通过社交媒体上的头像分析用户的幸福感

**Abstract：**

The profile images users post to their social media, as a stable content, are driven in part by their own characteristics，including gender, age and personality. In this paper, we have investigated individual differences in profile picture choice between wellbeing traits. We concluded that each wellbeing trait has a specific type of profile picture posting. For example, users that are high in positive emotion prefer to use pictures with soft, delicate and light style as their profile image. while users high in engagement do not like to use pictures with object texture, surface details and texture as their profile image.

**Introduction：**

In recent years，researchers have tried to make use of profile images to study users’ online behaviors and psychological states. On the one hand, researchers try to figure out factors driving users to choose their profile images. On the other hand, they attempted to obtain the characteristics of users from their profile images. Several studies have examined the differences of users' profile images are related to age, gender and personality. For example, literature [2] shows that users high in openness prefer more aesthetic photos. In general, appealing images tend to have increased contrast, sharpness, saturation and less blur. What attracts researchers is that users can choose profile images to enhance their physical attractiveness without being judged deceptive, this means profile images can reflect users' actual psychological states.

Wellbeing is more than simply positive emotion or mood. Psychologists, rganizations, and governments measuring wellbeing are now using multi-dimensional measures that include a range of factors including meaning in life, engagement in activities, and the state of one’s relationships, in addition to positive emotion [10]。

With the rise of social media, the way people express their feelings and opinions gradually shifts from offline to online. And the opinions and feelings expressed by users in social media can be spread and communicated quickly. Online platform allow users to create their own profiles with text, images, videos ,etc. This profiles may be different from the user's actual profile, but it can reflect Personal psychological characteristics [1]. On the one hand, Users' feeling and opinions is beneficial to the development of personalized search engines, recommendation systems and online markets. On the other hand, for the government, effective control of the focus of public opinion and emotional orientation is good to improve the decision-making mechanism and timely response to emergencies.

study [2] conclued that users with different personalities have differences in color, content and other characteristics when selecting profile image on personal social media, which also prompts us to shift our research goals to happiness and image style. However, literature [2] did not consider the abstract content contained in the image, such as atmosphere and tone, it focused on colors, facial expressions, image composition and other certain content. study [21] reveals that texture information is also important for human perception of image style. Convolution neural network was used to extract style features and content information to identify image style, this method makes better use of the poor explanatory and abstract characteristics of CNNs to obtain abstract style of image. In this paper, we used the CNNs model proposed by study[3] to extract image style(（**Detailed, Pastel, Melancholy, etc**), then we uncover their relationships with wellbeing traits from the PERMA.

PERMA[14] decomposes the wellbeing construct into five domains: Positive emotion, Engagement, Relationships, Meaning, and Accomplishment (PERMA). With these five elements of PERMA, people can have their own happiness and enjoy a vigorous life。In our study, we found high correlation between wellbeing traits and users‘ profile image style. we use existing state-of-the-art text prediction methods to estimated users’ wellbeing.

**Related work:**

study[19] demonstrated that Gender plays a role in determining the choice of profile picture。文献[20] 得出结论，性格可以在认证和识别工作中预测一些图像选择和线下行为，当然其它的预测变量对于个体特征也很重要。 conclude that personality can predict some image choices and behaviors that might be useful for future work on authentication and identification, although other predictor variables are potentially also important when considering the types of individual characteristics which might predict on- line behavior on SNSs。文献[2]发现用户头像的风格与其性格之间存在一定的关系，比如开放性较强的用户喜欢使用对比度、清晰度、饱和度更高的照片作为头像。现阶段年龄、性别、性格、幸福度的预测任务均利用了经典的自然语言方法，文献[4]建立了一个年龄、性别语料库（词典），并且使用一些带标记的数据集测试了其有效性。对于性别和年龄词典的训练，文章使用岭回归训练年龄模型，使用支持向量机训练性别模型，得到对应词汇的权重。文献[5]应用用户点赞的内容以及用户社交账号的基础数据对用户的个人信息，包括性别、年龄、宗教信仰、个人行为习惯等特征进行分析、建模。文献[7]提出了一种基于Facebook状态更新的社交媒体语言构建幸福感的统一方法。他们利用情感分析来生成用户的情感得分，并利用更新状态(推文)的情感得分和其他语言特征训练一个随机森林模型来预测用户的幸福感。文献[8]认为幸福感最终归因于人，因此他们在用户级的数据执行额外评估，发现使用消息级和用户级的多级级联模型表现最佳并且优于基于词典的幸福模型。文献[12]验证了图像风格与图像内容之间存在依赖关系，比如Macro风格的图像与图像中存在动物有着很强的相关性，而Long Exposure和Sunny风格的图像则很多包含着汽车。

