Personality-Assisted Multi-Task Learning for Generic and Personalized Image Aesthetics Assessment

Leida Li, Hancheng Zhu, Sicheng Zhao, *Senior Member, IEEE*, Guiguang Ding, and Weisi Lin, *Fellow, IEEE*

***Abstract*— Traditional image aesthetics assessment (IAA) approaches mainly predict the average aesthetic score of an image. However, people tend to have different tastes on image aesthetics, which is mainly determined by their subjective preferences. As an important subjective trait, personality is believed to be a key factor in modeling individual's subjective preference. In this paper, we present a personality-assisted multi-task deep learning framework for both generic and personalized image aesthetics assessment. The proposed framework comprises two stages. In the first stage, a multi-task learning network with shared weights is proposed to predict the aesthetics distribution of an image and Big-Five (BF) personality traits of people who like the image. The generic aesthetics score of the image can be generated based on the predicted aesthetics distribution. In order to capture the common representation of generic image aesthetics and people's personality traits, a Siamese network is trained using aesthetics data and personality data jointly. In the second stage, based on the predicted personality traits and generic aesthetics of an image, an inter-task fusion is introduced to generate individual's personalized aesthetic scores on the image. The performance of the proposed method is evaluated using two public image aesthetics databases. The experimental results demonstrate that the proposed method outperforms the state-of-the-arts in both generic and personalized IAA tasks.**

***Index Terms*— Image aesthetics assessment, generic and personalized image aesthetics, personality traits, multi-task deep learning, Siamese network**

Manuscript received July 18, 2019; revised January 10, 2020; accepted January 16, 2020. Date of publication January 27, 2020; date of current version February 4, 2020. This work was supported in part by the Natural Science Foundation of Jiangsu Province under Grant BK20181354, in part by the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant KYCX19_2142, in part by the Postgraduate Research & Practice Innovation Program of China, University of Mining and Technology, under Grant KYCX19_2142, in part by the National Natural Science Foundation of China under Grant 61701273, Grant 61771473, and Grant 61379143, in part by the Six Talent Peaks High-level Talents in Jiangsu Province under Grant XYDXX-063, and in part by the Qing Lan Project. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Joao M. Ascenso. *(Corresponding authors: Hancheng Zhu; Sicheng Zhao.)*

Leida Li is with the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China, and also with the School of Artificial Intelligence, Xidian University, Xi'an 710071, China (e-mail: reader1104@hotmail.com).

Hancheng Zhu is with the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China (e-mail: zhuhancheng@cumt.edu.cn).

Sicheng Zhao is with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94710 USA (e-mail: schzhao@gmail.com).

Guiguang Ding is with the School of Software, Tsinghua University, Beijing 100084, China (e-mail: dinggg@tsinghua.edu.cn).

Weisi Lin is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: wslin@ntu.edu.sg).

Digital Object Identifier 10.1109/TIP.2020.2968285

I. Introduction

WITH the prevalence of mobile internet, people have become used to sharing their daily lives using images in social networks. Many social networks, such as Flickr and Facebook, are becoming more reliant on using images to engage users. Hence, it is desirable to automatically assess the aesthetic quality of images. Image aesthetics assessment (IAA) aims at capturing people's aesthetic perception of images (e.g., through photographic rules) [1], which has many potential applications, e.g., photo ranking [2], photo aesthetic recommendation [3], photo enhancement [4], [5] and image retrieval [6]. Since aesthetics is a quite subjective attribute of an image [7], it is still a challenging task for IAA. There are two main categories for IAA: generic image aesthetics assessment (GIAA) and personalized image aesthetics assessment (PIAA) [8]. While GIAA is to infer the generic aesthetics of images perceived by the majority of people [1], PIAA aims at predicting individual's unique aesthetic perception of images [8]. Most of the previous studies have been focused on GIAA [2], [6], [9]–[20]. Meanwhile, with the ever-growing demand for users' unique aesthetics, several PIAA approaches [8], [21]–[24] have also been proposed toward personalized image recommendation [25].

In the last two decades, the research of image aesthetics mainly focuses on GIAA. Most early works map elaborately designed hand-crafted features, typically based on photographic rules such as the vivid color, the rule of thirds, and the symmetry [11]–[14], to evaluate generic image aesthetics. However, the conventional hand-crafted features can only incorporate limited aesthetic rules to model people's complex and abstract visual aesthetic perception, which could be problematic. Benefiting from the powerful learning ability of deep neural network based on big data [26], a mushrooming number of works have been proposed to learn deep representations for generic image aesthetics [2], [6], [15]–[20], [27]–[38]. These GIAA approaches focus on using the objective information of image, such as object, scene and attribute, for predicting humans' average aesthetic score of an image. While these computational approaches have achieved notable success, humans' aesthetic ratings on an image may

Fig. 1. Example images and their aesthetic scores rated by five different users as well as the corresponding average scores from FLICKR-AES database. The aesthetic scores are rated from 1 to 5.

vary significantly depending on their unique visual preferences [39]. As an example, Fig. 1 shows four images and the associated aesthetic scores rated by five different users as well as the corresponding average scores from the FLICKR-AES database [8]. As illustrated in Fig. 1, the aesthetic scores of an image differ significantly among different users. Although the higher the average score, the more appealing the image, it cannot effectively represent the potential difference of the human perception of aesthetics [40]. In contrast, aesthetics distribution prediction is a more reasonable way to evaluate the diversified aesthetics of images [34]–[38]. However, the inherent subjectiveness that leads to humans' diversity in image aesthetics is largely unknown, and more exploration is needed.

Recently, PIAA is also becoming increasingly more prevalent due to the demand for user-specific customization. Several works have been done towards this direction [8], [21], [22]. These methods aim to learn an individual's personalized aesthetic assessment from image content and attributes, which have shown to be correlated with their aesthetic preferences. While these approaches have obtained impressive performance, the subjective characteristics of users in aesthetic appreciation of images have rarely been incorporated. Studies have shown that individual users' preferences on images are mainly determined by their personality traits [41]–[45]. As an example, Fig. 2 shows two sets of images liked by users with different personality traits from the PsychoFlickr database [46]. As shown in Fig. 2 (a), extroverts tend to prefer image scenes with people playing together or engaging in exciting activities. Fig. 2 (b) shows images liked by a user with low extraversion, which manifests that introverts usually prefer quiet scenes. This indicates that personality traits are crucial in modeling an individual's subjective preference. Therefore, as an important subjective factor, humans' personality trait can be taken into account for PIAA.

According to [7], the generic aesthetics of an image is obtained by the general aesthetic judgment of human beings, which is correlated with their social or personality factors [47], [48]. Studies have also shown that the Big-Five personality traits are the general factors of aesthetic judgment for human beings and have a stable relationship with people's generic aesthetic experience for visual arts [49], [50]. Since people with similar personality traits typically like similar images, the personality traits can be inferred from their preferred images [51]. Therefore, personality traits are believed to the main influencing factors in people's image aesthetic preferences. When judging the generic aesthetics of an image, the personality traits of people who rate the image are very important auxiliary information. Besides, images aesthetic attributes (e.g., colorfulness, rule of thirds and depth of field) can be effectively used for predicting Big-Five personality traits [42], [46], which are also important factors in generic image aesthetic assessment [2], [11]. Therefore, it is obvious that generic aesthetics assessment and personality prediction from images are two related tasks. The relationship between generic aesthetic quality of an image and personality traits of people who like the image could be further investigated. Multi-task learning has been shown effective in capturing useful information contained in multiple related tasks, which can be used to improve the performances of all tasks [52]. Therefore, in this work we adopt a multi-task deep neural network for both aesthetics distribution prediction and personality prediction. To discover the common representation of aesthetics and personality in a unified network, the shared weights are trained using aesthetics data and personality data jointly. Since personality is a key subjective trait in influencing individual preferences on image aesthetics, we introduce an inter-task fusion based on our multi-task model to further delve the relationship between users' personality traits and their personalized image aesthetics.

In this paper, we propose a multi-task deep learning framework for generic and personalized image aesthetics assessment. The proposed framework consists of two stages. In the first stage, a multi-task convolutional neural network (MTCNN) with shared weights is proposed for both GIAA and personality prediction. In the second stage, based on the above two tasks, an inter-task fusion approach is further introduced for PIAA. We term the proposed framework Personality-Assisted Image Aesthetic Assessment (PA_IAA). The contributions of this work are as follows:

- We propose a novel personality-assisted multi-task learning network for both generic and personalized image aesthetics assessment, which outperforms the state-of-the-art methods.
- We train the proposed multi-task network jointly using aesthetics data and personality data, with the objective to capture the common representation features in a Siamese network. The trained model can predict both image aesthetics distribution and personality traits of people who like the image simultaneously.
- We propose an inter-task fusion approach to further infer individual user's aesthetic preference with the assistance of their personality traits. This can transfer image aesthetics from generic domain to personalized domain by only training a small group of images rated by individuals.

