IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 1, 2019, accepted February 20, 2019, date of publication March 5, 2019, date of current version May 22, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2902863

TensorFlow-Based Automatic Personality Recognition Used in Asynchronous Video Interviews

HUNG-YUE SUEN [1], KUO-EN HUNG [1], AND CHIEN-LIANG LIN [2]

[1]Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taipei City 106 , Taiwan
[2]Department of Management Information System, National Chengchi University, Taipei City 116, Taiwan

Corresponding author: Chien-Liang Lin (lin.chienliang@gmail.com)

This work was supported by the Ministry of Science and Technology, Taiwan, under Grant MOST-107-2511-H-003-040-MY2.

ABSTRACT With the development of artificial intelligence (AI), the automatic analysis of video interviews to recognize individual personality traits has become an active area of research and has applications in personality computing, human–computer interaction, and psychological assessment. Advances in computer vision and pattern recognition based on deep learning (DL) techniques have led to the establishment of convolutional neural network models that can successfully recognize human nonverbal cues and attribute their personality traits with the use of a camera. In this paper, an end-to-end AI interviewing system was developed using asynchronous video interview (AVI) processing and a TensorFlow AI engine to perform automatic personality recognition (APR) based on the features extracted from the AVIs and the true personality scores from the facial expressions and self-reported questionnaires of 120 real job applicants. The experimental results show that our AI-based interview agent can successfully recognize the ''big five'' traits of an interviewee at an accuracy between 90.9% and 97.4%. Our experiment also indicates that although the machine learning was conducted without large-scale data, the semisupervised DL approach performed surprisingly well with regard to APR despite the lack of labor-intensive manual annotation and labeling. The AI-based interview agent can supplement or replace existing self-reported personality assessment methods that job applicants may distort to achieve socially desirable effects.

INDEX TERMS Big five, convolutional neural network (CNN), personality computing, TensorFlow.

I. INTRODUCTION

Industrial and organizational (I/O) psychologists have found that personality is a global predictor used in employment selection [1]. Some employers use self-reported surveys to measure job applicants' personalities; however, job applicants may lie when self-reporting personality traits to gain more job opportunities [2]. Some employers evaluate the applicants' personalities from their facial expressions and other nonverbal cues during job interviews because applicants have considerable difficulty faking nonverbal cues [3]. However, it is not practical for every job applicant to attend a live job interview in person or participate in interviews conducted through telephone calls or web conferences due to the cost and time limitations [4]. One-way asynchronous video interview (AVI) software can be used to automatically interview job applicants at one point in time. This approach allows employers to review the audio-visual records at a later point in time [5]. When using AVI, human raters find it cognitively challenging to correctly assess applicants' personality traits based on video images [6]. Barrick *et al*. [7] found that human raters were unable to accurately assess an applicant's personality simply by watching recorded-video interviews.

Both I/O psychology and computer science scholars have suggested that artificial intelligence (AI) may surpass humans in recognizing or predicting an applicant's personality for screening job applicants because applying AI techniques to audio-visual datasets can achieve more reliable and predictive power than human raters [8]–[11]. ''AI is a branch of computer science that seeks to produce intelligent machines that respond in a manner similar to human intelligence'' [12], and it ''aims to extend and augment human capacity and efficiency of mankind in tasks of remaking nature'' [13]. Machine learning (ML) is a major approach for achieving AI, which ''gives computers the ability to learn without being explicitly programmed'' [14]. Deep learning (DL) is a

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Radhakrishnan.

technique to implement ML, and it can "mimic the human brain mechanism to interpret data such as images, sounds and texts" [15]. In contrast to traditional ML, DL feature extraction is automated rather than manual [12].

ML/DL can be divided into supervised learning, unsupervised learning, and semi-supervised learning [12]. Supervised learning tasks are commonly conducted by classification using predefined labeled training data (called "ground truth"), whereas unsupervised learning can automatically learn the correct answers from a large amount of data without requiring predefined labels [10], [16]. Semi-supervised learning combines those two approaches by using relatively smaller amounts of unlabeled data plus some labeled data for pattern recognition; therefore, this approach can reduce labeling efforts yet still achieve high accuracy.

