Human Behavior Identification Based on Graphology Using Artificial Neural Network

Shalaka Prasad Deore[*]

Department of Computer Engineering, M.E.S. College of Engineering, 411001 Pune, S. P. Pune University, India

[*] Correspondence: Shalaka Prasad Deore (anudeore9@gmail.com)

Abstract: Handwriting reflects a person's true nature, phobias, emotional outbursts, honesty, defenses and many more characteristics. Analysis of handwriting, also known as graphology, is a science that uses the strokes and patterns disclosed by handwriting to identify, evaluate, and analyze personality. It is the study of the patterns and physical characteristics of handwriting to identify the author, indicate the author's psychological state while writing, or analyze personality traits. Traditionally, professionals also called graphologists predict the behavior of the writer by analyzing their handwriting, but this procedure is tedious and expensive. Therefore, this paper focuses on developing an application for personality identification that can predict behavioral characteristics directly using a computer without any human involvement. Most of the existing applications use English as the primary language to identify the personality trait of the writer however, our approach uses Devanagari scripts for prediction, thereby eliminating the language barrier. Our proposed method uses a machine learning approach to predict personality by analyzing Devanagari samples using Artificial Neural Network. We have created our own Devanagari word dataset. There are almost 4000 images which belong to 5 classes namely Introvert, Extrovert, Optimistic, Pessimistic and Stable mind-set. The testing accuracy achieved by the proposed method is 94.75%.

Keywords: Graphology; Artificial Neural Network; Machine learning; Feature extraction; Personality traits; Prediction

1. Introduction

When we write something, our brain guides our hand which is because of the connection between the reflex muscles of our brain and our hand. Every individual has his distinctive handwriting that helps us determine the personality traits, character and some psychological attributes about that individual. These personality traits are illustrated by a neurological brain pattern. Each written stroke or gesture unfolds a specific quality of the writer. The study of differentiating these strokes and patterns seen in one's handwriting and identifying the corresponding personality traits is known as Graphology. Graphology is a science that uses the strokes and patterns disclosed by handwriting to identify, evaluate, and analyze personality. It is the study of the patterns and physical characteristics of handwriting to identify the author, indicate the author's psychological state while writing, or analyze personality traits.

Handwriting reflects a person's true nature, including emotional outbursts, phobias, honesty, defenses, and more. The use of handwriting analysis to predict personality and behavior is beneficial and trustworthy. Handwriting also shows a person's possible stress-related transactions. Handwriting analyzers use graphology to determine attitudes, traits, moods, and sentiments. Handwriting analysis is widely used in various fields such as psychology, diagnostics, recruiting, counselling, identification of writers, forensic studies, etc. Graphologists typically predict the behavior of an individual with the help of a bit of paper however; the accuracy of the results depends on the talents of the analyst. This manual handwriting analysis process is extremely expensive and time-consuming. Hence, here developed a method for developing an application that can predict the writer's personality or behavior directly using a computer without the need for human interference.

We have proposed an automated application that can recognize, evaluate and understand a person's personality based on the strokes, patterns and curves that can be recognized by handwriting. The proposed application does

not rely on areas such as gender, age, and characteristics of the person's appearance to predict human behavior and personality. The proposed approach removes the language barrier by developing an application that anticipates personality based on the Devanagari scriptures using machine learning. Artificial Neural Network (ANN) is used to predict personality. It can predict hidden connections on hidden data after learning from the initial inputs and their connections. In this way, it makes the model more generalized and predicts hidden information from it, therefore, increasing the accuracy to a great extent. The proposed method is predicting five personality traits like Introvert, Extrovert, Optimistic, Pessimistic and Stable mind-set.

