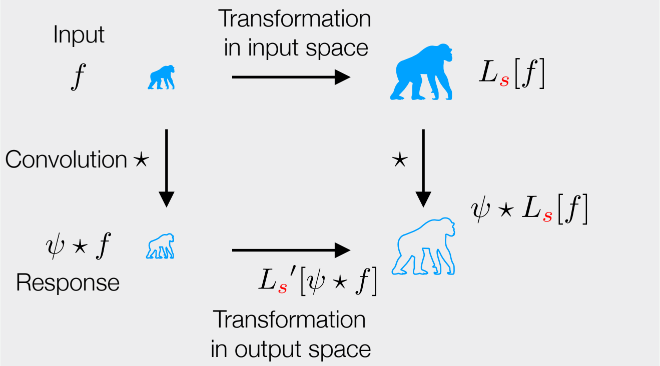
Deep Scale-spaces: Equivariance Over Scale

Daniel Worrall, Max Welling {d.e.worrall, m.welling}@uva.nl. https://deworrall92.github.io/

The Big Idea

Design a CNN, with built-in scale symmetry

What is symmetry?



Correlations: translation symmetry

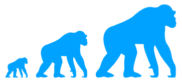
$$[\psi \star_{\mathbb{Z}} f](s) = \sum_{x \in \mathbb{Z}} \psi(x) f(x+s)$$

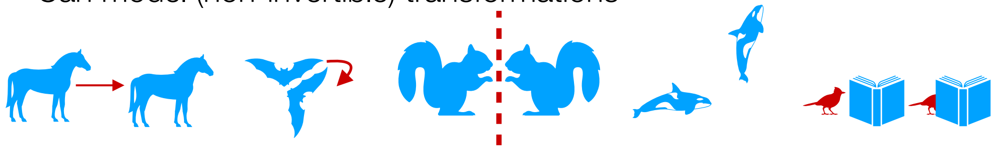
Response over shifts — Filter — Shifted image

Semigroups

Algebra: Closed set S with (associative) multiplication

- Can model (non-invertible) transformations





Semigroup correlations

$$[\psi \star_S f](s) = \sum_{x \in X} \psi(x) L_s[f](x)$$

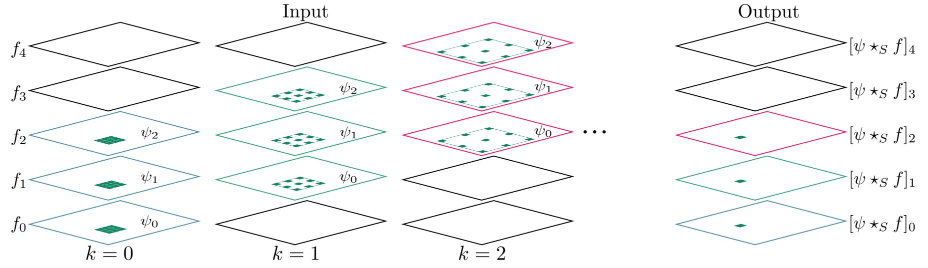
Response — Filter — Transformed image

Scale-space correlations

$$[\psi \star_S f](z, k) = \sum_{\ell \in \mathbb{Z}_{\geq 0}} \sum_{y \in \mathbb{Z}^d} \psi(y, \ell) f(2^k y + z, \ell + k)$$

Response extends over space + scale — Sum over scale + space — Dilate + shift — Filter extends over space + scale — Shift through scale

Implementation & Experiments



Architecture matched experiments

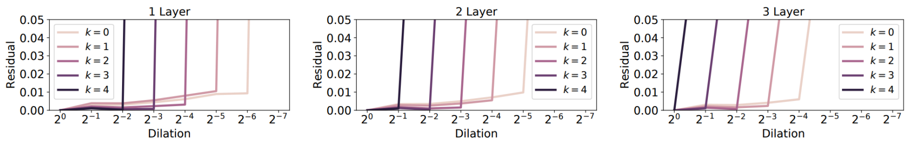
- Beat baselines, but yes we have handicapped them.

Table 1: Results on the Patch Camelyon and Cityscapes Dataset. Higher is better

PCam Model	Accuracy
DenseNet Baseline	87.0
S-DenseNet (Ours)	88.1
[Veeling et al., 2018]	89.8

Cityscapes Model	mAP
ResNet, matched parameters	45.66
ResNet, matched channels	49.99
S-ResNet, multiscale (Ours)	63.53
S-ResNet, no interaction (Ours)	64.78

Equivariance error (normalized L2): equivariance holds in practice



Scale-spaces

$$f(\mathbf{x}, t) = [G(\cdot, t) * f_0](\mathbf{x}), \quad t > 0$$

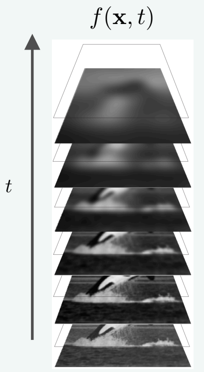
$$f(\mathbf{x}, 0) = f_0(\mathbf{x})$$ — Original image

Gauss-Weierstrass kernel

$$G(\mathbf{x}, t) = \frac{1}{(4\pi t)^{d/2}} \exp\left\{-\frac{\|\mathbf{x}\|^2}{4t}\right\}$$

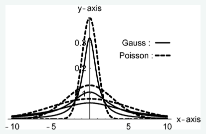
The "semigroup property"

$$G(\cdot, s) * G(\cdot, t) * f_0 = G(\cdot, s+t) * f_0$$



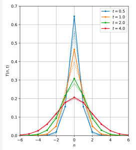
Gaussian scale-space

Continuous scale, continuous grid



Alpha SS

Continuous scale, discrete grid

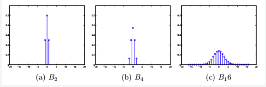


Discrete GSS

Discrete scale, continuous grid

NA

Discrete scale, discrete grid



Binomial SS

Limitations

Integer scale

Computationally inefficient (implementation, large kernels, extra scale dimension

Truncation of scale-space leads to boundary effects

Conclusions

Proposed new class of correlations for use in CNNs

Developed a correlation for scale-spaces

Demostrated scale-equivariance is achievable in a deep learning setting

Acknowledgements

We thank Koninklijke Philips N.V. for in-cash and in-kind support. We also give thanks to Rianne van den Berg, Patrick Forré, and our anonymous reviewers for the time and effort they put In to improving this paper