International Journal of Automation and Computing 14(4), August 2017, 386-395
DOI: 10.1007/s11633-017-1085-8

Physiognomy: Personality Traits Prediction by Learning

Ting Zhang[1,3] Ri-Zhen Qin[1] Qiu-Lei Dong[1,2,3] Wei Gao[1] Hua-Rong Xu[4] Zhan-Yi Hu[1,2,3]

[1]National Laboratory of Pattern Recognition, Institute of Automation,
Chinese Academy of Sciences, Beijing 100190, China
[2]Center for Excellence in Brain Science and Intelligence Technology,
Chinese Academy of Sciences, Beijing 100190, China
[3]University of Chinese Academy of Sciences, Beijing 100049, China
[4]Department of Computer Science & Technology, Xiamen Institute of Technology, Xiamen 361024, China

Abstract: Evaluating individuals′ personality traits and intelligence from their faces plays a crucial role in interpersonal relationship and important social events such as elections and court sentences. To assess the possible correlations between personality traits (also measured intelligence) and face images, we first construct a dataset consisting of face photographs, personality measurements, and intelligence measurements. Then, we build an end-to-end convolutional neural network for prediction of personality traits and intelligence to investigate whether self-reported personality traits and intelligence can be predicted reliably from a face image. To our knowledge, it is the first work where deep learning is applied to this problem. Experimental results show the following three points: 1) "Rule-consciousness" and "Tension" can be reliably predicted from face images. 2) It is difficult, if not impossible, to predict intelligence from face images, a finding in accord with previous studies. 3) Convolutional neural network (CNN) features outperform traditional handcrafted features in predicting traits.

Keywords: Personality traits, physiognomy, face image, deep learning, convolutional neural network (CNN).

1 Introduction

Since the ancient Chinese, Egyptian and Greek times, people have tried to establish relationships between facial morphological features and individual personality traits[1], which is known as physiognomy. People tend to evaluate others on their appearances and then interact with them based on their first impressions. Currently, psychological studies[2] have showed that faces play a central role in people′s everyday assessments of other people. Humans perform trait judgments from faces unconsciously, which may have a great influence on the results of important social events such as elections[3, 4] and court sentences[5]. The works in [6, 7] show that at least four personality traits (agreeableness, conscientiousness, extraversion and dominance) can be inferred reliably by humans from facial features.

To investigate whether the trait evaluations performed by humans can be learned automatically by computers, Rojas et al.[8, 9] constructed an automatic trait predictor based on facial structural and appearance descriptors. They found that all the analyzed personality traits could be predicted accurately.

Wolffhechel et al.[10] studied the relationship between self-reported personality traits and first impressions, and a normative self-reported questionnaire (cubiks in-depth personality questionnaire, CIPQ 2.0) was used to measure participants′ personality traits. The results showed that some personality traits could be predicted from faces to a certain extent. Kleisner et al.[11] used face photographs of 40 men and 40 women to investigate the relationship between measured intelligence and facial features, and little correlation was found.

In this work, we further investigate whether self-reported personality traits and measured intelligence can be predicted from face images by gathering more samples and by applying deep learning. In the previous works on predicting personality traits and intelligence, handcrafted textural descriptors were used to extract facial features from face images and typical machine learning methods were used to train the models. Instead, here we train an end-to-end convolutional neural network (CNN) to predict personality traits and intelligence where the facial features are automatically learned from the face image. We design a multi-task neural network to predict all the personality traits and intelligence jointly. Inspired by the fact that the performance of a CNN can be improved by transferring learning[12, 13]. The visual geometry group (VGG) Face[14], which performs well on the labeled faces in the wild (LFW)[15] dataset, is utilized as part of our proposed network.

Note that over the past twenty years, there has been huge progress in face recognition. Deep learning has achieved great success on face recognition[14, 16−22] and significantly

Research Article
Manuscript received September 19, 2016; Accepted February 27, 2017
This work was supported by National Natural Science Foundation of China (Nos. 61333015, 61421004 and 61375042) and Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDB02070002).
Recommended by Associate Editor Jangmyung Lee

outperformed the existing systems using low-level features[23−32]. However, so far as we know, there is no such a work that employs CNNs to predict personality traits and intelligence from the human face.

Our experimental results show that some personality traits can be predicted from face images reliably and may depend largely on genetic qualities, while some others may rely on the social environment largely, no reliable prediction is possible by face images for measured intelligence, no evident linear correlation between the predicted scores and the measured scores of the personality traits and intelligence is found in the results of regression experiments.

In summary, our work makes two major contributions. Firstly, we build an end-to-end neural network to extract features and implement classification or regression to predict personality traits and measured intelligence from the human face for the first time. Secondly, we construct a dataset for East-Asian race to investigate the correlations between personality traits, measured intelligence and face images. The dataset consists of face photographs, personality measurements and intelligence measurements.

2 Dataset

We construct a dataset named as physiognomy dataset to investigate the correlations between personality traits (also measured intelligence) and face images. It consists of face photographs, personality measurements and intelligence measurements. Our dataset is designed for the East-Asian race, different from the existing works[33] which target Caucasian race. Face photographs of 186 people (94 men and 92 women) are included in the dataset. The participants were photographed with neutral expression when sitting in front of a white background.

Ethics statement. This research was approved by the Institutional Review Board of the Institute of Automation of the Chinese Academy of Sciences. The participants were asked to give verbal consent to participate in the research and all data were collected after this consent was obtained. The consent is thereby documented by the recording of the data. The guidelines of the Ethics Committee, Ministry of Health of the People′s Republic of China state that written consent is only required if biological samples are collected, which was not the case in this study. In addition, the data for the self-reported personality traits and the measured intelligence were analyzed anonymously.

