Scale-Equivariant U-Net

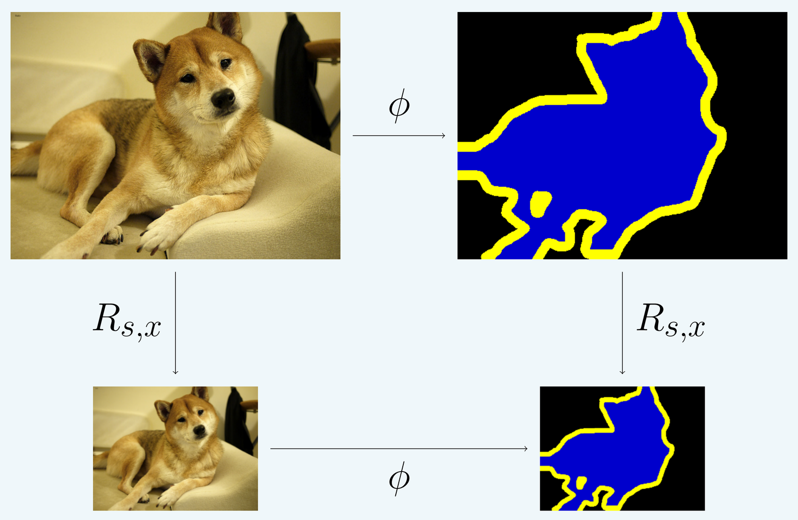
Mateus Sangalli Samy Blusseau Santiago Velasco-Forero Jesús Angulo

Centre for Mathemtical Morphology, Mines Paris, PSL University

Abstract

The U-Net [2] is not scale-equivariant in general. The scale-cross-correlation [4] is a linear scale-equivariant operator that generalizes the convolution for feature maps defined on a semigroup of scales and translations. In this work we propose the Scale-Equivariant U-Net (SEU-Net) which introduces scale-equivariance to the U-Net by using scale-cross-correlations instead of the convolution and by studying the equivariance of the pooling and upsampling operators that are used in the U-Net.

