A novel mobile application for personality assessment based on the five-factor model and graphology

Ahmed Remaida[1], Zineb Sabri[2], Benyoussef Abdellaoui[1], Chakir Fri[3], Yassine Lakhchaf[4], Younès El Bouzekri El Idrissi[1], Mohammed Amine Lafraxo[5,6], Aniss Moumen[1]

[1]Laboratory of Engineering Sciences, National School of Applied Sciences, Ibn Tofaïl University, Kenitra, Morocco
[2]Laboratory of Computer Science, Faculty of Sciences, Ibn Tofaïl University, Kenitra, Morocco
[3]Laboratory of Informatics and Information Science, Faculty of Sciences, Moulay Ismail University, Meknes, Morocco
[4]Laboratory of Engineering Science, Department of Interdisciplinary Computer Science and Physics, Faculty of Sciences and Technologies, Mohammed Ibn Abdallah University, Fez, Morocco
[5]Laboratory of Biology and Health, Faculty of Sciences, Ibn Tofaïl University, Kenitra, Morocco
[6]Higher Institute of Nursing Professions and Health Techniques, Oujda, Morocco

Article Info

Article history:

Received Mar 28, 2024
Revised Oct 26, 2024
Accepted Nov 11, 2024

Keywords:

Five factor model
Graphology
Handwriting analysis
Machine learning
Mobile application
Personality prediction

ABSTRACT

With the rising interest over the last decade, automated graphology has emerged as a promising filed of research, providing new insights on personality traits prediction on the basis of handwriting analysis. Although, few practical solutions to automate the extraction of handwriting features and personality prediction exist in the literature. This work aims to contribute to closing the gap in automated handwriting personality prediction by proposing a novel mobile application that uses robust feature extraction and machine learning models to predict big five personality traits. Our findings, based on high correlations between handwriting characteristics and personality traits, revealed convincing links. Notably, extraversion and extraversion have strong correlations with top margin feature, whereas agreeableness is expressed through line spacing. These findings emphasize the ability of automated graphology to properly interpret individual personalities. The proposed system achieved exceptional accuracy by using well known machine learning classifiers. The testing accuracy exceeded 92% in binary classification and 87% in multi-class case scenario, proving the adaptability and dependability of the system's architecture. Our Android app promises to provide users with unprecedented insights into their personalities, establishing a robust tool for psychological assessment and self-discovery.

Corresponding Author:

Ahmed Remaida
Laboratory of Engineering Sciences, National School of Applied Sciences, Ibn Tofaïl University
Kenitra 14000, Morocco
Email: ahmed.remaida@uit.ac.ma

1. INTRODUCTION

Understanding human personality, encompassing enduring patterns of thought, emotion, and behavior, has long been a primary objective in psychology [1]. Key theoretical frameworks like Raymond Cattell's 16 personality factors (16PF) [2], the five-factor model (FFM) or big five personality traits (OCEAN) [3], and the Myers Briggs type indicator (MBTI) [4] have contributed significantly to understanding individual differences. These models have been extensively used to examine how personality shapes various life domains, including social interactions, career decisions, and mental health outcomes. Handwriting, as a highly individualized mode of expression, has also been explored for its potential to reveal psychological characteristics. This study of handwriting as a reflection of personality known in the literature as graphology [5], has played an important role in this area, with applications in fields such as medical diagnosis, educational assessments, and employment selection [6]–[10].

Recent advancements in machine learning have further revived interest in handwriting analysis, automating the process and allowing for more robust personality predictions based on nuanced handwriting features that traditional methods might overlook [11], [12]. Despite these advancements, several unresolved problems persist. While various computational models have been proposed to predict personality traits from handwriting, the robustness of handwriting feature extraction techniques remains a challenge, as many models struggle to consistently capture subtle and complex patterns. Additionally, few studies have translated these models into practical applications for real-world use by professionals like graphologists, leaving room for improvement in creating accessible, reliable systems. Here in Table 1, we give a non exhaustive overview from the literature in constrast with handwriting personality prediction.

Table 1. Recent related works

Works	Summary	Number of features	Mobile application
[13]	A new three-layer method for identifying big five personality traits from handwriting was developed, validated using intrasubject and intersubject approaches. The latter yielded better accuracy rates, with 84.4% for neuroticism, extraversion, openness, and 77% for conscientiousness and agreeableness.	7	None
[14]	The study presents the largest handwriting dataset, analyzing 1,800 samples for detecting personalities based on graphological features. Using machine learning algorithms like support vector machine (SVM), Naïve Bayes (NB), and random forest (RF), the SVM achieved the highest prediction accuracy (97%).	5	None
[15]	The authors developed a new method for predicting personality traits from handwriting, using neo five-factor inventory-3 test results and handwriting samples. The accuracy ranged from 65% to 85%, with extraversion being the most accurate.	5	None
[16]	The authors propose a convolutional neural networks (CNNs)-based architecture for predicting personality from English and Hindi handwriting, despite low accuracy rates (average 43%), with plans to improve accuracy in future works.	None	None
[17]	The study developed a mobile application for identifying personality traits using graphology, specifically through handwriting analysis. Tested on 25 handwriting samples, the application extracts and analyses handwriting features, and achieves similar identification results to a graphology expert. The study demonstrates the feasibility of using a mobile application for handwriting analysis.	6	Android mobile application
[18]	The study aims to develop an Android app for handwriting analysis using deep learning techniques to predict extrovert or introvert MBTI personality types. The dataset consists of handwriting samples labeled with personality types obtained through an online MBTI test. The app uses a TensorFlow Lite MobileNet based on CNN to achieve an accuracy of 45% during training and testing. Authors suggested that more data can improve results.	None	Android mobile application

To address these gaps, we propose a novel mobile application for robust handwriting-based personality prediction. This application, designed for use by graphologists and other professionals, leverages machine learning algorithms to detect nuanced relationships between handwriting features and personality traits by offering an accessible practical tool grounded in scientific rigor. By this work we seek to advance the field of personality assessment and provide an innovative solution to previously unaddressed issues in handwriting analysis. The rest of this paper is organized as follows: the research method section outlines the study's approach and procedures, the results and discussion section presents key findings and situates them within the context of existing research, and the conclusion highlights the implications, limitations, and areas for future research.

2. METHOD

The main goal of this work is to developpe a mobile application capable of predicting the big five personality traits on the basis of a given handwriting sample. To achive this purpose we collected both big five inventory (BFI-10) and handwriting sample form each one of a 100 engineering students with their consent. Next, personality scores were calculated and handwriting features were extracted using features extraction techniques in order to be stored in the final dataset. The following phase consisted of analyzing the existing correlations between personality traits and the handwriting features, in addition to the training and testing of machine learning models to assess the predictability of personality traits on the basis of handwriting. Finally, a mobile application was developed to use both extraction techniques and trained models in order to predict personality scores from a given handwriting sample. The flowchart in Figure 1 illustrates the described phases in the research method of this work.