2. Related Work

Lemos et al. [1] have used Convolutional Neural Networks for their application. Their strategy includes 4 modules. The first module accepts the handwriting image through a website. For improved results, the second module is responsible for reducing noise, smoothing the image, and improving the contrast/sharpness of the image. They construct the dataset and train their CNN model in the third module, and they test the system in the fourth. They have 5 output labels. The likelihood of various output labels on an input image are observed via feature extraction. Chaudhari and Thakkar [2] have published a survey paper where they have considered around 15 handwriting features. Their research aims to promote the use of computer-assisted handwriting analysis for personality prediction using English handwritten documents. They had divided each handwriting feature into several categories that were further divided into sub-categories for precise prediction. Datasets used in other works were examined carefully. The paper also discusses the various fields in which an automated graphology system can be beneficial. After scanning the handwriting document and storing it in JEPG or PNG format Manoj Sachan, Hemlata and Singh [3] apply pre-processing techniques for enhancing the image quality. They divide the paragraphs of the document into lines, words and isolated characters in the segmentation unit. The essential and relevant handwriting features are extracted, and then classification is performed using classification algorithms, classifiers, or a rule-based approach.

Slant estimation, Baseline direction, arrangements and size of letters, margin, line and word spacing, and average upper the mid ratio are some of the attributes studied by Kumar et al. [4]. They collected samples from 100 different writers and used K-NN classification schemes. Djamal et al. [5] have extensively worked on signatures. They have proposed a method to identify the writer's behavior by analyzing his or her signature. The signature is segmented in three stages namely vertical, horizontal and line segmentation. ANN & multi-structure algorithm was used. A detailed study of twelve handwriting characteristics was considered by Varshney and Puri [6] in this survey paper. The main goal was to raise awareness about the importance of pursuing more efforts in this area. By thoroughly examining the handwriting attributes found in unique ways in English alphabets are written, Ghosh et al. [7] laid the hypothesis for three categories of human behavior namely negative, positive social behavior and personal behavior. CNN and SVM classifiers were used. Chanchlani et al. [8] created five labels for their dataset namely Persistent, Excitable, Criminal Intent, Narcissistic, Honest and self-centred. TensorFlow and OpenCV were used. Dataset was prepared for giving input to CNN. 20% of which was used for validation and the remaining was used for training.

Sen and Shah [9] got 95% accurate results by testing their system on 75 handwriting samples which were collected from people between the age group of 20 to 40 years old only. The system was developed in MATLAB GUIDE. Prashar [10] used feed-forward neural network that was created in MATLAB. After passing the values to ANN and all the computations, the Regression graph was plotted between actual and desired output. A dotted line was drawn where the actual output was exactly equal to the desired output. Salankara and Ghosh [11] made use of the renowned IAM Handwriting Database. The handwriting document is segmented into individual lines. The image processing software used by them was MATLAB. They used version R2014a(32bit). Using image processing tools [12] Feature Vector Matrix was generated of 100 samples of handwriting. After training the model with these 100 samples, the Feature Vector Matrix for the unknown sample is created. Using the similarity matrix method its similarity with the trained dataset would be calculated. Further, using the K-NN classifier the appropriate class of the unknown sample is identified. Rajani Suryavanshi, Aayush Agarwal, Shreya Kodolikar, Prachi Joshi and Ajinkya Dhavale implemented an incremental machine learning model where every new sample is added to the class to which it belongs in the training dataset.

Pen pressure, the stem of letter 't' and baseline were studied by Champa and AnandaKumar [13]. Methods like grey-level thresholding, template matching and polygonalization were used for extraction of respective features and ANN was used for prediction. Dhumal et al. [14] made use of SVM Classifier and Gabor filters. The features were extracted using a multi-channel Gabor filtering approach (MGF). A weighted Euclidean distance classifier was utilized to determine the writer's personality, and ten writers were requested to compose 10 paragraphs each using different pens. This way 100 samples were used for training purposes. Dang and Kumar [15] structured the paper into two sections, with the signature focusing on three features and the application style of the letters digit area. Performance is assessed using a Back Propagation Neural Network (BPNN) to calculate mean square error in this study. The neural network is used to analyze human behavior based on signature. Based on literature survey there is need of predicting person's behavioural characteristics. It has many significant applications like

psychology, medical diagnosis, recruitment of staff, career counselling, writer identification, forensics studies, etc. Mainly our contribution in this research is: Creation of new Devanagari word dataset and implementing effective personality traits prediction system using ANN.

