Research on Personality Prediction Technology Based on Self-Introduction Videos

Authors: Wen Yeye, Chen Deyuan, Li Baobin, Wang Xiaoyang, Xiaoqian Liu, Zhu Tingshao, Liu Xiaoqian

Date: 2020-03-08T00:00:00+00:00

Abstract

Personality influences individuals' work and lifestyle patterns, holding significant guiding value for psychological counseling, career development, and related domains. Traditional methods of assessing personality through questionnaires suffer from issues such as non-response and careless responding. In recent years, the development of machine learning has provided novel approaches for personality recognition. This study employs participants'self-introduction videos and Big Five Inventory scores, proceeding through keypoint extraction, feature dimensionality reduction, modeling, and iterative parameter tuning to obtain distinct predictive models for different personality dimensions. Experimental results demonstrate that the personality prediction model based on self-introduction videos achieves near-moderate or moderate correlations across all dimensions, enabling non-intrusive automatic personality identification and offering a new paradigm for personality assessment.

Full Text

Preamble

Research on Personality Prediction Technology Based on Self-Introduction Videos

Yeye Wen[1,2], Deyuan Chen[1], Baobin Li[3], Xiaoyang Wang[2], , Xiaoqian Liu[2*], Tingshao Zhu[2]

[1](School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China)
[2](Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China)
[3](School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China)

(Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract

Personality influences an individual' s work and lifestyle, providing important guidance for psychological counseling and career development. Traditional methods that evaluate personality through self-report questionnaires suffer from issues such as participant refusal and careless responding. Recent advances in machine learning have offered new approaches for personality recognition. This study utilizes participants' self-introduction videos and Big Five personality inventory scores, employing keypoint extraction, feature dimensionality reduction, modeling, and iterative parameter tuning to develop distinct prediction models for each personality dimension. The results demonstrate that personality prediction models based on self-introduction videos achieve near-moderate to moderate correlations across all dimensions, enabling non-intrusive automatic personality recognition and offering a novel approach to personality assessment.

Keywords: Self-Introduction, Big Five Personality, Machine Learning, Personality Prediction

1 Introduction

Personality represents the unified, stable, and enduring psychological characteristics of an individual that determine both explicit and implicit behaviors [1]. It influences how people perceive their personal circumstances and how they behave in specific situations, thereby shaping lifestyle, psychological states, and social roles to a considerable extent. Costa et al. [2] found that extraversion and neuroticism are primary determinants of subjective well-being through their study of 1,100 participants. Research by Liu Yuxin et al. [3] demonstrated that personality may affect college student stress both directly, by influencing individual perception, and indirectly, by affecting the probability of stressful events. Seibert et al. investigated the relationship between personality and career achievement, revealing that proactive personality positively correlates with innovation, political knowledge, and career initiative, which in turn relate positively to career development and satisfaction [4]. Additionally, the Big Five traits of extraversion, neuroticism, agreeableness, and openness show varying relationships with career satisfaction and salary levels [5].

Personality assessment finds applications in career planning, job candidate screening, and mental health evaluations for university admissions. Currently, personality questionnaires constitute the primary measurement tool. However, because respondents complete these instruments themselves, they may provide false answers in high-stakes contexts such as admissions or job applications to improve their chances [6]. Since individuals find it more difficult to disguise their nonverbal behaviors such as speech patterns and facial expressions, researchers have increasingly explored automatic personality recognition models using nonverbal cues following the rapid development of machine learning. These approaches have leveraged features from speech [7, 8], handwriting [9], and online behavior [10]. Given that stable personality traits manifest more prominently in physical and facial cues than in vocal characteristics [11], researchers have sought to predict personality from facial activity, aiming for more convenient, ecologically valid applications and improved predictive performance. Facial activity-based prediction of psychological indicators has been successfully applied to emotion recognition [12], depression detection [13], and autism diagnosis [14]. Modeling approaches typically involve either extracting statistical features from facial keypoint coordinates identified using open-source libraries such as OpenPose and OpenCV, then feeding these into traditional machine learning models, or directly inputting video frames into neural networks. While numerous studies have employed question-and-answer or speech videos to build personality prediction models [15, 16], these video materials constrain participant behavior, thereby reducing the ecological validity of the models. Self-introduction videos, compared to other video types, maintain relatively neutral emotional content, which minimizes facial activity caused by emotional states [17] and reduces variations in facial emotional expression attributable to personality differences [18], thus eliminating some confounding variables.

