Virtual Summer School on Deep Learning

July 5th - 9th, 2021

4th International Summer School on Deep Learning - Virtual Edition

#VSSonDL
#ISSonDL

Image Segmentation & Deep Learning

Alexandr Kalinin

Shenzhen Research Institute of Big Data, China
University of Michigan, USA
@alxndrkalinin

Organizers

GDAŃSK UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRONICS, TELECOMMUNICATIONS AND INFORMATICS
BIOMEDICAL ENGINEERING DEPARTMENT
aiBAY
Digital Innovation Hub dih4.ai
intel
SiMa ai

Sponsors

Diamond Sponsor
amazon

Platinum Sponsor
JIT
JitTeam™ the human factor of IT

GDAŃSK UNIVERSITY OF TECHNOLOGY
NVIDIA
FACULTY OF ELECTRONICS, TELECOMMUNICATIONS AND INFORMATICS
servodata elektronik sp. z o.o.
BIOMEDICAL ENGINEERING DEPARTMENT
Digital Innovation Hub dih4.ai
aiBAY
intel
SiMa ai

Introduction

My background:

- MSc in Applied Math and CS (Novosibirsk State Tech University, Russia)
- Fulbright Scholar in Statistics (UCLA)
- PhD in Bioinformatics (UMich)
- International postdoc at SRIBD (Shenzhen, China) and UMich (USA)

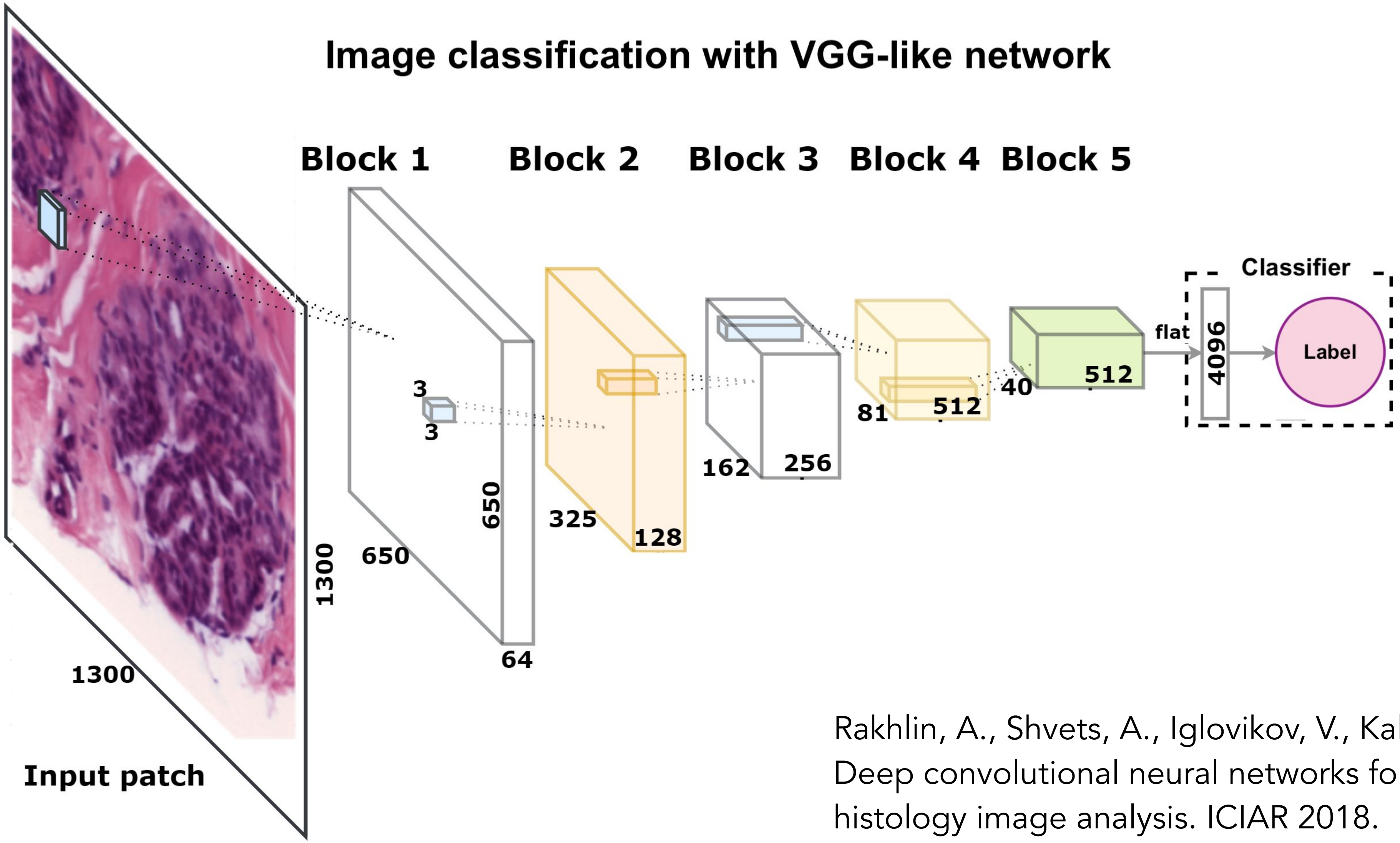
Research: biomedical image analysis, machine learning, visual analytics