3. Properties of Devnagari Script

India is the land of variety and so that we found several languages and cultures in one place. In India Devanagari is the most commonly used script. It consists of more than 120 languages making it the most adapted writing system [16]. Devanagari serves as the base language to many languages like Marathi, Hindi, Sanskrit, Gujarati and Nepali. Devanagari has evolved over a period of two thousand years and is highly continuous in nature. It is originated from Brahmi script. Devanagari represents phonetic script; this means each character is pronounced exactly the way it is written. Leftward to rightward direction is followed for writing Devanagari. Each word is separated by a line exist at the top called as "Shirorekha". Devanagari is an alphasyllabary writing system, such systems are based on consonants in which vowels are requisite but yet secondary. It is a syllabic script; it means Devanagari characters are written as combinations of vowels and consonants where a vowel can be added to the consonant in the form of character or in a form of modifier. The Devanagari script consists of 36 consonants, 12 vowels, 12 vowel modifiers and 10 digits as shown in Figure 1.
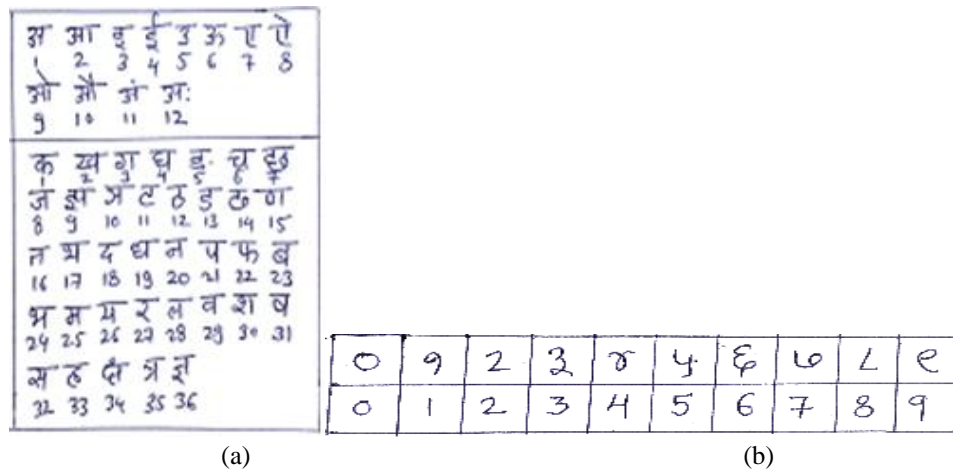

(a)   (b)

Figure 1. Handwritten samples: (a) Vowel & consonant characters; (b) Digits

4. Proposed Methodology

The manual technique of handwriting analysis is extremely expensive and time-intensive. Therefore, we have developed an automated application that predicts the personality of the writer with the help of a computer without any human interference. An automatic Handwriting Analysis Framework is introduced in the proposed approach, which extracts distinctive handwriting features. The related personality traits are seen based on these features. In an integrated GUI environment, the framework is modified.
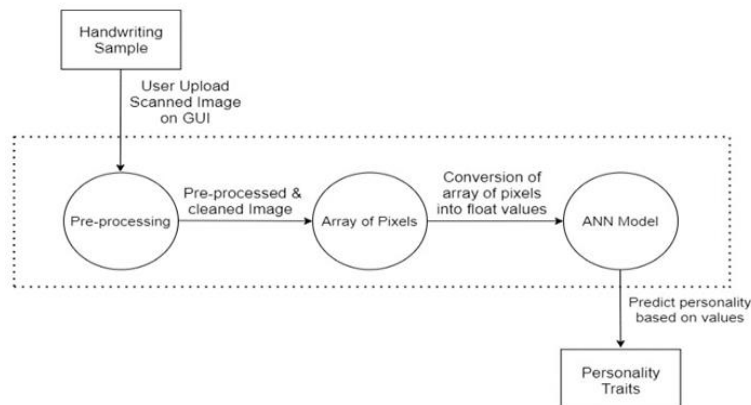

Figure 2. Block diagram

The proposed method predicts the results based on Graphology rules. The proposed technique extracts structural features like baseline, margins, letter size, slant, and word spacing to predict the personality dependent on Graphological rules. To eliminate the language barrier the proposed method focuses on Devanagari Scripts.

