Multimedia Tools and Applications (2024) 83:73583–73606
https://doi.org/10.1007/s11042-024-19101-2

1232: HUMAN-CENTRIC MULTIMEDIA ANALYSIS

Check for updates

Analysis and prediction of personality traits using a self-generated database of Moroccan instagram users: impact of gender on image content and quantity on prediction accuracy

Siham El Bahy[1] · **Noureddine Aboutabit**[1] · **Imad Hafidi**[1]

Received: 9 May 2023 / Revised: 22 February 2024 / Accepted: 27 March 2024 /
Published online: 30 April 2024

Abstract
The sharing of social media data has provided a significant amount of information about users. This has led to a growing interest in predicting user personality from social network data. Predicting user personality has potential applications in advertising and content recommendation systems, and can be a valuable source of insights for professionals seeking to understand and target their audiences on social media. This study focuses on the Instagram application and aims to extract and analyze the personality of Instagram users based on the content of their photos. Three categories of features are extracted from images: visual, emotional, and content. To describe users' personalities, we used the Big Five model. One of the main contributions of this study is the creation of a comprehensive database that includes 316 Moroccan Instagram users. This database is the first of its kind to belong to Moroccan users. Additionally, The study also analyzed the influence of personality traits and gender on photo content by independently assessing the personalities of female and male Instagram users. Moreover, the personalities of each gender were extracted separately, and the influence of the number of images on prediction accuracy was studied. The study evaluates prediction accuracy using the root mean square error (RMSE) on a [1,5] score scale. Commendable results were achieved across all personality traits, with standout performance in predicting conscientiousness for females (RMSE=0.59) and openness for males (RMSE=0.59) compared to previous studies.

Keywords Instagram · Personality traits · Big five · Emotions · Image content · Machine learning

✉ Siham El Bahy
siham.elbahy@gmail.com

Noureddine Aboutabit
n.aboutabit@usms.ma

Imad Hafidi
imad.hafidi@gmail.com

1 IPIM Laboratory, ENSA Sultan Moulay Slimane University, PO Box 523, Khouribga 23000, Morocco

1 Introduction

Social networks have become a means of expressing oneself. What we write and the photos we share reflect our inclinations, express our personality and allow others to build an idea about our personality.

Many researchers have been interested in studying the influence of the personality of social media users on their publications. Among these researchers were those interested in the Facebook application and could extract users' personality traits from several profile photos [1] and others who determined them from only one containing the face [2]. Moreover, others have focused on the Instagram app. Ferwerda and Tsakalidis [3] studied the relationship between user personality traits and the content of shared photos. Besides, there are others, who predicted personality from content [4] and visual characteristics [5, 6], and a combination of both [7, 8].

Numerous researchers have centered their attention on European Instagram users. Kim and Kim [7] tried to predict personality traits from visual characteristics, content, and emotions separately, and found that the combination of these categories gave more reliable results than when examined separately. Kim and Kim [6] analyzed color characteristics as an image attribute, and utilized the same database used in [7]. Another study [8] was conducted on this database analyzed visual and content characteristics separately to investigate their correlation with personality traits and utilized these features to predict said traits.

This study focuses on Moroccan Instagram users. The aim is to predict their personality traits based on the images they post on their accounts. Additionally, a study will be conducted on the impact of these traits on the content of the photos. The Big Five model was used to describe user personality, and 36 features were extracted from images, which can be classified into three categories: visual characteristics, content, and emotional features to describe the photos. The main contributions of our study are: firstly, the collecting of a large database in comparison to those used in previous works, which includes 316 users, and it is the first one dedicated to Instagram users in Morocco. Secondly, this study is the first to predict personality traits independently for each gender, taking into account the nuanced variations in interests and inclinations specific to each gender. Thirdly, it explores the impact of personality traits on the characteristics and content of posted images for each gender independently. Finally, a detailed study was conducted on the influence of the number of posted images on prediction accuracy.

The rest of this article is organized as follows; The second section consists of a presentation of related work, and in the third, we offer the methods used. The fourth section deals with the prediction of personality characteristics. In the last two sections, we examine the relationship between personality traits and image characteristics and discuss the results, which we compare to previous studies.

2 Related work

Researchers have shown that published images reflect the personality traits of social media users [1, 9]. Moreover, they have shown a correlation between images posted on social media and the five personality traits. This opens new ways to extract personality traits from social media and facilitate personalized systems. For example, a study on the Facebook platform [10] found that profile pictures convey comprehensive information about users. Furthermore, [10] showed that the information encoded in profile pictures could be exploited to classify the personality traits and interaction styles of Facebook users. Nevertheless, there are differences between the traits, for example, openness and conscientiousness, which have few examples and require more experiments. However, agreeableness and extraversion scored the highest among the personality traits. While dominance and affect reach slightly higher performances. However, neuroticism is the most difficult trait to predict. This is consistent with the results that human raters accurately estimate extraversion and agreeableness, and that extraversion is related to the expressiveness of the pictures.

Another study [2] was conducted on personality trait prediction on the same platform. They proposed an application of convolutional neural networks to model the personality prediction of individuals by analyzing their Facebook profile pictures. They showed that personality prediction models could predict users' personalities from a single profile picture containing a face.

Furthermore, research was conducted on 1290 Twitter users [11] using their profile pictures and tweets to predict their personalities automatically. They extracted facial features using a face detection model combined with smile detection and other features like hue, saturation, and value. The results show that image features can predict personality better than text features and the combination of text and image features.

Based on the user's choice of profile images, [1] extracted the users' personality traits. The semantics of the image can be classified into person identification, event semantics, concept semantics, and location semantics. Based on facial features, they found that nervous and open-minded users post images that contain fewer people and express mostly negative emotions. Moreover, conscientious users post what is necessary for a profile picture and express positive emotions. In addition, extraverts and agreeable users show aspects of the good life like pets, travel, and food. They like colorful pictures but not blurred and bright images. Their facial expressions are generally positive and show a positive mood.

On the Instagram platform, studies have been conducted to extract users' personality traits from the images posted on their accounts. For example, in [4], when they conducted a study on a sample of 113 users in the United States, they could predict personality using visual characteristics (hue, value, saturation) of Instagram images. The results show similar trends to previous work on personality extraction from social media [12]. However, they outperformed it in predicting most of the personality traits. The most successful predictions were openness, conscientiousness, and agreeableness. However, the most complex personality characteristics to identify were extraversion and neuroticism.

Based on the hue of the user's image, this study [13] aimed to infer the user's personality traits since applying filters to pictures before publishing them is considered a method of expressing personality traits. They explored the relationship between image hue and personality traits. In addition, they entered it into the SVM classifier and the CNN model. They succeeded in classifying personality traits.