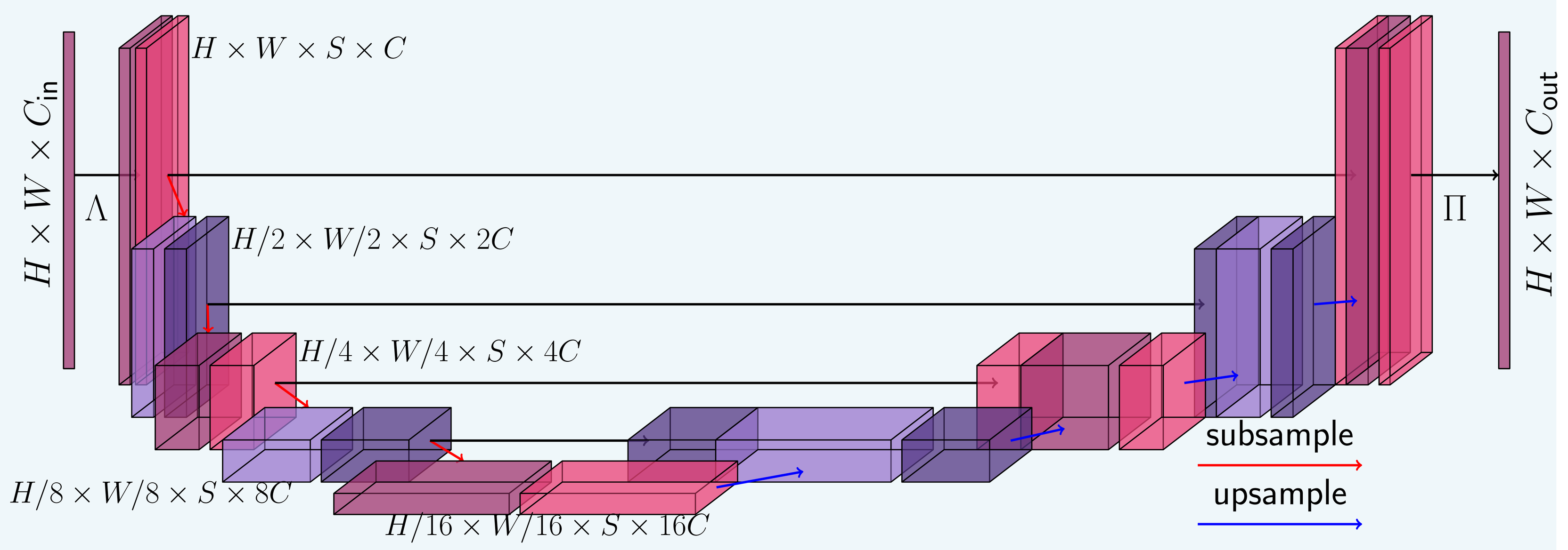
Scale-Equivariance



Let $S_\gamma = \{\gamma^0, \gamma^1, \gamma^2, \ldots\}$ with real multiplication be a semigroup of scales and $S_\gamma \times \mathbb{Z}^2$ a semigroup of scales and translations. Given $f : \mathbb{Z}^2 \to \mathbb{R}^n$ and $\bar{f} : S_\gamma \times \mathbb{Z}^2 \to \mathbb{R}^n$ we define the actions

$$R'_{s,x}f(y) = f(sy + x) \quad R_{s,x}f(y) = f(st, sy + x) \tag{1}$$

SEU-Net



- Images are lifted by a scale-space operator Λ to a semigroup of scales and translations
- In that space we perform scale-cross-correlations, pooling, upsampling and other operations used in the U-Net
- The feature maps are projected back to images using the projection Π

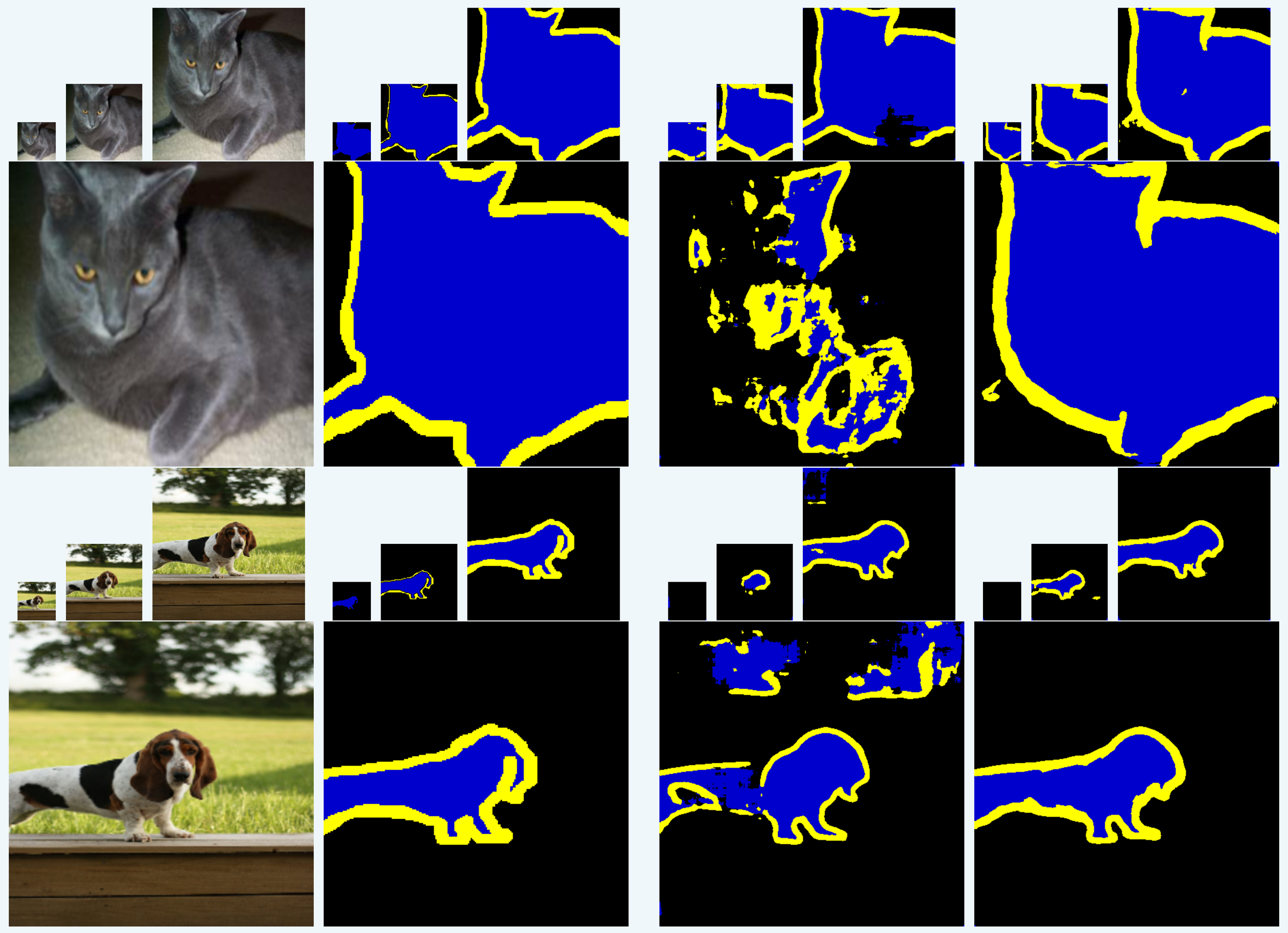
Equivariance of Subsampling and Upsampling

