OPO-FCM: A Computational Affection Based OCC-PAD-OCEAN Federation Cognitive Modeling Approach

Feng Liu, *Member, IEEE*, Han-Yang Wang, Si-Yuan Shen, Xun Jia, Jing-Yi Hu, Jia-Hao Zhang, Xi-Yi Wang, Ying Lei, Ai-Min Zhou, *Senior Member, IEEE*, Jia-Yin Qi, and Zhi-Bin Li

***Abstract*— In recent years, it is a difficult issue to integrate the deep cross-fertilization and interpretable cognitive modeling methods from the basic theory of emotional psychology with deep learning and other algorithms. To address this problem, a cognitive model that integrates the VGG-facial action coding system (FACS)-OCC model based on fer2013 expression features and the OCC-pleasure-arousal-dominance (PAD)-openness, conscientiousness, extraversion, agreeableness, and neuroticism (OCEAN) fusion of the basic theory of emotional psychology, namely, a computational affection-based OCC-PAD-OCEAN federation cognitive modeling (OPO-FCM), is constructed. By constructing this model and performing formal proof algorithms, it is shown that the OPO-FCM can acquire expression features in video streams, complete the acquisition of expression features in videos by training a deep neural network, map expressions to the PAD emotion space through the established expression–basic emotions–emotion space mapping relationship, and finally complete the mapping of the average emotion over a period time. The information of personality space is obtained through it. Finally, the experimental simulation of the model is conducted, and the results show that the average accuracy of the valid tested personalities is 79.56%. This article takes the knowledge-driven approach of emotional psychology as a starting point and combines deep learning techniques to construct interpretable cognitive models, thus providing new ideas for future cross-innovation between computer technology and psychology theory.**

***Index Terms*— Cognitive modeling, computational affection, emotional psychology, five-factor model (FFM), OCC-pleasure-arousal-dominance (PAD)-openness, conscientiousness, extraversion, agreeableness, and neuroticism (OCEAN), VGG-facial action coding system (FACS)-OCC.**

I. Introduction

As AN experimental science, the vast majority of psychological research advances are based on the psychological experimental paradigm, where subjective and objective data from subjects are collected and analyzed. The experiments require subjects to observe and record all aspects of their behavior with the strict control of variables and analyze the results of the collected data [1]. Even so, psychological experiments still suffer from a large number of problems such as poor reliability and validity of experimental results and nonreproducibility of experiments. These problems are partly due to the bystander effect and the limitations of laboratory experiments, which in turn lead to psychological experiments being confined to certain scenarios and cannot be extrapolated and extended. To address these problems, computer technology can play a role in precisely controlling and quantifying data, and this modern psychometric technique can have many benefits, including avoidance of complex reliability measures, improved construct validity, avoidance of exposure effects, and increased measurement efficiency [2]. The main purpose of most psychological experiments is to investigate the principles of human behavior or human cognitive patterns: based on the perspective of data acquisition, computer technology for the acquisition of observed data enables precise digital control of the entire experimental environment, such as accurate acquisition of video signals, audio signals, sensor data, and human motion information. Based on the perspective of constructing experimental environments, computer assistance can provide subjects with an immersive experience. For example, the use of virtual reality (VR) technology in the study of emotional psychology can enable emotions to be induced more effectively compared to general pictures or verbal stimuli. In addition, computer technology can simulate hypothetical models to explain observed behaviors, and computer simulations can also provide preliminary idealized validation

Manuscript received 18 October 2021; revised 16 December 2021 and 30 March 2022; accepted 11 August 2022. Date of publication 26 August 2022; date of current version 2 August 2023. This work was supported in part by the Science and Technology Commission of Shanghai Municipality under Grant 19511120601, in part by the Scientific and Technological Innovation 2030 Major Projects under Grant 2018AAA0100902, in part by the Research Project of Shanghai Science and Technology Commission under Grant 20dz2260300, and in part by the Fundamental Research Funds for the Central Universities. *(Corresponding authors: Ai-Min Zhou; Jia-Yin Qi.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University Committee on Human Research Protection under Application No. HR166-2021, and performed in line with the Declaration of Helsinki.

Feng Liu, Han-Yang Wang, Si-Yuan Shen, Jia-Hao Zhang, Ying Lei, Ai-Min Zhou, and Zhi-Bin Li are with the Institute of AI for Education, the School of Computer Science and Technology, the Shanghai Key Laboratory of Mental Health and Psychological Crisis Intervention, and the School of Psychology and Cognitive Science, East China Normal University, Shanghai 200062, China (e-mail: lsttoy@163.com; faceeyes@126.com; trert9811@gmail.com; zjh20000218@163.com; ly_20001007@163.com; amzhou@cs.ecnu.edu.cn; lizb@admin.ecnu.edu.cn).

Xun Jia and Xi-Yi Wang are with the School of Psychology and Cognitive Science, East China Normal University, Shanghai 200062, China (e-mail: jiaxundavid@163.com; sabri@sbsub.com).

Jing-Yi Hu is with the School of Business Administration, East China Normal University, Shanghai 200062, China (e-mail: jessy_7887@126.com).

Jia-Yin Qi is with the Institute of Artificial Intelligence and Change Management, Shanghai University of International Business and Economics, Shanghai 200336, China (e-mail: qijiayin@139.com).

This article has supplementary downloadable material available at https://doi.org/10.1109/TCSS.2022.3199119, provided by the authors.

Digital Object Identifier 10.1109/TCSS.2022.3199119

of hypotheses when experimental conditions are constrained. Machine learning approaches facilitate the understanding of numerous physiological processes underlying complex human mental states and behavior, leading to a new research direction named computational psychophysiology [3]. The study of human affective behavior has received increasing attention in the last decade [4]. Affective computing is precisely an interdisciplinary discipline based on psychology and computer science that studies how to use computers to recognize, model, and even express human emotions [5]. Then, the personality computing extended from affective computing can advance all the techniques related to the understanding and prediction of human behavior [6], [7], [8].

In psychological research on human behavior, prediction, and other directions, personality is a very important determinant that describes stable personal characteristics that can often be measured in a quantitative way to explain and predict observable behavioral differences [9]. The five-factor model (FFM) [10] is an important theory in personality psychology today and is one of the most influential models in psychological research [11]. Its five factors include openness, conscientiousness, extraversion, agreeableness, and neuroticism (OCEAN). Social psychologist Harry Reis described the FFM as “the most scientifically rigorous taxonomy in the behavioral sciences.” The FFM provides a structure for classifying personality traits in most people, with a set of highly replicable dimensions that provide a parsimonious and comprehensive description of most individual differences [12]. From a computer perspective, trait models represent personality in numerical form and can be suitable for computer processing. In contrast, most current personality assessments use a self-report format, in which personality is assessed by statements or adjectives in a scale. The self-report paradigm is simple and straightforward but does not control for the veracity of the subjects’ responses, and the experimental results are susceptible to large biases due to the influence of multiple extraneous factors. One of the major limitations of self-assessment is also that subjects may tend to bias ratings toward social expectations, especially when the assessment may have negative consequences and may hide negative traits, resulting in results that do not match true personality.

Overall, although some cross-innovation research is currently advancing the computation and quantification of psychological theories, psychological theories are still dominated by traditional qualitative findings, which make it difficult to provide direct quantitative model support for computer-based algorithmic implementations. In addition, algorithmic programs for computers cannot accurately represent emotion theories and emotion models in psychology, and a large barrier exists between them. Most of the existing studies usually consider only one of the perspectives from computer science or psychology rather than from a cross-fertilization perspective [13]. Meanwhile, although the current deep learning-based facial expression recognition techniques are relatively well-established [14], research on using deep learning techniques to process psychological signals is still in its infancy [15]. Thus, cognitive modeling methods that integrate deep learning and other algorithms for deep fusion from the basic theory of emotional psychology are still lacking, and how to improve the interpretability of the model while efficiently dealing with cognitive problems is also a key issue.

In this article, based on the existing research, we combine the emotion modeling approach of emotional psychology with computerized deep learning algorithms, use the deep learning framework VGG [16] to construct a facial action coding system (FACS)-OCC emotion modeling approach based on fer2013 expression features to obtain emotional features, and then combine the OCC-pleasure-arousal-dominance (PAD)-OCEAN federal cognitive modeling approach based on emotions to obtain the five-factor information of subjects. Specifically, the OCC-PAD-OCEAN federation cognitive modeling (OPO-FCM) cognitive model is shown in Fig. 1, which completes the acquisition of expression features in the video by training a deep neural network, maps expressions to the PAD emotion space in the established expression–basic emotions–emotion space mapping, and finally achieves the extraction of personality features through the established temperament-personality mapping. The OPO-FCM cognitive model achieves the extraction of personality features by mapping the average expressions and temperament over a period of time to finally obtain information with certain statistically significant confidence validity on the personality space.

