IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 30, No. 1, January, 2025, pp. 203-211
Journal homepage: http://iieta.org/journals/isi

Identifying the Personality Traits Using Handwriting Recognition in a Real-Time Environment

Shalaka Deore*, Pranjal Kalokhe, Sneha Patil, Sayali Nagarkar, Vidhi Oswal

Department of Computer Engineering, Modern Education Society's Wadia College of Engineering, Pune 411001, India

Corresponding Author Email: shalakasonawane25@gmail.com

https://doi.org/10.18280/isi.300117

Received: 1 July 2024
Revised: 13 November 2024
Accepted: 7 December 2024
Available online: 25 January 2025

Keywords:
personality, traits, detection, CNN, handwriting, recognition, real-time, deep learning

ABSTRACT

Traditional methods of personality assessment, such as questionnaires and interviews, rely on self-reporting and subjective interpretation, which can be influenced by biases and social desirability. A handwriting-based approach offers an alternative method that provides non-verbal cues and unconscious expressions, supplementing traditional methods and potentially offering more objective insights, especially with an automated approach, into personality traits. Non-real-time methods take longer due to manual analysis, and are often subjective and prone to bias, while real-time analysis with a convolutional neural network (CNN) model provides instant results. Real-time tools would have no human intervention, are more convenient and easily accessible. Reducing human intervention through real-time analysis with a CNN model enhances the reliability, objectivity, scalability, and speed of the handwriting assessment process, providing a significant advantage over traditional methods. The proposed system focuses on the use of deep learning (DL) techniques to determine a person's personality by analyzing their handwriting. With the use of deep learning, human intervention is much less, thus the variance in the results is also much less which ultimately increases precision in the final predictions or results. Not only does the project predict personality traits through handwriting analysis, but also does that in a real-time environment. Steps such as preprocessing, feature extraction, and label classification are involved in the prediction process. CNN model has been used in the proposed system, for the final Personality Prediction. The CNN model for handwriting analysis must balance speed and accuracy through efficient architecture design and parameter optimization, minimizing computational complexity and inference time. To achieve this challenge two models (or versions) were built, the second having better data augmentation with respect to image sizing to be given to the CNN Network. Performance using the evaluation metrics was calculated with a testing efficiency of 0.71 in model 1 and 0.74 in model 2 and 82% accuracy was obtained on this real-time data.

1. INTRODUCTION

The resurgence of handwriting analysis in personality detection holds promise for advancing research, practice, and innovation across various fields. By leveraging the unique insights provided by handwriting analysis, professionals can enhance their understanding of human behavior, improve diagnostic accuracy, and inform evidence-based interventions. Furthermore, the integration of handwriting analysis into digital platforms and technologies enables real-time assessment and feedback, expanding its accessibility and usability in diverse settings. Personality detection using deep learning is an emerging and interesting field of interest for many, due to its uniqueness of identifying a person's character traits using just his/her handwriting. An individual's distinct identity is defined by a set of distinguishing qualities, habits, and cognitive patterns that make up their personality. It is what defines who the person is, how they would behave or interact in public, how they would work in different environments, cope with various situations, manage their tasks, etc. According to graphology (or handwriting analysis), how someone writes, how they shape their letters, and how they space them can indicate a lot about who they are [1, 2]. Study of handwriting is essential to forensic investigations, where it is used to identify and authenticate handwritten documents, detect forgeries, and link suspects to crime scenes. Graphology frequently shares its findings in courtrooms, employment offices, and police stations, despite the fact that many people view it as a pseudoscience. The graphologist Michelle Dresbold [3] mentioned that "Every single character represents something related to your mental state" and "Handwriting analysis is not like tea leaves, despite what some people believe. Your brain is causing your body to move". Traditionally graphology has been going on for centuries. Additionally, since 1895, more than 2,200 researchers have presented their findings in articles related to medicine, education, and psychology [4]. A study carried out in the United States by the National Pen organization claims that an individual's writing style can provide information about 5,000 distinct personality features depending on how they sign their name, space their letters, and even connect them. The importance of graphology as a subject started to spread in the 1930s. Dr. Ludwig Klages, a well-known German philosopher and the founder of contemporary graphology, produced several publications. In 1965, Francis T. Hillige devised a method for assessing characteristics in a sample of handwriting. Graphology has also seen many advancements since then, such as not just predicting whose handwriting it is, but also, predicting the age and gender of the person, predicting any mental illness or disorder or imbalance one might possibly possess, predicting the state of mind of the person while they were writing and predicting their personality traits [5, 6].

This proposed system, focuses on the latter advancement of graphology, with the help of technology, namely deep learning (DL). With the use of DL, Human intervention is much less, thus the variance in the result is also much less which ultimately increases precision in the final predictions or results. Not only does our project predict personality traits through handwriting analysis, but it also does that in a real-time environment. Deep learning has been widely used to recognize handwriting [7, 8]. Text is evaluated after it has been written in off-line handwriting recognition. The binary output of a character on a background is the only data that can be studied. Even though switching to a digital stylus for writing now provides more details like pen stroke, pressure, and writing speed, offline solutions are still required when online is unavailable. For historical records, archives, or the widespread digitization of handwritten forms, it is especially important. Real-time tools for finding out the Personality Traits of a person, on the basis of their Handwriting, were not available. Real-time tools would make it much easier and more convenient for the person to know their personality quickly and without much hassle. Such a tool would have no human intervention. Also, inefficiencies caused by image inputs for such tools would no longer be a problem. Usually for someone to know their personality traits, they have to answer a questionnaire, but with this tool, just with the simple handwriting input, one could know their traits, in real-time.

