DTU Compute
Department of Applied Mathematics and Computer Science

Robust and generalizable nuclei segmentation using deep learning

Christian Eschen[1], MSc student in Medicine and Technology

Supervisors: Jeppe Thaagard[1,2], Thomas Ebstrup[2], Ole Winther[1] and Rasmus R. Paulsen[1]

[1]Technical University of Denmark, [2]Visiopharm A/S

Abstract

Identification of cell nuclei in biomedical images is of great importance for research, drug discovery and diagnosis of disease. Examination of cancer suspicious tissue on the microscopic level is the golden standard for diagnosis of almost any cancer type. One of the greatest difficulties in segmentation of cell nuclei in histopathology is separation of adjoining cell nuclei. Separation of adjoining cell nuclei in deep learning is difficult due to the fact that convolutional neural networks (CNN) typically use a categorical cross entropy, which only work on a pixel-wise level and therefore lacks more global information context.

Therefore, there is an increasing demand for robust segmentation algorithms which are designed for separation of cell nuclei.

We present improved segmentation results using a feature engineered weighted loss and an adverserial loss. The best results were obtained using a U-Net architecture and an adverserial loss yielding an AJI equal to 0.452 on the validation set on the MoNuSeg grand challenge. We denote this architecture Pix2pix U-Net.

Data

The MoNuSeg [1] data consists of 30 H&E stained images of cell nuclei from multiple organs with 40x maginification and associated ground truth annotations of cell nuclei. The images are separated into 16 training images and 14 images for validation. A third class is introduced describing boundary of the cell nuclei