Personality measurements. Cattell sixteen personality factors (16PF)[34], a normative self-reported questionnaire that scored 16 personality traits, was used to measure participants′ personality traits in our work. The traits measured by 16PF are warmth, reasoning, emotional stability, dominance, liveliness, rule-consciousness, social boldness, sensitivity, vigilance, abstractedness, privateness, apprehension, openness to change, self-reliance, perfectionism and tension. Assessing the responses on the questionnaire by the commercial software developed by Beijing Normal University Education Training Center, a discrete score ranging from 1 to 10 was obtained for each personality trait. Since the subjects of this Chinese questionnaire are the students of China, it is suitable for testing personality traits of Chinese people. Based on these sixteen personality factors, professor Cattell performed a second-order factor analysis and acquired the following four second-order factors: adaptation or anxiety, introversion or extroversion, impetuous action or undisturbed intellect, cowardice or resolution. The above 20 traits are used to describe the personality of each participant in our experiments. Fig. 1 shows an example of the 20 personality trait scores for one participant.

Intelligence measurements. To measure participants′ intelligence, each subject was instructed to fill out a Raven′s standard progressive matrices questionnaire (SPM)[35] which was compiled by the British Psychologist Raven in 1938. The participants′ observational ability and ability to think clearly were primarily measured in this test. To measure the intelligence level of the participant, the total score for the right answers was calculated and converted to a percentile score. This intelligence metric comprising of 60 questions is divided into five groups, A, B, C, D and E, with 12 questions in each group. Fig. 2 shows some examples of these questions. The difficulty of problem increases gradually from the first group to the last group. The internal problems within each group are also arranged sequentially by difficulty, from easy to difficult. The thought process required to complete the questions differs in each group. In our experiments, 186 participants completed the test and the percentile scores were used to indicate their intelligence levels.

Data preprocessing. First, we use the method described by Qin and Zhang[36] to detect the facial landmarks (including two pupil locations) from each face image. Then, based on the coordinates of these two points, a similarity transformation (a rotation + rescaling in our work) is performed for all the images in the dataset such that the two transformed pupils are horizontally positioned with a fixed distance. An original image and its transformed image are shown in Figs. 3 (a) and (b), respectively. In addition, for the sake of removing the redundant background information in the photographed images, we select the image region to make sure the eyes lie horizontally at the same height and leave a standard length of neck visible.

In order to remove irrelevant information such as the background, hair and clothing, researchers usually crop the images in facial analysis. In this work, we also cropped the sample images using the following steps in Fig. 4. First, the two pupils are connected by a line segment AB, then a downwardly perpendicular segment MC is drawn where MC $= \frac{1}{2}$ AB and M is the midpoint of AB. Finally, the cropped region is a square with each side = 2AB and centered at C. Fig. 3 (c) shows an example of a cropped image.

Remarks. Due to the expensive costs on collecting and labeling samples, Physiognomy dataset is not quite large. However, our dataset is larger than other datasets dealing

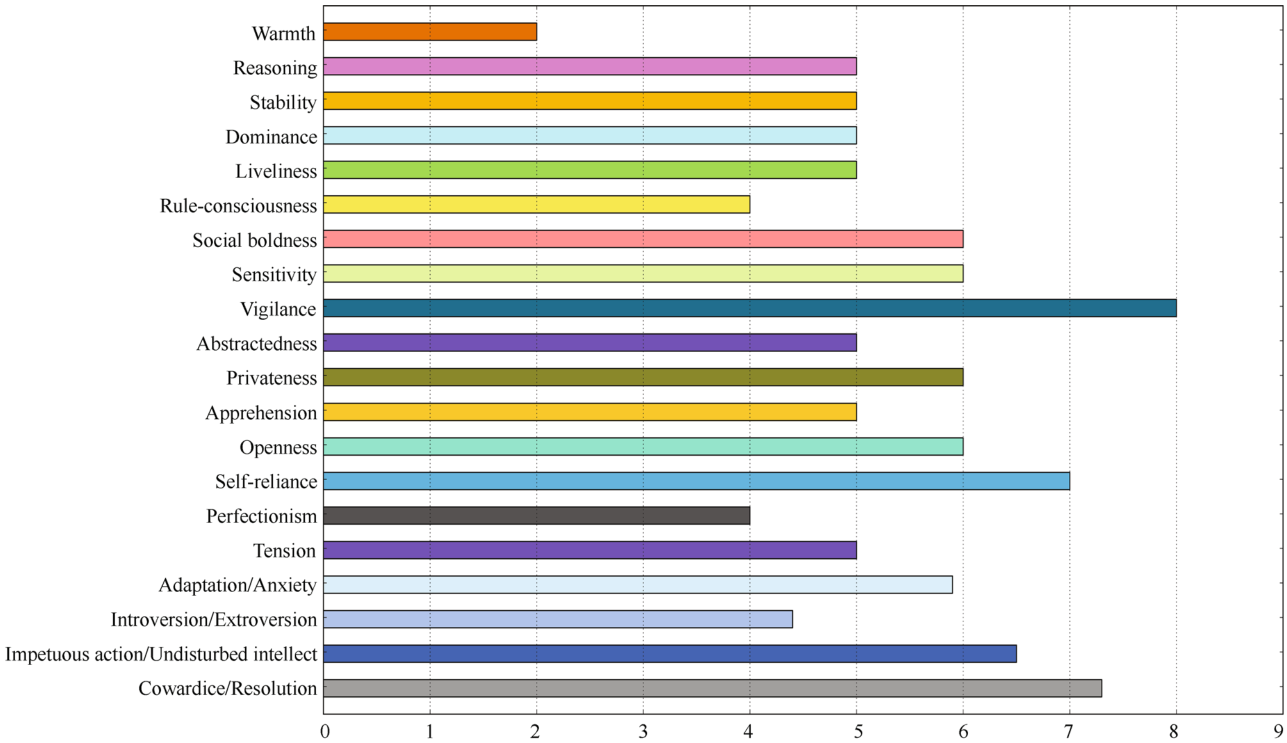


Fig. 1 Scores of 20 personality traits for one participant. Lengths of bars represent the scores of traits. The higher the score is, the more salient the trait is. (Color versions of the figures in this paper are available online at http://link.springer.com.)

