Zero-shot Personality Perception From Facial Images

Peter Zhuowei Gan[(✉)] [iD], Arcot Sowmya [iD], and Gelareh Mohammadi [iD]

University of New South Wales, Sydney, Australia
{zhuowei.gan,a.sowmya,g.mohammadi}@unsw.edu.au

Abstract. Personality perception is an important process that affects our behaviours towards others, with applications across many domains. Automatic personality perception (APP) tools can help create more natural interactions between humans and machines, and better understand human-human interactions. However, collecting personality assessments is a costly and tedious task. This paper presents a new method for zero-shot facial image personality perception tasks. Harnessing the latent psychometric layer of CLIP (Contrastive Language-Image Pre-training), the proposed PsyCLIP is the first zero-shot personality perception model achieving competitive results, compared to state-of-the-art supervised models. With PsyCLIP, we establish the existence of latent psychometric information in CLIP and demonstrate its use in the domain of personality computing. For evaluation, we compiled a new personality dataset consisting of 41800 facial images of various individuals labelled with their corresponding perceived Myers Briggs Type Indicator (MBTI) types. PsyCLIP achieved statistically significant results ($p<0.01$) in predicting all four Myers Briggs dimensions without requiring any training dataset.

Keywords: Personality · Personality perception · Personality computing · Data-driven approach · Computational modeling · Transfer learning

1 Introduction

When seeing a person's face for the first time, we instinctively form an impression of their personality [31]. While we may be taught at a young age to "not judge a book by its cover", psychology studies have shown that this innate first perception yields considerable accuracy [6,21,30,43]. Indeed, face-reading skills are critical in our daily lives. They are employed by salespersons to assess prospective clients, film directors to choose the optimal actor, and even while considering which stranger to ask for directions [36]. Regardless of its accuracy, personality perception influences our behaviour towards others [39].

It is a major research area to determine if a computer may gain this capability. Personality perception [42] is the automatic perception of a subject's personality based on their audio, visual, or other features. Rather than attempting

to recognise an individual's true personality as in personality recognition tasks, automated personality perception seeks to predict the perceived personality or how an individual's personality is perceived by others.

Automatic personality perception (APP) [42] may be a critical first step in developing a more affable conversational agents [34] that are perceived in certain ways. In theoretical psychology, APP can help better understand social interactions and group dynamics [20] In clinical psychology, it can serve as a coaching system to assist socially challenged individuals such as those diagnosed with autism spectrum disorder (ASD) and social anxiety disorder (SAD), in understanding how their behaviours affect others' perception, thus helping them in coping with social norms. Furthermore, the ability to reliably infer personality impressions from facial images enables us to employ it as a discriminator in an adversarial network [15] for creating face images based on certain personality attributes [12].

Currently, we can attain a high degree of object detection accuracy in some domains [22], partially owing to the vast number of training datasets available [46], such as the ImageNet [10] dataset with 14 million images. This is not the case, however, in areas such as personality computing [18]. The state-of-the-art model [19] of facial image personality perception has a training set of 28,230 images. In contrast, even the MNIST dataset [11], which is often used in beginner deep learning courses, has 60,000 images.

The size of datasets used by cutting-edge deep learning models is ever increasing, such as GPT-3 [8] with 410 billion training tokens. However, we have yet to replicate the success of large datasets in personality computing [42]. Creating an annotated dataset of psychometrics (i.e. personality measures) is far more complex and costly, since it is considerably more difficult to label psychological features [18].

This highlights the value and attraction of zero-shot classification in personality perception tasks [42]. The introduction of contrastive language-image pre-training (CLIP) [33] made zero-shot personality perception promising: Trained on 400 million image-caption pairs, CLIP has been shown to grasp abstract classification labels [4]. We hypothesise that there is a potential way for CLIP to comprehend cues meant to elicit personality attributes and hence be used in personality perception tasks, by accessing a latent psychometric layer within the CLIP model.