This study is an attempt to build a new cognitive modeling research paradigm by cross-fertilization innovation between computer and psychology basic theories, which provides new ideas for future cross-innovation between computer technology and psychology theories.

II. Related Works

Previous studies represented by cutting-edge research journals like *Nature* and *Science*, involving cognition and emotion, are the following: combining neurological and computational perspectives on social learning with an understanding of behavior of varying complexity [17]; discovering the micro-dynamics of networked emotions and revealing the role of emotion labeling [18]; proposing the existence of commonalities between face recognition and expression recognition and merging the two for recognition [19]; analyzing the potential profile of the emotional neuroscience personality inventory [20]; finding that complex emotional dynamics provide limited information to the prediction of mental health [21]; and investigating the neurocognitive dynamics of near-threshold speech signal detection and emotional speech evaluation [22]. And emotional computing [23] aims to make computers more intelligent by giving them the ability to recognize, understand, and express human emotions. Gebhard [24] proposed a layered emotion model called a layered model of affect (ALMA) based on the emotion engine [25], [26] that integrates emotions, temperament, and personality, applying psychological results from human behavior analysis to the concretization of virtual characters. To model human emotions, scholars have proposed methods in various forms, which can be generally classified into three categories: categorical methods, dimensional methods, and appraisal-based methods. The categorical approach assumes that there exists a small number of basic emotions [27] entrenched in the brain, and more complex

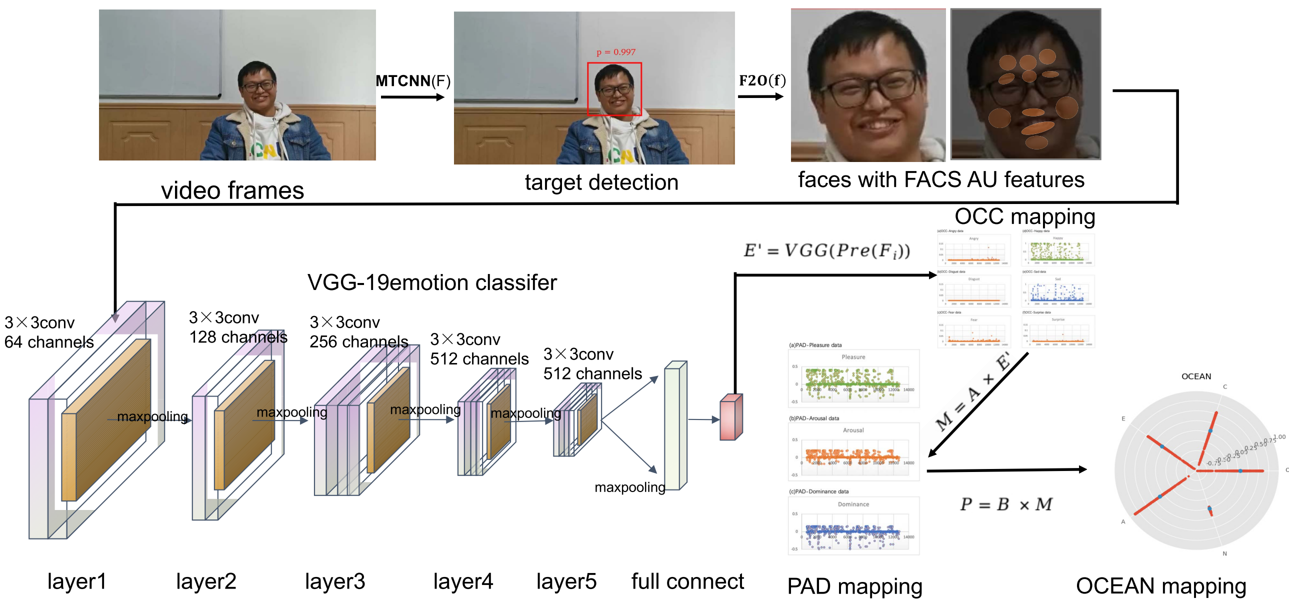

Fig. 1. Schematic of OPO-FCM cognitive model processing.

emotions can be combined based on these basic emotions. However, individual labels (or any small number of discrete categories) can hardly reflect the complexity of emotional states, and most researchers then began to advocate the use of dimensions to describe human emotions. In dimension-based modeling approaches, emotional states are not independent of each other but are interconnected in a systematic way. For example, Russell [28] proposed a circular configuration called circumplex of affect that uses two dimensions (arousal and valence) to describe emotions. Another commonly used theory is the PAD model of emotion [29], which includes three dimensions: pleasure–displeasure, arousal–nonarousal, and dominance–submissiveness. In contrast, in cognitive appraisal-based approaches, scholars have argued that emotions are generated through a continuous, recursive, subjective assessment of one's internal state and the state of the external world [30], and Ortony *et al.* [31] proposed an easily computable cognitive base model of emotion arousal, OCC, which allows the construction of an emotion model for a concretized character in a concretized conversational subject design [32].

Emotion recognition is an indispensable part of emotion computing [33], where audio signals can convey emotional information through explicit linguistic information and implicit acoustic information. There has been a large amount of work investigating how to map audio representations to dimensional models for emotion recognition through audio. Cowie *et al.* [34], [35] used valence-activation spaces to evaluate and model emotions in language; Wöllmer *et al.* [36] used acoustic low-level descriptors as acoustic features for emotion recognition based on the SAL database [37], [38], the regression using long short-term memory (LSTM) recurrent neural network (RNN) [39], and support vector machines; Yiling *et al.* [40] extracted temporal ignition sequences, ignition location information features, and mel-frequency cepstral coefficient (MFCC) 3-D emotion features for emotion recognition by audio, and mapped the results to the PAD 3-D emotion space.

Human emotions can be reflected not only in sound but also in facial movements, expressions, body postures, and behaviors. Therefore, several researchers have investigated how visual signals from various videos can be mapped to emotional dimensions. Russell [28] mapped facial expressions to different locations on the 2-D arousal valence plane, while Cowie *et al.* [41] examined the emotional and communicative significance of head nodding and head shaking in terms of arousal and valence dimensions as well as solidarity, confrontation, and agreement. Kipp and Martin [42] investigated the relationship between features of basic gesture shape (e.g., left- and right-hand preference, hand shape, and palm orientation) and the 3-D space of the PAD. Nicolaou *et al.* [43] used an association-related vector machine (OA-RVM) regression framework to map facial expressions, shoulder movements, and audio into a 2-D valence arousal space. In addition to visual and auditory, emotions can also be computed from physiological data such as ECG [44], EEG [45], [46], [47], or textual data [48]. Meanwhile, considering the relationship between personality and emotion, Fengfang *et al.* [49] proposed a personality-OCC model for emotion computation. Considering that one emotion can induce different concurrent emotions [50], Yang *et al.* [51] proposed a hybrid emotion computation model based on personality and event-driven personality association emotion model (PAEM). Moreover, some approaches tend to compute one's personality directly using a number of features from their online behavior [52].

In summary, the above related works have effectively studied the computer recognition and processing of human emotions in terms of audio, video, gesture, and behavior, respectively, but cognitive modeling approaches that integrate deep learning and other algorithms from psychological theories for deep integration remain lacking, and most cognitive modeling algorithms need to be improved in terms of interpretability. Therefore, in this study, starting from the basic theory of emotional psychology, video data of subjects in psychology experiments are taken as the original input, and

TABLE I
NOTATIONS

Notation	Meaning
I	The pixel matrix of an image
F	A frame in a video
t	Time
T	The time of the whole video
E	The emotion
M	The temperament
P	The personality
f	One layer of the neural network
h_t, w_t, p_t	Height threshold, width threshold, and confidence threshold
h, w, p	Height, width, confidence
A	The transformation matrix from basic emotions to emotion space
B	The conversion matrix of temperament and personality

a FACS-OCC emotion modeling method based on fer2013 expression features is constructed using the deep learning framework VGG in order to obtain emotional features. Second, a new cognitive modeling processing paradigm was constructed with the emotion-based OPO-FCM to obtain the five-factor information of the subjects.