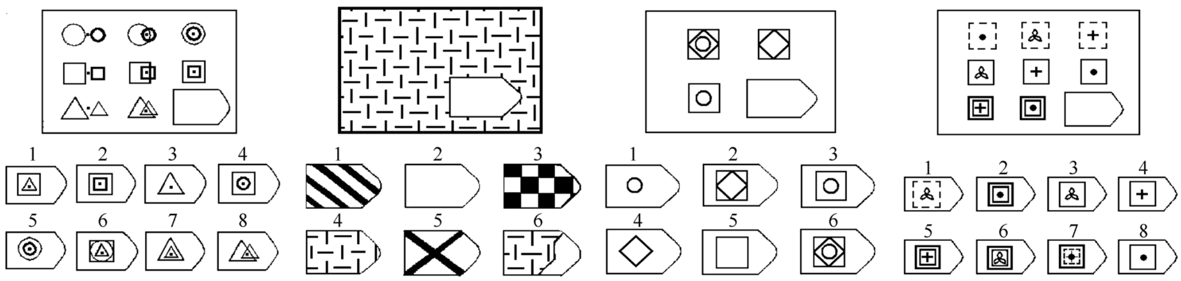


Fig. 2 Examples of Raven′s standard progressive matrices questionnaire

(a) Original image (b) Image after the similaritytransformation (c) Cropped image

Fig. 3 Results at different preprocessing stages with the same problem. For the sake of protection of privacy of the participants in our experiments, physiognomy dataset is not publicly released.

3 Personality traits prediction via CNNs

3.1 Overview

In this paper, trait prediction is cast as a classification problem and a regression problem respectively. To construct the classification labels and the regression targets, the personality traits scores and the measured intelligence scores are converted into appropriate values. We propose a CNN to tackle both problems. The inputs of the proposed CNN are facial images, while the corresponding traits and intelligence are used as labels in the training phase.

For the classification problem, we aim to classify each of the referred 21 traits into two categories. Because the score on each personality trait ranges from 1 to 10, we binarize the personality traits by setting the highest 5 as the "have-trait" class and the lowest 5 as the "do-not-have-trait" class. Similarly, to balance the number of samples in the two categories for the measured intelligence, we set the percentile scores less than or equal to 75% to one category and the scores greater than 75% to another category. The accuracy of classification is used to evaluate the classification model. For the regression problem, we directly set the score of each personality trait and the percentile value of the intelligence as the regression targets. The root mean square error (RMSE) is used to evaluate the regression model.

3.2 Model description

As shown in Fig. 5, we propose a convolutional neural network to explore whether self-reported personality traits and intelligence can be predicted reliably from a face image.

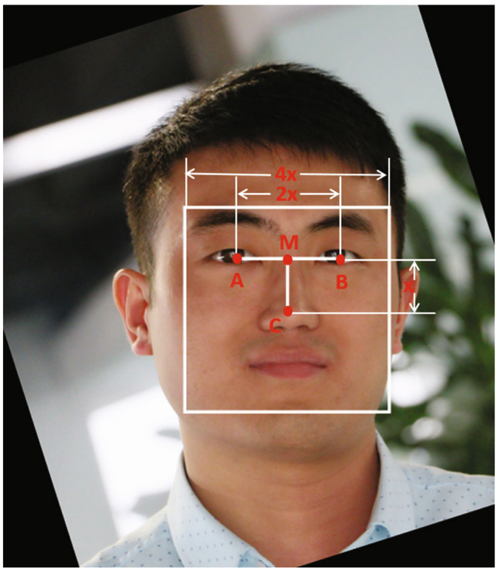


Fig. 4 An illustration of the cropping proceduce for an image

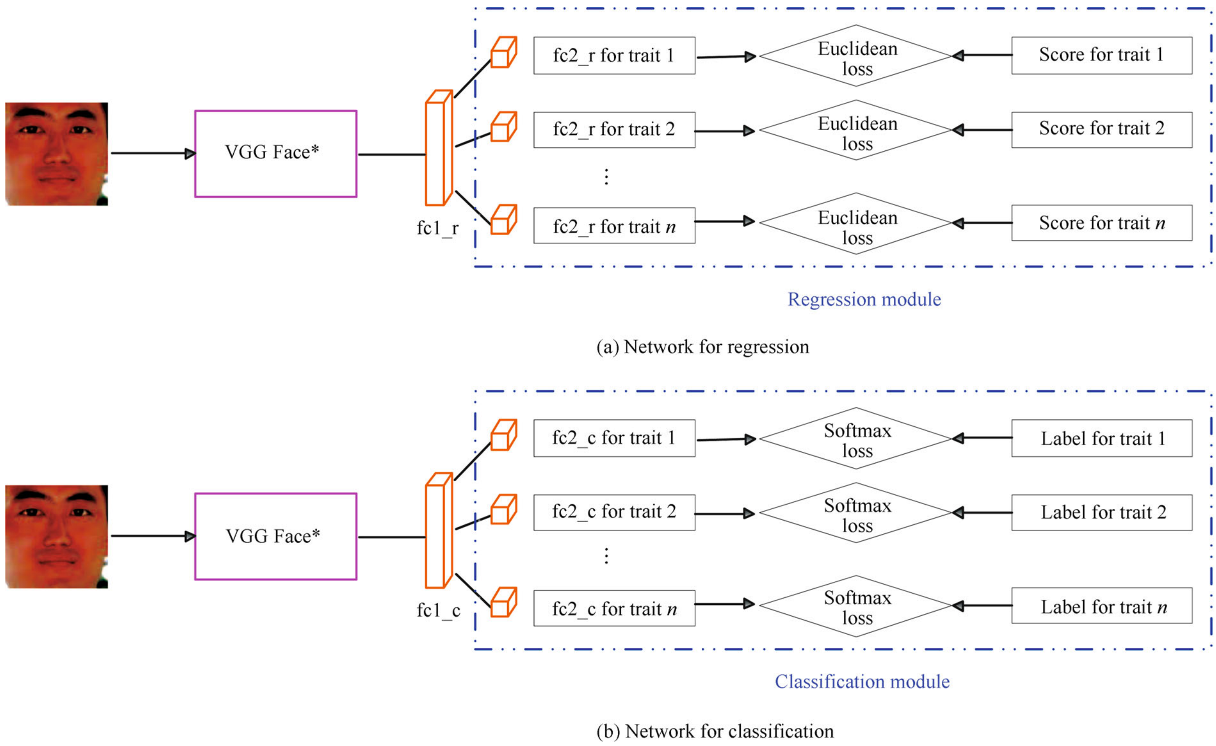


Fig. 5 Architecture of the proposed network. VGG Face* represents the network of VGG Face omitting the last fully-connected layer

Inspired by the fact that the performance of a CNN can be improved by transferring learning[12, 13], the VGG Face[14], which performs well on the LFW[15] dataset, is utilized as part of our proposed network.