This study employs participants' self-introduction videos as experimental material, using OpenPose to record facial keypoint coordinates during self-introductions. Statistical features of coordinate variations serve as input features, with Big Five inventory scores as labels. Support vector machine models are developed separately for each of the five personality dimensions, aiming to provide a novel method for personality measurement.

2.1 Participants

We recruited 240 participants from a university, including undergraduate students, graduate students, and faculty members. The sample comprised 110 males (45.8%) and 130 females (54.1%). In terms of education level: 70 participants (29.1%) held bachelor' s degrees or below, 155 (64.6%) were master' s students, and 15 (6.3%) held master' s degrees or above. The mean age was 23 $\pm$ 3 years.

2.2 Research Tools

(1) Big Five Inventory (BFI): Developed by John et al. in 1991 [19], this instrument measures the five dimensions of the Big Five personality model: extraversion, agreeableness, conscientiousness, neuroticism, and openness. The scale contains 44 items and is therefore abbreviated as BFI-44. Each item consists of a phrase containing one or two words that best describe a specific Big Five dimension. Participants respond using a five-point Likert scale, where 1 represents "strongly disagree" and 5 represents "strongly agree." The internal consistency reliability of each BFI-44 subscale is at or above 0.8, with test-retest reliability exceeding 0.8 after three months.

(2) OpenPose Real-Time System: An open-source project developed by Carnegie Mellon University [20] that enables keypoint detection for body torso, face, hands, and feet. OpenPose employs identical training methods for facial and hand keypoint detection [20], utilizing a multi-camera system to build the detection framework and clustering faces from the bottom up after identifying keypoints [21]. This study used the OpenPose system to record real-time coordinates of 70 facial keypoints, as illustrated in [Figure 2: see original paper]-1, with coordinate variations representing facial activity as input for the prediction models.

2.3 Data Collection Procedure

After participants completed personal information registration and provided informed consent, they were seated before a high-definition camera and instructed to deliver a self-introduction in Mandarin for no less than three minutes. Participants could be provided with a speech outline containing three questions: (1) Please introduce yourself and describe your hometown in detail; (2) Please elaborate on your major and research work during your studies; (3) Please describe your future plans and what kind of work you would like to pursue.

Following the self-introduction, participants completed the BFI-44 personality inventory. They were informed that the inventory, video recordings, and personal information would be used solely for research purposes, and confidentiality was assured.

3.1 Data Preprocessing

After data cleaning, we obtained self-introduction videos of equal length for each participant. Using the OpenPose open-source system, we extracted facial keypoint coordinates frame-by-frame from the videos. We then translated all keypoints to a two-dimensional coordinate system with keypoint 0 as the origin to balance differences in spatial positioning between participants and the camera. Frame-to-frame differences in each keypoint' s coordinates were calculated to replace the original coordinate data.

Subsequently, we applied equal-depth binning with mean values to smooth and denoise inter-frame differences, thereby reducing the impact of outliers and increasing granularity, with bin depth set as parameter a. After smoothing, we extracted both time-domain and frequency-domain features from coordinate variations for each facial keypoint. Time-domain features included 30 dimensions: dimensional features such as maximum, minimum, mean, standard deviation, and root mean square, as well as dimensionless features including skewness, kurtosis, and impulse factor. Frequency-domain features were obtained by applying Fourier transform to convert signals to the frequency domain, then selecting the amplitudes of the first five low-frequency components, yielding five dimensions. Following feature extraction for each keypoint, we obtained 4,900 features (70 keypoints × 2 coordinates × 35 features), which comprehensively captured temporal and frequency characteristics of facial activity.

After feature extraction, we employed both data standardization and normalization to balance the influence of keypoint distance from the coordinate origin, with the scaling method set as parameter *b*. Principal Component Analysis (PCA) was then applied for feature dimensionality reduction, with the number of dimensions after reduction set as parameter *c*. The data preprocessing workflow is illustrated in [Figure 3: see original paper]-1(a).

3.2 Model Construction

To develop predictive models for participants' Big Five personality traits, we employed regression rather than classification models. Compared to linear regression, Support Vector Regression (SVR), proposed by Drucker et al. in 1997 [22], maps nonlinear functions to higher-dimensional feature spaces through nonlinear transformation, obtaining solutions to the original nonlinear functions by solving linear functions in high-dimensional space. This approach is well-suited for nonlinear, high-dimensional problems. Given the high dimensionality of facial activity features and the nonlinear relationship between individual features and outcomes, we selected SVR to build separate prediction models for each of the five personality dimensions. The modeling workflow is shown in [Figure 3: see original paper]-1(b).