https://alxndrkalinin.github.io

Overview

1. Background and motivation
2. U-net and friends
3. Recent advancements
4. U-net implementation walkthrough

Recap: image classification

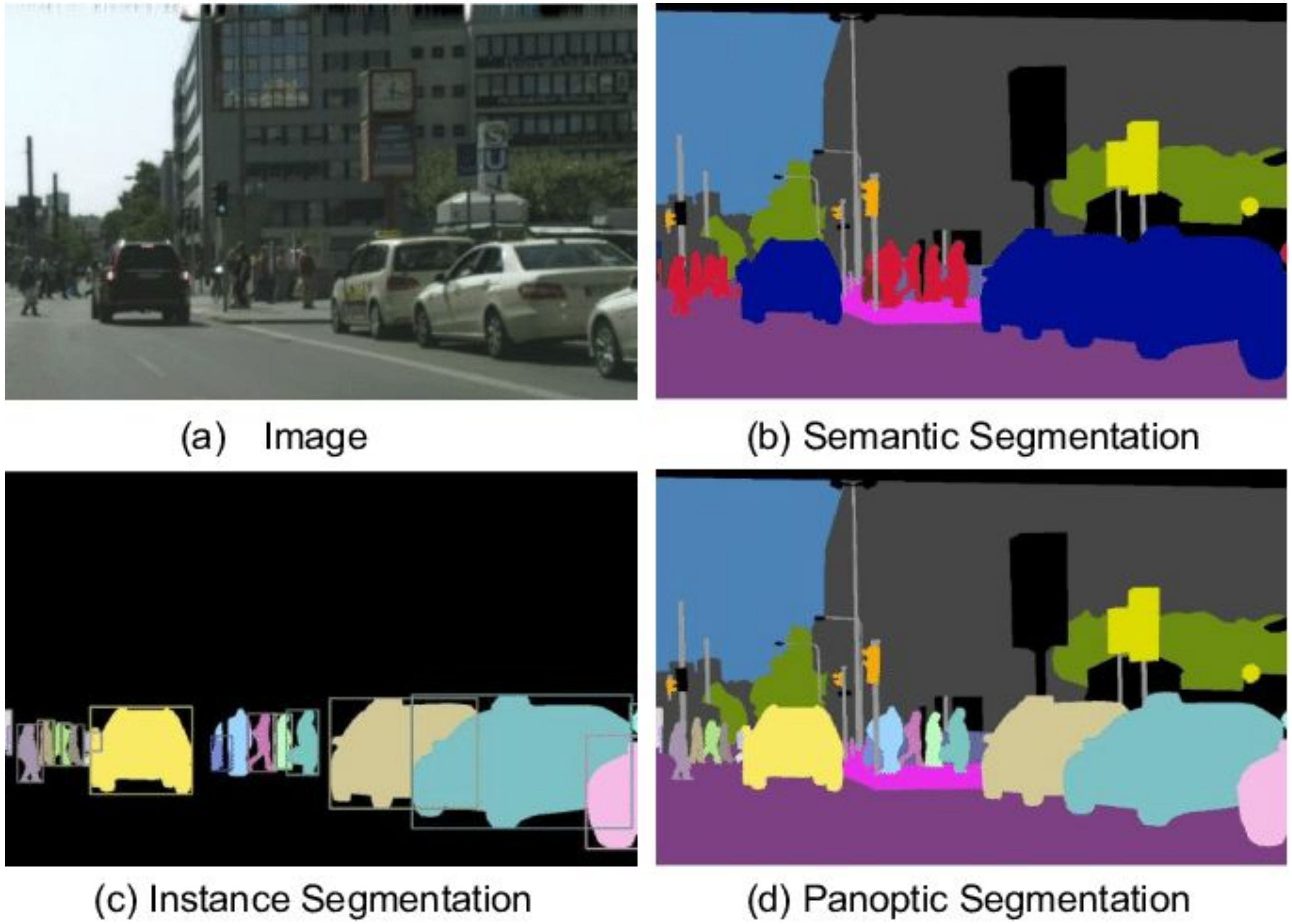


Cancer: 0.85

Healthy: 0.15

Rakhlin, A., Shvets, A., Iglovikov, V., Kalinin, A.A., 2018. Deep convolutional neural networks for breast cancer histology image analysis. ICIAR 2018.

