Burst Image Restoration and Enhancement

Fahad Khan
MBZUAI
Linköping University
Single Image Processing

- Single image quality is limited due to the imaging sensor.
- Thus, limited information in captured in single image.
- Restoration/enhancement approaches are limited to improve the quality of single image.
Burst Image Processing

```
Burst of low-quality Images
```

**Burst Image Processing**

- Multiple sub-sampled images of the same scene.
- More information than single image.
- Easy to capture burst image due to smartphone advancement. E.g., live photos.
- Multi-frame information helps to get precise high-quality image compared to the single image.

**Challenges:**

- Misalignment within multiple frames.
- How to merge multi-frame information to get single high-quality image.
- How to design a unified model that works in multiple scenarios.
- How to design compact and efficient model.
Burst Image Processing

Existing Architecture Designs

Explicit alignment and late feature fusion

Explicit flow-estimation → Feature Extraction → Feature Warping → Weighted Fusion → Up-sampling

Issues

- Error propagation
- No access to multi-frame information in up-sampling
- No communication within frames
Burst Image Processing

Existing Architecture Designs

Adaptive feature extraction, Implicit alignment

- Kernel prediction
- Adaptive Feature Extraction
- Feature Alignment
- Feature average
- Up-sampling

Issues
- Bulky in nature
- No communication within frames
- No access to multi-frame information in up-sampling
Proposed Approach: BIPNet

- An edge boosting alignment technique that removes spatial and color misalignment issues among the burst features.
- A novel pseudo-burst feature fusion mechanism to enable inter-frame communication and feature consolidation.
- An adaptive group upsampling module for progressive fusion and upscaling.
Proposed Approach: BIPNet

RAW LR Burst $B \times 4 \times H \times W$

Edge Boosting Feature Alignment $B \times f \times H \times W$

Feature Alignment $f \times f \times H \times W$

Feature Fusion $f \times f \times H \times W$

Pseudo Burst $f \times f \times H \times W$

Adaptive Group Up-sampling Module $1 \times f \times 8H \times 8W$

Output Image $1 \times 3 \times 8H \times 8W$
Proposed Approach: BIPNet

RAW LR Burst $B \times 4 \times H \times W$

- **Edge Boosting Feature Alignment**
- **Pseudo Burst Feature Fusion**
- **Adaptive Group Up-sampling**

Output HR Image $1 \times 3 \times 8H \times 8W$
Proposed Approach: BIPNet

 RAW LR Burst \( B \times 4 \times H \times W \)

 Edge Boosting
 Feature Alignment

 Pseudo Burst
 Feature Fusion

 Adaptive Group
 Up-sampling

 Output HR Image
 \( 1 \times 3 \times 8H \times 8W \)
Proposed Approach: BIPNet

- Generates feature tensors by concatenating corresponding channel-wise features from all burst feature maps.
- Each feature tensor in the pseudo-burst contains complimentary properties of all actual burst image features.
- Processing inter-burst feature responses simplifies the representation learning task and merges the relevant information by decoupling the burst image feature channels.
Proposed Approach: BIPNet

- Effectively utilizes the information available in multiple frames to get into HR space.
- Handles pseudo-bursts features in groups and progressively performs upscaling.
- Sequentially divide pseudo-burst features into groups, which are up-sampled with up-sampler.
Experiments

- Burst Super-resolution
- Burst Low-light Image Enhancement
- Burst Denoising
Experiments

Burst Image Super-resolution

**Synthetic Burst Super-resolution**
- Dataset: SyntheticBurst
- Train split (46839 synthetic bursts)
- Validation split (300 synthetic bursts)

**Real Burst Super-resolution**
- Dataset: BurstSR
- Train split (5405 real burst patches)
- Validation split (882 real burst patches)

**Training details**
- Loss function: $L_1$
- Optimizer: ADAM
- Schedular: Cosine annealing strategy
- Learning rate: $10^{-4}$ to $10^{-6}$
- Network parameters: 6.67 M
Experiments

Quantitative comparison for x4 Burst SR

<table>
<thead>
<tr>
<th>Methods</th>
<th>SyntheticBurst</th>
<th>(Real) BurstSR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR</td>
<td>SSIM</td>
</tr>
<tr>
<td>Single Image</td>
<td>36.17</td>
<td>0.909</td>
</tr>
<tr>
<td>HighRes-Net [1]</td>
<td>37.45</td>
<td>0.92</td>
</tr>
<tr>
<td>DBSR [2]</td>
<td>40.76</td>
<td>0.96</td>
</tr>
<tr>
<td>LKR [3]</td>
<td>41.45</td>
<td>0.95</td>
</tr>
<tr>
<td>MFIR [4]</td>
<td>41.56</td>
<td>0.96</td>
</tr>
<tr>
<td>BIPNet (Ours)</td>
<td>41.93</td>
<td>0.96</td>
</tr>
</tbody>
</table>

