CELL SEGMENTATION IN DIGITIZED PAP SMEAR IMAGES USING AN ENSEMBLE OF FULLY CONVOLUTIONAL NETWORKS

UNIVERSITY of DEBRECEN

Gergo Bogacsovics,
Andras Hajdu,
Balazs Harangi

OUR GOALS

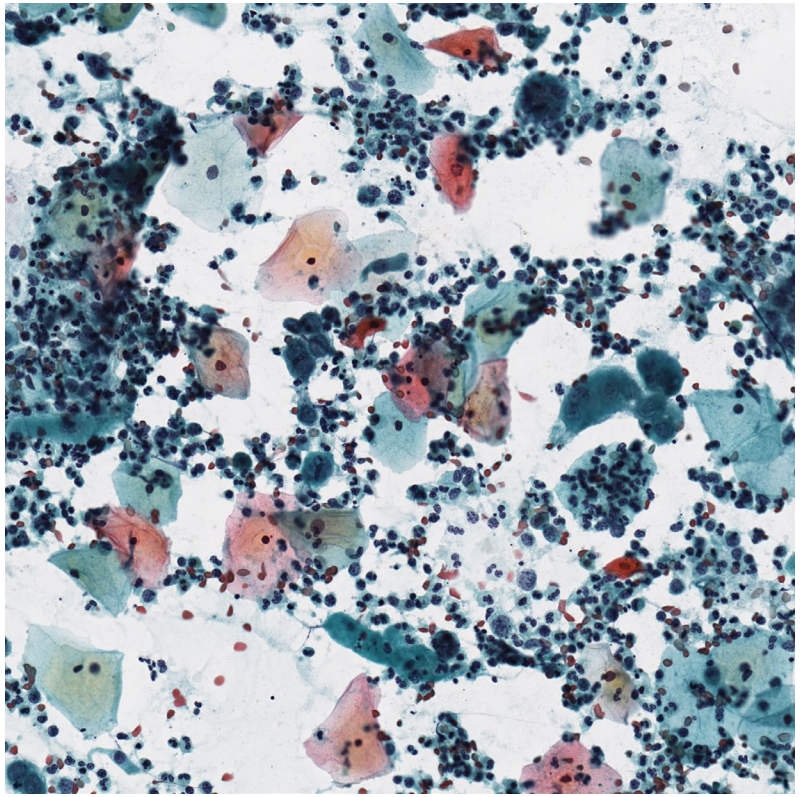
Final Goal

- Automatic screening system
- Early detection of cervical cancer

Current Goal

- Accurate segmentation of cells

FUTURE GOAL



Input image

- Patient DOES have cancer
- Patient DOES NOT have cancer