Previous automatic personality recognition (APR) studies were developed based on supervised ML, which involves manual labeling work and is time consuming [17]. Because convolutional neural networks (CNNs) have been proven to be high-performing models that can automatically process images and infer first impressions from camera images, this study implemented semi-supervised DL methods, including CNNs, to develop an AI-based interview agent that can automatically recognize a job applicant's personality by using relatively smaller datasets of the applicants' facial expressions [18].

The remainder of this article is structured as follows: In Section 2, we discuss the background of APR from audio-visual data. Section 3 describes our data processing approach. A detailed model and its results are presented in Section 4. Finally, we discuss and conclude our findings and future work in Section 5.

II. BACKGROUND

A. PERSONALITY TAXONOMY

Personality refers to "individual differences in characteristic patterns of thinking, feeling, and behaving" [19]. This construct is commonly used to predict whether a job candidate will perform well in a specific job role and engage well in a prospective cultural environment [20]. Although a variety of models can be used to assess personality, the "big five" traits, also called the five-factor model (FFM) or OCEAN model, provide researchers and practitioners with a well-defined taxonomy for selecting job applicants [20]. The core factors of the big five are categorized and applied in different cultural contexts; these factors are openness, conscientiousness, extraversion, agreeableness, and neuroticism (low emotional stability) [21]:

- **O**penness: the degree to which an individual is imaginative and creative.
- **C**onscientiousness: the degree to which an individual is organized, thorough, and thoughtful.
- **E**xtraversion: the extent to which an individual is talkative, energetic, and assertive.
- **A**greeableness: the degree to which an individual is sympathetic, kind, and affectionate.
- **N**euroticism: reflects the tension, moodiness, and anxiety an individual may feel.

Different approaches exist to measuring an individual's big five traits, including self-rating and observer-rating. Self-rating reflects self-image, whereas observer-rating reflects the subjective impressions perceived by others toward an individual's personality [7]. In the self-perspective approach, personality refers to a person's described motives, intensions, feelings, and past behaviors. From the observer's perspective, personality incorporates information about a person's social reputation, but valid observer-ratings should ideally be obtained by close acquaintances, such as partners, friends, or coworkers [19]. In the I/O psychology literature, when the valid observer-rated big five traits are difficult to assess, self-ratings are the foundational information used to predict individual workplace behaviors and performance. Self-ratings can also be used to predict whether a job candidate is a good fit for the job requirements and the organizational culture in a zero-acquaintance context, such as a job interview [19].

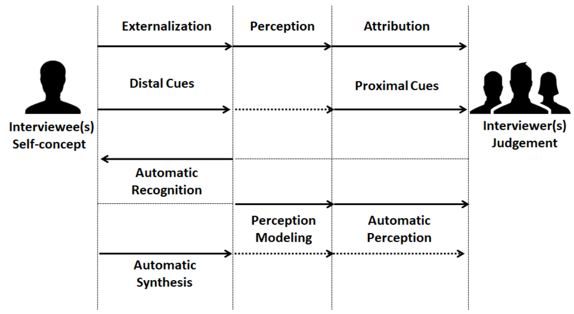


FIGURE 1. Brunswik's lens model describes personality externalization, perception, and attribution during job interviews. Revised and adapted from "A Survey of Personality Computing," by Vinciarelli and Mohammadi [6].

B. PERSONALITY COMPUTING

According to social information processing theory [22], people observe and interpret the cues exhibited by others and draw conclusions regarding their personalities during interactions such as interviews. Brunswik's lens model, depicted in Figure 1, illustrates how an interviewer uses cues to judge the interviewee's personality and to show the relationship between the interviewee's self-assessed personality and the interviewer's perceptual observations of personality regarding the interviewee [6].

The interviewees externalize their apparent personality through distal cues (i.e., any observable behaviors that can be perceived by the interviewer, such as facial expression, gaze, posture, body movement, speaking, and prosody). Alternatively, the interviewer uses a "lens" to attribute the unobservable personality traits of the interviewee through proximal cues (i.e., any interviewee behaviors that are actually perceived by the interviewer, including indirect observable cues); nonetheless, these cues can translate into perceptions by the interviewer [6], [23].

