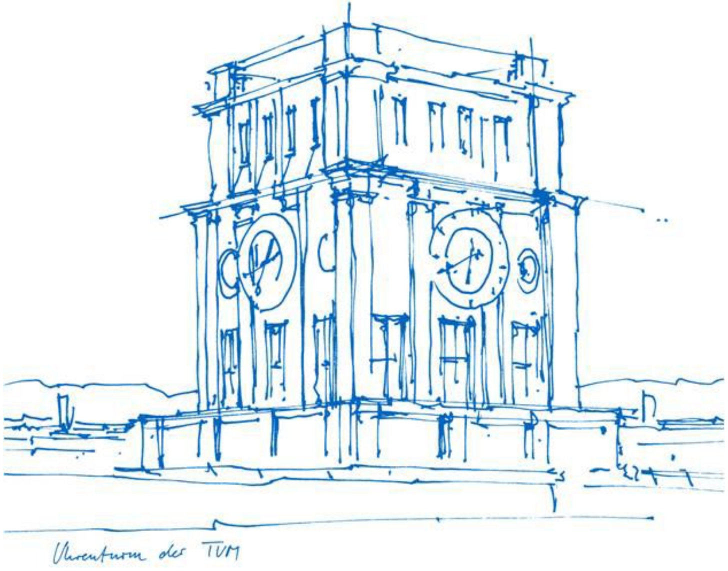
One-dimensional Path Convolution

<u>Xuanshu Luo</u>, Martin Werner
Technical University of Munich
TUM School of Engineering and Design
Professorship of Big Geospatial Data Management

ICML 2025



Motivation

Compared to 2D convolution, 1D convolutional kernels

- fail to preserve the spatial continuity of adjacent pixels in both directions.
+ scale linearly, providing better parameter efficiency;

Motivation

Compared to 2D convolution, 1D convolutional kernels

- fail to preserve the spatial continuity of adjacent pixels in both directions.
+ scale linearly, providing better parameter efficiency;

How to construct a vision model that exclusively utilizes 1D convolution to achieve superior parameter efficiency while simultaneously preserving the locality of images to match the capabilities of conventional 2D CNNs?

Motivation

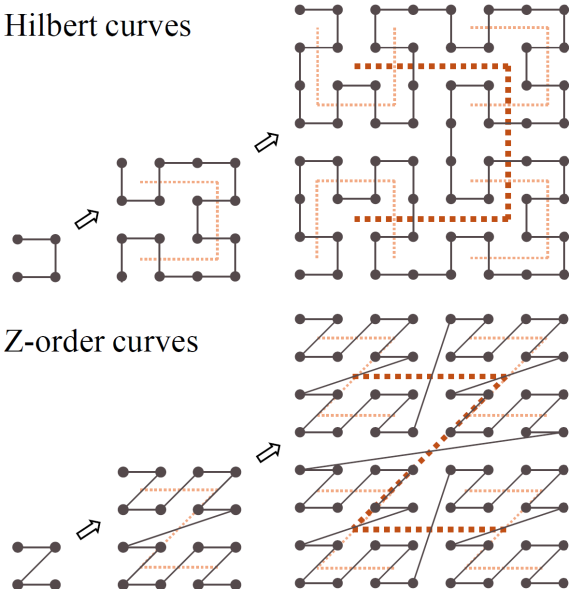
Compared to 2D convolution, 1D convolutional kernels

- fail to preserve the spatial continuity of adjacent pixels in both directions.
+ scale linearly, providing better parameter efficiency;

How to flatten a 2D image?

How to construct a vision model that exclusively utilizes 1D convolution to achieve superior parameter efficiency while simultaneously preserving the locality of images to match the capabilities of conventional 2D CNNs?

