Convolutional Neural Network with Inception Blocks for Image Compression Artifact Reduction

Purbaditya Bhattacharya, Udo Zölzer
Dept. of Signal Processing and Communication, Helmut Schmidt University, Hamburg, Germany
Email: bhattacp@hsu-hh.de
**Introduction**

- JPEG performs a lossy compression of an image based on quantization and coding of its BDCT coefficients.
- Different compression ratios invoke different quantization tables.
- For higher compression ratios, aggressive quantization leads to multiple artifacts.
JPEG Artifacts

- High compression ratios lead to the following major artifacts.

- **Blurring of image patches** - Suppression of high frequency information due to low bit allocation, leads to redundant pixels.

- **Blocking artifacts** – Non overlapping block based transform and independent quantization leads to visible discontinuities along block edges.

- **Ringing** – Poor quantization of high frequency information leads to a coarse transition across object boundaries and introduces artifacts.
### JPEG Artifacts

<table>
<thead>
<tr>
<th>Original</th>
<th>JPEG (Quality: 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.jpg" alt="Original Image" /></td>
<td><img src="image2.jpg" alt="JPEG Image (Quality: 10)" /></td>
</tr>
</tbody>
</table>

SPONSORS:
Example based deep learning models are efficient in reducing these artifacts and restoring the image.

Supervised convolutional neural networks (CNNs) trained particularly for image restoration problems have outperformed classical image restoration methods.

The basis network known as Denoising CNN (DnCNN) is a residual CNN model with batch normalization (BN) for image enhancement, introduced by Zhang et.al.
Proposed Network
Network Modules

**Filter Dilation**

- d0
- d1

**Inception Block (with BN)**

- 3x3xDxD, d0 → Conv + BN
- 3x3xDxD/2, d1 → Conv + BN
- 3x3xDxD/4, d2 → Conv + BN
- 3x3xDxD/8, d4 → Conv + BN
- 3x3xDxD/8, d6 → Conv + BN

- Concat
- ReLU
- 2D

- dn : Dilation with n zeros

**Conv Block (with BN)**

- 2D → Conv
- D → Batch Norm (BN) → ReLU → D

**SPONSORS:**
## Experiments

Performance over network depth

<table>
<thead>
<tr>
<th># Incp+Conv Blocks with BN</th>
<th>Classic5</th>
<th>LIVE1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR (dB)</td>
<td>SSIM</td>
</tr>
<tr>
<td>2</td>
<td>29.46</td>
<td>0.805</td>
</tr>
<tr>
<td>4</td>
<td>29.53</td>
<td>0.808</td>
</tr>
<tr>
<td>8</td>
<td>29.56</td>
<td>0.809</td>
</tr>
</tbody>
</table>

![Graph showing performance over network depth]
Experiments

Performance over feature map depth

![Graph showing PSNR (dB) over epochs for two architectures: Arch1 and Arch2.](image-url)
Results

![PSNR (dB) Chart]

- JPEG
- ARCNN
- DnCNN
- CNN (Proposed)

![SSIM Chart]

- JPEG
- ARCNN
- DnCNN
- CNN (Proposed)
Results

<table>
<thead>
<tr>
<th>Original</th>
<th>JPEG (Quality: 10)</th>
<th>DnCNN</th>
<th>CNN (Proposed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Original Image" /></td>
<td><img src="image2.png" alt="JPEG Image" /></td>
<td><img src="image3.png" alt="DnCNN Image" /></td>
<td><img src="image4.png" alt="CNN Image" /></td>
</tr>
<tr>
<td><img src="image5.png" alt="Original Image" /></td>
<td><img src="image6.png" alt="JPEG Image" /></td>
<td><img src="image7.png" alt="DnCNN Image" /></td>
<td><img src="image8.png" alt="CNN Image" /></td>
</tr>
</tbody>
</table>
Conclusion

Summary

- Introduction of inception blocks in the basis network improves the restoration task.
- Parallel convolutions with filters of multiple resolution aggregate multiscale context.
- Dilated convolution increases the receptive field but not the number of filter parameters.
- The performance improves with increasing network and feature map depths.

Future Work

- Further improvements can be obtained by introducing dense skip connections, recursive blocks and improved inception modules.
- Such an approach can be applied to other enhancement tasks like superresolution and inpainting.