Image segmentation–pixel level annotation

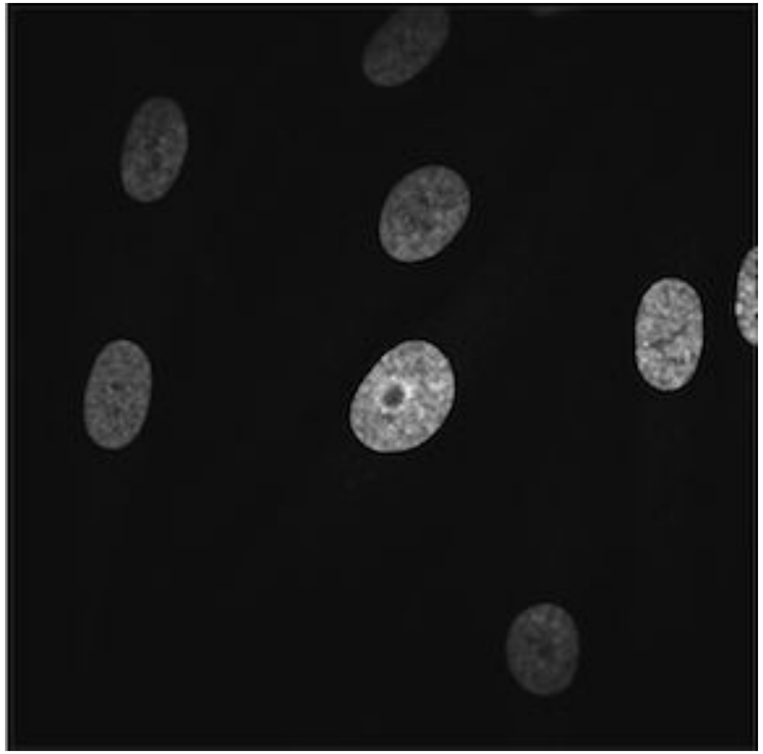


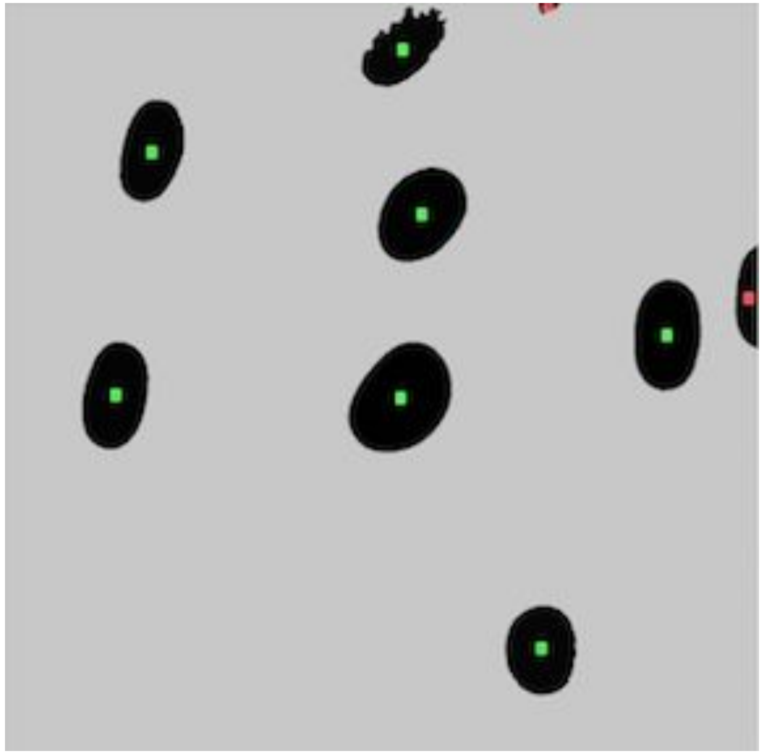
Chen, C., Wang, B., Lu, C.X., Trigoni, N. and Markham, A., 2020. A survey on deep learning for localization and mapping: Towards the age of spatial machine intelligence. arXiv:2006.12567.

