Predicting Personality Traits with Instagram Pictures

Bruce Ferwerda
Department of Computational Perception
Johannes Kepler University
Altenberger Str. 69
A-4040 Linz, Austria
bruce.ferwerda@jku.at

Markus Schedl
Department of Computational Perception
Johannes Kepler University
Altenberger Str. 69
A-4040 Linz, Austria
markus.schedl@jku.at

Marko Tkalcic
Department of Computational Perception
Johannes Kepler University
Altenberger Str. 69
A-4040 Linz, Austria
marko.tkalcic@jku.at

ABSTRACT

Instagram is a popular social networking application, which allows photo-sharing and applying different photo filters to adjust the appearance of a picture. By applying photo filters, users are able to create a style that they want to express to their audience. In this study we tried to infer personality traits from the way users take pictures and apply filters to them. To investigate this relationship, we conducted an online survey where we asked participants to fill in a personality questionnaire, and grant us access to their Instagram account through the Instagram API. Among 113 participants and 22,398 extracted Instagram pictures, we found distinct picture features (e.g., hue, brightness, saturation) that are related to personality traits. Our findings suggest a relationship between personality traits and the way users want to make their pictures look. This allow for new ways to extract personality traits from social media trails, and new ways to facilitate personalized systems.

Categories and Subject Descriptors

H.3.1 [**Information storage and retrieval**]: Content Analysis and Indexing; I.4.7 [**Image processing and computer vision**]: Feature Measurement

Keywords

Instagram, Personality, Photo filters, Picture features

1. INTRODUCTION

Instagram is a popular mobile photo-sharing, and social networking application with currently over 300 million active users. [1] Instagram lets users easily connect with other social networking platforms (e.g., Facebook, Twitter, Tumblr, and Flickr) to share the taken pictures on, and enables users to apply filters to their pictures. At this moment Instagram offers 25 predefined photo filters that soften and color shift picture properties, for users to customize and modify their pictures to create the desired visual style.

The ease with which a photo filter can be applied allow users to express a personal style and create a seeming distinctiveness with the customized pictures. Through the shared content and the way of applying filters, users are able to reveal a lot about themselves to their social network. With that, the question arises: What do Instagram pictures tell about the user? Or more specifically: What do Instagram pictures say about the personality of the user?

Personality traits have shown to consist of cues to infer users' behavior, preference, and taste (e.g., [3, 11, 13]). Hence, knowing one's personality can provide important cues for systems to cater to a personalized user experience. It can provide systems with estimations about user preferences without the use of extensive questionnaires or observations.

There has been an increased interest in how to use personality in systems (e.g., [2, 4, 14]), and how to automatically extract personality from online behavior trails (e.g., Facebook [1, 6, 9, 12], Twitter [5, 10]). In this work we join the personality extraction research by specifically focusing on the relationship between the picture features of an Instagram collection and the personality traits of the user.

Our work makes several contributions. We contribute to personality research by showing relationships between personality traits and the visual style of users' Instagram pictures. Additionally, we contribute to new ways to extract personality from social media (i.e., Instagram). To the best of our knowledge, we are the first to analyze (Instagram) pictures in relation with personality traits.

We conducted an online survey where we asked participants to fill in the widely used, Big Five Inventory (BFI) personality questionnaire, and grant us access to the content of their Instagram account. We extracted 22,398 Instagram pictures of 113 users, and analyzed them on several features (e.g., hue, brightness, saturation). Distinct correlations were found between personality traits and picture features.

In the remainder of the paper we will continue with related work, materials, features, results, discussion, and conclusion.

2. RELATED WORK

There is an increase of psychological research that investigate the relationship between personality and real-world behavior. Personality is known as an enduring factor and has shown to be related to a person's taste, preference, and interest (e.g., [3, 11, 13]). For example, Rawlings and Ciancarelli found relationships between personality traits and music genre preferences [11], while Tkalcic et al. found relation-

[1]https://instagram.com/press/ (accessed: 02/23/2015)

EMPIRE '15, September 16 - 20, 2015, Vienna, Austria

ISBN 978-1-4503-3615-4/15/09…$15.00
DOI: http://dx.doi.org/10.1145/2809643.2809644

ships between personality and classical music [13].

