Analyzing Personality through Social Media Profile Picture Choice

Leqi Liu[*†]
Bryn Mawr College
leqi.liu.11@gmail.com

Daniel Preoţiuc-Pietro[*]
Positive Psychology Center
University of Pennsylvania
danielpr@sas.upenn.edu

Zahra Riahi Samani[*†]
Faculty of Computer Science and Engineering
Shahid Beheshti University G.C
z_rsamani@sbu.ac.ir

Mohsen E. Moghaddam
Faculty of Computer Science and Engineering
Shahid Beheshti University G.C
m_moghadam@sbu.ac.ir

Lyle Ungar
Computer & Information Science
University of Pennsylvania
ungar@cis.upenn.edu

Abstract

The content of images users post to their social media is driven in part by personality. In this study, we analyze how Twitter profile images vary with the personality of the users posting them. In our main analysis, we use profile images from over 66,000 users whose personality we estimate based on their tweets. To facilitate interpretability, we focus our analysis on aesthetic and facial features and control for demographic variation in image features and personality. Our results show significant differences in profile picture choice between personality traits, and that these can be harnessed to predict personality traits with robust accuracy. For example, agreeable and conscientious users display more positive emotions in their profile pictures, while users high in openness prefer more aesthetic photos.

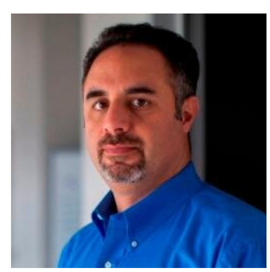
Introduction

Social media gives users the opportunity to build an online persona through posting of content such as text, images, links or through interaction with others. The way in which users present themselves is a type of behavior usually determined by differences in demographic or psychologic traits. Using large data sets of users and their online behaviors, recent studies have managed to successfully build models to predict a wide range of user traits such as age (Rao et al. 2010), gender (Burger et al. 2011), occupation (Preoţiuc-Pietro, Lampos, and Aletras 2015), personality (Schwartz et al. 2013), political orientation (Pennacchiotti and Popescu 2011) and location (Cheng, Caverlee, and Lee 2010). These studies used different types of information, ranging from social network connections which use the homophily hypothesis (Rout et al. 2013) to text from posts which are rooted in hypotheses about language use (Preoţiuc-Pietro, Lampos, and Aletras 2015).

The choice of content for posted images is a less studied online behavior. The picture of a user has been shown to be predictable of certain psychological traits by humans (Naumann et al. 2009). The study of profile images is particularly appealing as these are photos the users choose as representative for their online persona, and moreover, users can post pictures that do not stand for themselves. This choice is a type of behavior associated at least in part with personality, which is usually expressed by the five factor model (Digman 1990), (McCrae and John 1992) – the 'Big Five' – consisting of openness to experience, conscientiousness, extraversion, agreeableness and neuroticism.

For example, extraverts enjoy interacting with others, have high group visibility and are perceived as energetic. This could lead to extraverts using profile pictures involving other people or where they express more positive emotions. Users high in conscientiousness tend to be more orderly and prefer planned behaviors. This could lead users to conform to norms of what is expected from a profile picture i.e., a frontal photography of themselves. Conversely, users high in openness to experience may be more inclined to choose unconventional images and poses, as a general inclination of this type of people for art and novelty. Neuroticism is associated with negative emotions, which could also be reflected through users' choices of profile images. For example, Figure 1 illustrates sample profile images of users that score very high in extraversion and conscientiousness.

(a) Extraverted.



(b) Conscientious.

Figure 1: Example Twitter profile pictures for users scoring high in a personality trait.

The aim of this study is to analyze a broad range of interpretable image features from Twitter profile pictures, such

[*]These authors have contributed equally

[†]This work was carried out during a research visit at the University of Pennsylvania

as colors, aesthetics, facial presentation and emotions. Working with these features, we uncover their relationships with personality traits from the Big Five model. Previous studies (Celli, Bruni, and Lepri 2014) have shown that personality traits are predictable from images, demonstrating the existence of a correlation between personality and profile picture choice in social media. However, these fall short in some aspects. Foremost, the features of the models provide no interpretability and thus are not useful for psychologists who wish to understand the underlying correlations and generate hypotheses for further testing. Moreover, the data sets analyzed were very limited in size and user diversity, a problem that is very common as well in most psychology research.

To alleviate these problems, our main analysis uses a large sample of over 60,000 Twitter users with personality estimated using existing state-of-the-art text prediction methods. This offers a breath of subjects of various demographics and an orders of magnitudes larger sample than previous studies and traditional psychological research. Further, in order to compare and put our text-based personality assessments into context, we use a smaller sample of 429 users who filled in a standard personality questionnaire. Finally, we test the predictive performance of our interpretable features in held-out data prediction. With our analysis, we aim to present a procedure that can scale up psychological profiling without requiring users to undertake costly questionnaires and that better matches their online persona.

Related Work

Personality detection from appearance by humans has long been a topic of interest in the domain of psychology (Haxby, Hoffman, and Gobbini 2000), as it has deep implications in studying personal interaction and first impressions. Most of the studies in psychology have focused on facial expressions as people frequently use facial characteristics as a basis for personality attributions (Penton-Voak et al. 2006), while other studies additionally considered the pose of the person (Naumann et al. 2009). Human raters were able to correctly evaluate certain personality traits as assessed through questionnaires, for example extraversion (Penton-Voak et al. 2006). While human perception is important, psychologists also raise the possibility that computer vision algorithms would be able to predict personality automatically as a way to avoid collecting costly questionnaire data (Kamenskaya and Kukharev 2008).

With recent advances in computer science and a wider availability of inexpensive user generated data, automatic personality detection has become an important research topic. Personality influences a wide range of behaviors, many of which can be directly observed through social media usage. Therefore, methods using a range of modalities have been successfully developed: video (Subramanian et al. 2013), audio (Alam and Riccardi 2014), text (Schwartz et al. 2013) or social data (Van Der Heide, D'Angelo, and Schumaker 2012; Hall, Pennington, and Lueders 2014).