@alxndrkalinin, 2021

Do we need deep learning for segmentation?

- Requires dense annotation
- Amount of labelled data still matters
- Computationally expensive (especially in 3D)





Kalinin, A.A., et al., 2018. 3D Cell Nuclear Morphology: Microscopy Imaging Dataset and Voxel-Based Morphometry Classification Results. CVPRW 2018.

Recent case in point (ECCV 2020)

Histogram thresholding methods (like Otsu's) are still popular for binary segmentation of biomedical images.

Recent developments outperform DNNs

Barron, J.T., 2020. A Generalization of Otsu's Method and Minimum Error Thresholding. ECCV 2020.

Algorithm	ν	τ	κ	ω	$F_1 \times 100 \uparrow$	PSNR $\uparrow$	DRD [17] $\downarrow$
GHT (MET Case)	-	-	-	-	60.40 ± 20.65	11.21 ± 3.50	45.32 ± 41.35
Kefali *et al.* [20,24]					76.10 ± 13.81	15.35 ± 3.19	9.16 ± 4.87
Raza [20]					76.28 ± 9.71	14.21 ± 2.21	15.14 ± 9.42
GHT (wprctile Case)	-	-	10^{60}	$2^{-3.75}$	76.77 ± 14.50	15.44 ± 3.40	12.91 ± 17.19
Sauvola [20,25]					82.52 ± 9.65	16.42 ± 2.87	7.49 ± 3.97
Khan & Mollah [20]					84.32 ± 6.81	16.59 ± 2.99	6.94 ± 3.33
Tensmeyer & Martinez [16,27,28]					85.57 ± 6.75	17.50 ± 3.43	5.00 ± 2.60
de Almeida & de Mello [20]					86.24 ± 5.79	17.52 ± 3.42	5.25 ± 2.88
Otsu's Method [19,20]					86.61 ± 7.26	17.80 ± 4.51	5.56 ± 4.44
GHT (No wprctile)	$2^{50.5}$	$2^{0.125}$	-	-	87.16 ± 6.32	17.97 ± 4.00	5.04 ± 3.17
GHT (Otsu Case)	10^{60}	10^{-15}	-	-	87.19 ± 6.28	17.97 ± 4.01	5.04 ± 3.16
Otsu's Method (Our Impl.) [19]					87.19 ± 6.28	17.97 ± 4.01	5.04 ± 3.16
Nafchi *et al.* - 1 [20,30]					87.60 ± 4.85	17.86 ± 3.51	4.51 ± 1.62
Kligler [8,11,13]					87.61 ± 6.99	18.11 ± 4.27	5.21 ± 5.28
Roe & de Mello [20]					87.97 ± 5.17	18.00 ± 3.68	4.49 ± 2.65
Nafchi *et al.* - 2 [20,30]					88.11 ± 4.63	18.00 ± 3.41	4.38 ± 1.65
Hassaïne *et al.* - 1 [7,20]					88.22 ± 4.80	18.22 ± 3.41	4.01 ± 1.49
Hassaïne *et al.* - 2 [6,20]					88.47 ± 4.45	18.29 ± 3.35	3.93 ± 1.37
Hassaïne *et al.* - 3 [6,7,20]					88.72 ± 4.68	18.45 ± 3.41	3.86 ± 1.57
GHT	$2^{29.5}$	$2^{3.125}$	$2^{22.25}$	$2^{-3.25}$	88.77 ± 4.99	18.55 ± 3.46	3.99 ± 1.77
Oracle Global Threshold					90.69 ± 3.92	19.17 ± 3.29	3.57 ± 1.84