(a) (b)

Fig. 2. Example images liked by individual users with different personality traits from the PsychoFlickr database: (a) Images liked by a user with high extraversion; (b) Images liked by a user with low extraversion.

This intuitively explains the influence of personality traits on individuals' visual aesthetic preferences.

The rest of this paper is structured as follows. The related work is briefly introduced in Section II. In Section III, the proposed multi-task learning framework for both GIAA and PIAA is presented. Experimental results and discussions are given in Section IV, and the conclusions are drawn in Section V.

II. Related Work

This work addresses image aesthetics assessment, and it is based on personality computing and multi-task learning. In this section, we give a brief review of these techniques.

A. Image Aesthetics Assessment

1) Generic Image Aesthetics Assessment: Earlier works on image aesthetics assessment have mainly focused on mapping hand-crafted features into high-quality or low-quality aesthetic categories [11], [13], [14]. Datta *et al.* [11] proposed a learning-based GIAA approach, where the low-level and high-level visual features were combined to train a SVM model for binary aesthetics classification. Besides, generic content-based features were employed for image aesthetics classification, which were demonstrated to outperform hand-crafted features [14]. A large-scale Aesthetics Visual Analysis (AVA) database, which contains more than 250,000 labeled aesthetic images, was released in this work [14]. While these approaches have achieved great success in assessing the image aesthetics, the conventional features may fail to fully capture the aesthetic aspects of images, which are typically more abstract.

With the powerful abstract representation capability of deep Convolutional Neural Networks (CNNs) [26], many works have been done to learn deep models for GIAA [2], [6], [18]–[20], [28], [30]–[33], [36]. For example, a double-column Rating Pictorial Aesthetics using Deep Learning (RAPID) was presented in [6], [18], where both global and local patches of images were used for generic aesthetic binary classification. Lu *et al.* [19] proposed a Deep Multi-patch Aggregation (DMA) network for aesthetic quality categorization by aggregating multiple patches from an image. Similar to [18], Wang *et al.* [28] developed a multi-column Brain-inspired Deep Networks (BDN) model to predict multiple image attributes, which were then used for image aesthetic classification. In [33], Jin *et al.* proposed a novel CNN model, which combined Inception modules with connected Low and Global features (ILGNet) to classify image into high or low aesthetic quality. In addition to image aesthetics classification, image aesthetic regression and ranking have also attracted much attention. For example, Kong *et al.* [2] proposed to learn a ranking network with adaptive image attribute and content information for image aesthetic ranking. Recently, considering the difference among users' aesthetic perceptions on images, aesthetic distribution prediction have become another hot topic in IAA [34]–[38]. In [36], it has been shown that training CNN model to predict aesthetic distribution instead of average aesthetics score facilitates more effective IAA.

2) Personalized Image Aesthetics Assessment: More recently, PIAA has attracted increasingly more attention due to the fact that the aesthetic preferences of users on images are typically different [8], [21]–[24]. In [8], Ren *et al.* found that individual user's aesthetic preferences have a strong correlation with image content and aesthetic attributes, and proposed an active learning model for PIAA. In [21], a regression model and a ranking model based on Support Vector Machines (SVM) were jointly used to learn generic and personal aesthetic preferences on images. By considering the personal preference and user interaction, Lv *et al.* [22] proposed a user-friendly aesthetic ranking framework to automatically rank users' aesthetic preferences on images. These approaches attempt to leverage the objective visual features from images for modeling users' subjective aesthetic preferences. This may not be sufficient, because the subjective factors in rating image aesthetics are not fully investigated. Different from the aforementioned approaches, in this work we use personality traits as a subjective clue to assist the task of IAA.

B. Personality Computing

Personality computing aims at capturing individuals' stable preferences by exploring their observable behavioral cues [53].

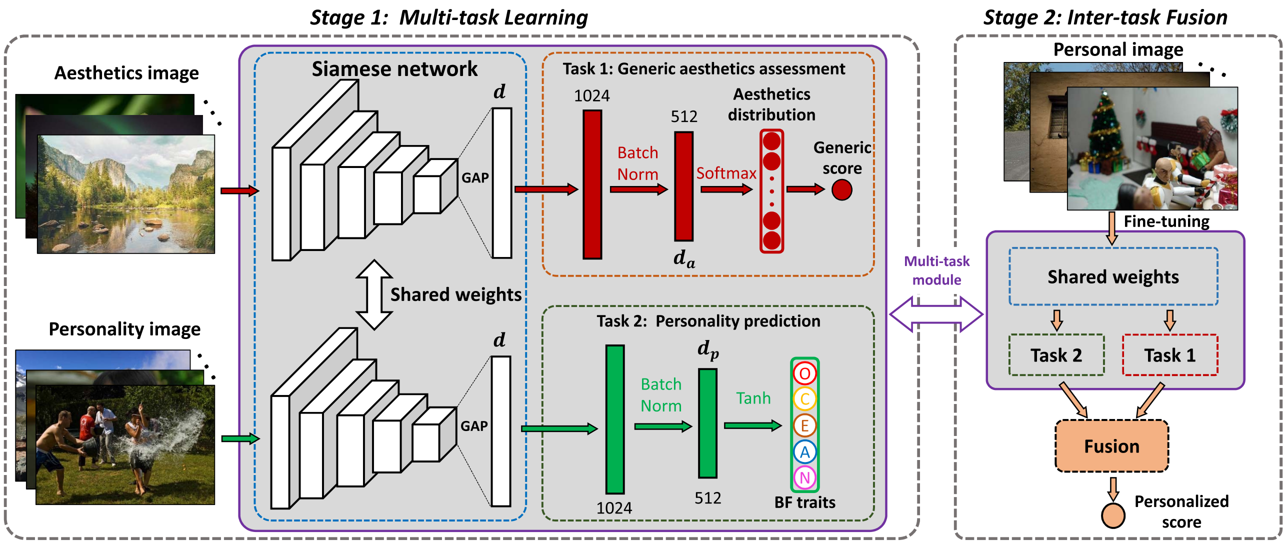


Fig. 3. The framework of the proposed personality-assisted multi-task learning for IAA.

Recently, a surging number of images are uploaded in social networks, which greatly promotes social media-based personality computing [54]. In [46], 300 Flickr individual users and their 60,000 liked images (200 images per user) were collected in the PsychoFlickr database. The Big-Five (BF) personality traits of each user, namely Openness (O), Conscientiousness (C), Extroversion (E), Agreeableness (A), and Neuroticism (N), were collected using the BIF-10 [55] questionnaire. Segalin *et al.* used a basic deep learning framework to classify an image into high and low categories according to each personality trait of people who like the image [51]. In [56], the preferences of individuals on local image regions were learned by a Weakly Supervised Dual Convolutional Network (WSDCN) for personality prediction. The user-level BF personality traits were used as the supervision of their liked images to train the personality prediction model, which can incorporate the subjective factor of users in aesthetic ratings.

C. Multi-Task Learning

Multi-task learning is a machine learning technique for improving the generalization performance of a task by using multiple domain knowledge contained in the shared information of related tasks [52], [57]. Recently, deep network-based multi-task learning [58] has achieved great success in computer vision [59]. Multi-task learning is inspired by the learning behavior of humans. In order to learn a new task, humans usually apply the knowledge obtained from learning related tasks. By sharing the hidden layers between related tasks, multi-task learning can obtain a common representation that captures all tasks, which can greatly decrease the risk of over-fitting [60]. The multi-task learning has shown to be an effective strategy to capture common features for multiple related tasks simultaneously. Personality traits are important subjective factors influencing their aesthetic preferences on images. Therefore, we believe personality prediction and aesthetics assessment from images are intimately related tasks. Motivated by this, in this work we propose to investigate the latent relationship between image aesthetics and people's personality traits using a multi-task deep network.

III. Proposed Method

In this paper, a novel multi-task learning framework is proposed to assess the aesthetic quality of images with the assistance of personality traits. By taking into account personality factors, more advanced image aesthetics assessment models can be developed. The architecture of the proposed approach is shown in Fig. 3. It comprises two stages. In the first stage, we design a multi-task learning network to predict both aesthetics distribution and BF personality traits simultaneously. The generic aesthetics score of the image is then generated based on the predicted aesthetics distribution. In order to capture the common information for both the aesthetic task and personality task, the multi-task module with shared weights is jointly trained using data from the aesthetics domain and personality domain. In the second stage, inter-task fusion is introduced for predicting the personalized aesthetic scores based on the generic aesthetics and personality traits information.