The goal of our study is to utilise CLIP to build a zero-shot model of personality perception from unlabelled images by harnessing latent psychometric information from the CLIP pre-trained model. PsyCLIP (Psychometric-CLIP) adopt the CLIP's text/image encoder structure. As CLIP was pre-trained on image-caption pairings, we translate each psychometric label into CLIP-style text prompts (i.e., image captions). To find the optimal text-prompt, we first generate a list of candidate prompts using GPT-3's text-davinci-002 text completion engine [1]. We then proceed to eliminate biased prompts that favour a particular personality trait and select the prompt that results in the highest accuracy. To evaluate the performance of PsyCLIP, we have created a large dataset of

41800 facial images, labelled with Myers Briggs [29] personality types. The result from PsyCLIP is encouraging: We achieved statistically significant results ($p < 0.01$) in all personality dimensions, which are comparable to those obtained by the state-of-the-art supervised model [19]. With PsyCLIP, we make the following contributions:

– Establish the existence of a latent psychometric layer in CLIP, and demonstrate how it can be harnessed in the domain of personality computing.
– Provide a new personality dataset consisting of 41800 facial images of various individuals labelled with their corresponding perceived MBTI personality.
– Introduce a novel approach in handling zero-shot personality perception tasks that produces results comparable to those of a state-of-the-art supervised model, without the need for any training sets.

PsyCLIP is significant because it provides a reasonable base model for computational social scientists, potentially capable of perceiving *any* psychological attribute [37]. It may serve as a playground for rapidly testing psychological theories and sparking new psychological discoveries.
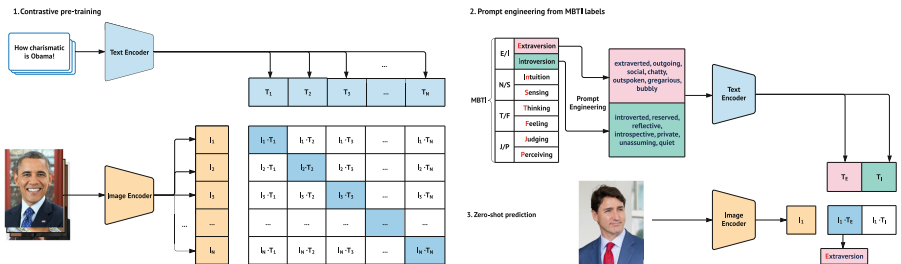

Fig. 1. Summary of our approach. We perform prompt engineering (i.e. translate MBTI subscale traits into CLIP-style text prompt) for each MBTI subscale and encode them using CLIP text encoders. We then assess the classification results of 16000 evaluation samples by encoding them with CLIP image encoders.

2   Related Work

2.1   CLIP

CLIP (Contrastive Language-Picture Pre-training) [33] combines image and text encoding to anticipate appropriate image-text pairing of training instances. Then, for zero-shot object classification, the classification labels are translated to captions such as "a photograph of an extroverted person," and CLIP predicts the caption class that most closely matches the provided photograph. Although CLIP is zero-shot, it outperforms some state-of-the-art supervised models.

However, CLIP has been mostly used for standard object classification tasks, and there is a dearth of research on CLIP's performance with psychological labels. Evidence suggests that CLIP might comprehend abstract prompts, as seen in BigSleep [23] and DeepDaze [24] CLIP functions as a discriminator in these projects, combining with a BigGAN [7] or SirenNetwork [38] to generate abstract artworks from arbitrary inputs.

This leads to the hypothesis that there might be hidden information about personality measures within the pre-trained CLIP model. This work explores the effectiveness of CLIP in personality perception tasks and intends to spark discussion on using pre-trained models in personality computing.

2.2   Personality Measures

Two psychometric instruments stand out among contemporary personality models: the Big Five [35] and the Myers-Briggs Type Indicator (MBTI) [29]. The Big Five (the five-factor usually assessed using NEO Personality Inventory) is more prominent in academia, whereas the MBTI is more prevalent in the consulting and training industries [14]. Big Five model describes each person's personality across five dimensions: Extraversion, Openness, Agreeableness, Conscientiousness, and Neuroticism that are revealed from semantic analysis of personality descriptors. However, MBTI indicates preferences in how people perceive the world and make decisions [29] with four categories: Extraversion-introversion, intuiting-sensing, thinking-feeling, and judging-perceiving. Studies show that there is a strong correlation [14] between the MBTI and the four dimensions of the Big Five, as shown in Table 1: The Big Five Extraversion is highly correlated with the MBTI Extraversion/Introversion (E-I) dimension; the Big Five Openness is highly correlated with the MBTI Intuition/Sensing (N-S) dimension; and the Big Five Agreeableness is only associated with MBTI thinking; The Big Five conscientiousness is associated with both the thinking-feeling (T-F) and judging-perceiving (J-P) dimensions; Neuroticism as measured by the NEO-PI is unrelated to any MBTI subscale score.