Output

FUTURE GOAL

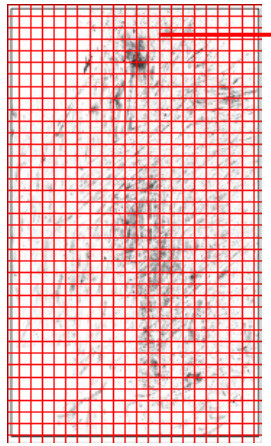
3D Histec Scanner

Input MRXS image
4-12 GB
cca. 100000 x 220000 px

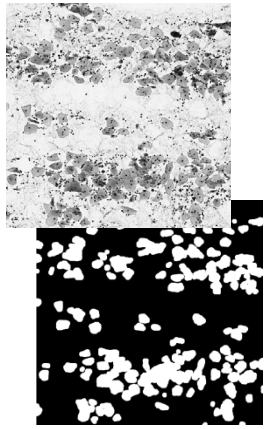
Internet

1. ROI identification

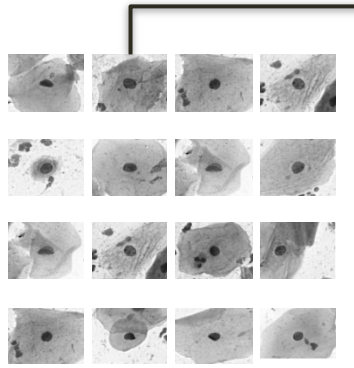
2. Smear slicing



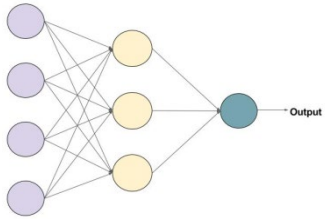
3. Preprocessing, cell segmentation



4. Extraction of the individual cells

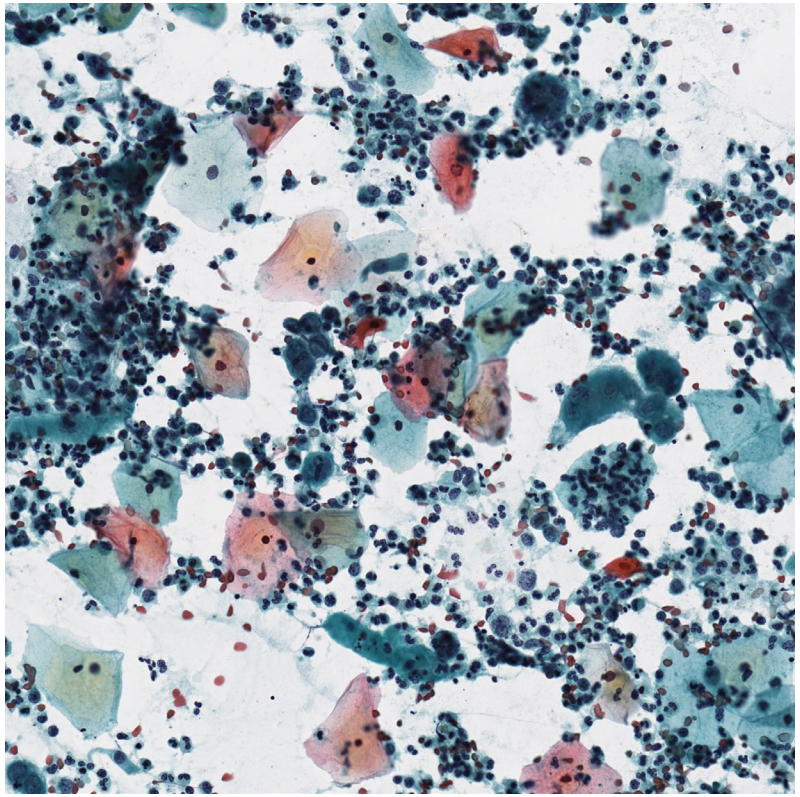