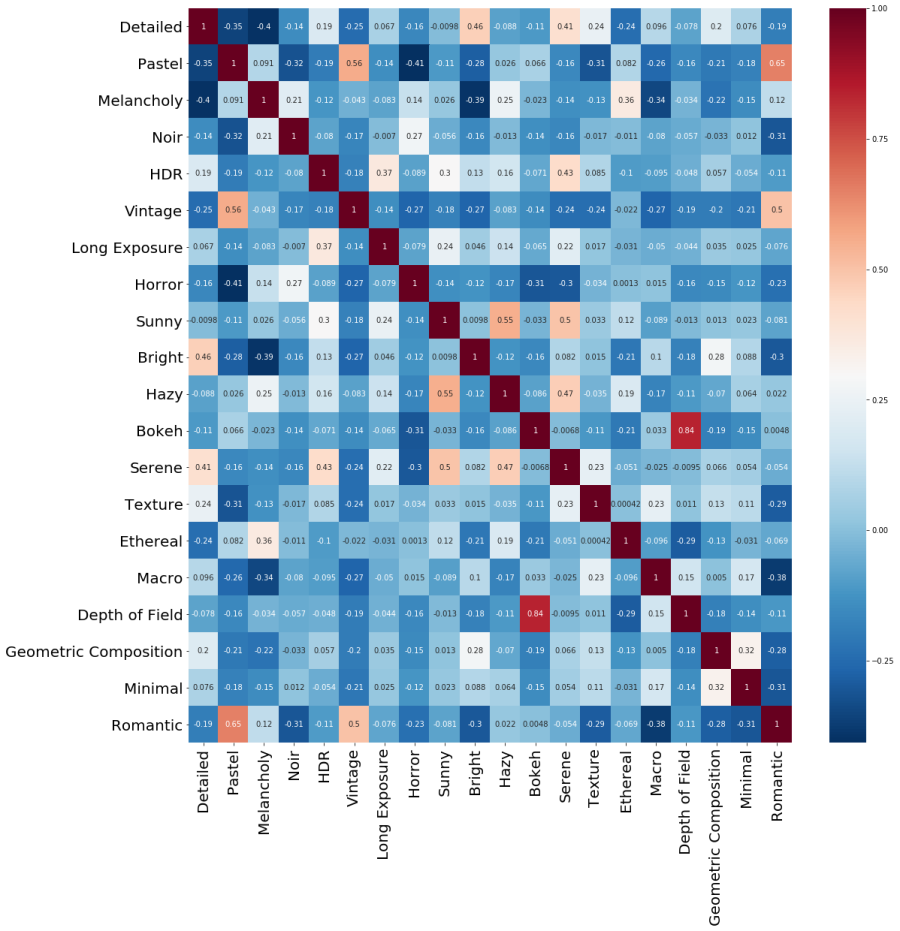
**Data**

twitter用户数据获取：

文献[11]研究英语、西班牙语和俄文tweets数据中主观语言使用的性别差异，并探讨男性和女性用户在表情符号和话题标签使用上的跨文化差异。我们考虑使用该文献中涉及到的twitter用户。我们使用python的Scrapy包爬取了1721名有效twitter用户从2010年1月1日到2018年12月31日发表的所有tweets，选取其中包含tweets数目在100条以上的用户，总共筛选出1523名用户，并获取他们的图像。我们使用DLATK[21]工具对tweets数据进行处理，包括筛选出英文tweets去除推文中的链接和@数据、重复数据。