In this study, we focus on static images, and in particular on self-selected profile pictures from social media. Although users can post other photos, studying profile pictures is particularly interesting as these reflect the impressions that the users want to convey to others. Although social media allows a user to shape his or her own personality and idealized view (the 'idealized virtual identity hypothesis'), evidence shows that social media behavior usually represents an extension of one's self (the 'extended real life hypothesis'), thus allowing others to observe the users' true personality (Back et al. 2010).

While most of the work in computer vision recognition has focused on object recognition, for personality prediction the subject of interest is usually a person or face. The typical computer vision framework for object recognition relies on thousands of low level features either pre-determined or, more recently, automatically extracted by deep neural networks. However, if using these for personality prediction, they would hardly offer any interpretability and insight into the image characteristics that reveal personality traits. A sub-category of work focuses on facial expression recognition (Pantic 2009), emotion recognition (Kim, Lee, and Provost 2013) and sentiment analysis (Borth et al. 2013; You et al. 2015) from images, all of which can disclose personality traits. Further, the separate area of computational aesthetics (Datta et al. 2006), aims to utilize features derived from photography theory to determine the factors that make a picture appealing.

Previous work on predicting personality from images has mainly focused on predictive performance. Recently, Celli, Bruni, and Lepri (2014) worked with profile pictures of 100 Facebook users with their self-assessed personalities and interaction styles. They used bag-of-visual-words features defined on local SIFT (Lowe 2004) features and combined different machine learning algorithms to test the effectiveness of classifying users as being high or low in each personality trait. They were able to classify personality traits with nearly 65% accuracy. In an attempt to interpret the results, they performed clustering on correctly classified images from each personality trait to find the most important characteristics of each personality trait and observed that extroverted and emotionally stable people tend to have pictures in which they are smiling or appear with other people. Al Moubayed, Noura and Vazquez-Alvarez, Yolanda and McKay, Alex and Vinciarelli, Alessandro (2014) used the FERET corpus consisting of 829 individuals whose personality was assessed by 11 independent judges. They used the first 103 eigenfaces as features for classification and reported around 70% accuracy in predicting personalities being above or below the median.

Data

We use two Twitter data sets in our experiments which differ in size and the set of available user traits.

TwitterText An orders of magnitude larger data set consists of 66,502 Twitter users with their self-reported gender information (31,307 males and 35,195 females). The labels were obtained by linking their Twitter accounts to accounts on other networks (e.g., MySpace, Blogger) where gender information was available (Burger et al. 2011; Volkova, Wilson, and Yarowsky 2013). For each user, we have collected up to 3,200 most recent tweets using the Twitter

REST API[1], leading up to a data set of 104,500,740 tweets.

TwitterSurvey An order of magnitudes smaller data set which contains 434 Twitter users whose Big Five personality scores were computed based on their completion of the International Personality Item Pool proxy for the NEO Personality Inventory Revised (NEO-PI-R) (Costa and McCrae 2008). We asked these users to self-report their gender as either male or female and age. All profile images were collected on the same day for all accounts in both data sets.

Text Analysis

We use posted tweets from an account as a different modality compared to images in order to predict user attributes and demographics. Text-based prediction methods have been successfully used to predict a wide range of traits including age (Rao et al. 2010), gender (Burger et al. 2011), political orientation (Pennacchiotti and Popescu 2011), location (Cheng, Caverlee, and Lee 2010), impact (Lampos et al. 2014), income (Preoţiuc-Pietro et al. 2015), occupation (Preoţiuc-Pietro, Lampos, and Aletras 2015), mental illnesses (De Choudhury, Counts, and Horvitz 2013) and personality (Schwartz et al. 2013).

As pre-processing, we tokenize all posts and filtered for English using the langid.py tool (Lui and Baldwin 2012). We then aggregate all user's posts and use state-of-the-art text prediction algorithms to estimate personality and age for users. In order to get reliable estimates for these traits, we only use the users who posted at least 50 tweets: only 254 users from TwitterSurvey and all users from TwitterText, as we filtered the users initially.

Personality We use the method developed by (Schwartz et al. 2013) to assign each user scores for personality from the popular five factor model of personality – 'Big Five' – (McCrae and John 1992), which consists of five dimensions: extraversion, agreeableness, conscientiousness, neuroticism and openness to experience. The model was trained on a large sample of around 70,000 Facebook users who have taken Big Five personality tests and shared their posts using a model using 1-3 grams and topics as features (Park et al. 2014; Schwartz et al. 2013).

In the original validation, the model achieved a Pearson correlation of $r > .3$ predictive performance for all five traits (Schwartz et al. 2013; Park et al. 2014), which is considered a high correlation in psychology, especially when measuring internal states (Meyer et al. 2001). However, in our use case, the text comes from a different social media (Twitter) and thus may suffer from some domain adaptation issues. A subset of 254 users in our TwitterSurvey data set have both taken the Big Five questionnaire and have predicted personality from their tweets. Figures 2a and 2b display the inter-correlations between the personality traits when the traits are assessed through questionnaires or tweets. We observe the same correlation patterns in text predictions as in the questionnaire based assessments, with neuroticism strongly anti-correlated with conscientiousness, extraversion and agreeableness in both cases, while conscientiousness and agreeableness have high positive correlation. Figure 2c shows the correlations between personality traits using the two different methods. Most importantly, all correlations between the same trait predicted by the two different methods are significantly correlated (0.123 critical value for $p < .05$, two tailed test), albeit smaller than what is reported in the original model using Facebook data.

Age We predict age from Twitter posts using the method introduced in (Sap et al. 2014). We use the estimated age in the TwitterText data set in order to control for the effects of basic demographics (gender and age) on the resulting correlations. The correlation of the outcome of the prediction model and true age, trained and tested on Facebook data, reaches .835 Pearson correlation (Sap et al. 2014).

Image Extraction

In this paper, we use public profile images as representative for a user. Although users can post other images, we focus on profile images, as the user has chosen these to represent their online persona and thus is most likely to contain important psychological cues (Penton-Voak et al. 2006).

