Foveation in the Era of Deep Learning

George Killick, Paul Henderson, Paul Siebert, Gerardo Aragon-Camarasa
School of Computing Science, University of Glasgow

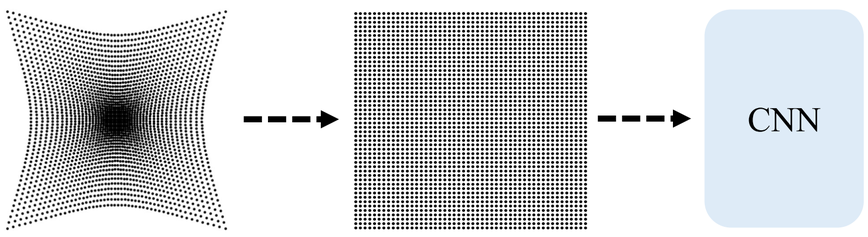
BMVC 2023

Foveated Images

- 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] ☒

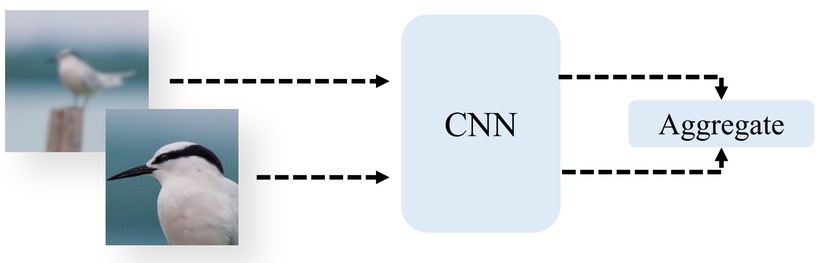
Previous Approaches

[2]



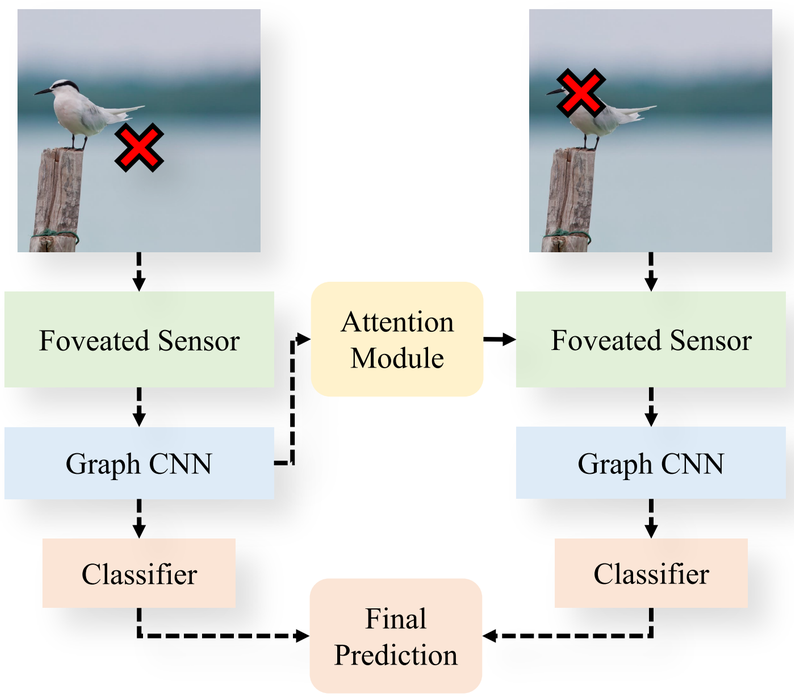
☒ Filter deformation ☑ No Image Discontinuities

[3]



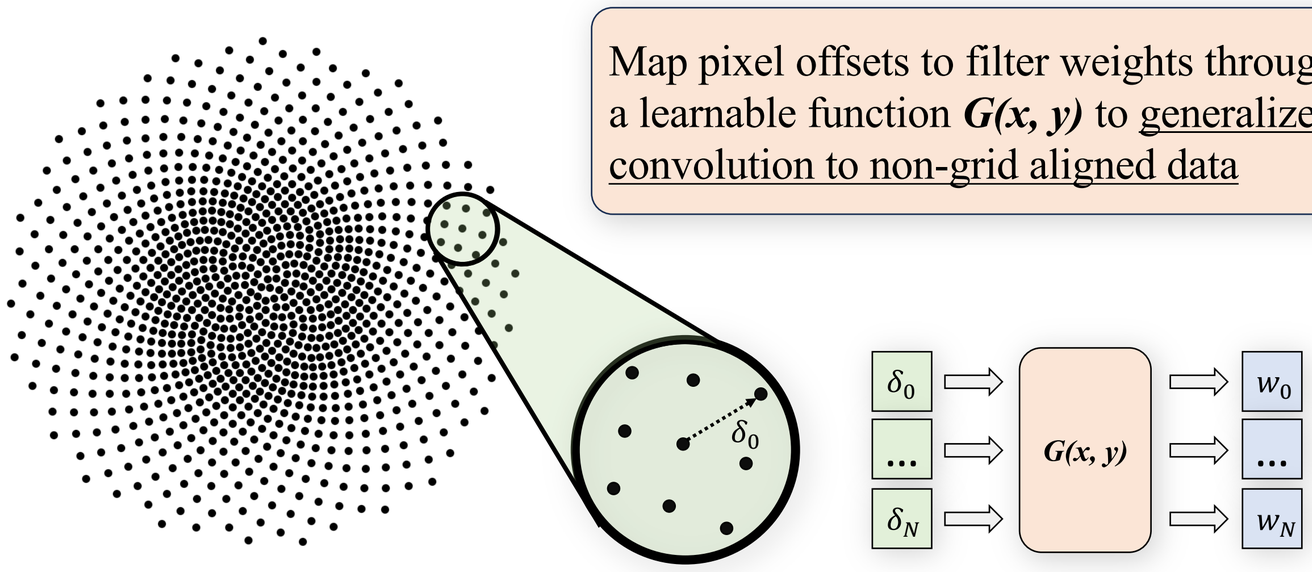
☑ No Filter deformation ☒ Image Discontinuities