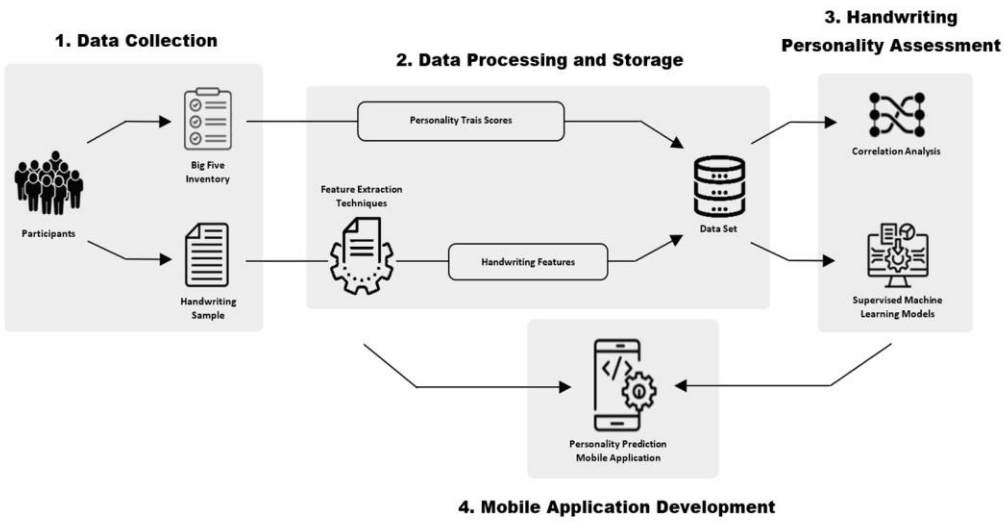


Figure 1. Flowchart of the research method

The experimental setup used for implementation and development is composed of a Lenovo ThinkPad laptop featuring an 8-core 2.30 GHz Intel i5 10th Gen CPU, 16 GB of RAM, an 8 GB Intel GPU, and 500 GB of storage. This configuration was utilized for both the feature extraction process and the training and testing of machine learning models. Python 3.7 was chosen due to its integration with the Scikit-Learn machine learning framework [19], OpenCV for feature extraction, and Flask for the RESTful API to routs requests through Ngrok tunneling. As for the mobile app development, the user interfaces were created in XML, and Java for handling event control, both within the Android Studio environment on the same setup.

2.1. Data collection

The first phase phase in the research method of this study consisted of collecting data from the participants where each one of them was asked to complete a personality questionnaire and to present a handwriting sample. For personality traits scoring, we used the five-factor model knwon as the big five personality traits inventory [20]. Recently, shorter versions with only 10 items were studied and validated in both English [21] and French [22] languages. The french version with only 10 items presented a suitable choice for faster personality assesement, therefor it was used in this study. As for the handwriting, we asked participants to rewrite the London letter model [23] as it is composed of alphabet letters in both uppercase and lowercase form, all numerals in different combinations and many other characteristics that made it suitable for the handwriting analysis. Further processing with an emphasis on characteristics such as margins, baseline, letter size, line spacing, and word spacing. A thorough comprehension of attributes can be achieved by integrating personality inventories with handwriting analysis, as suggested by existent literature [24].

2.2. Data processing and storage

2.2.1. Big five inventory

The next phase after data collection consisted of processing and storing both personality traits score and handwriting extracted feautres values. For personality traits scores, we used the formula presented in the paper presenting the BFI-10 [22]. Table 2 describes the score calculation for each one of the five personality dimensions.

Table 2. Big five personality traits inventory scoring

Personality traits	Item's number	Scoring formula
Extraversion	Q1 Q6	(Reversed (Q1)+Q6)/2
Agreeableness	Q2 Q7	(Q2+Reversed (Q7))/2
Conscientiousness	Q3 Q8	(Q3+Reversed (Q8))/2
Neuroticism	Q4 Q9	(Q4+Reversed (Q9))/2
Openness	Q5 Q10	(Q5+Reversed (Q10))/2

2.2.2. Handwriting samples

As for hadnwriting features values, we proposed and used robust feature extraction methods that implies pre-processing and optimizing handwriting samples. The significance of feature extraction in handwriting analysis is emphasized in the literature to enable successful predictive modeling [25]. In what follows, we present a detailed explanation of the feature extraction methodes that we proposed and used in this work.

A. Pre-processing

Grayscaling, thresholding and denoising are all tehniques used to enhance the image quality and make it suitable for the the next stpes in the feature extraction process. For this purpose, we used the OpenCV [26] library version in Python for reading image sample in RGB space and for conversion to Gray Scale. After that, comes thresholding with the Otsu’s technique [27] and denoising using median filter [28]. Here in Figure 2, we give the obtained results for a handwriting sample after the pre-processing steps, where Figure 2(a) represents the original handwriting sample, Figure 2(b) grayscaled sample, Figure 2(c) thresholded sample and Figure 2(d) as the denoised sample.

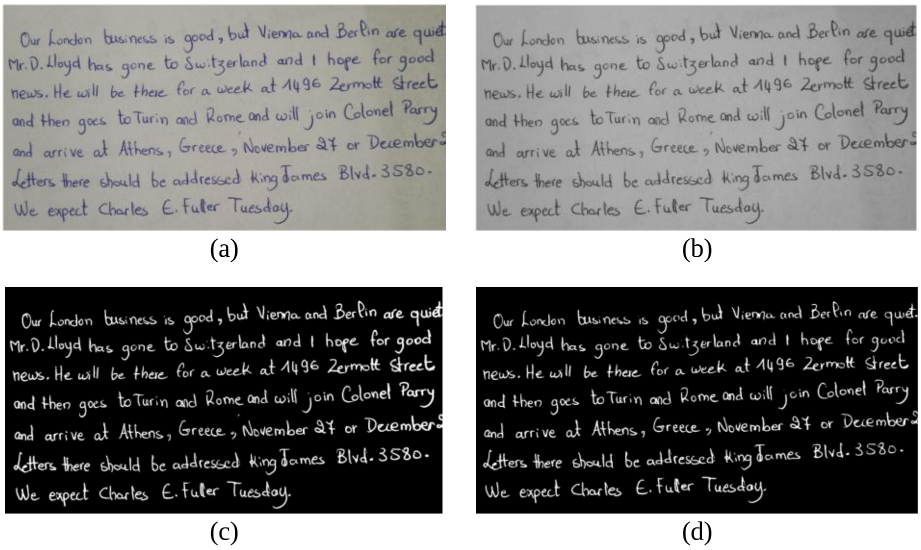


Figure 2. Pre-processing results for a handwriting sample; (a) original, (b) grayscaled, (c) threshed, and (d) filtered