A. Qualitative Analysis Between Generic Aesthetics and Personality

To our best knowledge, the relationship between generic image aesthetics and personality has not been fully investigated so far, and no public database has been simultaneously labeled with subjects' personality traits and their aesthetic ratings on images. In order to qualitatively demonstrate the relation between generic aesthetics and personality, we employ an effective GIAA method in [2][1] to measure the aesthetic of images from the PsychoFlickr database [46]. Fig. 4 shows

[1] https://github.com/aimerykong/deepImageAestheticsAnalysis

Aesthetics:	0.759	0.699	0.488	0.375
Personality:	High Conscientiousness	High Agreeableness	High Extroversion	High Neuroticism

Fig. 4. Example images liked by people with different personality traits as well as the corresponding generic aesthetic scores (ranges from 0 to 1) obtained by the GIAA method in [2].

some example images liked by people with different personality traits and the corresponding generic aesthetic scores. It can be seen from Fig. 4 that images liked by people with high conscientiousness and agreeableness have higher generic aesthetic scores, which indicates that the aesthetic preferences of these people may be easily accepted by others. Since neurotic people may have more extreme preferences, their preferred images may not be easily accepted to others. In order to further quantitatively analyze the relationship between people's personality traits and generic image aesthetics, we propose a multi-task learning framework with shared weights based on the Siamese network [61]. To address the lack of personality and aesthetic labels in a single image database, the multi-task learning model is trained using the aesthetics data and personality data jointly.

B. Multi-Task Learning

As aforementioned, in the multi-task learning framework, the personality dataset and aesthetics dataset are used to optimize the respective task of the network separately. This aims at seeking useful information of two related tasks with hard parameter sharing [58], which is generally implemented by sharing the hidden layers and keeping respective task output layers. Our multi-task model is built upon a CNN model that is pre-trained on the ImageNet dataset [26]. We remove the fully-connected layers after the last convolutional layer and employ the Global Average Pooling (GAP) operation to achieve a shared vector $\boldsymbol{d}$. The aesthetics data and personality data are leveraged to learn the parameters of shared weights.

1) Generic Aesthetics Assessment Task: In this work, we denote $\{I_a^i, \boldsymbol{s}_a^i\}_{i=1}^{N_a}$ as the aesthetics training set, where N_a is the number of images for training. $\boldsymbol{s}_a^i = \{s_{a_n}^i\}_{n=1}^{N}$ denotes the aesthetics distribution of i-th image and $\sum_{n=1}^{N} s_{a_n}^i = 1$, where a_n represents the n-th score bucket, and N denotes the number of score buckets. For AVA database, $N = 10, a_1 = 1$, $a_N = 10$, and for FLICKR-AES database, $N = 5, a_1 = 1$, $a_N = 5$. On the top of the shared weights, we add two Fully Connected (FC) layers, which contain 1,024 nodes and 512 nodes respectively, to obtain the features for generic aesthetic task $\boldsymbol{d}_a$. Following each FC layer, Batch Normalization (BN) [62], [63] is added to improve generalization and accelerate the convergence rate of training. Since our generic aesthetic assessment is a distribution prediction problem with a sum of 1, the Softmax operator is used to generate the estimated aesthetics distribution $\hat{\boldsymbol{s}}_a^i = \{\hat{s}_{a_n}^i\}_{n=1}^{N}$, which is defined as

$$\hat{s}_{a_n}^i = \frac{e^{\boldsymbol{w}_{a_n}^{\mathrm{T}} \boldsymbol{d}_a}}{\sum_{j=1}^{N} e^{\boldsymbol{w}_{a_j}^{\mathrm{T}} \boldsymbol{d}_a}}, \tag{1}$$

where $\boldsymbol{W_a} = \{\boldsymbol{w}_{a_n}\}_{n=1}^{N}$ denotes the weight of the aesthetic features $\boldsymbol{d}_a$. We employ the Euclidean loss to optimize the generic aesthetic task, and it is formulated as

$$\mathcal{L}_a = \frac{1}{N_a} \frac{1}{N} \sum_{i=1}^{N_a} \sum_{n=1}^{N} \|\hat{s}_{a_n}^i - s_{a_n}^i\|_2^2. \tag{2}$$

After obtaining the aesthetics distribution $\{\hat{s}_{a_n}^i\}_{n=1}^{N}$ of an image, the generic score is obtained by

$$\hat{s}_a^i = \sum_{n=1}^{N} n \times \hat{s}_{a_n}^i, \tag{3}$$

which can be used to quantitatively compare the aesthetic quality of images [36].

2) Personality Prediction Task: We denote $\{I_p^{u_m}\}_{m=1}^{M}$ as the set of training images liked by the u-th user, whose BF personality traits are $\boldsymbol{s}_p^u = \{s_p^{u_i}\}_{i=1}^{5}, u = 1, 2, \ldots, U$, where M is the number of images liked by a user and U is the number of training users. We introduce a personality prediction task in parallel with the generic aesthetics assessment. Another two FC layers, which also contain 1,024 nodes and 512 nodes respectively, are added to generate the features for personality task $\boldsymbol{d}_p$. Following each FC layer, BN is added to improve generalization and accelerate the convergence rate of training. For the m-th image liked by the u-th user, the Tanh operator is employed to produce the BF personality traits $\hat{\boldsymbol{s}}_p^{u_m} = \{\hat{s}_p^{u_{mi}}\}_{i=1}^{5}$, which is defined as

$$\hat{\boldsymbol{s}}_p^{u_m} = \frac{e^{\boldsymbol{w}_p^{\mathrm{T}} \boldsymbol{d}_p} - e^{-\boldsymbol{w}_p^{\mathrm{T}} \boldsymbol{d}_p}}{e^{\boldsymbol{w}_p^{\mathrm{T}} \boldsymbol{d}_p} + e^{-\boldsymbol{w}_p^{\mathrm{T}} \boldsymbol{d}_p}}, \tag{4}$$

where $\boldsymbol{w}_p$ indicates the weight of the personality features $\boldsymbol{d}_p$. We adopt the Euclidean loss to optimize the personality prediction task, and it is formulated as

$$\mathcal{L}_p = \frac{1}{5} \frac{1}{U} \frac{1}{M} \sum_{i=1}^{5} \sum_{u=1}^{U} \sum_{m=1}^{M} \|\hat{s}_p^{u_{mi}} - s_p^{u_i}\|_2^2. \tag{5}$$

In this way, we obtain the predicted BF traits of an image, which refer to five personality traits of people who would like the image.

C. Inter-Task Fusion

After obtaining the predicted aesthetics distribution and BF traits from an image, we begin to focus on the user-specific aesthetic preferences on the image. Since the personality of individual users is an important subjective factor that influences their unique preferences for images [44], the aesthetics difference among users can be modeled by the user-specific personality traits. In order to learn different responses of individual users with different personality traits to images, we introduce an inter-task fusion to fine-tune our multi-task learning model for generating personalized image aesthetic score. We denote $\{I_b^i, s_b^i\}_{i=1}^{N_b}$ as the training set of an individual user. For the i-th training image rated by the user, the aesthetic score $\hat{s}_b^i$ can be calculated by

$$\hat{s}_b^i = \hat{s}_a^i + \boldsymbol{w}_b \hat{\boldsymbol{s}}_p^i, \tag{6}$$

where $\boldsymbol{w}_b = \{w_b^1, w_b^2, \ldots, w_b^5\}$ represent the responses of the individual user to five personality traits extracted from the image, $\hat{s}_a^i$ and $\hat{\boldsymbol{s}}_p^i$ denote the predicted generic score and five personality traits of the i-th image from multi-task learning model, respectively. Then, the Euclidean loss is employed to optimize the inter-task fusion model, and it is defined as

$$\mathcal{L}_b = \frac{1}{N_b} \sum_{i=1}^{N_b} \|\hat{s}_b^i - s_b^i\|_2^2. \tag{7}$$

In this manner, the individual user's personalized score $\hat{s}_b^i$ of image can be predicted from the trained inter-task fusion model.

D. Network Training

The training process of our PA_IAA model is divided into two stages. In the fist stage, we leverage a large amount of data from two domains (aesthetics and personality) to jointly train the multi-task learning model by optimizing the task-specific loss function

$$\mathcal{L} = \mathcal{L}_a + \gamma \mathcal{L}_p, \tag{8}$$

where γ is set to 1 based on our experiments. In the second stage, since the aesthetic ratings on images are very limited for each individual user, it is a challenging task to directly learn user-specific image aesthetics by deep learning model, which needs a mass of annotated images for training. Hence, we use the trained multi-task learning module as a pre-trained model, and its network parameters are fixed. Then, we fine-tune the parameters of inter-task fusion with personalized aesthetic domain data by optimizing the loss function $\mathcal{L}_b$. Stochastic Gradient Descent (SGD) optimization is utilized to minimize these loss functions in both multi-task learning and inter-task learning. Finally, the generic and personalized image aesthetic scores can be generated by our PA_IAA method. The learning procedure is summarized in Algorithm 1.