We design a multi-task neural network to predict all the personality traits and intelligence jointly. The input of the network is a $224 \times 224 \times 3$ color face image. To adjust the network fitting for our goal, we append several fully-connected layers to the left part of the network. When training for regression, the network in Fig. 5 (a) is adopted. For classification, the network in Fig. 5 (b) is used.

In the classification task, all the layers except the last one are shared by all the traits, while trait-dependent layers (one fully-connected layer and one softmax layer for each trait k, $k = 1, \cdots, 21$) are stacked over them. During training, the loss for trait k will only propagate to its corresponding top fully-connected layer and lower shared layers. More specifically, there are 21 fully-connected layers for 21 traits and each fully-connected layer contains two neurons. Because each trait is classified into a binary category, every pair of neurons corresponds to one trait. Each fully-connected layer is fed to a two-way softmax which produces a distribution over the two class labels. The softmax loss is used to optimize the network.

In the regression task, the network is similar to that for classification. The lower layers are shared by all the traits and each fully-connected layer contains one neuron. The output of each neuron is the predicted score for one personality trait or intelligence. Euclidean loss is used to measure the distance between the predicted values and measured ones.

Note that training for the classification task is independent of that for the regression task and the experiment for each task is carried out independently. Our method is implemented with Caffe[37], which is one of the most popular deep learning frameworks. We train our models using stochastic gradient descent with momentum of 0.9 and weight decay of 0.004. The weights of fully-connected layers are initialized from a zero-mean Gaussian distribution with standard deviation 0.01. The learning rate is initialized with different values according to different tasks.

Network architecture using traditional features as input. We also design a network architecture using traditional features as input to conduct comparative experiments. Some handcrafted features have achieved good performance in face identification and verification. Five descriptors are used to represent facial features. Four are local descriptors (histogram of oriented gradients (HOG)[38], local binary patterns (LBP)[39], Gabor[40], and scale-invariant feature transform (SIFT)[41]), and one is a global descriptor (GIST[42]). The concatenation of the above five descriptors is used as our final feature. As shown in Fig. 6, the feature is used as the input of the designed network. A four-layer neural network is designed and the classification module and regression module are similar to those in Fig. 5.

4 Experimental results

In this section, we conduct experiments using Physiognomy dataset and discuss the results. Face images are aligned according to the position of pupils. The input to the network shown in Fig. 5 is a fixed-size $224 \times 224 \times 3$ face image, and the average face image computed on the training dataset is subtracted in advance.

To investigate whether the self-reported personality traits and measured intelligence can be evaluated from the facial features accurately, we conduct classification and regression experiments respectively. It is noted that Physiognomy dataset is not quite large due to the expensive costs on gaining information of personality traits and intelligence, hence, an N-fold cross-validation scheme is used to estimate the proposed method. The dataset is randomly divided into 10 mutually disjoint subsets. Nine subsets are used for training while the remaining subset is for testing. In order to reduce overfitting, we use label-preserving transformations (transformation of RGB channels and disturbance) to artificially enlarge the dataset 20 times. The experiments with and without data augmentation are all conducted, and the results are found to be close to each other. Only the results with data augmentation are reported here.

We also conduct some comparative experiments to validate the effectiveness of the proposed model using the network architecture shown in Fig. 6. The concatenation of the five descriptors is used as the input of the network.

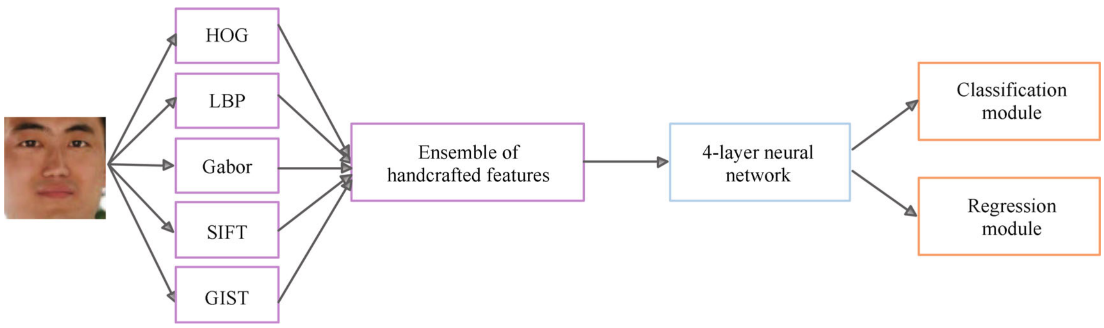


Fig. 6 Network architecture using traditional features as input

4.1 Classification of traits and intelligence

Table 1 shows the classification results of all the personality traits and measured intelligence using different features. For results using CNN, the accuracy scores for "Rule-consciousness" and "Vigilance" far surpass chance levels. The predicted accuracy of "Rule-consciousness" is higher than 82 % and the accuracy of "Vigilance" is higher than 77 %. The high predicted accuracy scores suggest that these two personality traits may closely correlate with the facial characteristics.

For measured intelligence, the classification accuracy slightly exceeds the level of chance. Considering the near-chance level predictions on measured intelligence, we may conclude that predicting measured intelligence from face images is difficult, if not impossible.

