Physiognomy in New Era: A Survey of Automatic Personality Prediction Based on Facial Image

Xu Jia[1,2(✉)], Weijian Tian[2], and Yangyu Fan[2]

[1] North China University of Science and Technology, Tangshan 063210, China
Xujiajia_2008@163.com
[2] Northwestern Polytechnical University, Xi'An 710129, China

Abstract. At present, personality computing technology facilitates the understanding, prediction, and management of human behavior. With the increasing importance of faces in personal daily assessments, establishing a relationship between facial morphological features and personality traits is a major breakthrough in personality computing technology. This paper is a survey of such technology of automatic personality prediction based on face and it aims at providing not only a solid knowledge base about the state-of-the-art in automatic personality prediction, but also to provide a conceptual model of automatic personality prediction, based on the literature. In addition, the analysis of the prediction results of the existing researches is emphasized, and there are still problems in the field, such as lack of information on research data, single age group of the sample population, incomplete design characteristics of the artificial design etc., and the potential applications and development directions are determined.

Keywords: Personality prediction · Face recognition · Machine learning

1 Introduction

1.1 Status Quo of Research on Personality Calculation

Since the Greek philosopher Theophrastus (371 BC - 287 BC), people have been keen on the study of personality [1]. As a construct, personality aims at capturing stable individual characteristics, typically measurable in quantitative terms, that explain and predict observable behavioral differences [2]. Current research on the personality model successfully predicts the relationship between 'thinking, emotions and behavioral patterns' [3] and the important aspects of life (including 'happiness, physical and mental health, quality of interpersonal relationships, career choices, professional satisfaction and professional performance, participation in social activities, criminal activities and political ideology,' etc. [4] have had some beneficial effects. Furthermore, attitude and social behavior towards a given individual depend to a significant extent, on the personality impression others develop about her [5].

Given all of this, the ability to reliably, effectively and efficiently evaluate personality is a valuable goal, and with the advent of the era of big data, modern computer

B. Li et al. (Eds.): IoTaaS 2018, LNICST 271, pp. 12–29, 2019.
https://doi.org/10.1007/978-3-030-14657-3_2

science has the practical potential to advance this effort. In recent years, research on personality has been moving in some interesting directions, such as the combination of personality analysis and human-computer interaction technology. Figure 1 shows the number of papers including the word "personality" in the title on IEEE Xplore and ACM Digital Library, probably the two most important repositories of computing oriented literature. While being only the tip of the iceberg, most articles revolving around personality do not mention it in the title, these papers clearly show that the interest for the topic is growing and that the trend promises to continue in the foreseeable future.

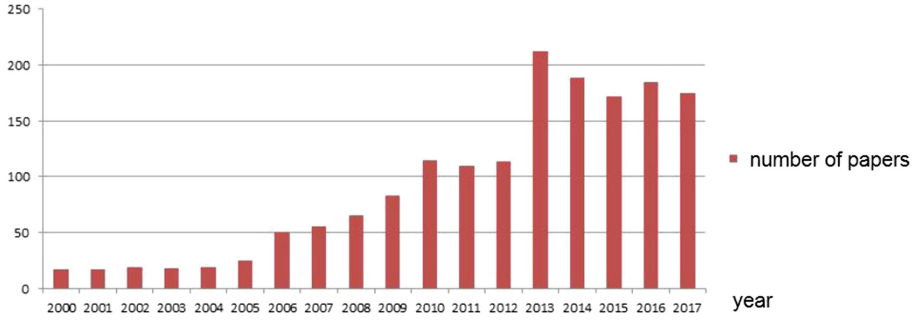

Fig. 1. The chart reports the number of papers per year with the word "personality" in their title (sum over IEEE Xplore and ACM Digital Library).

Some studies have shown that personality traits can predict user behavior by establishing an interaction between personality and automatic computation [6, 18, 32, 43, 52, 54]. Overall, personality is relevant to any computing area involving understanding, prediction or synthesis of human behavior. Still, while being different and diverse in terms of data, technologies and methodologies, all computing domains concerned with personality consider the same three main problems: That is, (1) selection of personality measurement methods, (2) selection and construction of data sets, and (3) establishment of calculation models. Through research, we found that there are many researches on automatic personality prediction, including: text-based [52], social media-based [6, 32, 54], mobile-based [16, 18], and computer games-based [43] and so on. However, there are few studies on automatic personality prediction based on human facial images. Figure 2 shows the number of papers containing the words 'face' and 'personality' in the headings of the IEEE Xplore and ACM digital libraries. Only a few articles have been searched since 2010. This paper mainly investigates the work of establishing the connection between personality and the use of computing technology based on face recognition. This is one of the main research issues in the development of personality calculation methods. Based on what we know, this is the first investigation of a solution to these problems.

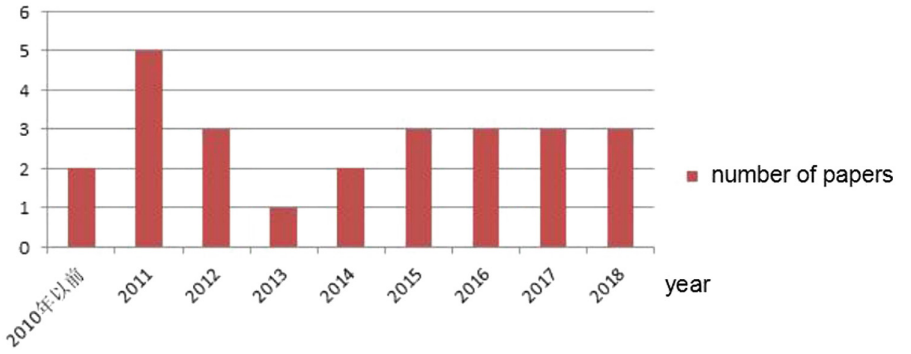

Fig. 2. The chart reports the number of papers per year with the word 'personality' and 'face' in their title (sum over IEEE Xplore and ACM Digital Library).