B. Extracting baseline angle feature

For extracting baseline feature (angle of rotation), we once again used OpenCV libraires for both dilating, contouring and rotating lines on the denoised hadnwriting sample [29]. The contouring seeks curve joining of all the boundary points that have same color in order to draw bounding boxes, which allowed us to rotate and calculate the baseline angle for each line of the handwriting sample. The baseline of the overall handwriting sample is calculated as the average value of calculated baselines for all lines. To rotate the lines we applied affine transformation with linear transformation and translation, as result a new handwriting samples with straightened lines is obtained. The Figure 3 illustrate the process of dilating Figure 3(a), drawing bounding boxes Figure 3(b) and rotating the lines to obtain the straigntened handwriting sample Figure 3(c).

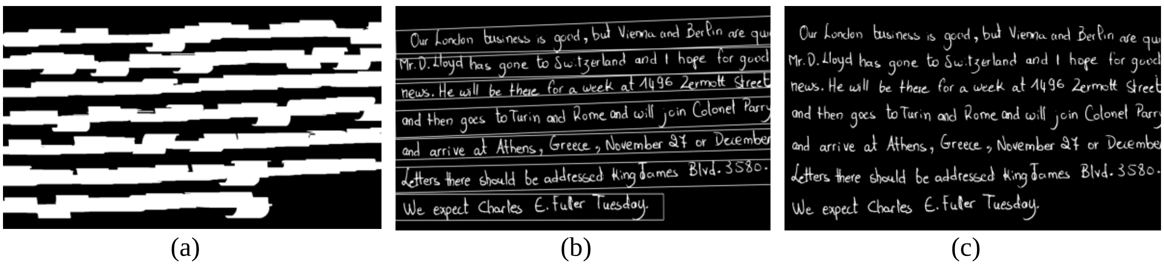


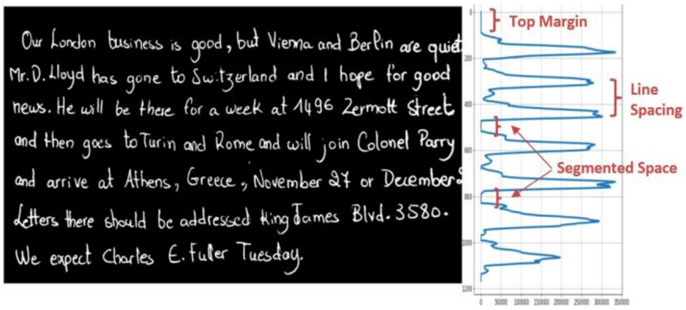
Figure 3. Baseline feature extraction process; (a) dilating, (b) contouring, and (c) rotating

C. Extracting top margin and line spacing features

For top margin and line spacing features, we applied horizontal projection technique [30] on the straightened handwriting sample. By counting the sum of white pixels horizentally we were able to trace zero values and values under a defined threshold to find the start and the end space indices for line segmentation. Additionally, we traced maximum values to find the middle baseline. Figure 4 illustrate the horizental projection algorithme used for top margin and line spacing feature extraction process Figure 4(a), where the top margin value is calculated based on the first two indices, and the line spacing values are calculated as the distance between two middle baseline values Figure 4(b). The average of all spaces between lines is considered as the line spacing of the overall handwriting sample.

```
Algorithm 1: Horizontal Projection
Input: img: The input image
Output: A list containing the sum of pixels in each row
Function horizontalProjection(img);
    (h, w) ← img.shape[: 2];
    sumRows ← [];
    for j ← 0 to h − 1 do
        row ← img[j : j + 1, 0 : w];
        sumRows.append(np.sum(row));
    return sumRows;
```

(a)



(b)

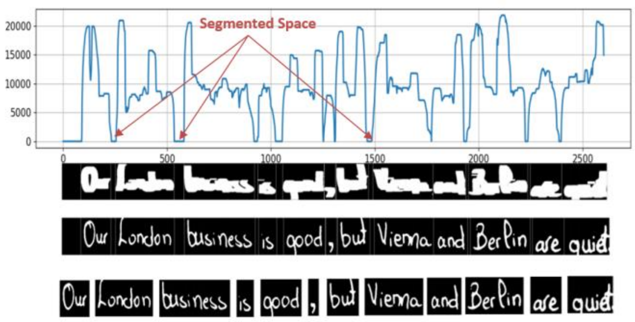
Figure 4. Top margin and line spacing features extraction algorithm and process (a) algorithm and (b) result

D. Extracting word spacing feature

To calculate only the space between words and not letters, we once again applied dilataion on segmented lines and used the vertical projection technique [30] on each dilated line to find spaces start and end indices. Figure 5 illustrate the vertical projection algorithme Figure 5(a) and the resulted segmented words based on spaces start and end indices Figure 5(b).

```
Algorithm 2: Vertical Projection
Input: img: The input image
Output: sumCols: A list containing the sum of pixels in each column
Function verticalProjection(img);
    h, w ← img.shape[: 2];
    sumCols ← [];
    for j ← 0 to w − 1 do
        col ← img[0 : h, j : j + 1];
        sumCols.append(np.sum(col));
    return sumCols;
```

(a)



(b)

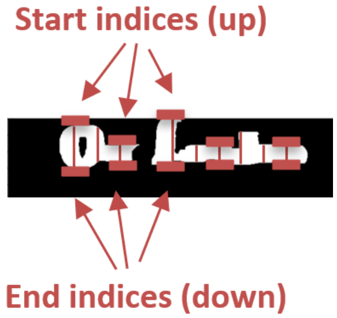
Figure 5. Word spacing features extraction algorithm and process (a) algorithm and (b) result