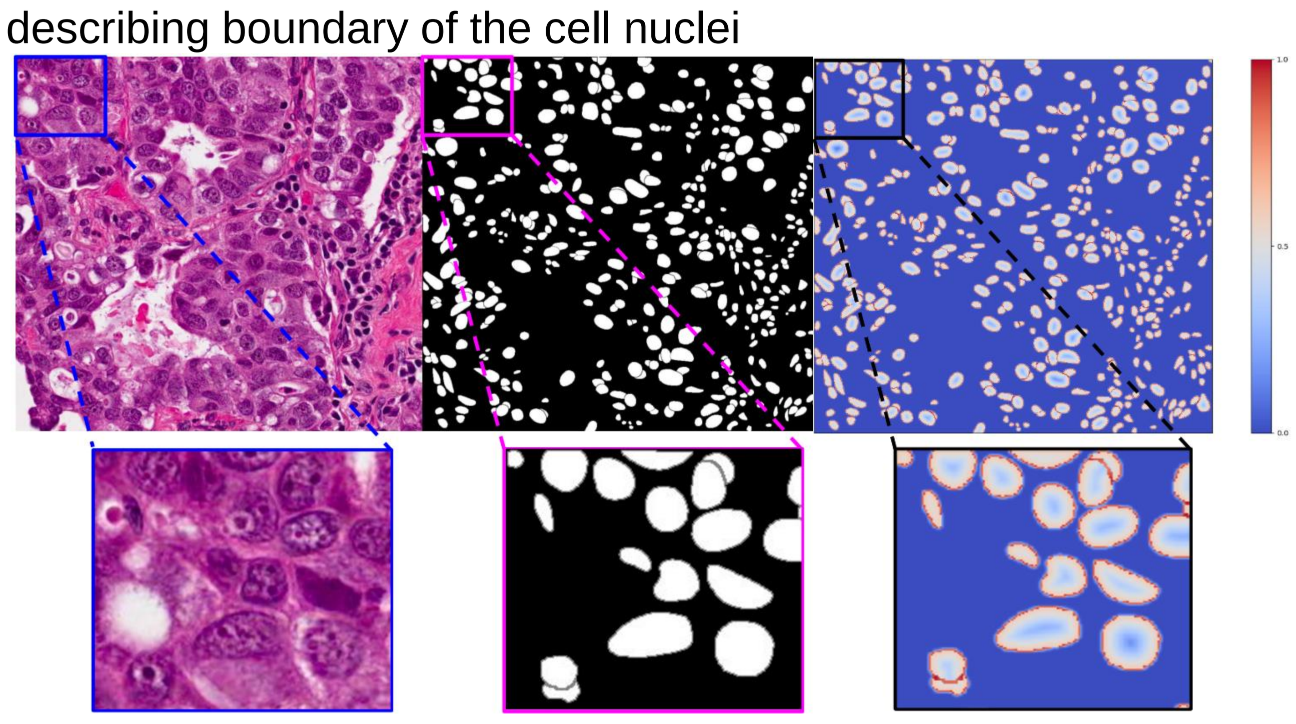

Figure 1: From left to right: H&E stained image of colon tissue. Ground truth image with three classes. Weight map used for training UNET SAW. The weight map is normalized for visualization only.

Methods

U-Net:

We employe a U-Net architecture [2] with a VGG encoder and a decoder with feature concatenation and bilinear upsampling, see Figure 2. The loss function used for training U-Net is:

$$\mathcal{L} = \sum_{x \in \Omega} \sum_{k'=1}^{K} W(\mathbf{x}) \cdot g(\mathbf{x}, k') \log(p_{k'}(\mathbf{x}))$$

where $g(\mathbf{x}, k')$ describes the ground truth image with pixels $\mathbf{x}$ and classes k', and the softmax activation is described as $p_{k'}(\mathbf{x})$ For the U-Net, we set the weight map to $W(\mathbf{x}) = 1$

U-Net SAW:

As described in [3], the separation border between cells can be enforced using a weight map in the cost function, which takes class frequencies into account and penalizes predictions close to the boarder of the nuclei:

$$W(\mathbf{x}) = W(\mathbf{x}, \tau, \sigma) = w_c(\mathbf{x}) + F_\sigma * w_0(\mathbf{x}, \tau)$$

The weight map $W(\mathbf{x})$ has now a contribution from the class areas $w_c(\mathbf{x})$ and the inverse distance transform $w_0(\mathbf{x}, \tau)$ smooted with a Gaussian kernel F_σ. The inverse distance transform is defined by the following equation with τ describing normalization, $\phi_{K\,bin}(\mathbf{x})$ is the euclidean distance transform of the foreground T_H:

$$w_0(\mathbf{x}, \tau) = \begin{cases} 1 - \frac{\phi_{K\,bin}(\mathbf{x})}{\tau} & x \in T_H \\ 0 & otherwise \end{cases}$$

The weight map is presented in Figure 1.

Pix2pix:

Instead of feature engineering the loss, one could define an objective "be indestinquishable from the ground truth". This is the goal in generative adverserial networks (GANS). We utilize the cGAN proposed in [4] for semantic segmentation.

This is achieved by a so-called min-max game in which a generator G tries to fool a discriminator D:

$$G^* = \arg\min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G)$$

The $\lambda \mathcal{L}_{L1}(G)$ term acts as a regularizer forcing the segmentations to be as close as possible to the ground truth. The Pix2pix U-Net framework is presented in Figure 3.

	U-Net	U-Net SAW	Pix2pix
Mean dice	0.769	0.761	0.819
Mean AJI	0.404	0.429	0.452
Mean Hausdorff	16.0	16.5	14.4