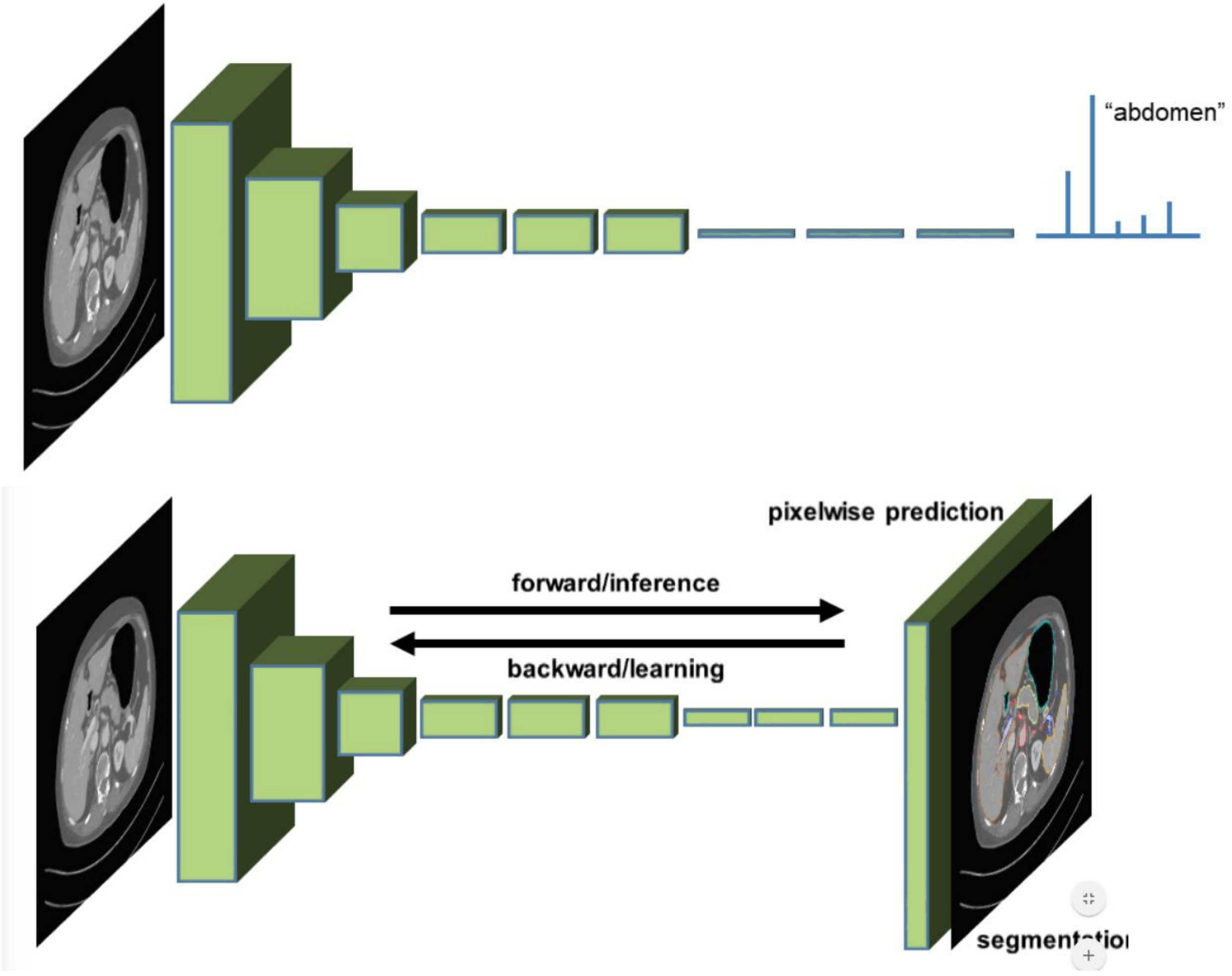
Deep networks for segmentation: FCN

Convert classification network into segmentation:

- removed dense (FC) layers
- added upsampling

- reduced # of parameters
- indifferent to input size

Long, J., Shelhamer, E. and Darrell, T., 2015. Fully convolutional networks for semantic segmentation. CVPR 2015.