III. Computational Affection-Based OCC-PAD-OCEAN Federal Cognitive Modeling Approach

This section describes the theoretical derivation process from images to personality, including notations in Table I, the deep learning models used, the modeling of related cognitive psychology, and formal reasoning and validation.

A. VGG-FACS-OCC Emotion Modeling Approach Based on fer2013 Expression Features

VGG is a kind of classical deep convolutional neural networks with powerful image feature extraction capability. In this article, VGG-19 trained on the fer2013 dataset is used as a video-emotion space inference model to derive the emotion space representation of subjects from their videos of natural conversation with VGG-19. The mapping can be considered as the arithmetic average of the mapping of each image to emotion space vectors in the video. An arbitrary natural talk video V can be represented as a sequence of frames $(F_1, F_2, \ldots, F_n)$, and then, VGG can transform any frame F into an OCC vector E. Let the frame preprocessing function be $\text{Pre}(F)$, then we have $E = \text{VGG}(\text{Pre}(F))$. Therefore, the process of deriving OCC vectors by VGG can be expressed as

$$E_V = \frac{1}{n}\sum_{i=1}^{n} \text{VGG}(\text{Pre}(F_i)). \tag{1}$$

The process of inferring the overall OCC vector by video has been given above, and the following discussion focuses on how to transform a single frame image into an OCC vector by VGG.

The model parameters of VGG are generally determined by the pretraining process, and its forward propagation inference process can be formalized as

$$\text{VGG}(I) = f_1{}^\circ f_2{}^\circ \ldots{}^\circ f_n(I) \tag{2}$$

where I denotes the pixel matrix of an image, $f_1, f_2, \ldots, f_n$ denote the functions corresponding to different levels of the neural network, and $^\circ$ denotes the function composite, meaning that the forward propagation of the model is completed by the function composite of different levels of the neural network. The implementation of VGG is VGG-19.

For the video frame F in real scenes, it generally cannot be used as the input of the VGG model directly. This is because the video frame in real scenes may not be able to recognize the face effectively and may possess too much interference information, so the preprocessed frame $\text{Pre}(F)$ is generally used as the input. The preprocessing function Pre() can be described by the following algorithmic procedure.

1) For any frame F in the video, the face target detection is performed using the multi-task cascaded convolutional neural network (MTCNN) face recognition algorithm [53], deriving the set of target frames $B = b_1, b_2, \ldots, b_m$. All these target frames can be represented in the form of an ordered quintet $b_i = (x_i, y_i, h_i, w_i, p_i)$, and these five elements represent the upper-left horizontal coordinate, upper-left vertical coordinate, height, width, and confidence (the probability that the algorithm considers this target as a face) of the target frame in turn.
2) Determine the hyperparameters $h_\text{t}, w_\text{t}, p_\text{t}$, which are height threshold, width threshold, and confidence threshold. For any target box $b_i \in B$, only the target boxes with height, width, and confidence exceeding the threshold are retained to prevent nonface targets from interfering with emotional reasoning. Finally, the set of filtered target frames B' is obtained

$$B' = \{h > h_\text{t}, w > w_\text{t}, p > p_\text{t} \mid (x, y, h, w, p) \in B\}. \tag{3}$$

3) From the filtered set of target frames B', the target frame b^* with the highest confidence p is obtained, F is cropped according to the target frame, and the size of F is resized to a specific size after cropping and finally used as the input to the VGG emotional inference model.

In order to adjust the parameters of VGG-19 so that it has good FACS feature extraction ability and emotion classification ability, the model is trained with the fer2013 dataset in this article. The fer2013 dataset provides OCC emotion annotation of face images based on FACS. Its FACS-OCC transformation is shown in Table II. Let the FACS-action unit (AU) intensity vector be f, each dimension of which represents the intensity of FACS and AU, which are associated with emotion recognition and take values in the range [0,1].

The above FACS-OCC transformation method is denoted as the function $\text{F2O}(f)$, then for the image I and FACS-AU

TABLE II
FACS-OCC SENTIMENT MAPPING TABLE BASED ON FER2013 EXPRESSION FEATURES

OCC emotion classification	OCC vector	FACS AU characteristics
anger	$[1,0,0,0,0,0]^T$	AU23, 24 exist simultaneously
disgust	$[0,1,0,0,0,0]^T$	AU9 or AU10 exist
fear	$[0,0,1,0,0,0]^T$	AU1, 2, 4 exist simultaneously, AU4 may not exist when AU5 is at E intensity level
happiness	$[0,0,0,1,0,0]^T$	\
sadness	$[0,0,0,0,1,0]^T$	AU1, 4, 15 exist simultaneously, or AU11 exists, or AU6, 15 exist at the same time
surprise	$[0,0,0,0,0,1]^T$	AU1, 2 exist at the same time, or AU5 whose intensity does not exceed B exists

feature label f in the training set, the optimization objective of the VGG model is

$$L(\mathrm{VGG}(I), \mathrm{F2O}(f)) = \mathrm{CrossEntropy}(\mathrm{VGG}(I), \mathrm{F2O}(f)) \quad (4)$$

where the cross-entropy loss is given by

$$\mathrm{CrossEntropy}(p, q) = -\sum_{i=1}^{n} p(x_i)\log(q(x_i)) \quad (5)$$

where n is the number of labels, and for the OCC emotion classification problem in this article, the labels are six emotion dimensions. The VGG model is trained with fer2013 dataset by batch gradient descent to minimize the objective function L. With the adjustment of the model parameters, the hidden layer of VGG can then perform FACS feature extraction and finally obtain the specific OCC sentiment vector.

B. Emotion-Based OCC-PAD-OCEAN Federal Cognitive Modeling Approach

The previous section discussed that the input pictures of facial expression can be transformed into 6-D continuous vectors of their corresponding basic emotions by the VGG model, and this section focuses on psychological modeling from basic emotions mapping to personality.

First, the meanings of basic emotion classification, space of emotion and temperament, space of personality, and their vector representations will be introduced, respectively.

1) Cognitive Evaluation Model of Emotion (OCC): Emotion reflects the short-term state of a person, which changes significantly in a short period of time due to receiving changes or stimuli from external environment. The OCC emotion model, as a standard model of emotion synthesis, specifies 22 emotion categories from the perspective of classification of emotion. Based on Ekman's [27] theory that all nonbasic emotions can be synthesized from basic emotions, the emotion space is constructed using the six basic emotions defined by him, namely, anger, disgust, fear, happiness, sadness, and surprise, which are represented in vector form as follows, where each element takes values in the range of [0,1], indicating the intensity of emotions:

$$E = \left[e_{\mathrm{angry}}, e_{\mathrm{disgust}}, e_{\mathrm{fear}}, e_{\mathrm{happy}}, e_{\mathrm{sad}}, e_{\mathrm{surprise}}\right]^{\mathrm{T}}. \quad (6)$$

2) 3-D Space of Temperament (PAD): Temperament is a typical and stable dynamic characteristic of a person's mental activity. According to Mehrabian's [29] theory, emotional trait refers to the average of an individual's emotional state in various life situations and can be used to describe temperament. However, since combinations of discrete emotional states (e.g., anger, disgust, fear, happiness, and sadness) cannot be meaningfully averaged, a conceptual system is needed to construct the basic dimensions of emotions. Thus, Mehrabian [54] introduced the PAD space of emotion consisting of three mutually independent dimensions, namely, pleasure, arousal, and dominance, from which the PAD space of temperament derives. The latter can be represented in the vector form as follows, with elements taking values in the range of [−1,1]:

$$M = [m_{\mathrm{P}}, m_{\mathrm{A}}, m_{\mathrm{D}}]^{\mathrm{T}}. \quad (7)$$

Pleasure describes the general relative strength of positive versus negative emotional states. Arousal measures the degree to which a person is aroused by "high information rate" (complex, rapid-changing, unexpected) stimuli and the speed of recovery to baseline levels. Dominance assesses a person's sense of control and influence over the environment in which he or she lives, or instead, the feeling of being controlled and influenced by others or events.

3) Space of Personality (OCEAN): Personality reflects stable differences in psychological characteristics among individuals, none of which change significantly over time. The FFM [10] is used to construct the personality space with five factors: OCEAN, which are expressed in vector form as follows, with elements takes values in the range of [−1,1]:

$$P = [p_{\mathrm{O}}, p_{\mathrm{C}}, p_{\mathrm{E}}, p_{\mathrm{A}}, p_{\mathrm{N}}]^{\mathrm{T}}. \quad (8)$$

According to the general interpretation of the five-factor [10], it is known that: openness describes one's cognitive style, seeking understanding of new experience, and tolerance and exploration of unfamiliar situations. Responsibility refers to the way in which one controls, manages, and regulates one's impulses and assesses the individual's organization, persistence, and motivation in goal-directed behavior. Extraversion indicates the amount and density of interpersonal interactions, the need for stimulation, and the ability to obtain pleasure. Agreeableness examines the attitudes held by the individual toward other people. Neuroticism (opposite to emotional stability) reflects the individual's emotion regulation process, the tendency to experience negative emotions, and emotional instability.