Higher extroversion was linked to a less logical decision-making style, but higher conscientiousness and agreeableness were significantly linked to a more logical decision-making style. While extroversion and openness to new experiences were strongly linked with a higher intuition style, higher conscientiousness, and agreeableness were highly associated with a lower intuition style. A more dependent decision-making style was significantly associated with greater levels of conscientiousness and agreeableness, whereas a less reliant decision-making style was strongly associated with higher levels of openness to new experiences. More agreeableness, conscientiousness, and neuroticism were significantly associated with a less spontaneous decision-making style [9]. When humans are taught how to write, they don't continue to write the same way all their lives. As time progresses, an individual’s handwriting keeps changing. This is due to the change in behavior or personality of an individual. Since handwriting, which is as distinctive as DNA, is the pattern of our psychology showed in symbols on paper. When you get to know someone's handwriting well enough, you can identify their writing just like you would a famous painting or photo. Graphology is based on the idea that each person's handwriting has a distinct personality that is entirely reflective of their individual personality. Therefore, in particular, the style and multi-modality highlight the significance and standing of graphology as a linguistic level of study [10]. Additionally, it is observed that handwriting based on graphology indicates an individual's inner feelings, even though these traits are not obvious in their actions [11].

By examining handwriting features psychologists can gain insights into clients' emotional states, cognitive patterns, and underlying psychological dynamics. This information can inform therapeutic interventions and treatment strategies. The purpose of Graphology is to identify various people from all walks of life, by just analyzing, observing, and studying their handwriting. Due to its extensive applicability in a variety of sectors, including education, psychology, medicine, criminal detection, marriage counseling, commerce, and recruitment, graphology-based behavioral analysis has become more and more popular in recent years [12]. Graphology lets us analyze a personality without considering extra factors such as lifestyle, appearance, social status, physical features, etc. Graphology focuses on handwriting features such as the baseline of the handwriting, the stroke of handwriting, the slant of letters and the words overall, the pressure of the pen, the word spacing, letter spacing, and overall sentence spacing and sometimes also paragraph spacing of the writing. A lot of methods are present for handwriting analysis but are often proven to be biased as it has human intervention, which introduces bias and variation in the results. Conventional personality assessment techniques may not be able to process results instantly and place a significant emphasis on subjective interpretation, which could introduce bias into the analysis. Furthermore, a lot of preprocessing and manual input is often needed for these approaches, which limits their scalability and accessibility. By utilizing a CNN model for real-time handwriting analysis, our suggested system seeks to overcome these drawbacks by offering objective insights into personality traits with little assistance from humans and efficient processing.

There are techniques that use a signature for identification in order to analyze behaviors including feelings, emotions, and human identifications [13, 14]. But unlike graphology, which takes into account a portion of a character, these techniques take into account the entire signature when making predictions. Furthermore, research on aesthetic analysis identifies human behaviors like beautiful, non-beautiful, great, and poor writing using handwritten letters or image attributes [15, 16]. These techniques are restricted to particular sorts of behavior, though. Techniques for evaluating a person's personality, including their interests, attitudes, and relationships with their family and community, have been offered [17, 18]. These techniques analyze user-typed texts; they do not analyze handwriting documents. Generally speaking, the techniques call for full-text words in order to evaluate personality. Therefore, graphology-based handwriting analysis becomes vital to forecast personal behaviors that may be imperceptible to the human eye [19]. The Big Five factor [20] is a popular model for characterizing a person's personality. Five fundamental personality traits serve as its foundation. Additionally, authors created the first dataset in the literature that includes both specified and arbitrary texts and correlates the Big Five personality handwriting types gathered from 128 people [21]. Alamsyah et. al. explained in his research three important features size, slant-ness and character stroke which is well-known features for any type of handwriting [22]. They had implemented CNN model to recognized Big five personality traits and achieved 80.88% of accuracy. The systematic review is performed between the handwritten writing and the main five personality traits. Various features for each category are explained by Elngar et al. [23]. Using deep learning CNN model the prediction is performed. The finding suggested the limitation of the diverse handwritten data samples and limited research in other regional languages for personality recognition.

Although behavioral assessments and personal questionnaires for self-reporting are the fundamental components of conventional personality assessment techniques, they are time-consuming, biased, and frequently unreliable. Moreover, the complicated and varied features of an individual's personality are often not well communicated by these methods. On the other hand, handwriting analysis offers a unique opportunity to gain an understanding of personality traits; yet, manual analysis is labor-intensive and requires ability. Thus, this research's primary goal is to develop a computerized system that can recognize traits of personality from handwriting with greater accuracy than existing methods. It improves the handwriting assessment process's speed, scalability, objectivity, and reliability. Based on the literature survey conducted for the proposed system, the conclusions were derived for the features of the 5 personality predictions as shown in Figure 1 using the CNN model. CNN is very effective for handling images, identifying universal traits, and sorting through them as compared to other DL models like RNN. Mainly CNN models are simple to design and our data is in image format and not required time series data like RNN. Hence our proposed system is implemented using CNN.