The authors of [5] examined the ability to predict personality characteristics based on Instagram image features. They used a database of 193 Instagram users. They exploited the images' visual features (e.g., hue, valence, saturation) and content characteristics (architecture, body parts, clothing, musical instruments, art, entertainment, botany, cartoons, animals, food, sports, vehicles, electronics, babies, hobbies, jewelry, weapons). As a result, they obtained a better prediction of certain personality traits based on visual and content features than [12, 14]. Furthermore, they showed that both visual and content characteristics can be used separately to predict personality and that they generally perform well. However, combining the two does not lead to an increase in predictive power. Therefore, they do not add more value than they already have independently. Nevertheless, [15] found features that could improve prediction when combined.

On the same platform [7], 179 university students participated. They are used as features of photos, content categories, counts of faces, emotions expressed on faces, and pixel features. The results of these studies confirm the ability to predict personality traits among the Big Five. Furthermore, they found that the content category is associated with users' extraversion and gender. Moreover, extraversion, agreeableness, and openness are related to the number of faces. However, users' extraversion, agreeableness, and openness were associated with the emotions expressed on the faces in their photos. In addition, some pixel characteristics were correlated with users' extraversion, agreeableness, conscientiousness, and gender. In another study [6], the same authors explored the relationships between Instagram users' characteristics and color characteristics using the same database. They reveal that the color features of Instagram photos are related to personality traits. The results indicate that agreeableness was the most relevant trait associated with all color features. However, the color characteristics of the images were different by gender. However, neuroticism was negatively associated with the color harmony of their photos. The study found that extraversion was positively correlated with color diversity. In contrast, openness was negatively correlated with color diversity and color harmony.

According to research [8], users' Instagram photos are able to predict not only their personality but also their age. This research examined the relationship between gender, age, and personality traits in the same database studied in [6, 7] and the Instagram images of these users. Moreover, gender, age, and personality traits are predicted using machine learning methods.

According to [16], images' features marked as favorites on Flickr can be used to predict self-reported and attributed personality traits. Nevertheless, they noticed a variation for traits but not self-reported traits. The authors explained that participants may have used information such as their personal history and life experiences to determine their traits [17]. However, their favorite pictures did not contain all these things. In addition, they differentiated between two activities that have fundamentally different goals. The first is production, the act of posting an image, and the second is the consumption of social media content, which is the act of liking the image [18].

This research [3] studied the reflection of personality traits of Instagram users on the content of shared images, they worked on a database of 193 Instagram users, and the number of images in the database was 54,962. Using the Google Vision API, they retrieved 4090 unique tags from the Instagram images, and using the doc2vec approach, they gathered them into 400 groups, which were then manually categorized into 17 groups. Their results showed that each user with a particular personality posts a specific type of photo. For example, those with a high score of openness to experience generally posted more images composed of musical instruments, while conscientious users more frequently shared images containing clothing and sports, unlike extroverts who tended to post pictures that included electronic devices. In addition, positive correlations were found between the agreeableness trait and the categories of clothing and hobbies, meaning that agreeable participants' Instagram photo collection consisted of clothing or hobbies. In contrast, those with high neuroticism scores tend to have fewer photographs of clothing and more jewelry.

3 Methodology

This section presents a comprehensive description of the database collection procedures employed in the study. The feature extraction process, including the use of deep learning models such as InceptionV3 and MTCNN, and the techniques applied to calculate correlations between image features and personality traits are explained. Additionally, the prediction methods used to estimate personality traits based on image data are detailed.

Figure 1 presents a flowchart that provides a comprehensive explanation of the presented work stages.

3.1 Database

To conduct this study, we created a database of 316 Moroccan Instagram users, comprising 133 males and 183 females aged between 18 and 38. The participants were diverse, including students from Moulay Ismail University in Morocco and other individuals from our circle of knowledge. They were offered the opportunity to participate in this research and willingly and voluntarily agreed to participate. The only requirement for participation was to have an active Instagram account.

The study utilized the Big Five model [19], which consists of five primary traits: openness, conscientiousness, extroversion, agreeableness, and neuroticism, to extract the personal characteristics of the participants. The model comprises a test that includes 60 questions related to various aspects of personality, presenting a percentage score for each of these five traits. This model is widely used for extracting personality traits [20].

To improve the participants' understanding, we translated the questionnaire into Moroccan dialect and presented it in a Google Docs file. The file contained fields for participants to enter their Instagram username, age, and gender. The file was shared with participants and emphasized the importance of their concentration during the test to avoid omissions or negligence. The file was linked to our email to receive participants' responses.

After receiving participants' responses via email, we entered them into the NEO FI website via the following link: https://truity.com/test/big-five-personality-test. As a result, we obtain scores for each of the five personality traits for every participant. Figure 2 shows the test results for two users.

Table 1 shows a sample of personality traits of database users

To confirm the correspondence between the results of the personality test and the participants' actual personality traits, we sent the results to them and asked for their feedback regarding the alignment of the results with their traits. This process was instrumental in validating the precision of the obtained results.

However, we created an Instagram account and requested participants to accept our subscription to enable us to download their published photos. We used an extension called 'Downloader for Instagram' to download their images into a file named after the partici-

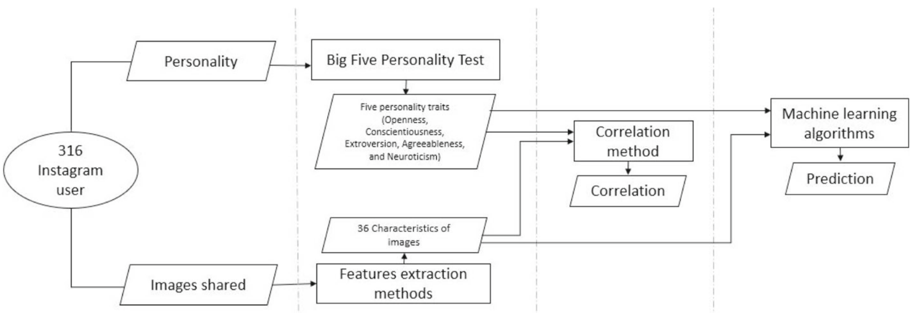

Fig. 1 Flowchart illustrating the stages of this study

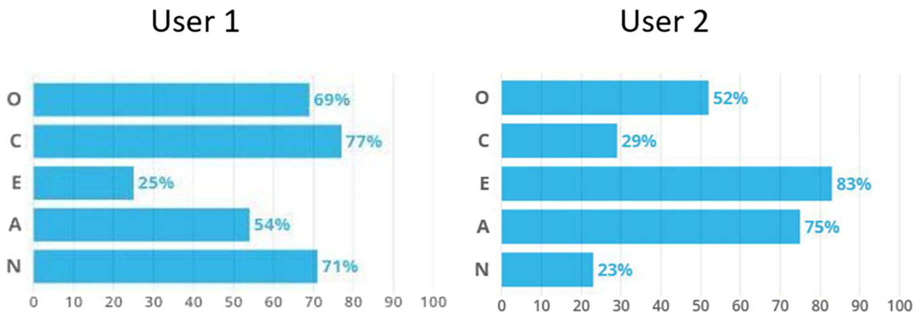

Fig. 2 An example of a test result of two users of dataset

pant's account. In total, we collected 19,898 photos. Figure 3 displays a sample of the photos in the database.