To categorize personality, several models have been developed. The five-factor model (FFM) is the most well known and widely used one. It categorizes personality into five general dimensions that describe personality in terms of: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism [8].

Based on psychological findings of personality in relation to real-world behavior, there is an emergent interest in how to use and implement these findings into applications (e.g., [2, 4, 14]). It can provide useful proxy measures for applications to cater to a more personalized service. For example, Tkalcic et al. proposes to use personality traits to enhance the nearest-neighborhood measurement for overcoming the cold-start problem (i.e., recommending items to new users) in recommender systems [14]. Ferwerda et al. provide a way to use personality traits for adjusting the user interface of music applications to fit user's music browsing styles [4].

Personality traits have shown to consist of valuable information for personalized system, but hard to acquire (e.g., extensive, time-consuming questionnaires). Therefore, as applications are increasingly interconnected (e.g., social media), research started to focus on how to extract personality information from (online) behavior trails. Current research on Facebook (e.g., [1, 6, 9, 12]) and Twitter (e.g., [5, 10]) have shown to consist of reliable cues to infer personality traits from. With our work we add to the personality extraction research by showing the relationship between personality traits and picture features on Instagram.

3. MATERIALS

To investigate the relationship between personality traits and picture features, we asked participants to fill in the 44-item BFI personality questionnaire (5-point Likert scale; Disagree strongly - Agree strongly [7]). The questionnaire include questions that aggregate into the five basic personality traits of the FFM. Additionally, we asked participants to grant us access to their Instagram account through the Instagram API in order to crawl their pictures. From hereon, we define the picture-collection term as all the Instagram pictures of a single user.

We recruited 126 participants through Amazon Mechanical Turk. Participation was restricted to those located in the United States, and also to those with a very good reputation to avoid careless contributions. Several comprehension-testing questions were used to filter out fake and careless entries. The Mahalanobis distance was calculated to check for outliers. This left us with 113 completed and valid responses. Age (18-64, median 30) and gender (54 male, 59 female) information indicated an adequate distribution.

4. FEATURES

For each picture in a picture-collection that was crawled, we extracted several features. The extracted features are discussed below. Most of the features are color-based, some are content-based. For color-based features we use the color space that is most closely related to the human visual system, i.e., the Hue-Saturation-Value (HSV) color space [15].

Brightness. For each picture, we calculated the average brightness and variance across all the pixels in the picture. Pictures that have a high average brightness tend to be bright, obviously. These features represent how light/dark a picture is and how much contrast there is in the picture, respectively. Pictures that have a high variance tend to have both dark and light areas, whereas pictures with a low variance tend to be equally bright across the picture. Furthermore, we divided the brightness axis into three equal intervals and counted the share of pixels that fall into each of these intervals (low/mid/high brightness). Pictures that have a high value in the *low brightness* feature tend to be darker, those that have a high value in the *mid brightness* feature tend to have mostly neither dark nor bright areas, while those pictures that have a high value in the *high brightness* feature tend to have lots of bright areas.

Saturation. We calculated the average saturation and the variance for each picture. Pictures with low average saturation tend to be bleak, colorless, while pictures with high saturation have more vivid colors. Pictures with a high saturation variance tend to have both bleak and vivid colors. Here we also divided the saturation axis into three equal intervals and calculated the share of pixels that fall into each interval (low/mid/high saturation). pictures that have a high value in the *low saturation* tend to have more bleak colors, those with a high value in the *mid saturation* feature tend to have neither bleak nor vivid colors while those pictures that have a high value in the *high saturation* feature tend to have vivid colors across most of the picture area.