Psychological researchers have conducted experiments among twins and found that approximately 50 % of a human′s personality traits are influenced by genetics. Some personality traits depend largely on genetic qualities, while some others mainly depend on the social environment. Meanwhile, biological studies demonstrate that humans′ facial characteristics are determined largely by gene. As a result, the traits more dependent on genetic factors may closely correlate with facial features and can be predicted from the face images more accurately.

For results using traditional features, principal component analysis (PCA) is used to conduct the dimension reduction and the results of both the original features and reduced ones are shown in the Table 1. Fig. 7 shows the classification results of the above three types of features for all the traits and intelligence.

Table 1 and Fig. 7 show that the original traditional features and reduced ones perform comparably. However, the accuracies of classification based on CNN features exceed that based on traditional features. It shows that the neural network taking face images directly as input outperforms the network taking traditional features as input.

4.2 Regression of traits and intelligence

In the regression experiments, we directly set the score of each personality trait and the percentile value of the intelligence as the regression targets.

The RMSE in (1) is used to evaluate the performance of regression.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{n}} \tag{1}$$

where X_i is the self-reported score, Y_i is the predicted score and n is the number of samples.

Table 2 is the performance of regression with respect to all the personality traits and measured intelligence using different features. The results using CNN show that the errors of "Rule-consciousness", "Openness", "Perfectionism" and "Tension" are smaller than the errors of other personality traits, while "Social boldness", "Vigilance" and "Introverted or Extroverted" are larger. This indicates that

Table 1 Mean accuracy of the classification results for the 20 traits and intelligence. "1" indicates the results using CNN, "2" indicates the results using traditional features, and "3" indicates the results using the reduced traditional features by PCA. Best results are written in bold

	Warm	Reas	Stab	Domin	Live	Cons	Soci	Sens	Vigil	Abst	Intell
1	**55.62**	**57.87**	**67.98**	**64.61**	**73.03**	**82.02**	**55.06**	**53.93**	**77.53**	**53.37**	**57.30**
2	52.81	55.62	64.04	58.99	65.17	79.08	46.57	44.94	72.84	47.75	52.81
3	50.99	51.69	56.74	59.55	67.42	79.21	49.44	48.31	71.91	45.51	52.25
	Priv	Appr	Open	Reli	Perf	Tens	Adap	Intro	Impet	Cowa	
1	**59.55**	**61.80**	**57.30**	**58.99**	**60.11**	**60.67**	**58.43**	**69.10**	**55.06**	**54.49**	
2	56.18	55.06	52.81	53.93	53.37	56.18	55.05	63.48	48.31	52.25	
3	57.87	55.62	49.44	51.69	54.92	50.56	56.18	64.04	48.88	49.44	

Table 2 RMSE of the regression results for the 20 traits and intelligence. "1" indicates the results using CNN, "2" indicates the results using traditional features, and "3" indicates the results using the reduced traditional features by PCA. Best results are written in bold

	Warm	Reas	Stab	Domin	Live	Cons	Soci	Sens	Vigil	Abst	Intell
1	**1.991 4**	**1.524 8**	**1.802 9**	**1.568 6**	**1.898 6**	**1.439 2**	**2.037 1**	**1.496 5**	**2.114 1**	**1.670 1**	**0.194 4**
2	2.183 1	1.592 7	1.881 5	1.710 1	2.049 2	1.561 3	2.151 9	1.621 9	2.265 2	1.764 7	0.208 2
3	2.121 3	1.801 0	2.017 1	1.778 2	2.075 5	1.611 5	2.360 9	1.703 7	2.345 2	1.952 6	0.343 2
	Priv	Appr	Open	Reli	Perf	Tens	Adap	Intro	Impet	Cowa	
1	**1.547 0**	**1.858 8**	**1.361 0**	**1.591 1**	**1.368 1**	**1.404 2**	**1.732 9**	**2.117 5**	**1.622 7**	**1.533 3**	
2	1.675 6	1.966 0	1.438 2	1.623 5	1.445 4	1.503 4	1.839 3	2.219 7	1.766 7	1.679 7	
3	1.714 9	1.994 3	1.554 4	1.888 5	1.554 1	1.604 5	1.861 3	2.322 0	1.946 0	1.786 5	

