Multi-Laplacian GAN with Edge Enhancement for Face Super Resolution

Shanlei Ko, Bi-Ru Dai
Outline

▪ Introduction
▪ Proposed Method
▪ Experiments
▪ Conclusions and Future Works
Introduction

- **Super Resolution (SR) Problem**
  - Generate a high-resolution (HR) face image from an LR one
  - Why is it important?
    - Applications: Remote sensing, medical diagnostic, intelligent surveillance etc.
    - Recover details of HR from LR

- **Face SR Major challenges**
  - Missing information: Identities, five senses, facial attributes, etc
  - Ill-posed problems: two different high resolution images are possible to be downsamped to two similar low resolution images.

![Diagram showing Super Resolution process]

- 8 × Downsampling
- 16 × 16 LR
- 8 × SR
- 128 × 128 SR

128 × 128 HR
Contribution

- Propose the Multi-Laplacian GAN with Edge enhancement (MLGE)

- MLGE
  - Model Generalization
    - No additional prior facial information is required in both training and testing phases for our model
  - Enhance the quality
    - MLGE achieves higher quantitative and qualitative performance.
Outline

- Introduction
- Proposed Method
- Experiments
- Conclusions and Future Works
MLGE Architecture-Discriminative Branch

- The general discriminator $D_{h^8}$ determines whether
  - the $8 \times$ SR $\hat{h}^8$ is similar to HR face images $h^8$
- The objective function for the general discriminator is expressed as:
  - $L_{D_{h^8}} = -E[\log D_{h^8}(h^8) + \log(1 - D_{h^8}(\hat{h}^8))]$
MLGE Architecture-Discriminative Branch

- Two edge discriminators $D_{e^s}$ determine whether
  - the gradients of SR images $\hat{e}^s$ are similar to the gradient of HR face images $e^s$ in each scale $s$.
- The objective function for general discriminator is expressed as:
  - $L_{D_{e^s}} = -E[\log D_{e^s}(e^s) + \log (1 - D_{e^s}(\hat{e}^s)), s = 4, 8$
Outline

- Introduction
- Related Works
- Proposed Method
- Experiments
- Conclusions and Future Works
Experiment Settings

▪ Datasets:
  ▪ Celebrity Face Attributes (CelebA) dataset
    ▪ 202,599 human face images with labeled 40 facial attributes, landmarks.

▪ Training data and testing data:
  ▪ 22K face images as training
  ▪ 2.6K face images as testing
  ▪ Resize to $128 \times 128$ as our HR images
  ▪ Downsample HR images to $16 \times 16$ as our LR images
Experiment Settings (cont.)

- **Comparison methods**
    - Traditional upsampling method
  - VDSR (CVPR 2016)
    - CNN based method
  - SRGAN (CVPR 2017)
    - GAN-based method
  - Yu et al (CVPR 2018) (attributes)
    - Conditional GAN based
    - Training images only include face region

- **Evaluate metrics**
  - PSNR (The average Peak Signal to Noise Ratio)
  - SSIM (The Structure Similarity Score)
Method comparison

- Quantitative Comparison
  - MLGE outperforms other methods with large margins in PSNR and SSIM

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>21.76</td>
<td>23.12</td>
<td>23.76</td>
<td>22.85</td>
<td>25.07</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.72</td>
<td>0.80</td>
<td>0.78</td>
<td>0.753</td>
<td>0.83</td>
</tr>
</tbody>
</table>

- Qualitative Comparison
Method comparison

- **Quantitative Comparison**
  - MLGE outperforms other methods with large margins in PSNR and SSIM

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>21.76</td>
<td>23.12</td>
<td>23.76</td>
<td>22.85</td>
<td><strong>25.07</strong></td>
</tr>
<tr>
<td>SSIM</td>
<td>0.72</td>
<td>0.80</td>
<td>0.78</td>
<td>0.753</td>
<td><strong>0.83</strong></td>
</tr>
</tbody>
</table>
Conclusions and Future Works

- **MLGE algorithm**
  - Generalization
    - No additional prior facial information is required in both training and testing phases
  - Enhance the quality
    - Outperform the state-of-the-arts both in quantitatively and qualitatively
    - Generates SR images which have almost the same facial attributes as HR images

- **Future works**
  - Explore better information extraction methods
    - Principal component analysis (PCA) and fast fourier transform
    - Extend MLGE to large scale SR task, such as 256 × 256 LR images
Thank you