Initially, users will upload the handwriting sample on the GUI (Website), then this sample will be cleaned and pre-processed by applying various techniques on it. Further, this sample will be converted into an array of pixels and different features will be extracted from it and passed to the ANN Model as input and classes of personality traits will be predicted as output which will then be displayed on the same GUI as shown in Figure 2.

4.1 User Interaction with GUI

In this primary module, the user's handwriting sample image is acquired from a website. The user uploads a handwriting sample to the website, and after that, the image is sent to the application for pre-processing.

4.2 Pre-processing

In the pre-processing module, the handwriting sample is converted using various techniques to remove noise like grains and make the sample smooth for the extraction of features. Pre-processing is the method of performing operations on the input sample and also upgrading the quality of the image for further processing. It is carried out using techniques like resizing, binarization and noise removal.

a. Resize: Resizing is the process of changing the size of your image without cropping anything out. The base size is set for resizing all the input samples into the same resolution.
b. Noise Removal: Noise removal techniques are applied to reduce the noise from the image andto improve the quality by smoothing the image. The median filter is used for noise removal fromthe input sample.
c. Binarization: Binarization is the method of converting any colored image into a binary image (i.e., black and white image). In the proposed method, we are using Otsu's Binarization method which is implemented using OpenCV.

The pre-processed handwriting sample is converted into an array of pixels which is further transformed into float values which are passed to feature extraction stage.

4.3 Feature Extraction

Based on the following handwriting features and their characteristics, we have identified and mapped corresponding personality traits by using graphology rules and created 5 different classes which will be the output of our application. The five classes are namely Introvert, Extrovert, Optimistic, Pessimistic and Stable Mindset. The Devanagari word sample of five classes are namely Introvert, Extrovert, Optimistic, Pessimistic and Stable Mind-set is depicted in Figure 3.

a. Baseline: The baseline might be ascending, descending, or straight. Baseline helps us to determine the inner outlay and the character of the writer. The individual with an ascending baseline is optimistic, enthusiastic, and cheery, stays busy and active, is excitable, and exhibits choleric conduct. The descending baseline indicates that the person is gloomy, has transient mental fatigue, and has issues in absorbing things. The straight baseline indicates that the individual has consistent external conduct, stableness, practicality, and discipline.
b. Margin: The amount of space left on the right or left side of the page, or at the top or bottom of the page is known as margin. Before beginning to write, the writer considers leaving some space on the piece of paper for margins. Margin variations include left margin, right margin, top margin, bottom margin and balanced margin. Margin can be used to characterize both the present and the past. The individual with the balanced margin reflects a calm and stable mindset. The top margin reflects personal happiness whereas the bottom margin reflects shy, feel foreign, depressed personality. The right margin indicates lack of confidence, fear of others and failure, worries about the future whereas the left margin indicates careless and inattentive behavior.
c. Slant: Slant of the handwriting is defined as the inclination of characters. It is calculated by the gradient of the down stroke of the letter. There are different types of slants namely, right slant, left slant and vertical or no slant. The right slant in the handwriting has the ability to express emotions and opinions, freedom of thought and pragmatic behavior. The left slant shows that the writer has difficulty in adapting to a new environment, likes to work alone and behind the scenes. Writers with no slant handwriting are independent and have the ability to control emotion in crises.
d. Letter Size: Letters can be large or small in size. The large letter size symbolizes an extroverted, socially confident individual who enjoys being the centre of attention. The small letter size symbolizes a timid or introverted attitude, as well as a studious and careful disposition.