## Experiments

### Visual comparison for SyntheticBurst SR

<table>
<thead>
<tr>
<th>Methods</th>
<th>SyntheticBurst</th>
<th>(Real) BurstSR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR</td>
<td>SSIM</td>
</tr>
<tr>
<td>Single Image</td>
<td>36.17</td>
<td>0.909</td>
</tr>
<tr>
<td>HighRes-Net [1]</td>
<td>37.45</td>
<td>0.92</td>
</tr>
<tr>
<td>DBSR [2]</td>
<td>40.76</td>
<td>0.96</td>
</tr>
<tr>
<td>LKR [3]</td>
<td>41.45</td>
<td>0.95</td>
</tr>
<tr>
<td>MFIR [4]</td>
<td>41.56</td>
<td>0.96</td>
</tr>
<tr>
<td>BIPNet (Ours)</td>
<td>41.93</td>
<td>0.96</td>
</tr>
</tbody>
</table>

### Visual comparison for RealBurst SR


## Experiments

### Quantitative comparison for x4 Burst SR

<table>
<thead>
<tr>
<th>Methods</th>
<th>SyntheticBurst</th>
<th>(Real) BurstSR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR</td>
<td>SSIM</td>
</tr>
<tr>
<td>Single Image</td>
<td>36.17</td>
<td>0.909</td>
</tr>
<tr>
<td>HighRes-Net [1]</td>
<td>37.45</td>
<td>0.92</td>
</tr>
<tr>
<td>DBSR [2]</td>
<td>40.76</td>
<td>0.96</td>
</tr>
<tr>
<td>LKR [3]</td>
<td>41.45</td>
<td>0.95</td>
</tr>
<tr>
<td>MFIR [4]</td>
<td>41.56</td>
<td>0.96</td>
</tr>
<tr>
<td>BIPNet (Ours)</td>
<td>41.93</td>
<td>0.96</td>
</tr>
</tbody>
</table>

### Visual comparison for x8 Burst SR

(a) Base frame (b) Ours (c) Ground-truth

---


Experiments

Burst Low-Light Image Enhancement

- Dataset: SID dataset (Sony subset)
- Train split (28k patches generated from 161 burst sequences)
- Validation split (50 burst sequences)

Training details

- Loss function: $L_1$
- Optimizer: ADAM
- Scheduler: Cosine annealing strategy
- Learning rate: $10^{-4}$ to $10^{-6}$
- Epochs: 50
- Network parameters: 6.67 M
### Experiments

#### Quantitative comparison for Burst low-light image enhancement

<table>
<thead>
<tr>
<th>Methods</th>
<th>PSNR</th>
<th>SSIM</th>
<th>LPIPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen et al. [1]</td>
<td>29.38</td>
<td>0.892</td>
<td>0.484</td>
</tr>
<tr>
<td>Maharajan et al. [2]</td>
<td>29.57</td>
<td>0.891</td>
<td>0.484</td>
</tr>
<tr>
<td>Zamir et al. [3]</td>
<td>29.13</td>
<td>0.881</td>
<td>0.462</td>
</tr>
<tr>
<td>Zhao et al. [4]</td>
<td>29.49</td>
<td>0.895</td>
<td>0.455</td>
</tr>
<tr>
<td>Karadeniz et al. [5]</td>
<td>29.80</td>
<td>0.891</td>
<td>0.306</td>
</tr>
<tr>
<td>BIPNet (Ours)</td>
<td>32.87</td>
<td>0.936</td>
<td>0.305</td>
</tr>
</tbody>
</table>

Experiments

Burst Image Denoising

Grayscale burst denoising

- Train split (20k from Open Images set)
- Validation split (73 Grayscale synthetic bursts)

Color burst denoising

- Train split (20k from Open Images set)
- Validation split (100 Color synthetic bursts)

Training details

- Loss function: $L_1$
- Optimizer: ADAM
- Scheduler: Cosine annealing strategy
- Learning rate: $10^{-4}$ to $10^{-6}$
- Epochs: 50
- Network parameters: 6.67 M
Experiments