TABLE III
3-D EMOTION PAD QUANTITATIVE MAPPING TABLE OF SPACE AND EMOTION

Emotion type	Pleasure	Arousal	Dominance
anger	-0.51	0.59	0.25
disgust	-0.40	0.20	0.10
fear	-0.64	0.60	-0.43
happy	0.40	0.20	0.15
sad	-0.40	-0.20	-0.50
surprise	0.20	0.45	-0.45

4) Mapping of Basic Emotions to Temperament Space: The mapping between basic emotions and PAD emotion space proposed by Ekman is shown specifically in Table III. Table III was developed and used to model hierarchical emotions by Gebhard [24], which is widely used. Table III defines the baseline coordinates of the six basic emotions in the PAD emotion space, where surprise has no corresponding PAD value in the original PAD scale. By observing the PAD values corresponding to emotions similar to surprise in the original scale, its PAD values are assumed to be 0.20, 0.45, and −0.45 after careful analysis.

In Table III, when the value of one emotion is the maximum value of 1 and the rest of the emotion values are 0, it can be mapped to the PAD value corresponding to the right side. Thus, the above scale is formalized and written as the following mapping relationship:

$$f_{\mathrm{e}}(e) = [m_{\mathrm{Pe}}, m_{\mathrm{Ae}}, m_{\mathrm{De}}]^{\mathrm{T}} \quad e \in \{\text{anger, disgust, fear, happy, sad, surprise}\}. \tag{9}$$

Further, it is written as a computer-friendly matrix multiplication form $M_{\mathrm{t}} = A \times E$, where

$$A = \begin{bmatrix} -0.51 & -0.40 & -0.64 & 0.40 & -0.40 & 0.20 \\ 0.59 & 0.20 & 0.60 & 0.20 & -0.20 & 0.45 \\ 0.25 & 0.10 & -0.43 & 0.15 & -0.50 & -0.45 \end{bmatrix} \tag{10}$$

is the transformation matrix from basic emotions to emotion space, and E is a binary vector. With the facial expression recognition technique, the emotion vector can be obtained as follows:

$$E = \left[e_{\text{angry}}, e_{\text{disgust}}, e_{\text{fear}}, e_{\text{happy}}, e_{\text{sad}}, e_{\text{surprise}}\right]^{\mathrm{T}}. \tag{11}$$

However, since only discrete correspondence is given in Table III, it needs to be transformed into a continuous mapping function to obtain the emotion space coordinate vector M corresponding to the emotion vector E. Thus, the formula is extended to obtain: $M_{\mathrm{t}} = A \times E'$, when E' is the emotion vector obtained by computer recognition. Averaging M_{t} over a longer time scale, we get $M = (\sum M_{\mathrm{t}}/T)$, which is the temperament vector M.

The mapping of basic emotion to temperament space based on this method simultaneously considers the intensity of continuous emotion vectors on six emotion categories, and the mapping results are obtained based on widely used scales with high credibility. By introducing the intermediate concept of emotion space between basic emotions and temperament, it reflects the mechanism how emotional states and temperament interact and why this kind of mapping works with high interpretability.

5) Mapping of Temperament Space to Personality Space: Mehrabian and O'Reilly [55] established the following transformation relationship from temperament space to personality space by linear regression:

$$\begin{aligned} \text{Sophistication} &= .16P + .24A + .46D \\ \text{Conscientiousness} &= .25P + .19D \\ \text{Extraversion} &= .29P + .59D \\ \text{Agreeableness} &= .74P + .13A - .18D \\ \text{Emotional stability} &= .43P - .49A. \end{aligned}$$

Here, sophistication is more reasonably described with openness and emotional stability is antonymous with neuroticism [56]. Namely,

$$\begin{aligned} \text{Sophistication} &= \text{Openness} \\ \text{Emotional stability} &= -\text{Neuroticism}. \end{aligned}$$

Thus, the transformation relation from the temperament space PAD to the personality space OCEAN is obtained

$$\begin{aligned} \text{Openness} &= .16P + .24A + .46D \\ \text{Conscientiousness} &= .25P + +.19D \\ \text{Extraversion} &= .29P + .59D \\ \text{Agreeableness} &= .74P + .13A - .18D \\ \text{Neuroticism} &= -.43P + .49A. \end{aligned}$$

Namely,

$$P = B \times M \tag{12}$$

where

$$B = \begin{bmatrix} 0.16 & 0.24 & 0.46 \\ 0.25 & 0.00 & 0.19 \\ 0.29 & 0.00 & 0.59 \\ 0.74 & 0.13 & -0.18 \\ -0.43 & 0.00 & 0.49 \end{bmatrix}$$

which is the conversion matrix of temperament and personality.

Based on the above analysis, this article proposes a cognitive modeling to identify five-factor by continuous facial expression data mediated by PAD emotion space and temperament space.

First, the videos from the whole conversation of the subjects in the experiment are sliced by time and sampled according to a fixed frequency. These sampled frames are fed into the trained deep neural network to obtain the corresponding expression data E

$$E_{\mathrm{t}} = \mathrm{VGG}(\mathrm{Image}_{\mathrm{t}}). \tag{13}$$

Then, the expression data are mapped to the emotion space to get the relationship $M_{\mathrm{t}} = A \times E_{\mathrm{t}}$ of emotion data for each

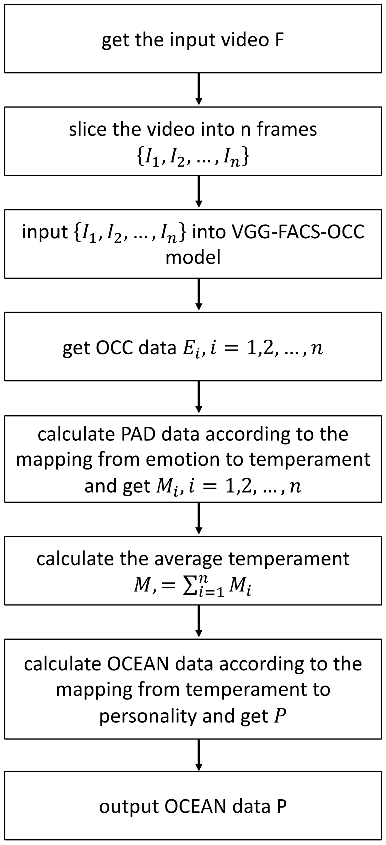

Fig. 2. Flow chart of data processing of the OPO-FCM cognitive model.

frame. Finally, the average temperament data in the whole video are found

$$M = \frac{\sum M_{\mathrm{t}}}{T}. \tag{14}$$

At this point, the temperament data can be mapped to the personality space to get the final personality information of the subject P. The whole brief process is shown in Fig. 2.

IV. Experimental Validation of the OPO-FCM Cognitive Modeling

To confirm the feasibility of the proposed cognitive modeling and its computer implementation, an experiment was conducted in this study with 31 college students (16 males and 15 females), who have no mental or physical disorders and with basic reading ability. We got an effective sample of 28 (14 males and 14 females) after selection. The experimental procedure was as follows: the participants were first interviewed for 5–10 min, and the content of the interview was based on stimulating memories and emotions in daily life, while the expressions of the participants were recorded with a video camera during the experiment. After the interview, participants filled out the big five personality scale [NEO Five-Factor Inventory including Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness (NEO-FFI)] revised by Morrison, 1996. The privacy of participants was well-protected with the confirmation of the Human Subject Protection Committee of ECNU (Ratification NO: HR166-2021). Finally, the validity of the model was empirically demonstrated by analyzing the results of the experiment and the results of the traditional big five personality inventory separately. The relevant experimental material, raw data, and analysis results are available in GitHub for open source access. Meanwhile, regarding the algorithm and result analysis processing of the video, the hardware environment for running is as follows: memory: 8GB; CPU: Intel[1] Core i5-7300HQ 2.50-GHz 4-core; and system running environment Debian 10.

The ethical approval number for this experiment is: HR 166-2021. The processing flow of the video data (12 782 frames and 1692 frames sampled for calculation) with the subject ID "26" is shown in Fig. 3.