Personality Label	Attributes/Features	Handwriting Sample
Neuroticism	**Baseline** - Level **Slant** - Moderate Right or Straight **Space** - Narrow	Don't blame the guide
Openness	**Baseline** - Ascending **Slant** - Moderate Right	Earth atmosphere mostly consist of nitrogen and oxygen
Extroversion	**Baseline** - Ascending **Slant** - Extreme Right **Space** - Normal	To develop industry
Agreeableness	**Baseline** - Ascending **Slant** - Moderate or Extreme Left **Space** - Normal	I am terribly hungry.
Consciousness	**Baseline** - Descending **Slant** - Moderate Right or Straight	Earth atmosphere mostly consists of nitrogen

Figure 1. The big five personalities and their features based on the literature survey

Further in the paper, Dataset Creation, Network Architecture, Execution Details, Comparative Analysis, and Results for the Proposed System have been discussed. The framework of the paper is as follows: Section 2 Details about Dataset Corpus Creation Section 3 Discussion about the Proposed System, the Preprocessing details and CNN Architecture Section 4 Experiments Results and Discussion contains Model Training details, Execution Details, and Comparative Analysis and Section 5 Conclusion.

2. DATASET CREATION

For the prediction of personality traits using handwriting, we collected a new dataset by asking individuals of different ages to write English sentences. All handwritten samples are collected in the English language. The age range for this group is limited to those aged 18 to 50. Hence, they are given a piece of paper on which to write separate handwritten sentences with a ballpoint pen. In order to promote diversity within the dataset, we intentionally avoided imposing any limitations on the quality of paper, ball pen, or writing style during the data collection process. After that, each document is individually scanned with a scanner to produce a dataset.

The database includes a distinct set of handwriting patterns for each of the five personality traits: extroversion, agreeableness, openness, conscientiousness, and neuroticism. The total dataset size is 2264 images, with around 450 images per personality trait. Each scanned image was labeled manually by its corresponding personality trait. The image size was 256 × 256 pixels, and each image was converted to grayscale.

The created dataset has the following features:

- Variation in writing style, ball pen, and quality of paper to ensure maximum diversity.
- There were no limitations imposed on the size, style, the text written, or placement of the written characters on the paper.
- Individuals like males and females aged between 18 and 50, encompassing students, professors, or professionals employed within academic institutions or college settings.

The dataset consists of 5 personality traits, and each trait has its own set of handwritten samples. Classification of images is based on features such as slant, baseline, and space. To ensure the reliability of the proposed model, we used both the newly generated dataset and traditional datasets to test our model.

3. PROPOSED SYSTEM

The suggested approach employs machine learning to analyze handwriting in order to predict personality. Essentially, it's an improvement on graphology. The use of machine learning reduces the amount of human intervention, which reduces the variance in findings and eventually improves the accuracy of the final predictions or results. Not only does the project predict personality traits through handwriting analysis, but it also does that in a real-time environment.

3.1 Preprocessing

Preprocessing is an important step in handwriting character recognition to prepare image data for model input. Preprocessing plays a crucial role, enabling the machine model to learn and extract the given features as output. As the CNN is fully connected layers, it needs all the input images in an array of the same size. Preprocessed images boost the speed of the model and work efficiently [24, 25]. In our system, Preprocessing on raw data or image data includes grey scaling, noise reduction, resizing, and background removal. The preprocessing procedure improves the quality of the visual input, allowing the model to understand and produce the best possible result.

3.1.1 Image background

As mentioned above, CNN takes the input images in an array of the same size; so, we pass the input image shown in Figure 2, on the white background canvas such that its aspect ratio will remain unchanged. By applying a white background, it reduces the unwanted noise and helps to enhance the network's ability to extract meaningful features as shown in Figure 3. Maintaining the aspect ratio is very important to avoid the deformations of features. Due to the whitening of the background, structures or patterns are boldly visible which will also improve the efficiency of feature extractions.

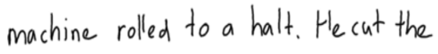

Figure 2. Raw input image

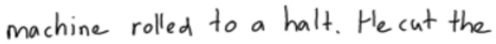

Figure 3. Image after pasting the raw input on a white canvas (background)

3.1.2 Gray scaling

Gray Scaling is an image conversion technique. It converts the colored image into a monochromatic image i.e., into a single-colored image. By applying grey scaling features to our image data it avoids the false classifications and reduces the color complexities to make it easier for the model training. Gray Scaling of image data does not reduce the size of the image as it takes the same space as the colored image as depicted in Figure 4.

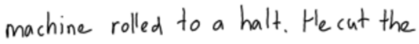

Figure 4. Input image after gray scaling

3.1.3 Noise reduction (median blur)

The process of reducing background noise from an image is known as noise reduction. Image noise reduction is very important as there are some factors that affect the quality of the image. Noise Reduction Algorithm alters the signal to some degree [26]. The Median Blur is used for noise reduction in our system. It reduces the background noise of images without affecting the original image or blurring it. The median filter is faster and provides superior outcomes in contrast to the other filters. The output of this is depicted in Figure 5.