1.2 Status Quo of Face-Based Personality Prediction Research

Even in ancient China, Egypt and Greece, people had tried to establish the relationship between facial morphological features and personality traits of an individual. The modern psychological studies revealed that people tend to evaluate others on their appearance and then move on to interact with them based on these first impressions. Nowadays, it has been a well-established fact the face plays a central role in the everyday assessments of other people. For example, humans perform trait judgments from faces and the results of this unconscious behavior can sometimes decisively affect the results of important social events, such as an electoral process [7, 23], crime tracking [53], etc.

Currently, the following two points hold on automatic face evaluation: Firstly, some self-reported personality traits and the intelligence can be evaluated by the human based on the facial features to a certain extent. Secondly, the commonalities existing in the evaluation behavior of the human can be mined by the machine learning methods. The following are some of the relevant work in the literature.

The research on the first point mainly focuses on the following points: In [32], the authors studied the human tendency to evaluate others on their faces and identified some important facial features that generate first impressions. In [11], Humans can make valid inferences for at least four personality traits from facial features. In [52], the relationship between self-reported personality traits and first impressions was studied. Study on the second point: To investigate whether the trait evaluations performed by humans can be learned by computers, [41] used machine learning methods to construct an automatic trait predictor based on facial structural descriptors and appearance descriptors. They found that all the analyzed personality traits could be predicted accurately.

The rest of this paper is organized as follows: Sect. 2 introduces the concept of personality (with particular emphasis on trait based models and the Big Five) and the techniques for its measurement. Section 3 introduces the establishment of data sets in the automatic prediction model in different literatures. Section 4 introduces the extraction of face personality traits. Section 5 investigates the algorithm of face-based

automatic personality prediction in different literatures. Section 6 aims to summarize the forecast results in the existing research and analyze the reasons; Sect. 7 analyze the problems that still exist in the field, and determine the potential applications and development directions.

2 Personality Model and Its Measurement

Personality psychology aims at predicting observable individual differences based on stable, possibly measurable, individual characteristics [2]. The theory of personality is representative of trait theory, type theory, and integration theory. Different theories use different' intrinsic attributes' as the basis of personality, including physiology (biological perspective), unconsciousness (psychological perspective), environment (behavioralist perspective), internal state (humanistic perspective), and mental (cognitive perspective), etc. (see extensive survey [8, 16]). However, the model that most effectively predicts the measurable aspects of people's lives is a model based on traits—the personality trait model, which is widely regarded as the 'structural one of the major achievements of psychology' [15]. On the other hand, several decades of research and experiments have shown that the same traits appear with surprising regularity across a wide spectrum of situations and cultures, suggesting that they actually correspond to psychologically salient phenomena [15]. This regularity and stability brings inspiration to the calculation of personality.

2.1 The Choice of Personality Trait Model

The trait theory holds that traits are the basic characteristics that determine individual behavior, are effective constituent elements of an individual, and are also the basic units commonly used for measuring personality. The main trait theories are: Alport's trait theory (common traits and personal traits), Cartel's personality trait theory (he believes that each person has these 16 traits), and Essen's three-factor model (outside (Pantastic, psychotic, and neurotic), the five-factor (large five personality model) model of Tarpez (external, pleasant, responsible, neurotic, and open), and the seven-factor model of Trigan (positive, negative, Titer, positive price, negative sentiment, reliability, appetite, and comparability). How to choose a trait model for our personality calculation is one of the issues discussed in many studies.

We found that the five factors of Tapes (the big five model can get relatively true and reliable personality traits, which laid the foundation for the accuracy of the experiment. Most of the works surveyed use the five-factor (Big Five) personality trait (More than 50% in the literature use the 'Big Five'). On the one hand, other theories can hardly reflect the dominance of feature-based models in personality psychology. On the other hand, the 'Five Factor Model (Big Five)' numerically expresses personality (see below), which is a form that is particularly suitable for computer processing. The Big-Five traits are as follows:

- Extraversion: Active, Assertive, Energetic, Outgoing, Talkative, etc.
- Agreeableness: Appreciative, Kind, Generous, Forgiving, Sympathetic, Trusting, etc.
- Conscientiousness: Efficient, Organized, Planful, Reliable, Responsible, Thorough, etc.
- Neuroticism: Anxious, Self-pitying, Tense, Touchy, Unstable, Worrying, etc.
- Openness: Artistic, Curious, Imaginative, Insightful, Original, Wide interests, etc.

Attempts were made to enrich the trait set with more dimensions, but 'Five- factor solutions were remarkably stable across studies, whereas more complex solutions were not' [16]. Other models considered less traits [13], but these still appeared to be linear combinations of theBig Five. In other words, the Big Five 'provide a set of highly replicable dimensions that parsimoniously and comprehensively describe most phenotypic individual differences' [43].