- Subsampling and Upsampling must be scale-equivariant in order for the network to be scale-equivariant
- Subsampling by strided scale-cross-correlations is scale-equivariant
- We denote by U_t an upsampling by a factor of t, i.e. an operator such that $U_t(f)(tx) = f(x)$ and $U_t \circ U_s = U_{t+s}$.
- Upsamplings are not scale-equivariant under some conditions it behaves as an equivariant operator

Partial Equivariance of upsampling

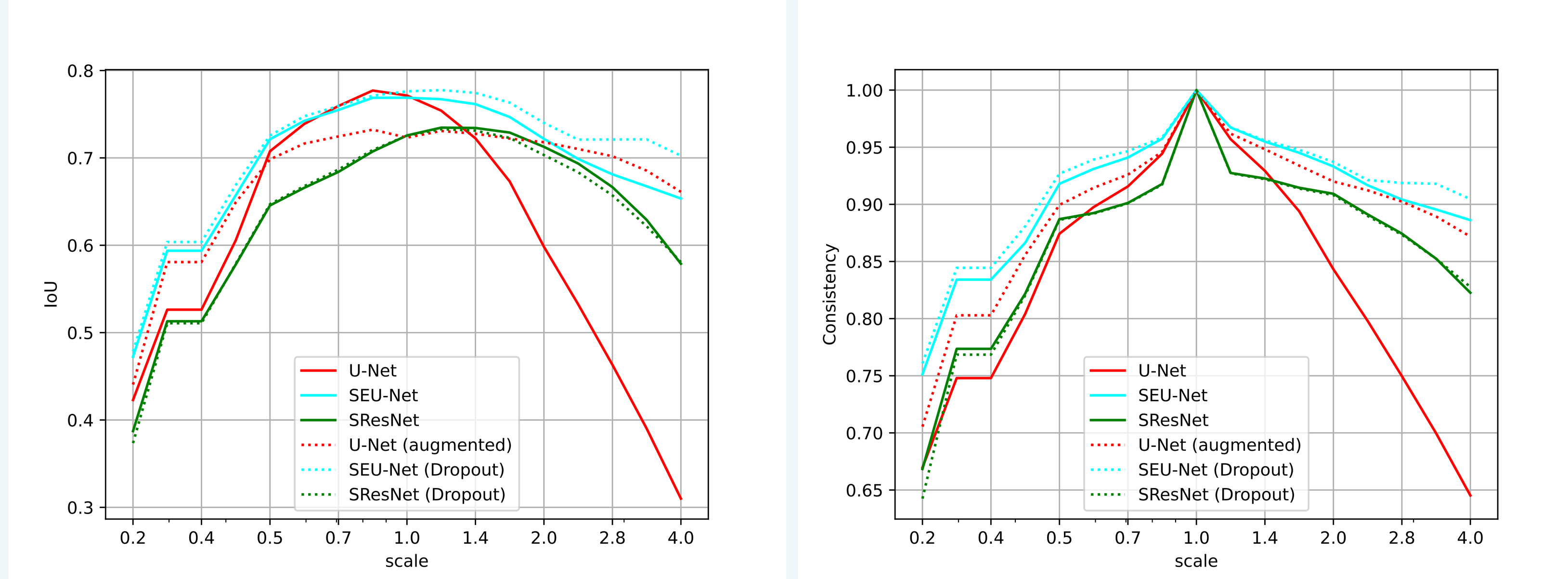
For $N \in \mathbb{N}^*$ and $i \in \{1, \ldots, N\}$ let $\mathcal{U}_i = \{U_{\gamma^{n_i+l}} f_i | l \in \mathbb{N}\}$, where each $f_i : G \to \mathbb{R}^n$ is a function on G and each n_i an integer. Let $n_0 \le \min\{n_i | i = 1, \ldots, N\}$ and $\mathcal{U} = \bigcup_{i=1}^{N} \mathcal{U}_i$. Then for all $f \in \mathcal{U}$, and $k, l \in \mathbb{N}$ such that $k - l \le n_0$, we have $U_{\gamma^l} R_{\gamma^k,x} f = R_{\gamma^k,\gamma^l x} U_{\gamma^l} f$.