3.2 Features extraction

Machine learning techniques were used to extract relevant features from images, predict personality traits based on the image data, and ultimately compute a correlation between the image features and personality traits.

3.2.1 Features extraction methods

Image features were measured in three categories: content, emotion, and visual characteristics.

Content category

16 contents were extracted from images, delineated as follows:

Firstly, the InceptionV3 method [21] was used to determine the percentage occupancy of each category (Abstract, Buildings, Dark, Object, People) in the image. InceptionV3 is a convolutional neural network architecture developed by Google for image recognition tasks. It utilizes a combination of convolutional layers, pooling layers, and fully connected layers to process image data and extract relevant features.

Secondly, the Multi-Task Cascaded Convolutional Networks (MTCCN) [22] model was used to detect faces in the image. It is a deep learning-based face detection algorithm that uses a series of convolutional neural networks. It is a multi-stage network that generates candidate face regions, refines them to eliminate false positives and adjust the bounding boxes, and extracts facial landmarks for each detected face.

Table 1 A sample of personality traits of database users, (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism

Username	O	C	E	A	N
User 1	69%	77%	25%	54%	71%
User 2	52%	29%	83%	75%	23%
User 3	79%	87%	65%	67%	19%
User 4	65%	37%	73%	98%	77%
User 5	71%	54%	60%	77%	52%
User 6	79%	56%	46%	75%	48%

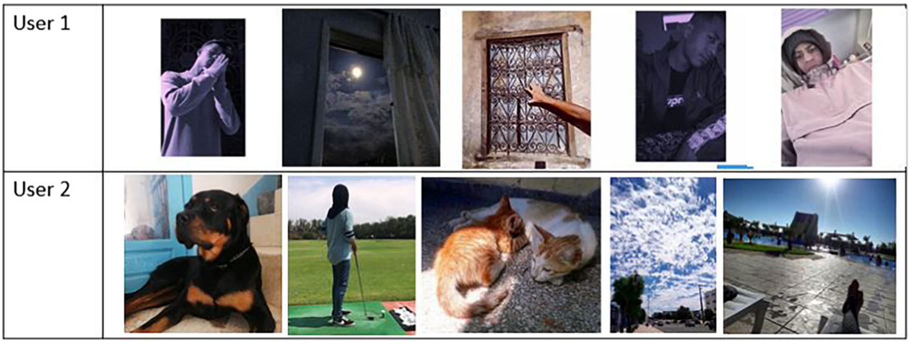

Fig. 3 A sample of the database's photos, (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism

Thirdly, the You Only Look Once (YOLO) method [23] was used to detect the following contents: Cat, Dog, Sports ball, TV, Laptop, Cup, Bottle, Horse, and Purse. It is a real-time object detection algorithm based on deep neural networks. It uses a single convolutional neural network to predict object classes and bounding boxes directly from full images in one evaluation. YOLO has been updated and improved over the years, with YOLOv2, YOLOv3, and YOLOv4 being the most recent versions. After detecting objects, the average of the images that contained each of these contents was calculated for each user.

Finally, As a method of detecting light sources in images, we converted each image into HSV format. Using the Value parameter, we calculate the average of pixels having a Value greater than 240 on each image. These values will be averaged over all photos posted by the user.

Emotion category

Convolutional Neural Networks (CNNs) [24] were used to detect emotional features. CNNs consist of multiple layers, including convolutional layers for feature extraction, pooling layers for reducing output resolution, and fully connected layers for input classification. Seven emotional features (Anger, Disgust, Fear, Joy, Neutral, Sadness, and Surprise) were identified using CNN. Additionally, the percentages of each emotion conveyed in the images were computed and then averaged across all of the user's photos.

Visual characteristic category

Nine colors were extracted from the images(Red, Orange, Yellow, Yellow-green, Green, Cyan, Blue, Violet, and Pink), in addition to two features: warm colors (Red, Orange, Yellow, Yellow-green, and Pink), and cold colors (Green, Cyan, Blue, Violet).

To extract these features, the images were converted into HSV and the Hue parameter range was divided into intervals corresponding to colors. For each interval, the total number of pixels having a value within the field was divided by the total number of pixels of the image. The average of these values for all images was then calculated.

3.2.2 Prediction methods

The following five machine learning methods were used to develop predictive models for personality traits:

Random forest [25]: A machine learning algorithm that belongs to the ensemble learning family. It is a collection of decision trees that are trained on random subsets of data and features. During the training process, multiple decision trees are generated, and their predictions are combined to make a final prediction. The algorithm is designed to reduce overfitting by introducing randomness in the selection of features and data samples. This leads to a more robust and generalizable model that can handle large amounts of data and high-dimensional feature spaces.

Decision tree [26]: A machine learning algorithm that builds a hierarchical tree-like model to represent decisions and their possible consequences. It partitions the data into smaller subsets based on a set of attributes and creates a tree structure in which each node represents a test on an attribute and each branch represents the outcome of the test. The goal is to create a model that can accurately predict the class label of new data based on the decision tree structure. Decision trees are easy to interpret and can handle a mix of continuous and categorical variables.

Radial Basis Function (RBF) [27]: A mathematical function used in machine learning to transform input data into a high-dimensional feature space. The RBF function is defined as a similarity function between pairs of input vectors, where the similarity is measured as the distance between the vectors in the input space.

Support Vector Regression (SVR) [28]: A machine learning algorithm that is used for regression analysis. It is an extension of Support Vector Machines (SVMs) and works by finding the optimal hyperplane that passes through a subset of the data points. The algorithm seeks to minimize the margin violations between the hyperplane and the training data while also minimizing the complexity of the model. SVR can handle both linear and non-linear data by transforming the data into a higher-dimensional space using kernel functions. The objective of SVR is to predict the output variable of a new data point by using the trained model.

Linear regression [29]: A statistical method for predicting a continuous output variable based on one or more input variables. It assumes that there is a linear relationship between the input variables and the output variable. The goal of linear regression is to find the best-fit line that represents the relationship between the input variables and the output variable. This line is obtained by minimizing the sum of the squared errors between the predicted and actual values.

3.2.3 Correlation methods

To calculate the correlation coefficient (r) between personality traits and image features, the Pearson correlation method [30] was utilized. It is a statistical measure that quantifies the degree of the linear relationship between two variables. It is commonly used to assess the strength and direction of the association between two continuous variables, X and Y, which can be calculated using the following mathematical expression.