In order to study and interpret people's personalities from their profile pictures, stylistic characteristics rather than traditional computer vision features of the profiles are more appropriate (Redi et al. 2015). Most profile images contain faces, which are known to reflect personality. We thus divide image features in two categories: general image features and stylistic facial features. The former contains basic color and facial information, while the later also includes facial expressions and postures extracted from the largest recognizable face from profile images.

We use two APIs based on deep learning methods – Face++[2] and EmoVu[3] – for facial feature extraction. We used Face++, which provides very accurate face recognition (Zhou et al. 2013), to indicate demographics and facial presentation. EmoVu offers more information about emotions expressed by the faces detected in the profile images.

We divide the features into the following categories:

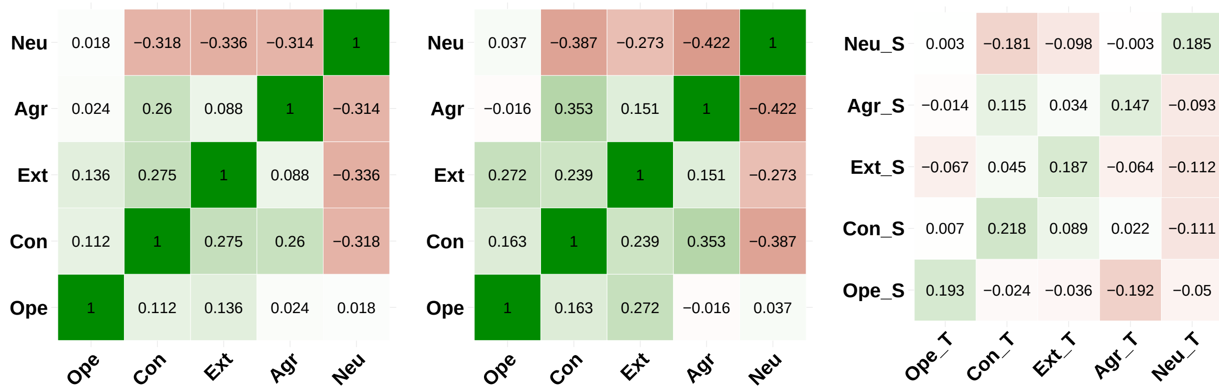
Color First, we divide images into grayscale images and color images. For color images, we have taken their normalized red, green and blue values and the average of the original colors. Colors are related to conceptual ideas like people's mood and emotion. Previous research showed that colors from images are related to psychologic traits (Wexner 1954): red with 'exciting-stimulating' and 'protective-defending'; green with 'calm-peaceful-serene'; and blue is connected with 'secure-comfortable' as well as 'calm-peaceful-serene'.

Human judgements of the attractiveness of images are influenced by color distributions (Huang, Wang, and Wu 2006) and aesthetic principles related to color composition (Datta et al. 2006). We thus compute brightness and contrast as the

[1] `https://dev.twitter.com/rest/public`

[2] `http://www.faceplusplus/`

[3] `http://emovu.com/`



(a) Survey Personality. (b) Text Personality. (c) Text (T) – Survey (S).

Figure 2: Pearson correlations between the Big Five personality traits assessed with both text-predicted and image-predicted personality on the same set of users. Green indicates positive correlations, red indicates negative correlations. All correlations are controlled for age and gender.

relative variations of luminance. We also represent images in the HSV (Hue-Saturation-Value) color space and extract the mean and variance for saturation and hue. High saturation indicates vividness and chromatic purity, which are more appealing to the human eye, although hue is not as clearly interpretable (Datta et al. 2006). Colorfulness is calculated as the difference against gray (San Pedro and Siersdorfer 2009) and naturalness measures the degree of correspondence between images and human perception (Huang, Wang, and Wu 2006). Sharpness is represented as mean and variance of the image Laplacian normalized by local average luminance. This aims to measure coarseness or the degree of detail contained in an image, which is a proxy for the quality of the photographing gear and photographer (Ke, Tang, and Jing 2006). Image blur is estimated using the method from (Ke, Tang, and Jing 2006). In this set of features, we do not explicitly detect the subject, but similarly to Geng et al. (2011) we use the saliency map (Ma et al. 2005) to compute a probability of each pixel to be on the subject and re-weight the image features by this probability. We compute all the above features for both the original image and the re-weighted image. Due to length limits, we only present correlations with these in the analysis section. Finally, we compute the affective tone of colors (Wei-ning, Ying-lin, and Sheng-ming 2006), represented by 17 color histogram features that are used to automatically annotate emotional image semantics for emotional image retrieval.

Image Composition We measure aesthetic features of basic photographic composition rules. First, we study the rule of thirds, where the main object in the picture lies at the border or inside an inner rectangle of a 3×3 grid. Professional photos strive for simplicity. We capture this using two methods. We first compute the spatial distribution of the high frequency edges of an image. In good quality photos, the edges are focused on the subject. We use the method from (Ke, Tang, and Jing 2006) to estimate the edge distribution between the subject and background. The number of unique hues of a photo is another measure of simplicity, based on the fact that good compositions have fewer objects, resulting in fewer distinct hues (Ke, Tang, and Jing 2006). Visual weight measures the clarity contrast between subject region and the whole image. Finally, the presence of lines in an image induces emotional effects (Arnheim 2004), therefore we compute the proportion of static and dynamic lines in the image (Machajdik and Hanbury 2010).

Image Type We extract basic face-related features for each profile picture as the number of faces it contains. If there is no face in the profile image, we look at whether the user uses the one of the default Twitter profile picture images. For profile images that contain faces, in addition to the face count we also create two binary features indicating whether there is exactly one face or multiple faces.

Image Demographics Age, gender and race (Asian, Black or White) are demographic features estimated from the profile images. When choosing profile pictures to represent themselves, these can either be of different people (e.g., children, friends), can include multiple people (e.g., spouse) or can use photos from their past or that make them appear younger.

Facial Presentation This category contains facial features related to the way a user chooses to present himself through his profile image. Features include the face ratio (the size of the face divided by the size of the profile picture), whether the face wears any type of glasses (reading or sunglassses), the closeness of the subject's face from the acquisition sensor provided by EmoVu's attention measurement (Eyeris 2016), the 3D face posture, which includes the pitch, roll and yaw angle of the face and eye openness. All these features try to capture the self-presentation characteristics of the user.

