A NEW COLOR AUGMENTATION METHOD FOR DEEP LEARNING SEGMENTATION OF HISTOLOGICAL IMAGES

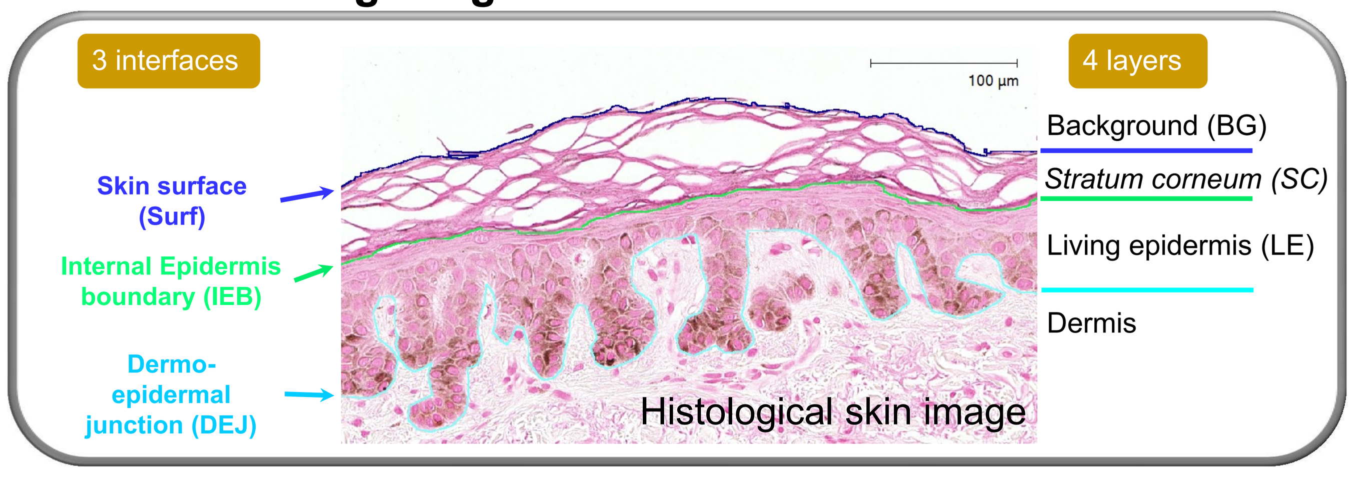
Yang Xiao[1] , Etienne Decencière[1] , Santiago Velasco-Forero[1] , Hélène Burdin[2]
Thomas Bornschlögl[3], Françoise Bernerd[3] , Emilie Warrick[3] and Thérèse Baldeweck[3]

[1] MINES ParisTech, PSL Research University, Centre for Mathematical Morphology, Fontainebleau, France
[2]ADCIS SA, 3 rue Martin Luther King, 14280 Saint-Contest, France
[3] L'Oréal Research and Innovation, Aulnay-sous-Bois, France