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

Here:

- n is the number of observations.
- X_i and Y_i are individual data points for variables X and Y.
- $\bar{X}$ and $\bar{Y}$ are the means of variables X and Y respectively.

The Pearson correlation coefficient (r) ranges from -1 to 1, indicating the strength and direction of a linear relationship between two variables. A coefficient of 1 indicates a perfect positive correlation, -1 indicates a perfect negative correlation, and 0 indicates no correlation.

4 Experiment result and discussion

This section presents two parts of our study: the prediction of personality from shared images and the correlation between image features and personality traits.

4.1 Personality prediction from shared images

This section focuses on the prediction of personality traits based on features extracted from images and presents the parameters of each experiment, followed by the evaluation method, results and discussion. Finally, a comparison with the results of previous work.

4.1.1 Experiments process

Tree experiments were conducted to predict personality traits. In each experiment, the number of accounts was modified by eliminating those with fewer than a certain number of photos. The database was divided into two sections, one for females and one for males, to predict the personality of each gender separately and understand the influence of gender on the prediction of each trait.

Each personality trait will be predicted separately, assuming that there is no overlap between them. We trained our prediction models using regression methods, including Random Forest, Decision Tree, Linear Regression, Support Vector Regression, and Radial Basis Function Networks. We repeated the experiment for ten iterations by periodically selecting 10% of the data for testing and 90% for training. Furthermore, we used the reported root mean square error (RMSE) (r∈[1.5]) to indicate the accuracy of predicting personality traits. Moreover, the accuracy obtained at each iteration was averaged.

Our experiments were as follows:

Experience 1

In this experiment, we used all database users and trained our predictive models from the extracted features (Table 2).

Experience 2

Our objective in this experience is to assess the effect of the number of images posted on an account on the accuracy of prediction traits. Therefore, we repeat the previous experience with only accounts with sufficiently shared images. In our case, we consider at least ten photos. Consequently, we have 270 Instagram users in our database after filtering.

Experience 3

In this experiment, we restricted the training of the predictive models to users with over 20 photos in their accounts. As a result, our database includes 210 Instagram users after filtering.

4.1.2 Evaluation metrics

To comprehensively evaluate the performance of our predictive model, we have chosen the Root Mean Squared Error (RMSE) as the primary metric. RMSE, widely employed in regression tasks, quantifies the average magnitude of errors between predicted and actual values. Its utilization aligns with the continuous nature of our predicted outcomes and offers a holistic perspective on prediction accuracy.

The selection of RMSE is justified by its ability to penalize larger errors more significantly, reflecting the importance we place on accurate predictions in our specific application. This

Table 2 Features categories

Content category	Emotion category	Visual characteristic category	Other features
abstract, buildings dark, object, people, cat, dog, sports ball, tv, laptop, cup, bottle, horse, handbag, percentage of photos with faces, face number.	angry, disgust, fear, happy, neutral, sad surprise.	brightness, red, orange, yellow, green_yellow, green, cyan, blue, violet, rose, warm color, cold color.	images number

metric has been consistently used in prior works within our domain, ensuring a standardized comparison basis for our results.

In interpreting RMSE values, lower scores signify improved predictive accuracy. In summary, our evaluation strategy centers on RMSE due to its relevance, alignment with established practices, and capacity to offer a comprehensive view of prediction accuracy. The following sections will present detailed results and interpretations derived from these chosen metrics.

The Root Mean Square Error (RMSE) is calculated using the following mathematical expression:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2}$$

Where:

- N is the total number of observations.
- Y_i is the actual observed value for the i-th observation.
- $\hat{Y}_i$ is the predicted value for the i-th observation.

4.1.3 Result and discussion

Table 3 presents the results of the predictive models for the first experiment.

It has been observed that the RMSE value for predicting males and females together is often higher than when predicting them separately. Additionally, certain traits are better predicted for females, while others are better identified for males. This suggests that each gender expresses specific traits more prominently than the other. For instance, in their pictures, females tend to exhibit more agreeableness and extraversion, whereas males tend to display more conscientiousness and neuroticism.

It has been observed that personality traits can be more accurately predicted using Radial-Based Functional Networks (RBF) regression compared to other regression methods.

Table 3 Experience 1: (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism

	Random forest			Decision tree			Linear regression			Support Vector Regression			Radial basis function network		
	M	F	All	M	F	All	M	F	All	M	F	All	M	F	All
O	0.71	0.75	0.75	0.78	0.78	0.76	0.80	0.77	0.80	0.72	0.77	0.76	0.63	0.61	0.61
C	0.73	0.81	0.79	0.82	0.87	0.81	0.85	0.90	0.84	0.73	0.82	0.80	0.61	0.67	0.67
E	0.84	0.86	0.85	0.91	0.88	0.88	1.07	0.92	0.88	0.82	0.81	0.83	0.95	0.73	0.77
A	0.74	0.69	0.72	0.80	0.73	0.77	0.82	0.76	0.73	0.78	0.72	0.77	0.67	0.71	0.75
N	0.84	0.90	0.88	0.91	0.92	0.90	0.95	0.94	0.90	0.88	0.86	0.87	0.84	0.95	0.95

Confirming the findings of previous works [4, 12], it is noted that the personality traits that are easier to predict are openness, conscientiousness, and agreeableness. In contrast, extraversion and neuroticism are more difficult to predict.

Table 4 presents the results of the predictive models for the second experiment.

In comparison to the previous experiment, more accurate results were observed. It was explained that the probability of predicting a user's personality increases with the number of photos they post. Furthermore, the obtained results confirm the findings of the previous experiment: males tend to express conscientiousness and neuroticism in their images, while females tend to express extroversion and agreeableness.

Table 5 presents the results of the predictive models for the third experiment.

The experiment's results confirm our previous conclusions: predicting user personality traits is more accurate with more published images.

Table 6 presents the best results from each experiment to provide a comprehensive discussion of the findings.

Our observations indicate that there are gender-based differences in the expression of personality traits through images. Specifically, pictures of females tend to convey higher levels of agreeableness and extraversion, whereas images of males are more likely to depict conscientiousness and neuroticism.

A study on gender differences in personality traits [31] found a significant difference in Agreeableness between genders, with females scoring higher than males This result was confirmed by another study [32] and is consistent with the findings of our previous three experiments. Additionally, the study by [31] revealed a slight but noteworthy gender discrepancy in terms of Extraversion, with females scoring higher than males, which is also consistent with our results.

This study found that personality traits influence the content of Instagram photos. Accurate results were obtained for all traits, although predicting extraversion and neuroticism proved challenging.