SVR models can be categorized into four types based on kernel function: rbf-SVR, poly-SVR, sigmoid-SVR, and Linear-SVR. Since the specific mapping relationship from low-dimensional to high-dimensional space cannot be determined a priori, we set the SVR type as parameter *d* to enable subsequent selection of appropriate kernel functions for each Big Five dimension based on cross-validation results. Three hyperparameters related to the kernel function —penalty coefficient, kernel coefficient, and maximum degree of the kernel function—were set as parameters *e*, *f*, and *g*, respectively. The parameter value ranges are shown in -1.

3.3 Evaluation Methods and Results

This study used Pearson correlation coefficients between participants' Big Five inventory scores and model predictions as the performance metric, employing cross-validation to mitigate the impact of training-test set partitioning. The Pearson correlation coefficient is a commonly used metric for evaluating regression models, with values ranging from -1 to 1; larger absolute values indicate stronger correlations.

After extensive experimentation, we identified optimal parameter configurations for each Big Five dimension prediction model. The selected optimal parameters are presented in -2. In addition to correlation coefficients, we conducted significance testing. Using five-fold cross-validation with 46 degrees of freedom, we obtained corresponding t-values and p-values. The performance evaluation of each Big Five dimension prediction model is shown in -3. The results indicate that prediction models for extraversion, agreeableness, conscientiousness, and neuroticism demonstrate significant positive correlations at the $p < 0.01$ level, while the openness model shows significant positive correlation at the $p < 0.05$ level.

4 Discussion

This study developed models based on participants' self-introduction facial videos to predict scores on the five Big Five dimensions: extraversion, agreeableness, conscientiousness, openness, and neuroticism, with separate models constructed for each dimension. Seven parameters were considered: bin depth for equal-depth binning, data standardization vs. normalization, number of features after PCA dimensionality reduction, SVR kernel function selection, SVM penalty coefficient, kernel coefficient (parameter gamma), and maximum kernel degree (parameter degree). The first three parameters relate to preprocessing, while the latter four pertain to model training.

The models output predicted scores for each personality dimension. Five-fold cross-validation was employed for training and evaluation given the sample size, with Pearson correlation between inventory scores and predicted scores serving as the evaluation metric. During initial modeling, parameter selection for different personality dimensions was coarse. Using preliminary models for prediction, we iteratively adjusted parameters based on average correlation coefficients from cross-validation, controlling variables and refining parameters within the ranges specified in -1 until stable models were obtained and parameters were finalized.

Compared to traditional questionnaire methods, personality prediction through video data modeling offers distinct advantages. It eliminates constraints from objective factors such as venue, time, and participant numbers, while avoiding subjective limitations like respondent aversion to questions or dishonest responding. This approach also substantially reduces resource expenditure, enabling large-scale, high-frequency personality testing.

The video data used in this study consisted of self-introductions, with effective modeling data derived from variations in facial keypoint coordinates during speech rather than speech content itself. Self-introductions were selected because participants generally maintain neutral, stable emotions during this task, minimizing interference from emotions elicited by dialogue, questioning, or contemplative scenarios, thereby enhancing prediction accuracy.

Optimal model parameters were iteratively selected for each personality dimension, with -2 showing the distinct parameters for each model. Developing five separate models for the five personality dimensions offers significant advantages over using a single model for all traits. Objectively, distinct differences exist among personality dimensions, necessitating trait-specific analysis. Subjectively, the relative prominence of different personality dimensions varies across individuals, requiring dimensional analysis of the same video data.

The personality prediction models output dimensional scores rather than categorical classifications, employing quantitative descriptions that provide more precise personality assessment with finer granularity. Additionally, since BFI-44 results are also presented as scores, this facilitates comparison, evaluation, and validation against the traditional method.

5 Conclusion and Outlook

Based on self-introduction video data, this study developed separate models for each personality dimension to predict personality scores, using correlation between inventory scores and predicted scores as the evaluation metric. We ultimately obtained optimal prediction models for all five personality dimensions. The correlation coefficients indicate moderate or better correlations for all dimensions except openness, which showed weaker correlation likely due to smaller sample size. The model predictions demonstrate high accuracy compared to inventory results, confirming that personality prediction models based on self-introduction videos are feasible and effective, offering a new approach to personality measurement.

