Personalized Image Aesthetics
(Supplementary Material)

1. Dataset

The images in Flickr Images with Aesthetics Annotation Dataset (FLICKR-AES) are downloaded from Flickr[1] and rated by 5 workers from Amazon Mechanical Turk[2](AMT). The average score of 5 workers, as shown in the main paper, servers as the ground truth and is normalized to the range of [0.2, 1]. The Real Album Curation Dataset (REAL-CUR) includes 14 albums and the aesthetics scores are given by the owners. Same as FLICKR-AES, the aesthetics scores range from 0.2 to 1. Figure 1 shows the score distributions of the two datasets, from which we can see the aesthetics scores of both datasets follow Gaussian approximately.

2. Qualitative Results

In this section, we use more visual examples to clearly illustrate our method.

To qualitatively analyze our generic aesthetics model, we display the images with high aesthetics and low aesthetics estimated by the model. As shown in Figure 2, the high aesthetics images typically present good aesthetic attributes, such as shallow depth of filed, rule of thirds, interesting content, etc. While images with low aesthetics are often containing defects like over saturation, image blur, etc.

As mentioned in the main paper, we propose an attributes network to provide aesthetic attributes scores. The ten aesthetic attributes include interesting content, object emphasis, good lighting, color harmony, vivid color, shallow depth of filed, rule of thirds, balancing element, repetition and symmetry. For each aesthetic attribute, the images from FLICKR-AES with high scores predicted by attributes network are shown in Figure 3a-3j.

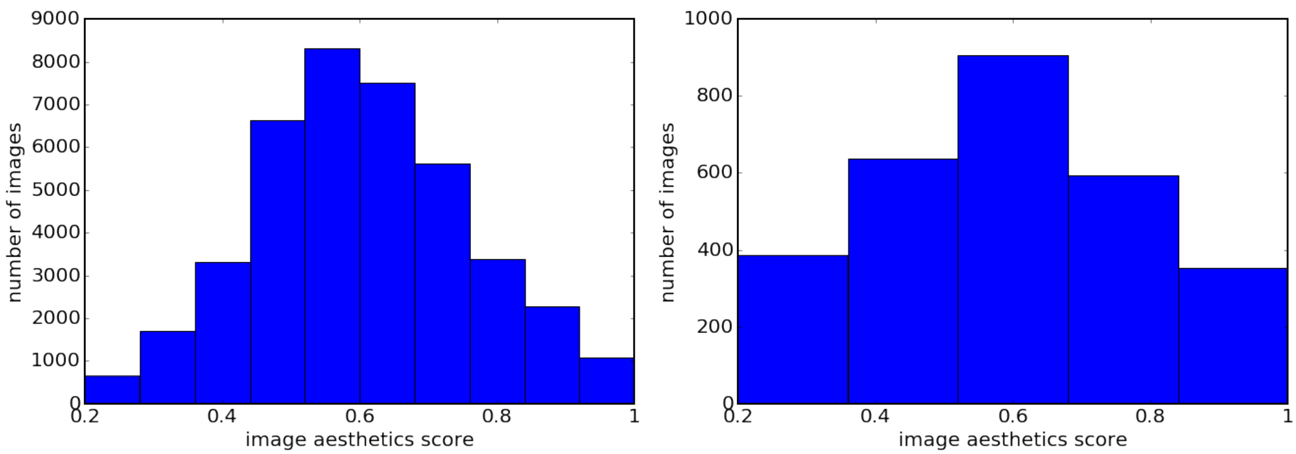
In the main paper, we mentioned the classes of training set for content network are generated by clustering the semantic features (avg pool) extracted from the classification network [1]. We show the examples of semantic content groups of FLICKR-AES in Figure 4. The images in different clusters share similar semantic content, such as landscape, plant, people, etc.

