AESTHETICS ASSESSMENT OF IMAGES CONTAINING FACES
Simone Bianco, Luigi Celona, Raimondo Schettini
University of Milano-Bicocca, {bianco, luigi.celona, schettini}@disco.unimib.it

Abstract
This paper introduces a method for aesthetic quality assessment of images with faces. We exploit three different Convolutional Neural Networks to encode information regarding perceptual quality, global image aesthetics, and facial attributes; then, a model is trained to combine these features to explicitly predict the aesthetics of images containing faces. Experimental results show that our approach outperforms existing methods for both binary, i.e. low/high, and continuous aesthetic score prediction on four different databases in the state-of-the-art.

Facial Image Aesthetic Estimation

Input image
Face detection
Features extraction
Learning/Inference

Aesthetic model
Facial features
Perceptual features
Global image aesthetics

Facial features
Alignment-Free Facial Attribute Classification Technique (AFFACT) [4], shortly FA: facial attributes encoding.

Perceptual features
• DeepBIQ model [2] (shortly IQ): perceptual quality metrics (noise, exposure, quality, JPEG quality, and sharpness);
• DeepIA model [3] (shortly IA): global image aesthetics concepts, such as composition, brightness, contrast, color, etc.

Genetic Algorithm (GA)

Experiments
Two sets of experiments:
• considering the whole image
• considering only the face regions

Evaluated in the following configurations:

For each experiment:
• 10-fold cross validation is performed by randomly selecting the training and testing images
• 10 repetitions to avoid sampling bias

Datasets
Four state-of-the-art databases:
• CUHKPQ: 3,148 photos annotated respectively with high and low aesthetic quality.
• HFS: 250 headshot photos (7 images of 20 subjects + 110 portrait images). Scores between 1 and 6 (average of 25 individual scores).
• FAVA: subset of the AVA dataset containing images with faces. Value between 1 and 10 (average of 210 individual scores).
• Flickr database: 500 images (portraits or group of faces). Scores range [0-10].

Results
Aesthetic quality estimation results for each database by extracting perceptual features from the whole image.

<table>
<thead>
<tr>
<th>#features</th>
<th>GA</th>
<th>FAVA</th>
<th>Flicker</th>
</tr>
</thead>
<tbody>
<tr>
<td>4,096</td>
<td>0.76</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>8,192</td>
<td>0.56</td>
<td>0.56</td>
<td>0.61</td>
</tr>
<tr>
<td>6,144</td>
<td>0.51</td>
<td>0.51</td>
<td>0.56</td>
</tr>
<tr>
<td>4,096</td>
<td>0.57</td>
<td>0.57</td>
<td>0.61</td>
</tr>
<tr>
<td>2,048</td>
<td>0.51</td>
<td>0.51</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Comparison with state-of-the-art methods for all the considered databases.

Methods

<table>
<thead>
<tr>
<th>Methods</th>
<th>CUHKPQ GCR (%)</th>
<th>HFS GCR (%)</th>
<th>LCC</th>
<th>FAVA GCR (%)</th>
<th>LCC</th>
<th>Flickr GCR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lienhard</td>
<td>94.8</td>
<td>79.3</td>
<td>0.75</td>
<td>65.1</td>
<td>0.75</td>
<td>49.9</td>
</tr>
<tr>
<td>Kairanbay</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Proposed</td>
<td>98.2</td>
<td>79.0*</td>
<td>0.76*</td>
<td>71.2</td>
<td>0.61</td>
<td>74.0 0.61</td>
</tr>
</tbody>
</table>

*These results are obtained by extracting perceptual features from face region.

References