E. Extracting letter size

Letter size feature could be determined as categories based on middle zone calculation compared to the average letter size [31]. Though, this methode doesn't provide a precise value of this handwriting feature. As an alternative solution, we used dilatation and vertical projection again on each one of the segmented lines to trace pixels with non-zero values that represents the start and end of the letters shape. The differences between those indices allowed us to calculate the average letter size for each line, and the average of all lines represents the letter size feature value of the overall handwriting sample. Figure 6 presents the used algorithme of this technique Figure 6(a) and the illustration of letter size feature values in a segmented dilated line Figure 6(b).

```
Algorithm 3: Extract Letter Size from Lines
  Input: lines: List of lines representing images
  Output: A list containing average letter sizes for each line
  Function Extract_Letter_Size(lines);
    all_letter_sizes ← [];
    foreach line ∈ lines do
        line_letter_sizes ← [];
        up ← 0, down ← 0;
        dilatedline ← cv2.dilate(line, np.ones(kernelSize, np.uint8), iterations = 1);
        (h, w) ← line.shape[: 2];
        for j ← 0 to w 10 do
            for i ← 0 to h − 1 do
                if dilatedline[i : i + 1, j : j + 1] > 0 then
                    up ← i;
                    break;
            for i ← h − 1 to 0 do
                if dilatedline[i : i + 1, j : j + 1] > 0 then
                    down ← i;
                    break;
            line[up : down, j : j + 1] ← [255, 0, 0];
            letter_size ← down − up;
            up ← 0;
            down ← 0;
            if letter_size ≠ 0 then
                line_letter_sizes.append(letter_size);
        all_letter_sizes.append(np.average(line_letter_sizes));
        cv2.imshow("Letter size of line", dilatedline);
    return all_letter_sizes;
```

(a)



(b)

Figure 6. Letter size features extraction algorithm and process (a) algorithm and (b) result

2.2.3. Dataset

Before storing data as the final data set, two main steps were crucial to enhance the classification task by the different classifiers, standardization and transformation. Standardizing the inputted data, representing the handwriting features, could be described as changing the values to fit in the form of a unified scale [32] mainly using a scaling techniques known as robust scaling [33]. As for the ouputed data, representing the personality traits scores, transformation to categories or classes is required for training and testing each one of the machine learning models, mainly two cases scenarios of classification for each one of the personality traits: a multiple (5) and a binary (2) classes classification.

2.3. Handwriting personality assessment

2.3.1. Correlation analysis

By using a statistical method such as correlation analysis we aim to investigate the existing links between the big five personality traits and handwriting feature values. Pearson correlation coefficient [34] was calculated for both values to investigate the nature of relationships between the predicting and predictors variables.

2.3.2. Machine learning models

At this stage, the extracted handwriting features are used to test the prediction of the big five personality traits using four well-known machine learning models. The efficiency of the selected models is measured in comparison with previous studies. This integration highlights how technology-driven data collection techniques can improve scalability and efficiency [35]. When it comes to evaluating machine learning models, selecting the right evaluation criteria is crucial, especially when working with limited datasets. These measures function as benchmarks to assess the effectiveness and resilience of models in many contexts. Accuracy, precision, recall, and F1-score [36] stand out as key tools for assessing a machine learning model performance among the multitude of existing metrics evaluation methods.

– Accuracy score: it is the fraction of predictions where the model has the right prediction.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

– Precision: it is the division of correct positive results by the predicted positive results.

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

– Recall: it is the division of correct positive results by all relevant samples.

$$Recall \ = TP\ (TP + FN) \tag{3}$$

– F1 score: it represents harmonic mean between precision and recall describing model's robustness.

$$F1 = \frac{2}{\frac{1}{Precision} * \frac{1}{Recall}} \quad (4)$$

2.4. Mobile application development

The developement of a mobile application will offer the possibility to automate handwriting feature extraction and big five personality traits prediction mainly to be used by graphologist. This solution will allow a faster and efficient assessment and interpretation of personality without manual analysis. Figure 7 illustrates the architecture design for this mobile app which encompasses an API for rooting requests and calling scripts for features extraction and predictive machine learning models in a local server that will be accessed using the Ngrok tunneling service.

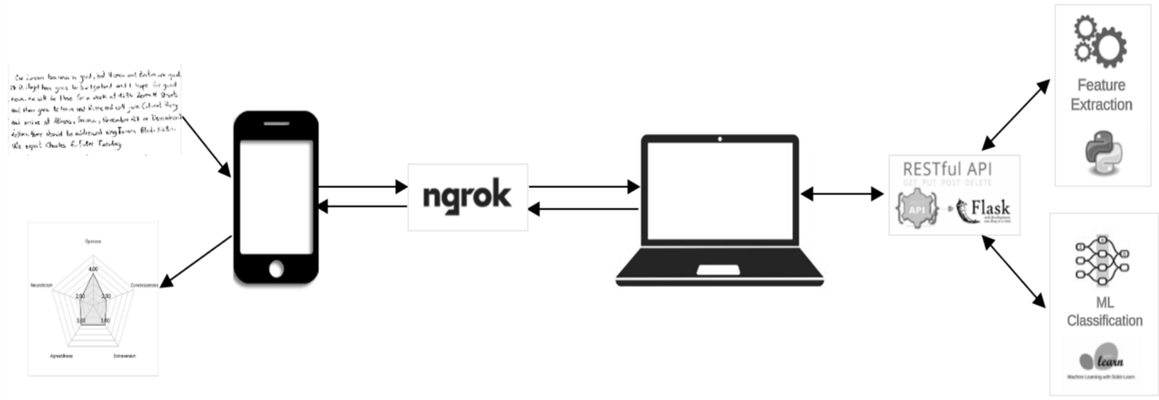


Figure 7. Architecture design of the mobile application for handwriting based personality prediction

3. RESULTS AND DISCUSSION

This section provides a comprehensive overview of our findings and their interpretation. Starting with a detailed description of collected data. Followed by correlation analysis results to validate the relationship between big five personality traits and handwriting features, and to support the methodology's effectiveness. Machine learning models developed for trait prediction are discussed, followed by an evaluation of their performance. Finally, we explore the deployment of these models in the mobile application, emphasizing its potential for automating graphology and enhancing personality trait prediction.

3.1. General description of collected data

Among the 100 participants in this study, the gender distribution revealed that 53 were female, accounting for 53% of the sample, and 47 were male, accounting for 47%. This reveals a fairly equal gender representation in the data set. As for the comparison of personality traits between male and females students participating, intriguing results were found. In terms of openness, females tend to score somewhat higher than males, and males had a slightly higher mean score for agreeableness than females. Both genders had equal mean scores on consciousness and extraversion. However, females score somewhat higher in neuroticism than males. Those findings are consistent with the current research on gender variations in personality characteristics. Previous researches have found similar results, with females rating higher in openness and neuroticism, while males score somewhat higher in agreeableness [37]. These persistent patterns lend credence to the concept of gender-specific variances in personality traits and strengthen the current study’s findings. A more detailed description of findings in contrast with the collected data are presented below in Table 3 as the resluted dataset of this study.