The primary distinction between MBTI and Big Five is that MBTI employs a binary classification system (e.g., either extrovert or introvert), whereas Big Five employs a linear scale (e.g., a number associated with each dimension) [9]. As a result, MBTI naturally lends itself to classification tasks, whereas Big Five dimensions lend themselves to regression. Therefore, MBTI was a more natural choice for evaluating PsyCLIP's performance, as CLIP was designed as a classification model.

Another reason we choose MBTI is that the large dataset we gathered was in MBTI. In practice, it is easier to collect big datasets of personality perceptions using MBTI. We posit that APP on dichotomy-based datasets (such as the MBTI) could be a necessary prelude to APP on scaling-based datasets (such as Big Five).

Table 1. The table shows how MBTI correlates to big five spectrums. [14]

MBTI	Big Five
E-I	Extraversion
N-S	Openness
T-F	Agreeableness, Conscientiousness
J-P	Conscientiousness

2.3   Personality Perception

Personality perception [42] is the automatic perception of a subject's personality based on their audio, visual, or other features [26]. Recent personality perception work includes textual personality perception [13,25,44], audio personality perception [27,41,45], visual perception from videos [5,16] and multimodal perception [28].

2.4   Facial Image Perception

In the field of personality perception, there are fewer studies simply on the basis of visual images. This can be ascribed in part to the difficulty inherent in gathering sufficiently large image datasets for personality computation [18, 19] described a supervised model based on ResNet and multi-layer Perceptron. It uses a person's face image to predict their Big Five traits. It was trained using 28,230 face images of 11,202 subjects. Although the connection between predicted and true scores is modest, it can correctly predict the relative standing of two randomly picked persons on a personality dimension in 58 % of situations (as against the 50 % expected by chance).

In the current study, instead of using conventional supervised models, we explores the possibility of Zero-Shot classificaton through large pretrained models like CLIP.

3   Method

To achieve zero-shot personality perception, We first translate MBTI subscale traits into CLIP-style text prompt for each MBTI subscale and encode them using CLIP text encoders. This is the primary distinction between PsyCLIP and CLIP: rather than utilising classification labels directly, as is customary in CLIP, we design psychological classification labels into text prompts that capture the relevant features for each category. We then assess the classification results of 16000 evaluation samples by encoding them with CLIP image encoders. This section describes the dataset we used and all steps to perform a zero-shot personality perception.

3.1   Dataset

Although the proposed method does not require any training, it needs a dataset to evaluate its performance. We have built a dataset from the largest online MBTI database [3]. The website contains 51800 profiles of famous people and characters. Each profile consists of a profile image of size 256×256. The profile has been scored by a number of voters for their perceived personality type. The personality type of each profile is determined by the perceived personality type with the highest vote. In post-processing, we took the top 1000 most voted non-fictional profiles for each 16 MBTI personality types (e.g., INPT type for Introvert, Intuition, Perceiving, Thinking) , resulting in a final sample size of 16000. The minimum number of votes per profile is 6, maximum is 5049, and average is 87.

3.2   CLIP Encoders

As explained earlier CLIP assigns each input image to the encoded text prompt that results in the highest similarity. In PsyCLIP we introduce prompts that are pertinent to each personality type. The prompts are engineered using GPT-3 as explained in the next section. The text and picture encoders in PsyCLIP were built using the ViT32 CLIP model,which has shown to achieve the best performance [33]. We retain the encoders in their current state in order to test CLIP's baseline performance against the MBTI evaluation dataset and to ascertain their potential for discriminating psychological features.

