Automatic Personality and Interaction Style Recognition from Facebook Profile Pictures

Fabio Celli
University of Trento
via sommarive 5
Trento, Italy
fabio.celli@unitn.it

Elia Bruni
University of Bozen-Bolzano
Dominikanerplatz 3
Bozen-Bolzano, Italy
elia.bruni@unibz.it

Bruno Lepri
FBK
via sommarive 18
Trento, Italy
lepri@fbk.eu

ABSTRACT

In this paper, we address the issue of personality and interaction style recognition from profile pictures in Facebook. We recruited volunteers among Facebook users and collected a dataset of profile pictures, labeled with gold standard self-assessed personality and interaction style labels. Then, we exploited a bag-of-visual-words technique to extract features from pictures. Finally, different machine learning approaches were used to test the effectiveness of these features in predicting personality and interaction style traits. Our good results show that this task is very promising, because profile pictures convey a lot of information about a user and are directly connected to impression formation and identity management.

Categories and Subject Descriptors

5.5 [**Emotional and Social Signals in Multimedia**]: Novel methods for the classification and representation of interactive social and/or emotional signals

General Terms

Facebook profiling personality pictures algorithms

Keywords

personality recognition, Facebook, data mining, machine learning, feature extraction

1. INTRODUCTION

People spend a considerable amount of effort in order to form and to manage impressions, especially in the initial stage of social interactions [10]. Nowadays, this fundamental process has been modified by the usage of new communication technologies. Social networking technologies, such as Facebook, offer new ways for self-presentations. Several studies reported that Facebook users engage in actively creating, maintaining and modifying an image of selves by adjusting their profiles, including descriptions and pictures, joining groups and displaying their likes and dislikes [12]. Hence, the Facebook profile page can be considered as a mediated representation of the Facebook user. Although users may be tempted to enhance their self-presentations, friends who are both offline and online keep Facebook users' self-presentations in check. Indeed, misrepresentation on profile pages can have serious offline consequences. Therefore, the online profile usually reflect the offline profile, although slightly enhanced [33]. Moreover, social psychology research has also highlighted that personality plays an important role in the way people manage the images they convey in self-presentations [17]. For this reason, some recent studies reported that users can make accurate personality impressions from the information displayed in social network user profiles [33], and have investigated the specific features from user profiles and photos that are more useful to create personality impressions [8]. For example, the results obtained by Hall *et al.* [13] indicate that observers could accurately estimate extraversion, agreeableness and conscientiousness of unknown profile owners. Interestingly, profile pictures were useful for estimating extraversion and agreeableness. Again, Utz [31] associated user extraversion with photo expressiveness. Finally, Van deer Heide *et al.* [32] have shown that in the context of Facebook, photos may have more impact on judgments of extraversion than textual self-disclosures. Based on these previous findings, we propose to automatically predict personality traits and interaction styles from Facebook profile images. Regarding personality, we resort to the big five model [7], a multifactorial approach which owes its name to the five traits it takes as constitutive of people's personality: extraversion, emotional stability/neuroticism, agreeableness, conscientiousness, and openness to experience. Moreover, we also target the prediction of people's interaction traits. To this end, we exploited the interpersonal circumplex, a model for organizing and assessing interpersonal traits [19]. Specifically, the interpersonal circumplex is defined by two orthogonal axes, a vertical axis of dominance or agency and a horizontal axis of affect or warmth. The main contributions of the paper are as follows: 1) We collect a dataset of profile images from 100 Facebook users and we annotate them with gold standard self-assessed personality and interaction style labels; 2) we exploit a bag-of-visual-word representation of images in order to extract relevant visual features; 3) we propose and validate a supervised learning approach, based on Support Vector Machines and logistic regression, to the automatic recognition of personality and interaction style traits from visual features extracted from profile pictures. The rest of the paper is structured as follows: Section 2 describes the relevant previous work on personality recognition, with a particular focus on personality recognition from social media data; Section 3 describes the data collected and used for our experiments; while the definition of the research task and the detailed information on the methodology (feature extraction, feature selection and classification) is provided in Section 4. In Section 5 we present the results and in Section 6 we draw our conclusions.

ACMMM 2014, Orlando, Florida

http://dx.doi.org/10.1145/2647868.2654977.

