Dual Variational Generation for Low Shot Heterogeneous Face Recognition

Chaoyou Fu1,2,3, Xiang Wu1,2, Yibo Hu1,2, Huabo Huang1,2, Ran He1,2,3,4
1 Center for Research on Intelligent Perception and Computing, CASIA
2 National Laboratory of Pattern Recognition, CASIA
3 University of Chinese Academy of Sciences
4 Center for Excellence in Brain Science and Intelligence Technology, CAS

Background

- Heterogeneous Face Recognition is a challenging issue because of the large domain discrepancy and a lack of heterogeneous data
- Previous image-to-image translation based methods face two challenges
  - Diversity: Given one image, a generator only synthesizes one new image of the target domain, resulting in limited number of images. Moreover, two images before and after translation have some attributes except for the spectral information, leading to limited intra-class diversity
  - Consistency: When generating large-scale samples, it is challenging to guarantee that the synthesized face images belong to the same identity of the input images

Objective

- Lean the joint distribution
  \[ L_{\text{rec}} = -D_{\text{G}(\delta)}(x_t | x_s, y_s) \log \hat{p}_G(x_t | x_s, y_s) \]
  \[ L_{\text{m}} = D_{\text{KL}(p_G(y_s | x_s | x_t) | p(x_s) + D_{\text{KL}(p_G(y_s | x_s | x_t) | p(x_t))} \]
- Align the distributions
  \[ L_{\text{id}} = \frac{1}{2} (\|y_0^1 - y_0^2\|^2 + \|y_1^1 - y_1^2\|^2) \]
- Pairwise Identity Preserving
  \[ L_{\text{pair}} = \|F_{\text{ID}}(x_t) - F_{\text{ID}}(x_s)\|^2 \]
  \[ L_{\text{pair}} = \|F_{\text{ID}}(x_t) - F_{\text{ID}}(x_s)\|^2 + \|F_{\text{ID}}(x_s) - F_{\text{ID}}(x_t)\|^2 \]

The purpose (left part) and training model (right part). It generates large-scale new paired heterogeneous images with the same identity from standard Gaussian noise, aiming at reducing the domain discrepancy for HFR. A distribution alignment in the latent space and a pairwise identity preserving in the image space are imposed to guarantee the identity consistency of the generated paired images

Dual Variational Generation

- Generate paired new heterogeneous data from noise
- Sample large-scale new images with abundant intra-class diversity
- Ensure the identity consistency of the generated paired images

Quantitative Results

<table>
<thead>
<tr>
<th>Method</th>
<th>CASIA NIR-VIS 2.0</th>
<th>Oulu CASIA NIR-VIS</th>
<th>BUAA-VIS-NIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank-1</td>
<td>82.1 ± 0.9</td>
<td>74.5</td>
<td>-</td>
</tr>
<tr>
<td>Rank-2</td>
<td>85.2 ± 0.9</td>
<td>75.8</td>
<td>-</td>
</tr>
<tr>
<td>Rank-3</td>
<td>89.6 ± 0.9</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

- We provide a new insight into the problems of HFR. That is, we consider HFR as a dual generation problem, and propose a novel dual variational generation framework. This framework generates new paired heterogeneous images with abundant intra-class diversity
- We can sample large-scale diverse paired heterogeneous images from noise. By constraining the pairwise feature distances of the generated paired images in the HFR network, the domain discrepancy is effectively reduced

Quantitative Results

- We can sample large-scale diverse paired heterogeneous images from noise. By constraining the pairwise feature distances of the generated paired images in the HFR network, the domain discrepancy is effectively reduced

More Experiments

- Tufts Face database: Baseline: Rank-1 = 37.3%, DVG: Rank-1 = 53.9%, Improving 15.6%
- HIB 0 Viewed Sketch database: Baseline: VPR=0.8%, DVG: VPR=96%, Improving 98.2%
- Multi-PF database: Baseline: Rank-1 = 65.4%, DVG: Rank-1 = 83.9 %, Improving 18.5%

Contributions

- We provide a new insight into the problems of HFR. That is, we consider HFR as a dual generation problem, and propose a novel dual variational generation framework. This framework generates new paired heterogeneous images with abundant intra-class diversity
- We can sample large-scale diverse paired heterogeneous images from noise. By constraining the pairwise feature distances of the generated paired images in the HFR network, the domain discrepancy is effectively reduced

Dual Variational Generation

- Generate paired new heterogeneous data from noise
- Sample large-scale new images with abundant intra-class diversity
- Ensure the identity consistency of the generated paired images

Image results

- Large-scale new images with same identity and abundant intra-class diversity

Visual Results

CASIA NIR-VIS 2.0 | Oulu | BUAA