Personality psychology researchers have found that despite only the head and torso of the applicants being visible in AVI, an interviewer or rater can still use nonverbal cues to judge the applicants' personality traits [24]. Some experimental studies have shown that individuals can attribute the valid personality traits of zero acquaintances based on brief video clips [23]. Personality computing, an emerging research area related to AI and personality psychology, is being used to automatically recognize, perceive, and synthesize human behavioral cues and personality based on the lens model. Three approaches in personality computing to auto-assess personality are APR, automatic personality perception (APP), and automatic personality synthesis (APS) [6].

By extracting features from the audio-visual data of AVI, APR is intended to auto-recognize an interviewee's self-assessed personality from distal cues [9], [25]–[27]. In contrast, APP is intended to auto-predict the observer-rated personality of an interviewee from proximal cues [28]. Because examining proximal cues is not easy, APP instead uses distal cues as an approximation, as described in [8], [10], [17], and [29]–[33]. In APS, artificial agents, avatars, or robots are used to display human-like distal cues and the goal is for these cues to be perceived and inferred by humans as predefined traits [34], [35]. To develop AI-based APR, AVI can be adapted and combined with APR to auto-recognize interviewees' "true scores" compared to their self-rated personality traits [6].

I/O psychology studies have found that an individual's dynamic facial expressions, such as facial dominance, smiling, or a tense aspect, reflects his or her self-rated big five traits (e.g., extraversion and agreeableness) [36]–[38]. Computer science studies have found that CNNs can be used to recognize a person's big five traits based on facial expressions extracted from video clips; in fact, these computing models have achieved greater predictive power than human raters [18], [39]–[41]. Human raters/observers may have biases (implicit or explicit) that impact how interviewee cues are interpreted, whereas a computer does not have implicit biases: we can expect that a computer will evaluate all interviewees using the same criteria and make personality judgments more consistent and fair compared with those of human raters (see [6]).

III. DATA PROCESSING

A. DATA COLLECTION

To establish our dataset in a real job interview context, we developed AVI cloud-based software, similar to the work in [42]. The AVI server uses Google cloud storage and can receive recorded video prompts, generate interview scripts, transmit the video prompts from the interview, and receive the video responses. The content of the video responses can then be used to conduct algorithmic analyses, including audio and visual data analyses of the video responses. During the AVI, interviewees' answers can be recorded at one point in time but later reviewed by an algorithm, human raters, or both at another point in time.

We conducted an experiment with sponsorship from a nonprofit human resources (HR) organization located in Hong Kong. Our sponsor hosted a web page accessible to their members that contained a job description for hiring 2–3 HR professionals from its affiliated company located in Shenzhen, Guangdong, China. Interested members submitted their resumes to the researchers, and the researchers screened the received resumes based on the job description. A total of 120 applicants were invited to login to the AVI, which delivered predefined interview questions via a browser to interviewees' web camera equipped mobile or computer devices at their leisure at any time of day. The interviewees' answers, including both audio and visual information, were recorded for analysis. The applicants were informed that the entirety of their interview processes and responses, including audio and video, would be recorded and analyzed by our algorithms and used as references for hiring recommendations.

The interview questions during the AVI were structured in a standard manner. All the applicants were provided with the same five questions, which were behaviorally oriented to assess the applicants' communication skills based on the job description. Each question was displayed on a new screen, and the audio of the text questions was automatically started when the applicants entered the screen. The questions were presented on-screen one at a time in sequence, and the applicants were given a maximum of 3 minutes to answer each question. The applicants could choose to skip to the next question within the 3-minute period. After 3 minutes, a new screen automatically appeared with the next question. Including one practice trial, the entire video interview process lasted approximately 20 minutes.