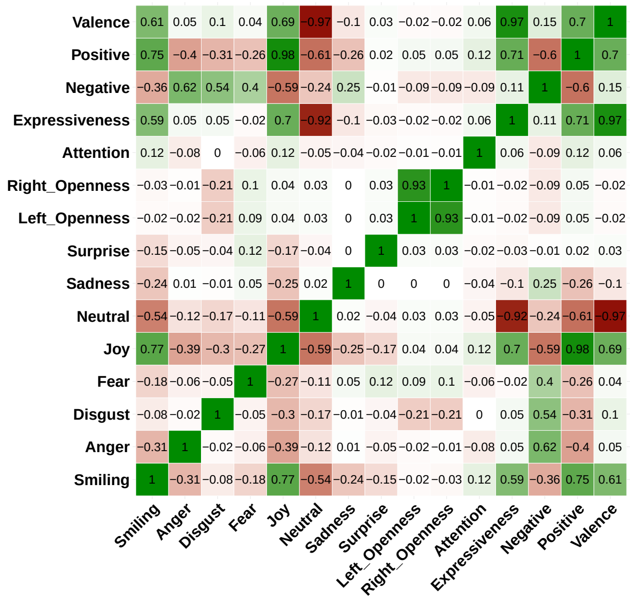


Figure 3: Inter-correlation table between facial emotion features.

Facial Expressions We adopt Ekman's model of six discrete basic emotions: anger, disgust, fear, joy, sadness and surprise (Ekman and Friesen 1971), which were originally identified based on facial expressions. We use the EmoVu API to automatically extract these emotions from the largest detected face in each profile image. Additionally, we also extract neutral expression (Batty and Taylor 2003). The six basic emotions can be categorized as either positive (joy and surprise) or negative (anger, disgust, fear, sadness). Along with the basic emotional expressions, EmoVu also gives composite features calculated from the basic ones (Eyeris 2016). Expressiveness, also referred to as the 'interaction' metric, is the highest value of the six basic emotions. Negative and positive mood are calculated as the maximum value of the positive and negative emotions respectively. Valence is the average of the negative mood and positive mood. Also, in this category we add the smiling degree provided by Face++. The features in this category should have strong correlations. Figure 3 presents the inter-correlation of the emotion features.

The correlations conform to what one would expect. Smiling, joy, positive mood and expressiveness are largely positively correlated and the four are significantly negatively correlated with anger, neutral and negative mood. Eye openness is correlated with fear and anti-correlated with disgust.

In total, Face++ was able to recognise 208 images with at least one face out of 429 from profile images for the TwitterSurvey data set and 36,402 out of 66,502 profile images for the TwitterText data set. EmoVu was able to detect facial emotions for 124 out of 429 images for the TwitterSurvey data set and 26,234 out of 66,502 profile images for the TwitterText data set. This was caused by different factors such as low image quality, very small face, or the face being obstructed by an object or not facing the camera. Similarly, we computed color and image composition features only to images that were not default or grayscale.

Analysis

In this section, we explore the correlations between personality measured through posted text and the image features introduced in the previous section. To this end, we compute univariate Pearson correlations between each visual feature and personality score. Demographic traits – most importantly age and gender – are known to affect both personality features (McCrae et al. 1999) as well as text-derived outcomes (Schwartz et al. 2013). In order for our correlations not to be an artifact caused by these demographic confounds, we control using partial correlation for gender (self-reported in both data sets) and age (predicted from an users' posts). We also include correlations with age and gender separately in Table 1 in order to highlight the important patterns caused by demographic factors. Due to space limitations, for 'Color Emotions' and 'Rule of Thirds' we show a single average correlation for all features of that type. The resulting correlations are presented in Table 1.

Openness The personality dimension of openness can be meaningfully separated into the distinct, but correlated subtraits of 'Intellect' and 'Openness to Experience'. The first reflects intellectual engagement while the latter indicates engagement with perceptual and aesthetic domains (DeYoung, Quilty, and Peterson 2007). Our large analysis of image type reveals that these users are most likely to have profile pictures other than faces, which reveals non-conformance with what is expected.

Most importantly, users high in openness are significantly correlated to the majority of features indicative of better aesthetic quality of their photos. In general, appealing images tend to have increased contrast, sharpness, saturation and less blur, which is the case for people high in openness. However, their photos are anti-correlated with color emotions and are less colorful. Naturalness is anti-correlated perhaps because of the artistic quality of images, fact reflected also by the correlation with the picture being grayscale. Image composition features confirm the color features findings. Edge distribution is the highest correlated feature, while smaller hue count, also indicative of simplicity is also correlated. Finally, the dynamic lines which should reflect emotion are significantly anti-correlated, again confirming that photos of users high in openness are low in emotion, albeit of artistic and aesthetic quality. For facial presentation, our results indicate these users display reading glasses, but not sunglasses and when a face is present, this is larger. In general, psychology research has shown that a person wearing reading glasses is more intelligent or has intellectual virtues (Hellström and Tekle 1994).

Facial emotions confirm the findings of the color features. Photos are higher in negative emotions, particularly anger, and lower in attention, smiling, valence and positive emotion, especially joy.