2.2 Existing Research Personality Model Selection

In the existing face character prediction works, their personality measurement methods are as follows: Qin et al. In the face-based personality and intelligence prediction research [37, 38], they uesd self-assessment, Cartel's Sixteen Personality Questionnaire (16PF) was used to measure participants' personality traits; Rojas et al. in the evaluation of personality based on facial feature points [40], they used others' assessments, each evaluator wrote an un-constrained description from a set of 66 standardized faces from the Karolinska [30] amateur actors face database. 1134 descriptions were collected, The researchers' classification of the unconstrained descriptions, resulted in 14 selected categories; In the study of Karin et al.'s interpretation of the impact of facial features on first impressions and personality [52], used the Cubiks Indepth Personality Questionnaire, CIPQ2.0, a normative self-report questionnaire scoring 17 personality traits covering the Big Five was used to measure the participants' personality traits; In his dissertation [54], Zeng also used the Big Five Personality Scale to test the Big Five personality traits using statistical analysis; In the study [6] of Al Moubayed et al., they used the method of others' assessments. The "Big Five Scale" was provided by 11 independent judges with white British ethnic background, normal vision, and no hearing problems; Oosterhof et al. [32] used Others' assessments, in the first phase of the project, they asked 55 participants to characterize the unrestricted faces, then classify these descriptions as feature dimensions, and extracted 14-dimensional feature dimensions, which are highly consistent with Rojas et al. [41]. The personality assessment methods used in all relevant studies are shown in Table 1.

Table 1. Personality assessment model in face-based personality prediction research.

Ref	Assessment type	Trait theory	Characteristic expression
[6]	Others' assessments	Big Five	Score(1–9)
[8]	Others' assessments	Manually select: Dominance, Warmth, Sociability and Credibility	Score
[9]	Others' assessments	Manually select: Intelligence, Maturity, Warmth, Sociality, Dominance and Credibility	Score(1–3)
[32]	Others' assessments	Manually select personality descriptors	Score(1–9)
[30, 41]	Others' assessments	Manually select personality descriptors	14-dimensional
[37, 38]	Self-assessments	16SP	Score(0–9)
[52]	Self-assessments	Big Five	Score(1–9)
[54]	Self-assessments	Big Five	Score(1–60)
[53]	Others' assessments	Manually select: Dominance, Attraction, Credibility and Extraversion	Score

3 Construction of Data Set

Focusing on face-based personality prediction, the database of experiments should include two parts: face image database and personality evaluation database. Their collection is the basis of the entire prediction process. The contents of this section will provide a brief summary of the current status of database construction in the field and the existing problems.

On the one hand, the current domestic and international research results on face-based personality prediction are few, and it is difficult to obtain open experimental data sets. On the other hand, since personality research involves participants' personal privacy, experimental data sets are not Suitable for public. Summing up the existing research results, we found that for the face database, the vast majority of them are self-built database with a small number of face databases [6, 8, 9, 23, 32, 37, 38, 41, 52–54]; For the personality evaluation results database, using the methods of self-assessment or others' assessment to measure and obtain the final personality data as described in the previous Sect. 2.2.

3.1 Construction of Face Database

The selection of face data is the key to verifying the face-based automatic personality prediction model and guaranteeing its generalization ability. Ideally, face databases should include face samples of different gender, race, and age under different

personality traits. However, so far, in the field of personality calculation, there is still no such a recognized database. In the field, different jobs often build independent databases based on their actual situation. The number of faces, ages, genders, races, postures and expressions of these face samples are different. Table 2 briefly describes the basic situation of the face database built in the existing personality calculation study.

From Table 2, it can be seen that the existing face database based on automatic personality prediction study contains the number of samples ranging from tens of samples to several thousand samples. Rojas et al. [41], developed a personality prediction model based on a set of 66 standardized faces in the database of amateur actors' faces of Karolinska [30] as a preliminary attempt on the personality calculation of face-based. Most of the research is based on young people in colleges and universities. For example, Qin et al. [37, 38] collected facial images of Xiamen University of Technology students. Zeng et al. [54] collected facial images in a college in Jiangxi province. Wolffhechel et al. [52], also collected images of college students' faces in Danish universities. This is because the research scholars are basically in scientific research institutes such as colleges and universities. They are more convenient when collecting data in this area, which facilitates the development of experiments, but at the same time it also brings a problem of a single age structure. It is worth mentioning that in order to enhance the generalization ability of personality calculations in real applications, some researchers have abandoned facial samples from laboratory environments and downloaded diverse face images from social networking sites. For the first time, White [51] and others selected thousands of images from the www.hotornot.com website for building a face database. This can be considered as a new attempt in the area of personality calculation. [8, 9] uses software-generated face data, using the popular composite software program FACES [20] produced by Inter Quest and Micro. Through FACES, it is possible to randomly generate a large number of unique faces by manipulating individual facial features. Images can be limited by many external conditions.