### Quantitative comparison for Grayscale Burst De-noising

<table>
<thead>
<tr>
<th>Method</th>
<th>Gain $\alpha_1$</th>
<th>Gain $\alpha_2$</th>
<th>Gain $\alpha_4$</th>
<th>Gain $\alpha_8$</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>KPN [1]</td>
<td>36.47</td>
<td>33.93</td>
<td>31.19</td>
<td>27.97</td>
<td>32.39</td>
</tr>
<tr>
<td>BPN [2]</td>
<td>38.18</td>
<td>35.42</td>
<td>32.54</td>
<td>29.45</td>
<td>33.9</td>
</tr>
<tr>
<td>MFIR [3]</td>
<td>39.37</td>
<td>36.51</td>
<td>33.38</td>
<td>29.69</td>
<td>34.74</td>
</tr>
<tr>
<td>BIPNet (Ours)</td>
<td>41.26</td>
<td>38.74</td>
<td>35.91</td>
<td>31.35</td>
<td>36.81</td>
</tr>
</tbody>
</table>

### Quantitative comparison for Color Burst De-noising

<table>
<thead>
<tr>
<th>Method</th>
<th>Gain $\alpha_1$</th>
<th>Gain $\alpha_2$</th>
<th>Gain $\alpha_4$</th>
<th>Gain $\alpha_8$</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>KPN [1]</td>
<td>38.86</td>
<td>35.97</td>
<td>32.79</td>
<td>30.01</td>
<td>34.41</td>
</tr>
<tr>
<td>BPN [2]</td>
<td>40.16</td>
<td>37.08</td>
<td>33.81</td>
<td>31.19</td>
<td>35.56</td>
</tr>
<tr>
<td>MFIR [3]</td>
<td>42.21</td>
<td>39.13</td>
<td>35.75</td>
<td>32.52</td>
<td>37.4</td>
</tr>
<tr>
<td>BIPNet (Ours)</td>
<td>42.28</td>
<td>40.2</td>
<td>37.85</td>
<td>34.64</td>
<td>38.74</td>
</tr>
</tbody>
</table>

[1] Ben Mildenhall et al., *Burst denoising with kernel prediction networks*. In CVPR, 2018
[2] Zhihao Xia et al., *Basis prediction networks for effective burst denoising with large kernels*. In CVPR, 2020
Summary

• A burst image restoration and enhancement framework which is developed to effectively fuse complimentary information from multiple burst frames.

• We propose the idea of pseudo-burst sequence that is created by combining the channel-wise features from individual burst frames.

• We introduce an edge-boosting burst alignment module that is robust to camera-scene movements.

• The pseudo-burst features are enriched using multi-scale information and later progressively fused to create up-sampled outputs.

• State-of-the-art results on three burst image restoration and enhancement tasks (super-resolution, low-light enhancement, denoising) corroborate the generality and effectiveness of BIPNet.
Akshay Dudhane, Syed Waqas Zamir, Salman Khan, Fahad Khan, Ming-Hsuan Yang: 
TUE-PM-150.*
Proposed Approach: Burstormer

- A transformer-based design, named Burstormer, for burst-image restoration and enhancement that leverages multi-scale local and non-local features for improved alignment and feature fusion.

- It enables inter-frame communication in the burst neighborhoods for information aggregation and progressive fusion while modeling the burst-wide context.

- Its flexible design allows processing bursts of variable sizes.

- Burstormer is efficient and outperforms state-of-the-art methods on burst super-resolution, burst denoising and burst low-light enhancement.
Proposed Approach: Burstormer

**Synthetic BurstSR. Input size: 14x4x48x48**

- **Burststormer (42.83 dB, 3.58M)**
- **MFIR (41.56 dB, 12.13M)**
- **DBSR (40.76 dB, 13.01M)**
- **AFCNet (42.21 dB, 47M)**
- **BIPNet (41.93 dB, 6.67M)**
- **High ResNet (37.45 dB, 34.78M)**

**Real BurstSR. Input size: 14x4x80x80**

- **Burststormer (48.82 dB, 3.58M)**
- **MFIR (48.33 dB, 12.13M)**
- **DBSR (48.05 dB, 13.01M)**
- **AFCNet (48.63 dB, 47M)**
- **BIPNet (48.49 dB, 6.67M)**
- **High ResNet (46.64 dB, 34.78M)**

---


**DBSR:** Bhat G et al., “Deep burst super-resolution”. In CVPR 2021.

**AFCNet:** Mehta N. et al., “Adaptive feature consolidation network for burst super-resolution”. In CVPRW-22.

**BIPNet:** Dudhane A. et al., “Burst image restoration and enhancement”. In CVPR-22.

**Proposed Approach: Burststormer**

**Contributions:**

- **Enhanced Deformable alignment:** Aligns burst features implicitly with base frame, boosts the edges.
- **Image reconstruction:** Establishes strong inter-frame communication and integrates complementary information available within multiple frames. Neighboring frame features are used for spatial upscaling.