[1]https://www.flickr.com
[2]https://www.mturk.com

Besides the visual examples in the main paper proving how our personalized image aesthetics model works, we show more examples from FLICKR-AES and REAL-CUR here. We present the results of seven albums from FLICKR-AES in Figure 5a-5g and seven albums from REAL-CUR in Figure 6a-6g. From the results, we can see our personalized model more accurately predicts user ratings than the generic model.

References

[1] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015. 1

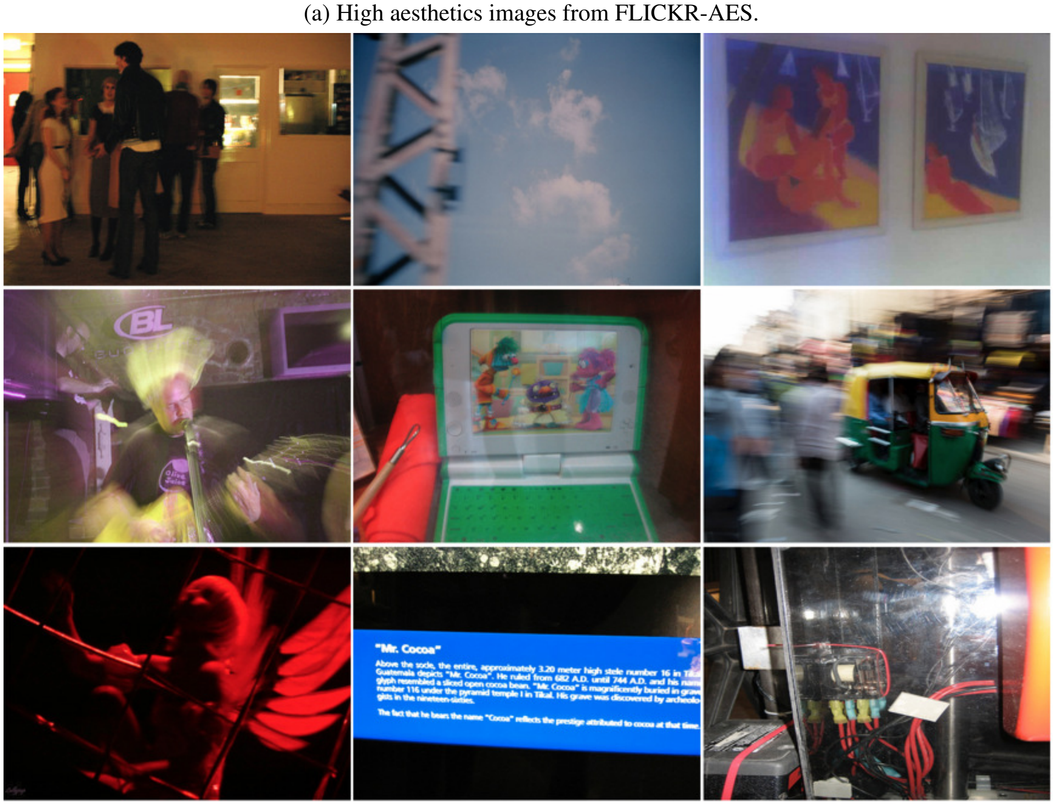


(a) Aesthetics score distribution of FLICKR-AES.

(b) Aesthetics score distribution of REAL-CUR.

Figure 1: The aesthetics scores of FLICKR-AES and REAL-CUR both fit in a Gaussian distribution.

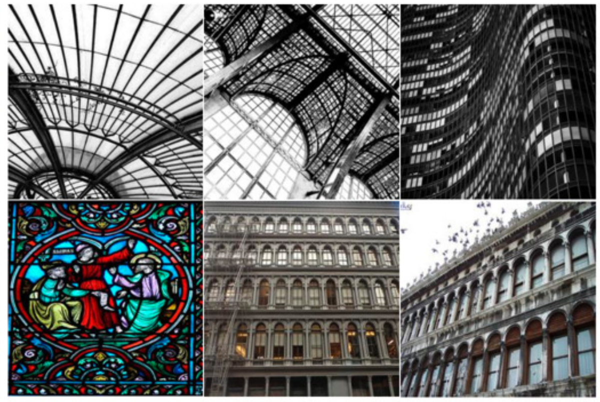
(a) High aesthetics images from FLICKR-AES.