Table 3. Data set predictive and predicted variables description

Predictive variables			Predicted variables		
Handwriting features	Mean	Range	Personality traits	Mean	Range
Baseline	-0.78	(-3.74)-3.22	Openness	2.5	1-5
Top margin	45.91	0-415	Conscientiousness	2.5	1-5
Line spacing	147.05	68.75-218.12	Extraversion	3	1-5
Word spacing	28.23	8.44-63.82	Agreeableness	3	1-5
Letter size	34.74	20.47-56.79	Neuroticism	3.5	1-5

3.2. Correlation analysis

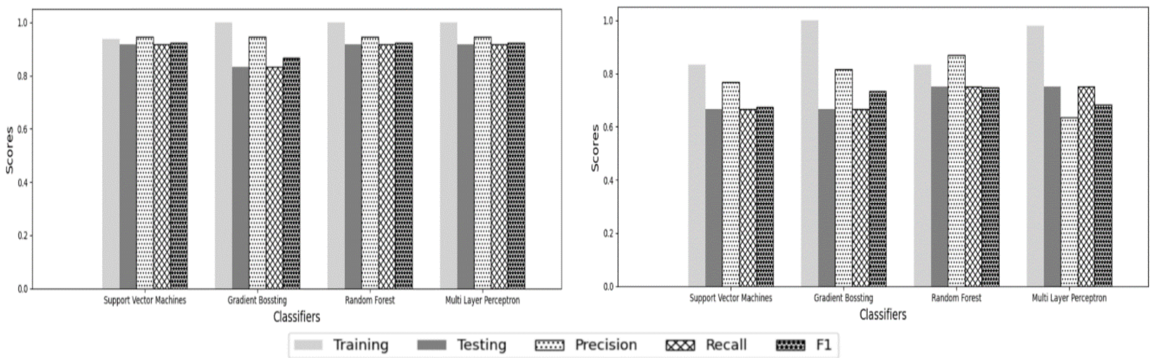
By all of five handwriting features variables, baseline showed a significant negative correaltion with agreeableness, suggesting that agreeableness tended to decrease as the baseline characteristic increases. For the other personality traits there were no statistically significant relationships found with baseline. Conversely, top margin and conscientiousness showed a strong positive correlation, indicating that greater levels of conscientiousness were correlated with higher levels of top margin. Moreover, correlations were found between extraversion and top margin suggesting a potential negative correlation, and between neuroticism and letter size indicating a potential decrease in neuroticism as letter size increased. These findings align with previous research conducted in the field of graphology and personality prediction, where similar relationships between handwriting features and personality traits have been observed [38]–[41]. Here in Table 4 we present the coefficients and their significances results of big five personality traits and handwriting features correlation analysis:

Table 4. Pearsonian correlation coefficients and significances of personality traits and handwriting features

		Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness
Baseline	Coeff	-0.055	-0.266	-0.102	0.027	0.094
	Sig	0.679	0.040	0.439	0.835	0.476
Top margin	Coeff	-0.225	-0.005	0.353	-0.150	0.130
	Sig	0.084	0.970	0.006	0.253	0.433
Letter size	Coeff	0.036	0.064	-0.196	-0.251	0.194
	Sig	0.783	0.627	0.133	0.053	0.138
Line spacing	Coeff	0.062	0.047	-0.073	-0.161	0.004
	Sig	0.637	0.719	0.582	0.218	0.975
Word spacing	Coeff	0.025	0.157	0.011	0.114	-0.170
	Sig	0.850	0.232	0.936	0.386	0.193

3.3. Machine learning

Before evaluating selected models prediction performences, we randomly splited the dataset into two sets: 80% of data as training set and the remaining 20% as testing set. The test set has for the purpose to explore the classifier prediction accuracy on unseen handwriting features data. A two case scenario of multiple trials and errors were conducted for each one of the classifiers, a five classes and two classes predictions. In Figure 8, we present comparison of classifiers performances in predicting binary classes Figure 8(a) and multiclasses Figure 8(b) for openness personality traits based on handwriting features variables.



(a) (b)

Figure 8. Evaluation and comparison of classifiers performances in predicting openness personality traits based on handwriting features (a) binary classification and (b) multiclasses classification

In the binary class case scenario, SVM, RF, and multilayer perceptron (MLP) all performed well in predicting openness, with all metrics scores above 0.90. Similarly, in predicting conscientiousness, all classifiers had excellent training accuracies of 0.90 to 1.00, with testing accuracies, precision, recall, and F1-scores higher than 0.83. However, for extraversion, while SVM, RF and MLP performed well especially MLP outperforming the other classifiers, GB had lower testing accuracies and precision scores. In predicting agreeableness, all classifiers scored impressive testing accuracies ranging from 0.83 to 0.92, with RF and MLP outperforming the others. As for neuroticism, all classifiers had excellent testing accuracies and recall scores of 0.92, with SVM excelling in performance. Overall, RF and MLP consistently outperformed across various measures, presenting an excellent alternatives for predicting personality dimensions. For a more nuanced evaluation of our best models' performances in comparison to previous relevant studies, we present in Table 5 a comparison with findings described in the existing studies that explored the binary classification of personality traits employing the five-factor model (OCEAN).

Table 5. Comparison of handwriting based personality prediction studies in contrast with five factors model

Study (year)	Dataset size	Handwriting features	ML models	Prediction accuracies according to the five factors model				
				O	C	E	A	N
Gavrilescu and Vizireanu [13]	128	Baseline angle word slant pen pressure connecting strokes lower case "t" lower case "f" line spacing	Feed forward neural network (FFNN)	84	77	84	77	84
Elngar *et al.* [15]	125	Baseline pen pressure word spacing lLne spacing lower case "t	FFNN	85	65	85	75	70
Chaubey and Arjaria [16]	110	No feature selection	CNN	43	46	41	40	46
This study (2024)	60	Top margin baseline angle line spacing word spacing letter size	MLP and RF	92	92	92	92	92

As for the multi class case scenario, both RF and MLP classifiers had a testing accuracy and recall of 0.75 for predicting openness traits, whereas SVM and GB classifiers had a testing accuracy and F1 score of 0.67 and 0.73, respectively for predicting conscientiousness. RF and MLP classifiers also excelled in predicting extraversion traits, with testing accuracy and F1 scores higher than 0.60. Moreover, both classifiers led in agreeableness traits, with RF having testing accuracy of 0.75, precision and recall of 0.83 and 0.75, respectively. GB classifiers excelled at predicting neuroticism characteristics, with testing accuracy and precision of 0.58, while RF was more precise classifier with 0.68 score. Overall, while performance varies among personality dimensions, RF and MLP classifiers regularly emerge as good competitors for predicting personality traits when evaluated using the presented metrics scores. On the basis of these experimental findings, we propose the architecture presented in Figure 9, as the prediction system that will be used in the mobile application handwriting personality prediction.