As seen in Fig. 1, we evaluated the effectiveness of PsyCLIP across the four MBTI dimensions. We apply a prompt engineering technique, detailed in the next section, to determine the ideal prompt that best describes each Myers Briggs subscale feature. For instance, we discovered that the prompt that best captures the extroversion attribute is "extraverted, outgoing, sociable, talkative, outspoken, gregarious, effervescent." We next repeated the prompt engineering procedure for each class of the four categories, resulting in a total of eight subscale feature sets.

3.3   Prompt Engineering

Prompt Generation. We produce prompts for each dimension using Generative Pre-trained Transformer 3 (GPT-3) [8]. GPT-3 is the state-of-the-art text generation model, trained on 499 billion tokens. We hypothesise that GPT-3 might assist us in converting psychological labels to CLIP-style instructions. We employed a temperature of 0.7 and the text-davinci-002 text completion model [1], and use the text completeion engine to complete the following: "list a series of adjectives that describes MBTI extroversion." This results in a list of potential candidates that capture the psychological qualities, as shown in Table 2. We then evaluated the performance of the generated prompts against a small test set of 100 samples per personality for prompt selection.

Table 2. Sample prompts generated by GPT-3 text-davinci-002 engine with the input "list a series of adjectives that describes MBTI Extraversion/Introversion/Thinking/ Feeling".

Trait	Generated prompts from GPT3
Extraversion	Extraverted, Outgoing, Social, Chatty, Outspoken, Gregarious, Bubbly
Introversion	Introverted, Reserved, Reflective, Introspective, Private, Unassuming, Quiet
Thinking	Analytical, Logical, Rational, Objective, Introspective, Thoughtful
Feeling	Empathetic, Compassionate, Sympathetic, Cooperative, Caring

Prompt Selection. After generating prompts, we proceed to finding the optimal prompts. We begin by determining the accuracy of each prompt candidate's categorisation in a test pool of 100 randomly chosen candidates for each personality type. The samples are randomly chosen amongst the set of 35800 profiles that are not in the evaluation set. The findings are then utilised for eliminating biased prompts. Biased prompts are prompts that result in skewed results in favour of a certain sub-scale. For instance, if we use the raw GPT-3-generated prompt "analytical, logical, rational, objective, introspective, thoughtful" for the MBTI Thinking type and the raw GPT-3-generated prompt "empathetic, compassionate, sympathetic, cooperative, caring" for the MBTI Feeling type, the result would heavily favour the thinking type. While it is 90.3% accurate in identifying the thinking attribute of an INTP (introverted, intuiting, thinking, perceiving), it is only 20.0 percent accurate in identifying the feeling trait of an INFP (introverted, intuiting, feeling, perceiving). In this case, although the average accuracy of INPs is 55.6% in this scenario, the result cannot be considered statistically significant. As a consequence, we reject prompts that result in skewed outcomes and retain only those that result in above-expectation accuracy in both sub-scales. After rejecting biased prompts, we then find the prompts that would result in highest overall perception accuracy.

4   Results

The prediction accuracy for each MBTI category is shown in Tables 3, 4, 5 and 6. Predictions for each category conditioned on other categories are also reported to help better understand the model behaviour. Overall, PsyCLIP performed above the 50% chance level in all four categories and is statistically significant at $p < 0.01$ on the 16000-person sample size. This corroborates the hypothesis that CLIP do contain latent psychometric information.

4.1   Comparison to Similar Models

In aggregate, the average accuracy is 56.95%. There is a dearth of research on MBTI-based face personality perception with which to make direct comparisons. However, we may still make comparisons to models based on Big Five [19]. In

Table 3. Accuracy score for Thinking/Feeling classification. Result is significant at p < 0.001 against the baseline of 50% (chance level). For example, the first entry means the model has 68.8% accuracy in classifying INTPs as Thinking amongst 1000 INTP samples.

	T(%)	F(%)	Overall(%)
INP	68.8	53.8	57.1
INJ	68.4	46.0	61.3
ENP	66.1	55.1	60.6
ENJ	57.1	54.0	55.6
ISJ	63.8	52.7	58.2
ISP	65.6	47.7	56.7
ESJ	54.3	68.1	61.2
ESP	54.4	60.5	57.2
Overall			58.5

Table 4. Accuracy score for Judging/Perceiving classification. Result is significant at p < 0.001 against the baseline of 50% (chance level). For example, the first entry means the model has 51.6% accuracy in classifying INTJs as Judging amongst 1000 INTJ samples.