4.1.4 State of art comparison

A comparative analysis between our most outstanding outcomes and those of previous works[4–8] on the prediction of personality traits using Instagram images was conducted in Table 7. The characteristics that were employed in previous studies of [4–8] are mentioned in Table 8.

An improvement in predicting certain personality traits compared to previous works is observed.

The trait of openness was predicted with higher accuracy, achieving an RMSE value of 0.59, surpassing all previous work for males [4–8]. For females, the value was 0.63, which means that the accuracy of predicting this trait exceeded the work of [4–6]. Additionally, our findings surpass those reported in previous studies [4–6, 8] when considering both genders. Furthermore, we achieved a higher accuracy in predicting the trait of consciousness for females compared to all previous works [4–8]. For males and mixed genders, we obtained an RMSE of 0.67, which is consistent with the values reported in previous studies [4, 8]. Moreover, for males, the prediction of the extraversion trait outperforms the work [4, 5, 8], and for females, the literature work [4, 5]. Nevertheless, we were unable to surpass previous studies in predicting the trait of agreeableness, except work ([4]).

In general, this study has succeeded in extracting most personality traits of Instagram users, surpassing the most results of previous studies on this topic. Moreover, a prediction

Table 4 Experience 2: (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism

	Random forest			Decision tree			Linear regression			Support vector regression			Radial basis function network		
	M	F	All	M	F	All	M	F	All	M	F	All	M	F	All
O	0.69	0.74	0.72	0.73	0.78	0.74	0.79	0.82	0.79	0.70	0.75	0.73	0.68	0.59	0.61
C	0.73	0.84	0.80	0.78	0.91	0.86	0.97	0.98	0.87	0.72	0.84	0.81	0.59	0.74	0.69
E	0.85	0.85	0.88	0.91	0.86	0.89	0.99	1.03	0.92	0.83	0.84	0.85	0.96	0.77	0.81
A	0.73	0.65	0.70	0.82	0.72	0.73	0.94	0.80	0.75	0.77	0.70	0.75	0.86	0.68	0.75
N	0.85	0.93	0.88	0.86	0.93	0.93	1.03	0.97	0.90	0.85	0.91	0.88	0.87	0.89	0.92

Table 5 Experience 3: (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism, M(ales), F(emales), All(Combine the two)

	Random forest			Decision tree			Linear regression			Support vector regression			Radial basis function network		
	M	F	All	M	F	All	M	F	All	M	F	All	M	F	All
O	0.70	0.73	0.71	0.78	0.72	0.76	0.80	0.82	0.81	0.70	0.72	0.72	0.72	0.62	0.61
C	0.70	0.81	0.79	0.74	0.86	0.88	0.97	1.07	0.88	0.66	0.84	0.79	0.53	0.72	0.67
E	0.86	0.85	0.90	0.92	0.90	0.92	1.11	1.14	0.95	0.86	0.85	0.87	0.88	0.80	0.82
A	0.67	0.66	0.71	0.79	0.71	0.75	0.91	0.82	0.75	0.73	0.67	0.75	0.80	0.65	0.78
N	0.74	0.96	0.86	0.75	0.99	0.93	1.15	1.1	0.91	0.75	0.95	0.86	0.81	0.99	0.98

Table 6 The best results were obtained in the three experiments for each regression method used: (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism, M(ales), F(emales), All(Combine the two)

	Random forest			Decision tree			Linear regression			Support vector regression			Radial basis function network		
	M	F	All	M	F	All	M	F	All	M	F	All	M	F	All
O	0.69	0.73	0.71	0.73	0.72	0.74	0.76	0.77	0.79	0.70	0.72	0.72	0.63	0.59	0.61
C	0.70	0.81	0.79	0.74	0.81	0.81	0.85	0.90	0.83	0.66	0.82	0.79	0.53	0.67	0.67
E	0.84	0.84	0.85	0.91	0.86	0.88	0.98	0.92	0.88	0.82	0.81	0.83	0.88	0.73	0.77
A	0.67	0.65	0.71	0.79	0.71	0.73	0.82	0.76	0.73	0.73	0.67	0.75	0.67	0.65	0.75
N	0.74	0.89	0.86	0.75	0.92	0.90	0.95	0.94	0.90	0.75	0.86	0.86	0.81	0.89	0.92

Table 7 Comparison of our results with those of previous work: (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism, M(ales), F(emales), All(Combine the two)

Year	[4]			[5]									[7]			
Number of users	113			193									179			
Population													European (German University Students)			
Number of photos	22 398			54 962									25 394			
Features type	Pixel features			Visual Prop			Content Prop			Comb			Content category	Face and emotion	Pixel features	All
Features	[A]			[A]			[B]			[A],[B]			[C]	[D]	[E]	[C] [D] [E]
Number of features	24			23			17			40			15	8	22	45
method	RBFN	RF	M5 Rules	M5 Rules	RF	RBF	M5 Rules	RF	RBF	M5 Rules	RF	RBF	RF			
O	0.68	0.71	0.77	0.77	0.73	0.72	0.72	0.71	0.71	0.76	0.75	0.71	0.62	0.64	0.64	0.59

Table 7 continued

Year	[4]			[5]									[7]			
C	0.66	0.67	0.73	0.62	0.65	0.61	0.60	0.63	0.63	0.74	0.61	0.62	0.60	0.61	0.58	0.55
E	0.90	0.95	0.96	1.02	1.06	0.99	0.95	1.04	0.97	0.99	1.05	0.98	0.61	0.56	0.61	0.59
A	0.69	0.71	0.78	0.64	0.60	0.59	0.57	0.58	0.62	0.61	0.62	0.61	0.58	0.58	0.56	0.57
N	0.95	0.95	0.97	0.79	0.88	0.89	0.87	0.81	0.89	0.87	0.89	0.89	0.67	0.65	0.68	0.66
Year	[6]	[8]		Our results												
Number of users	179	179		316												
Population	European (German University Students)	European (German University Students)		Moroccan												
Number of photos	25 394	25 394		19 898												
Features type	Pixel features	Content category	Pixel features	Content category	emotion	Pixel features										
features	[F]	[G]	[H]	[I]	[J]	[K]										
Number of features		29	15	16	7	12										
method	RF	RF		RF or DT or SVC or SVR or LR												
O	0.64	0.64		F: 0.63 , M: 0.59 , All: 0.61												
C	0.61	0.61		F: 0.59, M: 0.71, All: 0.67												
E	0.65	0.65		F: 0.82, M: 0.71, All: 0.79												
A	0.56	0.56		F: 0.67, M: 0.65, All: 0.70												
N	0.74	0.74		F: 0.63, M: 0.59, All: 0.61												