2. RELATED WORK

Pioneering work addressing automatic recognition of personality was done, among others, by Mairesse *et al.* [21], who investigated systematically the usefulness of different sets of acoustic and textual features, extracted from self-reports and observed data. They reported that the best prediction performance, using text as feature, is on openness to experience. In more recent years, the interest in personality recognition has grown and several studies have started exploring the wealth of behavioral data made available by cameras, microphones [26, 2, 24, 18, 1], wearable sensors [25, 15], and mobile phones [30, 6, 9]. Again, researchers have also focused on personality prediction from small corpora of social network data, like Twitter and Facebook, exploiting either linguistic features in status updates, social features such as friends count, and daily activity [11, 27, 5]. Our paper also uses data from a social network; however it adopts a novel approach using visual features extracted from Facebook profile pictures in order to classify subjects' traits. Interestingly, two recent works addressed the automatic recognition of Big Five traits from self-presentation videos registered in a lab setting [2] and from Youtube conversational vlogs [3]. However, they used different data compared with ones proposed in the current paper. Another novelty introduced in our paper is the classification of interaction style traits and as defined by the interpersonal circumplex [23]. To the best of our knowledge, no previous works in computational domain have targeted the task of interaction style classification from pictures using the interpersonal circumplex.

3. DATA COLLECTION

As we mentioned in section 1, in our study we resorted to the big five factor model, a very popular and widely used model for personality assessments [7], and to the interpersonal circumplex, a model for the assessment of interpersonal behavior that has been used in psychology and psychopathology to identify interpersonal problems and interpersonal values [14, 19]. While the big five factor model measures the individual tendencies, the interpersonal circumplex measures the attitude towards others. The big five factor

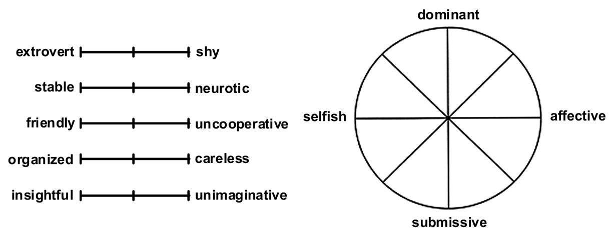

Figure 1: big 5 and interpersonal circumplex

model defines personality along 5 bipolar scales that can be turned into binary classes: extroversion (sociable *vs* shy); emotional stability (secure *vs* neurotic); agreeableness (friendly *vs* ugly); conscientiousness (organized *vs* careless); openness to experience (insightful *vs* unimaginative). The interpersonal circumplex instead defines two orthogonal axes, dominance and affect, as represented in figure 1. We recruited 112 volunteers among Facebook users, most of them among friends and friends-of-friends, in order to have a better control over the data. Using Facebook graph APIs, we collected a dataset of their profile images. Then, we asked to the participants to fill two surveys about individual traits, in order to have gold standard self-assessed personality and interaction style labels. The two surveys are: 1) the big five personality test (BFI-10) [28] and 2) the interpersonal circumplex (IPIP-IPC-32) for interaction styles [23]. We collected the assessments online and removed the data of users with incomplete tests. In the end, we collected the personality and interaction styles of about 100 Facebook users. Both tests are designed to be used when time is limited and take about 5 minutes in total. This allows us to have the full attention of the users.

4. METHODOLOGY

The goal of the paper is to automatically recognize personality traits and interaction styles from Facebook profile images. In the following subsections we will describe the feature extraction and the classification tasks.

4.1 Feature Extraction

In order to extract features from profile images we exploited a Bag-of-Visual-Word (BoVW) representation of images, a popular technique for object recognition inspired by the traditional bag-of-words technique in NLP [29]. Bag-of-Words is a dictionary-based