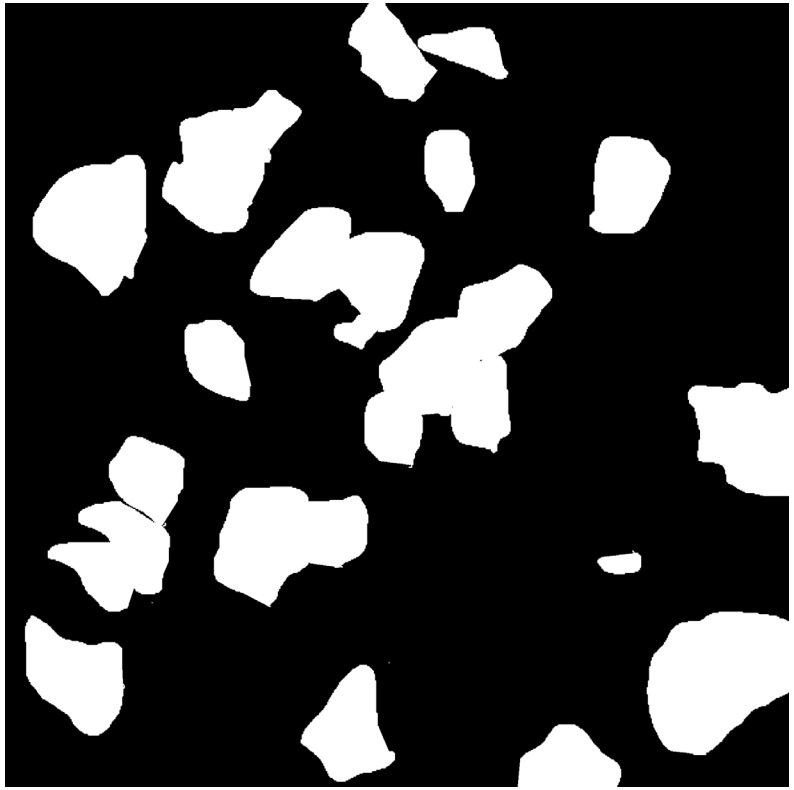
5. Cell classification



6. Final decision

CURRENT GOAL



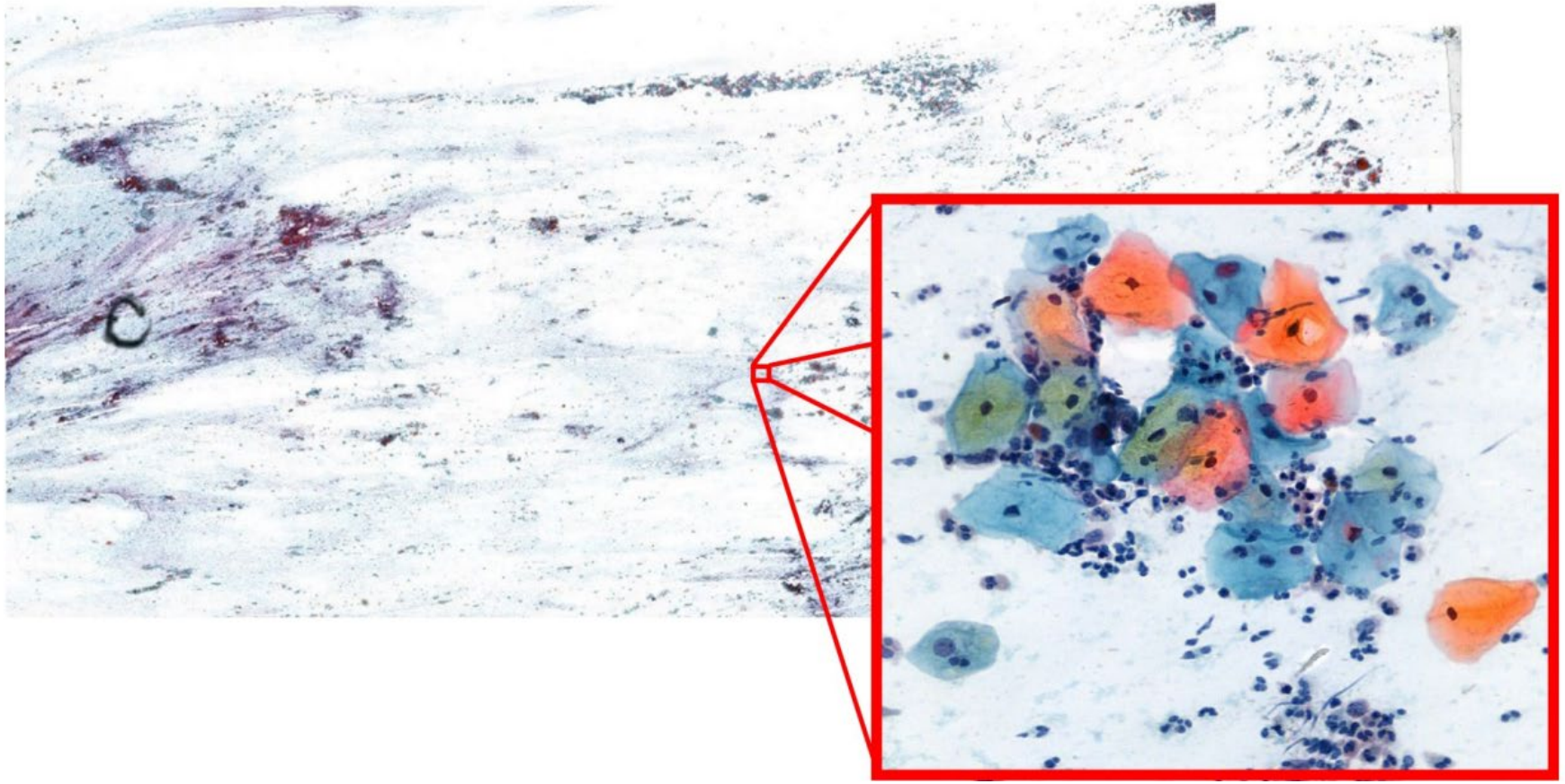


Input image

Output image

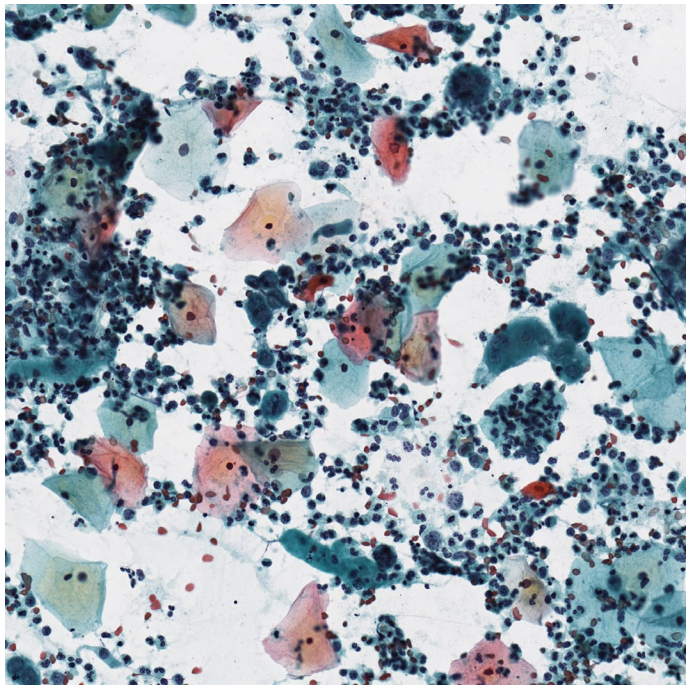
DATASET

TRAINING SET AND TEST SET



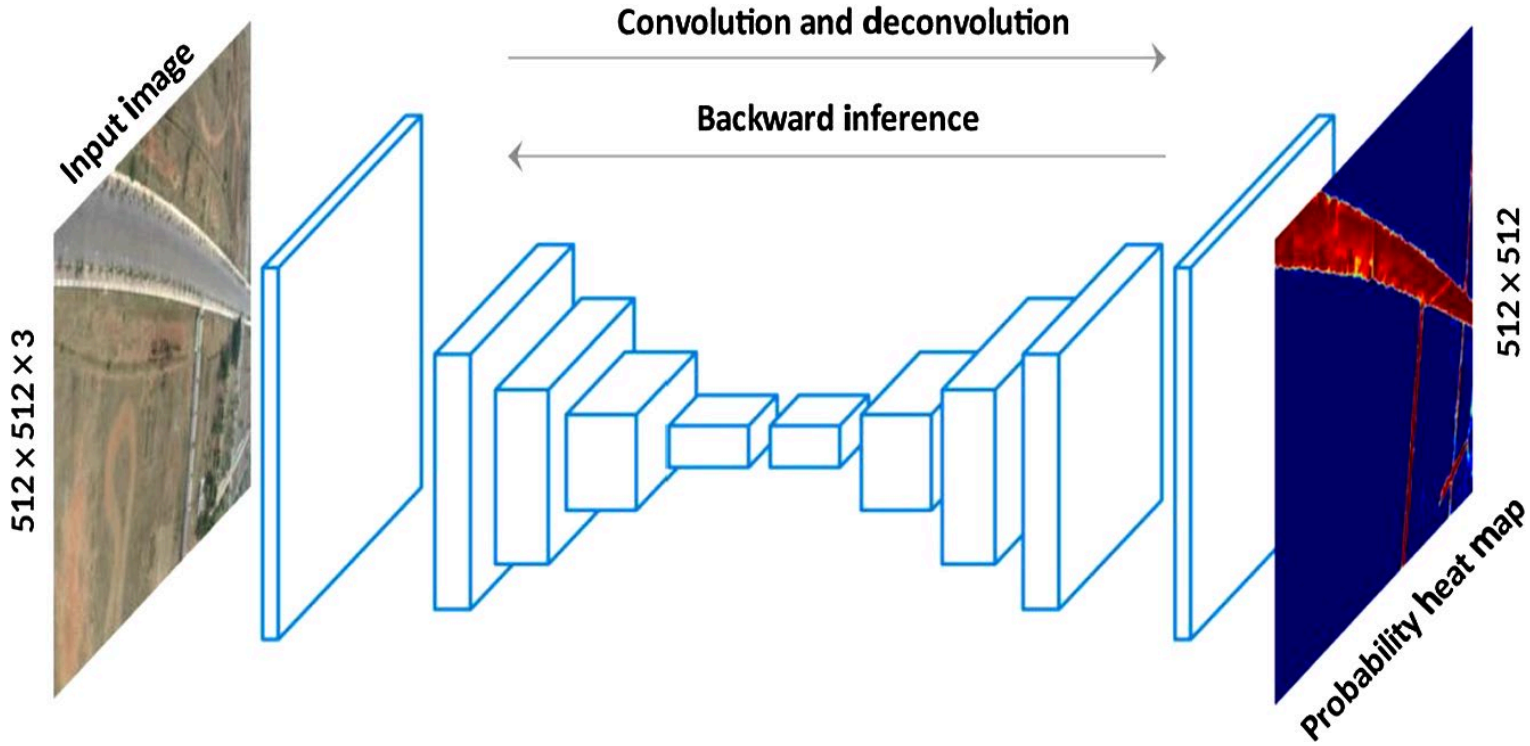
TRAINING SET AND TEST SET

- Pap smear images
- Image-mask pairs
- Annotated by clinical experts
- 2 257 images in total
- 500x500 pixels