**Applications:**

- Proposed modules are validated for Burst Super-resolution, Burst Low-light Image Enhancement, Burst De-noising.
Proposed Approach: Burstormer

**Contributions:**

- **Enhanced Deformable alignment**: Aligns burst features implicitly with base frame, boosts the edges.
- **Image reconstruction**: Establishes strong inter-frame communication and integrates complementary information available within multiple frames. Neighboring frame features are used for spatial upscaling.

**Applications:**

- Proposed modules are validated for Burst **Super-resolution**, Burst **Low-light Image Enhancement**, Burst **De-noising**.
Proposed Approach: Burstormer

- **Feature Alignment**: Aligns burst features implicitly with base frame, boosts the edges.
- **Image Reconstruction**: Establishes strong inter-frame communication and integrates complementary information available within multiple frames. Neighboring frame features are used for spatial upscaling.

**Contributions:**

- Proposed modules are validated for Burst Super-resolution, Burst Low-light Image Enhancement, Burst De-noising.

**Applications:**
Proposed Approach

- Feature Alignment (FA).
- Reference based feature Enrichment (RBFE).

Enhanced Deformable Alignment (EDA)
Proposed Approach

Enhanced Deformable Alignment (EDA)

Feature Alignment (FA)

RAW LR Burst $B \times 4 \times H \times W$

Conv

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$

Feature Alignment

RBFE

$B \times f \times H \times W$
Proposed Approach

RAW LR Burst $B \times 4 \times H \times W$

Feature Alignment

RBFE

Feature Alignment

RBFE

Feature Alignment

RBFE

Reference-based feature enrichment (RBFE)

Base Frame $1 \times f \times H \times W$

Repeat, dim=0

Burst Features $B \times f \times H \times W$

Conv

Burst Feature Fusion (BFF)

Enhanced Deformable Alignment
Proposed Approach

Image reconstruction

Adaptive Burst Feature Pooling

$B \rightarrow 8$

$8 \times f \times H \times W$

NRFE

$8 \times f \times H \times W$

Reshape

$2 \times 4 \times f \times H \times W$

$\times 2$

Upsample

$2 \times f \times 2 \times H \times 2W$

NRFE

$2 \times f \times 2 \times H \times 2W$

Reshape

$1 \times 2 \times f \times 2 \times H \times 2W$

$\times 4$

Upsample

$1 \times f \times 4 \times H \times 4W$

Image
Proposed Approach

Image reconstruction

Adaptive Burst Feature Pooling

[Diagram showing the process with symbols and equations]

B → 8

Reshape 8xfH×W

NRFE 8xfH×W

Reshape 2×4xfH×W

×2 Upsample 2×f×2H×2W

NRFE 2×f×2H×2W

Reshape 1×2×2H×2W

×4 Upsample 1×f×4H×4W

Adaptive Burst pooling

Reshape H×W×Bf

1D avg pool H×W×8f

Permute 8xfH×W
Proposed Approach

Image reconstruction

Adaptive Burst Feature Pooling $3x3\times H\times W$

NRFE $8\times f\times H\times W$

Reshape $2\times 4\times H\times W$

Upsample $2\times f\times 2\times H\times 2W$

NRFE $2\times f\times 2\times H\times 2W$

Reshape $1\times 2\times f\times 2\times H\times 2W$

Upsample $1\times f\times 4\times H\times 4W$

Cyclic Burst Sampling

$\hat{B}\times f\times H\times W$

$\hat{B}\times f\times H\times W$

BFF

$\hat{B}\times f\times H\times W$
Proposed Approach

Image reconstruction

Adaptive Burst Feature Pooling

NRFE \(8 \times f \times H \times W\)

Reshape \(2 \times f \times H \times W\)

\(\times 2\) Upsample \(2 \times f \times H \times W\)

NRFE \(2 \times f \times 2H \times 2W\)

Reshape \(1 \times 2f \times 2H \times 2W\)

\(\times 4\) Upsample \(1 \times 4f \times 4W\)

Frame number

Burst Neighbourhoods

Cyclic Burst Sampling

\(\hat{B} \times \hat{f} \times \hat{H} \times \hat{W}\)

\(\hat{B} \times \hat{f} \times \hat{H} \times \hat{W}\)