Table 8 Features used in the experiments described in Table 8

A	red, green, blue, yellow, orange, violet, saturation mean, saturation variance, saturation low, saturation mid, saturation high, value mean, value variance, value low, value mid, value high, warm, cold, pleasure, arousal, dominance, of faces, of people
B	architecture, body parts, clothing, music instruments, art, performances, botanical, cartoons, animals, foods, sports, vehicles, electronics, babies, leisure, jewelry, weapons.
C	abstract, animal, building, dark, drink, food, indoor, others, outdoor, people, plant, object, sky, text, and transportation.
D	face mean, anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise.
E	B mean, B_var, Gmean, G_var, R_mean, R_var, S_mean, S_var, V mean, V_var, Pleasure, arousal, dominance, red share, orange, share yellow, share green, share blue, share violet, share warm, share cold and share hue_peaks.
F	Colorfulness, Color diversity and Color harmony.
G	Food and drinks, Art, Animals, Clothing and accessoies, Botanical, Body parts, Home and interior, Architecture, Fantasy and fiction, Beauty and care, Sports, Landscape and nature, Music, Colors, Dance and performance, Fabric and material, People, Weapons, Vehicles and transport, Tools and machine, Electronics, Events, holidays and gatherings, Human emotions and behavior; Leisure and play, Business and education, Crafts, Services, Jewelry and Humans.
H	Average saturation, Saturation variance, Red, Orange, Yellow, Green, Blue, Purple, Warm, Cold, Normalized brightness, Contrast, Pleasure, Arousal and Dominance.
I	abstract, buildings, dark, object, people, cat, dog, sports ball, tv, laptop, cup, bottle, horse, handbag, percentage of photos with faces and face number.
J	angry, disgust, fear, happy, neutral, sad, surprise.
K	luminosité, red, orange, yellow, green_yellow, green, cyan, blue, violet, rose, warm colors, cold colors.

of the personality traits of each gender separately gives a clearer insight into the personality, since each gender differs in their preferences and published content.

Personality traits were successfully predicted. The aim is to identify the specific characteristics that enable accurate prediction of each trait. To achieve this, we will study the relationship between image traits and personality traits.

4.2 Correlation between image features and personality traits

This section examines the relationship between the five personality traits and the characteristics of Instagram users' images. Machine learning techniques are used to determine the influence of the dependent variables (value of each personality trait) on the independent variables (content category, emotion, pixels, and the number of images). A scatter plot of each feature against each parameter was plotted to identify the types of dependencies between the parameters. Figure 4 illustrates examples of the relationship between the openness trait and the dark feature.

It was observed that the majority of the curves exhibited linear dependencies. As a result, the Pearson correlation method was utilized.

4.2.1 Correlation results for females

For females, Table 9 shows a Pearson correlation between image characteristics and personality traits.

We have obtained connections between image features and personality traits, suggesting that the content of Instagram images is directly impacted by the personality traits of Instagram users.

Openness

The results of our analysis suggest that open users are more likely to post images of objects and animals other than people [1]. Among their publications are images of dogs, televisions, and abstract objects. However, this type of user usually posts pictures full of negative emotions [1], notably sadness and fear.

Conscientiousness

Conscientious users more frequently share images colored yellow. This type of user tends to post pictures of animals, especially dogs, equipment such as televisions and laptops, and objects such as cups and sports balls. Higher consciousness levels are associated with more images posted by users. In addition, disgust emotions and conscience were positively correlated [1].

Extraversion

Extraverts users publish images colored in cold colors, especially blue, with a higher percentage of images with faces [1]. In addition, we found that the higher the level of extraversion of users, the higher the number of photos they publish. Dogs and mugs have a positive correlation with this trait.

Agreeableness

In general, agreeable users post images that contain abstract objects, things, and warmer colors, such as orange. According to research [1], agreeable people tend to have pictures that contain positive emotions.

Neuroticism

Neurotic users tend to share bright images [4], and their photos do not contain animals. This type of user represents the emotion of anger and less disgust in their photos.

4.2.2 Correlation results for males

For males, Table 10 shows a Pearson correlation between image characteristics and personality traits.

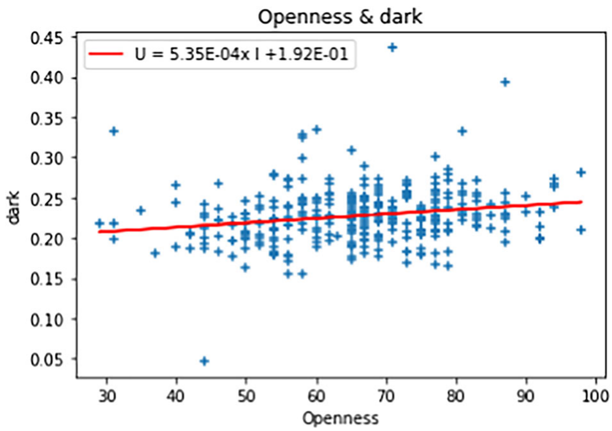

Fig. 4 Exemple of scatter plot between the dark feature and the openness trait

Table 9 Pearson correlation matrix between image characteristics and personality traits for females: (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism, M(ales), F(emales), All(Combine the two)

	O	C	E	A	N
Abstract	^^0.14	0.007	−0.04	^^0.15	0.06
Buildings	0.04	0.02	−0.08	−0.06	−0.03
Dark	0.07	−0.09	−0.06	−0.03	−0.11
Object	0.05	−0.02	−0.04	∗0.20	0.04
People	^^−0.15	0.03	0.09	−0.09	−0.02
Cat	0.04	0.06	0.04	−0.03	^^−0.16
Dog	^0.12	^0.13	^^0.14	0.002	−0.08
Sports ball	0.12	∗∗0.24	0.02	0.07	0.04
Tv	^0.13	^^0.14	0.008	−0.05	−0.08
Laptop	−0.02	0.07	0.08	−0.06	−0.02
Cup	0.07	^^0.16	^0.13	0.06	−0.06
Bottle	0	0.08	0.07	0	−0.06
Horse	0.09	0.01	0.08	0.04	0.04
Handbag	0.09	0.07	0.03	0.08	0.03
Percentage of the images contains faces	−0.07	0.0007	^^0.14	0.06	0.05
Number of faces	0.03	0.08	0.11	0.01	0.01
Angry	0.08	0.01	−0.02	−0.06	−0.02
Disgust	0.08	^^0.15	0.03	0.10	^−0.13
Fear	0.09	−0.02	0.07	0.02	0.008
Happy	−0.09	0.07	−0.02	0.04	−0.002
Neutral	0.08	^−0.12	−0.02	0.06	−0.008
Sad	0.10	−0.07	0.04	−0.10	0.04
Surprise	−0.09	0.08	−0.05	0.07	−0.03
Red	0.04	−0.06	^−0.12	−0.02	−0.03
Orange	0.03	0.03	−0.02	^0.13	−0.09
Yellow	0.05	^^0.16	−0.04	0	−0.008
Yellow green	−0.03	0.05	0.01	0.05	−0.07
Green	−0.05	−0.01	0.01	−0.04	−0.07
Cyan	−0.01	−0.007	0.03	^−0.12	−0.08
Blue	−0.01	−0.02	^0.14	−0.05	0.09
Violet	−0.08	0.05	0.04	0.04	0.05
Rose	−0.07	0.10	−0.05	0.008	0.10
Warm	0.05	0.01	^^−0.16	0.08	−0.05
Cold	−0.05	−0.01	^^0.16	−0.08	0.05
lumino	−0.01	−0.06	−0.10	0.08	0.08
N of photos	0.10	^^0.15	^0.12	0.06	−0.03