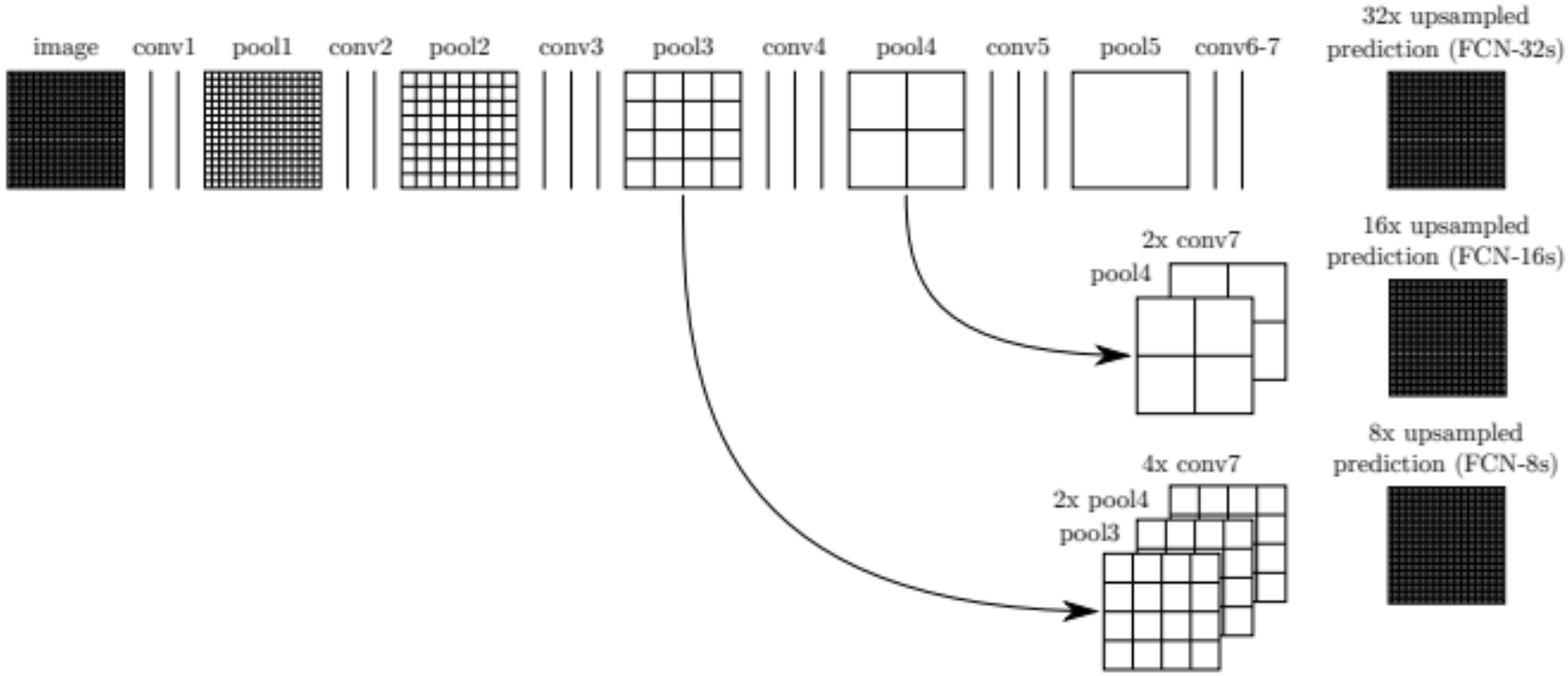
FCNS

FULLY CONVOLUTIONAL NETWORKS[1]



Ramesh Kestur, Shariq Farooq, Rameen Abdal, Emad Mehraj, Omkar Subbaramajois Narasipura, and Meenavathi Mudigere
"UFCN: a fully convolutional neural network for road extraction in RGB imagery acquired by remote sensing from an unmanned aerial vehicle,,,
Journal of Applied Remote Sensing 12(1), 016020 (13 February 2018). https://doi.org/10.1117/1.JRS.12.016020