(b) Low aesthetics images from FLICKR-AES.

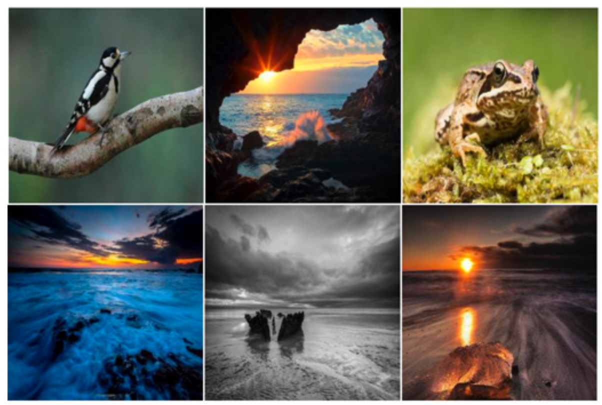
Figure 2: Examples of high aesthetics images and low aesthetics images from FLICKR-AES.

(a) Color harmony.



(b) Repetition.

(c) Object emphasis.

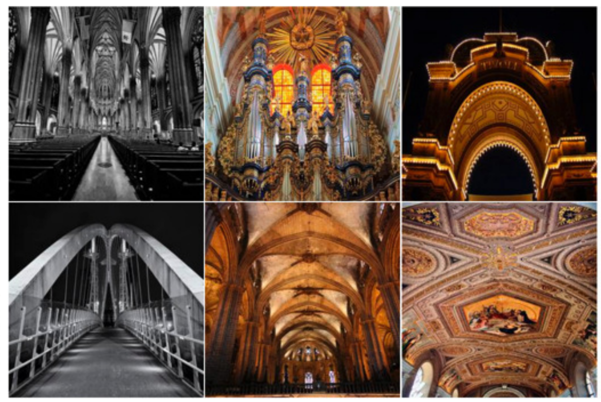


(d) Interesting content.

(e) Shallow depth of field.

(f) Vivid color.

(g) Rule of thirds.



(h) Symmetry.

(i) Good lighting.

(j) Balancing element.

Figure 3: Examples of images from FLICKR-AES with high aesthetic attributes scores, grouped by different attributes.