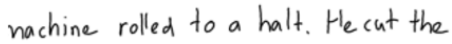

Figure 5. Input image after performing noise reduction

3.1.4 Resizing

Maintaining a consistent size of data before feeding it to the CNN results in fewer distortions of features and patterns. We have taken image resolution in the range of 256 × 256 pixels for training CNN. There are two ways of resizing the image: one is by scaling down the image and another is by cropping the image. Though there is an option of cropping the image, it poses risks of losing some important features or patterns. Rescaling the image will distort the image features and patterns. Hence resizing or rescaling the image is more suitable for our data inputs. Figure 6 shows the output of resizing.

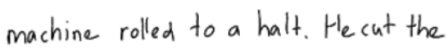

Figure 6. Input image after resizing

3.2 Analysis of features

The dataset includes a distinct set of handwritten patterns for each of the 5 traits of personality. Classification of these samples is based on the following features:

Baseline: A word's baseline. Baseline can be ascending, descending straight, etc.

Space: Spaces between words like normal, wide or narrow.

Slant: A writing way like moderate right/left, Extreme right/left, straight, etc.

Some samples are shown in Figure 1 based on each personality trait. The equilibrium between the ego and knowledge is shown by the baseline. The distance that the writer wishes to keep between themselves and society i.e., their boundaries is symbolized by the spaces between words. Different types of slanting information will explain the different personalities like extroverts, the openness of the person.

3.3 CNN architecture

3.3.1 Brief

A system composed of millions of artificial neurons connected to one another is called an artificial neural network (ANN). Each node is called a perceptron which computes and processes inputs and forwards the intermediate output to the next layer.

ANN comprises 3 layers, first is the input layer which is responsible for taking information as input from the environment and forwards it to the next layers which are called hidden layers. Hidden layers are sets of layers that are responsible for processing the input and extracting information from them. They process the information provided by input layers and forward that to the output layer. The output layer is responsible for giving out the output to the system. The output varies with the type of problem for which the model is built. In the case of regression problems and probabilities in the case of classification problems, CNN is a type of feed-forward Neural Network that needs and processes input in matrices-like structures like images. It is widely used in computer vision tasks, including image and speech classification, object detection, and image segmentation [27]. One of the important advantages of using a CNN model is its ability to learn features by itself. CNN comes with automatic feature extractions. Before CNN, researchers had to do manual feature engineering which used to be very time-consuming and resource intensive. But with the help of CNN feature extraction is done automatically [28].

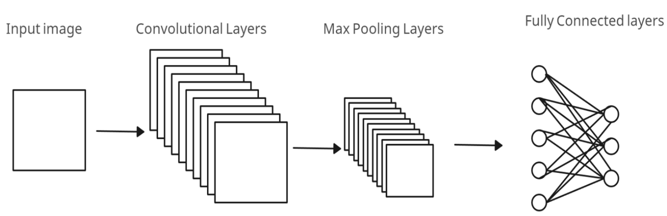

Figure 7. CNN architecture

CNN network consists of mainly three types of layers as shown in Figure 7:

Convolutional layer: This layer is the core of the CNN models. Here input data is processed through a tiny filter, often known as a kernel (typically a 3×3 matrix) [29].

Pooling layer: The major responsibility of a pooling layer is Dimension reductions mainly known as downsampling. It reduces the number of parameters of the input. Max pooling and average pooling are frequently used pooling techniques.

Fully-connected (FC) layer: Each neuron is connected to the subsequent layer neurons. These layers serve as a classifier to the CNN model. They perform the task of classifying the inputs based on the previously learned features and their filters. One of the important aspects of CNN models is the activation functions used. They introduced non-linearity to the model which enables it to learn and represent complex relationships of the data used [30]. Various activation functions can be used such as Sigmoid, Tanh, ReLU, SoftMax, etc. Choosing the most optimal activation function depends on the type of problem the model is built for. Various factors influence the choice of activation function like non-linearity, gradient Vanishing and Exploding problem, Computational Efficiency, output range, the problem domain of the model, etc. [31].

The next important task while fitting the CNN model is to choose the optimal optimizer. An optimizer is an algorithm or technique used to adjust a machine learning model's parameter as during training. An optimizer seeks to minimize the loss function or maximize the objective function of the model. Popular optimizers include Adam, Stochastic Gradient Descent (SGD), RMSProp, Adagrad, and others [32].

3.3.2 Proposed architecture

The proposed system consists of 4 pairs of Convolutional Layers and Max Pooling Layers followed by flattened and dense layers as shown in Figure 8. All the Convolutional Layers are with the kernel size 3×3 with ReLU as the activation function. The MaxPooling is with a pool size of 2×2.

