# How to Transform Kernels for Scale-Convolutions

Ivan Sosnovik, Artem Moskalev, Arnold Smeulders
UvA-Bosch Delta Lab, University of Amsterdam

## Scale-Convolution

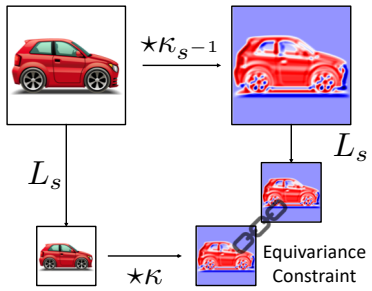
The mapping defined as:

$$[f \star_H \kappa](s,t) = \sum_{s'} [f(s',\cdot) \star \kappa_s(s^{-1}s',\cdot)](\cdot,t)$$

is scale-equivariant if and only if the following is true with respect to the kernels.

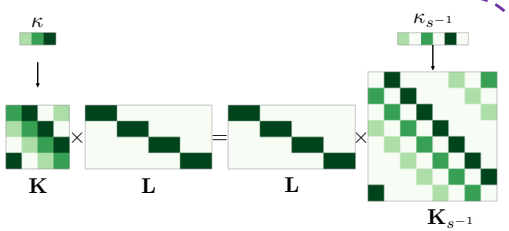
$$L_s[f] \star \kappa = L_s[f \star \kappa_{s^{-1}}], \quad \forall f, s$$

When it is not satisfied an equivariance error appears




## Equivariance Constraint

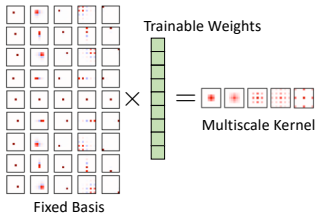
The equivariance constraint can be rewritten in the matrix form as follows:

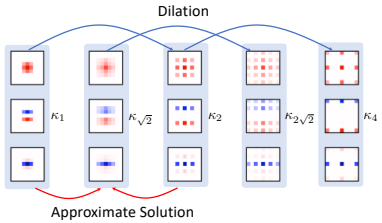



$$\boldsymbol{KLf} = \boldsymbol{LK}_{s^{-1}}\boldsymbol{f}, \ \forall \boldsymbol{f} \quad \iff \quad \boldsymbol{KL} = \boldsymbol{LK}_{s^{-1}}$$

It has a solution only if the downscaling is performed by an **integer scale factor**. In this cases, the solution is a dilated filter. In all other cases, only an **approximated solution** can be computed as follows, for example:

$$\kappa_{\sqrt{2}} = \arg\min \mathbb{E}_f \|L[f] \star \kappa_1 - L[f \star \kappa_{\sqrt{2}}]\|_F^2 + \|L[f] \star \kappa_{\sqrt{2}} - L[f \star \kappa_2]\|_F^2$$

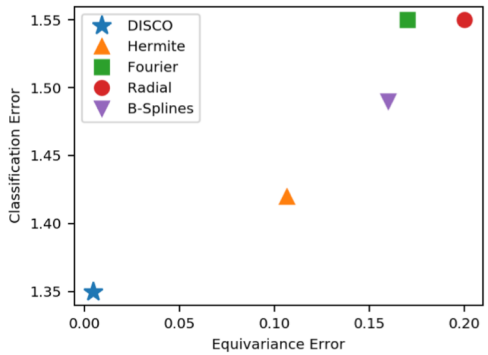



DISCO Kernels are computed via multiplying a fixed multi-scale basis with trainable weights. $\sqrt{2}$ is the smallest non-integer scale, for which the kernel is approximated by minimizing the equivariance constraint, the rest of the kernels can be obtained with dilation.

## Equivariance Error vs Classification Error

Lower equivariance error leads to a lower classification error on the MNIST-scale dataset.




## Large-Scale

| Model | Basis | STL-10 | Time, s |
|---|---|---|---|
| WRN | - | 11.48 | 10 |
| SiCNN | - | 11.62 | 110 |
| SI-ConvNet | - | 12.48 | 55 |
| DSS | Dilation | 11.28 | 40 |
| SS-CNN | Radial | 25.47 | 15 |
| SESN | Hermite | 8.51 | 165 |
| DISCO | Discrete | **8.07** | 50 |

| Model | AUC | FPS |
|---|---|---|
| SiamFC | 0.61 | - |
| TriSiam | 0.62 | - |
| SiamFC+ | 0.67 | 56 |
| SE-SiamFC+ | **0.68** | 14 |
| DISCO-SiamFC+ | **0.68** | 28 |

The proposed DISCO filters allow for more accurate classification on STL-10 (left) and faster inference when it is extremely important, such as on OTB-13 (Right).

[SESN] Sosnovik I., Szmaja M., Smeulders A., Scale-equivariant steerable networks. ICLR 2020
[SiamFC+] Sosnovik I., Moskalev A., Smeulders A. W. M., Scale equivariance improves siamese tracking. WACV 2021
[DISCO] Sosnovik I., Moskalev A., Smeulders A. How to Transform Kernels for Scale-Convolutions ICCV 2021, VIPriors