Sequential Attention Architecture



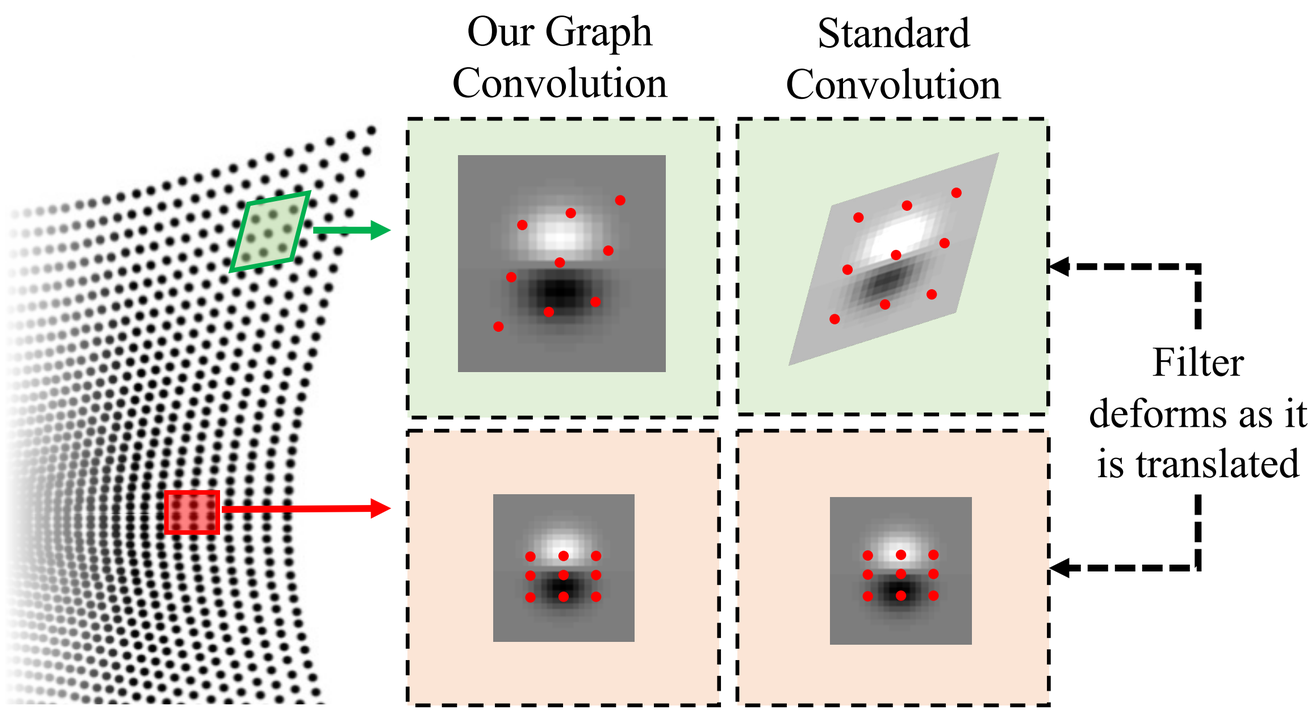
Our Approach (Graph Convolution)

Map pixel offsets to filter weights through a learnable function ***G(x, y)*** to <u>generalize convolution to non-grid aligned data</u>



☑ No Filter deformation ☑ No Image Discontinuities

Visualization of Filter Deformations



[2]

Classification Accuracy on Imagenet-100

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

- 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%.

References

1. Torabian, Parsa, et al. "Comparison of foveated downsampling techniques in image recognition." *Journal of Computational Vision and Imaging Systems* 6.1 (2020): 1-3.
2. Wang, Binxu, et al. "On the use of cortical magnification and saccades as biological proxies for data augmentation." *arXiv preprint arXiv:2112.07173* (2021).
3. Sermanet, Pierre, Andrea Frome, and Esteban Real. "Attention for fine-grained categorization." *arXiv preprint arXiv:1412.7054* (2014).
4. Esteves, Carlos, et al. "Polar transformer networks." *arXiv preprint arXiv:1709.01889* (2017).
5. Lukanov, Hristofor, Peter König, and Gordon Pipa. "Biologically inspired deep learning model for efficient foveal-peripheral vision." *Frontiers in Computational Neuroscience* 15 (2021): 746204.

Acknowledgements

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

2182951k@student.gla.ac.uk, {Paul.Henderson, Paul.Siebert, Gerardo.AragonCamarasa}@glasgow.ac.uk
Linkedin.com/in/george-killick-47315a158/