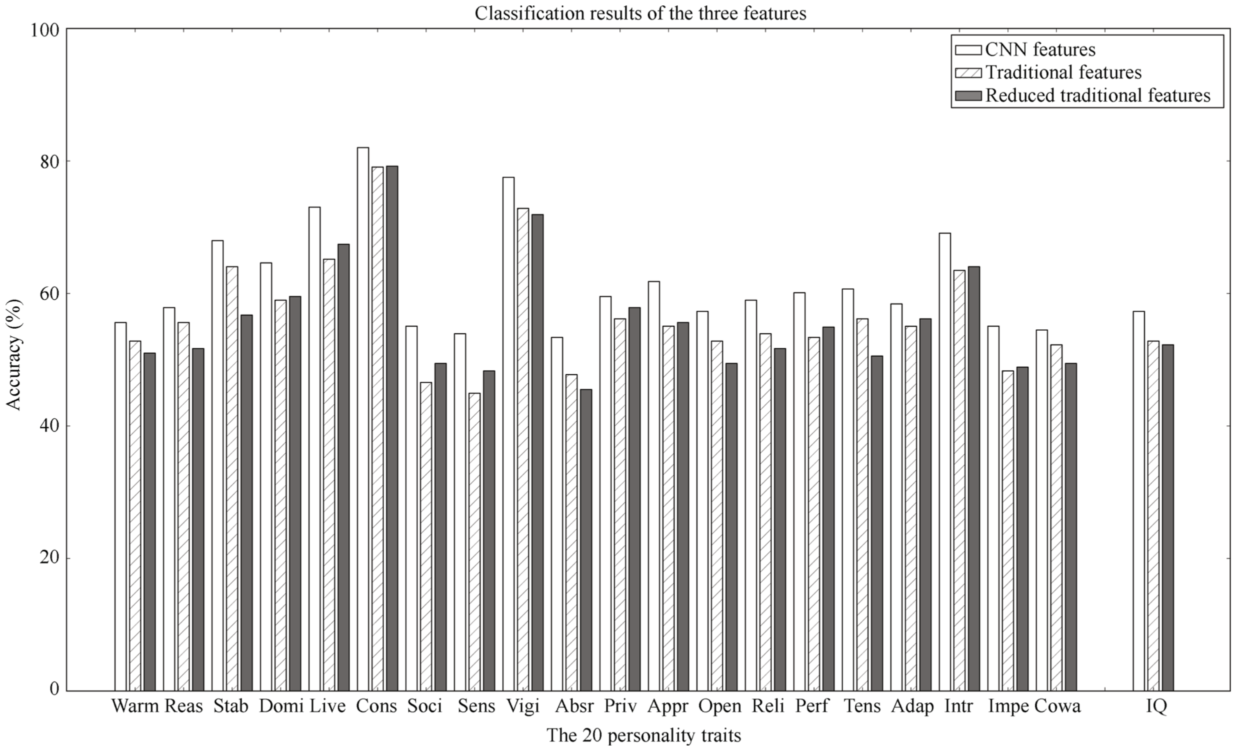


Fig. 7 Classification results by using the CNN features and traditional features for all the personality traits as well as measured intelligence

little relationship exists between the latter three personality traits and face images. The fitting error of intelligence is somewhat high, which indicates that it is difficult to predict a person′s intelligence score from the face image accurately.

People generally do not favor a clear-cut choice, e.g., people often remark that he or she is a bit controlling or somewhat sensitive. Because physiognomy dataset is not quite large, and most scores for the personality traits are around median, the regression models trained on our dataset are more suitable for the prediction of the median scores.

Fig. 8 shows the regression results of the above three types of features for all the personality traits and measured intelligence. Table 2 and Fig. 8 show that the results of original traditional features and reduced ones are comparable. However, the errors of regression model based on CNN features are smaller than those based on traditional features. It shows that features directly extracted from the face images using CNN outperforms the traditional handcrafted features.

4.3 Application of predicting personality traits from internet images

In the above sections, we used CNN to investigate whether personality traits and intelligence could be predicted from the human face. The regression and classification experimental results show that, certain personality traits could be predicted with high reliability. Therefore, we build a new dataset to further assess our experimental observations. As shown in Fig. 9, the dataset contains two groups of frontal face images with neutral expression collected from the Internet. The categories of the two groups are entertainers and teachers respectively. In order to investigate whether people with similar social behaviors share similar personality traits, the proposed model is used to conduct the experiment. Physiognomy dataset is used to train the network, and the newly constructed dataset is to test.

The experimental results show that samples in the same category share several similar personality traits. The prediction results of entertainers show that they get high scores in "Reasoning" and "Liveliness", while low scores in "Emotional Stability". They usually learn fast and react quickly, which verifies the high scores in "Reasoning". The results in "Emotional stability" and "Liveliness" indicate that they are emotionally less stable, animated and enthusiastic, which exactly reflect the personality traits of the entertainers in the dataset. Teachers get high scores in "Liveliness" and "Extrovert" in the prediction results. They are lively, expressive and good at communicating with students, which verifies the results.

The above analyses show that there exists some kind of correlation between the predicted personality traits and the real characters. Especially for people in the same group, the proposed model seems able to verify their general characters.

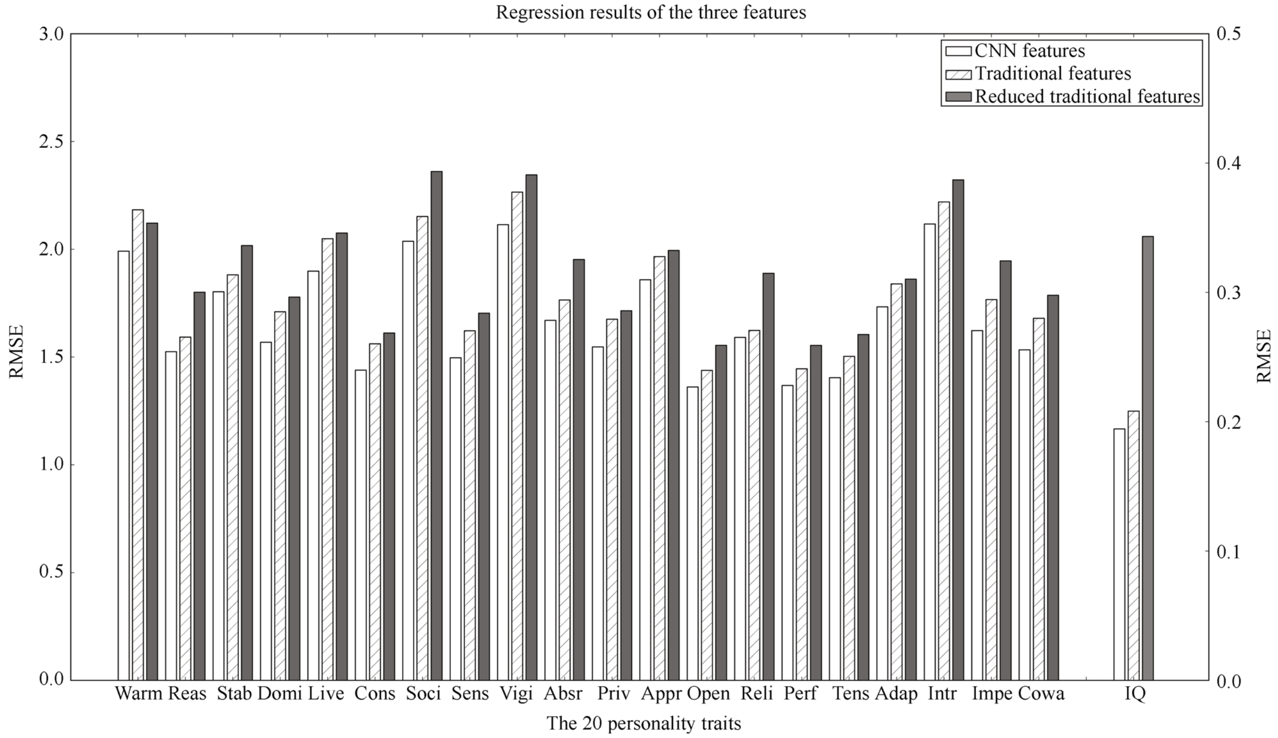


Fig. 8 Regression results by using the CNN features and traditional features for all the personality traits as well as measured intelligence

Fig. 9 Two groups of face images. From top to down: entertainers and teachers.