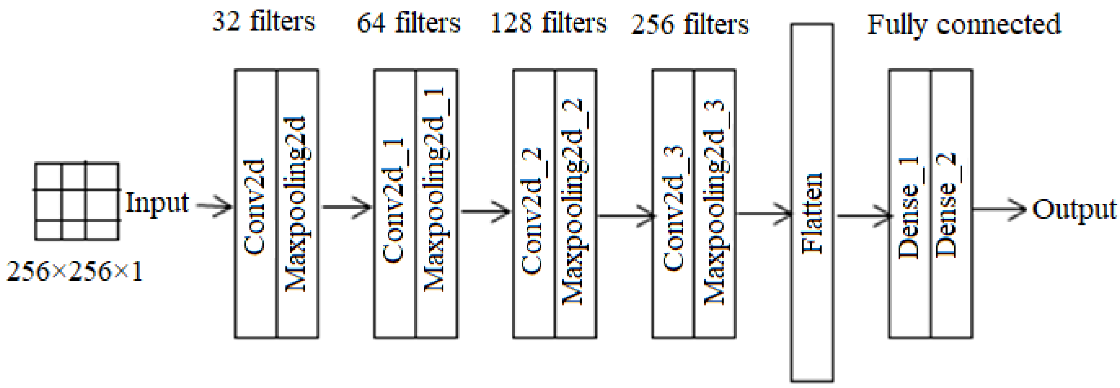

Figure 8. Proposed model architecture

The first layer of conv2D is with input size (None, 256, 256, 32) and output shape (None, 254, 254, 32). It consists of a total of 32 kernel filters. It is followed by max_pooling2d with an output shape of (None, 127, 127, 32). The second convolutional layer conv2d_1 (Conv2D) is with the output shape of (None, 125, 125, 64). It contains a total of 64 kernel filters. It is again followed by a second max pooling layer max_pooling2d_1 with an output shape of (None, 62, 62, 64). The third convolutional layer conv2d_2 (Conv2D) is with the output shape of (None, 60, 60, 128). It contains a total of 128 kernel filters. It is then followed by the third max pooling layer max_pooling2d_2 with an output size of (None, 30, 30, 128). The last as in the fourth convolutional layer consists of 256 kernel filters with an output shape of (None, 28, 28, 256) and it is followed by the fourth max pooling layer max_pooling2d_3 with an output shape of (None, 14, 14, 256).

After this, the input is sends to the Flatten layer with an output shape of (None, 50176), which converts the two-dimensional input data from the last max pooling layer to one dimensional as the dense layers also known as fully connected layers process upon one dimensional data.

Then the flatten output is then propagated through the first dense layer with an output shape of (None, 256). Dropout layer of 0.5 is added which indicates that during training, 50% of the input units will be randomly set to 0 at each update, which aids to avoid overfitting problems.

Last layer is the final dense layer with an output shape of (None, 5) which means it has output a vector of length 5. (As our model is built to classify input into one of the 5 classes, namely, Extroversion, Agreeableness, Conscientiousness, Openness and Neuroticism.). Softmax is used as the activation function as our model is built for multi-class classification. It takes a vector as an input and gives the probability distribution of various classes involved. The sum of all the probabilities is always one. The loss calculated here is Categorical Cross entropy which is usually used in case of multi-class classification problems, where each sample can belong to one and only one class. It calculates the difference between the expected probability and the true labels. It is given by Eq. (1).

$$\text{Cross Entropy}(y,y) = -\sum yi \log(yi) \qquad (1)$$

This loss function also penalizes the model more when the predicted probabilities deviate from the true labels. During training, the goal is to decrease the loss produced by adjusting the values of the weights and bias until optimal values are not found.

The Adam optimizer is used to execute the model. It keeps track of the rapidly dropping mean of the previous gradient (u_t), which represents the mean of the first moment, and the previous squared gradient (a_t), which represents the variance of the second moment The a_t and u_t are calculated is shown in Eqs. (2) and (3) [32].

$$a_t = \beta_1 a_{t-1} + (1-\beta_1)\, d_t \qquad (2)$$

$$u_t = \beta_2 u_{t-1} + (1-\beta_2)\, d_t^2 \qquad (3)$$

where, β represents the decaying rate of the gradients and dt is the gradient calculated at the time t. Adam optimizer uses Gradient descent (GD) algorithm as a base during training to modify training parameters and is also useful in self-learning [33]. Adam optimizer comes with many advantages, mainly having an adaptive learning rate, and high efficiency as it merges the features of AdaGrad and RMSProp. It is robust to noisy gradients and requires very little manual tuning [34].

4. EXPERIMENTS RESULTS AND DISCUSSIONS

4.1 Model training

We collected a new dataset of handwriting samples from individuals, each providing a sample of their handwriting on a blank sheet of paper. Each sample was meticulously labeled by hand with scores corresponding to the Big Five personality traits:

Openness: Usually showing a handwriting with ascending baseline and moderate right slant.

Conscientiousness: Descending slant majorly indicating introversion, emotional depth, or a tendency to reflect inwardly.

Extroversion: Ascending direction of the letters usually reflect a tendency towards enthusiasm, energy, and a forward-looking attitude.

Agreeableness: Handwriting with a leftward slant exhibits characteristics such as empathy, compassion, and cooperativeness.

Neuroticism: Display a narrow spacing as a reflection of their tendency towards anxiety, worry, or inner tension.