3.2 Collection of Personality Trait Data

A major problem in the study of personality calculations is the lack of real personality traits data, and other relevant research, such as crime prediction [53], beautiful attraction calculation [28, 29], expression recognition [21, 50], gender [34], true and false face recognition [7], etc., the classification tag is easy to obtain, and the acquisition of personality traits is much more difficult. In the attractive attraction calculations, such as [28, 29], it is only necessary to find enough observers to score face samples. Each face sample corresponds to a score; In facial expression recognition [21, 50], the definition of expression is also fixed, such as 'happy', 'sad', etc. The observer also only needs to give a label to the face sample; Crime judgments, gender judgments, and true/false face judgments simplify calculations to a two-category problem, making it easier to classify. However, the collection of personality traits data, whether self-evaluation or other people's evaluation, must be evaluated by the personality rating scale. Many scales have a large number of questions, and it is possible to measure a single sample for several tens of minutes to obtain the results of personality traits. Each sample must correspond to the personality traits of multiple dimensions and need to be

Table 2. Attributes of Face data set in existing studies.

Ref	Num	Age	Gender	Race	Posture	Expression	Data sources
[1]	829	20–39	men& women	Mix	Positive	Neutral	Color-FERET Data set
[3]	220	Synthesis	men& women	Mix	Positive	Neutral	Synthesis by FACES software
[4]	480	Synthesis	men& women	Mix	Positive	Neutral	Synthesis by FACES software
[19]	650	30–50	men& women	Mix	Positive	Neutral	Image of politician
[29]	66	20–30	men& women	White	Positive	Neutral (Accessories and cosmetics are visible)	Karolinska's [26] amateur actor's face database
[35, 36]	186	18–22	men& women	Asia	Positive	Neutral	Students of Xiamen University of Technology
[39]	66	20–30	men& women	White	Positive	Neutral (Accessories and cosmetics are visible)	Karolinska's [26] amateur actor's face database
[51]	3998	18–25	men& women	Mix	Positive	different	Download in Social website
[52]	244	18–37	men& women	White	Positive	Neutral	Students of Danish Technical University
[53]	1856	18–55	men	Asia	Positive	Neutral	Contains two subsets of non-criminals and criminals
[54]	608	18–22	men& women	Asia	Positive	Neutral	Students of one University in Jiangxi

evaluated one by one. The process is complicated and slow, but of low precision. Therefore, the acquisition of personality traits data is very difficult, which is one of the important reasons for little research on automatic personality prediction.

In the existing studies, the assessors usually evaluate the personality of the face database with reference to different Likert scales, and finally calculate the scores of different personality traits through statistical calculations. Value-based, including 3-point system [9], 9-point system [6, 32, 41, 52], 10-point system [39, 40] and 60-point system [54]. Finally, the average score of each personality trait in each sample face is often used as a benchmark score for the machine prediction criteria. Table 1 summarizes the various attributes of the evaluators in the existing study including the scoring system.

4 Expression of Facial Character Traits

The expression of face personality traits is a crucial step in the face-based automatic personality prediction framework. On the one hand, because the pixel value feature of the face image is usually high in dimension and the amount of data is large, on the other hand, since the face similarity is high, the original pixel value data contains more redundant components. Before the experiment based on the face image is performed,

Table 3. Characterization of face character traits.

Ref.	Type of feature	Range of feature	Expression of feature
[6]	Geometric	Eyes, nostrils, chin tips, lip corners	34-dimensional area
	Holistic	Full face	Eigenface
[8]	Texture, geometry and overall three feature vectors	Full face	Pixels, PCA
[9]	Texture	Full face	Gabor, LBP, Pseudo-sliding window
[23]	Geometric	Head, eyebrows, eyes, nose and mouth	76 feature points and distances, ratios
	Texture	Full face	HOG
	Color	Full face	RGB
[30]	Texture	Full face	HOG
[32]	Holistic	Full face	Pixels, PCA
[37]	Texture	Full face	HOG, LBP, Gabor, GIST, SIFT
[38]	Geometric	Face profile, eyebrows, eyes, cheeks, nose, mouth	21 points, distance and ratio of 1134 dimensions
[41]	Geometric	Face profile, eyebrows, eyes, cheeks, nose, mouth	21 points, distance and ratio of 1134 dimensions
[52]	Texture	Full face	Pixels, PCA
	Holistic	Full face	PCA
			(32 PCs for male face, 35 PCs for female face)
[53]	Geometric	Face profile, eyebrows, eyes, cheeks, nose, mouth	FGM feature generation machine
	Texture	Full face	LBP, HOG
	Holistic	Full face	PCA
	Synthesis	Full face	Mosaic of
[54]	Geometric	Face profile, eyebrows, eyes, cheeks, nose, mouth	Improved ASM, distance and ratio of 32 feature points

features of the face image need to be extracted to reduce the dimension of the face feature, so as to extract information that can discriminate between faces.

In the existing research literature on face-based automatic personality prediction, the expression of face features can be divided into two categories: Feature-based Representation and Holistic Representation. Feature-based methods use geometric features, texture features, color features, and other local features to represent faces. The geometric features can be composed of coordinate values of the calibrated facial feature points, the distance between feature points, distance ratio, etc. [37, 38, 41]; the texture features can be Gabor filters [41] or local binary values. Local Binary Patterns (LBP) operator [39] to extract. Active Shape Model (ASM) and Active Appearance Model (AAM) parameters combine geometric shape and texture information and are also extracted as potential features [54]. The global expression method uses the original grayscale image to represent the human face, splicing it in rows as an input feature vector, and obtaining it through dimension reduction methods such as principal component analysis (PCA) or manifold learning. Required low-dimensional features such as feature faces [22, 51], manifolds [34], etc. Table 3 summarizes representative works of expression of face character traits.