Feature	Demographics		Personality Trait				
Color	Gender	Age	Ope	Con	Ext	Agr	Neu
Grayscale	-.050	-.014	.050	-.031	-.012		.014
Red	.026				-.041		
Green	-.021	.012			.021		.011
Blue	-.022				.045		
Average RGB	.030	.015		.025	.033	.019	
Brightness	.024	.015		.028	.012	.023	
Contrast	.012			.016		.019	-.011
Saturation	.038	.015	.017			-.016	.014
Hue		.012	-.021	-.015	.022		.013
Colorfulness			-.017	.013	.040	.029	-.036
Naturalness		.029	-.015	.013	-.036	.011	
Sharpness	-.056		.025	-.022	.015	-.021	.014
Blur		.057	-.011	.036		.023	
Average Color Emotions	.018		-.021			.021	-.017
Image Composition	Gender	Age	Ope	Con	Ext	Agr	Neu
Average Rule of Thirds	.034	-.056	-.033	-.021	.032	.033	-.034
Edge Distribution	-.034	.018	.047			-.048	.038
Hue Count		.028					
Visual Weight	.010			-.014			
Static Lines	.058				.017	.018	
Dynamic Lines	.042	.016	-.020			.033	
Image Type	Gender	Age	Ope	Con	Ext	Agr	Neu
Default Image			-.022		-.043	.015	-.023
Is Not Face	-.072	-.021	.061	-.121	-.108	-.070	.071
One Face	.054	.029	-.016	.102	.081	.046	-.057
Multiple Faces	.040	-.019	-.102	.043	.058	.053	-.032
No. Faces	.072		-.092	.106	.103	.078	-.067
Image Demographics	Gender	Age	Ope	Con	Ext	Agr	Neu
Age	-.310	.306	.050	.105	-.036		
Gender	.795	-.041			.035	.034	
Asian	.064	-.150	-.072	-.042			
Black	-.034	-.061	.047	.050	.085	-.055	-.096
White	-.033	.169	.031		-.066	.026	.071
Facial Presentation	Gender	Age	Ope	Con	Ext	Agr	Neu
No Glasses	.145	-.036		.027	.085	.026	-.065
Reading Glasses	-.141	.054	.020		-.099	-.017	.071
Sunglasses	-.034	-.020	-.017	-.028		-.019	
Pitch Angle	-.043						
Roll Angle	.017						
Yaw Angle							
Face Ratio	.034	.036	.038	-.039	-.097	-.039	.057
Facial Expressions	Gender	Age	Ope	Con	Ext	Agr	Neu
Smiling	.229	.141	-.089	.190	.050	.148	-.104
Anger	-.108	-.019	.037	-.080	-.042	-.055	.056
Disgust	-.142	.048					
Fear		-.017	.018	-.029		-.043	.018
Joy	.191	.119	-.093	.180	.061	.140	-.107
Sadness	-.122	-.032	.023	-.051		-.034	.026
Surprise	.038	-.064		-.041		-.031	
Left Eye Openness	.093			.025			
Right Eye Openness	.091			.027			
Attention	-.055	.061	-.047	.049	.018	.040	-.048
Expressiveness	.101	.123	-.072	.140	.054	.106	-.089
Neutral	-.064	-.133	.068	-.128	-.047	-.093	.081
Positive Mood	.198	.111	-.093	.175	.065	.137	-.107
Negative Mood	-.164		.043	-.079	-.029	-.067	.044
Valence	.101	.132	-.075	.140	.053	.105	-.090

Table 1: Pearson correlations between profile image and Big Five personality controlled for age and gender and with age and gender (coded as 1 – female, 0 – male) separately. Positive correlation is highlighted with green (paler green $p < .01$, deeper green $p < .001$, two-tailed t-test) and negative correlation with red (paler red $p < .01$, deeper red $p < .001$, two-tailed t-test).

Conscientiousness Conscientiousness is the personality trait associated with orderliness, planned behavior and self-discipline. The image type features are strongest correlated to this trait. These indicate that profile images with faces, especially with only one face, are good indicators of higher conscientiousness. This behavior can be caused by the fact that users high in this trait prefer the expected behavior (i.e., posting a picture of themselves).

In terms of colors, conscientious users prefer pictures which are not grayscale and are overall more colorful, natural and bright. Despite this, their pictures are not more aesthetic, being anti-correlated with sharpness and positively correlated with blur. Image composition shows a strong correlation only with not respecting the rule of thirds. These users show correlations for facial presentation with not wearing any type of glasses and a smaller face ratio. By analyzing demographics inferred from images, we observe that there is a strong correlation with predicted age, even if the data set is controlled for age. This, together with preference for a single face, indicate that conscientious users might display pictures that make them seem older.

Facial expressions are very indicative of conscientious people. The facial emotions of smiling, positive mood and valence (mostly influenced by joy) are all highly positively correlated, while negative mood, especially anger and sadness, are anti-correlated. We observe negative correlations with negative mood, disgust and fear and strong positive correlations with positive mood, joy and smiling. In general, conscientious people express the most emotions (highest expressiveness, lowest neutral) across all five traits. This does not align with what is generally known about conscientious people, but is explainable by taking in account that in a profile picture, a person is expected to smile and appear happy.

Extraversion Extraversion is a trait marked by engagement with the outside world. These type of users are correlated the highest out of all traits with colorful images (both colorfulness and high average RGB). Their photos do not have any correlation with the color attributes that make a photo aesthetically pleasing (contrast, saturation, lack of blur), with the exception of a positive correlation with sharpness. The number of static lines indicates a small positive correlation with emotion. In other composition features, extroverts are only correlated with the use of the rule of thirds.

Similar to conscientiousness, extraversion is largely related to the number of faces of the profile pictures, albeit extraverts slightly prefer images with more people. Different from all other personalities, extraversion is negatively correlated with the age of the presenting faces, which means that users either appear younger in their profiles or are photographed with other young(er) people.

With the strongest correlation compared to other personalities, extraverts have a small face ratio, perhaps caused by the multiple people present in the picture or showing of more of their body or environment. Extroverted people are also strongly associated with not displaying reading glasses, which was shown to be associated with introverts (Harris, Harris, and Bochner 1982). For facial expressions, extraverts display the same positive emotion trends as conscientious people, although weaker across the board.

Agreeableness The agreeableness trait is characterized by social harmony and cooperation. Users high in this trait like to have profile pictures with faces in them. For colors, the correlations are almost all opposite to those for openness, even though the two traits are uncorrelated in both surveys and text predictions. Agreeable people use colorful pictures (but to a lesser extent than extraverts) which are low in sharpness, blurry and bright. They tend to respect the rule of thirds, but the edge distribution is strongly negatively correlated, hinting their pictures are cluttered as opposed to simple. Color emotions are highest across all traits, a fact also indicated by the presence of static and dynamic lines. This leads to the conclusion that, although bright, colorful and color emotive, pictures of agreeable users are not the most aesthetically pleasing. Facial presentation features show very low magnitude correlations.

