.	They can be applied in handwritten analysis of character and recognition of object in general.
[6]	A system that analyses features and gives a detailed analysis employing the principles of Graphology. Extraction of six different English handwriting features are proposed and then further implemented.	Image Processing was mainly chosen for feature extraction using MATLAB.
[7]	Identification the character parameters in training stage, The Minnesota Multiphase Personality Inventory. Selection of the similar images by comparing input image's context after eliciting its property with all the exiting patterns in dataset.	The test output regarding the personality of the image of the target handwritten text which is recommended by the system will be provided as the parameters of output character.
[8]	A method to predict the personality of a person from the lower loop of letter 'y', the baseline, the pen pressure, the slant of the writing and the letter 't' as found in an individual's handwriting.	The parameters that are the inputs to a Rule-Base, outputs the personality trait of the writer.
[9]	A system is presented for writer identification. From handwritten text, twelve features are extracted which are then used to recognize personality traits, based on their handwriting features.	The features extracted mainly correspond to visible characteristic traits of the writing, for example, the slant, the width and the height of the three important writing zones.
[10]	A neural-network based system, inspired by the optical structure of human cortex area and its application to pattern recognition.	The system's most relevant feature is its insensibility to spatial shift variance. The system's behaviour is based on the identification and decomposition that will be used to achieve classification.

From the paper surveyed and based on the observations listed in Table I, the problem statement can be framed as follows:

To implement a system to automatically identify user's behaviour which may be Analytical, Team player skill or Leadership skill through handwriting using hybridized approach to check employability of person.

The above problem statement can be achieved through following objectives:

1. To use OpenCV library for preprocessing of Image.
2. To implement Convolution Neural Networks (CNN) for Classification.
3. To implement Support Vector Machine (SVM) learning algorithm for classification.
4. To show behavior of user (Analytical, Team Player, Leadership skill etc.) using PyQt Graphical User Interface (GUI) application.
5. To test and validate implemented system with CNN and SVM.

IV. PROPOSED SYSTEM

The proposed system is illustrated in Figure no: 1.

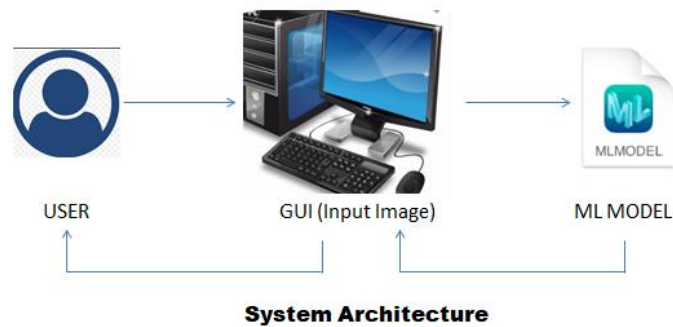


Figure no 1: System Architecture

System architecture is a conceptual model that defines the structural, behavioural and other views of a system. In this system, Admin has primary access to the system. He can Authenticate Login Process and can add and remove the user. He is also liable to modify and maintain the dataset. System pre-processes all the dataset and classifies the images based on the slope, slant, height, pen pressure of words through CNN. Then the User registers to the system; if he has already registered, he can login to system.

The figure illustrated below contains the system flow of the proposed system.

