Our Graph Convolution

- Wide field of view
- High visual acuity
- Fewer pixels than uniform images
- Appealing for efficient computer vision systems
- Difficult to process with convolutional layers
- In practice, have shown minimal improvement over uniform sampling [1]

Map pixel offsets to filter weights through a learnable function $G(x, y)$ to generalize convolution to non-grid aligned data

No Filter deformation
No Image Discontinuities

Previous Approaches

- Filter deformation
No Image Discontinuities

Sequential Attention Architecture

- No Filter deformation
Image Discontinuities

Visualization of Filter Deformations

Classification Accuracy on Imagenet-100

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Operator</th>
<th># Fixations</th>
<th># Input Pixels</th>
<th>GFLOPs</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>Conv</td>
<td>-</td>
<td>112x112</td>
<td>0.20</td>
<td>70.0</td>
</tr>
<tr>
<td>Foveated (ours)</td>
<td>Graph Conv</td>
<td>-</td>
<td>112$^2$</td>
<td>0.20</td>
<td>72.5</td>
</tr>
<tr>
<td>FCG [5]</td>
<td>Conv</td>
<td>2</td>
<td>112x112</td>
<td>0.41</td>
<td>70.2</td>
</tr>
<tr>
<td>Log-Polar [4]</td>
<td>Conv</td>
<td>2</td>
<td>80x160</td>
<td>0.41</td>
<td>70.4</td>
</tr>
<tr>
<td>Multi-FoV Crops [3]</td>
<td>Conv</td>
<td>2</td>
<td>2x80x80</td>
<td>0.41</td>
<td>72.8</td>
</tr>
<tr>
<td>Foveated (ours)</td>
<td>Graph Conv</td>
<td>2</td>
<td>112$^2$</td>
<td>0.41</td>
<td>73.8</td>
</tr>
<tr>
<td>Foveated (ours)</td>
<td>Graph Conv</td>
<td>3</td>
<td>112$^2$</td>
<td>0.61</td>
<td>76.5</td>
</tr>
</tbody>
</table>

- Comparisons of different methods on Imagenet-100, a natural image dataset comprised of 130,000 examples across 100 classes
- Methods with “...” fixations indicate no attention is used.

Conclusion

- We presented a novel graph convolutional approach to processing foveated images and incorporated it into a sequential attention architecture.
- We showed that our approach outperforms previous foveated CNN architectures by at least 1% and a uniform CNN by 2.5%.

Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), grant number 2443519.

References


2182951k@student.gla.ac.uk, {Paul.Henderson, Paul.Siebert, Gerardo.AragonCamarasa}@glasgow.ac.uk