Figure 4: Comparison of the performance of the models

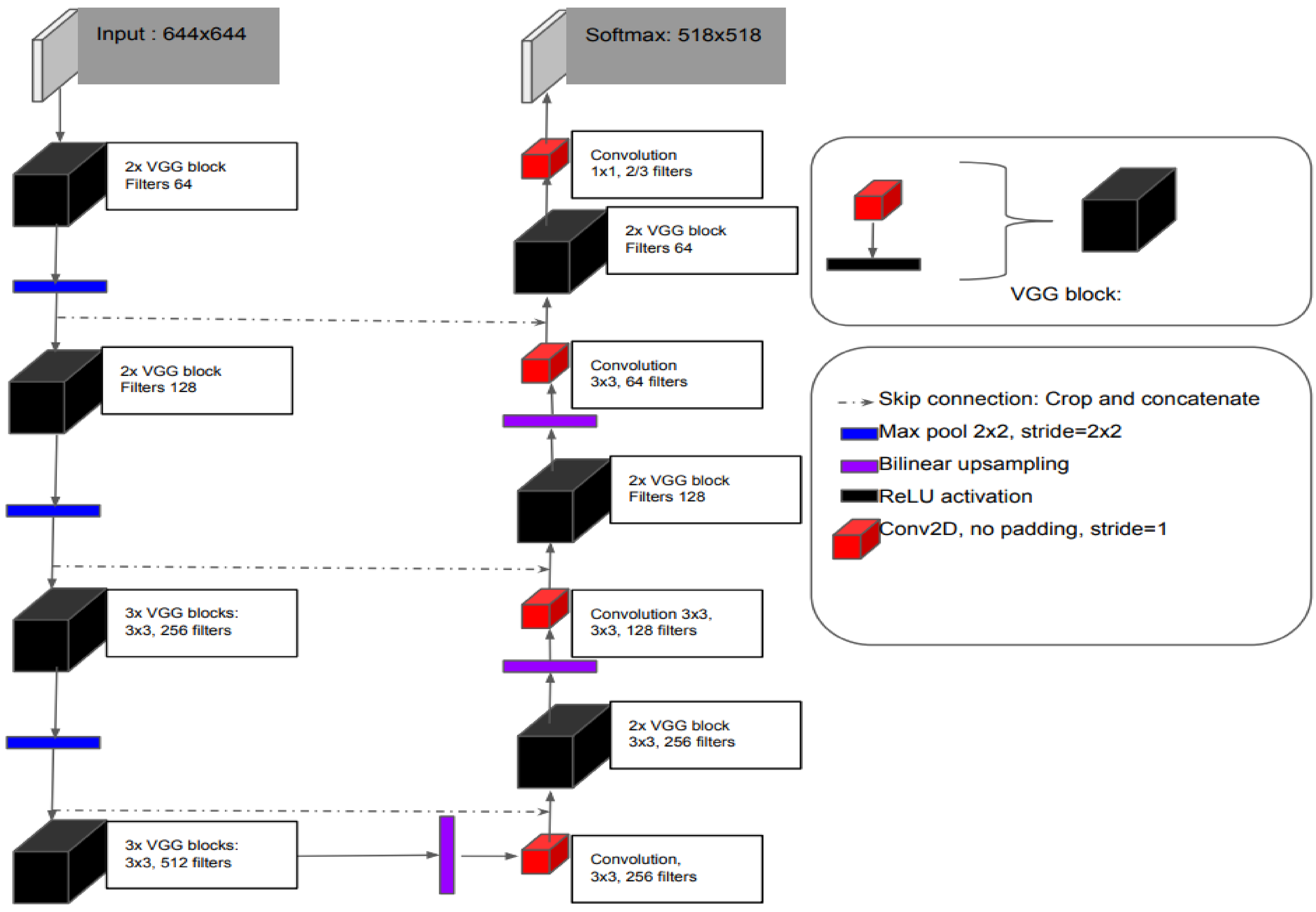

Figure 2: VGG U-Net

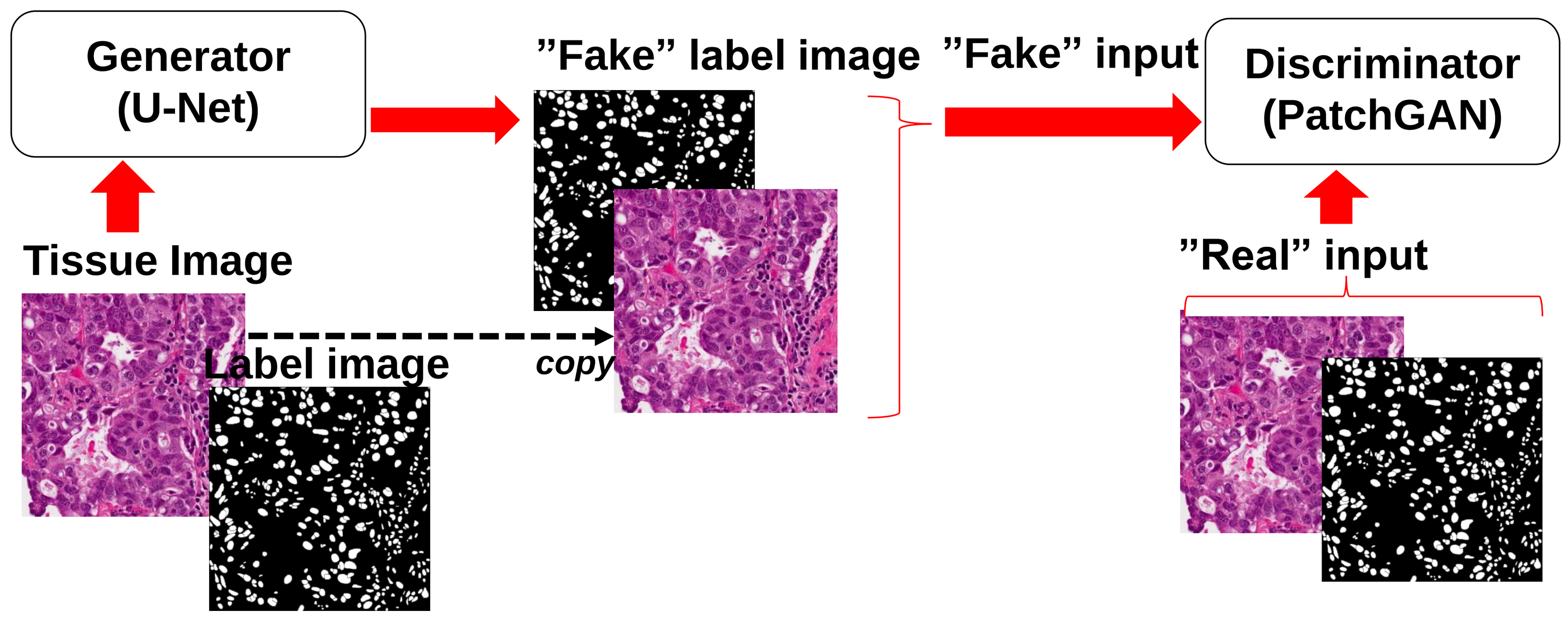

Figure 3: Pix2pix segmentation framework

Results

The segmentation performance is measured using the Dice score, Hausdorff distance and the AJI index [1]. The validation metrics for the models are presented in Figure 4 and the segmentation maps are presented in Figure 5.