The dataset contains a total of 2264 handwriting samples, with around 450 samples for each personality trait. We preprocessed the images by first converting them to grayscale and then resizing them to a uniform size of [256×256]. Subsequently, we partitioned the dataset into training, testing and validation datasets, in 70:20:10 ratios, respectively. Real-time validation is also performed on user input collected in a real-time setting. A total of 50 inputs were collected, 10 from each of the big five personalities. 82% of accuracy was obtained on this real-time validation data. We used a deep convolutional neural network (DCNN) that was optimized for this purpose in order to predict people's personality attributes based on their handwriting.

From Table 1 we can say Model 2 demonstrated superior performance compared to Model 1 across multiple evaluation metrics, including higher accuracy on testing and training data, as well as a higher precision score. Additionally, Model 2 achieved a slightly higher accuracy on the testing dataset. Therefore, based on these comprehensive results, it can be concluded that Model 2 outperformed Model 1 in identifying personality traits using handwriting recognition in a real-time environment using a CNN model.

Table 1. Performance metrics comparison for Model 1 and Model 2

Performance Metrics	Model 1 (%)	Model 2 (%)
Accuracy (testing)	0.71	0.74
Accuracy (training)	0.94	0.95
Accuracy (validation)	0.78	0.79
Real-time dataset (validation)	0.80	0.82
Precision	0.70	0.74
Recall	0.70	0.70
F1-score	0.70	0.70
Average speed per epoch	3.6 s 68ms/step	4 s 68ms/step

Several convolutional layers and fully connected layers are incorporated in the model. The following parameters were used to train the proposed model:

Optimizer: Adam optimizer. Compared to other optimizers, it works better on our sequential model. It is a faster and more adaptive learning optimizer [32].

Learning rate: 0.001

Total epochs: 25

Dropout: 0.5

Training and testing batch size: 32 and 20 samples respectively.

4.2 Execution details

The prediction of personality traits of a person using handwriting was implemented using Python and several deep learning libraries including Keras and TensorFlow. The code was executed on Google Colaboratory. This is a free cloud service for creating apps using deep learning. For faster calculations, the environment offers a single 12 GB NVIDIA Tesla K80 GPU.

The dataset used in this project consisted of a collection of handwriting samples from various individuals. The dataset was preprocessed to extract relevant features from the handwriting samples and label them with corresponding personality traits. Following the aforementioned process, the data was divided into training, testing and validation sets, with 70% of the data allocated for training purposes and the remaining 20% reserved for testing and 10% for validation.

Once the training phase was completed, the model underwent evaluation on the testing set to measure its accuracy in predicting personality traits based on the provided handwriting samples. The model exhibited an impressive accuracy rate of 82%, indicating its effectiveness in accomplishing this task.

4.3 Results and discussions

The developed model for predicting personality traits using handwriting achieved a classification accuracy of 82%. The observed accuracy of 82% suggests a moderate correlation between handwriting features and personality traits. It is important to note that the model was trained using a dataset comprising 2264 handwriting samples obtained from individuals across various age groups and genders. This diverse dataset contributes to the model's ability to generalize and make predictions across different demographics. In terms of feature importance, the model revealed that the baseline of sentences was one of the features for predicting personality traits. This finding is consistent with previous studies that have shown a relationship between these features and personality traits. For example, a descending baseline is associated with conscientiousness, while a level baseline has been linked to neuroticism. It is worth noting that the model had difficulty in accurately predicting some personality traits such as agreeableness and conscientiousness. This suggests that handwriting may not be a reliable predictor for all personality traits. Real-time validation testing was performed on inputs taken in a real-time environment from users. A total 50 inputs were taken, 10 belonging to each class (The Big 5 Personalities). The Accuracy achieved on this set was 82%.

Figure 9. Neuroticism

Figure 10. Extraversion

As shown in Figures 9 and 10, the given inputs in real-time were predicted as neuroticism and extroversion as based on the literature survey table.

The overall findings of this study indicate that handwriting analysis holds promise in providing valuable insights into an individual's personality traits. Although the model's accuracy is not flawless, it still presents potential usefulness in various applications, including but not limited to recruitment and forensic analysis. Future research endeavors could concentrate on enlarging the dataset to encompass a wider range of individuals, as well as exploring additional handwriting features that can further enhance the prediction of personality traits. By doing so, we can improve the model's accuracy and broaden its applicability in practical scenarios.

4.4 Comparative analysis

Table 2 compares the effectiveness of our suggested CNN model with other methodologies in terms of its real-time validation/testing accuracy.

Table 2. Proposed model performance with other methods

Ref. No.	Methods	Accuracy (%)
[18]	MLP NN	76
[22]	CNN	80.88
[23]	PersonaNet CNN	70.73
[35]	ANN	70
[36]	SVM, KNN and AdaBoost classifiers	72.8
[37]	Set of rule-based classifiers	68
Proposed CNN model 2	CNN	82

The proposed CNN model demonstrated superior performance as shown in Table 2. The specific performance metrics and results are presented in the accompanying table, showcasing the CNN model's superiority over the other models in terms of performance.

5. CONCLUSIONS

In this approach, a system for evaluating personality traits of an individual based their handwritten samples has been put forth. These kinds of systems have already been the subject of past research. But they were lacking a real-time environment. This method is faster and more user-friendly thanks to real-time analysis. With image inputs, due to lighting or blur effects, the system is not very accurate in detection, hence with a real-time system, this drawback is overcome. In determining certain personality traits, handwriting recognition systems have demonstrated promising results. For example, the personality of a person can be identified by characteristics like baseline, slant, and spacing. 2264 handwriting samples were used to train the model, which were collected from people of all ages and genders. Following that, the results for predicting personality traits based on handwriting had a classification accuracy of 82%.