Algorithm 1 Multi-Task Learning Algorithm of Our Proposed PA_IAA

Input: Generic aesthetic training set $\{I_a^i, \boldsymbol{s}_a^i\}_{i=1}^{N_a}$, personality training set $\{\{I_p^{u_m}\}_{m=1}^{M}, \boldsymbol{s}_p^u\}_{u=1}^{U}$, personalized aesthetic training set $\{I_b^i, s_b^i\}_{i=1}^{N_b}$, max-epochs ϵ, test image I

Output: Predicted generic aesthetic score $\hat{s}_a$ and personalized aesthetic score $\hat{s}_b$ for I

Initialize the network weights and learning rates;
for $i = 1$; $i <= \epsilon$; $i++$ **do**
  Sample a batch of K samples from aesthetic database and personality database respectively;
  for $j = 1$; $j <= N_a/K$; $j++$ **do**
    Generate aesthetics distribution $\{\hat{\boldsymbol{s}}_a^i\}_{k=1}^{K}$ and personality traits $\{\hat{\boldsymbol{s}}_p^k\}_{k=1}^{K}$;
    Update multi-task model by optimizing the multi-task loss function $\mathcal{L}$;
  end for
end for
Obtain the trained multi-task model;
for $i = 1$; $i <= \epsilon$; $i++$ **do**
  Sample a batch of K samples from personalized aesthetic database;
  for $j = 1$; $j <= N_b/K$; $j++$ **do**
    Generate personalized aesthetic scores $\{\hat{s}_b^i\}_{k=1}^{K}$;
    Update inter-task learning model by optimizing the loss function $\mathcal{L}_b$;
  end for
end for
Obtain the trained inter-task model;
Input I into the trained PA_IAA model;
return $\hat{s}_a$, $\hat{s}_b$.

IV. Experimental Results and Discussion

A. Experimental Settings

1) Aesthetic Databases: We evaluate the performance of our approach on two public image aesthetic databases, including AVA [14] and FLICKR-AES [8]. Each image in the AVA database is only labeled with the aesthetic scores rated by multiple raters, while the FLICKR-AES database also provides each rater's ID when labeling the aesthetic ratings on images. Therefore, the AVA database is used for GIAA, and the FLICKR-AES database is used for PIAA.

AVA database [14] contains more than 250,000 images, each of which is rated by 78 $\sim$ 594 individual users. The aesthetic scores range from 1 to 10. For the task of GIAA, the aesthetics distribution of images are used as the supervisor for model training. Similar to [32], [33], more than 230,000 images are selected for aesthetic model training and the remaining about 20,000 images are used for testing.[2]

FLICKR-AES database [8] contains 40,000 images, each of which is labeled with aesthetic score by about five individual users. The aesthetic scores are rated from 1 to 5. Similar to [8], we first use 35,263 images with aesthetics distribution for generic model training, and then the remaining 4,737 test images labeled by 37 users are used to learn the personalized aesthetic model of each individual user. The number of testing images each individual user rated ranges from 105 to 171.

[2]The training/testing splits of AVA database can be found at: https://github.com/BestiVictory/ILGnet

2) Personality Database: We leverage the PsychoFlickr database [46] to infer the personality traits of people who like an image. In the PsychoFlickr database, 300 Flickr individual users and their 60,000 liked images (200 images per user) are collected. Each individual's BF personality traits, Openness (O), Conscientiousness (C), Extroversion (E), Agreeableness (A), and Neuroticism (N), are obtained by BIF-10 [55] questionnaire. Users' liked images as well as their BF personality traits are used to learn the task for personality prediction.

3) Implementation Details: We adopt two popular deep networks, i.e., Inception-v3 [64] and DenseNet121 [65], as the backbone of our multi-task module, which is pre-trained on ImageNet [26]. Images are resized to 224 × 224 × 3 (DenseNet121) and 299 × 299 × 3 (Inception-v3) to feed into the proposed network. In the multi-task learning process, the initial learning rates of shared layers and prediction layers are set to $1e-5$ and $1e-4$, respectively. In the inter-task learning process, the initial learning rate is $1e-4$. The learning rate drops to a factor of 0.9 every epoch. The other hyper-parameters are set as follows: weight decay of $1e-5$, momentum of 0.9, and total epoch of 50. We use Pytorch [66] to implement the proposed method.

4) Performance Criteria: For comparison with the existing GIAA and PIAA methods, the overall Accuracy (ACC), Spearman Rank Order Correlation Coefficient (SROCC) and squared Earth Movers Distance (EMD) are employed to evaluate the performance of the proposed method.

The ACC is the most popular criteria for evaluating the performance of classification [1], and is computed by

$$ACC = \frac{TP + TN}{P + N}, \tag{9}$$

where P and N denote the number of positive and negative sample images, respectively. TP and TN are the number of correctly classified positive and negative sample images. The ACC is in the range [0, 1], and higher ACC value represents better classification performance.

Similar to [2], [22], [36], the SROCC is employed to evaluate the ranking correlation between the predicted aesthetics scores and the ground-truth aesthetics scores. Assuming s_i and $\hat{s}_i$ represent the score ranks of the i-th test image in actual and predicted aesthetic scores respectively, the difference between the ranks of the actual and aesthetic scores can be calculated as $d_i = s_i - \hat{s}_i$, and the SROCC is defined by

$$SROCC = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N(N+1)}, \tag{10}$$

where N is the number of the testing samples. The SROCC ranges from -1 to 1, and higher SROCC value represents better prediction performance.

Similar to [36], we employ EMD to evaluate the closeness of predicted aesthetics distribution and the ground-truth aesthetics distribution. The EMD is defined by

$$EMD = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{K}\sum_{k=1}^{K}|CDF_{\hat{s}^i}(k) - CDF_{s^i}(k)|^r\right)^{\frac{1}{r}}, \tag{11}$$

where N is the number of the testing samples and $CDF_s(k)$ denotes the cumulative distribution function. Similar to [36], r is set to 2 and lower EMD value represents better prediction performance.

B. Generic Image Aesthetics Assessment

1) Baseline Methods: There are three kinds of approaches for GIAA: aesthetic binary classification, aesthetic score regression and aesthetic distribution prediction. For aesthetic distribution prediction, we use EMD to evaluate the prediction performance of the proposed approaches and NIMA [36]. For aesthetic score regression, the average score of the aesthetics distribution is computed by Eq. 3. We conduct comparative experiments with the state-of-the-art methods [2], [22], [36] in term of the ranking correlation (SROCC). For aesthetic binary classification, representative works include AVA hand-crafted features [14], RAPID [18], RAPID (improved version) [6], DMA [19], Wang *et al.* [20], Kao *et al.* [27], BDN [28], Kao *et al.* [29], Zhang *et al.* [30], Schwarz *et al.* [31], Kucer *et al.* [32] and ILGNet [33]. For fair comparison with these existing classification accuracy results (ACC) reported on the AVA database, we simply threshold the average score to produce a binary classification. The threshold of low and high aesthetic scores is set to 5 in AVA database.

2) Performance Evaluation on AVA Database: Table I summarizes the prediction performances of the proposed method and baseline methods on AVA database, and the best results are highlighted in bold font. For aesthetic binary classification, PA_IAA achieves the highest accuracy among all the methods. Furthermore, when aesthetics task is used individually, the performances of the proposed approaches are close to those of the state-of-the-arts. After incorporating the personality task, the classification accuracy increases by 2.4%(DenseNet121) and 2.8%(Inception-v3). In terms of aesthetic score regression and aesthetic distribution prediction, DenseNet121(aesthetics) and Inception-v3(aesthetics) deliver competitive performance with NIMA [36]. When the aesthetics task is jointly trained with the personality task, PA_IAA yields the best performance. This demonstrates the effectiveness of multi-task network for jointly learning the aesthetics distribution of an image and the personality traits of people who like the image. Our multi-task learning approach can leverage useful information learned from two related tasks (aesthetics and personality) to improve the performance of GIAA task.

C. Personality Prediction

1) Baseline Methods: In order to verify whether the proposed multi-task learning module is also effective for personality prediction, we compare the PA_IAA with several state-of-the-art methods for personality prediction, including Segalin *et al.* [42], Guntuku *et al.* [43]. In this experiment, 300 users with their liked images are divided into two subsets with 90% for training and 10% for testing, and ten-fold cross-validation is used to avoid bias. The average of testing results is reported.