**图像风格提取**

Different image styles bring different feelings to people. For different image styles, people can feel subtle differences, but it is quite difficult to describe them，这是因为风格涉及到的内容较多，One reason is that the style information of an image may be encoded in any component of the image including local and global, shape and color features. Another reason is that the definition of a style is often subjective, especially for photography works, which may vary from person to person. Literature [12] USES low-level statistics, color choices, composition, and content to construct the style prediction model, which has a good explanation of what is the image style. In addition, through comparative experiments, the literature found that the performance of their classifier was quite accurate compared with that of trained artificial markers. study[13] propose a CNN-based model that integrates both content (or object) information and texture information for image style recognition. And their work using the CNN-based model improves over the state-of-the-art results. In their study, three style recognition datasets are experimented: WikiPainting, Flickr Style, and AVA Style. The Flickr Style was also introduced by Karayev et al. [12]. It contains 80,000 images with 20 style labels. The labels were grouped into several categories (Optical techniques, Atmosphere, Mood,Composition styles, Color and Genre) and each category contains two to five labels. For example, the Atmosphere category has two labels, Hazy and Sunny. We used Flickr dataset established to train the style recognition model.



头像风格之间的相关性

文中，我们使用Pearson相关系数来分析图像风格与幸福度得分之间的相关性。

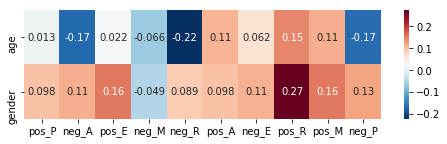
**Text analysis**

**wellbeing：**

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 (Preot¸iuc-Pietro et al. 2015), occupation (Preot¸iucPietro, Lampos, and Aletras 2015), mental illnesses (De Choudhury, Counts, and Horvitz 2013) and personality (Schwartz et al. 2013). We use the method developed by (Schwartz 2016[8]) to assign each user scores for wellbeing，According to study [9], this model divides happiness into five categories, including Positive emotion, Engagement, Relationships, Meaning, and Accomplishment, and each category has two polarities. for example, The Engagement category contains two categories: Engagement and lack of accomplishment.

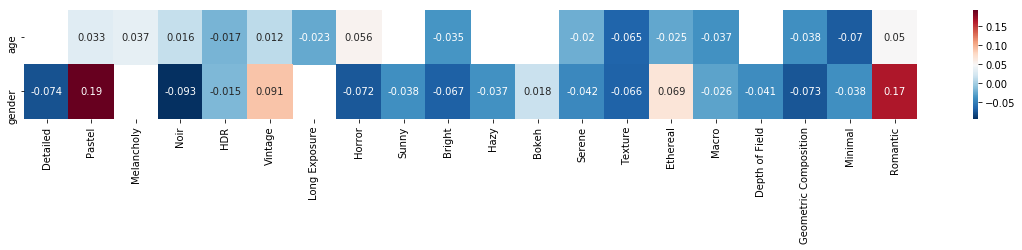
**age&gender：**

我们通过基于tweets文本分析的模型预测用户的年龄和性别，研究了在不同性别、年龄段的用户的幸福度的区别。结果显示随着年龄的增长，正向的PERMA均表现出增长的趋势，而反向的幸福感却随着年龄的增长而呈现下降的趋势。尤其对Relationship来说，女性和年龄越大的用户越容易表现出对他人的信任和肯定。其次，年龄越大的用户的Accomplishment越强。女性较于男性更容易表现出Engagement。