4.1 Personality Prediction Based on Feature Expression

Face character traits are characterized and described by a diverse set of features, including geometry, textures, colors, shapes, etc., where geometric features are the most widely used features in the field. The basis for measuring the face geometry information lies in the accurate positioning of the critical areas of the face (i.e., face contours, eyebrows, eyes, nose, and mouth) and their feature points. Feature points can be automatically retrieved by the feature location method [27], or manually extracted from the graphical user interface [27]. The number of feature points selected in different research work varies from 21 points [30, 37, 38, 41], 39 points [54] to 76 points [27].

4.2 Personality Prediction Based on Overall Expression

The overall expression method is proposed to overcome the limitations of artificial design features based on feature expression methods. The feature-based method extracts features that are often local and discrete face features, while the holistic approach extracts global features from the entire face and studies the spatial relationships between features. Since the human brain's perception of personality is a holistic processing model, such methods are also widely used in face analysis applications.

In the overall expression method, a typical face expression is to splice the original grayscale image into a high-dimensional feature vector. As an effective subspace projection method, PCA reduces the dimensions of the original high-dimensional human face space to obtain a low-dimensional human face feature, such as a feature face [6, 41]. The literature [41] applied eigenfaces to personality prediction studies for the first time. The eigenface method was used to analyze pixel information [47]. Based on information theory, and intended to find the main components of face distribution, that is, the algorithm projects the image across the group. The most significant change in the feature space is in the image. [41] Rojas et al. used two kinds of feature

extraction schemes to determine whether the information conveyed by the overall characterization is complementary to the information expressed by the structural information. The conclusion is positive. In [6], the eigenface employed is the principal components extracted from a set of training images, which form the basis in the space where each point corresponds to a different facial image. The eigenfaces are sorted by the amount of difference they occupy in the original data. In the experiment, the extracted contribution rate was 90% before the 103 feature faces. [8, 32] used principal component analysis (PCA) to extract components that account for most of the changes in facial shape.

Face character prediction based on the overall expression depends on the apparent features of the face, and does not require manual calibration, so that the prediction process is fully automated.

5 Face-Based Character Prediction Algorithm

Based on the previous research, almost all studies have obtained the necessary data, including the characteristics of face data and the scores of personality traits of different standards. Next we need to use these data to accurately infer a person's personality Traits.

5.1 Personality Inference Experiment Based on Classification Model

Classification of Personality Trait Scores. From Sect. 3.2, we summarize the acquisition of personality traits data in relevant literature. Whether it is the 'Big Five Personality' or '16PF' assessment method, the study will be divided into different dimensions of personality traits, including a 3-point system. [8], 9 points [6, 32, 41, 52], 10 points [37, 38], and 60 points [54]. Corresponding to the discrete scores within each interval, most of the studies are to N-value the score of personality traits (N generally takes 2 or 3) in order to construct a classification problem. For example, in [37, 38], for the 20 kinds of personality traits of '16PF', the score between 6–10 is defined as a category, indicating that the performance of such personality traits is more significant, and the score is between 1–5. The other category shows that this personality trait is not obvious. Similarly,in [41], first, the five with the highest personality traits are classified as having "traits", and the lowest 5 points as 'non-characteristic' categories, which then form two categories. [54] differs from the previous studies in that personality traits are divided into three categories (e.g., low-amenity, medium-amenity and high-amenity). In [6], in order to facilitate the classification, only the extreme values (highest and lowest) of the personality trait scores are retained, which are divided into two categories. The main reason is that a higher consensus was reached between the evaluation members of the extreme values. In [23], for the evaluation of the corresponding personality traits of politicians, a positive and negative two-classification method was performed. The prediction of criminality in [53] itself is a question of a two-category (criminal?).

Selection of Classifiers. Different studies have also used different classification methods to study the problem of inferring personality traits based on human facial features. Qin et al. [37, 38] used the Parzen window [35], the Decision Tree [17], K-Nearest Neighbor(KNN) [38], Naive Bayes [38] and Random Forest [10]. In Rojas et al. [30, 41], the classification in the experiment was performed using the most advanced classifier. Five algorithms are used: GentleBoost as an example of an iterative method, support vector machine(SVM) has a radial basis function kernel as an example of a non-linear classifier, and K-nearest neighbor as an example of a nonparametric classifier with Parzen window and binary decision tree of Random Subspace. The system was evaluated using a 20x cross validation strategy and the results were supported by confidence intervals calculated with 95% confidence. Zeng et al. [54] adopted a deep confidence network classification algorithm based on BP algorithm. A deep confidence network structure model including a 5-layer RBM is used, where the fifth RBM is the output layer and the output is three types (low, medium, and high of the personality trait). Al Moubayed et al. [1] also used a support vector machine(SVM) binary classification method. Jungseock et al. [19] used the (RankSVM) [18] method to predict the social dimension in order to train the model. Unlike the general SVM algorithm, RankSVM is more suitable for their task because it aims to retain the training examples. The pre-specified pairwisesort order. This advantage makes RankSVM very popular in the literature of information retrieval (web search) and the recent related properties in the field of computer vision [34]. Wu et al. [53] constructed four classifiers to identify the relationship between criminality and human face: logistic regression [14], K-nearest neighbor(KNN), SVM, and convolutional neural network [26]. In [9], Brahnam and Nanni et al. use SVM and neural networks as decision rules (Table 4).