e. Word Spacing: The space present between two consecutive words is referred to as word spacing. Word spacing might help you comprehend how the writer interacts with others. The writer's intelligence and closeness to others are conveyed through word spacing. Wide word spacing indicates excellent taste, exclusivity, elitism, discrimination, independence, pride, a clear mind, and the capacity to arrange one's work. Narrow word spacing indicates an inability to be alone, bad taste, friendliness, and obtrusiveness.
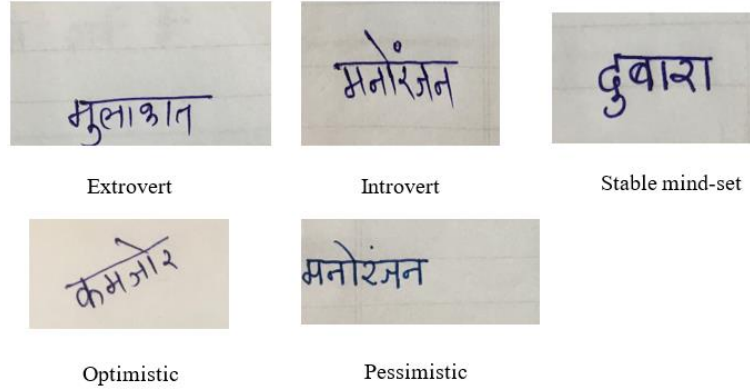

Figure 3. Sample images of five classes used for prediction

4.4 Proposed Personality Recognition Model Using ANN

The input layer, the hidden layer, and the output layer are the three types of layers in a neural network. The proposed study focuses on values derived for the baseline, slant, margin, word spacing, and letter size as inputs to the ANN that predicts personalities. Each ANN model output represents a specific behavioral characteristic found in the handwriting sample. We created our Devanagari word dataset with five classes representing five distinct personality traits: Introvert, Extrovert, Optimistic, Pessimistic, and Stable Mindset. The dataset is split into two sets at random, a training set and a testing set, at a 4:1 ratio. In addition, the training set is randomly divided into training and validation sets at random in a 3:1 ratio. A training dataset will be used to train the ANN model, and a testing dataset will be utilized to test the trained model. The output of the ANN model depicts many features or personality traits observed in handwriting. These personality traits will be displayed on the GUI. The **Algorithm** 1 is used to detect human personality based on handwriting.

Algorithm 1
Input: Devanagari handwritten samples
Output: Predicted personality class
Step 1: Read the file path of the dataset
Step 2: Append all the samples into a list
Step 3: Extract the target label set from the file path and store it in another list
Step 4: Apply various pre-processing operations as mentioned in section 3.2
Step 5: Convert the samples into an array of pixels
Step 6: Use Label Encoder to normalize the target labels into numerical labels
Step 7: Normalize the values from [0, 255] range to [0, 1] range
Step 8: Split the dataset into training, validation and testing set
Step 9: Pass the training set to the ANN Model for training
Step 10: Pass the validation set to tune the parameters of the model
Step 11: Test the trained model using the testing set to assess the performance of the model
Step 12: Plot graphs using the training and validation accuracy and loss values to analyze the performance of the model

5. Results and Discussions

Once the dataset is trained, the testing set is passed to the trained model for evaluating the performance of the model. The accuracy of the model is approximately 94% which implies that our model is predicting correctly on unseen handwriting samples. When the application is tested on unknown handwriting samples, it gives personality traits as output.

Figure 4 depicts that the accuracy increases rapidly in the first few epochs indicating that the network is learning

fast. Later, the curve flattens indicating that not too many epochs are required to train the model further, which shows that the model has started learning the data. In the Figure 5, the loss on the training and validation set decreases rapidly in the first few epochs. Afterward, the curve remains almost flat for multiple epochs which indicate that our model is generalizing well on unknown data. The Figure 6 compares the validation accuracy with the validation loss.
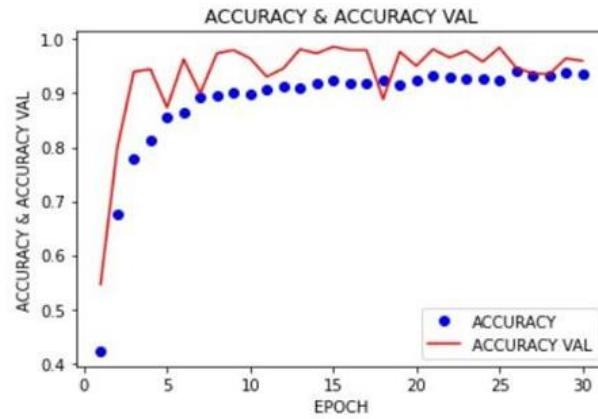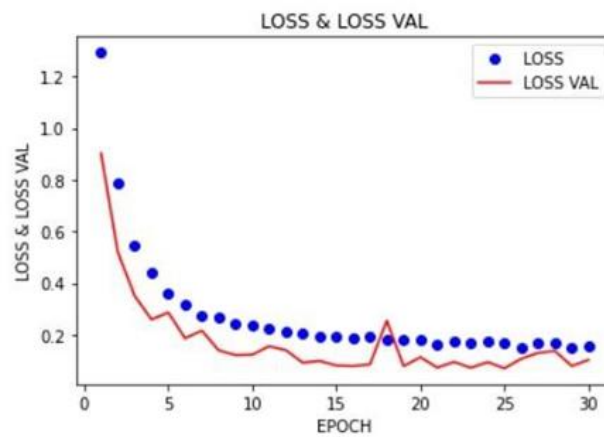