Pleasure-Arousal-Dominance (PAD). As the filters on Instagram intend to create a certain expression, we adopted the PAD model of Valdez and Merhabian [16]. They created general rules of the expression of pleasure, arousal, and dominance in a picture as a combination of brightness and saturation levels:

1. Pleasure = .69 Brightness + .22 Saturation
2. Arousal = -.31 Brightness + .60 Saturation
3. Dominance = -.76 Brightness + .32 Saturation

Hue-related features We extracted features that represented the prevalent hues in pictures. We chose many features that represent various aspects of the hues. For each of the basic colors (red, green, blue, yellow, orange, and violet) we counted the share of pixels that fall into each color. As the discrete color clustering of the hue dimension is nonlinear and subjective, we also divided the hue into 10 equal intervals and calculated the share of pixels for each interval. These intervals are hard to describe with subjective color descriptions. Furthermore, we calculated the share of pixels that fall into cold (violet, blue, green) and warm (yellow, red, orange) colors.

Content-based features. Beside the color-centric features we also performed picture content analysis. We counted the number of faces and the number of people in each picture. We used the standard Viola-Jones algorithm [17]. A manual inspection of the Viola-Jones face detector results revealed some false positives (e.g., a portrait within the picture) and false negatives (e.g., some rotated and tilted faces). However, in general the users who tended to take pictures of people (e.g., selfies) had a higher number of average number of faces/people per picture than those users who tended to take mostly still photographs.

5. RESULTS

We crawled 22,398 pictures, and extracted all the features per picture for each picture-collection. As the features in the picture-collections show a symmetrical distribution, we calculated mean values for each feature to create a measurement of central tendency. The mean values of the features were used to calculate the correlation matrix (see Table 1). Pearson's correlation ($r \in [-1,1]$) is reported to indicate the linear relationship between personality and picture features.

The correlation matrix shows several features related with personality traits. We will discuss the results related to each personality trait below. Besides significant correlations of $p<.05$, we decided to report marginally significant results as well (i.e., significant levels of $.1 > p > .05$).

Openness to experience. The openness factor correlates positively with the feature *green*, meaning that open users tend to take pictures with a lot of green or applied a filter to express more greenness. We observed a negative correlation with the feature *brightness mean*, which means that open users tend to upload pictures that are low on brightness. This was further confirmed by the positive correlation on *brightness low*, and the negative correlation on *brightness high*. These correlations show that the pictures of open users show more dark areas, and less bright areas.

Openness is also correlated with the *saturation mean*. This indicate that the pictures of open users consist of more saturated, vivid colors. We also observed a positive correlation with the feature *saturation variance*, which means that open users upload pictures that have both vivid and bleak colors.

A marginally significant correlation was observed for the warm/cold features. Pictures of open users contained less warm colors (i.e., red, orange), but more cold colors (i.e., blue, green). Also, the pictures of open users tend to express less pleasure, but more arousal and dominance. Additionally, their pictures consist of less faces and people.

Conscientiousness. A marginally significant (positive) correlation was found between the *saturation variance* feature and conscientiousness. This indicate that conscientious users upload pictures consisting of bleak and vivid colors.

Extraversion. We mostly found marginally significant correlations with the picture features and the extraversion. Extraverts tend to upload pictures with less *red* and *orange*, but with more *green* and *blue* tones. Additionally, their pictures tend to be darker (*brightness low*), but tend to consist of both vivid and bleak colors (*saturation variance*). Additionally, the emotion that the pictures of extraverts consist, are low on *pleasure*, but high on *dominance*.

Agreeableness. A marginally significant correlation was found between agreeableness and the *brightness mid* feature. This means that the pictures of agreeable users do not show emphasized bright or dark areas, but are more in between.

Neuroticism. A marginally significant correlation was found on *brightness mean*, *brightness low*, and *brightness high*. The positive correlation with *brightness mean* indicate that extravert users tend to upload pictures that are high on brightness. This is also reflected in the *brightness low* (negative correlation) and *brightness high* (positive correlation) features. Additionally, correlations were found in the emotion expression of the pictures of extraverts. They show to adjust their pictures to express more pleasure but less dominance.