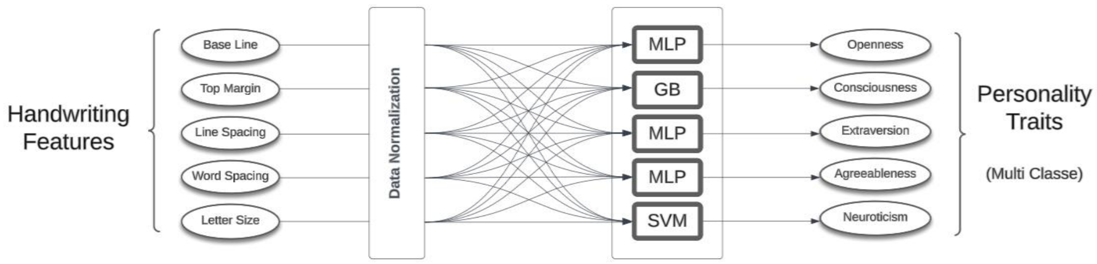


Figure 9. Architecture of the multi class classification system

3.4. The mobile application

Based on the handwriting feature extraction algorithms and the predicting system encompassing best machine learning models, we developed a novel android mobile application to offer fast and robust handwriting-based personality prediction mainly for graphologist. The application was developed under android studio integrated development environment using XML language for coding user interfaces and Java for event handling. In Figure 10, we demonstrate each activity role and use case where; Figure 10(a) this activity is the start graphical interface, where the user is asked to open his camera for capturing a photo of the handwriting sample, Figure 10(b) this activity shows the captured photo and allows user to crop for better prediction results, Figure 10(c) this activity is the same activity Figure 10(a), it shows the cropped image and give the user possibility to go back to take another photo of a new handwriting, Figure 10(d) this activity is the results activity, first it shows a loading panel to allow user to wait for the server response, Figure 10(e) This activity is the same as activity Figure 10(d) after getting results from the server the loading panel is swapped with the chart panel in order to show the predicted scores for the big five personality traits scores.

Compared to earlier studies, we advanced the field by integrating machine learning models into a mobile app for real-time personality prediction. While similar applications are rare [17], [18], our approach covers both automating handwriting feature extraction and big five personality traits prediction. Although our work imitated by is the smaller number of handwriting samples and number of extracted features, our predictive system performed unexpectedly well in predicting big five personality traits. In summary, we developed a mobile solution that predicts big five traits from handwriting. While promising, further work is needed to automate the extraction of more features, scale the models to larger datasets, and explore additional personality frameworks for broader applicability.

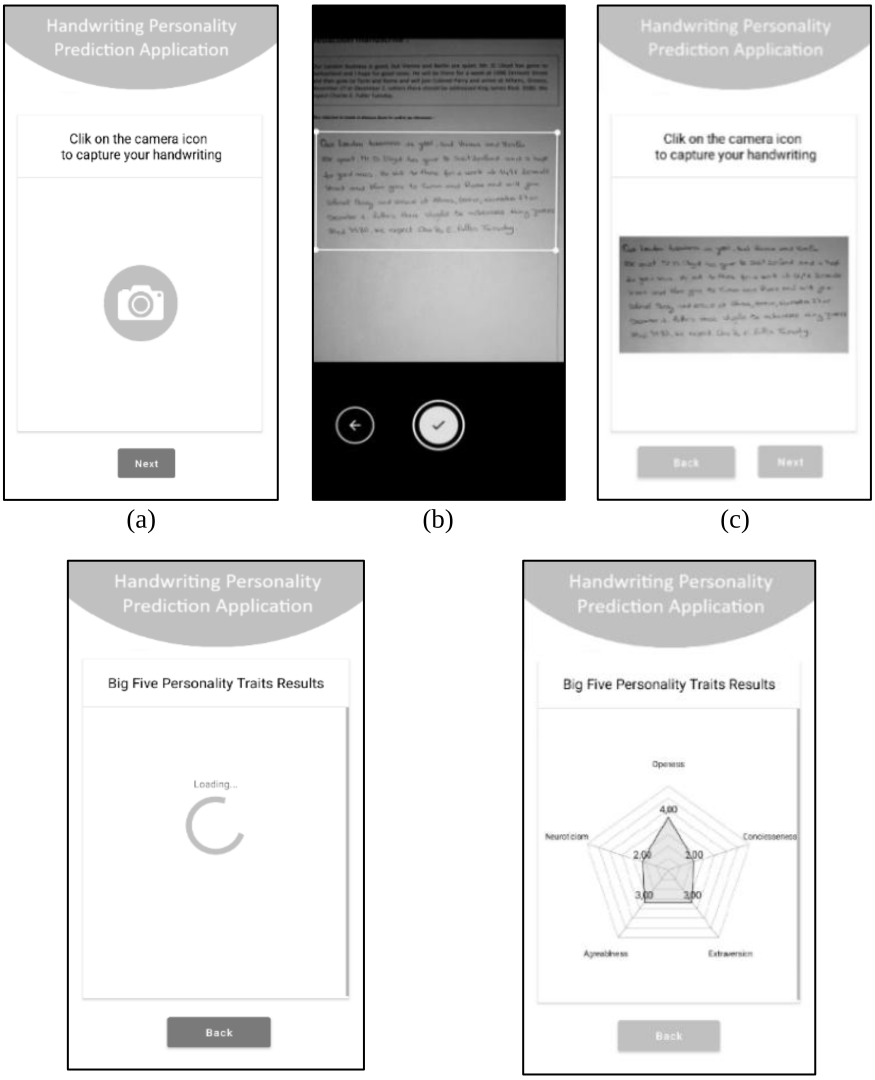


Figure 10. Illustration of the developed application for handwriting personality prediction; (a) scan, (b) crop, (c) resulted image, (d) loading, and (e) results

4. CONCLUSION

In this work, we successfully established a novel approach to personality prediction through handwriting analysis. While correlation analysis revealed significant associations between certain features and personality dimensions, notably agreeableness with baseline and conscientiousness with top margin, machine learning algorithms achieved satisfactory accuracies in predicting all personality dimensions. These findings suggest potential in using advanced computational approaches for personality prediction from handwriting features, contributing to understanding individual differences in personality assessment. Moreover, a mobile application was developed based on robust handwriting feature extraction techniques and machine learning models to bridge theoretical study with practical application in handwriting based personality assessment. Several limitations should be acknowledged, such as the limited number of handwriting features, the exclusive relevance to the big five model, reliance on predefined thresholding techniques for segmentation, and the need for more evaluation of machine learning models. Future research avenues could address these limitations by exploring a broader range of handwriting features, examining alternative trait models such as the MBTI, refining segmentation approaches using machine learning, conducting further validation of machine learning models, and increasing sample size diversity. Such perspectives would contribute to advancing the field of handwriting-based personality prediction and offer a more comprehensive understanding of handwriting and personality correlations.