FULLY CONVOLUTIONAL NETWORKS[1]



Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation.„ *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

THE PROPOSED ALGORITHM

THE PROPOSED ALGORITHM - INTUITION

Input image FCN-32 FCN-16 FCN-8

THE PROPOSED ALGORITHM - INTUITION

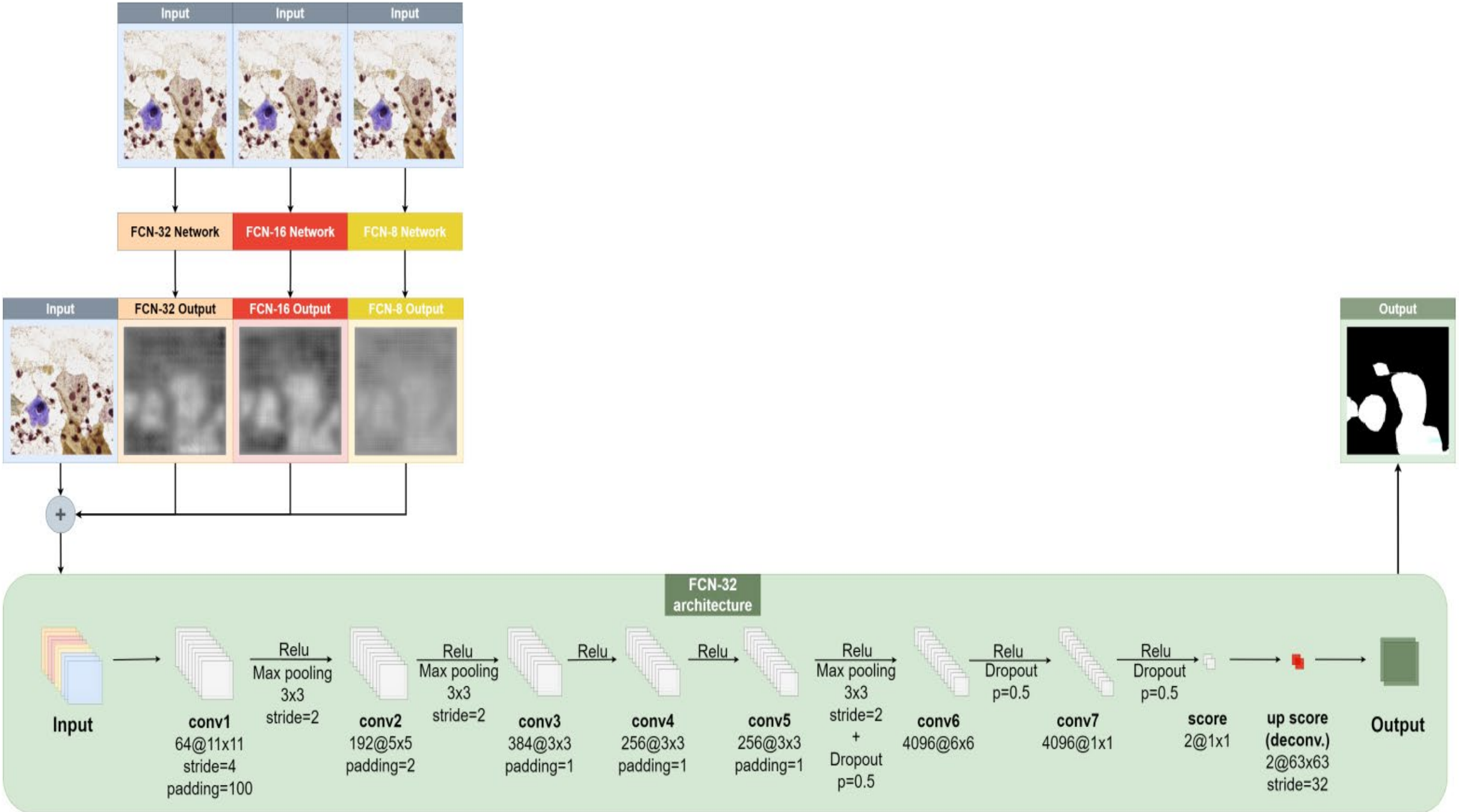
- An ensemble approach would probably yield **better results**
- Building this ensemble is **not trivial**