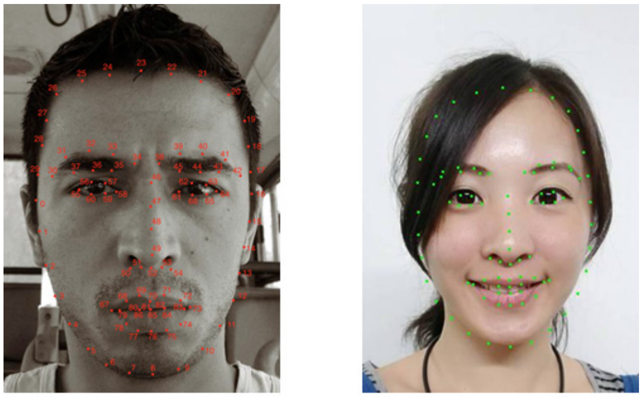
B. DATA LABELING

To collect the true ratings for the individual big five traits [6], we used a 50-item international personality item pool (IPIP) inventory developed in [43] to measure the applicants' self-rated big five traits. Prior to participating in the AVI, all applicants were required to complete the IPIP survey online and informed that the survey results would be delivered to researchers only and that they would be irrelevant to the hiring recommendation. This procedure was conducted to reduce the effects of social desire, which may distort the self-rated personality traits in an effort to gain the job opportunity [44].

C. FEATURE EXTRACTION

To capture the applicants' facial expressions, we started with the pretrained Inception-v3 dataset collected for ImageNet, which includes more than 14 million images grouped into 1,000 classes. Additionally, we trained our facial detection model based on OpenCV and Dlib while tracking 86 facial landmark points per frame, as shown in Figure 2. Moreover, we used landmark point 47 on the nasal root (see the picture of training data in Figure 2) as the anchor point position during feature extraction to reduce background noise and minimize errors such as head motion because this landmark point is little affected by facial expressions [45].

To develop the feature extractor, we extracted the images frame by frame from our AVI dataset using FFmpeg.



Training data Real case

FIGURE 2. Image annotation.

FIGURE 3. Extracted video frames.

The width of all the images was normalized to 640 pixels, while the height of each image was determined by the pixel ratio of the vision device. We extracted the features of the 86 landmark points from each frame within a 5-second period from among all the AVI records for each applicant, as shown in Figure 3. To improve image classification and reduce background interference from hair and cosmetics, we converted all the images to grayscale. The test cases used in this experiment comprised more than 10,000 images.

IV. MODELING AND RESULTS

A. MODEL BUILDING

We combined the personality-labeled data of the 120 applicants with their extracted features to train our APR model, which was constructed of an advanced CNN built using Python and the TensorFlow DL engine [46]. Prior to inputting the images into the neural network model, we normalized the features by rescaling the range of feature values to [0, 1]. In the neural network model, the first several layers, which included more than 72 thousand neurons, were intended to automatically extract features from the images preprocessed as described in the previous section. Then the extracted features were concatenated with other features and fed to the output layer for final classification. As illustrated in Figure 4, our CNN structure consisted of four convolutional layers, three pooling layers, ten mixed layers, a fully connected layer and a softmax layer as the output.

In the neural network, the extracted features of the applicants' facial expressions were used as the inputs, and their

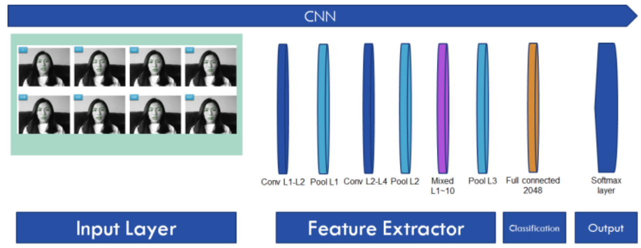


FIGURE 4. The CNN structure.

self-assessed big five personality trait scores were used as the output. The fully connected network was built using the relationships between different nodes, as shown in Figure 4. The input was represented by a grayscale image. Each of the four convolutional layers used 3×3 filter functions. The number of convolution filters increased from 32 in convolutional layer 1 to 64 in convolutional layer 2. Each convolutional layer was followed by a pooling layer. Both pooling layers (one average-pooling and two max-pooling) had a stride of 2×2 and the dropout rate was set to 0.1. The final fully connected layer included 2,048 neurons with dropouts of 0.4 and 0.5. The final layer of the proposed CNN was a softmax layer with 50 possible outputs (10 interval-scale classes from 1.1 to 6.0 that reflected the big five personality classifications). To prevent overfitting, we added dropout starting with a probability of 0.5 after the fully connected layers at the end of our convolutional network and then gradually reduced the dropout rate until the performance was maximized.