Facial emotion patterns are similar to psychology theory: very strong correlation with smiling, joy and overall positive emotion and low in all negative emotion expressions. This corresponds to the color correlations. Intriguingly, this is different to conscientious people who are highest in facial positive emotions, but do not express this through the overall color tone of the image as agreeable people do.

Neuroticism Neuroticism is associated with the experience of negative emotions and emotional instability. It is usually anti-correlated with agreeableness and extraversion. Notably, photos of neurotic people are perhaps unsurprisingly anti-correlated with colorfulness. The average color emotion correlations are also negative. In terms of composition, neurotic people display simpler images and do not respect the rule of thirds. This shows that overall, neurotic people display simple, uncolorful images with negative color emotions. Although this is similar for openness, the photos of neurotic people do not display the aesthetic features that characterize openness.

Neurotic people have a strong tendency not to present faces. When a face is present, they have the strongest positive correlation with displaying reading glasses across all traits and, when a face is present, it is significantly larger. Presence of reading glasses have been associated with perceived introversion and a decrease in attractiveness (Terry and Kroger 1976). In terms of facial emotions, neuroticism displays, as expected, both a lack of positive emotions and, to a lower extent, the presence of negative emotions. Higher correlations than for negative emotions are obtained with features related to the absence of emotions (neutral and expressiveness). Therefore, the lack of emotion expression is what characterizes their profile pictures, which aligns with the strong social norm against a very sad or angry appearance in profile pictures. In general, when examining facial emotion correlation patterns, we highlight two well aligned clusters: openness and neuroticism in one, and conscientiousness, extraversion and agreeableness in the other.

The same set of experiments on personality correlations on

Feature set	# Feat	Ope	Con	Ext	Agr	Neu
Colors	44	.071	.060	.089	.057	.045
Image Composition	10	.053	.031	.084	.051	.039
Image Type	5	.112	.122	.117	.082	.078
Demographics	5	.065	.086	.066	.044	.065
Facial Presentation	7	.046	.034	.099	.037	.064
Facial Expressions	14	.068	.114	.045	.090	.072
All	85	.162	.189	.180	.150	.145

(a) TwitterText data set.

Feature set	# Feat	Ope	Con	Ext	Agr	Neu
Colors	44	(.0)	(.0)	(.0)	(.002)	.122
Image Composition	10	(.03)	(.026)	(.0)	(.0)	(.043)
Image Type	5	(.0)	(.086)	(.0)	(.030)	(.0)
Demographics	5	(.011)	(.091)	(.0)	(.037)	.128
Facial Presentation	7	.147	(.042)	(.040)	(.0)	.033
Facial Expressions	14	.139	.125	(.041)	(.0)	.101
All	85	.190	.134	.095	(.046)	.151

(b) TwitterSurvey data set.

Table 2: Predictive performance using Linear Regression, measured in Pearson correlation over 10-fold cross-validation. Correlations in brackets are not significant ($p < .05$, two-tailed t-test).

the TwitterSurvey dataset unveiled a total of only 3 significant correlations at $p < .01$ and none at $p < .001$. Given that these were obtained from a total of 260 tests, we cannot consider any of these correlations as being robust to randomness. This shows the need for this type of behavior to be studied using very large sample sizes of social media personality. This also hints at the possibility that personality of social media users is better measured through other social media behaviors (here, tweets) and may not be equal to offline personality, a question which we leave for future work.

Prediction

Finally, we investigate the accuracy of using interpretable visual features to predict the personality traits. We use linear regression with Elastic Net regularization (Zou and Hastie 2005) as our prediction algorithm. We report results on 10 fold cross-validation. We test the prediction performance of each independent group as well as a model that uses all features. As in our analysis section, to avoid demographic confounds, the personality outcomes are the residual of each trait after adjusting for the effect of age and gender. When features could not be extracted i.e., in the case of facial presentation and facial expressions when there is no face detected, we replace these with the sample mean. Results, measured using Pearson correlation over the 10 folds and both data sets are presented in Table 2a. Similar patterns can be observed using Root Mean Squared Error (RMSE) and are omitted for brevity.

On the TwitterText data set we observe that the most useful category for prediction is the type of image, despite containing only 5 features. Colors, image composition, and facial presentation are the most useful features for predicting extraversion, while facial expressions are the least predictive for this trait. Conscientiousness is most distinctive through facial expressions and overall easiest to predict.

For the TwitterSurvey data set, we can predict with significant accuracy all traits except agreeableness, despite the very small sample size. On this dataset, openness is easiest to predict, especially through facial presentation and expression. Similarly to the TwitterText dataset, conscientiousness is very accurately revealed through facial expressions. Neuroticism is highly predictive through colors, demographics and facial expressions, although these overlap substantially, causing the combined performance to be only slightly higher than the individual accuracies.

Using the TwitterText data set, we observe an overall correlation of $r > .145$ across all traits with conscientiousness the most predictive at $r = .189$. In order to put this into context, psychological variables typically have a 'correlational upper-bound' around $.3 - .4$ correlation (Meyer et al. 2001). We also note that the method for personality prediction using text reports a Pearson correlation of $r => .3$ for all five traits. However, their method uses thousands of features extracted from hundreds of posts per person. Our method uses a single profile image to make the personality prediction.

Conclusion

We presented the first large-scale study of profile photos on social media and personality that allows for psychological insight. To this end, we used a range of interpretable aesthetic and facial features. Our personality assessment method used the tweets of the users and was compared to a smaller data set collected using a standard psychological questionnaire. While experiments on the latter data set did not offer statistical power to uncover any strong relationships, our large scale experiments allowed us to find correlations with personality that are in line and complement psychological research.

We concluded that each personality trait has a specific type of profile picture posting. Users that are either high in openness or neuroticism post less photos of people and when these are present, they tend not to express positive emotions. The difference between the groups is in the aesthetic quality of the photos, higher for openness and lower for neuroticism. Users high in conscientiousness, agreeableness or extraversion prefer pictures with at least one face and prefer presenting positive emotions through their facial expressions. Conscientious users post more what is expected of a profile picture: pictures of one face that expresses the most positive emotion out of all traits. Extraverts and agreeable people regularly post colorful pictures that convey emotion, although they are not the most aesthetically pleasing, especially for the latter trait. Finally, we tested the predictive performance of our features, showing relatively robust accuracy.