	O	C	E	A	N
Red	-0.06	0.02	**-0.17**^	-0.05	0.03
Green	**0.17**^	0.14	**0.23**^^	0.03	-0.12
Blue	-0.01	0	**0.17**^	0.02	-0.01
Yellow	0.01	0.04	0.01	0.14	-0.07
Orange	-0.03	-0.07	**-0.16**^	-0.02	0.06
Violet	0	-0.06	-0.09	-0.07	0.06
Bright.mean	**-0.25***	-0.1	**-0.19**^	-0.07	**0.22**^
Bright.var.	0.06	0	0	-0.07	0.05
Bright.low	**0.28****	0.09	**0.16**^	-0.05	**-0.16**^
Bright.mid	-0.09	0.06	0.04	**0.15**^	-0.06
Bright.high	**-0.2**^	-0.12	**-0.18**^	-0.08	**0.21**^
Sat.mean	**0.16**^	0.06	0.03	-0.04	0
Sat.var.	**0.2**^^	**0.16**^	**0.19**^^	0.1	-0.05
Sat.low	-0.08	-0.02	0.02	0.07	0.01
Sat.mid	0.08	-0.09	0.02	0.07	0.01
Sat.high	0.13	0.1	0.04	-0.01	0.01
Warm	**-0.05**^^	-0.04	-0.2	0	0.03
Cold	**0.05**^^	0.04	0.2	0	-0.03
Pleasure	**-0.19**^^	-0.08	**-0.18**^	-0.09	**0.22**^^
Arousal	**0.23***	0.09	0.1	0	-0.08
Dominance	**0.28****	0.11	**0.17**^	0.05	**-0.18**^^
# of faces	**-0.16**^	0.03	0.11	-0.11	-0.03
# of people	**-0.22**^^	-0.05	-0.07	-0.01	0.07

Note. ^$p<0.1$, ^^$p<.05$, *$p<.01$, **$p<.001$.

Table 1: Correlation Matrix of the picture features against the personality traits: (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism.

6. PERSONALITY PREDICTION

Given that we found significant correlations between picture features and personality traits, we explored personality prediction based on these features. We trained our predictive model with the *radial basis function network* classifier in Weka, with a 10-fold cross-validation. We report the *root-mean-square error* (RMSE) in Table 2. The RMSE of each personality trait relates to the [1,5] score scale.

Personality	RMSE
Openness to experience	0.73
Conscientiousness	0.69
Extraversion	0.95
Agreeableness	0.74
Neuroticism	0.95

Table 2: Personality prediction with the root-mean-square error (RMSE).

The RMSE values that we found are low and comparable with previous work on personality extraction from social media trails. For example, Quercia et al. [10] looked at the relationship between personality traits and Twitter usage and reported RSME scores of 0.69, 0.76, 0.88, 0.79, and 0.85, respectively for, openness to experience, conscientiousness, extraversion, agreeableness, and neuroticsm. Our results, as well as the results of prior work show that the

Personality	Picture properties
Openness to experience	Green, low brightness, high saturation, cold colors, few faces
Conscientiousness	Saturated and unsaturated colors
Extraversion	Green and blue tones, low brightness, saturated and unsaturated colors
Agreeableness	Few dark and bright areas
Neuroticism	High brightness

Table 3: Picture properties in relation to personality traits. The properties apply for the pictures of users who score high in the respective personality trait.

most difficult traits to predict are the extraversion and the neuroticism personality traits.

7. DISCUSSION

We found Instagram picture features to be correlated with personality. Results show that the most strongly significant correlations are found in the openness to experience personality trait. Although we found weaker significant levels for the other personality traits, we were still able to find distinct correlations. See Table 3 for a summary of our findings.

Based on the found correlations, we also explored the prediction of the personality traits based on the picture features. Compared with the findings of prior work (i.e., [10]), we were able to find similar results and patterns. The most successful personality traits to predict are openness to experience, conscientiousness, and agreeableness, whereas the more difficult traits are extraversion and neuroticism.

8. LIMITATIONS AND FUTURE WORK

Our study contains limitations that need to be considered. Although we were able to obtain a fair amount of Instagram pictures (n=22,398), our personality measurement was limited to 113 participants. Given that we only had personality information of 113 participants to find relationships with picture features, we decided to give attention to the marginally significant results as well (i.e., significant levels of $.1 > p > .05$). A bigger sample size to assess personality traits should provide more conclusive results about the marginally significant effects that we found in this study.