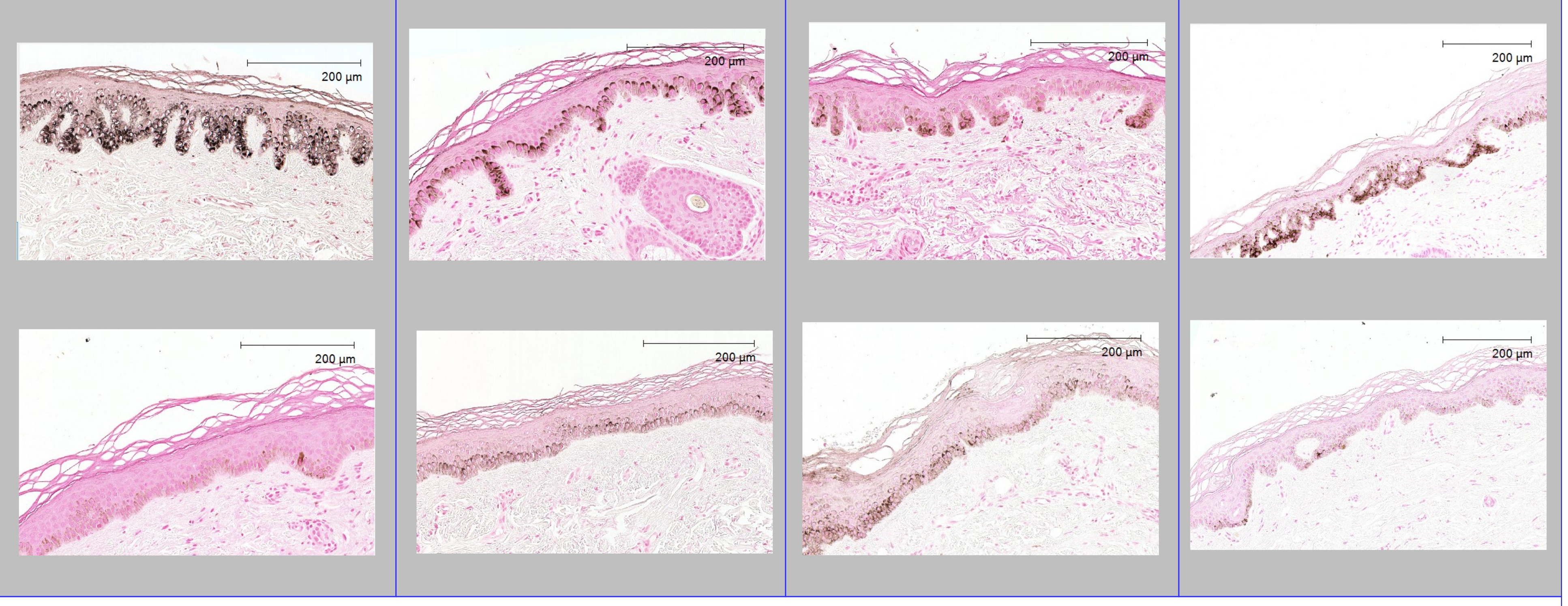
INTRODUCTION

Challenges in histological image segmentation

- With the recent advent of digital whole slide imaging, the number and the size of acquired images are growing up and there is a need of finding ways to also adapt the throughput of image quantification.
- Deep neural networks usually require large training sets to achieve an acceptable performance, while the generation of the segmentation ground-truth is very time-consuming.
- Given complex tissue structures and inconsistencies in sample preparation, **network generalization** is crucial.
- **We show that a deep neural network can learn a satisfactory segmentation model with relatively few data thanks to a convenient image augmentation method.**



MATERIALS

Data base	Data base n°1			Data base n°2
	Training	Validation	Test	Generalization Test
N=	26 images	9 images	41 images	52 images
Study	Clinical studies n°1 & 2			Clinical study n°3
				
Typical image size	1 million pixels (~1000 x 1000 pixels) to 10 million pixels (~3000 x 3000 pixels)			

METHODS

TRAINING

Topological pre-processing [1]

Conversion to Lab color space

512x512 crops extraction

Fully convolutional network training (U-Net)
- **Color augmentation**
- Geometric augmentation
- Jaccard loss [2]

$$L_J = 1 - \frac{\sum_{i,j,k}(t_{ijk}p_{ijk})}{\sum_{i,j,k} t_{ijk}^2 + \sum_{i,j,k} p_{ijk}^2 - \sum_{i,j,k}(t_{ijk}p_{ijk})}$$

Transformation of 3-channel output into labelled image

Post-processing: removal of spurious components [1]

TESTING / PREDICTION

Topological pre-processing [1]

Conversion to Lab color space

U-Net application to complete histological image

Transformation of 3-channel output into labelled image

Post-processing: removal of spurious components [1]