Acknowledging possible limitations of this study, we consider this represents a necessary experiment in analyzing social media profile images using interpretable features on a data set orders of magnitude larger than previously. Future work will analyze a more diverse set of psychological traits by looking at a wider set of photos that users post, curate or engage with using social media.

Acknowledgments

The authors acknowledge the support of the Templeton Religion Trust, grant TRT-0048.

References

Al Moubayed, Noura and Vazquez-Alvarez, Yolanda and McKay, Alex and Vinciarelli, Alessandro. 2014. Face-based automatic personality perception. In *Proceedings of the 22nd ACM international conference on Multimedia*, MM, 1153–1156.

Alam, F., and Riccardi, G. 2014. Predicting Personality Traits using Multimodal Information. In *Proceedings of the 22nd ACM international conference on Multimedia, Workshop on Computational Personality Recognition*, WCPR, 15–18.

Arnheim, R. 2004. *Art and Visual Perception: A Psychology of the Creative Eye*. University of California Press.

Back, M. D.; Stopfer, J. M.; Vazire, S.; Gaddis, S.; Schmukle, S. C.; Egloff, B.; and Gosling, S. D. 2010. Facebook Profiles Reflect Actual Personality, not Self-Idealization. *Psychological Science* 21:372–374.

Batty, M., and Taylor, M. J. 2003. Early Processing of the Six Basic Facial Emotional Expressions. *Cognitive Brain Research* 17(3):613–620.

Borth, D.; Ji, R.; Chen, T.; Breuel, T.; and Chang, S.-F. 2013. Large-scale visual sentiment ontology and detectors using adjective noun pairs. In *Proceedings of the 21st ACM International Conference on Multimedia*, MM, 223–232.

Burger, D. J.; Henderson, J.; Kim, G.; and Zarrella, G. 2011. Discriminating Gender on Twitter. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, EMNLP, 1301–1309.

Celli, F.; Bruni, E.; and Lepri, B. 2014. Automatic personality and interaction style recognition from facebook profile pictures. In *Proceedings of the 22nd ACM International Conference on Multimedia*, MM, 1101–1104.

Cheng, Z.; Caverlee, J.; and Lee, K. 2010. You are where you Tweet: A Content-Based aAproach to Geo-Locating Twitter Users. In *Proceedings of the 19th ACM Conference on Information and Knowledge Management*, CIKM, 759–768.

Costa, P., and McCrae, R. 2008. *The Revised NEO Personality Inventory (NEO-PI-R)*. SAGE Publications Ltd.

Datta, R.; Joshi, D.; Li, J.; and Wang, J. Z. 2006. Studying Aesthetics in Photographic Images Using a Computational Approach. In *Proceedings of the 9th European Conference on Computer Vision*, ECCV, 288–301.

De Choudhury, M.; Counts, S.; and Horvitz, E. 2013. Social media as a measurement tool of depression in populations. In *Proceedings of the 5th Annual ACM Web Science Conference*, WebScience, 47–56.

DeYoung, C.; Quilty, L.; and Peterson, J. 2007. Between facets and domains: 10 aspects of the big five. *Journal of Personality and Social Psychology* 93(5):880–896.

Digman, J. M. 1990. Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology* 41(1):417–440.

Ekman, P., and Friesen, W. V. 1971. Constants across Cultures in the Face and Emotion. *Journal of Personality and Social Psychology* 17(2):124.

Eyeris. 2016. Eyeris EmoVu Web API documentation. `http://emovu.com/docs/html/web_api.htm`.

Geng, B.; Yang, L.; Xu, C.; Hua, X.-S.; and Li, S. 2011. The role of Attractiveness in Web Image Search. In *Proceedings of the 19th ACM International Conference on Multimedia*, MM, 63–72.

Hall, J. A.; Pennington, N.; and Lueders, A. 2014. Impression Management and Formation on Facebook: A Lens Model Approach. *New Media & Society* 16(6):958–982.

Harris, M. B.; Harris, R. J.; and Bochner, S. 1982. Fat, Four-Eyed, and Female: Stereotypes of Obesity, Glasses, and Gender. *Journal of Applied Social Psychology* 12(6):503–516.

Haxby, J. V.; Hoffman, E. A.; and Gobbini, M. I. 2000. The Distributed Human Neural System for Face Perception. *Trends in Cognitive Sciences* 4(6):223–233.

Hellström, Å., and Tekle, J. 1994. Person Perception through Facial Photographs: Effects of Glasses, Hair, and Beard on Judgments of Occupation and Personal Qualities. *European Journal of Social Psychology* 24(6):693–705.

Huang, K.-Q.; Wang, Q.; and Wu, Z.-Y. 2006. Natural Color Image Enhancement and Evaluation Algorithm Based on Human Visual System. *Computer Vision and Image Understanding* 103(1):52–63.

Kamenskaya, E., and Kukharev, G. 2008. Recognition of Psychological Characteristics from Face. *Metody Informatyki Stosowanej* (1):59–73.

Ke, Y.; Tang, X.; and Jing, F. 2006. The Design of High-level Features for Photo Quality Assessment. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR, 419–426.

Kim, Y.; Lee, H.; and Provost, E. M. 2013. Deep Learning for Robust Feature Generation in Audiovisual Emotion Recognition. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, ICASSP, 3687–3691.

Lampos, V.; Aletras, N.; Preoţiuc-Pietro, D.; and Cohn, T. 2014. Predicting and Characterising User Impact on Twitter. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, EACL, 405–413.

Lowe, D. G. 2004. Distinctive image features from scale-invariant keypoints. *International journal of computer vision* 60(2):91–110.

Lui, M., and Baldwin, T. 2012. Langid.Py: An Off-the-shelf Language Identification Tool. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, System Demonstrations*, ACL, 25–30.

Ma, Y.-F.; Hua, X.-S.; Lu, L.; and Zhan, H.-J. 2005. A Generic Framework of User Attention Model and its Application in Video Summarization. *IEEE Transactions on Multimedia* 7(5):907–919.