The main data processing process of OPO-FCM cognitive modeling is as follows. In the big five personality test, the subjects were first interviewed in order to obtain their video stream information. After that, the video stream data were processed into emotional OCC feature data using VGG-19 based on FACS-OCC emotion modeling, and then processed by OCC-OCEAN cognitive modeling to get the subjects' time-series personality data. Finally, the weighted personality data can be obtained.

The specific processing process is as follows: first, the six expression dimensions of OCC temporal data are obtained from the subjects' video stream information by FACS-OCC emotion modeling based on fer2013 expression features. The 6-D OCC emotion temporal data after the process are shown in Fig. 4.

The subjects' OCC emotion activation during the experiment fluctuated between happy and sad more frequently and more intensely, while the activation of angry, disgust, fear, and surprise were relatively limited.

According to the above output OCC 6-D emotion timing data, the timing data of PAD are obtained through the mapping process from emotion space to temperament space as shown in Fig. 5.

According to Fig. 6, it can be seen that the traditional big five personality data represented by the blue dots fall within the algorithm's credible recognition region; meanwhile, in order to quantify the deviation of each personality, the accuracy of OPO-FCM cognitive modeling is evaluated by calculating the deviation rate of the five personalities for all subjects.

The specific deviation rate calculation method is illustrated by the example of openness

$$\mathrm{Deviation}_{\mathrm{rate}_{\mathrm{Openness}}} = \frac{\left|\frac{1}{n}\left(\sum_{t=1}^{n} \mathrm{Openness}_{\mathrm{r}}\mathrm{ec}_t\right) - \mathrm{Openness}\right|}{\mathrm{Openness}_{\mathrm{r}}\mathrm{ec}_{\max} - \mathrm{Openness}_{\mathrm{r}}\mathrm{ec}_{\min}} \tag{15}$$

where $\mathrm{Openness}_{\mathrm{r}}\mathrm{ec}_t$ denotes the openness value obtained by the algorithm for frame t, n denotes the total number of frames, Openness denotes the openness value calculated by the personality scale, $\mathrm{Openness}_{\mathrm{r}}\mathrm{ec}_{\max}$ denotes the maximum openness value calculated by the algorithm in all frames, and $\mathrm{Openness}_{\mathrm{r}}\mathrm{ec}_{\min}$ denotes the minimum openness

[1]Trademarked.

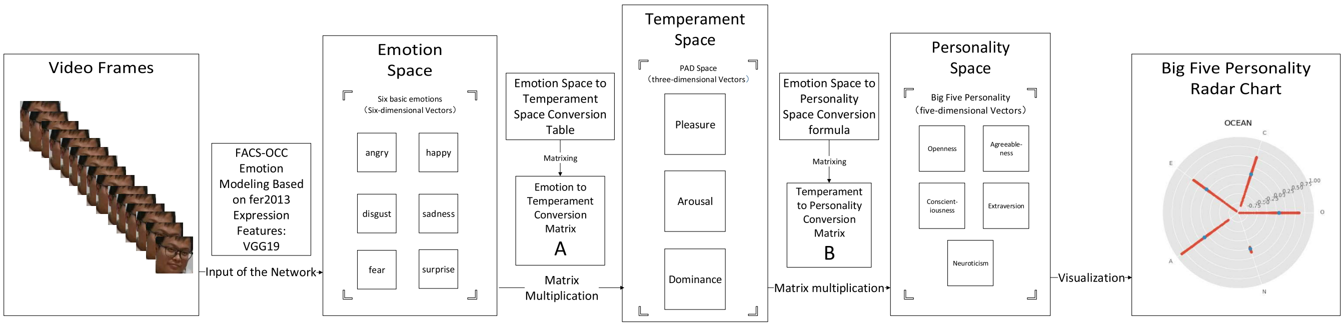

Fig. 3. Schematic representation of the data processing process of the OPO-FCM cognitive model.

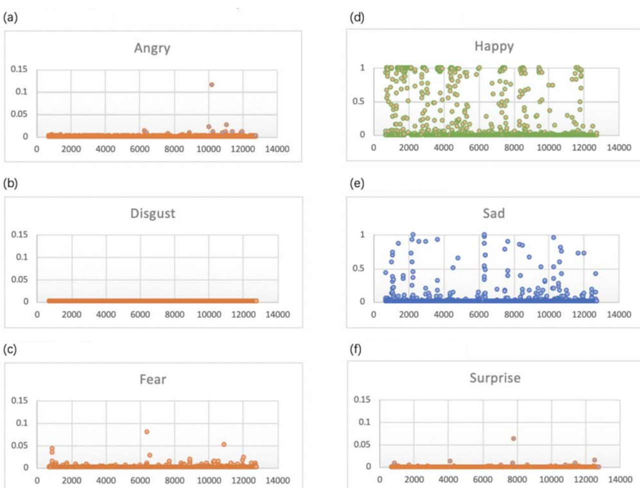

Fig. 4. Timing data plot of the six emotional dimensions of the OCC in all frames of the video: (a) OCC-Angry data; (b) OCC-Disgust data; (c) OCC-Fear data; (d) OCC-Happy data; (e) OCC-Sad data; and (f) OCC-Surprise data.

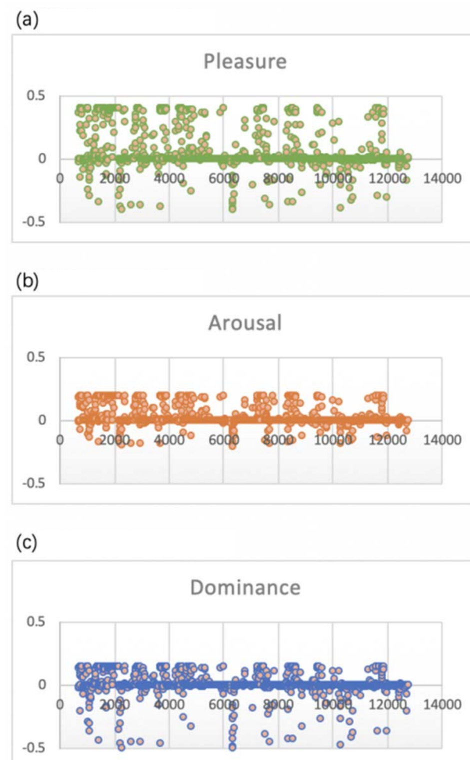

Fig. 5. Timing data diagram of the PAD emotion space in all frames of the video: (a) PAD-Pleasure data; (b) PAD-Arousal data; and (c) PAD-Dominance data.

TABLE IV
COMPARISONS WITH OTHER METHODS

	Time cost	Generalization	Interpretability	Reproducibility
Scales	High	Good	High	Not good
Pure deep learning methods [57-62]	Low	Not good	Low	Good
Ours	**Low**	**Good**	**High**	**Good**

value calculated by the algorithm in all frames. Additionally, to avoid negative numbers, the numerator is taken as absolute value.

Equation (22) represents the difference between the arithmetic average of openness values obtained by the algorithm and openness values obtained by the personality scale, and the difference between the minimum and maximum openness values after taking the absolute value, that is, the openness deviation rate. Similarly, the deviation rates of conscientiousness, extraversion, agreeableness, and neuroticism can be obtained as follows:

$$\text{Deviation}_{\text{rate}_{\text{Conscientiousness}}} = \frac{\left|\frac{1}{n}\left(\sum_{t=1}^{n}\text{Conscientiousness}_{\text{r}}\text{ec}_{t}\right) - \text{Conscientiousness}\right|}{\text{Conscientiousness}_{\text{r}}\text{ec}_{\max} - \text{Conscientiousness}_{\text{r}}\text{ec}_{\min}} \tag{16}$$

$$\text{Deviation}_{\text{rate}_{\text{Extraversion}}} = \frac{\left|\frac{1}{n}\left(\sum_{t=1}^{n}\text{Extraversion}_{\text{r}}\text{ec}_{t}\right) - \text{Extraversion}\right|}{\text{Extraversion}_{\text{r}}\text{ec}_{\max} - \text{Extraversion}_{\text{r}}\text{ec}_{\min}} \tag{17}$$

$$\text{Deviation}_{\text{rate}_{\text{Agreeableness}}} = \frac{\left|\frac{1}{n}\left(\sum_{t=1}^{n}\text{Agreeableness}_{\text{r}}\text{ec}_{t}\right) - \text{Agreeableness}\right|}{\text{Agreeableness}_{\text{r}}\text{ec}_{\max} - \text{Agreeableness}_{\text{r}}\text{ec}_{\min}} \tag{18}$$