	J(%)	P(%)	Overall(%)
INT	51.6	60.9	56.3
ENT	53.9	64.4	59.2
ESF	31.7	72.7	52.0
ISF	43.2	70.9	57.1
EST	49.7	66.5	58.1
IST	54.3	60.3	57.3
ENF	38.0	74.2	56.1
INF	46.4	65.8	56.1
Overall			56.5

58% of situations (as opposed to the 50% anticipated by chance), the supervised model [19] can correctly predict the relative standing of two randomly picked persons on a personality dimension, which could be used as a reference point for comparison. Unfortunately, the dataset used in [19] is not publicly available, making direct comparisons of PsyCLIP's performance difficult. We plan to make the evaluation dataset for PsyCLIP publicly accessible, so that other researchers can test their model and we can make a direct comparison to PsyCLIP performance.

Table 5. Accuracy score for introversion/extraversion classification. Result is significant at p < 0.001 against the baseline of 50% (chance level). For example, the first entry means the model has 55.1% accuracy in classifying INTJs as Introverted amongst 1000 INTJ samples.

	I(%)	E(%)	Overall(%)
NTJ	55.1	44.9	50.0
NTP	54.3	63.3	57.5
SFJ	29.4	82.3	55.3
SFP	42.5	61.4	51.8
STP	35.4	68.3	51.9
STJ	65.6	47.7	55.6
NFJ	54.3	68.1	64.9
NFP	54.4	60.5	67.8
Overall			56.9

Table 6. Accuracy score for Sensing/Intuiting classification. Result is significant at p < 0.001 against the baseline of 50% (chance level). For example, the first entry means the model has 65.8% accuracy in classifying INTJs as Intuitive amongst 1000 INTJ samples.

	N(%)	S(%)	Overall(%)
ITJ	65.8	41.3	53.6
ETJ	60.0	49.1	54.6
ITP	58.8	61.5	60.2
ETP	55.4	59.9	57.7
IFP	49.4	61.6	55.5
EFP	45.1	63.2	54.2
EFJ	57.3	52.5	54.9
IFJ	62.0	50.4	56.2
Overall			55.9

4.2   Comparison Amongst Predictors in Each Dimension

The top performing dimension is the thinking/feeling subscale 2, where Psy-CLIP attained an overall accuracy level of 58%. With a prediction accuracy of 55.9%, the intuiting/sensing dimension is the lowest predictor. Interestingly this is consistent with the supervised model [19], which attained the highest accuracy scores for Big Five conscientiousness and the lowest accuracy score for openness. (For inter-scale relationships, see Table 1.) According to their work [19], the attributes most associated with cooperation (conscientiousness and agreeableness) should be more easily represented in the human face from an evolutionary standpoint. Our results add to the evidence supporting this theory.
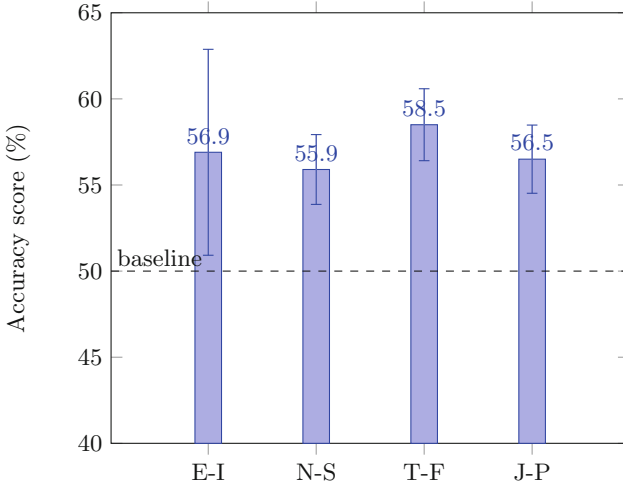
Fig. 2. Percentage accuracy of PsyCLIP with respect to each MBTI dimension. The classification accuracy exceeds the 50% prediction baseline in all personality dimensions.