Figure 2: SIFT features on a profile image

method to represent a document in terms of a cluster of words which are entries of the dictionary. BoVW transfers this idea to images, describing them as a collection of discrete regions called **visual words**. BoVW downplays word arrangement (i.e. the spatial information in the image) and describes the image as a histogram of the frequency of visual words. The set of visual words forms a visual vocabulary, which is constructed by clustering a large corpus of low-level features. The low-level features are handcrafted attributes designed to find local image structures in a repeatable fashion and to encode them in ways that are robust to image transformations such as translation, rotation and affine deformation. There is a large variety of low-level features to use in the BoVW pipeline. In our case, we use the Scale Invariant Feature Transform (SIFT), one of the most popular and effective feature extraction technique used for object recognition [20]. SIFT is invariant to several image transformations such as image scale, orientation, noise and it is also partially invariant to changes in illumination. A SIFT descriptor, as depicted in figure 2 is obtained by concatenating the contents of 4×4 sampling subregions which are explored around each pixel-wise step across the image. For each of the 16 samples, 8 gradient orientations are calculated to obtain a 128-dimensional feature vector. Our resulting BoVW feature vectors have 4096 dimensions which were subsequently reduced in dimensionality to about 90 dimensions by using Singular Value Decomposition [22]. To perform the entire visual pipeline we use VSEM, an open library for visual semantics [4]. VSEM contains a large variety of off-the-shelf low level features, but it does not allow to backtrack to feature types.

Since this is a pilot experiment, we used the largest possible number of features, keeping the number of instances very limited in order to run a manual qualitative analysis (see section 5) of the profile pictures. In the future we will repeat the experiment with a very large dataset.

4.2 Classification and Evaluation

First of all we ran a t-test to compare personality trait scores of our small set to the scores of a set of about 100000 Facebook users randomly sampled from myPersonality [16]. It resulted that the probability of having differences in scores due to chance in a small and large dataset is $p < 0.001$ for all traits. Details are reported in table 1. We could not run a t-test for dominance and affect since there are not such dimensions in myPersonality. We formulated the automatic recognition of personality traits and interaction styles from profile pictures as 7 binary classification tasks, one per each trait of personality and interaction style. To this purpose, we balanced the classes of each personality trait with a median split, turning scores into classes and reducing each task to a binary classification. We have few instances of unconscientious and unimaginative people, due to a bias in the participation of Facebook users. Because of the novelty of the task, we experimented with different classification algorithms: naïve bayes (nb), support vectors (svm), decision trees (dt), logistic regression(lr), nearest neighbors (nn), rule learner (rip). We ran a 60% training and 40% test split evaluation (f1-s) and provide f1-scores. We computed two baselines for the calssification tasks: the majority baseline (maj) and the averaged majority baseline (avg maj). The majority baseline is close to 0.5, because we balanced the two classes for each task. The averaged majority baseline is the average of the performance obtained classifying all the instances before in one class and then in the other. Both the baselines were computed using f1-score. We argue that averaged majority is the meaningful baseline for these classification tasks, given the bipolar nature of each personality and interaction style dimension.

trait	t	df	p-value
ext.	11.4	99.1	< 0.001
stab.	-5.1	99.1	< 0.001
agree.	12.6	99.1	< 0.001
consc.	12.1	99.1	< 0.001
open.	8.2	99	< 0.001

Table 1: Detailed results of t-test for the 5 personality traits.

task	alg	inst	maj	a-maj	f1-s	acc
ext.	nb	33	.51	.35	**.615**	61.5%
stab.	rip	57	.508	.343	**.609**	60.8%
agree.	svm	30	.5	.333	**.667**	66.6%
consc.	svm	9	.55	.397	**.733**	75%
open.	nn	11	.545	.385	**.732**	74.5%
dom.	lr	51	.509	.344	**.615**	65%
aff.	lr	21	.509	.344	**.619**	62.5%

Table 2: results of classification. alg=algorithm, inst=number of instances, maj=majority, a-maj=averaged majority, f1-s=f1-score with 60% training and 40% test split evaluation, acc=accuracy.