\(\hat{B} \times \hat{f} \times \hat{H} \times \hat{W}\)

\(\hat{B} \times \hat{f} \times \hat{H} \times \hat{W}\)

\(\hat{B} \times \hat{f} \times \hat{H} \times \hat{W}\)

\(\hat{B} \times \hat{f} \times \hat{H} \times \hat{W}\)

\(\hat{B} \times \hat{f} \times \hat{H} \times \hat{W}\)

\(\hat{B} \times \hat{f} \times \hat{H} \times \hat{W}\)
Experiments

### Burst Super-resolution

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Synthetic Burst</th>
<th>BurstSR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>46k</td>
<td>5405</td>
</tr>
<tr>
<td>Validation</td>
<td>1204</td>
<td>882</td>
</tr>
</tbody>
</table>

### Burst De-noising

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Color set</th>
<th>Grayscale set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>20k</td>
<td>20k</td>
</tr>
<tr>
<td>Validation</td>
<td>100</td>
<td>73</td>
</tr>
</tbody>
</table>

### Burst Low-light Image Enhancement

<table>
<thead>
<tr>
<th>Datasets</th>
<th>SONY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>6.5k</td>
</tr>
<tr>
<td>Validation</td>
<td>93</td>
</tr>
</tbody>
</table>
Experiments

<table>
<thead>
<tr>
<th>Methods</th>
<th>PSNR ↑</th>
<th>SSIM ↑</th>
<th>LPIPS ↓</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen et al. [6]</td>
<td>29.38</td>
<td>0.892</td>
<td>0.484</td>
</tr>
<tr>
<td>Maharjan et al. [28]</td>
<td>29.57</td>
<td>0.891</td>
<td>0.484</td>
</tr>
<tr>
<td>Zamir et al. [43]</td>
<td>29.13</td>
<td>0.881</td>
<td>0.462</td>
</tr>
<tr>
<td>Zhao et al. [45]</td>
<td>29.49</td>
<td>0.895</td>
<td>0.455</td>
</tr>
<tr>
<td>LEED [18]</td>
<td>29.80</td>
<td>0.891</td>
<td>0.306</td>
</tr>
<tr>
<td>BIPNet [9]</td>
<td>32.87</td>
<td>0.936</td>
<td>0.305</td>
</tr>
<tr>
<td><strong>Ours</strong></td>
<td><strong>33.34</strong></td>
<td><strong>0.941</strong></td>
<td><strong>0.285</strong></td>
</tr>
</tbody>
</table>

[6]: Chen C et al., “Learning to see in the dark”. In CVPR 2018.
Experiments

Low-light Enhancement

Ground Truth  RAW Input Patch  LEED  BIPNet  Ours  Ground Truth Patch
Summary

- A transformer-based framework for burst image processing, capable of generating a single high-quality image from a given burst of noisy images having pixel misalignments among them.

- Burstormer employs a multi-scale hierarchical module EDA that, at each scale, first generates denoised features encoding local and non-local context, and then aligns each burst frame with the reference frame.

- To fix any remaining minor alignment issues, we incorporate a reference-based feature enrichment RBFE module in EDA that enables additional interaction of the features of each frame with the base frame features.
Summary

- A transformer-based framework for burst image processing, capable of generating a single high-quality image from a given burst of noisy images having pixel misalignments among them.

- Burstormer employs a *multi-scale hierarchical module EDA* that, at each scale, first generates denoised features encoding local and non-local context, and then aligns each burst frame with the reference frame.

- To fix any remaining minor alignment issues, we incorporate a reference-based feature enrichment RBFE module in EDA that enables additional interaction of the features of each frame with the base frame features.

- In the *image reconstruction stage*, we repeatedly apply the no-reference feature enrichment NRFE and upsampling modules in tandem until the final image is obtained. NRFE progressively and adaptively fuses each pair of frame features that are obtained with the proposed cyclic burst sampling.
Future Directions

- To develop a reference frame selector algorithm which will adaptively select the reference-frame for a given burst rather than assuming first frame as a reference frame.

- To come up with a self-supervised burst processing technique which overcome the pre-training on synthetic data and domain adaptation.

- To validate the proposed network modules for other multi-frame applications such as, satellite image processing, medical image analysis etc., where feature alignment, fusion and reconstruction plays a crucial role in systems performance.
Acknowledgements

- Akshay Dhudhane
- Syed Waqas Zamir
- Salman Khan
- Nancy Mehta
- Subrahmanyam Murala
- Ming-Hsuan Yang
Visit Us During Poster Sessions


Thanks