4.3   Significance

The significance of the result can be interpreted in three ways:

– **It is competitive to SOTA model, without any training set.** PsyCLIP is competitive to the state-of-the-art model out of the box without any fine-tuning, hence it has much potential when datasets are fed into it.
– **It is highly generalisable.** The model is not conditioned on a particular set of psychometric prompts nor designed specifically for MBTI. This means the model has a potential to be a good base model for **any** image-based psychometric classification task.
– **It is statistically significant as a proof of concept.** It proves the existence of a psychometric layer within contrastive language-image pretraining models. We hope it can inspire more affective computing research utilising large pretrained models.

5   Ethical Impact

5.1   Societal Value

On the positive side, automatic personality perception is of significant societal value:

– In affective computing (AC), automatic personality perception is a necessary step in creating a social AI. To be social, an AI must understand how human perceive one another. The perception can be used to create a social avatar, or

to build social conversational agents that are perceived in certain way, among other things.

– In clinical psychology, it can serve as a coaching system to assist socially challenged individuals such as those diagnosed with autism spectrum disorder (ASD) and social anxiety disorder (SAD), in understanding how their behaviours affect others' perception, thus helping them in coping with social norms.

– In theoretical psychology, computational models of such complex perception processes could potentially provide new insights or evidences into psychology theories. For example, as elaborated in Result section, our paper provided a data point to theory that the attributes most associated with cooperation (conscientiousness and agreeableness) should be more easily represented in the human face from an evolutionary standpoint [19]

5.2 Potential Misuses

The abuse of personality computing and its repercussions have been graphically described in several fictions [17] involving a dystopian society in which people are mercilessly evaluated and classified by a computer system.

Beyond fiction, there have been reports [32,40] of HR departments use AI to analyse a candidate's personality based on their web footprint. It would be devastating if PsyCLIP or a similar technology were utilised in this manner to assess a person's personality based on their appearances.

PsyCLIP was not designed for such purposes. One argument is that since PsyCLIP was trained for perception rather than recognition, it is only capable of predicting an applicant's perceived personality and hence has little use for candidate screening.

However, as individual researchers, we have little influence over whether a third party will recognise the delicate distinction between personality detection and apparent personality perception, or how third parties would use such technology.

Does this, however, imply that we should never do research on computer modelling of the human psychological traits? Is this to indicate that AIs are meant to be heartless machines forbidden the knowledge of human emotions, personality, or psychology? One may argue that if research into automated personality perception is halted, we will never be able to build a social AI [2].

We call upon the community to come together and come up with ethical frameworks and regulations on the usage of personality computing technologies, especially in sensitive areas such as recruitment, user profiling and surveillance.

We are also concerned about the potential biases in the dataset. Although by theory [29] all sixteen personalities are equal in value and none are preferable to another, the dataset is labelled by humans, who could be typing a person based on their racial or cultural stereotype. We attempted to mitigate this issue by only evaluating data points that received at least ten votes. Additionally, we would make the dataset and model available to the public upon publication, as we believe that increased transparency and openness are critical in identifying and combating such biases.

6 Conclusion

Based on our experiments, we provide new evidence on the correlation between personality and the facial image. With a sample size of 16000, the findings are statistically significant at $p < 0.01$ and consistently better than the baseline across all four dimensions.

The effectiveness of CLIP in personality perception, along with its zero-shot nature, offers up new possibilities for personality computing applications. It is a complement to conventional supervised models and opens a new direction in study of personality perception phenomenon. With some improvements, computational psychologists now can have a simple model that can be used to predict *any* perceived personality attributes and use it to better understand the semantic associations between words/phrases and personality types.

One area for future study is to investigate how consensus among personality voting influences PsyCLIP's performance. There may be a difference between forecast outcomes for persons with a high personality voting consensus and those with a low voting consensus. Another possibility is to broaden the scope of PsyCLIP's examination beyond MBTI to include additional psychological qualities. Finally, the concept of fine-tuning the PsyCLIP model against a certain personality scale is intriguing and worth exploring.