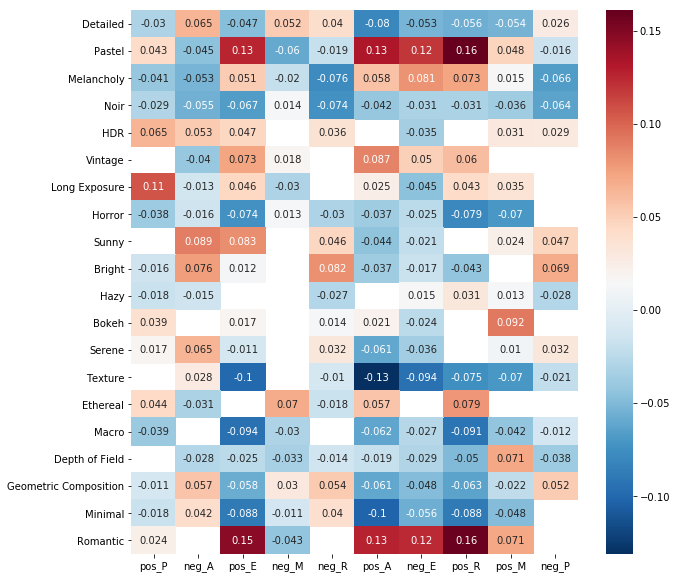
性别、年龄与wellbeing的相关性

（gender的值大于0时标识女性，age的值为连续的正值，标识用户的预测年龄）

****

性别、年龄与风格的相关性

图像风格之间的相关性，由于往往一张图片属于多个风格分类，也就是不同风格之间具有相同的特征，所以风格之间的相关性也存在较大的差异。比如Bokeh与Depth of Filed表现出极强的相关性，原因在于前者由镜头产生的图像离焦部分的模糊的美学效果，后者在图像中为凸显最近物体而模糊背景物体，两者均存在模糊的图像内容。

**分析**

**头像风格与幸福感PERMA之间的相关性**

**Positive Emotion：**

**Positive Emotion**是人们对幸福感研究最多的一个方面，积极情绪 [14]includes positively valenced emotions such as joy, contentment, and excitement。正向positive emotion得分较高的用户与Long Exposure、HDR、Pastel(上面三种均属于Optical techniques大类下)、Ethereal具有较强的相关性，Long Exposure 风格分类下是体现运动物体在某一时刻表现出来的状态，而Pastel和Ethereal风格的图片则体现的是柔和、细腻、轻盈的颜色或者是色调。(这一类图像的特点是模糊、飘渺、空灵)也就是说，表现出积极情绪的用户更喜欢使用更加柔和、细致的图片作为个人头像。文献[12]在研究中发现Long Exposure风格的图像中经常出现汽车内容，而Macro风格则与图像中出现动物具有较强的相关性。Positive Emotion与Macro具有最强的负相关性，Macro风格分类下的图片为使用微距镜头拍摄的比正常物体大很多的图像，比如较正常大小5倍的蜜蜂的照片。其次，Positive Emotion还与Horror呈现负相关，该风格分类下的图片为恐怖、血腥以及怪异的样式。也就是说，**Positive Emotion** 的用户更倾向于柔和、细腻、清晰、包含运动内容的头像。排斥恐怖、血腥、内容异常的图像。

对消极情绪的用户而言，他们头像的风格与Bright、Geometric Composition以及Sunny表现出较强的相关性。Bright风格的图片呈现明亮、强烈、迷幻的色彩，而Geometric Composition风格则描述了圆形、三角形、矩形等对称物体以及局部重复图案。值得注意的是，对于表现消极情绪的用户，他们往往还喜欢使用包含太阳的图片作为头像(Sunny风格的图像中包含的内容大多数有太阳)。

**Engagement：**

**Engagement**是包括行为、认知和情感的多维结构。它可以指对参与活动表现出的热情和兴趣、对工作的投入感和奉献、以及对手头任务的专注度等[15]。Romantic风格与Engagement表现出很强的正相关性。Romantic的一个描述"A mysterious or fascinating quality or appeal, as of something adventurous, heroic, or strangely beautiful: “These fine old guns often have a romance clinging to them” (Richard Jeffries)."。Pastel和Sunny则表现出较强的相关性。Engagement的用户不喜欢使用图片包含物体纹理、表面细节、质地的图片作为头像。相反的，这类用户喜欢使用场景较为宏大的图像。