In these experiments, we separated the data as follows: the test set was 50% and the validation set was 50% based on the same sampling. Each applicant had 5 different features (big five traits) in this dataset. We conducted 4,000 training iterations. The learning rate was 0.01, the evaluation frequency was 10, and the training batch size was 256.

B. CRITERIA FOR THE ASSESSMENT

We used concurrent validity to assess the performance of our APR so that we could measure how well the new measurement procedure (the APR) correlated with the well-established measurement procedure (the self-reported inventory). Following [6] and [47], we used the Pearson correlation coefficient (r) to measure the concurrent validity in this experiment. Moreover, we followed [8] for measuring the coefficient of determination (R^2) and the mean square error (*MSE*). The R^2 indicates the variance in the dependent variable (y) that can be predicted or explained by the predictor. The higher the R^2, the better the model is. The *MSE* measures the goodness of fit of the regression model; the larger this number is, the larger the error is.

C. RESULTS

Prior to assessing our APR's performance, we used IBM's statistical package for the social sciences (SPSS v 23) to test the construct validity and internal consistency reliability for the self-reported personality traits. In this study, the construct validity was satisfactory because a confirmatory factor analysis showed that each factor loading was greater than 0.6, while the Kaiser-Meyer-Olkin (KMO) value was more than 0.8. The internal consistency reliability was good because the Cronbach's alpha (α) values were all larger than 0.7 as follows: Openness to experience ($\alpha = .75$), conscientiousness ($\alpha = .83$), extraversion ($\alpha = .88$), agreeableness ($\alpha = .80$), and neuroticism ($\alpha = .84$).

TABLE 1. Experimental results.

Personality Traits	*R*	R^2	*MSE*	*ACC%*
Openness to experience	0.966	0.933	0.053	97.4
Conscientiousness	0.976	0.952	0.094	96.7
Extraversion	0.974	0.949	0.120	97.0
Agreeableness	0.971	0.943	0.069	90.9
Neuroticism	0.968	0.937	0.092	94.8

$p < .01$

As indicated in Table 1, all the dimensions of the big five traits were learned and predicted successfully by the AI TensorFlow engine. All the true big five personality self-assessment scores could be predicted by APR. The Pearson correlation for each dimension was between 0.966 and 0.976. The R^2 for each dimension was between 0.933 and 0.952. All the correlations were found to be significant (p < 0:01), while the MSE for each dimension was between 0.053 and 0.120. The higher the R^2 is (100% is perfect), the better the estimator is. Conversely, the lower the MSE is (0 is perfect), the smaller the estimator error is. Additionally, the classification accuracy results show that the average accuracy of the classifiers (ACC) was 95.36%.

V. CONCLUSIONS AND FUTURE DIRECTIONS

This study is a response to the call for research into personality computing [40], [48], [49]. In traditional personality computing, validating APR using manually labeled features from any possible detectable distal cues was quite complicated [6]. Thus, some recent studies have adopted DL-based architectures to predict personality based on third-party datasets, such as Amazon's Mechanical Turk or ChaLearn's First Impressions dataset [40]. However, most of these studies used APP, in which the DL engines mimicked human raters as observers detecting an interviewee's nonverbal cues and made inferences concerning the interviewees' personality traits in the context of zero-acquaintance judgements. In other words, these experiments used subjective personality impressions rather than true personality scores [6] as the independent variables, which may have introduced existing biases [18].

This paper developed an AVI embedded with a TensorFlow-based semi-supervised DL model to accurately auto-recognize an interviewee's true personality based on only 120 real samples of job applicants. Our APR approach achieved an accuracy above 90%, outperforming previous related laboratory studies whose accuracy ranged between 61% and 75% in the context of nonverbal communication [6]. The high-performing APR used in this AVI can be adopted to supplement or replace self-reported personality assessment methods that can be distorted by job applicants due to the effects of social desire to be selected for employment.

Previous related studies have found that multimodal features (image frames and audio) learned by deep neural networks can deliver better performances in predicting the big five traits than can unimodal features. In future work, we may combine our visual approach with prosodic features to learn how to recognize an interviewee's personality. Moreover, this study utilized a specific type of professional as participants, which may limit the generalizability of these experimental results. Future research should include a more diverse participant population.