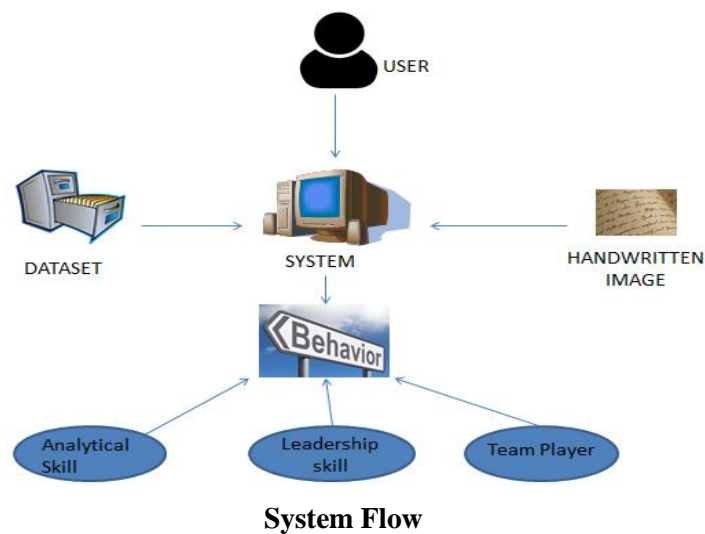


Figure no 2: System Flow

1. The whole architecture is made with the help of PyQT library used in python language. PyQT library gives all the necessary stuff related to GUI design. PyQT provides us display screen, buttons and so on. So, in this way, PyQT helps us in designing GUI.
2. After designing of GUI, another task is to authenticate valid user for operating application. To deal with this task, we are using MySQL database to store data of username and password and through this, user can authenticate easily.

3. Another task is to pre-process the input image which can be done by OpenCV library of python. By using this library, image is converted into grayscale image, contour image and smoothen image.
4. For implementation, CNN algorithm is applied for classification of images in terms of slant, slope, pen pressure and height for model training.
5. After providing an input image by user, SVM algorithm is applied for mathematical calculation to predict employability in a person.

V. RESULTS

The Employability Analysis System was created in PyQT Application using Python Language. The dataset used is “iamhandwritingdb” database from where approximately 500 images are selected and sorted into 4 categories namely upper-zone, lower-zone, slant-height and pen-pressure. This dataset is then preprocessed and fed to the CNN so that the model could be created. After the training the Neural network model, an accuracy of up to 98% was obtained. Images of trainable parameters and accuracy of the CNN model is as shown in Figure no: 3.a. and 3.b.:

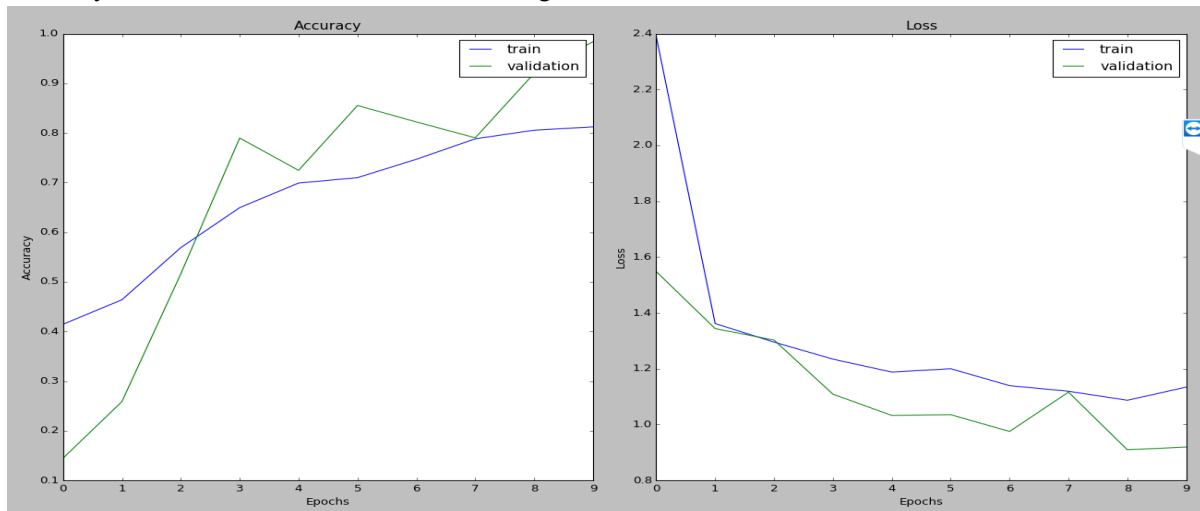


Figure no 3.a: Accuracy of CNN Model – About 98%

```

Layer (type)                Output Shape                Param #
-----
conv2d_1 (Conv2D)           (None, 198, 198, 32)       320
max_pooling2d_1 (MaxPooling2 (None, 99, 99, 32)         0
conv2d_2 (Conv2D)           (None, 97, 97, 32)         9248
max_pooling2d_2 (MaxPooling2 (None, 48, 48, 32)         0
dropout_1 (Dropout)         (None, 48, 48, 32)         0
flatten_1 (Flatten)         (None, 73728)              0
dense_1 (Dense)             (None, 150)                11059350
dropout_2 (Dropout)         (None, 150)                0
dense_2 (Dense)             (None, 150)                22650
dropout_3 (Dropout)         (None, 150)                0
dense_3 (Dense)             (None, 150)                22650
dense_4 (Dense)             (None, 4)                  604
-----
Total params: 11,114,822
Trainable params: 11,114,822
Non-trainable params: 0
Model has been saved.
    
```