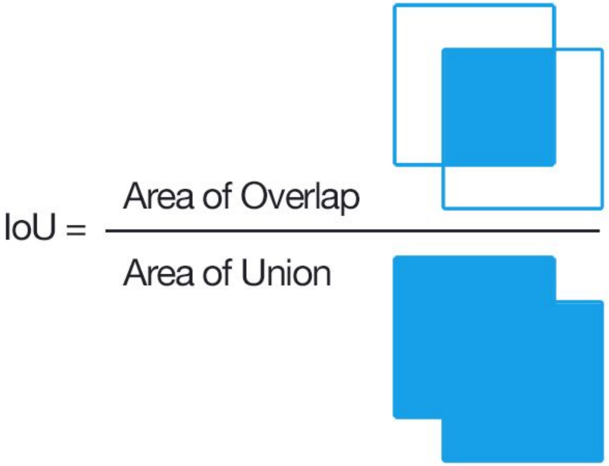


Deep networks for segmentation: Loss functions

Intersection over Union

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

$$J = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{y_i \hat{y}_i}{y_i + \hat{y}_i - y_i \hat{y}_i} \right)$$



PyImageSearch, 2019. https://www.pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/

Cross-Entropy

$$H = -\frac{1}{n} \sum_{i=1}^{n} [y \log(\hat{y}) + (1-y) \log(1-\hat{y})]$$

Combined losses, e.g.

$$L = H - \log J$$

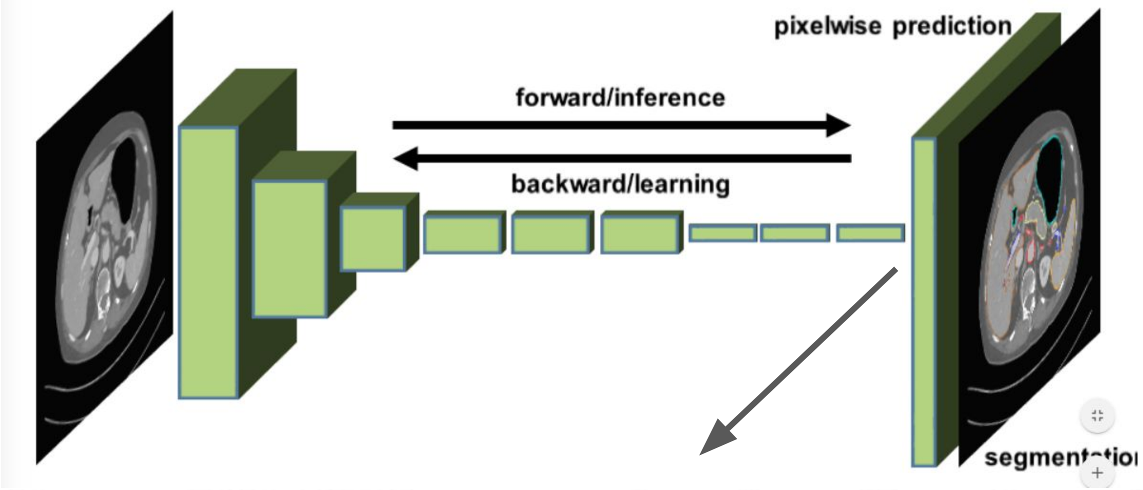
Iglovikov, V., et al., 2017. Satellite imagery feature detection using deep convolutional neural network: A kaggle competition. arXiv preprint arXiv:1706.06169.