REFERENCES

[1] G. W. Allport, *Personality: a psychological interpretation*. APA PsycNet, 1937.

[2] H. E. P. Cattell and A. D. Mead, "The sixteen personality factor questionnaire (16PF)," in *The SAGE Handbook of Personality Theory and Assessment: Volume 2 — Personality Measurement and Testing*, 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom: SAGE Publications Ltd, 2008, pp. 135–159. doi: 10.4135/9781849200479.n7.

[3] R. R. McCrae and O. P. John, "An introduction to the five-factor model and its applications," *Journal of Personality*, vol. 60, no. 2, pp. 175–215, Jun. 1992, doi: 10.1111/j.1467-6494.1992.tb00970.x.

[4] I. B. Myers, M. H. Mccaulley, N. L. Quenk, and A. L. Hammer, *MBTI manual: a guide to the development and use of the myers-briggs type indicato*. Consulting Psychologists Press, 1998.

[5] L. Durga and R. Deepu, "Handwriting analysis through graphology: a review," in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Sep. 2018, pp. 1160–1166. doi: 10.1109/ICACCI.2018.8554416.

[6] D. E. Thorpe, "'Nonsense rides piggyback on sensible things': the past, present, and future of graphology," in *Medical Paratexts from Medieval to Modern*, 2018, pp. 139–155. doi: 10.1007/978-3-319-73426-2_9.

[7] C. De Stefano, F. Fontanella, D. Impedovo, G. Pirlo, and A. Scotto di Freca, "Handwriting analysis to support neurodegenerative diseases diagnosis: a review," *Pattern Recognition Letters*, vol. 121, pp. 37–45, Apr. 2019, doi: 10.1016/j.patrec.2018.05.013.

[8] H. S. Friedman, M. L. Kern, S. E. Hampson, and A. L. Duckworth, "A new life-span approach to conscientiousness and health: combining the pieces of the causal puzzle.," *Developmental Psychology*, vol. 50, no. 5, pp. 1377–1389, 2014, doi: 10.1037/a0030373.

[9] A. E. Poropat, "A meta-analysis of the five-factor model of personality and academic performance.," *Psychological Bulletin*, vol. 135, no. 2, pp. 322–338, Mar. 2009, doi: 10.1037/a0014996.

[10] T. A. Judge and D. M. Cable, "Applicant personality, organizational culture, and organization attraction," *Personnel Psychology*, vol. 50, no. 2, pp. 359–394, Jun. 1997, doi: 10.1111/j.1744-6570.1997.tb00912.x.

[11] A. Remaida, A. Moumen, Y. El Bouzekri El Idrissi, and B. Abdellaoui, "Handwriting personality recognition with machine learning: A comparative study," in *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, Dec. 2020, pp. 1–6. doi: 10.1109/ICECOCS50124.2020.9314529.

[12] K. Chaudhari and A. Thakkar, "Survey on handwriting-based personality trait identification," *Expert Systems with Applications*, vol. 124, pp. 282–308, Jun. 2019, doi: 10.1016/j.eswa.2019.01.028.

[13] M. Gavrilescu and N. Vizireanu, "Predicting the big five personality traits from handwriting," *EURASIP Journal on Image and Video Processing*, vol. 2018, no. 1, Dec. 2018, doi: 10.1186/s13640-018-0297-3.

[14] P. Joshi, P. Ghaskadbi, and S. Tendulkar, "A machine learning approach to employability evaluation using handwriting analysis," in *Advanced Informatics for Computing Research*, 2019, pp. 253–263. doi: 10.1007/978-981-13-3140-4_23.

[15] A. A. Elngar, N. Jain, D. Sharma, H. Negi, A. Trehan, and A. Srivastava, "A deep learning based analysis of the big five personality traits from handwriting samples using image processing," *Journal of Information Technology Management*, vol. 12, pp. 3–35, 2020.

[16] G. Chaubey and S. K. Arjaria, "Personality prediction through handwriting analysis using convolutional neural networks," in *Proceedings of International Conference on Computational Intelligence*, 2022, pp. 59–70. doi: 10.1007/978-981-16-3802-2_5.

[17] K. A. Prasetyo, N. Ramadijanti, and A. Basuki, "Mobile application for identifying personality of person using graphology," in *2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC)*, Sep. 2017, pp. 212–219. doi: 10.1109/KCIC.2017.8228589.

[18] P. B. Nair, A. M. Johnson, A. M. Alex, A. Sebastian, and D. Sunil, "Android app for handwriting analysis using deep learning," *Journal of Communication Engineering and Its Innovations*, vol. 5, no. 3, pp. 16–21, 2019.

[19] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2012.

[20] O. Plaisant, J. Guertault, R. Courtois, C. Réveillère, G. A. Mendelsohn, and O. P. John, "Big five history: OCEAN of personality factors. Introduction of the French big five inventory or BFI-Fr *(in France: Histoire des big five : OCEAN des cinq grands facteurs de la personnalité. Introduction du big five inventory Français ou BFI-Fr)*," *Annales Médico-psychologiques, revue psychiatrique*, vol. 168, no. 7, pp. 481–486, Sep. 2010, doi: 10.1016/j.amp.2009.04.016.

[21] S. D. Gosling, P. J. Rentfrow, and W. B. Swann, "A very brief measure of the big-five personality domains," *Journal of Research in Personality*, vol. 37, no. 6, pp. 504–528, Dec. 2003, doi: 10.1016/S0092-6566(03)00046-1.

[22] R. Courtois *et al.*, "Validation of the French version of the 10-item big five inventory *(in France: Validation Française du big five inventory à 10 items (BFI-10))*," *L'Encéphale*, vol. 46, no. 6, pp. 455–462, Dec. 2020, doi: 10.1016/j.encep.2020.02.006.

[23] "Front Matter," in *Forensic Handwriting Identification*, Elsevier, 2021, pp. i–ii. doi: 10.1016/B978-0-12-409602-8.09989-2.

[24] F. Chiarello, P. Belingheri, and G. Fantoni, "Data science for engineering design: state of the art and future directions," *Computers in Industry*, vol. 129, Aug. 2021, doi: 10.1016/j.compind.2021.103447.