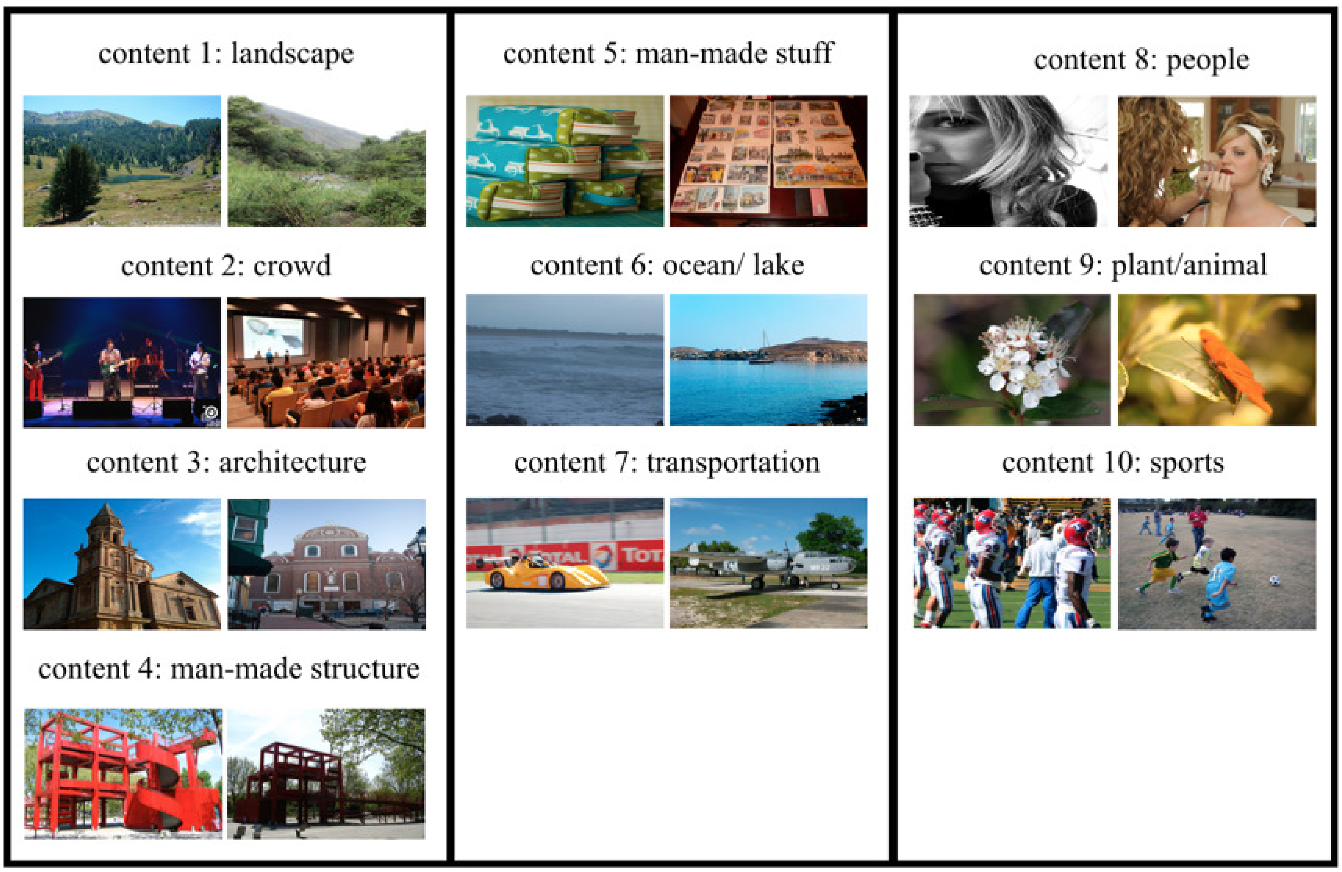
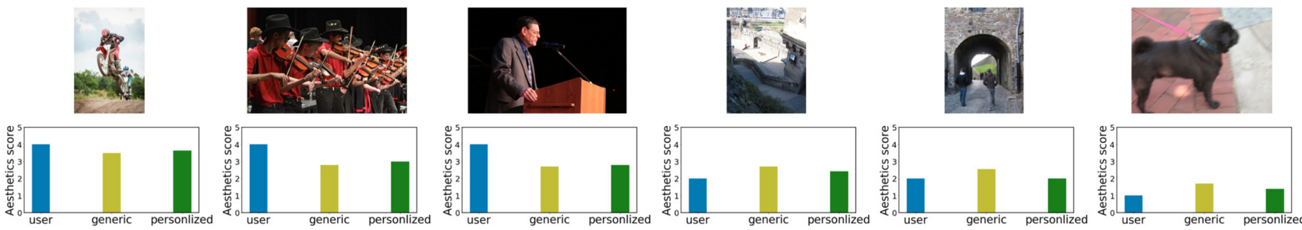
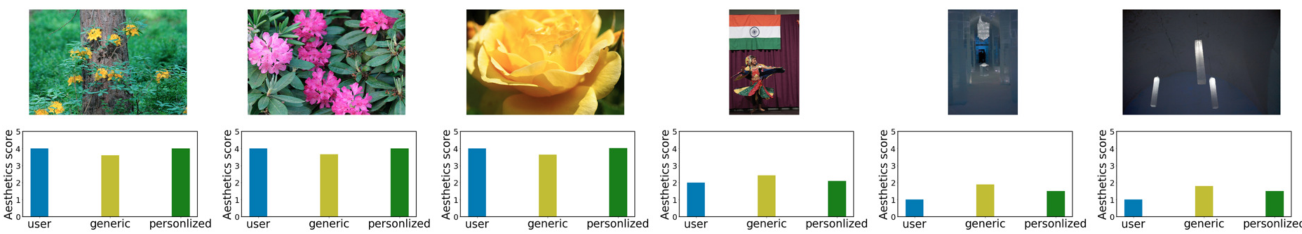