TABLE I
PERFORMANCE COMPARISON BETWEEN OUR METHOD AND BASELINE METHODS ON AVA DATABASE

Method	ACC(%)↑	SROCC↑	EMD↓
AVA handcrafted features [14]	68.0	-	-
RAPID [18]	74.5	-	
RAPID (improved version) [6]	75.4	-	-
DMA [19]	75.4	-	-
Wang *et al.* [20]	76.9	-	-
Kao *et al.* [27]	76.2	-	-
BDN [28]	78.1	-	-
Kao *et al.* [29]	79.1	-	-
Zhang *et al.* [30]	78.8	-	-
Schwarz *et al.* [31]	75.8	-	-
Kucer *et al.* [32]	81.9	-	-
ILGNet [33]	82.7	-	-
AlexNet_FT_Conf [2]	71.5	0.481	-
Reg+Rank+Att [2]	75.5	0.545	-
Reg+Rank+Cont [2]	73.4	0.541	-
Reg+Rank+Att+Cont [2]	77.3	0.558	-
USAR_PPR [22]	72.4	0.600	-
USAR_PAD [22]	77.7	0.545	-
USAR_PPR&PAD [22]	78.1	0.578	-
NIMA(VGG16) [36]	80.6	0.592	0.054
NIMA(Inception-v2) [36]	81.5	0.612	0.050
DenseNet121(aesthetics)	80.5	0.630	0.051
Inception-v3(aesthetics)	80.9	0.638	0.050
PA_IAA(DenseNet121)	82.9	0.666	0.049
PA_IAA(Inception-v3)	**83.7**	**0.677**	**0.047**

TABLE II
PERFORMANCE COMPARISON (SROCC) OF PERSONALITY PREDICTION ON PSYCHOFLICKR DATABASE

Method	O	C	E	A	N
Segalin *et al.* [42]	0.354	0.535	0.625	0.476	0.613
Guntuku *et al.* [43]	0.398	0.552	0.679	0.525	0.636
DenseNet121(personality)	0.548	0.647	0.711	0.638	0.698
Inception-v3(personality)	0.536	0.654	0.703	0.651	0.709
PA_IAA(DenseNet121)	**0.567**	0.659	**0.722**	0.646	0.708
PA_IAA(Inception-v3)	0.555	**0.668**	0.715	**0.662**	**0.717**

2) Performance Evaluation on PsychoFlickr Database: In Table II, we list the comparative performance of our method and the state-of-the-art methods reported on PsychoFlickr database, where the best results are highlighted in bold font. When personality task is used individually, DenseNet121(personality) and Inception-v3(personality) outperform the other two methods. In addition, PA_IAA(DenseNet121) and PA_IAA(Inception-v3) can achieve the highest SROCC values when combined with aesthetic task. It is worth noting that PA_IAA outperforms the model trained with single task for personality prediction, which demonstrates that the proposed multi-task learning module also contributes helpful information for personality prediction task.

D. Cross Database Evaluation

To verify the generalization of our method for GIAA task, we train models on one database and test them on other databases. In this experiment, Inception-v3 is used as the backbone network. The cross database evaluation results are listed in Table III. Although training and testing on FLICKR-AES leads to the best results (0.716), the PA_IAA model trained on AVA achieves higher average results than PA_IAA model trained on FLICKR-AES. This demonstrates that PA_IAA model trained on AVA database has more effective generalization ability for GIAA task, which is mainly because that the training images on AVA receive more aesthetic labels and the number of training images on AVA is far greater than the number of training images on FLICKR-AES.

TABLE III
THE CORRELATION (SROCC) OF PA_IAA(INCEPTION-V3) FOR GENERIC IAA WHEN TRAINING AND TESTING ON DIFFERENT DATABASE

Train Database	Test Database		Average
	AVA [14]	FLICKR-AES [46]	
AVA [14]	0.677	0.528	**0.603**
FLICKR-AES [46]	0.234	0.716	0.475

E. Discussion Between Generic Aesthetics and Personality Traits

The average aesthetic scores can be used as the generic aesthetics of images. Fig. 5 shows some example images from AVA database and the corresponding predicted aesthetic distributions using our PA_IAA(Inception-v3) model. The aesthetics distributions and the normalized average scores of predicted (and ground-truth) are shown beside each image. The predicted personality traits, which are normalized into the range $[-1, 1]$, are also shown. As shown in Fig. 5, the proposed PA_IAA model can predict aesthetics distributions consistently with the ground-truth aesthetics distributions. From the predicted personality traits of images, we have the following observations. (1) Images liked by people with high Conscientiousness and Agreeableness may have high average scores, while the images with low aesthetic quality may be liked by people with low Conscientiousness and Agreeableness (Fig. 5(a)). This confirms the findings in Section III-A that responsible and friendly people are more likely to be recognized in the aesthetic aspect of images. (2) Images with low generic aesthetics are mainly preferred by people with high Neuroticism (Fig. 5(d)), which indicates that neurotic people often like somewhat enclosed and dark image scenes that are not easily accepted by others.

To further investigate the correlation between the generic aesthetics of an image and the personality traits of people who like the image, we employ SROCC to measure the ranking correlation between the predicted average scores and personality traits from the images on FLICKR-AES [8] and AVA [14] databases. The test results are summarized in Table IV, where the highest positive and negative correlations for each database are marked in bold font. From this table, we can find that "Neuroticism" have the highest negative correlation with generic scores for two databases, which demonstrates that the images preferred by neurotic people are very likely to have low aesthetic quality. "Conscientiousness" and "Agreeableness" have positive correlation with generic scores, which indicates that if an image is mostly liked by people with these two personality traits, its aesthetic score tends to be rated high by the majority of people. In addition, "Openness"

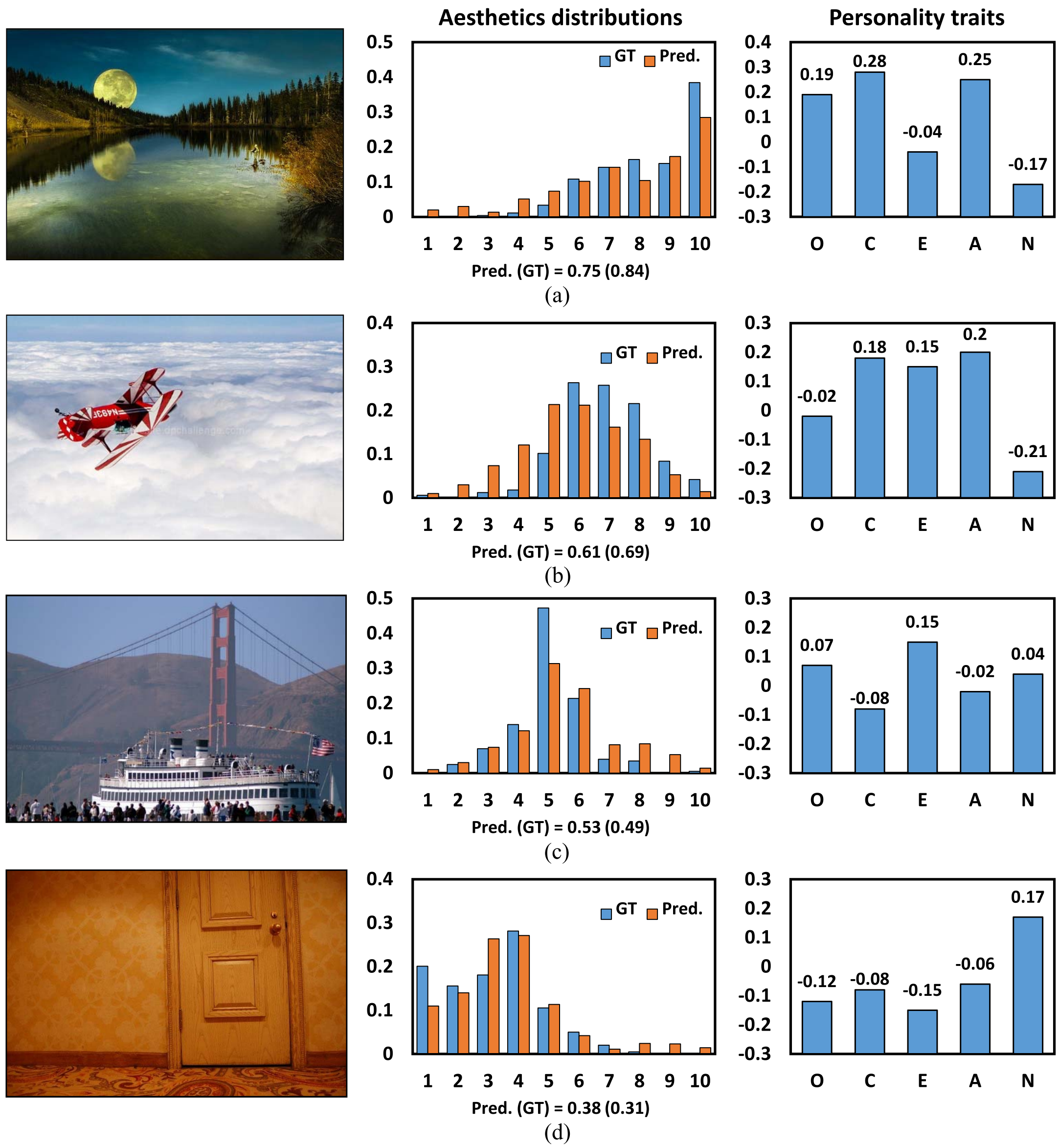


Fig. 5. Some example images from AVA database tested by our PA_IAA(Inception-v3) model. The aesthetics distributions of predicted (and ground-truth) and the five predicted personality traits are shown beside each image. The average score of each testing image is also shown. The aesthetic scores range from 0 to 1 and the personality traits range from -1 to 1. (a)-(d) Four example images and the predicted (and ground-truth) aesthetic distributions and five personality traits.