References

1. Engines - openai api. https://beta.openai.com/docs/engines/gpt-3. Accessed 22 Apr 2022
2. The ethics of artificial intelligence: Issues and initiatives: Think tank: European parliament (2020). https://europarl.europa.eu/thinktank/en/document/EPRS_STU634452
3. Personality database. www.personality-database.com/vote (2022). Accessed 06 Jan 2022
4. Ali, S., Parikh, D.: Telling creative stories using generative visual aids. arXiv preprint arXiv:2110.14810 (2021)
5. Biel, J.I., Teijeiro-Mosquera, L., Gatica-Perez, D.: Facetube: predicting personality from facial expressions of emotion in online conversational video. In: Proceedings of the 14th ACM International Conference on Multimodal Interaction, pp. 53–56 (2012)
6. Borkenau, P., Brecke, S., Möttig, C., Paelecke, M.: Extraversion is accurately perceived after a 50-ms exposure to a face. J. Res. Pers. J. Res. Pers. **43**(4), 703–706 (2009). https://doi.org/10.1016/j.jrp.2009.03.007
7. Brock, A., Donahue, J., Simonyan, K.: Large scale gan training for high fidelity natural image synthesis (2019)
8. Brown, T.B., et al.: Language Models are Few-Shot Learners. arXiv (2020). arxiv.org/2005.14165v4
9. Celli, F., Lepri, B.: Is big five better than mbti? a personality computing challenge using twitter data. In: CLiC-it (2018)
10. Deng, J., et al.: ImageNet: a large-scale hierarchical image database. In: CVPR 2009, pp. 248–255 (2009)

11. Deng, L.: The mnist database of handwritten digit images for machine learning research. IEEE Sign. Process. Mag. **29**(6), 141–142 (2012)
12. Durupinar, F.: Personality-Driven Gaze Animation with Conditional Generative Adversarial Networks. arXiv (2020). arxiv.org/2012.02224v1
13. Farnadi, Get al.: Computational personality recognition in social media. User Model. User-Adapt. Interact. **26**(2–3), 109–142 (2016)
14. Furnham, A.: The of big five versus the big four: the relationship between the myers-briggs type indicator (mbti) and neo-pi five factor model personality. Pers. Individ. Differ **21**(2), 303–307 (1996)
15. Goodfellow, I., et al.: Generative adversarial nets. Adv. Neural Inf. Process. Syst. **27** (2014)
16. Gürpınar, F., Kaya, H., Salah, A.A.: Combining deep facial and ambient features for first impression estimation. In: Hua, G., Jégou, H. (eds.) ECCV 2016. LNCS, vol. 9915, pp. 372–385. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-49409-8_30
17. Hermann, I.: Artificial intelligence in fiction: between narratives and metaphors. AI Soc. 1–11 (2021)
18. Junior, J.C.S.J., et al.: First impressions: a survey on vision-based apparent personality trait analysis. IEEE Trans. Affect. Comput. 1 (2019). https://doi.org/10.1109/TAFFC.2019.2930058
19. Kachur, A., Osin, E., Davydov, D., Shutilov, K., Novokshonov, A.: Assessing the big five personality traits using real-life static facial images. Sci. Rep. **10**(8487), 1–11 (2020). https://doi.org/10.1038/s41598-020-65358-6
20. Kenny, D.A.: Person: a general model of interpersonal perception. Pers. Soc. Psychol. Rev. **8**(3), 265–280 (2004)
21. Kramer, R.S., King, J.E., Ward, R.: Identifying personality from the static, nonexpressive face in humans and chimpanzees: evidence of a shared system for signaling personality. Evol. Hum. Behav. **32**(3), 179–185 (2011)
22. Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., Pietikäinen, M.: Deep learning for generic object detection: a survey. Int. J. Comput. Vis. **128**(2), 261–318 (2019). https://doi.org/10.1007/s11263-019-01247-4
23. lucidrain: Github - lucidrains/big-sleep. http://github.com/lucidrains/big-sleep (2021). Accessed 06 Jan 2022
24. lucidrain: Github - lucidrains/deep-daze. http://github.com/lucidrains/deep-daze (2021). Accessed 06 Jan 2022
25. Majumder, N., Poria, S., Gelbukh, A., Cambria, E.: Deep learning-based document modeling for personality detection from text. IEEE Intell. Syst. **32**(2), 74–79 (2017)
26. Mohammadi, G., Vinciarelli, A.: Automatic personality perception: prediction of trait attribution based on prosodic features. IEEE Trans. Affect. Comput. **3**(3), 273–284 (2012)
27. Mohammadi, G., Vinciarelli, A.: Automatic personality perception: prediction of trait attribution based on prosodic features extended abstract. In: 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 484–490. IEEE (2015)
28. Mohammadi, G., Vuilleumier, P.: A multi-componential approach to emotion recognition and the effect of personality. IEEE Trans. Affect. Comput. (2020)
29. Myers, I.B.: The myers-briggs type indicator: Manual (1962) (1962)
30. Naumann, L.P., Vazire, S., Rentfrow, P.J., Gosling, S.D.: Personality judgments based on physical appearance. Pers. Soc. Psychol. Bull. **35**(12), 1661–1671 (2009). https://doi.org/10.1177/0146167209346309