Figure no 3.b.: Trainable parameters extracted from the Dataset

The input image is then passed on to SVM Classifier for feature extraction and the output is compared with the CNN model to predict the behaviour or personality traits in a person. The system showed about 98% accuracy in identifying the handwriting features and assigning the correct trait according to principles of Graphology. The image below in Figure no: 4.a. and 4.b. shows the GUI on which the behaviour is shown as result along with the other parametric values extracted from the respective input images.

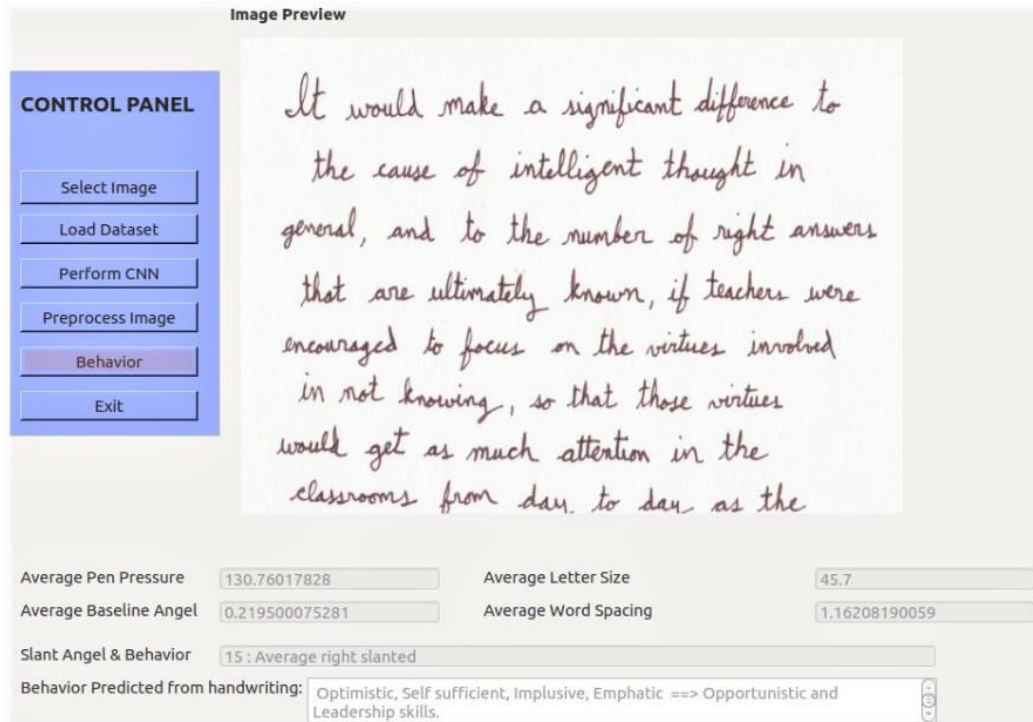


Figure no 4.a.: Employability Prediction Output 1 on GUI Application

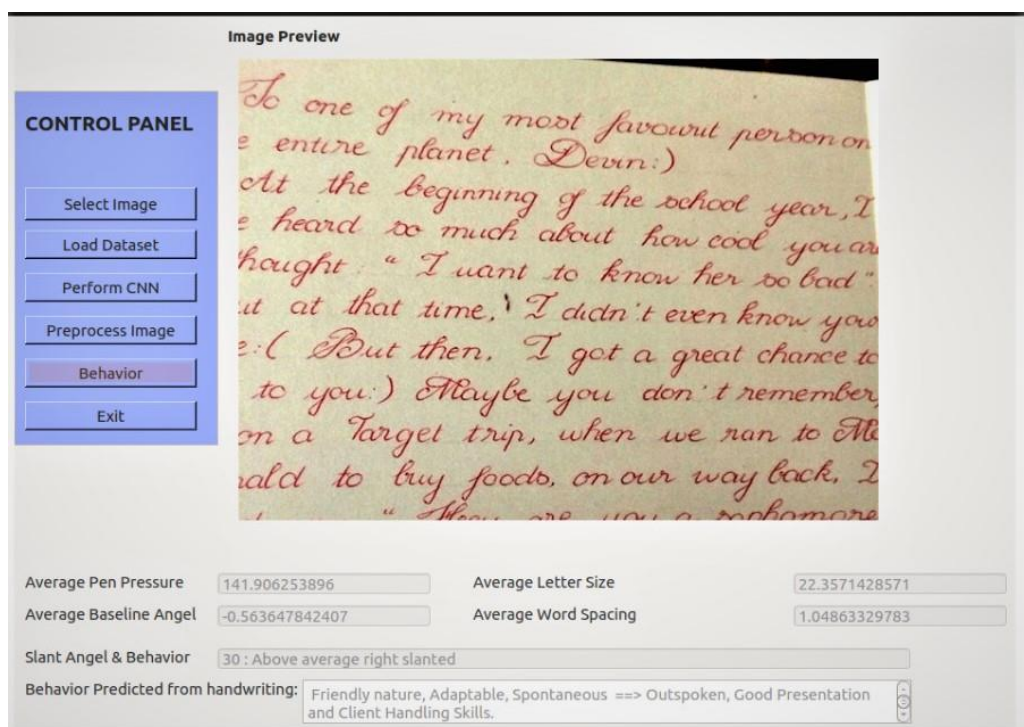


Figure no 4.b.: Employability Prediction Output 2 on GUI Application

VI. CONCLUSION

This paper illustrates Employability analysis which is to be done for finding employability skills in a person which may be analytical skills, leadership skills, team player skills etc. This can be done using hybridized approach which takes both classifier CNN and SVM in order to achieve the dream accuracy (about 98%).

VII. FUTURE SCOPE

The futuristic scope of the paper can be the usage of the system for various career counseling processes. This system is economical and can be easily deployed in professional recruitment processes. Employability of person can be predicted by using further more parameters apart from analytical skills, leadership skills, and team player skills etc. Graphology is very much in demand for written reports which determine forgery and fraud cases. Given the small number of graphologists and professionally certified handwriting experts in countries like India, their services are in high demand. However, these services are very expensive. This system aims to provide economical and reliable handwriting analysis. Personality traits in a person can easily be predicted and used in various fields such as medicine, psychology, human resources, forensics, elementary schools etc.

