Detail Fusion GAN: High-Quality Translation for Unpaired Images with GAN-based Data Augmentation

Ling Li¹  Yaochen Li¹  Chuan Wu¹  Hang Dong²  and Peilin Jiang¹

¹ School of Software Engineering, Xi’an Jiaotong University, Xi’an, China
² Institute of Artificial Intelligence and Robotics, Xi’an Jiaotong University, Xi’an, China
Goal

- High-Quality Translation for Unpaired Images
Challenge

INPUT

TRANSLATED
Challenge

Original image \((I_{ori})\)

Translated image \((I_{tra})\)

Structural Information
Related Work

CycleGAN

UNIT

MUNIT

DRIT
Contribution

- Introducing super-resolution loss as a guidance
- Two branches and a filter module
- Generating better results and converging faster
Network Architecture

Transfer Branch

Detail Branch

Filter Module

Reconstruction Module
Datasets

SYNTHIA

GTA

BDD

Cityscapes
## Ablation Study

<table>
<thead>
<tr>
<th>Model</th>
<th>Transfer Branch</th>
<th>Detail Branch</th>
<th>Filter Module</th>
<th>Reconstruction Module</th>
<th>Input</th>
<th>Output</th>
<th>FID / SSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CycleGAN</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>256×256</td>
<td>256×256</td>
<td>62.53 / 0.62</td>
</tr>
<tr>
<td>Model-1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>256×256</td>
<td>512×512</td>
<td>56.30 / 0.76</td>
</tr>
<tr>
<td>Model-2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>256×256</td>
<td>512×512</td>
<td>54.51 / 0.79</td>
</tr>
<tr>
<td>DFGAN-s</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>128×128</td>
<td>256×256</td>
<td>48.87 / 0.84</td>
</tr>
<tr>
<td>DFGAN</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>256×256</td>
<td>512×512</td>
<td>47.76 / 0.86</td>
</tr>
</tbody>
</table>

- The lower FID, the better
- The higher SSIM, the better.
Evaluation

(a) Input EPOCH / SSIM
7 / 0.50
(b) CycleGAN [6]
49 / 0.63
(c) UNIT [9]
30 / 0.79
(d) MUNIT [10]
35 / 0.74
(e) DRIT [13]
1199 / 0.36
(f) DFGAN(ours)
7 / 0.85
## Evaluation

### Table

<table>
<thead>
<tr>
<th>Method</th>
<th>Input</th>
<th>Output</th>
<th>acc. of FCN</th>
<th>AP of YOLO</th>
<th>FID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$I_{ori}$</td>
<td>$I_{hr}$</td>
<td>$I_{tra}$</td>
<td>$I_{sr}$</td>
<td>Day</td>
</tr>
<tr>
<td>Daytime Only</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.7946</td>
</tr>
<tr>
<td>CycleGAN</td>
<td>256×256</td>
<td>-</td>
<td>256×256</td>
<td>-</td>
<td>0.8223</td>
</tr>
<tr>
<td>UNIT</td>
<td>256×256</td>
<td>-</td>
<td>256×256</td>
<td>-</td>
<td>0.7983</td>
</tr>
<tr>
<td>MUNIT</td>
<td>256×256</td>
<td>-</td>
<td>256×256</td>
<td>-</td>
<td>0.8393</td>
</tr>
<tr>
<td>DRIT</td>
<td>256×256</td>
<td>-</td>
<td>256×256</td>
<td>-</td>
<td>0.8389</td>
</tr>
<tr>
<td>DFGAN-s</td>
<td>128×128</td>
<td>256×256</td>
<td>128×128</td>
<td>256×256</td>
<td>0.8421</td>
</tr>
<tr>
<td>DFGAN</td>
<td>256×256</td>
<td>512×512</td>
<td>256×256</td>
<td>512×512</td>
<td>0.8518</td>
</tr>
</tbody>
</table>

- The lower FID, the better.
- The higher ACC. of FCN or AP of YOLO, the better.
Image translation cases