Note.^p<0.1, ^^p<.05, ∗p<01, ∗∗p<.001

Openness

These users post images other than faces, consistent with previous studies [1, 4]. This type of user prefers to post images containing abstract things, sports balls, and buildings. In addition, they publish fewer images containing people [1]. Users with a large aperture tend to have dark photos, rich in red color, high in negative emotions [1], especially sadness, and lower in positive emotions, such as happiness.

Conscientiousness

Table 10 Pearson correlation matrix between image characteristics and personality traits for males: (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism, M(ales), F(emales), All(Combine the two)

	O	C	E	A	N
Abstract	∗0.26	0.09	0.004	−0.10	0.14
Buildings	ˆ0.16	0.10	ˆˆ−0.17	0.10	−0.06
Dark	∗∗0.35	0.08	0.03	−0.06	0.02
Object	0.06	−0.007	−0.04	ˆ−0.14	0.13
People	∗∗−0.34	−0.12	0.08	0.04	−0.05
Cat	0.13	−0.03	−0.09	0.08	0.02
Dog	0.02	0.06	−0.01	ˆ0.16	−0.10
Sports ball	ˆ−0.16	ˆˆ0.20	0.06	ˆ0.16	−0.02
Tv	0.03	∗0.23	ˆˆ0.21	0	ˆ−0.15
Laptop	−0.09	ˆˆ0.18	0.04	0	−0.13
Cup	−0.08	ˆ0.16	0.04	0.13	−0.10
Bottle	−0.11	−0.008	−0.07	0.07	−0.02
Horse	−0.005	−0.01	−0.13	0.09	−0.04
Handbag	−0.04	0.03	0.02	0.02	−0.13
percentage of the images contains faces	∗∗−0.29	0.03	0.10	−0.04	ˆ−0.16
Number of faces	ˆ−0.15	0	−0.01	0.06	−0.08
Angry	0.07	−0.08	0.03	−0.11	0.08
Disgust	−0.04	−0.01	−0.04	−0.10	ˆˆ0.18
Fear	0.02	∗∗0.29	0.12	−0.09	−0.14
Happy	−0.08	−0.14	−0.10	−0.05	0.11
Neutral	−0.01	0.08	−0.11	−0.02	ˆ0.16
Sad	ˆ0.16	0.13	0.05	ˆ0.14	−0.01
Surprise	−0.11	−0.11	−0.007	−0.05	−0.09
Red	ˆ0.16	−0.06	0.13	ˆˆ−0.18	0.10
Orange	−0.07	0.11	−0.08	−0.01	ˆˆ−0.18
Yellow	−0.07	0.04	−0.08	−0.10	−0.06
Yellow green	−0.03	−0.05	−0.10	−0.12	−0.08
Green	0.06	0.04	ˆˆ−0.19	−0.11	−0.009
Cyan	0.04	0.10	−0.08	0.01	ˆˆ−0.22
Blue	ˆ−0.16	−0.09	−0.06	ˆ0.16	0.06
Violet	0.01	0.02	ˆ0.16	∗∗0.28	0.12
Rose	0.02	0.11	0.13	∗0.26	−0.02
Warm	0.11	0.04	0.10	ˆˆ−0.19	−0.01
Cold	−0.11	−0.04	−0.10	ˆˆ0.19	0.01
lumino	0.10	−0.07	0.04	−0.05	0.13
N of photos	−0.03	−0.02	−0.03	0.02	−0.06

Note.ˆp<0.1, ˆˆp<.05, ∗p<01, ∗∗p<.001

Images posted by conscientious users contain a large percentage of the objects, particularly TV, laptop, sports balls, and cups. Moreover, there is an inverse relationship between the trait of awareness and the characteristic of joy, which indicates that conscious users express negative emotions such as fear [1].

Extraversion

Extroverted users tend to have images with warm colors; notably, it's rich in red, purple, and pink. Users who are less extroverted tend to post images colored in green. Contain buildings, and their pictures do not contain TVs. In general, it remains a difficult trait to correlate with other characteristics.

Agreeableness

Users with this trait publish colorful images in cold colors, especially blue, purple, and pink. In addition, their pictures show aspects like animals (dogs), sports balls, and buildings.

Neuroticism

The photos published by nervous users are rich in negative emotions [1], such as disgust. The images posted by them are not colorful [1]. Nervous users like to post objects rather than faces [1].

5 Conclusion and future work

This study has produced satisfactory results in predicting the personality traits of Moroccan Instagram users, surpassing previous studies [4–8] in predicting openness for males and Neuroticism traits for females and males. Furthermore, it has been demonstrated that the prediction is more accurate when the personality traits of each gender are separately predicted. The study has also shown that the accuracy of personality trait prediction is affected by the number of images published by the user. Additionally, the accuracy of the prediction increased with the number of photos posted by the user. The study found that the prediction of Agreeableness and Extraversion traits was more accurate for females than males, while the opposite was observed for the Conscientiousness and Neuroticism traits.

The result of this study can be used in the development of personalized marketing strategies, recognizing that users' visual preferences can be influenced by their personalities. This understanding could enable brands and marketers to better tailor their messages and visual campaigns to elicit positive responses from their target audience.

Although our results offer valuable insights into the correlation between personality traits and Instagram content in Morocco, it is important to recognize that cultural differences may affect image content and expression. Therefore, we plan to replicate our study using a dataset that includes users from various countries to expand the scope of our research. It can serve as a basis for understanding how personality traits are manifested in diverse visual representations. Additionally, applying our study about personality prediction to a large database can contribute to the development of a more universally applicable predictive model.

In future work, we propose, in addition to the image properties, to integrate the properties of the legend of the image, such as (language style, text length, feeling expressed, hashtag, using emojis), the video features, such as classifying the type of music used, as well as the visual properties of the video, such as tracking movement. Additionally, we can analyze the type of publication made by the user, such as whether they usually publish only images, videos, or both. In addition, the number of followers can be considered. Combining all of these characteristics in one study can produce a more accurate predictive model.