REFERENCES

[1] A. M. Ryan *et al.*, "Culture and testing practices: Is the world flat?" *Appl. Psychol.*, vol. 66, no. 3, pp. 434–467, Jul. 2017.
[2] M. J. W. McLarnon, A. C. DeLongchamp, and T. J. Schneider, "Faking it! Individual differences in types and degrees of faking behavior," *Personality Individual Differences*, vol. 138, pp. 88–95, Feb. 2019.
[3] T. DeGroot and J. Gooty, "Can nonverbal cues be used to make meaningful personality attributions in employment interviews?" *J. Bus. Psychol.*, vol. 24, no. 2, pp. 179–192, Feb. 2009.
[4] I. Nikolaou and K. Foti, "Personnel selection and personality," in *The SAGE Handbook of Personality and Individual Differences*, V. Zeigler-Hill and T. K. Shackelford, Eds., London, U.K.: Sage, 2018, pp. 659–677.
[5] F. S. Brenner, T. M. Ortner, and D. Fay, "Asynchronous video interviewing as a new technology in personnel selection: The applicant's point of view," *Frontiers Psychol.*, vol. 7, p. 863, Jun. 2016.
[6] A. Vinciarelli and G. Mohammadi, "A survey of personality computing," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 273–291, Jul. 2014.
[7] M. R. Barrick, G. K. Patton, and S. N. Haugland, "Accuracy of interviewer judgments of job applicant personality traits," *Personnel Psychol.*, vol. 53, no. 4, pp. 925–951, Dec. 2000.
[8] O. Celiktutan and H. Gunes, "Automatic prediction of impressions in time and across varying context: Personality, attractiveness and likeability," *IEEE Trans. Affect. Comput.*, vol. 8, no. 1, pp. 29–42, Jan. 2017.
[9] N. Y. Asabere, A. Acakpovi, and M. B. Michael, "Improving socially-aware recommendation accuracy through personality," *IEEE Trans. Affective Comput.*, vol. 9, no. 3, pp. 351–361, Jul./Sep. 2018.
[10] I. Naim, M. I. Tanveer, D. Gildea, and M. E. Hoque, "Automated analysis and prediction of job interview performance," *IEEE Trans. Affective Comput.*, vol. 9, no. 2, pp. 191–204, Apr./Jun. 2018.
[11] M. Langer, C. J. König, and K. Krause, "Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings," *Int. J. Selection Assessment*, vol. 25, no. 4, pp. 371–382, Dec. 2017.
[12] Y. Xin *et al.*, "Machine learning and deep learning methods for cybersecurity," *IEEE Access*, vol. 6, no, pp. 35365–35381, 2018.
[13] J. Liu *et al.*, "Artificial intelligence in the 21st century," *IEEE Access*, vol. 6, pp. 34403–34421, 2018.
[14] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.
[15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[16] K. K. Htike and O. O. Khalifa, "Comparison of supervised and unsupervised learning classifiers for human posture recognition," in *Proc. Int. Conf. Comput. Commun. Eng. (ICCCE)*, Kuala Lumpur, Malaysia, 2010, pp. 1–6.
[17] B. Aydin, A. A. Kindiroglu, O. Aran, and L. Akarun, "Automatic personality prediction from audiovisual data using random forest regression," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Cancun, Mexico, 2016, pp. 37–42.
[18] H. J. Escalante *et al.* (2018). "Explaining first impressions: Modeling, recognizing, and explaining apparent personality from videos." [Online]. Available: https://arxiv.org/abs/1802.00745

[19] M. Mikulincer, P. R. Shaver, M. L. Cooper, and R. J. Larsen, *APA Handbook of Personality and Social Psychology. Personality Processes and Individual Differences* (APA Handbooks in Psychology). Washington, DC, USA: APA, 2015.

[20] L. M. Hough and S. Dilchert, "Personality: Its measurement and validity employee selection," in *Handbook of Employee Selection*, J. L. Farr and N. T. Tippins, Eds. New York, NY, USA: Routledge, 2017, pp. 298–325.

[21] D. P. Schmitt, J. Allik, R. R. McCrae, and V. Benet-Martínez, "The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations," *J. Cross-Cultural Psychol.*, vol. 38, no. 2, pp. 173–212, Mar. 2007.