TABLE V
FIVE PERSONALITY BIAS RATES WITH DIFFERENT BACKBONE MODELS

Method	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism*
VGG11 [16]	34.59%	31.62%	25.89%	21.40%	167.83%
VGG13 [16]	30.08%	29.38%	22.30%	16.99%	153.13%
VGG16 [16]	28.06%	25.69%	24.95%	**15.29%**	140.17%
ResNet18 [63]	34.17%	32.46%	24.33%	20.40%	144.67%
DenseNet [64]	34.15%	33.06%	**21.91%**	26.45%	**135.50%**
SqueezeNet [65]	36.33%	34.56%	25.71%	19.63%	160.97%
MobileNetV3 [66]	46.41%	47.59%	29.24%	29.74%	187.89%
VGG19(ours)	**22.54%**	**20.63%**	22.42%	16.03%	152.64%

TABLE VI
FIVE PERSONALITY BIAS RATES WITH DIFFERENT SAMPLE RATES

Sample Rate (figures per second)	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism*
1	27.78%	18.26%	27.26%	24.12%	250.40%
2	27.64%	18.12%	26.88%	24.22%	243.20%
5	22.54%	20.63%	22.42%	**16.03%**	152.64%
10	**21.41%**	**16.90%**	**22.33%**	17.40%	**139.98%**

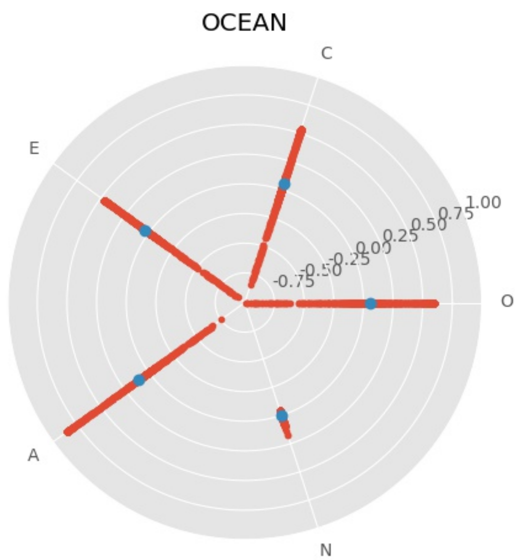

Fig. 6. Personality radar map based on the output of the OPO-FCM cognitive model.

$$\mathrm{Deviation}_{\mathrm{rate}_{\mathrm{Neuroticism}}} = \frac{\left|\frac{1}{n}\left(\sum_{t=1}^{n}\mathrm{Neuroticism}_{\mathrm{r}}\mathrm{ec}_{t}\right) - \mathrm{Neuroticism}\right|}{\mathrm{Neuroticism}_{\mathrm{r}}\mathrm{ec}_{\max} - \mathrm{Neuroticism}_{\mathrm{r}}\mathrm{ec}_{\min}}. \quad (19)$$

To the best of our knowledge, we are the first to combine both deep learning methods and traditional psychology formulas. Thus, our method ensures generalization and interpretability compared with deep learning methods. Compared with using scales to obtain the personality, our method does not cost much time and the results can be easily reproduced. The comparisons are shown in Table IV.

We conduct ablation studies on different backbone methods as well as different sample rates to assess the proposed method. As the main contribution of our method is the cognitive modeling framework, we simply use a classic effective backbone. With better backbone methods or training methodologies, the accuracy still can be improved. For different sample rates, we can see from Table VI that a higher sample rate leads to a lower bias rate. However, increasing the sampling rate from 1 to 10 means that the computational overhead is also increased by a factor of ten.

According to the above calculation method, the five personality deviation rates for all subjects of the OPO-FCM cognitive model shown in Table VII can be obtained.

V. DISCUSSION

The OPO-FCM cognitive model theoretically bridges the entire process from visual information input to the final five-factor data output and incorporates an exhaustive derivation of the formalization. In the experimental phase, the model proved that, except for one personality, neuroticism, which was objectively not validated, the average deviation rate of the personality results of the four valid tests was around 20.41%, i.e., the average accuracy rate was 79.56%.

For the bias results of neuroticism, the results of neuroticism and the bias rate of the subjects in the subsequent personality comparison results were influenced by the fact that it is relatively difficult to truly capture the relatively rare negative emotions in life through recall during the actual test. It is known from basic psychological theory that neuroticism is associated with a large number of negative emotions, but in the standard five-factor test experiment, there is no deliberately stimulated negative emotions, so the

TABLE VII
FIVE PERSONALITY BIAS RATES FOR ALL SUBJECTS ON THE OPO-FCM COGNITIVE MODEL

Subject ID	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism*
01	0.1738	0.2248	0.1236	0.0025	0.0523
02	0.1344	0.4677	0.3875	0.2981	5.4053
03	0.0789	0.0428	0.1627	0.4348	2.2964
04	0.2237	0.2469	0.4025	0.1868	1.2911
05	0.0255	0.1210	0.0048	0.1223	0.6033
06	0.3485	0.1260	0.1511	0.0904	0.2094
07	0.1306	0.2725	0.1407	0.2611	1.7200
08	0.3593	0.0063	0.2089	0.1016	1.9890
09	0.0419	0.0486	0.6177	0.0230	1.6676
10	0.3332	0.3269	0.0954	0.3358	0.2096
11	0.5965	0.7348	0.2372	0.3832	1.3015
12	0.1468	0.0840	0.1906	0.3095	3.4109
13	0.3015	0.1337	0.1820	0.0700	1.3184
14	0.1900	0.0131	0.1032	0.1425	0.5898
15	0.3586	0.4296	0.4893	0.0678	1.1078
16	0.1191	0.2293	0.1232	0.3294	0.1249
17	0.0842	0.0289	0.5237	0.2676	0.2039
18	0.3710	0.3459	0.1700	0.0027	3.4547
19	0.0273	0.0934	0.1960	0.1316	0.0281
20	0.1374	0.1296	0.1093	0.2014	0.3408
21	0.1135	0.0064	0.1162	0.1637	1.7027
22	0.1893	0.0214	0.0153	0.0429	2.2077
23	0.3751	0.0415	0.0514	0.0574	1.9540
24	0.3072	0.2136	0.1237	0.1690	4.8395
25	0.1898	0.2917	0.2324	0.0066	0.1539
26	0.3245	0.5364	0.4021	0.1784	2.7528
27	0.2293	0.2131	0.1818	0.0053	1.4537
28	0.4016	0.3470	0.5339	0.1021	0.3501
Deviation rate	22.54%	20.63%	22.42%	16.03%	152.64%

* Neuroticism was not effectively tested in the Big Five personality experimental environment

subject's neuroticism characteristics are not accurately tested and recorded in the experiment, which objectively causes the abnormal results of neuroticism personality. If the model is applied to a real-life scenario where the subjects are unconsciously recorded and analyzed, it is theoretically possible to effectively test neuroticism. Therefore, the next stage of the study is to conduct observations and experiments on a large number of scenarios with subjects' permission to verify the validity of the OPO-FCM cognitive model under large-scale observations.

VI. CONCLUSION

This study constructs a cognitive model that integrates the VGG-FACS-OCC model based on fer2013 expression features with the OCC-PAD-OCEAN fusion of the basic theory of emotion psychology, namely, the emotion-based OCC-PAD-OCEAN federal cognitive modeling approach. By constructing the model and performing formal proof algorithms, it is shown that the OPO-FCM cognitive model was able to complete the acquisition of expression features in videos through a deep neural network, followed by mapping them to the PAD emotion space through the established expression–basic emotions–emotion space mapping relationship, and finally mapping the average emotional state over time through the established temperament–personality mapping relationship to obtain the information in the personality space. The results of the experimental simulation of the model show that the average accuracy of the validly tested personality is at 79.56%.

Meanwhile, due to the time, ethical, and environmental constraints during the experimental process of the five-factor test, some personality traits are difficult to be elicited and identified, especially the neurotic personality that characterizes the tendency toward negative emotions. The reason for this may be that the Hawthorne effect resulted in the difficulty for the subjects to avoid the interference of related factors during the interview process. Therefore, the next step of the study is to conduct a massive data observation and experiment on the unconscious state of the subjects in a massive scenario with their permission, in order to verify the validity of the OPO-FCM cognitive model under large-scale observation and to continue to optimize the OPO-FCM cognitive model driven by the massive data.