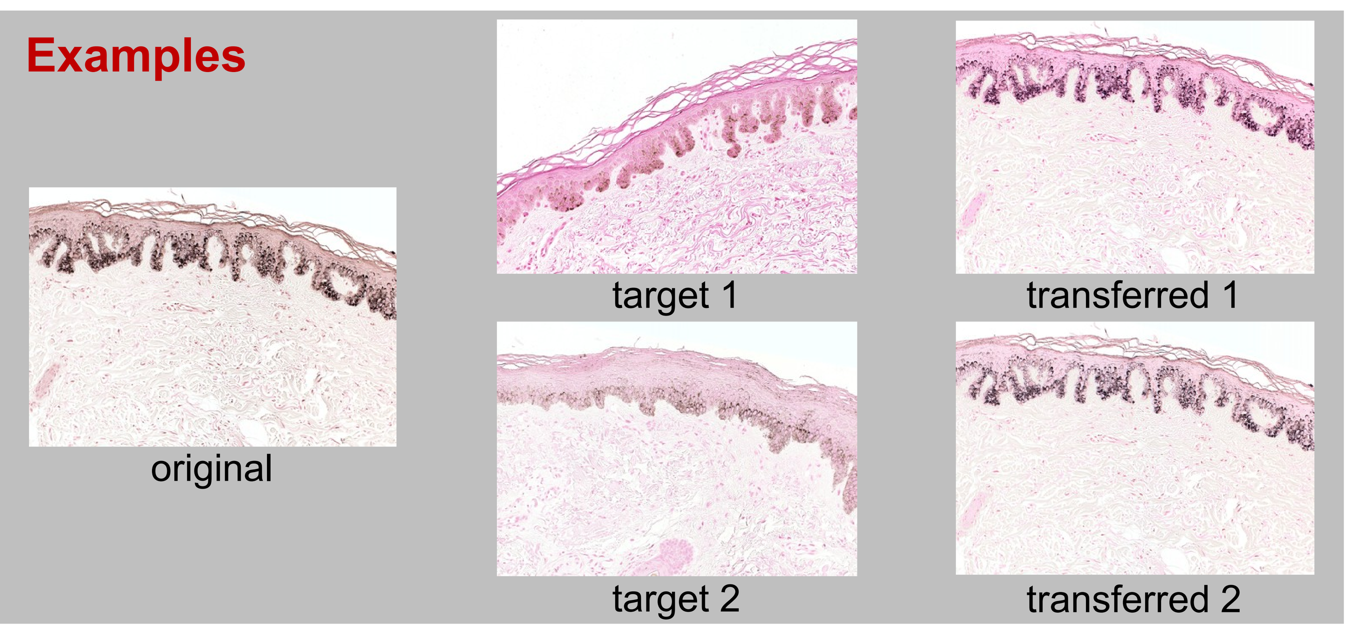
COLOR AUGMENTATION IN Lab COLOR SPACE

At training time, before feeding a crop (the *original* crop) into the network, another one (the *target* crop) is chosen and each channel of the original one is transformed:

$$C_{\text{transferred}} = C_{\text{original}} - \overline{C}_{\text{original}} + \overline{C}_{\text{target}}$$

This transformation is only applied to the pixels considered as belonging to the biological sample (threshold $0.86L_{max}$ applied to the L channel)

Examples



RESULTS

Train: the more augmentation, the less overfitting

Method	Best train loss	Best val loss
No augm	0.0060	0.0307
Geom	0.0133	0.0243
Color	0.0114	0.0272
Mixed	0.0178	0.0211