5. DISCUSSION

Results, reported in table 2, show that the information encoded in profile pictures can be successfully exploited for classification of personality and interaction style traits of users in Facebook. Nevertheless there are differences between traits: apart from openness and conscientiousness, that have few instances and require more experiments on a larger scale, agreeableness and extraversion achieve the best performances among personality traits, while dominance and affect achieve performances slightly above f1=.6. Emotional stability is the most difficult trait to predict. This is consistent with the findings that extroversion and agreeableness are accurately estimated by human raters [13] and that extroversion is related to photo expressiveness [31]. More experiments are required to test the finding that observers could accurately estimate conscientiousness from pictures. Unfortunately, the Bag-of-Visual-Words

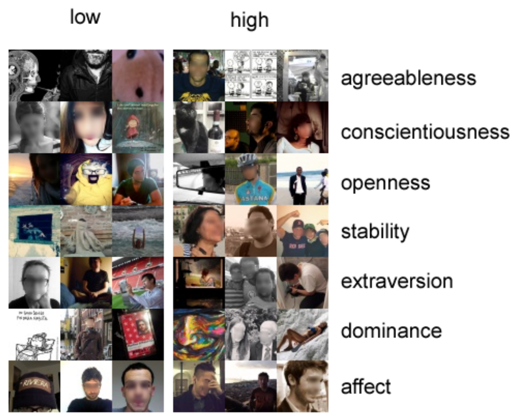

Figure 3: Examples of profile pictures per trait.

method adopted for the extraction of features from images is not suitable for the interpretation of the pictures' attributes. Nevertheless, we performed clustering on the classified data, using a simple k-means algorithm to automatically retrieve the more similar pictures that were correctly classified, and we checked their similarities manually. This revealed that extrovert and stable people tend to have pictures where they are smiling and they appear with other people. Introverts tend to appear alone, neurotics tend to have images without humans and close-up faces. Not agreeable persons tend to use profile pictures with few colors. Profile pictures of people that is not open to experience tend to have strong spots of light and sometimes darkened figures. Affective people tend to smile in profile pictures. Dominant people tend to have very bright, visible or attractive pictures. Conscientiousness seems to be predictable by eye gazing direction. We included some examples the pictures in the paper, anonymized with out-of-focus spots, see figure 3.

6. CONCLUSION AND FUTURE WORK

In this paper, we successfully exploited a Bag-of-Visual-Words technique to automatically predict personality and interaction styles from profile pictures in Facebook. This is a new and promising task, because profile pictures convey a lot of information about a user and are directly connected to her/his identity. The good, although preliminary, results of the computational classification are in line with the previous findings in the literature about impression formation and personality and with alternative approaches to personality classification. In the future, we would like to test the proposed approach on a larger dataset of Facebook profile pictures.

7. REFERENCES

[1] O. Aran and D. Gatica-Perez. Cross-domain personality prediction: from video blogs to small group meetings. In *Proceedings of the 15th ACM on International conference on multimodal interaction*, pages 127–130. ACM, 2013.

[2] L. M. Batrinca, N. Mana, B. Lepri, F. Pianesi, and N. Sebe. Please, tell me about yourself: automatic personality assessment using short self-presentations. In *Proceedings of the 13th international conference on multimodal interfaces*, pages 255–262. ACM, 2011.

[3] J. Biel and D. Gatica-Perez. The youtube lens: Crowdsourced personality impressions and audiovisual analysis of vlogs. *Multimedia, IEEE Transactions on*, 15(1):41–55, 2013.

[4] E. Bruni, U. Bordignon, A. Liska, J. Uijlings, and I. Sergienya. Vsem: An open library for visual semantics representation. In *Proceedings of ACL*, Sofia, Bulgaria, 2013.

[5] F. Celli and L. Polonio. Relationships between personality and interactions in facebook. In *Social Networking: Recent Trends, Emerging Issues and Future Outlook*, pages 41–54. Nova Science Publishers, Inc., 2013.

[6] G. Chittaranjan, J. Blom, and D. Gatica-Perez. Mining large-scale smartphone data for personality studies. *Personal and Ubiquitous Computing*, 17(3):433–450, 2013.

[7] P. T. Costa and R. R. McCrae. The revised neo personality inventory (neo-pi-r). *The SAGE handbook of personality theory and assessment*, 2:179–198, 2008.

[8] S. Counts and K. B. Stecher. Self-presentation of personality during online profile creation. In *ICWSM*, 2009.

[9] Y.-A. de Montjoye, J. Quoidbach, F. Robic, and A. S. Pentland. Predicting personality using novel mobile phone-based metrics. In *Social Computing, Behavioral-Cultural Modeling and Prediction*, pages 48–55. Springer, 2013.