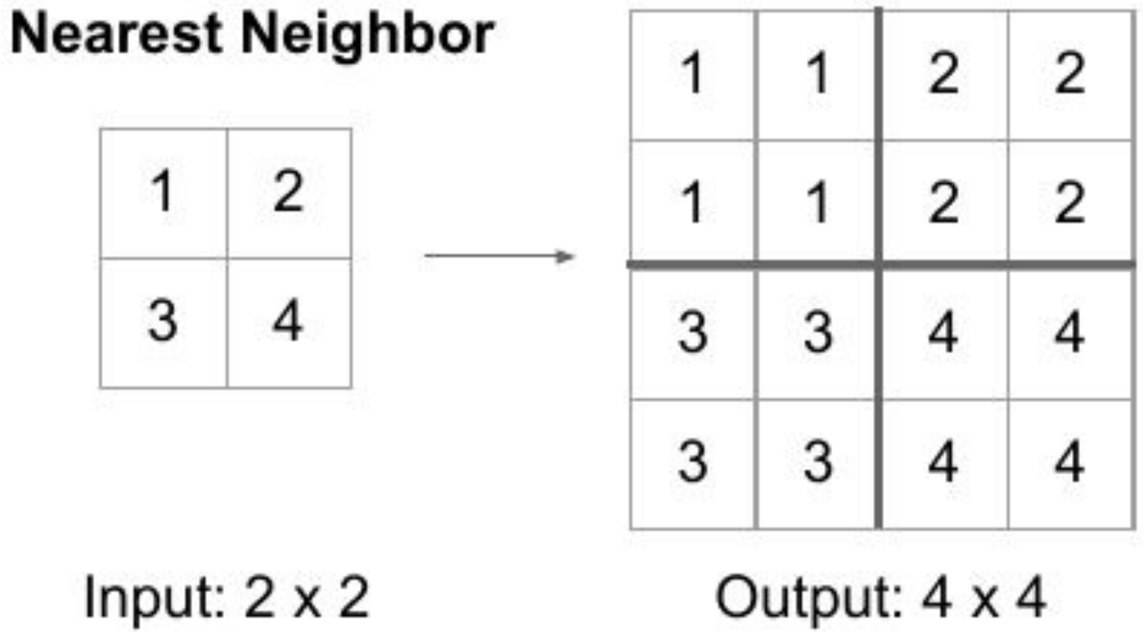
Deep networks for segmentation: FCN

Convert classification network into segmentation:

- removed dense (FC) layers
- added **upsampling**

- reduced # of parameters
- indifferent to input size





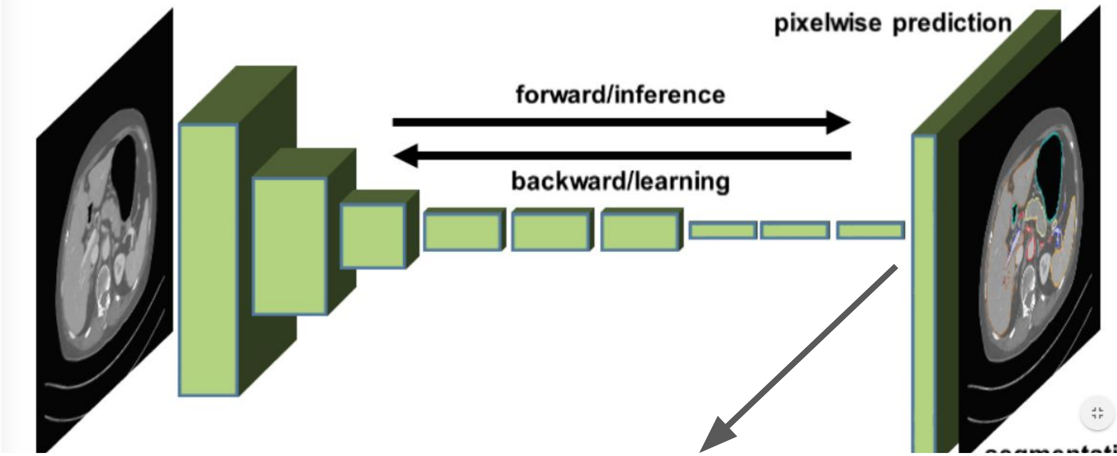
Roth, H.R., Shen, C., Oda, H., Oda, M., et al., 2018. Deep learning and its application to medical image segmentation. Medical Imaging Technology.

Deep networks for segmentation: FCN

Convert classification network into segmentation:

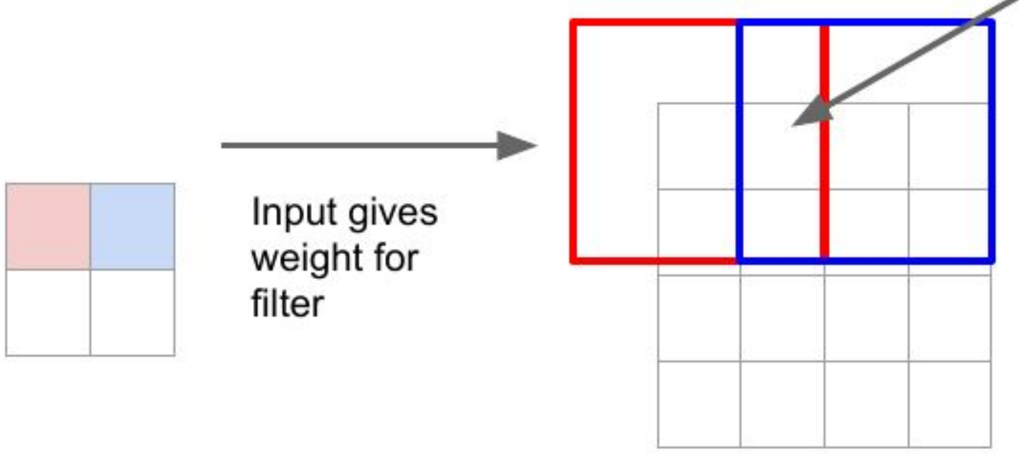
- removed dense (FC) layers
- added **upsampling**