Video data in this study were collected at a designated location. To further overcome constraints of time, venue, and resources, participants could be instructed to record self-introduction videos using their own devices (e.g., smartphones, computers) and submit them to researchers. Although this may increase preprocessing workload, it would substantially expand sample size and potentially improve model accuracy and ecological validity. Furthermore, the model could be deployed on a web platform, enabling general users to upload personal video data to the cloud for rapid personality assessment, achieving self-testing effects comparable to traditional inventories.

References

[1] 黄希庭, 再谈人格研究的中国化. 西南师范大学学报：人文社会科学版, 2004. 30(6).
[2] Costa, P.T., Jr. and R.R. McCrae, Influence of extraversion and neuroticism on subjective well-being: happy and unhappy people. Journal of personality and social psychology, 1980. 38(4): p. 668-78.
[3] 刘玉新, 社会支持与人格对大学生压力的影响. 心理学报, 2005. 37(1).
[4] Seibert, S.E., M.L. Kraimer, and J.M. Crant, What do proactive people do? A longitudinal model linking proactive personality and career success. Personnel Psychology, 2001. 54(4): p.
[5] Seibert, S.E. and M.L. Kraimer, The five-factor model of personality and career success. Journal of Vocational Behavior, 2001. 58(1): p. 1-21.
[6] McLarnon, M.J.W., A.C. DeLongchamp, and T.J. Schneider, Faking it! Individual differences in types and degrees of faking behavior. Personality and

Individual Differences, 2019. 138: p. 88-95.
[7] Fallahnezhad, M., M. Vali, and M. Khalili. Automatic Personality Recognition from reading text speech. in 2017 Iranian Conference on Electrical Engineering (ICEE). 2017.
[8] Solera-Urena, R., et al., Acoustic-Prosodic Automatic Personality Trait Assessment for Adults and Children, in Advances in Speech and Language Technologies for Iberian Languages, Iberspeech 2016, A. Abad, et al., Editors. 2016. p. 192-201.
[9] Fallah, B. and H. Khotanlou. Identify human personality parameters based on handwriting using neural network. in 2016 Artificial Intelligence and Robotics (IRANOPEN). 2016.
[10] Bai, S., et al. Predicting Big Five Personality Traits of Microblog Users. in 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT). 2013.
[11] Harrigan, J.A., K. Wilson, and R. Rosenthal, Detecting state and trait anxiety from auditory and visual cues: A meta-analysis. Personality and Social Psychology Bulletin, 2004. 30(1): p. 56-66.
[12] Rao, K.P., M.V.P.C.S. Rao, and N.H. Chowdary, An integrated approach to emotion recognition and gender classification. Journal of Visual Communication and Image Representation, 2019. 60: p. 339-345.
[13] Wang, Q., H. Yang, and Y. Yu, Facial expression video analysis for depression detection in Chinese patients. Journal of Visual Communication and Image Representation, 2018. 57: p.
[14] Jiang, M., et al., Classifying Individuals with ASD Through Facial Emotion Recognition and Eye-Tracking. Conference proceedings : ···Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2019. 2019: p. 6063-6068.
[15] Bekhouche, S.E., et al. Personality Traits and Job Candidate Screening via Analyzing Facial Videos. in 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2017.
[16] Weninger, F., et al. Speaker trait characterization in web videos: Uniting speech, language, and facial features. in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. 2013.
[17] Darwin, C., Expression of the emotions in man and animals. 2003. 123(1): p. 146.
[18] Mergl, R., et al., Facial expressions and personality: A kinematical investigation during an emotion induction experiment. Neuropsychobiology, 2006. 54(2): p. 114-119.
[19] John, O.P., E.M. Donahue, and R.L. Kentle, The big five inventory–versons 4a and 54. 1991, Berkeley: University of California, Berkeley, Institute of Personality and Social Research.
[20] Cao, Z., et al., OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields.
[21] Simon, T., et al. Hand Keypoint Detection in Single Images using Multiview Bootstrapping. arXiv e-prints, 2017.

[22] Drucker, H., et al., Support vector regression machines, in Advances in Neural Information Processing Systems 9: Proceedings of the 1996 Conference, M.C. Mozer, M.I. Jordan, and T. Petsche, Editors. 1997. p. 155-161.

(Corresponding author: Xiaoqian Liu, E-mail: liuxiaoqian@psych.ac.cn)