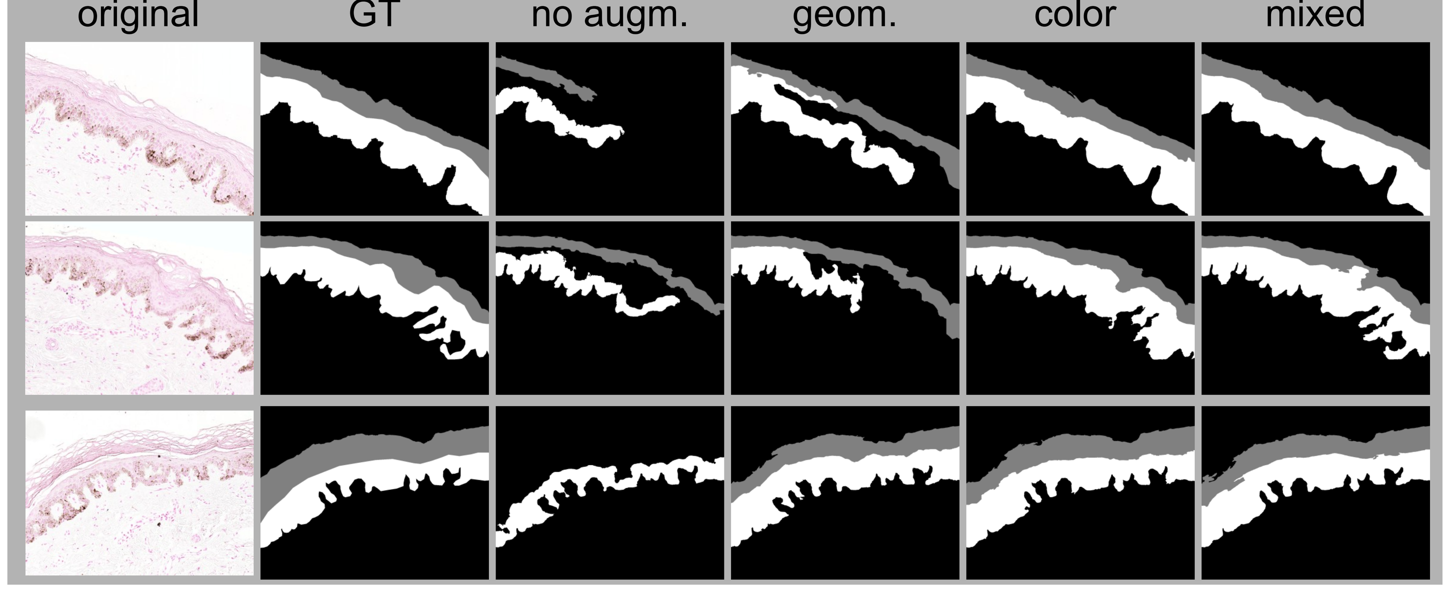
Results on Database n°1: "Test"
Color and mixed augmentation clearly enhance results

Method	Jaccard index (per class)			Mean dist. GT / prediction		
	BG	SC	LE	Surf.	IEB	DEJ
No augm	0.97	0.77	0.87	54.6	311.5	7.9
Geom	**0.99**	0.82	0.89	16.3	30.9	5.1
Color	**0.99**	**0.89**	0.91	2.1	**9.7**	4.3
Mixed	**0.99**	**0.89**	**0.92**	**2.0**	**9.7**	**3.6**