There were no real-time techniques for predicting a personality trait of an individuals based on their handwriting. Hence our system proposed real-time tools which makes it much simpler and more convenient for the person to swiftly and easily determine their personality. There would be no human involvement with such an instrument. Additionally, problems brought on by picture input inefficiencies for such technologies would no longer exist. With this technique, one might know their attributes in real-time with just a simple handwritten input as opposed to the usual questionnaire required to determine personality traits. As currently proposed, personality detection systems can only reveal some aspects of a person; they cannot reveal their true nature. To broaden the scope and uses of this study, research can be done on determining a person's age, gender, and any mental diseases or illnesses using a real-time environment. This would provide this application greater functional variety. The model architecture for this project might be improved in the future, along with the generalization and prediction accuracy, by looking at new handwriting traits and using more sophisticated methods like transfer learning or Ensemble learning.

In conclusion, our effort has been effective in creating a CNN model for personality prediction based on handwriting samples, providing insightful information about people's personalities. Applications in areas including career counseling, psychology research, forensic investigation, and personal growth are made possible by this. Despite the difficulties, the project's successes set a firm platform for future developments and as a tool for understanding human psychology based on his/her handwriting.

In future, we would like to implement our model for various regional languages and scripts. Incorporating additional scripts would enable the research to be utilized globally. More handwriting traits and other popular deep learning models based on multi-task can be explore in future to enhance the accuracy of the model. This multi-task DL model contain task specific layer to identify each individual traits and gives more benefits in recognition.

REFERENCES

[1] Taru, S., Mehta, V., Shinde, P., Deore, S. (2022). Graphology based human behavior and personality identification using artificial neural networks. In ICCCE 2021: Proceedings of the 4th International Conference on Communications and Cyber Physical Engineering, pp. 701-708. https://doi.org/10.1007/978-981-16-7985-8_73

[2] Deore, S.P. (2022). Human behavior identification based on graphology using artificial neural network. Acadlore Transactions on AI and Machine Learning, 1(2): 101-108. https://doi.org/10.56578/ataiml010204

[3] Picard, C. (2020). What the style of Your Handwriting Says About You, According to Graphology Experts. https://www.goodhousekeeping.com/life/g5057/handwriting-analysis/.

[4] Facts and History of Graphology. http://www.handwriting.com/facts/history.html.

[5] Dorosh, N., Fenenko, T. (2020). Recognition of MNIST handwritten digits and character set research. In Information Technologies in Metallurgy and Machine Building-ITMM, pp. 299-302. https://doi.org/10.34185/1991-7848.itmm.2020.01.032