[22] J. B. Walther, "Theories of computer-mediated communication and interpersonal relations," in *The Handbook of Interpersonal Communication*, M. L. Knapp and J. A. Daly, Eds. Thousand Oaks, CA, USA: Sage, 2011, pp. 443–479.

[23] S. Nestler and M. D. Back, "Applications and extensions of the lens model to understand interpersonal judgments at zero acquaintance," *Current Directions Psychol. Sci.*, vol. 22, no. 5, pp. 374–379, Sep. 2013.

[24] C. A. Gorman, J. Robinson, and J. S. Gamble, "An investigation into the validity of asynchronous Web-based video employment-interview ratings," *Consulting Psychol. J.*, vol. 70, no. 2, pp. 129–146, Jun. 2018.

[25] L. Batrinca, B. Lepri, N. Mana, and F. Pianesi, "Multimodal recognition of personality traits in human-computer collaborative tasks," in *Proc. 14th ACM Int. Conf. Multimodal Interact.*, Santa Monica, CA, USA, 2012, pp. 39–46.

[26] L. M. Batrinca, N. Mana, B. Lepri, F. Pianesi, and N. Sebe, "Please, tell me about yourself: Automatic personality assessment using short self-presentations," in *Proc. 13th Int. Conf. Multimodal Interfaces*, Alicante, Spain, 2011, pp. 255–262.

[27] A. V. Ivanov, G. Riccardi, A. J. Sporka, and J. Franc, "Recognition of personality traits from human spoken conversations," in *Proc. 12th Annu. Conf. Int. Speech Commun. Assoc. (INTERSPEECH)*, Florence, Italy, 2011, pp. 1549–1552.

[28] Y. Güçlütürk *et al.*, "Multimodal first impression analysis with deep residual networks," *IEEE Trans. Affective Comput.*, vol. 9, no. 3, pp. 316–329, Jul./Sep. 2018.

[29] L. Chen, F. Zaromb, Z. Yang, C. W. Leong, and M. Martin-Raugh, "Can a machine pass a situational judgment test measuring personality perception?" in *Proc. 7th Int. Conf. Affect. Comput. Intell. Interact. Workshops Demos (ACIIW)*, San Antonio, TX, USA, 2017, pp. 7–11.

[30] L. Chen, R. Zhao, C. W. Leong, B. Lehman, G. Feng, and M. E. Hoque, "Automated video interview judgment on a large-sized corpus collected online," in *Proc. 7th Int. Conf. Affective Comput. Intell. Interact. (ACII)*, San Antonio, TX, USA, 2017, pp. 504–509.

[31] L. Chen *et al.*, "Automatic scoring of monologue video interviews using multimodal cues," in *Proc. 17th INTERSPEECH Annu. Conf. Int. Speech Commun. Assoc.*, Princeton, NJ, USA, 2016, pp. 32–36.

[32] L. S. Nguyen and D. Gatica-Perez, "Hirability in the wild: Analysis of online conversational video resumes," *IEEE Trans. Multimedia*, vol. 18, no. 7, pp. 1422–1437, Jul. 2016.

[33] A. T. Rupasinghe, N. L. Gunawardena, S. Shujan, and D. A. S. Atukorale, "Scaling personality traits of interviewees in an online job interview by vocal spectrum and facial cue analysis," in *Proc. 16th Int. Conf. Adv. ICT Emerg. Regions (ICTer)*, Negombo, Sri Lanka, 2016, pp. 288–295.

[34] M. McRorie, I. Sneddon, G. McKeown, E. Bevacqua, E. de Sevin, and C. Pelachaud, "Evaluation of four designed virtual agent personalities," *IEEE Trans. Affective Comput.*, vol. 3, no. 3, pp. 311–322, Jul./Sep. 2012.

[35] S. K. Ötting and G. W. Maier, "The importance of procedural justice in human–machine interactions: Intelligent systems as new decision agents in organizations," *Comput. Hum. Behav.*, vol. 89, pp. 27–39, Dec. 2018.

[36] L. P. Naumann, S. Vazire, P. J. Rentfrow, and S. D. Gosling, "Personality judgments based on physical appearance," *Personality Social Psychol. Bull.*, vol. 35, no. 12, pp. 1661–1671, Sep. 2009.