5 Conclusions

In this paper, we investigate whether self-reported personality traits and measured intelligence could be predicted from face images. An end-to-end convolutional neural network is proposed to deal with the task. Both the classification and regression experiments are conducted to evaluate the prediction. The results show that certain personality traits could be predicted from face images rather reliably but it is difficult to predict intelligence from face images. Comparative experiments show that features extracted from face images using CNN perform better than traditional handcrafted features in predicting traits. In the future, we will combine 3D facial features and deep learning for traits prediction. In addition, we would also enlarge our database to include more diverse groups of people rather currently only of college students.

Acknowledgement

We are grateful to all the students of Xiamen Institute of Technology in China who participated in this study.

References

[1] D. McNeill. *The Face: A Natural History*, New York, USA: Back Bay Books, 2000.

[2] A. Todorov, C. P. Said, A. D. Engell, N. N. Oosterhof. Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, vol. 12, no. 12, pp. 455–460, 2008.

[3] C. C. Ballew II, A. Todorov. Predicting political elections from rapid and unreflective face judgments. *Proceedings of*

the National Academy of Sciences of the United States of America, vol. 104, no. 46, pp. 17948–17953, 2007.

[4] A. C. Little, R. P. Burriss, B. C. Jones, S. C. Roberts. Facial appearance affects voting decisions. *Evolution and Human Behavior*, vol. 28, no. 1, pp. 18–27, 2007.

[5] I. V. Blair, C. M. Judd, K. M. Chapleau. The influence of afrocentric facial features in criminal sentencing. *Psychological Science*, vol. 15, no. 10, pp. 674–679, 2004.

[6] D. R. Carney, C. R. Colvin, J. A. Hall. A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality*, vol. 41, no. 5, pp. 1054–1072, 2007.

[7] R. S. S. Kramer, J. E. King, R. Ward. Identifying personality from the static, nonexpressive face in humans and chimpanzees: Evidence of a shared system for signaling personality. *Evolution and Human Behavior*, vol. 32, no. 3, pp. 179–185, 2011.

[8] Q. M. Rojas, D. Masip, A. Todorov, J. Vitriä. Automatic point-based facial trait judgments evaluation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, San Francisco, USA, pp. 2715–2720, 2010.

[9] Q. M. Rojas, D. Masip, A. Todorov, J. Vitria. Automatic prediction of facial trait judgments: Appearance vs. structural models. *PLoS One*, vol. 6, no. 8, Article number e23323, 2011.

[10] K. Wolffhechel, J. Fagertun, U. P. Jacobsen, W. Majewski, A. S. Hemmingsen, C. L. Larsen, S. K. Lorentzen, H. Jarmer. Interpretation of appearance: The effect of facial features on first impressions and personality. *PLoS One*, vol. 9, no. 9, Article number e107721, 2014.

[11] K. Kleisner, V. Chvátalová, J. Flegr. Perceived intelligence is associated with measured intelligence in men but not women. *PLoS One*, vol. 9, no. 3, Article number e81237, 2014.

[12] J. Yosinski, J. Clune, Y. Bengio, H. Lipson. How transferable are features in deep neural networks? In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, NIPS, Montréal, Canada, pp. 3320–3328, 2014.

[13] M. S. Long, Y. Cao, J. M. Wang, M. I. Jordan. Learning transferable features with deep adaptation networks. In *Proceedings of the 32nd International Conference on Machine Learning*, JMLR, Lille, France, 2015.

[14] O. M. Parkhi, A. Vedaldi, A. Zisserman. Deep face recognition. In *Proceedings of British Machine Vision Conference*, Swansea, UK, vol. 41, pp. 1–12, 2015.

[15] G. B. Huang, M. Ramesh, T. Berg, E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, USA, 2007.

[16] Y. Taigman, M. Yang, M. A. Ranzato, L. Wolf. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Columbus, USA, pp. 1701–1708, 2014.

[17] Y. Sun, X. G. Wang, X. O. Tang. Deep learning face representation from predicting 10,000 classes. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Columbus, USA, pp. 1891–1898, 2014.

[18] Z. Y. Zhu, P. Luo, X. G. Wang, X. O. Tang. Deep learning identity-preserving face space. In *Proceedings of IEEE International Conference on Computer Vision*, IEEE, Sydney, Australia, pp. 113–120, 2013.

[19] Y. Sun, Y. H. Chen, X. G. Wang, X. O. Tang. Deep learning face representation by joint identification-verification. In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, NIPS, Montréal, Canada, pp. 1988–1996, 2014.

[20] Y. Taigman, M. Yang, M. A. Ranzato, L. Wolf. Web-scale training for face identification. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Boston, USA, pp. 2746–2754, 2015.

[21] T. Zhang, Q. L. Dong, Z. Y. Hu. Pursuing face identity from view-specific representation to view-invariant representation. In *Proceedings of IEEE International Conference on Image Processing*, IEEE, Phoenix, USA, pp. 3244–3248, 2016.

[22] B. Zhao, J. S. Feng, X. Wu, S. C. Yan. A survey on deep learning-based fine-grained object classification and semantic segmentation. *International Journal of Automation and Computing*, vol. 14, no. 2, pp. 1–17, 2017.

[23] N. Kumar, A. C. Berg, P. N. Belhumeur, S. K. Nayar. Attribute and simile classifiers for face verification. In *Proceedings of IEEE International Conference on Computer Vision*, IEEE, Kyoto, Japan, pp. 365–372, 2009.