Oxford-IIIT Pet [1]



(a) Image (b) Ground Truth (c) U-Net (d) SEU-Net

Figure 1. Sample test image at different scales and ground truth from the Oxford-IIIT Pet dataset, along with the U-Net and SEU-Net predictions. The scales present are $0.25, 0.5, 1$ and 2 times the training scale.

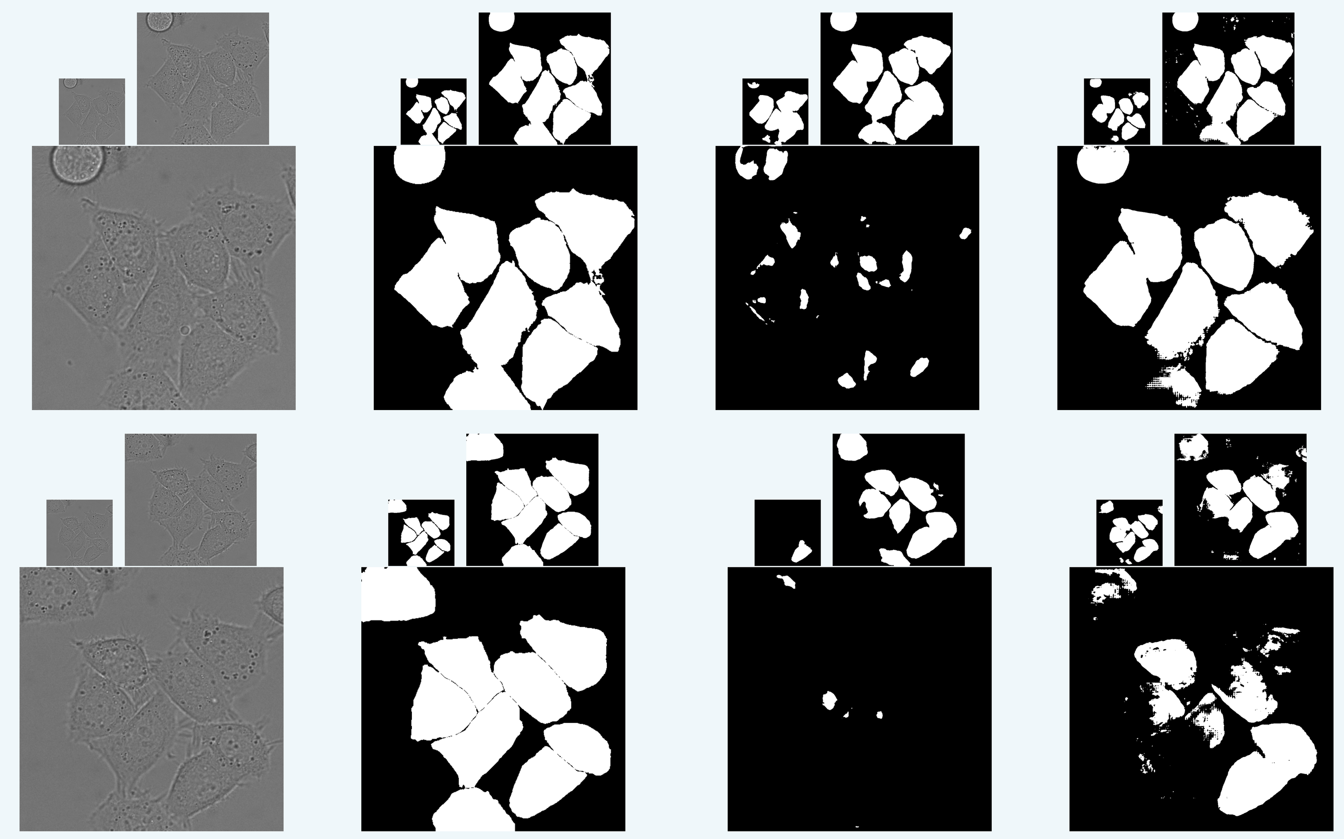


(a) IoU (b) Consistency

Figure 2. Overall results in terms of IoU and Consistency for each scale of the Pet dataset. Consistency is defined as the probability that the re-scaled samples are classfied in the same class as the original.