Table 4. Summarizes the selection of classification methods in different literature.

Ref.	Classification type	Classifier selection
[6]	Binary classification	SVM
[8]	Binary classification	SVM
[9]	Binary classification	SVM & neural network
[23]	Binary classification	RankSVM
[30, 41]	Binary classification	GentleBoost, SVM, K-Nearest Neighbor, Parzen Window and Binary Decision Tree
[37, 38]	Binary classification	Parzen Window, Decision Tree, K Nearest Neighbor, Naive Bayes and Random Forest
[53]	Binary classification	Logistic regression, K-nearest neighbor, support vector machine and convolutional neural network
[54]	Three classification	Deep confidence network based on BP algorithm

In the classification experiments, the classification accuracy rate used as the evaluation criteria, we can also see that most of the literature classification results are good (see below).

5.2 Individual Inference Experiment Based on Regression Model

Section 5.1 mainly introduces the experiments based on the morphological characteristics of different faces in the existing literature, using a variety of classification methods to infer human personality traits. In this section, the regression method used in the literature is analyzed to estimate the actual values of personality traits based on the morphological characteristics of the face. Qin Rizhao and other scholars are the first to introduce regression models in personality prediction experiments. In [37, 38], the scores of each personality trait obtained by the 16PF test are integer discrete scores ranging from 1–10. They did not make any changes and they directly used these values as regression targets. In regression experiments, Mean Square Error (RMSE) was used to measure the performance of regression experiments. Qin Rixi and others used the facial features and texture features to perform regression experiments, respectively. Compared with the classification experiments, the regression algorithms obtained errors are relatively large, indicating that the use of human facial images to predict the specific score of personality traits is still difficult of.

6 Forecast Analysis

The question of how much personality trait information can be learned from human face images can be studied around from a computer science perspective. Table 5 summarizes the conclusions of different research results. In general, the existing studies

Table 5. Summary of accuracy results of personality prediction in different literatures

Ref.	Characteristics with better prediction results	The highest accuracy
[26, 27]	Responsibility(male)	81.56%
	Responsibility (female)	82.22%
	Skeptical (male)	72.64%
	Skeptical (female)	82.22%
[28, 29]	Dominant	91.23%
	Threatening	90%
	Extraversion	
[30]	Open, Striving, Dominant (female)	63%
	Reliability, friendliness, responsibility (male)	65%
[2]	Neuroticism	82.35%
	Extroverted	84.31%
	Rigor	84.31%
[31]	Openness, Extraversion, Neuroticism	65%
[42]	Criminality	89.51%
[53]	Warm	76%
	Reliability	81%
[54]	Intelligence, maturity, sociality, dominance, warmth, credibility	80%

based on the automatic personality prediction of human face mainly have the following conclusions:

1. The calculation and prediction of personality traits based on facial images is reliable to a certain extent.
2. Partial personality traits ('responsibility', 'dominant', 'threatening', 'attractiveness', 'openness', 'extroversion' and 'nervousness', etc.) are more relevant to face images. These personality traits can achieve better prediction results. The other part of the environmental impact of personality trait prediction results are not satisfactory.
3. The application of the latest deep learning method can automatically learn more advanced face expressions from face images. The prediction results are more breakthrough than the traditional whole and feature methods, but the interpretation of extracted features is not intuitive and easy to understand [27].

7 Limitations and Prospects

Automatic personality prediction based on face images is a new research topic in the field of computer vision. Since 2008, some scholars have started to engage in this research, and there are few research results. However, after a small amount of research work, some preliminary and valuable results have also been obtained. However, there are still many problems worthy of deep thinking and exploration. In particular, the most important question in face personality prediction, namely 'what factors do people face contribute to the judgment of personality, and how much contribution', is still far from being explained better.

7.1 Existing Problems in Existing Research

Summarizing the existing research, the prominent problems in the field of automatic personality prediction based on face images at the current stage are mainly reflected in the following aspects:

1. Lack of information on research data: In the current research on personality prediction, all studies reviewed in this paper are based on two-dimensional human faces, especially two-dimensional frontal face images. However, simply relying on positive images will lose a lot of personality-related information. From ancient times to the present, Chinese facial studies mentioned that personality-related descriptions of 'five features', 'three courts', and 'twelve houses' are related to certain prominent facial regions (e.g., Forehead (rich in the vestibule), nose (big and tall), cheekbones (broad and wide), chin (ground radius, etc.). The described key features can only be accurately located from the side face. Recent studies have found that lateral faces can describe face features well and can be used for identity recognition [20], gender and race recognition [43], and face recognition [34]. Therefore, future face character analysis based on 2.5D or 3D is an inevitable trend in future face character prediction research.

2. Sampling of the age structure of the sample population: In all the existing studies, the convenience of the experimental implementation and the authenticity of the data were fully taken into account. The majority of the sample population were college

students. The age range was basically 18–35. Although many studies have fully considered the different disciplines [54], different genders [37, 38], and different occupations [30, 41], the diversity of samples has been reflected, but the well-known Chinese psychologist Song Chubo In the 'Heart to Heart' [45], it is said that 'there is no heart, no phase, no heart, no heart, no heart. The words are simple, and the truth of the human relations program', the face of the people above the age of 35 and their personality. Therefore, it is desirable to collect data for more than 35-year-olds and predict personality, in the hope that there will be more objective results in predicting accuracy.