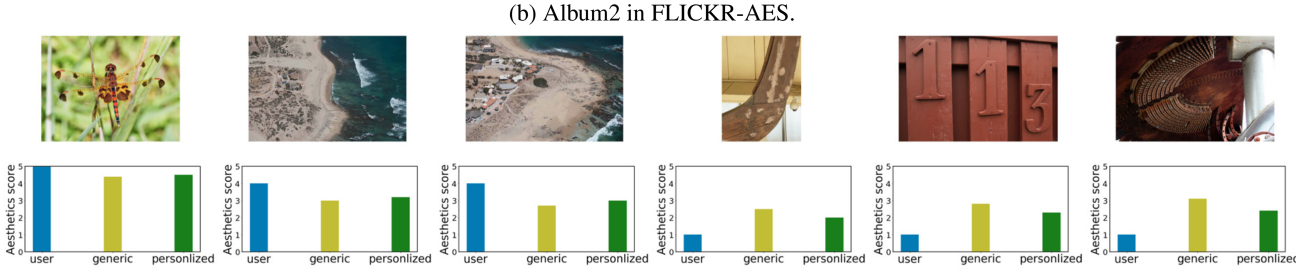
Figure 4: Examples from semantic content groups in the training set. Images in FLICKR-AES with similar thematic categories of content are clustered together without additional annotations.



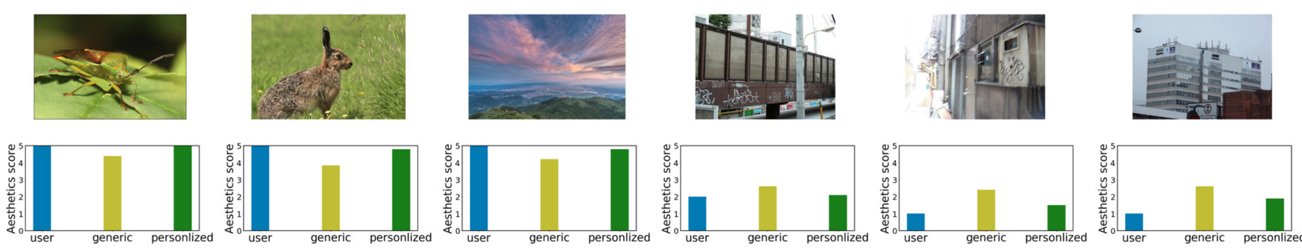
(a) Album1 in FLICKR-AES.



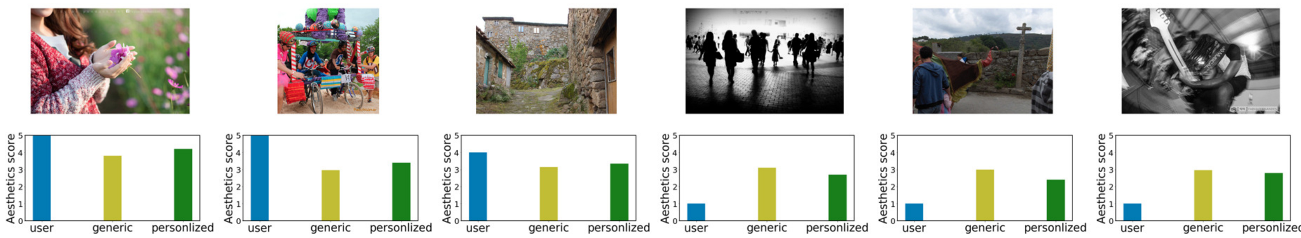
(b) Album2 in FLICKR-AES.