[25] N. R. Soora and P. S. Deshpande, "Review of feature extraction techniques for character recognition," *IETE Journal of Research*, vol. 64, no. 2, pp. 280–295, Mar. 2018, doi: 10.1080/03772063.2017.1351323.

[26] G. Bradski, "The Opencv library," *Dr. Dobb's Journal of Software Tools*, vol. 120, pp. 122–125, 2000.

[27] N. Otsu, "A Tlreshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.

[28] M. Hee Lee and S. Y. Shin, "Modified pixels based fast median filter in impulse noise environments," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 9, no. 3, pp. 731–741, Mar. 2018, doi: 10.11591/ijeecs.v9.i3.pp731-741.

[29] S. Suzuki and K. Be, "Topological structural analysis of digitized binary images by border following," *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 1, pp. 32–46, Apr. 1985, doi: 10.1016/0734-189X(85)90016-7.

[30] R. Ptak, B. Żygadło, and O. Unold, "Projection–based text line segmentation with a variable threshold," *International Journal of Applied Mathematics and Computer Science*, vol. 27, no. 1, pp. 195–206, Mar. 2017, doi: 10.1515/amcs-2017-0014.

[31] A. H. Garoot, M. Safar, and C. Y. Suen, "A comprehensive survey on handwriting and computerized graphology," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Nov. 2017, pp. 621–626. doi: 10.1109/ICDAR.2017.107.

[32] V. N. G. Raju, K. P. Lakshmi, V. M. Jain, A. Kalidindi, and V. Padma, "Study the influence of normalization/transformation process on the accuracy of supervised classification," in *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Aug. 2020, pp. 729–735. doi: 10.1109/ICSSIT48917.2020.9214160.

[33] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median," *Journal of Experimental Social Psychology*, vol. 49, no. 4, pp. 764–766, Jul. 2013, doi: 10.1016/j.jesp.2013.03.013.

[34] P. Sedgwick, "Pearson's correlation coefficient," *BMJ*, vol. 345, no. jul04 1, pp. e4483–e4483, Jul. 2012, doi: 10.1136/bmj.e4483.

[35] I. H. Sarker, M. M. Hoque, M. K. Uddin, and T. Alsanoosy, "Mobile data science and intelligent apps: concepts, AI-based modeling and research directions," *Mobile Networks and Applications*, vol. 26, no. 1, pp. 285–303, Feb. 2021, doi: 10.1007/s11036-020-01650-z.

[36] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, no. 2, pp. 01–11, Mar. 2015, doi: 10.5121/ijdkp.2015.5201.

[37] M. Del Giudice, "Gender differences in personality and social behavior," in *International Encyclopedia of the Social & Behavioral Sciences*, Elsevier, 2015, pp. 750–756. doi: 10.1016/B978-0-08-097086-8.25100-3.

[38] S. Broschk, "Graphology and personality: a correlational analysis," Master of Arts (MA), University of Johannesburg, 2008.

[39] E. A. Lemke and J. H. Kirchner, "A multivariate study of handwriting, intelligence, and personality correlates," *Journal of Personality Assessment*, vol. 35, no. 6, pp. 584–592, Dec. 1971, doi: 10.1080/00223891.1971.10119716.

[40] D. A. Rosenthal and R. Lines, "Handwriting as a correlate of extraversion," *Journal of Personality Assessment*, vol. 42, no. 1, pp. 45–48, Feb. 1978, doi: 10.1207/s15327752jpa4201_6.

[41] B. Thiry, "Graphology and personality according to the five-factor model," *(in French) Psychologie Française*, vol. 53, no. 3, pp. 399–410, Sep. 2008, doi: 10.1016/j.psfr.2008.04.002.

BIOGRAPHIES OF AUTHORS

Ahmed Remaida is a Ph.D. student at the Laboratory of Engineering Science, Department of Informatics, Logistics and Mathematics at the National School of Applied Sciences at Ibn Tofail University in Kenitra. He has a master's degree in educational technology from the "École Normale Supérieure, ENS" of Martil, Morocco. His research interests include deep machine learning and personality assessment. He can be contacted at email: ahmed.remaida@uit.ac.ma.

Zineb Sabri is a Ph.D. student at the Laboratory of Research in Computer Science at Faculty of Sciences, Ibn Tofail University in Kenitra. She has obtained her master's degree in information systems engineering from University Cadi Ayad, Morocco. Her research disciplines are education, games, machine learning, and employability. She can be contacted at email: zineb.sabri@uit.ac.ma.

Benyoussef Abdellaoui is a Ph.D. student at the Laboratory of Engineering Science, Department of Informatics, Logistics and Mathematics at the National School of Applied Sciences at Ibn Tofail University in Kenitra. He has a master's degree in internet of things (IoT) from the National School of Computer Science and Systems Analysis (ENSIAs) one of the ten establishments of the University Mohammed V, Morocco. His research interests include face detection, emotional state, and engagement with deep learning. He can be contacted at email: benyoussef.abdellaoui@uit.ac.ma.

Chakir Fri is a Ph.D. Student at the Laboratory of Informatics and Information Science, Department of Computer Science at the Faculty of Sciences at Moulay Ismail University in Meknes. He has a master's degree in educational technology from the "Ecole Normale Supérieure, ENS" of Martil, Morocco. His research interests include deep machine learning and E-learning. He can be contacted at email: chakir.fri@gmail.com.

Yassine Lakhchaf is a Ph.D. student at the Laboratory of Engineering Science, Department of Interdisciplinary Computer Science and Physics at the Sciences and Technologies Faculty at Mohammed Ibn Abdallah University in Fez. He has a master's degree in educational technology from the "École Normale Supérieure, ENS" of Martil, Morocco. His research interests to integrate artificial intelligence into the educational system. He can be contacted at email: yassine.lakhchaf@usmba.ac.ma.

Younès El Bouzekri El Idrissi is a Ph.D. research scholar, he is a member of the Laboratory of Engineering Science, Department of Informatics, Logistics and Mathematics at the National School of Applied Sciences at Ibn Tofail University in Kenitra. He can be contacted at email: y.elbouzekri@gmail.com.

Mohammed Amine Lafraxo is a Ph.D. Research scholar in Health Sciences at the Higher Institute of Nursing Professions and Health Techniques Oujda, Ministry of Health. He can be contacted at email: lafraxo.ma@gmail.com.

Aniss Moumen is a Ph.D. research scholar, he is a member of the Laboratory of Engineering Science, Department of Informatics, Logistics and Mathematics at the National School of Applied Sciences at Ibn Tofail University in Kenitra. He can be contacted at email: aniss.oumoumen@uit.ac.ma.