Hilbert and Z-order paths



- space-filling
- topologically self-organizing
- recursively defined

provide scale-invariant 1D-to-2D mappings with spatial proximity preservation capability[1,2]

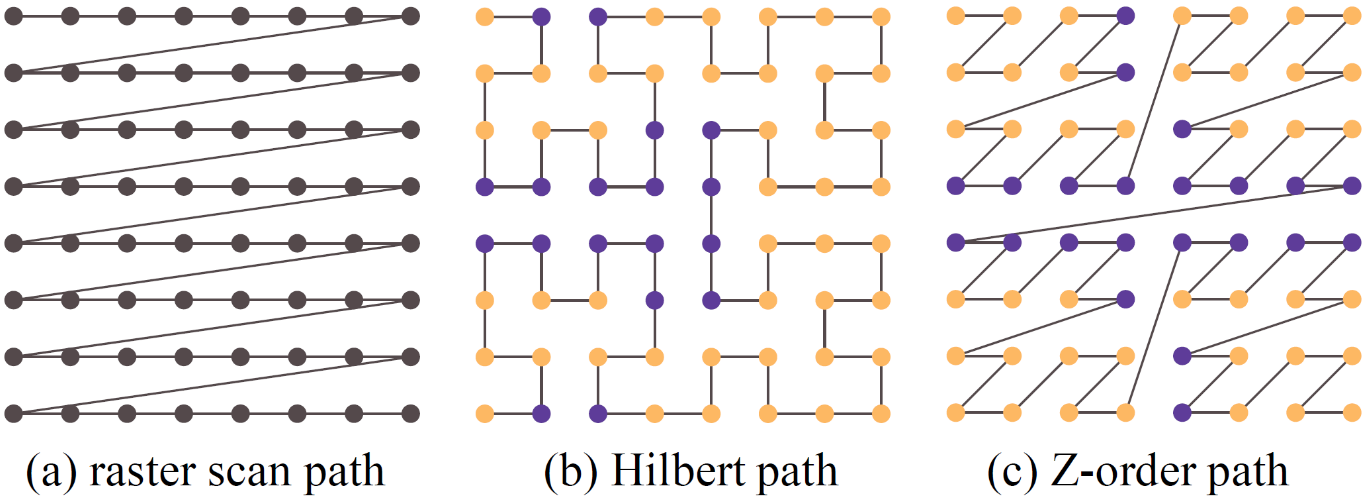
image compression[3]
database[4]
parallel computing[5]
point cloud processing[6]
…

[1] Jagadish, H. V. Linear clustering of objects with multiple attributes. In Proceedings of the 1990 ACM SIGMOD International Conference on Management of Data
[2] Dai, H. and Su, H.-C. On the locality properties of space-filling curves. In International Symposium on Algorithms and Computation. Springer, 2003.
[3] Wang, H., Gupta, K., Davis, L., and Shrivastava, A. Neural space-filling curves. In European Conference on Computer Vision. Springer, 2022
[4] Kamel, I. and Faloutsos, C. Hilbert r-tree: An improved rtree using fractals. In VLDB, volume 94, pp. 500–509. Citeseer, 1994
[5] Böhm, C., Perdacher, M., and Plant, C. A novel Hilbert curve for cache-locality preserving loops. IEEE Transactions on Big Data, 7(2):241–254, 2018.
[6] Wu, X., et al. Point transformer v3: Simpler faster stronger. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024

Hilbert and Z-order paths

provide scale-invariant **1D-to-2D** mappings with spatial proximity preservation capability[1,2],

but we are considering **2D-to-1D** mappings…



Pixels with shorter cumulative distances to all their 8 neighbors than (a) are in apricot, otherwise purple.

[1] Jagadish, H. V. Linear clustering of objects with multiple attributes. In Proceedings of the 1990 ACM SIGMOD International Conference on Management of Data

[2] Dai, H. and Su, H.-C. On the locality properties of space-filling curves. In International Symposium on Algorithms and Computation. Springer, 2003.