[24] Y. Taigman, L. Wolf, T. Hassner. Multiple one-shots for utilizing class label information. In *Proceedings of British Machine Vision Conference*, London, UK, vol. 2, pp. 1–12, 2009.

[25] M. Guillaumin, J. Verbeek, C. Schmid. Is that you? Metric learning approaches for face identification. In *Proceedings of IEEE International Conference on Computer Vision*, IEEE, Kyoto, Japan, pp. 498–505, 2009.

[26] Q. Yin, X. O. Tang, J. Sun. An associate-predict model for face recognition. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Colorado Springs, USA, pp. 497–504, 2011.

[27] C. Huang, S. H. Zhu, K. Yu. Large scale strongly supervised ensemble metric learning, with applications to face verification and retrieval. arXiv:1212.6094, 2012.

[28] D. Chen, X. D. Cao, L. W. Wang, F. Wen, J. Sun. Bayesian face revisited: A joint formulation. In *Proceedings of the 12th European Conference on Computer Vision*, Florence, Italy, pp. 566–579, 2012.

[29] T. Berg, P. N. Belhumeur. Tom-vs-pete classifiers and identity-preserving alignment for face verification. In *Proceedings of British Machine Vision Conference*, Guildford, UK, vol. 129, pp. 1–11, 2012.

[30] D. Chen, X. D. Cao, F. Wen, J. Sun. Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Portland, USA, pp. 3025–3032, 2013.

[31] X. D. Cao, D. Wipf, F. Wen, G. Q. Duan, J. Sun. A practical transfer learning algorithm for face verification. In *Proceedings of IEEE International Conference on Computer Vision*, IEEE, Barcelona, Spain, pp. 3208–3215, 2013.

[32] F. K. Zaman, A. A. Shafie, Y. M. Mustafah. Robust face recognition against expressions and partial occlusions. *International Journal of Automation and Computing*, vol. 13, no. 4, pp. 319–337, 2016.

[33] N. N. Oosterhof, A. Todorov. The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 32, pp. 11087–11092, 2008.

[34] S. Karson. *A Guide to the Clinical Use of the 16 pf*, Savoy, USA: Institute for Personality and Ability Testing, 1976.

[35] R. M. Kaplan, D. P. Saccuzzo. *Psychological Testing: Principles, Applications, and Issues*, Boston, USA: Wadsworth Publishing, 2012.

[36] R. Z. Qin, T. Zhang. Shape initialization without ground truth for face alignment. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, Shanghai, China, pp. 1278–1282, 2016.

[37] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell. Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of ACM International Conference on Multimedia*, ACM, Orlando, USA, 2014.

[38] N. Dalal, B. Triggs. Histograms of oriented gradients for human detection. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, San Diego, USA, pp. 886–893, 2005.

[39] T. Ahonen, A. Hadid, M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[40] C. J. Liu, H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image Processing*, vol. 11, no. 4, pp. 467–476, 2002.

[41] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[42] A. Oliva, A. Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, vol. 42, no. 3, pp. 145–175, 2001.

Ting Zhang received the B. Sc. degree in communication engineering from Beijing Jiaotong University, China in 2013. She is currently a Ph. D. degree candidate in the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China.

Her research interests include deep learning and face recognition.

E-mail: ting.zhang@nlpr.ia.ac.cn
ORCID iD: 0000-0001-9145-5913

Ri-Zhen Qin received the B. Sc. degree in automation from the Xidian University, China in 2013, received the M. Sc. degree from the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China in 2016.

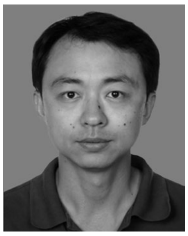
His research interests include machine learning and face recognition.

E-mail: rizhen.qin@nlpr.ia.ac.cn

Qiu-Lei Dong received the B. Sc. degree in automation from the Northeastern University, China in 2003, received the Ph. D. degree from the Institute of Automation, Chinese Academy of Sciences, China in 2008. Currently, he is a professor in the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China.

His research interests include motion analysis, 3D computer vision and pattern classification.

E-mail: qldong@nlpr.ia.ac.cn (Corresponding author)
ORCID iD: 0000-0003-4015-1615



Wei Gao received the B. Sc. degree in computational mathematics, the M. Sc. degree in pattern recognition and intelligent system from Shanxi University, China, and the Ph. D. degree in pattern recognition and intelligent system from Institute of Automation, Chinese Academy of Sciences, China in 2002, 2005 and 2008, respectively. Since July 2008, he has joined Robot Vision Group of National Laboratory of Pattern Recognition where he is currently an associate professor.

His research interests include 3D reconstruction from images and SLAM technology.

E-mail: wgao@nlpr.ia.ac.cn

Hua-Rong Xu received the M. Sc. and the Ph. D. degrees in computer science from Xiamen University, China in 2003 and 2011, respectively. Now he is a professor of Xiamen Institute of Technology, China. He has worked on computer vision and pattern recognition.

His research interests include 3D computer vision and driverless-navigation.

E-mail: hrxu@xmut.edu.cn

Zhan-Yi Hu received the B. Sc. degree in automation from the North China University of Technology, China in 1985, and received the Ph. D. degree in computer vision from the University of Liege, Belgium, in 1993. Since 1993, he has been with the National Laboratory of Pattern Recognition at Institute of Automation, Chinese Academy of Sciences, China. From 1997 to 1998, he was a visiting scientist with the Chinese University of Hong Kong, China. From 2001 to 2005, he was an executive panel member with the National High-Tech Research and Development Program (863 Program). From 2005 to 2010, he was a member of the Advisory Committee, National Natural Science Foundation of China. He is currently a research professor of computer vision, the deputy editor-in-chief of the *Chinese Journal of CAD and CG*, and an associate editor of *Science China*, and *Journal of Computer Science and Technology*. He was the Organization Committee Co-Chair of the ICCV2005, and the Program Co-Chair of the ACCV2012.

His research interests include biology-inspired vision and large scale 3D reconstruction from images.

E-mail: huzy@nlpr.ia.ac.cn