31. Oosterhof, N.N., Todorov, A.: The functional basis of face evaluation. Proc. Natl. Acad. Sci. **105**(32), 11087–11092 (2008)
32. Patil, S.M., Singh, R., Patil, P., Pathare, N.: Personality prediction using digital footprints. In: 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1736–1742. IEEE (2021)
33. Radford, A., et al.: Learning transferable visual models from natural language supervision. arXiv preprint arXiv:2103.00020 (2021)
34. Ruane, E., Farrell, S., Ventresque, A.: User Perception of Text-Based Chatbot Personality. In: Følstad, A., Araujo, T., Papadopoulos, S., Law, E.L.-C., Luger, E., Goodwin, M., Brandtzaeg, P.B. (eds.) CONVERSATIONS 2020. LNCS, vol. 12604, pp. 32–47. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-68288-0_3
35. Salgado, J.F.: The big five personality dimensions and counterproductive behaviors. Int. J. Select. Assess. **10**(1–2), 117–125 (2002)
36. Shevlin, M., Walker, S., Davies, M., Banyard, P., Lewis, C.A.: Can you judge a book by its cover? evidence of self-stranger agreement on personality at zero acquaintance. Pergamon-Elsevier (2003). http://irep.ntu.ac.uk/id/eprint/16819
37. Sijtsma, K.: Introduction to the measurement of psychological attributes. Measurement **44**(7), 1209–1219 (2011)
38. Ruane, E., Farrell, S., Ventresque, A.: User perception of text-based chatbot personality. In: Følstad, A., et al. (eds.) CONVERSATIONS 2020. LNCS, vol. 12604, pp. 32–47. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-68288-0_3
39. Uleman, J.S., Adil Saribay, S., Gonzalez, C.M.: Spontaneous inferences, implicit impressions, and implicit theories. Annu. Rev. Psychol. **59**, 329–360 (2008)
40. Upadhyay, A.K., Khandelwal, K.: Applying artificial intelligence: implications for recruitment. Strat. HR Rev. (2018)
41. Valente, F., Kim, S., Motlicek, P.: Annotation and recognition of personality traits in spoken conversations from the AMI meetings corpus. In: Thirteenth Annual Conference of the International Speech Communication Association (2012)
42. Vinciarelli, A., Mohammadi, G.: A Survey of Personality Computing. IEEE Trans. Affect. Comput. **5**(3) (2014). DOI: https://doi.org/10.1109/TAFFC.2014.2330816
43. Walker, M., Vetter, T.: Changing the personality of a face: perceived big two and big five personality factors modeled in real photographs. J. Pers. Soc. Psychol. **110**(4), 609–624 (2016). https://doi.org/10.1037/pspp0000064
44. Yu, J., Markov, K.: Deep learning based personality recognition from facebook status updates. In: 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST), pp. 383–387. IEEE, Taichung (2017). https://doi.org/10.1109/ICAwST.2017.8256484, https://ieeexplore.ieee.org/document/8256484/
45. Yu, M., Gilmartin, E., Litman, D.: Identifying personality traits using overlap dynamics in multiparty dialogue. arXiv preprint arXiv:1909.00876 (2019)
46. Zou, Z., Shi, Z., Guo, Y., Ye, J.: Object Detection in 20 Years: A Survey. ResearchGate (2019). www.researchgate.net/publication/333077580