Hilbert and Z-order paths

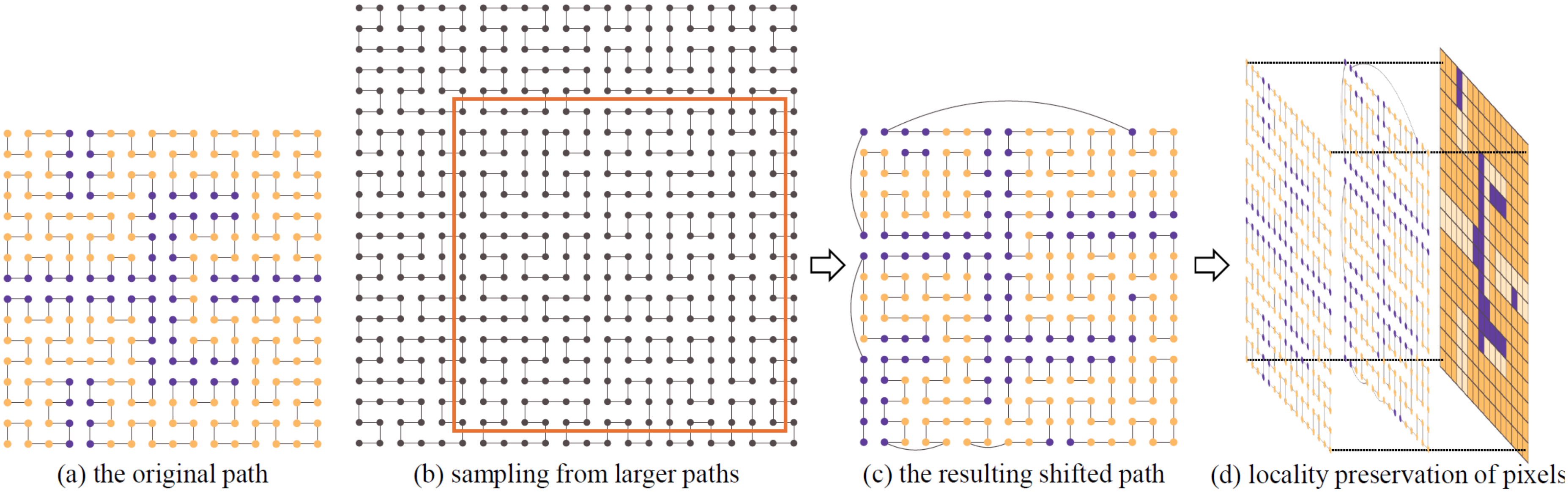
Table 1. Locality measurements of the raster scan, Hilbert, and Z-order paths for multiple resolutions.

Path	Resolution	Total Distance	P_{sd}
raster scan	32×32	1.88×10^5	-
Hilbert		2.15×10^5	79.10%
Z-order		1.80×10^5	74.61%
raster scan	64×64	1.54×10^6	-
Hilbert		1.80×10^6	80.96%
Z-order		1.49×10^6	81.35%
raster scan	128×128	1.24×10^7	-
Hilbert		1.48×10^7	86.51%
Z-order		1.22×10^7	84.50%
raster scan	256×256	1.00×10^8	-
Hilbert		1.20×10^8	89.04%
Z-order		0.99×10^8	89.17%

P_{sd} = proportion of pixels with shorter distances to their neighbors at the same positions than raster scan paths.

Directly applying Hilbert/Z-order paths yields suboptimal spatial locality preservation.

Path Shifting



The path shifting approach effectively relocates purple pixels.
Combining multiple shifted paths can meet the locality constraint (P_{sd} =100%).

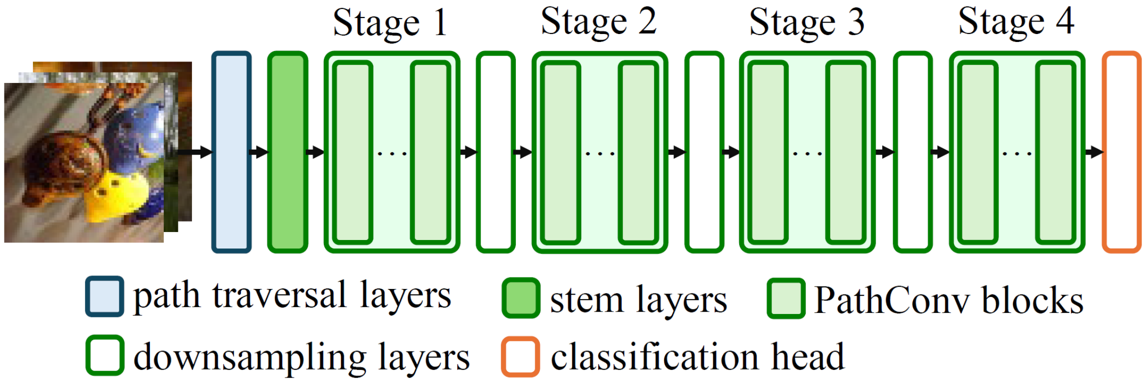
The minimal set of paths

To find the minimal set of paths satisfying the locality constraint is polynomial-time reducible to the NP set cover problem[1] (Appendix C provides proof).