3. Artificial design features are not comprehensive: In the existing research work, researchers usually manually design a set of features (geometry, texture, color, global appearance features, etc.) based on heuristic criteria to predict personality traits. A large number of diversified features have been constructed and analyzed for different databases, but so far there is no unified conclusion to prove which features constitute the main factors affecting personality traits. The artificial design features that have been proposed so far have generally been unsatisfactory in character prediction research, and it is therefore reasonable to conclude that these features are not comprehensive. Is there a specific area related to facial feature assessment on the person's face? How much is the contribution to character prediction? This important aspect of the issue is not well explained. This problem was only attempted in Rojas et al. [41], trying to identify the most important area of facial assessment of personality traits, but the results did not have deep guidance. significance. Therefore, whether or not we can find a more comprehensive description of face features and feature extraction methods, and unify them under an effective framework for overall research, is the key to final performance prediction.

7.2 Development Prospects of Automatic Personality Prediction

So far, the interest in computerized automatic personality prediction is still at a relatively early stage (see Fig. 2). Most of the work is devoted to establishing the field, collecting data, developing methods, and identifying related tasks. Research shows that the application of personality calculation in many fields brings convenience to human society, first in daily life [49]: the user's personality score improves the performance of the recommendation system [46, 47]; synthetic speech based on individual preferences is The proof can improve the acceptance of the GPS system [31, 44]; the correct positioning of advertising campaigns on potential users [25]; and retrieval techniques that match the user's personality [14]. Secondly, in terms of medical care: With advances in technology involving autism spectrum problems and other mental problems [42], personality calculations may be based on techniques designed to detect psycho-psychiatric disorders such as paranoia and schizophrenia that often interfere with personality. Plays an important role in [13]; Finally, in terms of social behavior, personality calculations may help establish the link between features and behaviors, and so far it has not been possible to achieve this connection for the prediction of major social events such as elections [23], criminal identification [53] and so on.

Future research is expected to establish a more extensive data set of age structure (35 years old or older) that satisfies the requirements of 2.5D-3D research, and collect scores of personality traits from reliable observers to fill existing limitations. At the

same time, it should also try to introduce some new research models for face character prediction. For example, consider it as a marker distribution learning problem to cope with the problem of lack of training samples; explore the impact of group classification variables on personality scores, and customize the character prediction model of the corresponding group.

References

1. Edmonds, J.M.: The Characters of Theophrastus. W. Heinemann, London (1967)
2. Matthews, G., Deary, I., Whiteman, M.: Personality Traits, 3rd edn, pp. 1–568. Cambridge University Press, Cambridge (2009)
3. Funder, D.C.: Personality. Annu. Rev. Psychol. **52**(1), 197–221 (2001). https://doi.org/10.1146/annurev.psych.52.1.197. pMID: 11148304
4. Ozer, D.J., Benetmartinez, V.: Personality and the prediction of consequential outcomes. Annu. Rev. Psychol. **57**(1), 401–421 (2006)
5. Uleman, J.S., Saribay, S.A., Gonzalez, C.: Spontaneous inferences, implicit impressions, and implicit theories. Annu. Rev. Psychol. **59**(1), 329–360 (2008)
6. Al Moubayed, N., Vazquez-Alvarez, Y., McKay, A., Vinciarelli, A.: Face-based automatic personality perception. In: Proceedings of the 22nd ACM International Conference on Multimedia, pp. 1153–1156. ACM (2014)
7. Ballew, C.C., Todorov, A.: Predicting political elections from rapid and unreflective face judgments. Proc. Natl. Acad. Sci. **104**(46), 17948–17953 (2007)
8. Brahnam, S.: A study of artificial personality from the perspective of the observer, June 2018
9. Brahnam, S., Nanni, L.: Predicting trait impressions of faces using local face recognition techniques. Expert Syst. Appl. **37**(7), 5086–5093 (2010)
10. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
11. Carney, D.R., Colvin, C.R., Hall, J.A.: A thin slice perspective on the accuracy of first impressions. J. Res. Pers. **41**(5), 1054–1072 (2007)
12. Cheng, X.H.: Research on human face detection and recognition. Masters thesis, Xidian University, September 2006
13. Corr, P.J., Matthews, G.: The Cambridge Handbook of Personality Psychology. Cambridge University Press, Cambridge (2009)
14. Cristani, M., Vinciarelli, A., Segalin, C., Perina, A.: Unveiling the multimedia unconscious: implicit cognitive processes and multimedia content analysis. In: Proceedings of the 21st ACM International Conference on Multimedia, pp. 213–222. ACM (2013)
15. Deary, I.J.: The trait approach to personality. In: The Cambridge Handbook of Personality Psychology, pp. 89–109 (2009)
16. Digman, J.M.: The curious history of the five-factor model (1996)
17. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. Wiley, Hoboken (2012)
18. Eysenck, H.J.: Dimensions of personality: 16, 5 or 3? - Criteria for a taxonomic paradigm. Personality Individ. Diff. **12**(8), 773–790 (1991)
19. Freedman, D.A.: Statistical Models: Theory and Practice. Cambridge University Press, Cambridge (2009)
20. Freierman, S.: Constructing a Real-life Mr. Potato Head. Faces: The Ultimate Composite Picture. The New York Times 6 (2000)
21. Huang, J., Li, W.S., Gao, Y.J.: Research advance of facial expression recognition. Comput. Sci. (z2), 123–126 (2016)