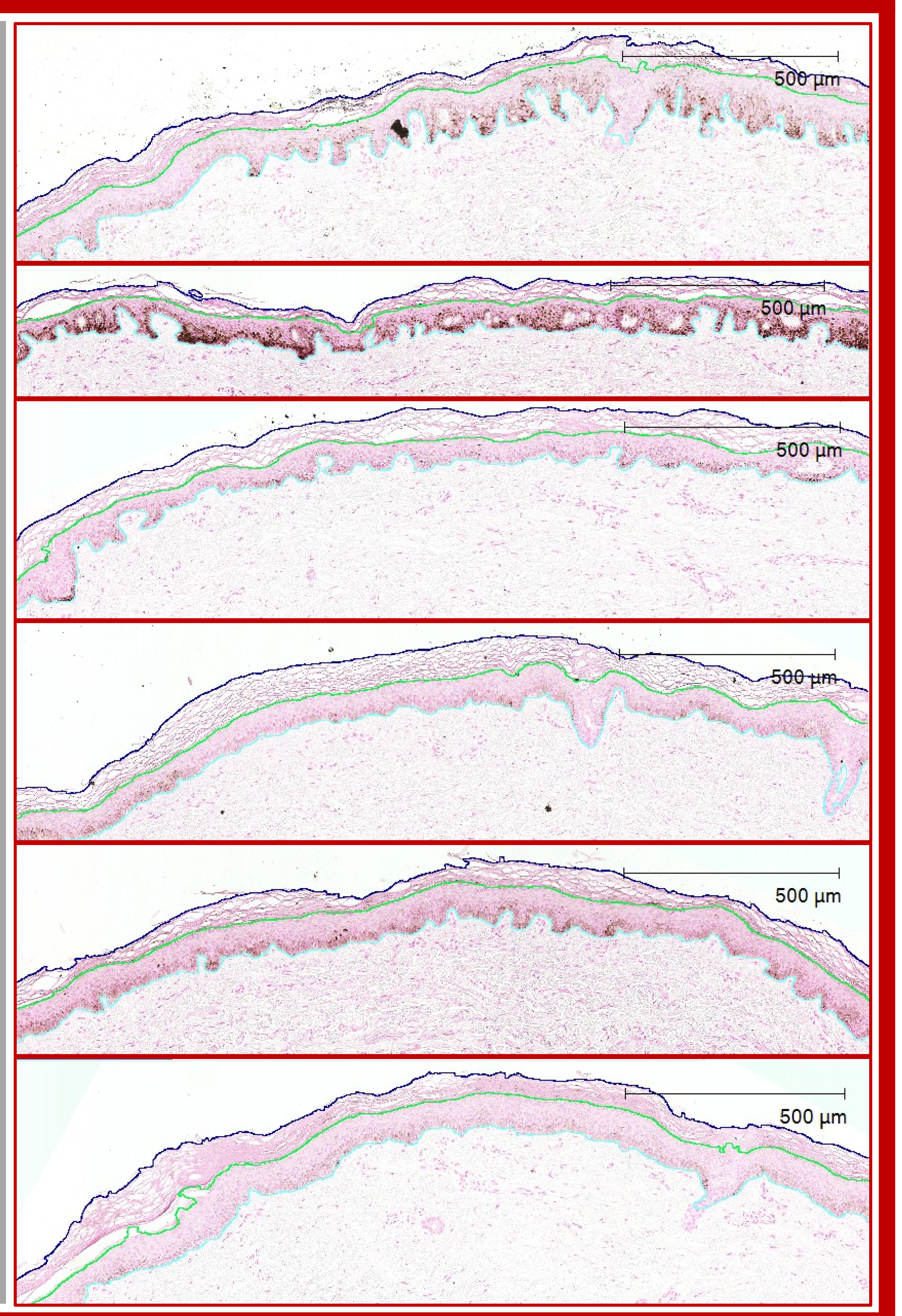
Results on Database n°2: "Generalization test"
Thanks to color and mixed augmentation the model performance is the same

Method	Jaccard index (per class)			Mean dist. GT / prediction		
	BG	SC	LE	Surf.	IEB	DEJ
No augm	0.92	0.47	0.70	156.6	992.5	58.5
Geom	0.95	0.74	0.70	19.2	321.5	72.8
Color	**0.99**	0.91	0.91	1.8	9.0	5.1
Mixed	**0.99**	**0.92**	**0.92**	**1.4**	**8.0**	**4.2**

Examples



Segmentation results on Database n° 2



CONCLUSIONS & PERSPECTIVES

- A new image augmentation method has been proposed
- We have shown that it helps improving the generalization capacity of a fully convolutional neural network trained to segment histological images
- It has been integrated in an industrial software

Perspectives:

- Generative adversarial networks might help improving the augmentation method
- Data augmentation considering elastic deformation could be combined with the proposed color augmentation method to increase the variety

REFERENCES

[1] E. Decencière et al., "Dealing with topological information within a fully convolutional neural network," in ACIVS, 2018.
[2] Y. Yuan et al. "Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance," IEEE TMI, vol. 36, pp. 1876–1886, 2017.