S2A: Wasserstein GAN with Spatio-Spectral Laplacian Attention for Multi-Spectral Band Synthesis

CVPR-EarthVision 2020

Litu Rout      Indranil Misra      S Manthira Moorthi      Debajyoti Dhar

Space Applications Centre
Indian Space Research Organisation
Super-resolution as conditional band synthesis

- Direct super-resolution is intractable.
- Lack necessary geometric attributes.
- Reformulate as conditional band synthesis.
- Geometry from existing high resolution bands: HR-NIR, R, G.
- Radiometry from corresponding low resolution band: LR-SWIR.
- FCC: NIR (R), R (G), G(B)
Super-resolution as conditional band synthesis

- Direct super-resolution is intractable.
- Lack necessary geometric attributes.

- Reformulate as conditional band synthesis.
- Geometry from existing high resolution bands: HR-NIR, R, G.
- Radiometry from corresponding low resolution band: LR-SWIR.

FCC: NIR (R), R (G), G(B)
Traditional Approach

LR-SWIR
Traditional Approach
Traditional Approach

LR-Upsampled-SWIR

LR-SWIR

HR-NIR,R,G
Traditional Approach
Traditional Approach
Over dependence on upsampled coarse resolution band results in unpleasant artifacts.

- Geometric distortion
- Radiometric imbalance
Over dependency on upsampled coarse resolution band results in unpleasant artifacts.

- Geometric distortion
- Radiometric imbalance
Traditional Approach

Over dependency on upsampled coarse resolution band results in unpleasant artifacts.

- Geometric distortion
- Radiometric imbalance
Over dependency on upsampled coarse resolution band can be suppressed by replacing it with spatial attention map.
Traditional Approach

Proposed Approach

FCC: SWIR (R), NIR (G), Red (B)
Spatial Attention from Discriminator
Spatial Attention from Discriminator
Spatial Attention from Discriminator

\[ A_s(x) = \mathcal{N}(D_s(x)), \]

\[ D_s(x) = \sum_{i=1}^{K} \mathcal{N} \left( \sum_{j=1}^{C} |A_{ij}(x)| \right) \]
Spatial Attention from Discriminator

Spatial Attention Loss

\[ \mathcal{L}_{sa} = \mathbb{E}_{\hat{x} \sim P_x, y \sim P_y} \left[ \| A_s(\hat{x}) - A_s(y) \|_2^2 \right] \]

Domain Adaptation Loss

\[ \mathcal{L}_{da} = \mathbb{E}_{\tilde{y} \sim P_y, y \sim P_y} \left[ \| A_s(\tilde{y}) - A_s(y) \|_2^2 \right] \]
Spatial Attention from Discriminator

Spatial Attention Loss
\[ \mathcal{L}_{sa} = \mathbb{E}_{\hat{x} \sim P_x, y \sim P_y} \left[ \| A_s(\hat{x}) - A_s(y) \|_2^2 \right] \]

Domain Adaptation Loss
\[ \mathcal{L}_{da} = \mathbb{E}_{\tilde{y} \sim P_{\tilde{y}}, y \sim P_y} \left[ \| A_s(\tilde{y}) - A_s(y) \|_2^2 \right] \]

Discriminator Objective
\[
\min_D \mathbb{E}_{\hat{x} \sim P_x} [D(\hat{x})] - \mathbb{E}_{x \sim P_x} [D(x)] \\
+ \lambda_{gp} \mathbb{E}_{\hat{x} \sim P_x} \left( \| \nabla_x D(\hat{x}) \|_2 - 1 \right)^2 \\
+ \lambda_{sa} \mathcal{L}_{sa} + \lambda_{da} \mathcal{L}_{da}
\]
Spatial Attention from Discriminator

Spatial Attention Loss
\[ L_{sa} = \mathbb{E}_{\hat{x} \sim P_{\hat{x}}, y \sim P_{y}} \left[ \| A_s(\hat{x}) - A_s(y) \|_2^2 \right] \]