Using *randomized rounding algorithm*[2], we find that **three** shifted paths are sufficient to meet the locality constraint for multiple resolutions.

[1] Vazirani, V. V. Set Cover, pp. 15–26. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.
[2] Bertsimas, D. and Vohra, R. Rounding algorithms for covering problems. Mathematical Programming, 1998.

Path Convolution Model



- The CUDA-optimized path traversal layer provides up to 73-fold acceleration compared to a single-thread CPU implementation.
- We introduce **path-aware channel attention** (PACA) to capture both path-specific and cross-path dependencies.

Experiments

Datasets:

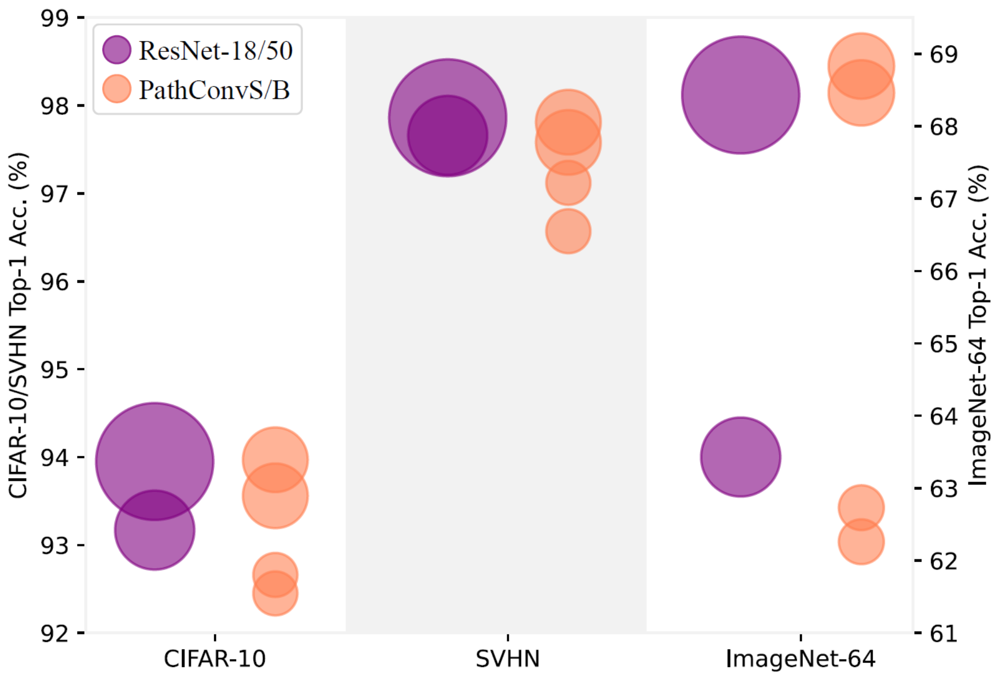
CIFAR-10[1], SVHN[2], ImageNet-64[3]

Models:

`PathConvS/B`

`ResNet18/50`[4] (comparable FLOPs)

PathConv models achieve ResNet-level accuracy using only 1/3 parameters.



[1] Krizhevsky, A., Hinton, G., et al. Learning multiple layers of features from tiny images. 2009.
[2] Netzer, Y., et al. Reading digits in natural images with unsupervised feature learning. In NIPS workshop on deep learning and unsupervised feature learning, Granada, 2011.
[3] Chrabaszcz, P., Loshchilov, I., and Hutter, F. A downsampled variant of imagenet as an alternative to the cifar datasets.
[4] He, K., et al. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

Thank you!

For more details, please refer to our paper.

Code at GitHub

More research in our group