REFERENCES

- [1] Rohan Vaidya , Darshan Trivedi , Sagar Satra , Prof. Mrunalini Pimpale, “*Handwritten Character Recognition Using Deep Learning*” 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 2018.
- [2] Latypov, R., & Stolov, E. “*A New Method for Slant Calculation in Off-Line Handwriting Analysis*”. 2018 41st International Conference on Telecommunications and Signal Processing (TSP), 2018.
- [3] Asra, S., & Shubhangi, D. C. “*Human behavior recognition based on hand written cursives by SVM classifier*”. 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), 2017
- [4] Varshney, A., & Puri, S. “*A survey on human personality identification on the basis of handwriting using ANN*”. 2017 International Conference on Inventive Systems and Control (ICISC), 2017
- [5] Kieu, V.-C., Stutzmann, D., & Vincent, N. “*Vacuity Measure for Handwritten Character Analysis*”. 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), 2017
- [6] Sen, A., & Shah, H. “*Automated handwriting analysis system using principles of graphology and image processing*”. 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017.
- [7] Fallah, B., & Khotanlou, H. “*Identify human personality parameters based on handwriting using neural network*”. 2016 Artificial Intelligence and Robotics (IRANOPEN), 2016
- [8] Champa, H. N., & Ananda Kumar, K. R. “*Automated Human Behavior Prediction through Handwriting Analysis*”. 2010 First International Conference on Integrated Intelligent Computing, 2010.
- [9] Marti, U.-V., Messerli, R., & Bunke, H. (n.d.). “*Writer identification using text line-based features*”. Proceedings of Sixth International Conference on Document Analysis and Recognition, 2001.
- [10] Santos, N. M. P., & da Costa Oliverira, E. (n.d.). “*A neural network for handwritten pattern recognition*”. Proceedings of 8th Mediterranean Electrotechnical Conference on Industrial Applications in Power Systems, Computer Science and Telecommunications (MELECON 96), 1996.