Domain Adaptation Loss
\[ L_{da} = \mathbb{E}_{\tilde{y} \sim P_{\tilde{y}}, y \sim P_y} \left[ \| A_s(\tilde{y}) - A_s(y) \|_2^2 \right] \]

Discriminator Objective
\[ \min_D \mathbb{E}_{\hat{x} \sim P_{\hat{x}}} \left[ D(\hat{x}) \right] - \mathbb{E}_{x \sim P_x} \left[ D(x) \right] \]
\[ + \lambda_{gp} \mathbb{E}_{\hat{x} \sim P_{\hat{x}}} \left[ (\| \nabla_{\hat{x}} D(\hat{x}) \|_2 - 1)^2 \right] \]
\[ + \lambda_{sa} L_{sa} + \lambda_{da} L_{da} \]
Spatial Attention from Discriminator

Generator Objective

\[
\min_G - \mathbb{E}_{z \sim P_S, \tilde{y} \sim P_{\tilde{y}}} \left[ D \left( G(z, A_S(\tilde{y})) \right) \right] \\
+ \lambda_p \mathbb{E}_{z \sim P_S, \tilde{y} \sim P_{\tilde{y}}, y \sim P_y} \left[ \|G(z, A_S(\tilde{y})) - y\|^2 \right] \\
= G(z, A_S(x))
\]
Spatio-Spectral Laplacian Attention

\[ \hat{x} = G(z, A_s(x)) \]
Spatio-Spectral Laplacian Attention

Spectral attention coefficients

\( \hat{x} = G(z, A_s(x)) \)
Combining Spatial Attention with Source Bands

- Attention module latches on to bright targets.
- Synthesized band contains blocky artifacts.
Table: Image Quality Metrics for Different Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>RMSE</th>
<th>SSIM(%)</th>
<th>SRE(dB)</th>
<th>PSNR(dB)</th>
<th>SAM(deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AeroGAN [31]</td>
<td>21.62</td>
<td>86.03</td>
<td>44.62</td>
<td>36.50</td>
<td>12.15</td>
</tr>
<tr>
<td>DSen2 [21]</td>
<td>14.14</td>
<td>93.85</td>
<td>50.04</td>
<td>41.94</td>
<td>7.88</td>
</tr>
<tr>
<td>DeepSWIR [33]</td>
<td>13.75</td>
<td>94.02</td>
<td>50.35</td>
<td>42.27</td>
<td>7.66</td>
</tr>
<tr>
<td>ALERT [32]</td>
<td>12.97</td>
<td>94.54</td>
<td>50.81</td>
<td>42.80</td>
<td>7.48</td>
</tr>
<tr>
<td>S2A (ours)</td>
<td>11.74</td>
<td>95.08</td>
<td>50.83</td>
<td>42.76</td>
<td>6.87</td>
</tr>
</tbody>
</table>
- Learns to attend to relevant parts of source imagery.
- Homogeneous and heterogeneous targets are discernible.
- Similar features have similar attention coefficients
Wetland Delineation

Water Segmentation

(a) NIR(R),R(G),G(B)  (b) GT-MNDWI (IoU)  (c) S2A (99.117)
Wetland Delineation

Water Segmentation

(a) NIR(R), R(G), G(B)  
(b) GT-MNDWI (IoU)  
(c) S2A (99.117)
Additional Value Product Generation

India

Hilly Terrain

Main land

Desert

Coastal
Summary

- Formulated super resolution as conditional band synthesis

- Regulated band synthesis through spatial and Laplacian spectral channel attention

- Introduced two new cost functions for the discriminator:
  - Spatial attention loss
  - Domain adaptation loss

- Experimented on multiple datasets:
  - LISS-3
  - LISS-4
  - WorldView-2

- Demonstrated real world applications of synthesized band:
  - Wetland delineation
  - Index based water segmentation
  - Additional value product generation/ Large area mosaic