TABLE IV
THE CORRELATION (SROCC) BETWEEN GENERIC IMAGE AESTHETICS AND EACH PERSONALITY TRAIT (I.E. O, C, E, A, N) ON AVA AND FLICKR-AES DATABASES

Database	O	C	E	A	N
AVA	-0.0753	**0.6752**	-0.1034	0.4633	**-0.3864**
FLICKR-AES	-0.1463	**0.8421**	-0.2347	0.6457	**-0.6355**

and "Extroversion" have relatively weak negative correlation with generic aesthetics through the testing results. Therefore, GIAA and personality prediction are two related tasks and the relationship between them can be learned by the multi-task module of PA_IAA.

F. Personalized Image Aesthetics Assessment

1) Baseline Methods: To validate the performance of PA_IAA model for PIAA, we conduct comparative experiments with the state-of-the-art methods, including PAM [8], USAR [22], FPMF [67], which also reported the test results on FLICKR-AES database. In this experiment, Inception-v3 is used as the backbone network. We fist leverage the training sets of FLICKR-AES and PsychoFlickr databases to learn the multi-task module of PA_IAA, and then conduct 37 personalized aesthetic experiments on the testing set of FLICKR-AES database. For each individual, the images he/she rated are split into two sets, i.e., k images for model fine-tuning and the remaining images for test. In order to avoid bias, the experiments are conducted 50 times for each individual, and the average results as well as the standard deviation are reported. In order to verify the prediction performance with both a few training images and many training images and compare with the reported results in [8], [22], we set $k = 10$ and $k = 100$, respectively.

2) Performance Evaluation on FLICKR-AES Database: To verify the effectiveness of using personality data, we replace

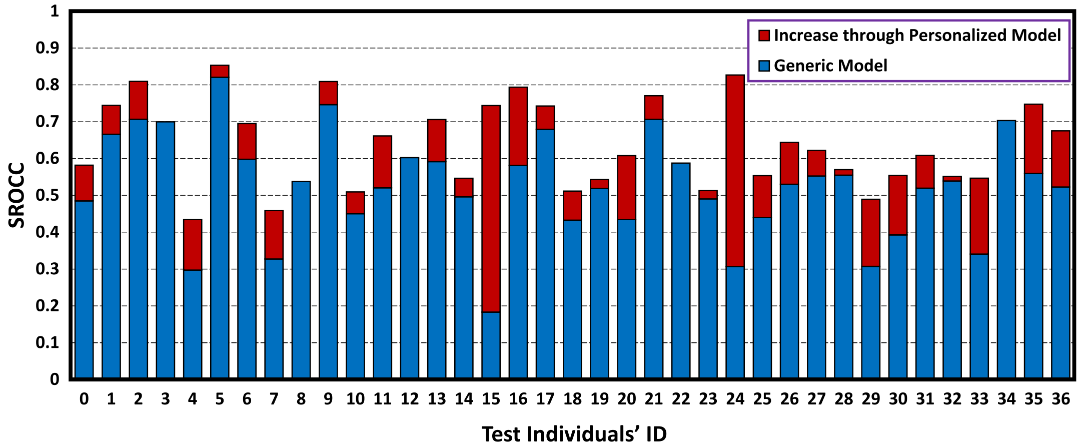


Fig. 6. Performance comparison of 37 test individuals from the FLICKR-AES database by directly using our generic aesthetic model and personalized aesthetic model when $k = 100$. The blue bars show the SROCC values of 37 individuals using the generic model, and the red bars show the increase in SROCC values after using the proposed personalized model.

TABLE V
PERFORMANCE RESULTS (SROCC) OF OUR METHOD AND BASELINE METHODS ON FLICKR-AES DATABASE FOR PERSONALIZED IMAGE AESTHETICS ASSESSMENT

Method	10 images	100 images
FPMF (only attribute) [67]	0.511±0.004	0.516±0.003
FPMF (only content) [67]	0.512±0.002	0.516±0.010
FPMF (content and attribute) [67]	0.513±0.003	0.524±0.007
PAM (only attribute) [8]	0.518±0.003	0.539±0.013
PAM (only content) [8]	0.515±0.004	0.535±0.017
PAM (content and attribute) [8]	0.520±0.003	0.553±0.012
USAR_PPR [22]	0.521±0.002	0.544±0.007
USAR_PAD [22]	0.520±0.003	0.537±0.003
USAR_PPR&PAD [22]	0.525±0.004	0.552±0.015
MT_IAA	0.523±0.004	0.582±0.014
PA_IAA	**0.543±0.003**	**0.639±0.011**

the personality prediction on PsychoFlickr in the multi-task module with aesthetic distribution prediction on AVA. We call this approach MT_IAA. Table V summarizes the performances of our method and the baseline methods for PIAA, where the best results are highlighted in bold font. From this table, it can be observed that our PA_IAA method outperforms all the other state-of-the-art models. In addition, PA_IAA outperforms MT_IAA by a large margin, which indicates that the personality prediction task of the proposed model has made a significant contribution to PIAA. In particular, when 10 images are selected for training, PA_IAA has better performance than the other methods. By contrast, our approach can achieve significantly better performance than the other methods with 100 training images. This demonstrates the effectiveness of our PA_IAA model, which can leverage the predicted personality traits extracted from images for modeling individual's personalized aesthetics.

To further demonstrate the effectiveness of our PA_IAA model for learning each individual's personalized aesthetic preferences, we first directly use the generic model to compute the prediction performance (SROCC) of 37 test individuals on FLICKR-AES database, and then calculate the increase in SROCC values by using our personalized model when $k = 100$. As shown in Fig. 6, we observe that the prediction performance for almost all individuals has increased significantly through our personalized model. For example, the prediction performance has increased by about 0.56 (from 0.18 to 0.74) for individual 15. The average increase in SROCC for 37 test individuals is about 0.11 (from 0.527 to 0.639). Hence, our inter-task personalized model can effectively utilize individual different response to the predicted personality traits of images for observably improving the overall performance for personalized aesthetics assessment.

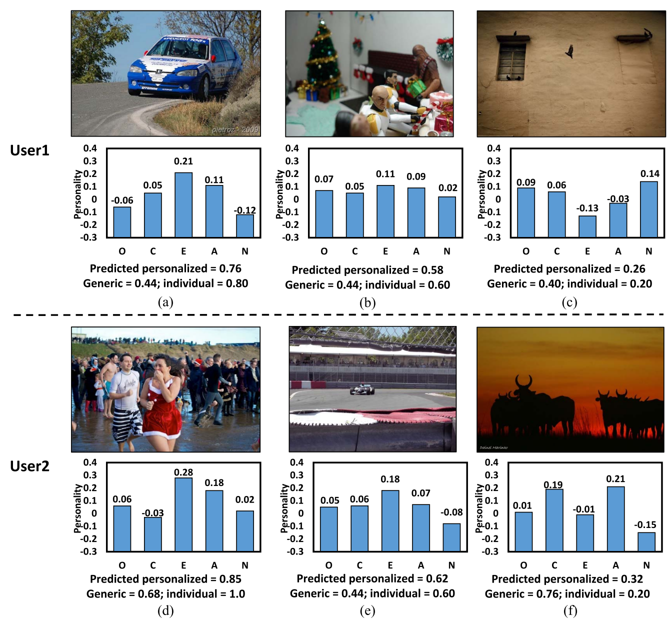


Fig. 7. Some example images rated by two users from FLICKR-AES database. The predicted personality traits and personalized aesthetic scores are shown below each image. The ground truth generic and individual aesthetic scores are also shown below each image. The aesthetic scores range from 0 to 1 and the personality scores range from -1 to 1.

3) Influence of Personality Traits for Personalized Image Aesthetics: In order to explore the influence of personality traits for individual users' aesthetic preferences on images, we show two sets of example images rated by two testing individual users from FLICKR-AES database in Fig. 7.

The average scores of aesthetics distribution is used as the generic quality of images. For images in the first row, the generic scores are almost the same (0.44, 0.44 and 0.40). However, User1 has significantly different aesthetic ratings on these images (0.80, 0.60 and 0.20). Particularly, User1 assigns a high aesthetic score on Fig. 7(a), which is most likely to be liked by people with higher extroversion (with probability 0.21) and lower neuroticism (-0.12). In contrast, people with lower extroversion (-0.13) and higher neuroticism (0.14) have preferences on Fig. 7(c), which is perceived by User1 with low aesthetic quality. This indicates that User1 may be a person with high extroversion and low neuroticism. For images in the second row, the generic aesthetic scores are inconsistent with the personal aesthetic scores rated by User2. This is mainly because that User2 is an exocentric and neurotic person, who has high aesthetic perception of images liked by people with the similar personality traits and dislikes those images preferred by people with the opposite personality traits. In addition, the predicted personalized scores of our PIAA model are closer to the users' individual rating scores compared to the generic scores.