Machajdik, J., and Hanbury, A. 2010. Affective Image Classification Using Features Inspired by Psychology and Art Theory. In *Proceedings of the 18th ACM International Conference on Multimedia*, MM, 83–92.

McCrae, R. R., and John, O. P. 1992. An introduction to the five-factor model and its applications. *Journal of Personality* 60:175–215.

McCrae, R. R.; Costa, P. T.; de Lima, M. P.; Simes, A.; Ostendorf, F.; Angleitner, A.; Marusic, I.; Bratko, D.; Caprara, G. V.; Barbaranelli, C.; Chae, J.-H.; and Piedmont, R. L. 1999. Age Differences in Personality across the Adult Life Span: Parallels in Five Cultures. *Developmental Psychology* 35(2):466–477.

Meyer, G. J.; Finn, S. E.; Eyde, L. D.; Kay, G. G.; Moreland, K. L.; Dies, R. R.; Eisman, E. J.; Kubiszyn, T. W.; and Reed, G. M. 2001. Psychological Testing and Psychological Assessment: A Review of Evidence and Issues. *American Psychologist* 56(2):128–165.

Naumann, L. P.; Vazire, S.; Rentfrow, P. J.; and Gosling, S. D. 2009. Personality Judgments based on Physical Appearance. *Personality and Social Psychology Bulletin* 35(12):1661–1671.

Pantic, M. 2009. Machine analysis of facial behaviour: Naturalistic and dynamic behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1535):3505–3513.

Park, G.; Schwartz, H. A.; Eichstaedt, J. C.; Kern, M. L.; Kosinski, M.; Stillwell, D. J.; Ungar, L. H.; and Seligman, M. E. 2014. Automatic Personality Assessment through Social Media Language. *Journal of Personality and Social Psychology* 108(6):934–952.

Pennacchiotti, M., and Popescu, A.-M. 2011. A machine learning approach to twitter user classification. In *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*, ICWSM, 281–288.

Penton-Voak, I. S.; Pound, N.; Little, A. C.; and Perrett, D. I. 2006. Personality Judgments from Natural and Composite Facial Images: More Evidence for a "Kernel of Truth" in Social Perception. *Social Cognition* 24(5):607–640.

Preoţiuc-Pietro, D.; Volkova, S.; Lampos, V.; Bachrach, Y.; and Aletras, N. 2015. Studying User Income through Language, Behaviour and Affect in Social Media. *PLoS ONE* 10(9).

Preoţiuc-Pietro, D.; Lampos, V.; and Aletras, N. 2015. An Analysis of the User Occupational Class through Twitter Content. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, ACL, 1754–1764.

Rao, D.; Yarowsky, D.; Shreevats, A.; and Gupta, M. 2010. Classifying Latent User Attributes in Twitter. In *Proceedings of the 2nd International Workshop on Search and Mining User-generated Contents*, SMUC, 37–44.

Redi, M.; Quercia, D.; Graham, L. T.; and Gosling, S. D. 2015. Like Partying? Your Face Says It All. Predicting the Ambiance of Places with Profile Pictures. In *Proceedings of the 9th International AAAI Conference on Weblogs and Social Media*, ICWSM, 347–356.

Rout, D.; Preoţiuc-Pietro, D.; Kalina, B.; and Cohn, T. 2013. Where's @wally: A Classification Approach to Geolocating Users based on their Social Ties. HT, 11–20.

San Pedro, J., and Siersdorfer, S. 2009. Ranking and classifying attractiveness of photos in folksonomies. In *Proceedings of the 18th International Conference on World Wide Web*, WWW, 771–780.

Sap, M.; Park, G.; Eichstaedt, J.; Kern, M.; Ungar, L.; and Schwartz, H. A. 2014. Developing Age and Gender Predictive Lexica over Social Media. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, EMNLP, 1146–1151.

Schwartz, H. A.; Eichstaedt, J. C.; Kern, M. L.; Dziurzynski, L.; Ramones, S. M.; Agrawal, M.; Shah, A.; Kosinski, M.; Stillwell, D.; Seligman, M. E.; et al. 2013. Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS ONE* 8(9).

Subramanian, R.; Yan, Y.; Staiano, J.; Lanz, O.; and Sebe, N. 2013. On the Relationship between Head Pose, Social Attention and Personality Prediction for Unstructured and Dynamic Group Interactions. In *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*, ICMI, 3–10.

Terry, R. L., and Kroger, D. L. 1976. Effects of Eye Correctives on Ratings of Attractiveness. *Perceptual and Motor Skills* 42(2):562–562.

Van Der Heide, B.; D'Angelo, J. D.; and Schumaker, E. M. 2012. The Effects of Verbal versus Photographic Self-presentation on Impression Formation in Facebook. *Journal of Communication* 62(1):98–116.

Volkova, S.; Wilson, T.; and Yarowsky, D. 2013. Exploring Demographic Language Variations to Improve Multilingual Sentiment Analysis in Social Media. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, EMNLP, 1815–1827.

Wei-ning, W.; Ying-lin, Y.; and Sheng-ming, J. 2006. Image Retrieval by Emotional Semantics: A Study of Emotional Space and Feature Extraction. In *IEEE International Conference on Systems, Man and Cybernetics*, SMC, 3534–3539.

Wexner, L. B. 1954. The Degree to which Colors (hues) are Associated with Mood-tones. *Journal of Applied Psychology* 38(6):432.

You, Q.; Luo, J.; Jin, H.; and Yang, J. 2015. Robust image sentiment analysis using progressively trained and domain transferred deep networks. In *The Twenty-Ninth AAAI Conference on Artificial Intelligence*, AAAI, 381–388.

Zhou, E.; Fan, H.; Cao, Z.; Jiang, Y.; and Yin, Q. 2013. Extensive Facial Landmark Localization with Coarse-to-fine Convolutional Network Cascade. In *Proceedings of the 2013 IEEE International Conference on Computer Vision, Workshops*, ICCV, 386–391.

Zou, H., and Hastie, T. 2005. Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society, Series B* 67.