[6] Gulli, A., Pal, S. (2017). Deep Learning with Keras. Packt Publishing Ltd.
[7] Impedovo, D., Pirlo, G. (2018). Dynamic handwriting analysis for the assessment of neurodegenerative diseases: A pattern recognition perspective. IEEE Reviews in Biomedical Engineering, 12: 209-220. https://doi.org/10.1109/RBME.2018.2840679
[8] Pal, K.K., Sudeep, K.S. (2016). Preprocessing for image classification by convolutional neural networks. In 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, pp. 1778-1781. https://doi.org/10.1109/RTEICT.2016.7808140
[9] El Othman, R., El Othman, R., Hallit, R., Obeid, S., Hallit, S. (2020). Personality traits, emotional intelligence and decision-making styles in Lebanese universities medical students. BMC Psychology, 8: 46. https://doi.org/10.1186/s40359-020-00406-4
[10] Jiménez, E.M.G. (2015). An introduction to graphology: Definition, theoretical background and levels of analysis. Miscelánea: A Journal of English and American Studies, (51): 71-85.
[11] Brewer, J.F. (1999). Graphology. Complementary Therapies in Nursing and Midwifery, 5(1): 6-14. https://doi.org/10.1016/S1353-6117(99)80065-1
[12] Tett, R.P., Palmer, C.A. (1997). The validity of handwriting elements in relation to self-report personality trait measures. Personality and Individual Differences, 22(1): 11-18. https://doi.org/10.1016/S0191-8869(96)00183-3
[13] Hafemann, L.G., Sabourin, R., Oliveira, L.S. (2016). Analyzing features learned for offline signature verification using deep CNNs. In 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, pp. 2989-2994. https://doi.org/10.1109/ICPR.2016.7900092
[14] Miguel-Hurtado, O., Guest, R., Stevenage, S.V., Neil, G.J. (2014). The relationship between handwritten signature production and personality traits. In IEEE International Joint Conference on Biometrics, Clearwater, FL, USA, pp. 1-8. https://doi.org/10.1109/BTAS.2014.6996245
[15] Gattupalli, V., Chandakkar, P.S., Li, B. (2016). A computational approach to relative aesthetics. In 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, pp. 2446-2451. https://doi.org/10.1109/ICPR.2016.7900003
[16] Majumdar, A., Krishnan, P., Jawahar, C.V. (2016). Visual aesthetic analysis for handwritten document images. In 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, pp. 423-428. https://doi.org/10.1109/ICFHR.2016.0085
[17] Kedar, S.V., Bormane, D.S. (2015). Automatic personality assessment: A systematic review. In 2015 International Conference on Information Processing (ICIP), Pune, India, pp. 326-331. https://doi.org/10.1109/INFOP.2015.7489402
[18] Fallah, B., Khotanlou, H. (2016). Identify human personality parameters based on handwriting using neural network. In 2016 Artificial Intelligence and Robotics (IRANOPEN), Qazvin, Iran, pp. 120-126. https://doi.org/10.1109/RIOS.2016.7529501
[19] Ghosh, S., Shivakumara, P., Roy, P., Pal, U., Lu, T. (2020). Graphology based handwritten character analysis for human behaviour identification. CAAI Transactions on Intelligence Technology, 5(1): 55-65. https://doi.org/10.1049/trit.2019.0051
[20] Costa, P.T., McCrae, R.R. (2008). The revised neo personality inventory (neo-pi-r). The SAGE Handbook of Personality Theory and Assessment, 2(2): 179-198.
[21] Gavrilescu, M., Vizireanu, N. (2018). Predicting the Big Five personality traits from handwriting. EURASIP Journal on Image and Video Processing, 2018: 57. https://doi.org/10.1186/s13640-018-0297-3
[22] Alamsyah, D., Widhiarso, W., Hasan, S. (2022). Handwriting analysis for personality trait features identification using CNN. In 2022 International Conference on Data Science and Its Applications (ICoDSA), Bandung, Indonesia, pp. 232-238. https://doi.org/10.1109/ICoDSA55874.2022.9862910
[23] Elngar, A.A., Jain, N., Sharma, D., Negi, H., Trehan, A., Srivastava, A. (2020). A deep learning based analysis of the big five personality traits from handwriting samples using image processing. Journal of Information Technology Management, 12: 3-35. https://doi.org/10.22059/jitm.2020.78884
[24] Kumar, G., Bhatia, P.K., Banger, I. (2013). Analytical review of preprocessing techniques for offline handwritten character recognition. International Journal of Advances in Engineering Sciences, 3(3): 14-22.
[25] Hashemi, M. (2019). Enlarging smaller images before inputting into convolutional neural network: Zero-padding vs. interpolation. Journal of Big Data, 6(1): 98. https://doi.org/10.1186/s40537-019-0263-7
[26] Zhu, Y., Huang, C. (2012). An improved median filtering algorithm for image noise reduction. Physics Procedia, 25: 609-616. https://doi.org/10.1016/j.phpro.2012.03.133
[27] Jaswal, D., Vishvanathan, S., Kp, S. (2014). Image classification using convolutional neural networks. International Journal of Scientific and Engineering Research, 5(6): 1661-1668. https://doi.org/10.14299/ijser.2014.06.002
[28] Jogin, M., Madhulika, M.S., Divya, G.D., Meghana, R.K., Apoorva, S. (2018). Feature extraction using convolution neural networks (CNN) and deep learning. In 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, pp. 2319-2323. https://doi.org/10.1109/RTEICT42901.2018.9012507
[29] Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., et al. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 8: 53. https://doi.org/10.1186/s40537-021-00444-8
[30] Nanni, L., Brahnam, S., Paci, M., Ghidoni, S. (2022). Comparison of different convolutional neural network activation functions and methods for building ensembles for small to midsize medical data sets. Sensors, 22(16): 6129. https://doi.org/10.3390/s22166129
[31] Sharma, S., Sharma, S., Athaiya, A. (2017). Activation functions in neural networks. International Journal of Engineering Applied Sciences and Technology, 6(12): 310-316.
[32] Poojary, R., Pai, A. (2019). Comparative study of model optimization techniques in fine-tuned CNN models. In

2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA), Ras Al Khaimah, United Arab Emirates, Ras Al Khaimah, United Arab Emirates, pp. 1-4. https://doi.org/10.1109/ICECTA48151.2019.8959681

[33] Cui, N. (2018). Applying gradient descent in convolutional neural networks. Journal of Physics: Conference Series, 1004: 012027. https://doi.org/10.1088/1742-6596/1004/1/012027

[34] Bera, S., Shrivastava, V.K. (2020). Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification. International Journal of Remote Sensing, 41(7): 2664-2683. https://doi.org/10.1080/01431161.2019.1694725

[35] Asra, S., Shubhangi, D.C. (2015). Personality trait identification using unconstrained cursive and mood invariant handwritten text. International Journal of Education and Management Engineering, 5(5): 20-31. https://doi.org/10.5815/ijeme.2015.05.0

[36] Chen, Z., Lin, T. (2017). Automatic personality identification using writing behaviours: An exploratory study. Behaviour & Information Technology, 36(8): 839-845. https://doi.org/10.1080/0144929X.2017.1304994

[37] Champa, H.N., AnandaKumar, K.R. (2010). Automated human behavior prediction through handwriting analysis. In 2010 First International Conference on Integrated Intelligent Computing, Bangalore, India, pp. 160-165. https://doi.org/10.1109/ICIIC.2010.29