V. Conclusion

In this paper, we have presented a novel personality-assisted multi-task learning framework for both generic and personalized image aesthetics assessment (PA_IAA). Particularly, the personality data and aesthetics data have been jointly used to train a multi-task learning network, which can capture common features to predict both aesthetics distribution and personality traits simultaneously. By integrating data from two related domains in the same network, PA_IAA can improve the performance of two tasks. In addition, an inter-task fusion is introduced to learn the influence of BF personality traits in individuals' aesthetic preferences on images. Extensive experimental results on two public databases have demonstrated that our proposed PA_IAA model is superior to the existing state-of-the-art GIAA and PIAA approaches.

References

[1] Y. Deng, C. C. Loy, and X. Tang, "Image aesthetic assessment: An experimental survey," *IEEE Signal Process. Mag.*, vol. 34, no. 4, pp. 80–106, Jul. 2017.

[2] S. Kong, X. Shen, Z. Lin, R. Mech, and C. Fowlkes, "Photo aesthetics ranking network with attributes and content adaptation," in *Proc. Eur. Conf. Comput. Vis.*, May 2016, pp. 662–679.

[3] W.-T. Sun, T.-H. Chao, Y.-H. Kuo, and W. H. Hsu, "Photo filter recommendation by category-aware aesthetic learning," *IEEE Trans. Multimedia*, vol. 19, no. 8, pp. 1870–1880, Aug. 2017.

[4] R. Hong, L. Zhang, and D. Tao, "Unified photo enhancement by discovering aesthetic communities from Flickr," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1124–1135, Mar. 2016.

[5] W. Wang, J. Shen, and H. Ling, "A deep network solution for attention and aesthetics aware photo cropping," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 7, pp. 1531–1544, Jul. 2019, doi: 10.1109/tpami.2018.2840724.

[6] X. Lu, Z. Lin, H. Jin, J. Yang, and J. Z. Wang, "Rating pictorial aesthetics using deep learning," *IEEE Trans. Multimedia*, vol. 17, no. 11, pp. 2021–2034, Sep. 2015.

[7] W. Kim, J. Choi, and J. Lee, "Objectivity and subjectivity in aesthetic quality assessment of digital photographs," *IEEE Trans. Affective Comput.*, to be published.

[8] J. Ren, X. Shen, Z. Lin, R. Mech, and D. J. Foran, "Personalized image aesthetics," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 638–647.

[9] Y. Ke, X. Tang, and F. Jing, "The design of high-level features for photo quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2006, pp. 419–426.

[10] M.-C. Yeh and Y.-C. Cheng, "Relative features for photo quality assessment," in *Proc. 19th IEEE Int. Conf. Image Process.*, Sep. 2012, pp. 2861–2864.

[11] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Studying aesthetics in photographic images using a computational approach," in *Proc. Eur. Conf. Comput. Vis.*, Oct. 2006, pp. 288–301.

[12] Y. Luo and X. Tang, "Photo and video quality evaluation: Focusing on the subject," in *Proc. Eur. Conf. Comput. Vis.*, Jun. 2008, pp. 386–399.

[13] X. Tang, W. Luo, and X. Wang, "Content-based photo quality assessment," *IEEE Trans. Multimedia*, vol. 15, no. 8, pp. 1930–1943, Dec. 2013.

[14] N. Murray, L. Marchesotti, and F. Perronnin, "AVA: A large-scale database for aesthetic visual analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2408–2415.

[15] X. Tian, Z. Dong, K. Yang, and T. Mei, "Query-dependent aesthetic model with deep learning for photo quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 11, pp. 2035–2048, Nov. 2015.

[16] L. Mai, H. Jin, and F. Liu, "Composition-preserving deep photo aesthetics assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 497–506.

[17] Z. Liu, Z. Wang, Y. Yao, L. Zhang, and S. Ling, "Deep active learning with contaminated tags for image aesthetics assessment," *IEEE Trans. Image Process.*, to be published.

[18] X. Lu, Z. Lin, H. Jin, J. Yang, and J. Z. Wang, "RAPID: Rating pictorial aesthetics using deep learning," in *Proc. ACM Int. Conf. Multimedia*, Nov. 2014, pp. 457–466.

[19] X. Lu, Z. Lin, X. Shen, R. Mech, and J. Z. Wang, "Deep multi-patch aggregation network for image style, aesthetics, and quality estimation," in *Proc. IEEE Int. Conf. Comput. Vis.*, May 2015, pp. 990–998.

[20] W. Wang, M. Zhao, L. Wang, J. Huang, C. Cai, and X. Xu, "A multi-scene deep learning model for image aesthetic evaluation," *Signal Process., Image Commun.*, vol. 47, pp. 511–518, Sep. 2016.

[21] K. Park, S. Hong, M. Baek, and B. Han, "Personalized image aesthetic quality assessment by joint regression and ranking," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, May 2017, pp. 1206–1214.

[22] P. Lv *et al.*, "USAR: An interactive user-specific aesthetic ranking framework for images," in *Proc. ACM Int. Conf. Multimedia*, Nov. 2018, pp. 1328–1336.

[23] X. Deng, C. Cui, H. Fang, X. Nie, and Y. Yin, "Personalized image aesthetics assessment," in *Proc. ACM Conf. Inf. Knowl. Manage. (CIKM)*, Nov. 2017, pp. 2043–2046.

[24] G. Wang, J. Yan, and Z. Qin, "Collaborative and attentive learning for personalized image aesthetic assessment," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Jul. 2018, pp. 957–963.

[25] J. Chen, C. Wang, J. Wang, X. Ying, and X. Wang, "Learning the personalized intransitive preferences of images," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4139–4153, Sep. 2017.

[26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, Dec. 2012, pp. 1097–1105.

[27] Y. Kao, K. Huang, and S. Maybank, "Hierarchical aesthetic quality assessment using deep convolutional neural networks," *Signal Process., Image Commun.*, vol. 47, pp. 500–510, Sep. 2016.

[28] Z. Wang, F. Dolcos, D. Beck, S. Chang, D. Liu, and T. S. Huang, "Brain-inspired deep networks for image aesthetics assessment," 2016, *arXiv:1601.04155*. [Online]. Available: https://arxiv.org/abs/1601.04155

[29] Y. Kao, R. He, and K. Huang, "Deep aesthetic quality assessment with semantic information," *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1482–1495, Mar. 2017.

[30] C. Zhang, C. Zhu, X. Xu, Y. Liu, J. Xiao, and T. Tillo, "Visual aesthetic understanding: Sample-specific aesthetic classification and deep activation map visualization," *Signal Process., Image Commun.*, vol. 67, pp. 12–21, Sep. 2018.

[31] K. Schwarz, P. Wieschollek, and H. P. A. Lensch, "Will people like your image? Learning the aesthetic space," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, May 2018, pp. 2048–2057.

[32] M. Kucer, A. C. Loui, and D. W. Messinger, "Leveraging expert feature knowledge for predicting image aesthetics," *IEEE Trans. Image Process.*, vol. 27, no. 10, pp. 5100–5112, Oct. 2018.

[33] X. Jin *et al.*, "ILGNet: Inception modules with connected local and global features for efficient image aesthetic quality classification using domain adaptation," *IET Comput. Vis.*, vol. 13, no. 2, pp. 206–212, Mar. 2019.

[34] B. Jin, M. V. O. Segovia, and S. Susstrunk, "Image aesthetic predictors based on weighted CNNs," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2016, pp. 2291–2295.

[35] X. Jin *et al.*, "Predicting aesthetic score distribution through cumulative jensen-Shannon divergence," in *Proc. 32nd AAAI Int. Conf. Artif. Intell.*, Oct. 2018, pp. 77–84.

[36] H. Talebi and P. Milanfar, "NIMA: Neural image assessment," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 3998–4011, Aug. 2018.

[37] C. Cui, H. Liu, T. Lian, L. Nie, L. Zhu, and Y. Yin, "Distribution-oriented aesthetics assessment with semantic-aware hybrid network," *IEEE Trans. Multimedia*, vol. 21, no. 5, pp. 1209–1220, May 2019.

[38] X. Zhang, X. Gao, W. Lu, and L. He, "A gated peripheral-foveal convolutional neural network for unified image aesthetic prediction," *IEEE Trans. Multimedia*, vol. 21, no. 11, pp. 2815–2826, Nov. 2019, doi: 10.1109/tmm.2019.2911428.

[39] Y. Zhu, S. C. Guntuku, W. Lin, G. Ghinea, and J. A. Redi, "Measuring individual video QoE: A survey, and proposal for future directions using social media," *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 14, no. 2s, 2018, Art. no. 30.

[40] N. Ma, A. Volkov, A. Livshits, P. Pietrusinski, H. Hu, and M. Bolin, "An universal image attractiveness ranking framework," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Jan. 2019, pp. 657–665.