THE PROPOSED ALGORITHM

- Train the FCN architectures **separately**
- Train a modified FCN architecture that receives the outputs of the other networks as input
  - The network receives both the input image **and** these outputs as input
  - It can not only combine the outputs but also come to its **own decisions**

THE PROPOSED ALGORITHM

- We propose **multiple versions** of the architecture
  - Depending on the number of extra input channels
  - E.g. $C_{32\text{-}8}$ receives the input image and the outputs of pre-trained FCN-32 and FCN-8 networks


Input
Input
Input
FCN-32 Network
FCN-16 Network
FCN-8 Network
Input
FCN-32 Output
FCN-16 Output
FCN-8 Output
+
Output
FCN-32 architecture
Input
conv1
64@11x11
stride=4
padding=100
Relu
Max pooling
3x3
stride=2
conv2
192@5x5
padding=2
Relu
Max pooling
3x3
stride=2
conv3
384@3x3
padding=1
Relu
conv4
256@3x3
padding=1
Relu
conv5
256@3x3
padding=1
Relu
Max pooling
3x3
stride=2
+
Dropout
p=0.5
conv6
4096@6x6
Relu
Dropout
p=0.5
conv7
4096@1x1
Relu
Dropout
p=0.5
score
2@1x1
up score
(deconv.)
2@63x63
stride=32
Output

THE TRAINING PROCEDURE

- We divided the dataset into **three parts.** We used the
  - **1st part:** for training the FCN algorithms
  - **2nd part:** for training the combined network
  - **3rd part:** for evaluation.
- We used **cross-validation:**
  - We shuffled the three parts around

EVALUATION

BASELINES

- FCN networks (FCN-32, FCN-16, FCN-8) [1]
- Sota [2]
- DeepLab (v3) [3]
- U-Net [4]
- GSCNN [5]
- Our previous ensemble [6]

METRICS

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$IoU = \frac{TP}{TP + FP + FN}$$

$$DSC = \frac{2TP}{2TP + FP + FN}$$

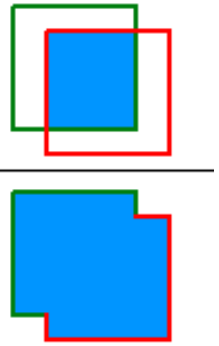
METRICS

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$IoU = \frac{TP}{TP + FP + FN}$$

$$DSC = \frac{2TP}{2TP + FP + FN}$$

$$IOU = \frac{\text{area of overlap}}{\text{area of union}} =$$



Padilla, Rafael, Sergio L. Netto, and Eduardo AB da Silva.
"A survey on performance metrics for object-detection algorithms.,,
2020 International Conference on Systems, Signals and Image Processing (IWSSIP).
IEEE, 2020.

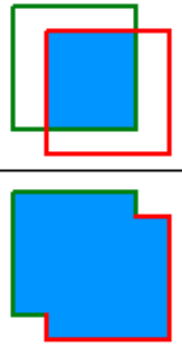
METRICS

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$IoU = \frac{TP}{TP + FP + FN}$$

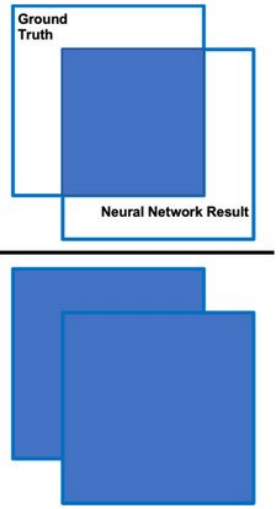
$$DSC = \frac{2TP}{2TP + FP + FN}$$

$$IOU = \frac{\text{area of overlap}}{\text{area of union}} =$$



Padilla, Rafael, Sergio L. Netto, and Eduardo AB da Silva.
"A survey on performance metrics for object-detection algorithms.,,
2020 International Conference on Systems, Signals and Image Processing (IWSSIP).
IEEE, 2020.