[37] J. C. S. J. Junior *et al.* (2018). "First impressions: A survey on vision-based apparent personality trait analysis." [Online]. Available: https://arxiv.org/abs/1804.08046

[38] R. Petrican, A. Todorov, and C. Grady, "Personality at face value: Facial appearance predicts self and other personality judgments among strangers and spouses," *J. Nonverbal Behav.*, vol. 38, no. 2, pp. 259–277, Jan. 2014.

[39] A. Basu, A. Dasgupta, A. Thyagharajan, A. Routray, R. Guha, and P. Mitra, "A portable personality recognizer based on affective state classification using spectral fusion of features," *IEEE Trans. Affective Comput.*, vol. 9, no. 3, pp. 330–342, Jul./Sep. 2018.

[40] S. Escalera, X. Baró, I. Guyon, and H. J. Escalante, "Guest editorial: Apparent personality analysis," *IEEE Trans. Affective Comput.*, vol. 9, no. 3, pp. 299–302, Jul./Sep. 2018.

[41] G. Pons and D. Masip, "Supervised committee of convolutional neural networks in automated facial expression analysis," *IEEE Trans. Affective Comput.*, vol. 9, no. 3, pp. 343–350, Jul./Sep. 2018.

[42] S. Rasipuram, S. B. P. Rao, and D. B. Jayagopi, "Asynchronous video interviews vs. face-to-face interviews for communication skill measurement: A systematic study," in *Proc. 18th ACM Int. Conf. Multimodal Interact.*, Tokyo, Japan, 2016, pp. 370–377.

[43] L. R. Goldberg, "The development of markers for the big-five factor structure," *Psychol. Assessment*, vol. 4, no. 1, pp. 26–42, Mar. 1992.

[44] B. W. Swider, M. R. Barrick, and T. B. Harris, "Initial impressions: What they are, what they are not, and how they influence structured interview outcomes," *J. Appl. Psychol.*, vol. 101, no. 5, pp. 625–638, May 2016.

[45] S. Yang and B. Bhanu, "Understanding discrete facial expressions in video using an emotion avatar image," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 42, no. 4, pp. 980–992, Aug. 2012.

[46] K. Bregar and M. Mohorčič, "Improving indoor localization using convolutional neural networks on computationally restricted devices," *IEEE Access*, vol. 6, pp. 17429–17441, 2018.

[47] L. Zheng, L. R. Goldberg, Y. Zheng, Y. Zhao, Y. Tang, and L. Liu, "Reliability and concurrent validation of the IPIP big-five factor markers in China: Consistencies in factor structure between Internet-obtained heterosexual and homosexual samples," *Personality Individual Differences*, vol. 45, no. 7, pp. 649–654, Nov. 2008.

[48] D. Xue *et al.*, "Personality recognition on social media with label distribution learning," *IEEE Access*, vol. 5, pp. 13478–13488, 2017.

[49] M. M. Tadesse, H. Lin, B. Xu, and L. Yang, "Personality predictions based on user behavior on the facebook social media platform," *IEEE Access*, vol. 6, pp. 61959–61969, 2018.

HUNG-YUE SUEN received the Ph.D. degree in management information systems. He is currently an Assistant Professor with the Technology Application and Human Resource Development Department, National Taiwan Normal University. His main research interests include social computing, human–computer interaction, data analytics, and artificial intelligence in human resources development and management.

KUO-EN HUNG received the master's degree in computer science from National Tsing Hua University, Taiwan. He is currently with the Department of Technology Application and Human Resource Development, National Taiwan Normal University, where he leads the Industry–Academia Cooperation Project. His research interests include computer vision, machine learning, pattern recognition, machine learning, human–computer interaction, and facial expression and emotion understanding.

CHIEN-LIANG LIN is currently pursuing the Ph.D. degree with the Department of Management Information Systems, National Cheng-Chi University, Taipei City, Taiwan. His research interests include applied deep learning, information technology adoption, electronic commerce, virtual communities, and e-learning. He has published numerous research papers in these fields. He was a recipient of the Highly Commended Award for Information Technology and People, in 2015.