In this study we solely focused on participants based in the United States. However, color interpretation and meaning could be influenced by cultural factors. Therefore, cultures could engage in different behavior of picture taking and applying filters. Future work should address this.

9. CONCLUSION

Our results suggest that there is a relation between personality traits and the way of taking and applying filters to pictures. By analyzing pictures of Instagram, personality traits can be inferred. Implicitly extracting personality from social media trails makes it possible to facilitate systems in order to provide a personalized experience. For example, it can help recommender systems to overcome the cold-start problem [14], or provide a personalized user interface [4].

10. ACKNOWLEDGMENTS

This research is supported by the Austrian Science Fund (FWF): P25655; and by the EU FP7/2013-2016 through the PHENICX project under grant agreement no.601166.

11. REFERENCES

[1] M. D. Back, J. M. Stopfer, S. Vazire, S. Gaddis, S. C. Schmukle, B. Egloff, and S. D. Gosling. Facebook profiles reflect actual personality, not self-idealization. *Psychological science*, 2010.

[2] B. Ferwerda and M. Schedl. Enhancing music recommender systems with personality information and emotional states: A proposal. In *Proceedings of the 2nd EMPIRE Workshop*, pages 1613–0073, 2014.

[3] B. Ferwerda, M. Schedl, and M. Tkalcic. Personality & emotional states: Understanding users' music listening needs. *UMAP 2015 Extended Proceedings*, 2015.

[4] B. Ferwerda, E. Yang, M. Schedl, and M. Tkalcic. Personality traits predict music taxonomy preferences. In *CHI'15 Extended Abstracts*. ACM, 2015.

[5] J. Golbeck, C. Robles, M. Edmondson, and K. Turner. Predicting personality from twitter. In *IEEE Third Conference on Social Computing*, pages 149–156, 2011.

[6] S. D. Gosling, S. Gaddis, S. Vazire, et al. Personality impressions based on facebook profiles. *ICWSM*.

[7] O. P. John, E. M. Donahue, and R. L. Kentle. The big five inventory: institute of personality and social research. *UC Berkeley, CA*, 1991.

[8] R. R. McCrae and O. P. John. An introduction to the five-factor model and its applications. *Journal of personality*, 60(2):175–215, 1992.

[9] G. Park, H. A. Schwartz, J. C. Eichstaedt, M. L. Kern, M. Kosinski, D. J. Stillwell, L. H. Ungar, and M. E. Seligman. Automatic personality assessment through social media language. 2014.

[10] D. Quercia, M. Kosinski, D. Stillwell, and J. Crowcroft. Our twitter profiles, our selves: Predicting personality with twitter. In *IEEE Third Conference on Social Computing*, pages 180–185, 2011.

[11] D. Rawlings and V. Ciancarelli. Music preference and the five-factor model of the neo personality inventory. *Psychology of Music*, 25(2):120–132, 1997.

[12] C. Ross, E. S. Orr, M. Sisic, J. M. Arseneault, M. G. Simmering, and R. R. Orr. Personality and motivations associated with facebook use. *Computers in Human Behavior*, 25(2):578–586, 2009.

[13] M. Tkalcic, B. Ferwerda, D. Hauger, and M. Schedl. Personality Correlates for Digital Concert Program Notes. *UMAP*, pages 1–6, 2015.

[14] M. Tkalcic, M. Kunaver, A. Košir, and J. Tasic. Addressing the new user problem with a personality based user similarity measure. In *DEMRA*, 2011.

[15] M. Tkalcic and J. Tasic. Colour spaces: perceptual, historical and applicational background. *IEEE EUROCON. Computer as a Tool*, pages 304–308, 2003.

[16] P. Valdez and A. Mehrabian. Effects of color on emotions. *Journal of Experimental Psychology: General*, 123(4):394, 1994.

[17] P. Viola and M. J. Jones. Robust Real-Time Face Detection. *International Journal of Computer Vision*, 57(2):137–154, May 2004.