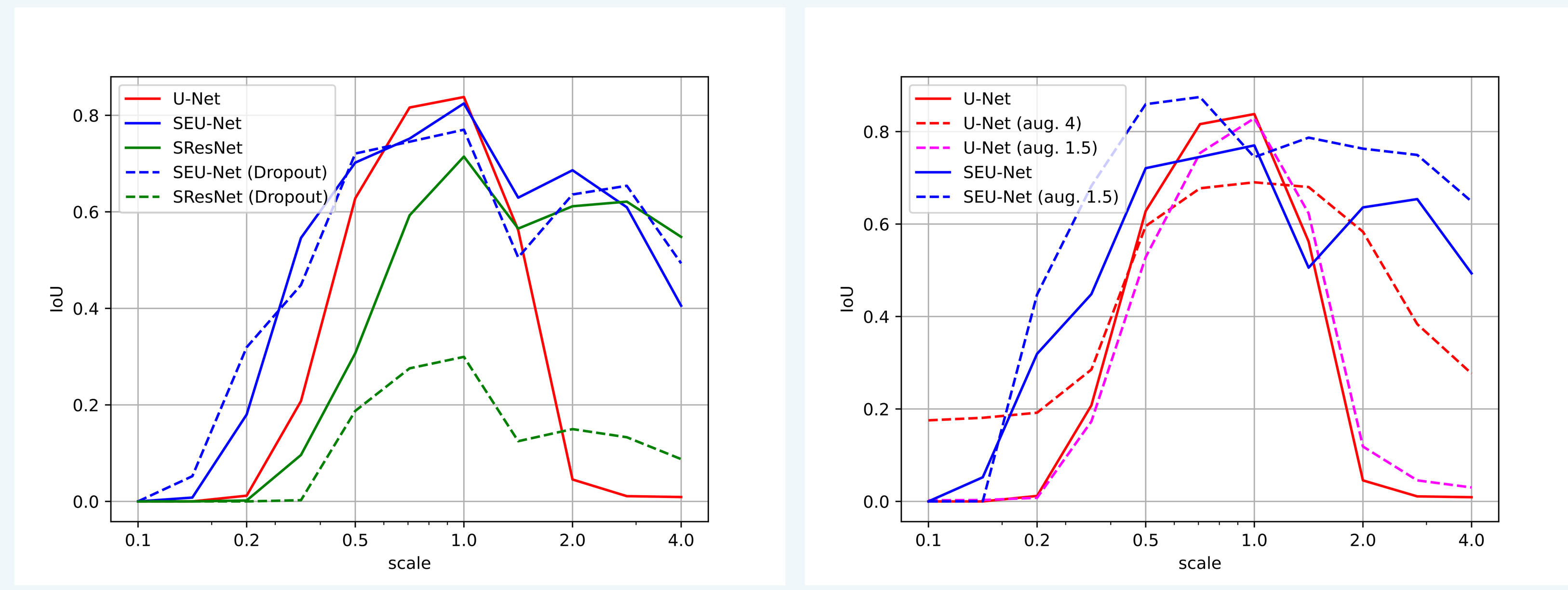
The code for the experimental section is available at https://github.com/mateussangalli/ScaleEquivariantUNet

DIC-HeLa [3]



(a) Image (b) Ground truth (c) U-Net (d) SEU-Net

Figure 3. Predictions from DIC-HeLa at different scales, namely scales 0.5, 1 and 2.



(a) IoU per scale. (b) Augmentation comparison.

Figure 4. IoUs of the cell segmentation experiment with comparisons with U-Net, SResNet. Aug. 4 refers to the U-Net trained with scale jittering with range 4 and aug. 1.5 with range 1.5.

Conclusions

- Generalization to unseen scales of the SEU-Net is improved compared to the U-Net, even when the latter is trained with scale-jittering to simulate the test scales.
- The SEU-Net also improves generalization compared to the SResNet [4], which uses scale-cross-correlations but does not perform upsampling operators inside the equivariant pipeline.
- It would be interesting apply the SEU-Net with a scale-convolution layer that supports a non-integer scale discretization.

References

[1] Omkar M. Parkhi et al. "Cats and Dogs". In: *IEEE Conference on Computer Vision and Pattern Recognition.* 2012.

[2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional networks for biomedical image segmentation". In: *International Conference on Medical image computing and computer-assisted intervention.* Springer. 2015, pp. 234–241.

[3] Vladimı´r Ulman et al. "An objective comparison of cell-tracking algorithms". In: *Nature methods* 14.12 (2017), pp. 1141–1152.

[4] Daniel Worrall and Max Welling. "Deep scale-spaces: Equivariance over scale". In: *Advances in Neural Information Processing Systems.* 2019, pp. 7364–7376.