**Relationship：**

**Relationship**指的是信任他人、肯定他人的存在、不仅接受来自社会的给予还能够给予他人[16]。大量证据表明了积极的**Relationship**对健康、长寿和其他重要生活品质的重要性。relationship感强的用户更偏向于使用Romantic和Pastel这类风格的图像。这类用户的头像风格往往与黑白色调、氛围沉郁的头像呈现负相关。而相反的，缺乏Relationship的用户的头像更喜欢使用Bright、Geometric Composition风格的图像。Relationship和Positive Emotion两种特征与风格表现出相似的特性，不同的是**Relationship**与Romantic表现更强的相关性。

**Meaning:**

**Meaning** 诠释了生命的目的和重要性，以及对生活的理解[18]。Meaning得分较高的用户喜欢Broken和Depth of Field风格，两种风格的Pearson相关性指数高达0.84。这两种风格的共同点在于图像中的物体有不同的层次结构，比如物体的远近、清晰和模糊。另外Meaning与Minimal表现出负相关。

对于缺乏Meaning的用户，更偏向于使用构图细致、透明或者是半透明、空灵的Ethereal风格。

**Accomplishment：**

**Accomplishment**通常根据奖励、荣誉和其他客观成就标志来定义。就个人成就感而言，它包括掌控力、感知能力和目标达成感。Accomplishment与Romantic表现出较强的正相关性，与Texture、Detailed负相关。也就是说，对于accomplishment越强的用户就越不喜欢使用包含细节、纹理清晰、物体的结构和质地等内容，他们更偏向于具有较高质量、异常美丽的图像。

缺乏成就感的人则偏好Sunny和Bright风格的头像，这一点与我们普通的认知有些相悖，在Sunny和Bright这两个风格中的图片包含大量强烈、大胆、几近疯狂的色彩，就如同阳光一样。与之对应的，该群体不喜欢Pastel和Melancholy这种颜色轻柔、细腻、色调柔和的风格的头像。

**总结：**

我们使用卷积神经网络提取用户图像的风格，用Pearson相关度量图像风格和用户幸福度之间的关系。该研究揭示了用户的幸福度在一定程度上影响着用户图像的选择。能够对相关心理学研究提供线索和依据。

该研究是首次分析了图像风格与用户wellbeing之间的相关性，发现不同wellbeing的用户所选用的头像存在一定的差异，positive emotion的用户更喜欢使用柔和、细腻风格的头像。engagement的用户则不喜欢图像中包含纹理、细节，而是偏向于场景较大的头像。Relationship 则与PE表现出相似的趋向，negative emotion和缺乏relationship的用户反而喜欢使用色彩更加鲜艳、明亮，包含几何内容的头像。meaning的用户更喜欢具有层次感的头像。Accomplishment的用户则偏好Romantic风格的图像。

**References**

[1] Profile pictures on social media: Gender and regional differences.

[2] Analyzing personality through social media profile picture choice

[3] Convolution neural network with two pathways for image style recognition

[4] Developing Age and Gender Predictive Lexica over Social Media

[5] Private traits and attributes are predictable from digital records of human behavior

[6] <https://www.faceplusplus.com>

[7] Building a profile of subjective wellbeing for social media users

[8] Predicting individual wellbeing through the language of social media

[9] M. J. C. Forgeard, E. Jayawickreme, M. Kern and M. E. P. Seligman, International Journal ofWellbeing 1, 49 (2011).

[10] A Visionary New Understand of Happiness and Wellbeing

[11] Exploring Demographic Language Variations to Improve Multilingual Sentiment Analysis in Social Media

[12] Recognizing Image Style

[13] Convolution Neural Networks With Two Pathways for Image Style Recognition

[14] Doing the right thing : Measuring wellbeing for public policy

[15] M. Csikszentmihalyi, Creativity: Flow and the psychology of discovery and invention (Harper Perennial, 1997)

[16] The Oxford Handbook of Health Psychology

[17] Applied Psychology: Health and Wellbeing (2012)

[18] Man’s search for ultimate meaning

[19]Profile pictures on social media: Gender and regional differences

[20] A picture tells a thousand words: What Facebook and Twitter images convey about our personality

[21] DLATK: Differential Language Analysis ToolKit