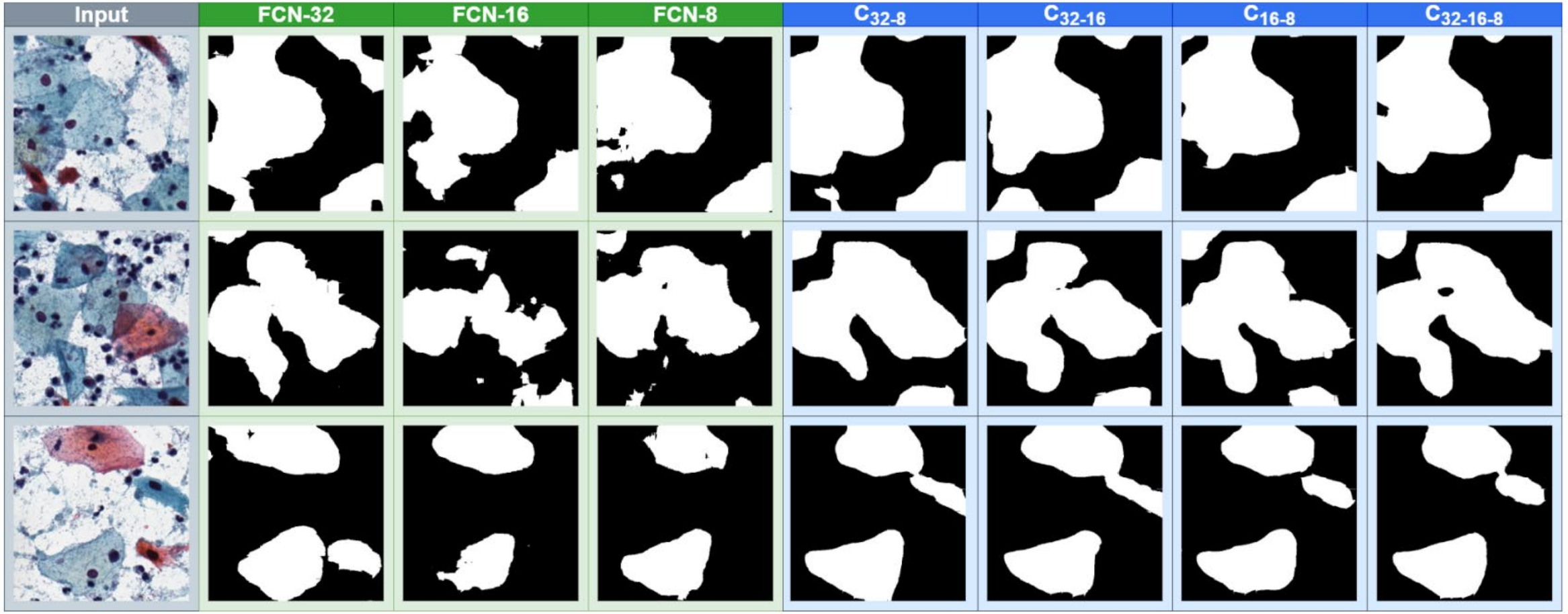
$$\text{Dice Score} = \frac{2 \times \text{Area of Overlap}}{\text{Area of Union}}$$



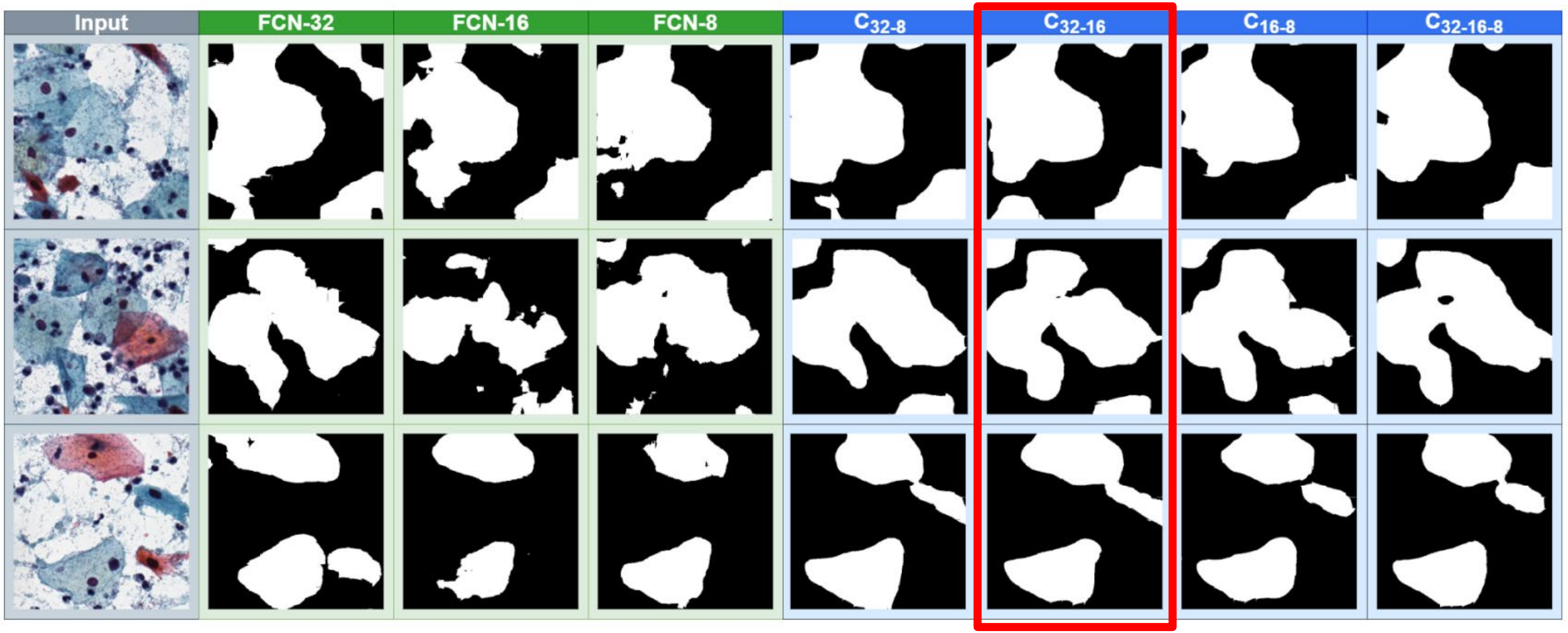
Bardis, Michelle, et al.
"Deep learning with limited data: Organ segmentation performance by U-Net.,,
Electronics 9.8 (2020): 1199.