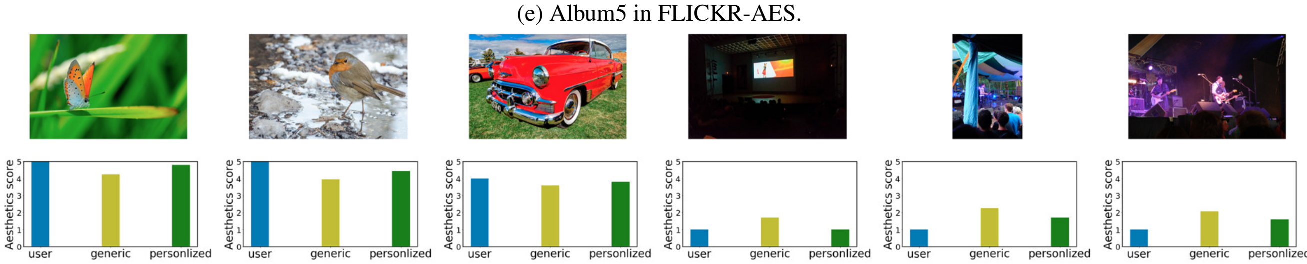
(c) Album3 in FLICKR-AES.



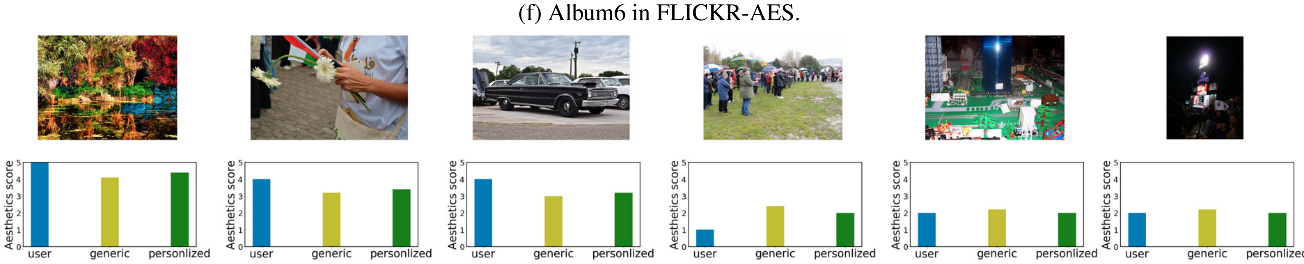
(d) Album4 in FLICKR-AES.



(e) Album5 in FLICKR-AES.

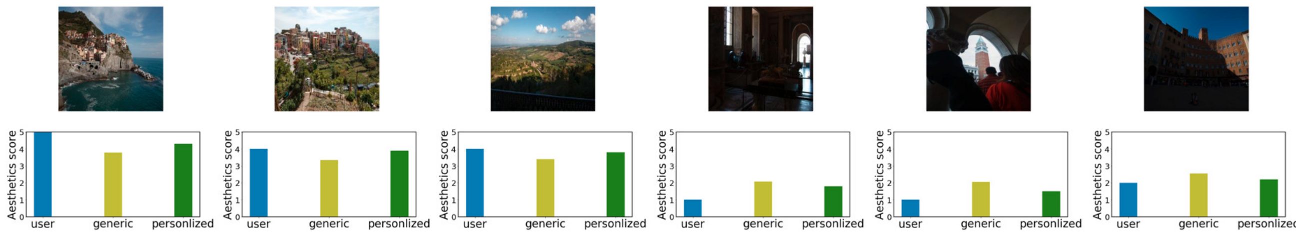


(f) Album6 in FLICKR-AES.