[10] E. Goffman. *The presentation of self in everyday life*. Garden City, NY Double Day, 1959.

[11] J. Golbeck, C. Robles, and K. Turner. Predicting personality with social media. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 253–262. ACM, 2011.

[12] A. L. Gonzales and J. T. Hancock. Identity shift in computer-mediated environments. *Media Psychology*, 11(2):167–185, 2008.

[13] J. A. Hall, N. Pennington, and A. Lueders. Impression management and formation on facebook: A lens model approach. *New Media & Society*, page 1461444813495166, 2013.

[14] L. M. Horowitz. *Interpersonal foundations of psychopathology.* American Psychological Association, 2004.

[15] K. Kalimeri, B. Lepri, and F. Pianesi. Going beyond traits: Multimodal classification of personality states in the wild. In *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*, ICMI '13, pages 27–34, New York, NY, USA, 2013. ACM.

[16] M. Kosinski, Y. Bachrach, P. Kohli, D. Stillwell, and T. Graepel. Manifestations of user personality in website choice and behaviour on online social networks. *Machine Learning*, pages 1–24, 2013.

[17] N. C. Krämer and S. Winter. Impression management 2.0: The relationship of self-esteem, extraversion, self-efficacy, and self-presentation within social networking sites. *Journal of Media Psychology*, 20(3):106, 2008.

[18] B. Lepri, R. Subramanian, K. Kalimeri, J. Staiano, F. Pianesi, and N. Sebe. Connecting meeting behavior with extraversion: A systematic study. *Affective Computing, IEEE Transactions on*, 3(4):443–455, 2012.

[19] K. D. Locke. Circumplex scales of interpersonal values: Reliability, validity, and applicability to interpersonal problems and personality disorders. *Journal of Personality Assessment*, 75(2):249–267, 2000.

[20] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 2004.

[21] F. Mairesse, M. A. Walker, M. R. Mehl, and R. K. Moore. Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research*, 30(1):457–500, 2007.

[22] C. Manning, P. Raghavan, and H. Schütze. *Introduction to Information Retrieval*. Cambridge University Press, Cambridge, UK, 2008.

[23] P. M. Markey and C. N. Markey. A brief assessment of the interpersonal circumplex the ipip-ipc. *Assessment*, 16(4):352–361, 2009.

[24] G. Mohammadi and A. Vinciarelli. Automatic personality perception: Prediction of trait attribution based on prosodic features. *Affective Computing, IEEE Transactions on*, 3(3):273–284, 2012.

[25] D. O. Olguın, P. A. Gloor, and A. S. Pentland. Capturing individual and group behavior with wearable sensors. In *Proceedings of the 2009 aaai spring symposium on human behavior modeling, SSS*, volume 9, 2009.

[26] F. Pianesi, N. Mana, A. Cappelletti, B. Lepri, and M. Zancanaro. Multimodal recognition of personality traits in social interactions. In *Proceedings of the 10th international conference on Multimodal interfaces*, pages 53–60. ACM, 2008.

[27] D. Quercia, R. Lambiotte, D. Stillwell, M. Kosinski, and J. Crowcroft. The personality of popular facebook users. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pages 955–964. ACM, 2012.

[28] B. Rammstedt and O. P. John. Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german. *Journal of Research in Personality*, 41(1):203–212, 2007.

[29] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In *Proceedings of ICCV*, pages 1470–1477, Nice, France, 2003.

[30] J. Staiano, B. Lepri, N. Aharony, F. Pianesi, N. Sebe, and A. Pentland. Friends don't lie: inferring personality traits from social network structure. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 321–330. ACM, 2012.

[31] S. Utz. Show me your friends and i will tell you what type of person you are: How one's profile, number of friends, and type of friends influence impression formation on social network sites. *Journal of Computer-Mediated Communication*, 15(2):314–335, 2010.

[32] B. Van Der Heide, J. D. DŠAngelo, and E. M. Schumaker. The effects of verbal versus photographic self-presentation on impression formation in facebook. *Journal of Communication*, 62(1):98–116, 2012.

[33] S. Vazire and S. D. Gosling. e-perceptions: personality impressions based on personal websites. *Journal of personality and social psychology*, 87(1):123, 2004.