Research Data Policy and Data Availability Statements The datasets generated and analyzed in the current study are not publicly accessible due to the presence of personal data and images of Instagram users. However, they are available from the corresponding author on reasonable request.

Declarations

Compliance with Ethical Standards The participants agreed to participate in this study by granting me access to their publicly available photos on Instagram and by completing the Big Five questionnaire sent to them.

Competing Interests This study did not receive any financial support, and the authors of this article have no financial or non-financial interests to disclose.

References

1. Naik A, Kulkarni A (2018) Personality prediction using profile pictures. Int J Eng Res Comp Sci Eng 5(7):221–226
2. Akshat D (2016) Application of convolutional neural network models for personality prediction from social media images and citation prediction for academic papers. Comp Sci
3. Ferwerda B, Tsakalidis A (2018) You are what you post: what the content of instagram pictures tells about users' personality. In: Proceedings of the 23rd international conference on intelligent user interfaces. ACM, pp 477–482. https://doi.org/10.1145/3172944.3172981
4. Ferwerda B, Tkalcic M (2016) Using instagram picture features to predict users' personality. In: International conference on multimedia modeling. Springer, pp 595–606. https://doi.org/10.1007/978-3-319-27671-7_71
5. Ferwerda B, Tsakalidis A (2018) Predicting users' personality from instagram pictures: using visual and/or content features? In: Proceedings of the 26th conference on user modeling, adaptation and personalization. ACM, pp 157–161. https://doi.org/10.1145/3209219.3209248
6. Kim JH, Kim Y (2019) Instagram user characteristics and the color of their photos: colorfulness, color diversity, and color harmony. Inf Process Manag 56(2):306–317. https://doi.org/10.1016/j.ipm.2018.10.018
7. Kim Y, Kim JH (2018) Using computer vision techniques on instagram to link users' personalities and genders to the features of their photos: an exploratory study. Inf Process Manag 54(1):320–330. https://doi.org/10.1016/j.ipm.2018.07.005
8. Branz L, Brockmann P, Hinze A (2020) Red is open-minded, blue is conscientious: predicting user traits from instagram image data. In: Proceedings of the 2020 CHI conference on human factors in computing systems. ACM, pp 1–14. https://doi.org/10.1145/3313831.3376693
9. Celli F, Benedetto E, Lepri B (2014) Automatic personality and interaction style recognition from facebook profile pictures. In: Proceedings of the 22nd ACM international conference on multimedia. ACM, pp 965–968. https://doi.org/10.1145/2647868.2654977
10. Hall JA, Pennington N, Lueders A (2013) Impression management and formation on facebook: a lens model approach. J Media Psychol 25(1):1–15. https://doi.org/10.1027/1864-1105/a000085
11. Christian G, Kamal DSMF, Suryaningrum KM (2021) Automatic personality prediction using deep learning based on social media profile picture and posts. In: 2021 4th International seminar on research of information technology and intelligent systems (ISRITI). https://doi.org/10.1109/ISRITI54043.2021.9702873
12. Quercia D, Kosinski M, Stillwell D, Crowcroft J (2011) Our twitter profiles, our selves: predicting personality with twitter. In: 2011 IEEE Third international conference on privacy, security, risk and trust and 2011 IEEE Third international conference on social computing. IEEE, pp 180–185. https://doi.org/10.1109/PASSAT/SocialCom.2011.26
13. Wu C-C, Hsu N-XM-S-CP-Y, Chen Y-Y (2022) Prediction of personality traits through instagram photo hsv. In: Human-computer interaction. Technological innovation. Springer, pp 279–287. https://doi.org/10.1007/978-981-16-5324-2_25

14. Ferwerda B, Yang E, Tkalcic M (2015) Personality traits predict music taxonomy preferences. In: Proceedings of the 33rd annual ACM conference extended abstracts on human factors in computing systems, pp 2241–2246. https://doi.org/10.1145/2702613.2732754
15. Skowron M, Tkalčič M, Ferwerda B, Schedl M (2016) Fusing social media cues: personality prediction from twitter and instagram. In: Proceedings of the 25th international conference companion on world wide web. International world wide web conferences steering committee, pp 107–108. https://doi.org/10.1145/2872518.2889368
16. Segalin PACMC, Vinciarelli A (2017) The pictures we like are our image: continuous mapping of favorite pictures into self-assessed and attributed personality traits. IEEE Trans Affect Comput 8(3):268–285. https://doi.org/10.1109/TAFFC.2016.2516994
17. Wright AGC (2014) Current directions in personality science and the potential for advances through computing. IEEE Trans Affect Comput 5(3):292–296. https://doi.org/10.1109/TAFFC.2014.2332331
18. Lay A, Ferwerda B (2018) Predicting users' personality based on their 'liked' images on instagram. IEEE Trans Affect Comput 9(3):268–285. https://doi.org/10.1109/TAFFC.2017.2700578
19. McCrae RR, Costa PT (1997) Personality trait structure as a human universal. Am Psychol 52(5):509–516. https://doi.org/10.1037/0003-066X.52.5.509
20. Costa Jr PT, McCrae RR (1992) Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI): professional manual. Psychological Assessment Resources, ???
21. Zhang K, Zhang Z, Li Z, Qiao Y (2016) Joint face detection and alignment using multi-task cascaded convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), Las Vegas, NV, USA, pp 5297–5305. https://doi.org/10.1109/CVPR.2016.573
22. Zhang K, Zhang Z, Li Z, Qiao Y (2016) Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Process Lett 23(10):1499–1503. https://doi.org/10.1109/LSP.2016.260334
23. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), Las Vegas, NV, USA, pp 779–788. https://doi.org/10.1109/CVPR.2016.91
24. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324. https://doi.org/10.1109/5.726791
25. Breiman L (2001) Random forests. Mach Learn 45(1):5–32. https://doi.org/10.1023/A:1010933404324
26. Quinlan JR (1986) Induction of decision trees. Mach Learn 1(1):81–106. https://doi.org/10.1007/BF00116251
27. Broomhead DS, Lowe D (1988) Radial basis functions, multi-variable functional interpolation and adaptive networks. Technical report, Royal Signals and Radar Establishment, Malvern, England
28. Vapnik V (1995) The nature of statistical learning theory. Springer, ???
29. Seber GAF, Wild CJ (2003) Nonlinear regression. Wiley-Interscience, ???
30. Pearson K (1895) Notes on regression and inheritance in the case of two parents. Proc R Soc Lond 58:240–242
31. Costa PT, Terracciano A, McCrae RR (2001) Gender differences in personality traits across cultures: robust and surprising findings. J Pers Soc Psychol 81(2):322–331
32. Weisberg YJ, DeYoung CG, Hirsh JB (2011) Gender differences in personality across the ten aspects of the big five. Front Psychol 2:178