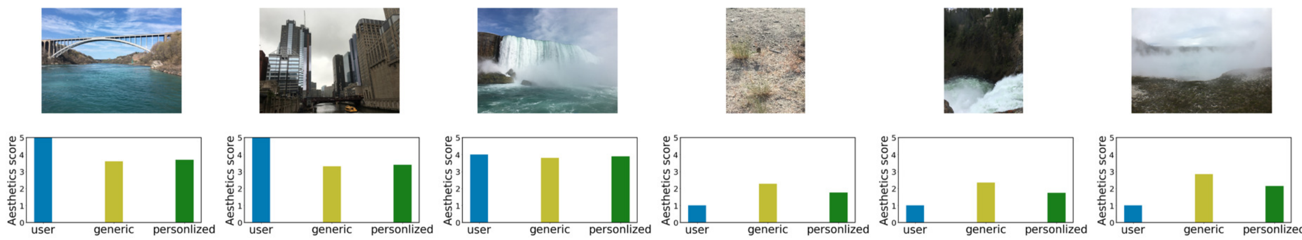


(g) Album7 in FLICKR-AES.

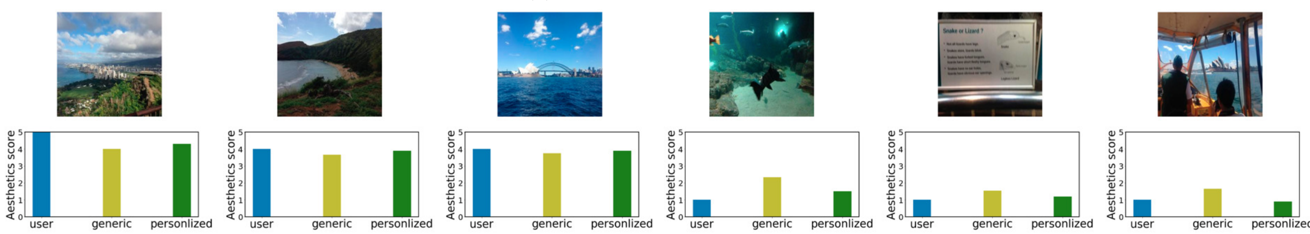
Figure 5: Example results of personalized aesthetics prediction from seven users. The examples come from FLICKR-AES. The blue bar is the user rating, the yellow bar is the generic aesthetics prediction and the green bar is the personalized aesthetics prediction for this user. As can be seen, our personalized model more accurately predicts user ratings than the generic model.



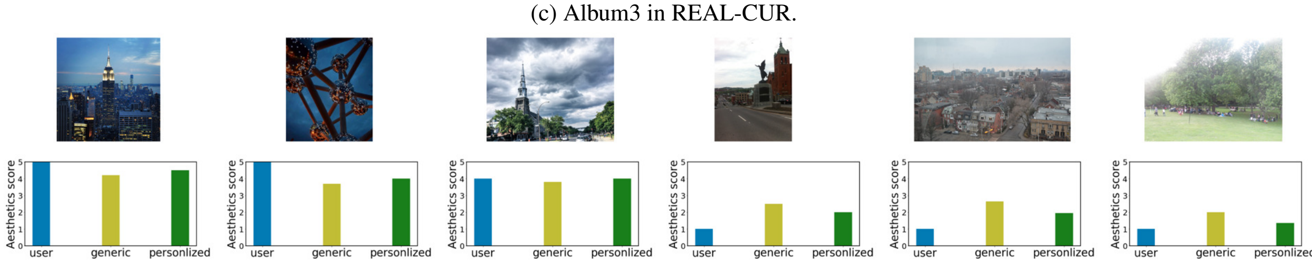
(a) Album1 in REAL-CUR.