In addition, besides the recognition of five-factor through facial expressions, physiological data such as human posture and behavior, speech and tone of voice, and ECG and blood oxygen can directly or indirectly reflect five-factor, therefore, how to integrate more dimensional elements and build a multimodal five-factor-based cognitive model is one of the main research directions for future cognitive modeling.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Feng Liu: conceptualization, writing original draft, review and editing, supervision, and project administration; Han-Yang Wang: conceptualization, writing original draft, methodology, software, and visualization; Si-Yuan Shen: conceptualization and writing original draft; Xun Jia: conceptualization and writing original draft; Jing-Yi Hu: conceptualization and writing original draft; Jia-Hao Zhang: conceptualization, methodology, software, and visualization; Xi-Yi Wang: conceptualization and writing; Ying Lei: software and visualization; Ai-Min Zhou: review and editing, and supervision; Jia-Yin Qi: review and editing, and supervision; and Zhi-Bin Li: review and editing, and supervision.

REFERENCES

[1] P. Ekman and W. V. Friesen, *Unmasking the Face: A Guide to Recognizing Emotions From Facial Clues*. Malor Books, Dec. 2003.

[2] P. M. C. Harrison, T. Collins, and D. Müllensiefen, "Applying modern psychometric techniques to melodic discrimination testing: Item response theory, computerised adaptive testing, and automatic item generation," *Sci. Rep.*, vol. 7, no. 1, pp. 1–18, Dec. 2017.

[3] W. Zheng, H. Cai, Z. Yao, X. Zhang, X. Li, and B. Hu, "Modelling mental states via computational psychophysiology: Benefits and challenges," in *Proc. Int. Conf. Hum. Centered Comput.* Cham, Switzerland: Springer, 2019, pp. 659–670.

[4] M. Moghimi, R. Stone, and P. Rotshtein, "Affective recognition in dynamic and interactive virtual environments," *IEEE Trans. Affect. Comput.*, vol. 11, no. 1, pp. 45–62, Jan. 2020.

[5] J. Marín-Morales *et al.*, "Affective computing in virtual reality: Emotion recognition from brain and heartbeat dynamics using wearable sensors," *Sci. Rep.*, vol. 8, no. 1, pp. 1–15, Dec. 2018.

[6] A. Vinciarelli and G. Mohammadi, "A survey of personality computing," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 273–291, Jun. 2014.

[7] H. Ning, S. Dhelim, and N. Aung, "PersoNet: Friend recommendation system based on big-five personality traits and hybrid filtering," *IEEE Trans. Computat. Social Syst.*, vol. 6, no. 3, pp. 394–402, Jun. 2019.

[8] S. Dhelim, H. Ning, N. Aung, R. Huang, and J. Ma, "Personality-aware product recommendation system based on user interests mining and metapath discovery," *IEEE Trans. Computat. Social Syst.*, vol. 8, no. 1, pp. 86–98, Feb. 2021.

[9] G. Matthews, I. J. Deary, and M. C. Whiteman, *Personality Traits*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[10] J. Wiggins, *The Five-Factor Model of Personality: Theoretical Perspectives*. New York, NY, USA: The Guilford Press, 1996.

[11] G. Saucier and L. R. Goldberg, "The language of personality: Lexical perspectives on the five-factor model," in *The Five-Factor Model of Personality: Theoretical Perspectives*. New York, NY, USA: Guilford Press, 1996, pp. 21–50.

[12] R. R. McCrae, "The five-factor model of personality traits: Consensus and controversy," in *The Cambridge Handbook of Personality Psychology*, P. J. Corr and G. Matthews, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2009, pp. 148–161.

[13] A. El Bolock, "What is character computing?" in *Character Computing* (Human-Computer Interaction Series), A. El Bolock, Y. Abdelrahman, and S. Abdennadher, Eds. Cham, Switzerland: Springer, 2020, pp. 1–16.

[14] S. Li and W. Deng, "Deep facial expression recognition: A survey," *IEEE Trans. Affect. Comput.*, early access, Mar. 17, 2020, doi: 10.1109/TAFFC.2020.2981446.

[15] J. Comas, D. Aspandi, and X. Binefa, "End-to-end facial and physiological model for affective computing and applications," in *Proc. 15th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, Nov. 2020, pp. 93–100.

[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[17] A. Olsson, E. Knapska, and B. Lindström, "The neural and computational systems of social learning," *Nature Rev. Neurosci.*, vol. 21, no. 4, pp. 197–212, 2020.

[18] R. Fan *et al.*, "The minute-scale dynamics of online emotions reveal the effects of affect labeling," *Nature Hum. Behav.*, vol. 3, no. 1, pp. 92–100, Jan. 2019.

[19] A. J. Calder and A. W. Young, "Understanding the recognition of facial identity and facial expression," *Nature Rev. Neurosci.*, vol. 6, no. 8, pp. 641–651, Aug. 2005.

[20] M. Orri *et al.*, "Identifying affective personality profiles: A latent profile analysis of the affective neuroscience personality scales," *Sci. Rep.*, vol. 7, no. 1, pp. 1–14, Dec. 2017.

[21] E. Dejonckheere *et al.*, "Complex affect dynamics add limited information to the prediction of psychological well-being," *Nature Hum. Behav.*, vol. 3, no. 5, pp. 478–491, May 2019.

[22] H. Swanborough, M. Staib, and S. Frühholz, "Neurocognitive dynamics of near-threshold voice signal detection and affective voice evaluation," *Sci. Adv.*, vol. 6, no. 50, Dec. 2020, Art. no. eabb3884.

[23] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press, 2000.

[24] P. Gebhard, "ALMA: A layered model of affect," in *Proc. 4th Int. Joint Conf. Auto. Agents Multiagent Syst. (AAMAS)*, 2005, pp. 29–36.

[25] P. Gebhard, M. Kipp, M. Klesen, and T. Rist, "Adding the emotional dimension to scripting character dialogues," in *Proc. Int. Workshop Intell. Virtual Agents*. Berlin, Germany: Springer, 2003, pp. 48–56.

[26] P. Gebhard, M. Klesen, and T. Rist, "Coloring multi-character conversations through the expression of emotions," in *Proc. Tutorial Res. Workshop Affect. Dialogue Syst.* Berlin, Germany: Springer, 2004, pp. 128–141.

[27] P. Ekman, "An argument for basic emotions," *Cognit. Emotion*, vol. 6, nos. 3–4, pp. 169–200, 1992.

[28] J. A. Russell, "A circumplex model of affect," *J. Personality Social Psychol.*, vol. 39, no. 6, pp. 1161–1178, 1980.

[29] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychol.*, vol. 14, no. 4, pp. 261–292, 1996.

[30] D. Grandjean, D. Sander, and K. R. Scherer, "Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization," *Consciousness Cognition*, vol. 17, no. 2, pp. 484–495, 2008.

[31] A. Ortony, G. L. Clore, and A. Collins, *The Cognitive Structure of Emotions*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[32] C. Bartneck, "Integrating the OCC model of emotions in embodied characters," in *Proc. Workshop Virtual Conversational Characters, Appl., Methods, Res. Challenges*, 2002, pp. 1–5.

[33] G. L. Zhao, J. J. Song, Y. Ge, Y. Liu, L. Yao, and T. Wen, "Advances in emotion recognition based on physiological big data," *J. Comput. Res. Develop.*, vol. 53, no. 1, pp. 80–92, 2016.

[34] R. Cowie, E. Douglas-Cowie, S. Savvidou, E. McMahon, M. Sawey, and M. Schröder, "'FEELTRACE': An instrument for recording perceived emotion in real time," in *Proc. ISCA Tuts. Res. Workshop (ITRW) Speech Emotion*, 2000, pp. 1–6.

[35] R. Cowie *et al.*, "Emotion recognition in human-computer interaction," *IEEE Signal Process. Mag.*, vol. 18, no. 1, pp. 32–80, Jan. 2001.

[36] M. Wöllmer *et al.*, "Abandoning emotion classes—Towards continuous emotion recognition with modelling of long-range dependencies," in *Proc. 9th Interspeech, 12th Australas. Int. Conf. Speech Sci. Technol.*, Brisbane, QD, Australia, Sep. 2008, pp. 597–600.

[37] E. Douglas-Cowie *et al.*, "The HUMAINE database: Addressing the needs of the affective computing community," in *Proc. 2nd Int. Conf. Affect. Comput. Intell. Interact.*, 2007, pp. 488–500.