Figure 4. Training and validation accuracy graph

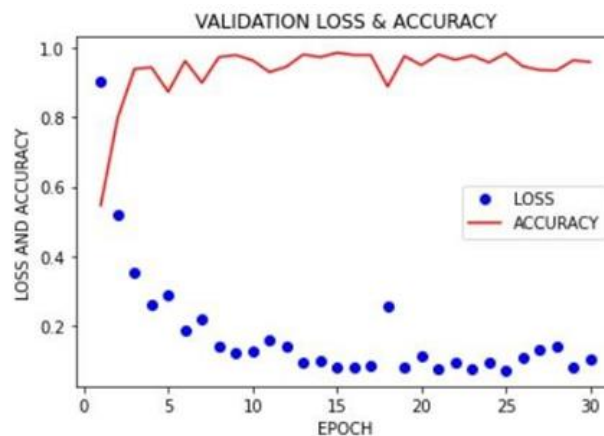**Figure 5.** Training and validation loss graph

Figure 6. Validation loss and accuracy graph

6. Conclusion and Future Work

Many systems are predicting personality and behavior by analyzing English handwriting samples. But then a question arises, what about people who cannot write or do not know how to write in English? By considering this, our method eliminates the language barrier by developing an application where we can predict personality by analyzing. Devanagari samples using artificial neural networks. Since Devanagari is the root of all languages, by working on the Devanagari script, we can open the door to predicting the behaviour of a writer in regional languages.

The traditional procedure used by graphologists is time-consuming and costly. Human efforts, on the other hand, are reduced with the help of automation. As a result, our strategy is not only efficient in terms of money but also in terms of time and human effort. Our method avoids human intervention and hence chances of error-prone situations are reduced. We have created our dataset which contains thousands of images consisting of single words or short phrases/sentences in Devanagari. We have divided the dataset into five personality classes. With the help of an automated graphology-based human behaviour prediction application, we are able to define the maturity of the writer without depending on features like gender, age, weight, hair, height, skin tone, etc. of a person. Certain handwriting attributes help us understand the tendency of the writer to the past or future, but graphology cannot foresee the future or justify the underlying cause, actions, or specific circumstances. In order to build more usable and functional applications, current techniques need to be enhanced in terms of feature extraction processes and precise evaluation of personality characteristics.

For future scope, we can advance the existing application to predict the writer's personality and behavior based on different regional languages. More scripts could be included so that the application can be universal and used worldwide. Our application considers 5 classes of personality which are dependent on 15 different features of handwriting. More handwriting features could be considered in the future. Our application predicts personality by examining only single words and short phrases, future scope could include prediction by analyzing paragraphs.

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Conflicts of Interest

The authors declare conflicts of interest.

References

[1] N. Lemos, K. Shah, R. Rade, and D. Shah, "Personality prediction based on handwriting using machine learning," In 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems, (CTEMS 2018), Belgaum, India, December 21-22, 2018, IEEE, pp. 207-207. https://doi.org/10.1109/CTEMS.2018.8769221.