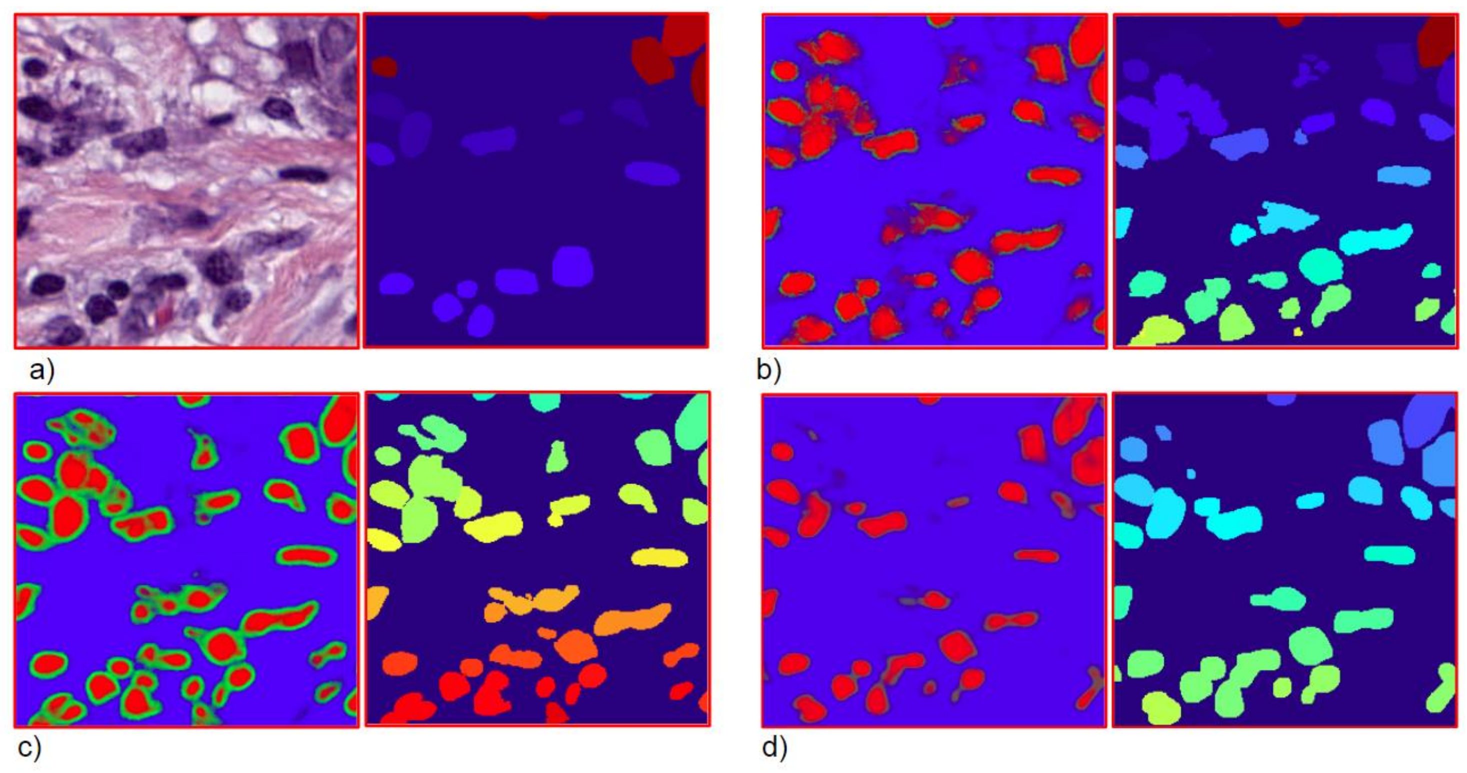

Figure 6: a) H&E stained image and ground truth image of breast tissue. b) Softmax activation and segmentations using U-Net. c) Softmax activation and segmentations using U-Net SAW. d) Softmax activation and segmentations using Pix2pix U-Net.

Discussion

We present improved discriminative properties of the CNN regarding ajoining cell nuclei separation using U-Net SAW. The best segmenation performance was achieved using the Pix2pix U-Net yielding a mean AJI equal to 0.452. However, the Pix2pix U-Net has not learned the boundary class sufficiently. Future work will investigate methods for successfully learning the boundary class.

References

[1] Kumar, Neeraj, et al. "A dataset and a technique for generalized nuclear segmentation for computational pathology." *IEEE transactions on medical imaging* 36.7 (2017): 1550-1560.

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

[3] Guerrero-Pena, Fidel A., et al. "Multiclass Weighted Loss for Instance Segmentation of Cluttered Cells." *arXiv preprint arXiv:1802.07465* (2018).

[4] Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." *arXiv preprint* (2017).

Contact information

Christian_eschen@hotmail.com
Ph. 20652915

DTU Compute
Department of Applied Mathematics and Computer Science