[38] B. Schuller, M. Wimmer, L. Mosenlechner, C. Kern, D. Arsic, and G. Rigoll, "Brute-forcing hierarchical functionals for paralinguistics: A waste of feature space?" in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 4501–4504.

[39] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[40] C. Yiling *et al.*, "Speech emotion recognition in PAD three-dimensional emotion space," *J. Harbin Inst. Technol.*, vol. 50, no. 11, pp. 160–166, 2018.

[41] R. Cowie, H. Gunes, G. McKeown, L. Vaclau-Schneider, J. Armstrong, and E. Douglas-Cowie, "The emotional and communicative significance of head nods and shakes in a naturalistic database," in *Proc. LREC Int. Workshop Emotion*, 2010, pp. 42–46.

[42] M. Kipp and J.-C. Martin, "Gesture and emotion: Can basic gestural form features discriminate emotions?" in *Proc. 3rd Int. Conf. Affect. Comput. Intell. Interact. Workshops*, Sep. 2009, pp. 1–8.

[43] M. A. Nicolaou, H. Gunes, and M. Pantic, "Output-associative RVM regression for dimensional and continuous emotion prediction," *Image Vis. Comput.*, vol. 30, no. 3, pp. 186–196, Mar. 2012.

[44] Y.-L. Hsu, J.-S. Wang, W.-C. Chiang, and C.-H. Hung, "Automatic ECG-based emotion recognition in music listening," *IEEE Trans. Affect. Comput.*, vol. 11, no. 1, pp. 85–99, Jan. 2020.

[45] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Trans. Affect. Comput.*, vol. 11, no. 3, pp. 532–541, Jul./Sep. 2020.

[46] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 327–339, Jul. 2014.

[47] S. M. Alarcão and M. J. Fonseca, "Emotions recognition using EEG signals: A survey," *IEEE Trans. Affect. Comput.*, vol. 10, no. 3, pp. 374–393, Jul./Sep. 2019.

[48] C. X. Wan, T. J. Jiang, M. J. Zhong, and H. R. Bian, "Sentiment computing of web financial information based on the part-of-speech tagging and dependency parsing," *J. Comput. Res. Develop.*, vol. 50, no. 12, pp. 2554–2569, 2013.

[49] S. Fengfang *et al.*, "A new approach to personality_OCC emotion modeling," *J. Jilin Univ., Sci. Ed.*, vol. 50, no. 1, pp. 106–110, 2012.

[50] M. S. El-Nasr, J. Yen, and T. R. Ioerger, "Flame-fuzzy logic adaptive model of emotions," *Auton. Agents Multi-Agent Syst.*, vol. 3, no. 3, pp. 219–257, 2000.

[51] S. Yang, R. Huang, and J. Ma, "A computational personality-based and event-driven emotions model in PAD space," in *Proc. Int. Conf. Intell. Sci. Intell. Data Eng.* Berlin, Germany: Springer, 2012, pp. 825–831.

[52] A. A. Marouf, M. K. Hasan, and H. Mahmud, "Comparative analysis of feature selection algorithms for computational personality prediction from social media," *IEEE Trans. Computat. Social Syst.*, vol. 7, no. 3, pp. 587–599, Jun. 2020.

[53] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016.

[54] A. Mehrabian, "Analysis of the big-five personality factors in terms of the PAD temperament model," *Austral. J. Psychol.*, vol. 48, no. 2, pp. 86–92, Aug. 1996.

[55] A. Mehrabian and E. O'Reilly, "Analysis of personality measures in terms of basic dimensions of temperament," *J. Personality Social Psychol.*, vol. 38, no. 3, pp. 492–503, Mar. 1980.

[56] R. R. McCrae and O. P. John, "An introduction to the five-factor model and its applications," *J. Pers.*, vol. 60, no. 2, pp. 175–215, 1992.

[57] R. Subramanian, J. Wache, M. K. Abadi, R. L. Vieriu, S. Winkler, and N. Sebe, "ASCERTAIN: Emotion and personality recognition using commercial sensors," *IEEE Trans. Affect. Comput.*, vol. 9, no. 2, pp. 147–160, Apr./Jun. 2018.

[58] F. Anselmi, N. Noceti, L. Rosasco, and R. Ward, "Genuine personality recognition from highly constrained face images," in *Proc. Int. Conf. Image Anal. Process.* Cham, Switzerland: Springer, 2019, pp. 421–431.

[59] K. Ilmini and T. Fernando, "Persons' personality traits recognition using machine learning algorithms and image processing techniques," *Adv. Comput. Sci., Int. J.*, vol. 5, no. 1, pp. 40–44, 2016.

[60] L. Wu, D. Zhang, Q. Liu, S. Li, and G. Zhou, "Speaker personality recognition with multimodal explicit many2many interactions," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2020, pp. 1–6.

[61] S. Song, S. Jaiswal, E. Sanchez, G. Tzimiropoulos, L. Shen, and M. Valstar, "Self-supervised learning of person-specific facial dynamics for automatic personality recognition," *IEEE Trans. Affect. Comput.*, early access, Mar. 9, 2021, doi: 10.1109/TAFFC.2021.3064601.

[62] P. Antonova and D. Popov, "Comparison of personality recognition algorithms," in *Proc. 4th Sci. School Dyn. Complex Netw. Appl. Intellectual Robot. (DCNAIR)*, 2020, pp. 34–36.

[63] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[64] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2017, pp. 4700–4708.

[65] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size," 2016, *arXiv:1602.07360*.

[66] A. Howard *et al.*, "Searching for MobileNetV3," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Oct. 2019, pp. 1314–1324.

Feng Liu (Member, IEEE) is currently pursuing the Ph.D. degree with the School of Computer Science and Technology, East China Normal University, Shanghai, China.

His current research interests include computational affection, deep learning, and blockchain technology.

Mr. Liu is a Senior Member of China Computer Federation and a member of the Chinese Psychological Society.

Han-Yang Wang received the B.E. degree from the School of Computer Science and Technology, East China Normal University, Shanghai, China, in 2021, where he is currently pursuing the master's degree with the College of Computer Science and Technology.

His current research interests include computational affection, deep learning, and generative adversarial networks.

Si-Yuan Shen received the B.E. degree from the School of Computer Science and Technology, East China Normal University, Shanghai, China, in 2021, where he is currently pursuing the master's degree with the College of Computer Science and Technology.

His current research interests include computational affection, deep learning, and speech emotion recognition.

Xun Jia is currently pursuing the B.E. degree with the School of Psychology and Cognitive Science, East China Normal University, Shanghai, China.

His current research interests include computational affection and emotional psychology.

Jing-Yi Hu received the B.E. degree from the School of Business Administration, East China Normal University, Shanghai, China, in 2021, where she is currently pursuing the master's degree.

Her current research interests include computational affection and consumer psychology.

Jia-Hao Zhang is currently pursuing the B.E. degree with the School of Computer Science and Technology, East China Normal University, Shanghai, China.

Her current research interests include computational affection and computer vision.

Xi-Yi Wang is currently pursuing the B.E. degree with the School of Psychology and Cognitive Science, East China Normal University, Shanghai, China.

Her current research interests include computational affection and emotional psychology.

Ying Lei is currently pursuing the B.E. degree with the School of Computer Science and Technology, East China Normal University, Shanghai, China.

Her current research interests include computational affection, deep learning, and EEG emotion recognition.

Ai-Min Zhou (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in computer science from Wuhan University, Wuhan, China, in 2001 and 2003, respectively, and the Ph.D. degree in computer science from the University of Essex, Colchester, U.K., in 2009.

He is currently a Professor with the School of Computer Science and Technology and the Shanghai Institute of AI for Education, East China Normal University, Shanghai, China. He has authored over 80 peer-reviewed articles. His current research interests include evolutionary computation and optimization, machine learning, image processing, and their applications.

Dr. Zhou is an Associate Editor of *Swarm and Evolutionary Computation* and an Editorial Board Member of *Complex & Intelligent Systems*.

Jia-Yin Qi is currently a Professor and also the Dean of the Institute of Artificial Intelligence and Change Management, Shanghai University of International Business and Economics, Shanghai, China.

Her research interests are social network analytics, big data modeling, artificial intelligence, and blockchain.

Prof. Qi was also recognized by Elsevier as one of the Highest Cited Chinese Scholars from 2014 to 2020.

Zhi-Bin Li is currently a Professor with East China Normal University, Shanghai, China. His current research interests include symbolic calculation on analytical solutions of mathematical physics equations, and computer symbolic calculation and its application in nonlinear science.