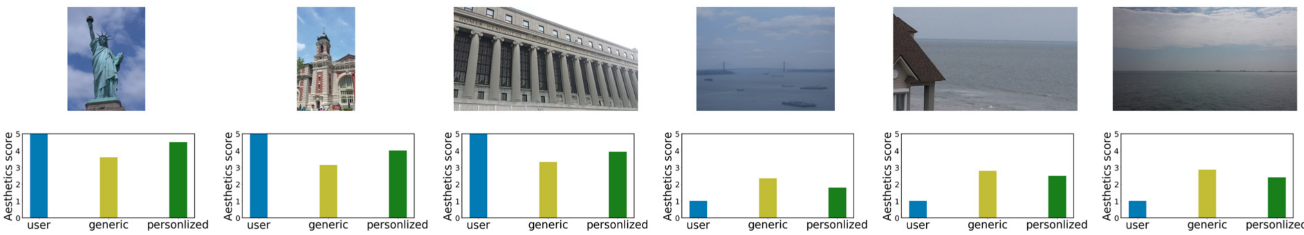
(b) Album2 in REAL-CUR.



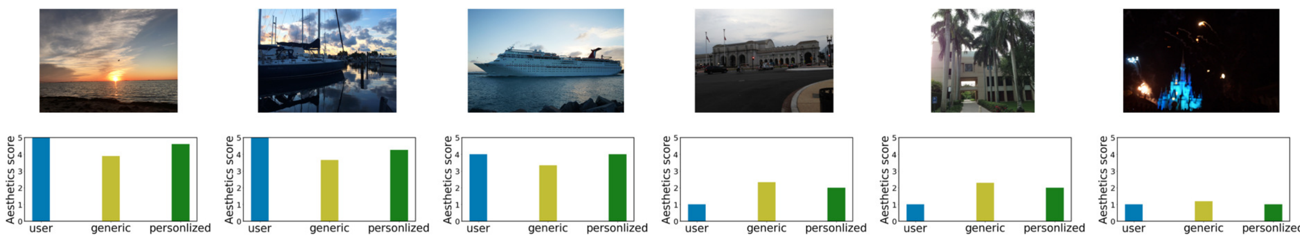
(c) Album3 in REAL-CUR.



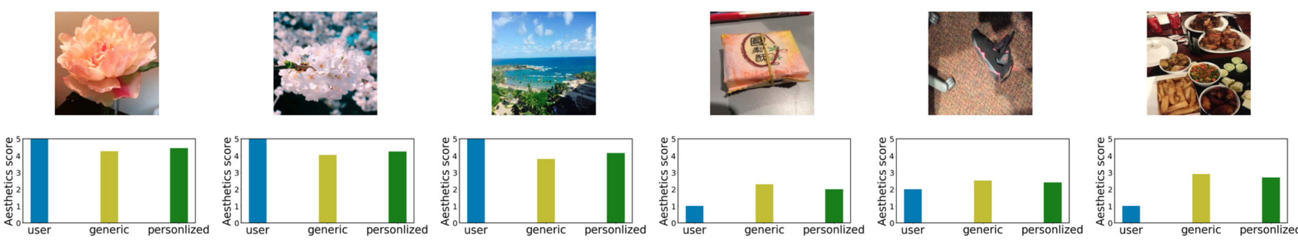
(d) Album4 in REAL-CUR.



(e) Album5 in REAL-CUR.



(f) Album6 in REAL-CUR.



(g) Album7 in REAL-CUR.

Figure 6: Example results of personalized aesthetics prediction from seven users. The examples come from REAL-CUR. The blue bar is the user rating, the yellow bar is the generic aesthetics prediction and the green bar is the personalized aesthetics prediction for this user. As can be seen, our personalized model more accurately predicts user ratings than the generic model.