[2] K. Chaudhari and A. Thakkar, "Survey on handwriting-based personality trait identification," *Expert Syst. Appl.,* vol. 124, pp. 282-308, 2019. https://doi.org/10.1016/j.eswa.2019.01.028.

[3] S. Hemlata and S. Singh, "Personality detection using handwriting analysis: Review," In Seventh International Conference on Advances in Computing, Electronics and Communication, Kuala Lumpur, Malaysia, August 18-19, 2018, SEEK, pp. 85-89. https://doi.org/10.15224/978-1-63248-157-3-33.

[4] R. Kumar, K. Ravulakollu, and R. Bhat, "Fuzzy-membership based writer identification from handwritten Devnagari script," *J. Inf. Proces. Syst.*, vol. 13, pp. 893-913, 2017. https://doi.org/10.3745/JIPS.02.0018.

[5] E. Djamal, R. Darmawati, and S. Ramdlan, "Application image processing to predict personality based on structure of handwriting and signature," In International Conference on Computer, Control, Informatics and Its Applications, (IC3INA 2013), Jakarta, Indonesia, November 19-21, 2013, IEEE, pp. 163-168. https://doi.org/10.1109/IC3INA.2013.6819167.

[6] A. Varshney and S. Puri, "A survey on human personality identification on the basis of handwriting using ANN," In International Conference on Inventive Systems and Control, (ICISC 2017), Coimbatore, India, January 19-20, 2017, IEEE, pp. 1-6. https://doi.org/10.1109/ICISC.2017.8068634.

[7] A. Ghosh, P. Shivakumara, P. Roy, U. Pal, and T. Lu, "Graphology based handwritten character analysis for human behaviour identification," *CAAI T. Intell. Techno.*, vol. 5, no. 1, pp. 55-65, 2020. https://doi.org/10.1049/trit.2019.0051.

[8] A. Chanchlani, P. Kharade, R. Kapase, S. Janvalka, and A. Jaitly, "Predicting human behavior through handwriting," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 5, pp. 624-628, 2017. https://doi.org/10.22214/ijraset.2017.10093.

[9] A. Sen and H. Shah, "Automated handwriting analysis system using principles of graphology and image processing," In International Conference on Innovations in Information, Embedded and Communication Systems, (ICIIECS), Coimbatore, India, March 17-18, 2017, IEEE, pp. 1-6. https://doi.org/10.1109/ICIIECS.2017.8276061.

[10] D. Prashar, "Behavior prediction through handwriting analysis," *IJCST*, vol. 3, no. 2, pp. 520-523, 2012.

[11] M. Salankara and I. D. Ghosh, "Feature extraction from text images to study individuality of handwriting," In 4th International Conference on Computational Intelligence and Networks, (CINE 2020), Kolkata, India, February 27-29, 2020, IEEE, pp. 1-6. https://doi.org/10.1109/CINE48825.2020.234387.

[12] P. M. Joshi, A. Agarwal, A. Dhavale, R. Suryavanshi, and S. Kodolikar, "Handwriting analysis for detection of personality traits using machine learning approach," *Int. J. Comput. Appl.*, vol. 130, pp. 40-45, 2015. https://doi.org/10.5120/ijca2015907189.

[13] H. N. Champa and K. R. AnandaKumar, "Artificial neural network for human behavior prediction through handwriting analysis," *Int. J. Comput. Appl.*, vol. 2, no. 2, pp. 36-41, 2010. http://dx.doi.org/10.5120/629-878.

[14] Y. Dhumal, V. Sawant, A. Surwase, and P. Sonnad, "Human behaviour recognition and person identification by handwriting analysis using neural network," *Int. J. Electron. Comput. Sci. Eng.*, vol. 5, no. 3, pp. 2454-1222, 2018.

[15] S. Dang and M. Kumar, "Handwriting analysis of human behaviour based on neural network," *Int. J. Adv. Res. Comput. Sci. Softw. Engi.*, vol. 4, no. 9, 2014.

[16] P. K. Singh, R. Sarkar, and M. Nasipuri, "Offline Script Identification from multilingual Indic-script documents: a state-of-the-art," *Comput. Sci. Rev.*, vol. 15, pp. 1-28, 2015.