22. Joachims, T.: Optimizing search engines using clickthrough data. In: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 133–142. ACM (2002)
23. Joo, J., Steen, F.F., Zhu, S.C.: Automated facial trait judgment and election out-come prediction: social dimensions of face. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3712–3720 (2015)
24. Kakadiaris, I.A., Abdelmunim, H., Yang, W., Theoharis, T.: Profile-based face recognition. In: 8th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2008, pp. 1–8. IEEE (2008)
25. Kosinski, M., Stillwell, D.: myPersonality Research Wiki (2012)
26. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
27. Liu, S.: Research on the prediction of facial beauty from 2D to 3D. Ph.D. thesis, Northwestern Polytechnical University (2017)
28. Liu, S., Fan, Y.Y., Samal, A., Guo, Z.: Advances in computational facial attractiveness methods. Multimedia Tools Appl. **75**(23), 16633–16663 (2016)
29. Liu, Shu, Fan, Yangyu, Guo, Zhe, Samal, Ashok: 2.5D facial attractiveness computation based on data-driven geometric ratios. In: He, Xiaofei, et al. (eds.) IScIDE 2015. LNCS, vol. 9242, pp. 564–573. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-23989-7_57
30. Lundqvist, D., Flykt, A., Ohman, A.: Karolinska directed emotional faces (Department of Neurosciences, Karolinska hospital, Stockholm, Sweden). Karolinska Directed Emotional Faces 2(4) (1998)
31. Nass, C., Brave, S.: Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship. MIT Press, Cambridge (2005)
32. Oosterhof, N.N., Todorov, A.: The functional basis of face evaluation. Proc. Natl. Acad. Sci. **105**(32), 11087–11092 (2008)
33. Pantic, M., Patras, I., Rothkruntz, L.: Facial action recognition in face profile image sequences. In: 2002 IEEE International Conference on Multimedia and Expo, 2002. ICME 2002. Proceedings, vol. 1, pp. 37–40. IEEE (2002)
34. Parikh, D., Grauman, K.: Relative attributes. In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 503–510. IEEE (2011)
35. Parzen, E.: On estimation of a probability density function and mode. Ann. Math. Stat. **33** (3), 1065–1076 (1962)
36. Pei, W.W.: Gender studies based on the feature extraction of face images analysis. Masters thesis, Taiyuan University of Technology, February 2010
37. Qin, R.: Personality analysis based on face image. Master's thesis, Chinese Academy of Sciences University (2016)
38. Qin, R., Gao, W., Xu, H., Hu, Z.: Modern physiognomy: an investigation on predicting personality traits and intelligence from the human face. Sci. China Ser. F Inf. Sci. **61**(5), 058105 (2018)
39. Ren, S., Cao, X., Wei, Y., Sun, J.: Face alignment at 3000 fps via regressing local binary features. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1685–1692 (2014)
40. Rish, I.: An empirical study of the naive Bayes classifier. In: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, vol. 3, pp. 41–46. IBM (2001)
41. Rojas, M., Masip, D., Todorov, A., Vitria, J.: Automatic prediction of facial trait judgments: appearance vs. structural models. PLoS ONE **6**(8), e23323 (2011)

42. Saint-Georges, C., et al.: Do parents recognize autistic deviant behavior long before diagnosis? Taking into account interaction using computational methods. PLoS ONE **6**(7), e22393 (2011)
43. Saucier, G., Goldberg, L.R.: The language of personality: lexical perspectives. In: The Five-Factor Model of Personality: Theoretical Perspectives, pp. 21–50 (1996)
44. Schwab, K., Marcus, A., Oyola, J., Hoffman, W., Luzi, M.: Personal data: the emergence of a new asset class. In: An Initiative of the World Economic Forum (2011)
45. Tariq, U., Hu, Y., Huang, T.S.: Gender and ethnicity identification from silhouetted face profiles. In: 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 2441–2444. IEEE (2009)
46. Tkalcic, M., Tasic, J., Košir, A.: Emotive and personality parameters in multimedia recommender systems. In: Affective Computing and Intelligent Interaction, ACII 2009, vol. 33 (2009)
47. Tkalcic, M., Tasic, J., Košir, A.: The LDOS-PerAff-1 corpus of face video clips with affective and personality metadata. In: Proceedings of the Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality (Malta 2010), LREC, p. 111 (2010)
48. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cogn. Neuro- Sci. **3**(1), 71–86 (1991)
49. Vinciarelli, A., Pantic, M., Bourlard, H.: Social signal processing: survey of an emerging domain. Image Vis. Comput. **27**(12), 1743–1759 (2009)
50. Wang, Z.L., Chen, F.J., Xue, W.M.: A survey of facial expression recognition. Comput. Appl. Softw. (12), 63–66 (2003)
51. White, R., Eden, A., Maire, M.: Automatic prediction of human attractiveness. UC Berkeley CS280A Project 1, 2 (2004)
52. Wolffhechel, K., et al.: Interpretation of appearance: the effect of facial features on first impressions and personality. PLoS ONE **9**(9), e107721 (2014)
53. Wu, X., Zhang, X.: Automated inference on criminality using face images. arXiv preprint arXiv:1611.04135 (2016)
54. Zeng, Z.: An analysis of students' personality traits based on face features and deep learning. Masters thesis, Jiangxi Normal University, March 2017