[41] S. C. Guntuku, J. T. Zhou, S. Roy, W. Lin, and I. W. Tsang, "Understanding deep representations learned in modeling users likes," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3762–3774, Aug. 2016.

[42] C. Segalin, A. Perina, M. Cristani, and A. Vinciarelli, "The pictures we like are our image: Continuous mapping of favorite pictures into self-assessed and attributed personality traits," *IEEE Trans. Affective Comput.*, vol. 8, no. 2, pp. 268–285, Apr. 2017.

[43] S. C. Guntuku, J. T. Zhou, S. Roy, W. Lin, and I. W. Tsang, "'Who likes what and, why?' Insights into modeling Users' personality based on image 'Likes,'" *IEEE Trans. Affective Comput.*, vol. 9, no. 1, pp. 130–143, Jan. 2018.

[44] H. Zhu, L. Li, S. Zhao, and H. Jiang, "Evaluating attributed personality traits from scene perception probability," *Pattern Recognit. Lett.*, vol. 116, pp. 121–126, Dec. 2018.

[45] H. Zhu, L. Li, and H. Jiang, "Inferring personality traits from user liked images via weakly supervised dual convolutional network," in *Proc. ACM Int. Conf. Multimedia Workshop (ASMMC-MMAC)*, Oct. 2018, pp. 63–69.

[46] M. Cristani, A. Vinciarelli, C. Segalin, and A. Perina, "Unveiling the multimedia unconscious: Implicit cognitive processes and multimedia content analysis," in *Proc. ACM Int. Conf. Multimedia*, Nov. 2013, pp. 213–222.

[47] C. Cui, W. Yang, C. Shi, M. Wang, X. Nie, and Y. Yin, "Personalized image quality assessment with Social-Sensed aesthetic preference," *Inf. Sci.*, vol. 512, pp. 780–794, Feb. 2020.

[48] P. Galanter, "Computational aesthetic evaluation: Past and future" in *Computers and Creativity*. New York, NY, USA: Springer, 2012, pp. 255–293.

[49] V. Swami and A. Furnham, "Personality and aesthetic experiences," in *Cambridge Handbooks in Psychology*. Cambridge, U.K.: Cambridge Univ. Press, 2014, pp. 540–561.

[50] T. Chamorro-Premuzic, S. Reimers, A. Hsu, and G. Ahmetoglu, "Who art thou? Personality predictors of artistic preferences in a large UK sample: The importance of openness," *Brit. J. Psychol.*, vol. 100, no. 3, pp. 501–516, Aug. 2009.

[51] C. Segalin, D. S. Cheng, and M. Cristani, "Social profiling through image understanding: Personality inference using convolutional neural networks," *Comput. Vis. Image Understand.*, vol. 156, pp. 34–50, Mar. 2017.

[52] R. Caruana, "Multitask learning," *Mach. Learn.*, vol. 28, no. 1, pp. 41–75, 1997.

[53] G. Matthews and I. J. Deary, *Personality Traits*, vol. 28, no. 27. Cambridge, U.K.: Cambridge Univ. Press, 2009, p. 325.

[54] A. Vinciarelli and G. Mohammadi, "A survey of personality computing," *IEEE Trans. Affective Comput.*, vol. 5, no. 3, pp. 273–291, Jul. 2014.

[55] 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," *J. Res. Personality*, vol. 41, no. 1, pp. 203–212, Feb. 2007.

[56] H. Zhu, L. Li, H. Jiang, and A. Tan, "Inferring personality traits from attentive regions of user liked images via weakly supervised dual convolutional network," *Neural Process. Lett.*, to be published, doi: 10.1007/s11063-019-09987-7.

[57] Y. Zhang and Q. Yang, "A survey on multi-task learning," 2017, *arXiv:1707.08114*. [Online]. Available: https://arxiv.org/abs/1707.08114

[58] S. Ruder, "An overview of multi-task learning in deep neural networks," 2017, *arXiv:1706.05098*. [Online]. Available: https://arxiv.org/abs/1706.05098

[59] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1440–1448.

[60] J. Baxter, "A Bayesian/information theoretic model of learning to learn via multiple task sampling," *Mach. Learn.*, vol. 28, no. 1, pp. 7–39, Jul. 1997.

[61] S. Chopra, R. Hadsell, and Y. Lecun, "Learning a similarity metric discriminatively, with application to face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jun. 2005, pp. 539–546.

[62] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning requires rethinking generalization," in *Proc. Int. Learn. Representations*, Apr. 2017, pp. 1–15.

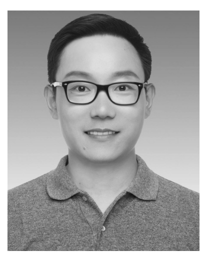
[63] S. Santurkar, D. Tsipras, A. Ilyas, and A. Madry, "How does batch normalization help optimization?" in *Proc. Adv. Neural Inf. Process. Syst.*, Dec. 2018, pp. 2488–2498.

[64] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2818–2826.

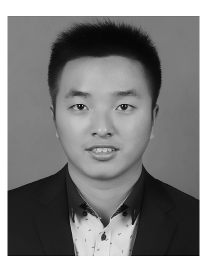
[65] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Sep. 2017, pp. 2261–2269.

[66] A. Paszke *et al.*, "Automatic differentiation in pytorch," in *Proc. Int. Conf. Neural Inf. Process. Syst. Workshop*, 2017.

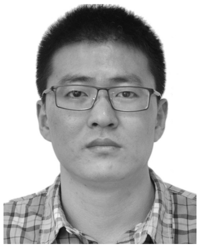
[67] P. O'Donovan, A. Agarwala, and A. Hertzmann, "Collaborative filtering of color aesthetics," in *Proc. Workshop Comput. Aesthet.*, 2014, pp. 33–40.



Leida Li received the B.S. and Ph.D. degrees from Xidian University, Xi'an, China, in 2004 and 2009, respectively. In 2008, he was a Research Assistant with the Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Taiwan. From 2014 to 2015, he was a Visiting Research Fellow with the Rapid-Rich Object Search (ROSE) Laboratory, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he was a Senior Research Fellow from 2016 to 2017. He is currently a Full Professor with the School of Information and Control Engineering, China University of Mining and Technology, and the School of Artificial Intelligence, Xidian University. His current research interests include multimedia quality assessment, affective computing, information hiding, and image forensics. He is also an Associate Editor of the *Journal of Visual Communication and Image Representation* and the *EURASIP Journal on Image and Video Processing*.

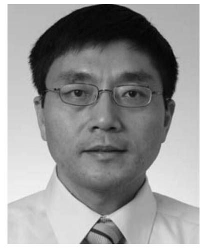


Hancheng Zhu received the B.S. degree from the Changzhou Institute of Technology, Changzhou, China, in 2012, and the M.S. degree from the China University of Mining and Technology, Xuzhou, China, in 2015, where he is currently pursuing the Ph.D. degree with the School of Information and Control Engineering. His research interests include affective computing and image aesthetics assessment.



Sicheng Zhao (Senior Member, IEEE) received the Ph.D. degree from the Harbin Institute of Technology, Harbin, China, in 2016. He has been a Visiting Scholar with the National University of Singapore from July 2013 to June 2014 and a Research Fellow with Tsinghua University from September 2016 to September 2017. He is currently a Research Fellow with the University of California at Berkeley, USA. His research interests include affective computing, multimedia, and computer vision.

Guiguang Ding received the Ph.D. degree in electronic engineering from Xidian University, China, in 2004. Before joining the School of Software in 2006, he was a Post-Doctoral Research Fellow with the Department of Automation, Tsinghua University. He is currently an Associate Professor with the School of Software, Tsinghua University. He has published 80 articles in major journals and conferences, including the IEEE TIP, TMM, TKDE, SIG IR, AAAI, ICML, IJCAI, CVPR, and ICCV. His current research centers on the areas of multimedia information retrieval, computer vision, and machine learning.



Weisi Lin (Fellow, IEEE) received the Ph.D. degree from King's College, London University, U.K. He is currently a Professor with the School of Computer Science and Engineering, Nanyang Technological University. His areas of expertise include image processing, perceptual signal modeling, video compression, and multimedia communication. He has published over 200 journal articles, over 230 conference articles, filed seven patents, and authored two books in these areas. He is a fellow of IET and an Honorary Fellow of the Singapore Institute of Engineering Technologists. He has been an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON MULTIMEDIA, and the IEEE SIGNAL PROCESSING LETTERS. He has been a Technical Program Chair of the IEEE ICME 2013, PCM 2012, QoMEX 2014 and IEEE VCIP 2017. He has been an invited/panelist/keynote/tutorial speaker in over 20 international conferences, as well as a Distinguished Lecturer of the IEEE Circuits and Systems Society from 2016 to 2017, and the Asia-Pacific Signal and Information Processing Association (APSIPA) from 2012 to 2013.