RESULTS

RESULTS



RESULTS



$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$IoU = \frac{TP}{TP + FP + FN}$$

$$DSC = \frac{2TP}{2TP + FP + FN}$$

RESULTS

Algorithm	*ACC*	*IoU*	*DSC*
FCN-32 [5]	0.915 ± 0.054	0.497 ± 0.161	0.660 ± 0.144
FCN-16 [5]	0.913 ± 0.063	0.503 ± 0.143	0.666 ± 0.127
FCN-8 [5]	0.919 ± 0.037	0.507 ± 0.180	0.668 ± 0.158
sota [13]	0.775	0.343	
Ens_1 [8]	0.923 ± 0.022	0.534 ± 0.239	0.688 ± 0.205
Ens_2 [8]	0.923 ± 0.020	0.534 ± 0.243	0.688 ± 0.208
DeepLabv3[14]	0.889 ± 0.117	0.487 ± 0.039	0.655 ± 0.035
U-Net [9]	0.917 ± 0.063	0.504 ± 0.224	0.662 ± 0.199
GSCNN [10]	0.909 ± 0.091	0.514 ± 0.185	0.674 ± 0.162
C_{32-8}	0.926 ± 0.034	0.530 ± 0.195	0.688 ± 0.168
C_{32-16}	0.928 ± 0.036	0.534 ± 0.191	0.691 ± 0.163
C_{16-8}	**0.928 ± 0.031**	**0.537 ± 0.203**	**0.693 ± 0.173**
$C_{32-16-8}$	0.927 ± 0.040	0.531 ± 0.175	0.687 ± 0.147

REFERENCES

1. **J. Long, E. Shelhamer, and T. Darrell,** "*Fully convolutional networks for semantic segmentation*" in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.
2. **Lu, Zhi, Gustavo Carneiro, and Andrew P. Bradley**. "*Automated nucleus and cytoplasm segmentation of overlapping cervical cells.*" International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Berlin, Heidelberg, 2013.
3. **Chen, Liang-Chieh, et al.** "*Rethinking atrous convolution for semantic image segmentation.*" arXiv preprint arXiv:1706.05587 (2017).

REFERENCES

4. **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.
5. **Takikawa, Towaki, et al.** "*Gated-scnn: Gated shape cnns for semantic segmentation.*" Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.
6. **Harangi, Balazs, et al.** *"Cell detection on digitized pap smear images using ensemble of conventional image processing and deep learning techniques."* 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA). IEEE, 2019.

THANK YOU FOR YOUR ATTENTION!

Új Nemzeti
Kiválóság Program

Innovációs és Technológiai
Minisztérium

RESEARCH REPORTED IN THIS PUBLICATION WAS SUPPORTED BY THE ***ÚNKP-21-3-I-DE-99*** *AND THE* ***ÚNKP-20-5-DE-31*** *NEW NATIONAL EXCELLENCE PROGRAM OF THE MINISTRY FOR INNOVATION AND TECHNOLOGY FROM THE SOURCE OF THE NATIONAL RESEARCH, DEVELOPMENT AND INNOVATION FUND.*

RESEARCH WAS ALSO SUPPORTED IN PART BY THE ***JANOS BOLYAI RESEARCH SCHOLARSHIP*** *OF THE HUNGARIAN ACADEMY OF SCIENCES AND THE* ***GINOP-2.2.1-18-2018-00012*** *SUPPORTED BY THE EUROPEAN UNION, CO-FINANCED BY THE EUROPEAN SOCIAL FUND.*

ANY QUESTIONS?