- reduced # of parameters
- indifferent to input size



3 x 3 **transpose** convolution, stride 2 pad 1

Sum where output overlaps



Input gives weight for filter

Filter moves 2 pixels in the output for every one pixel in the input

Stride gives ratio between movement in output and input

Input: 2 x 2

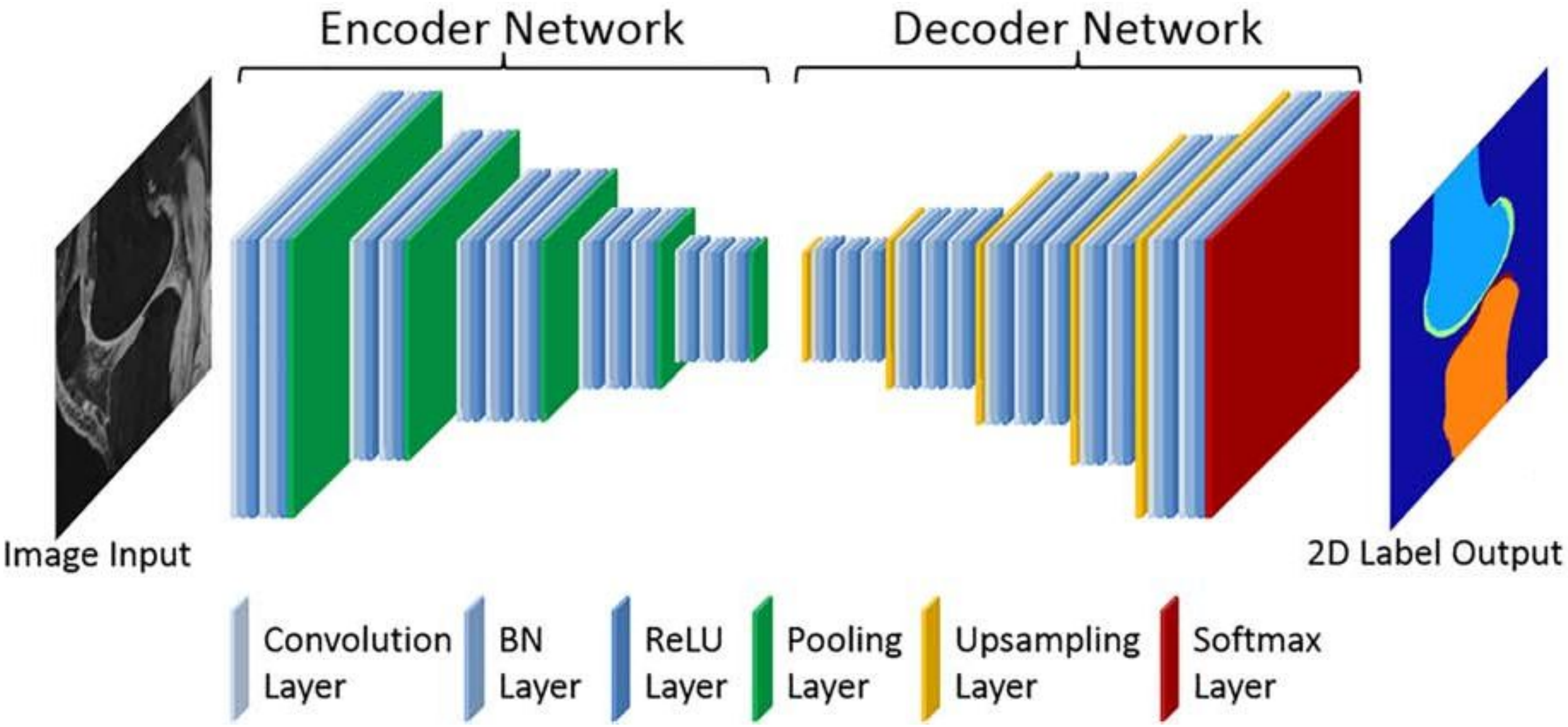
Output: 4 x 4

Autoencoder: Downsampling and Upsampling. 2019. Technical Fridays. https://kharshit.github.io/blog/2019/02/15/autoencoder-downsampling-and-upsampling

Deep networks for segmentation: SegNet

Improves the results over FCN by:

- progressive upsampling
- added convolutions to improve upsampling results

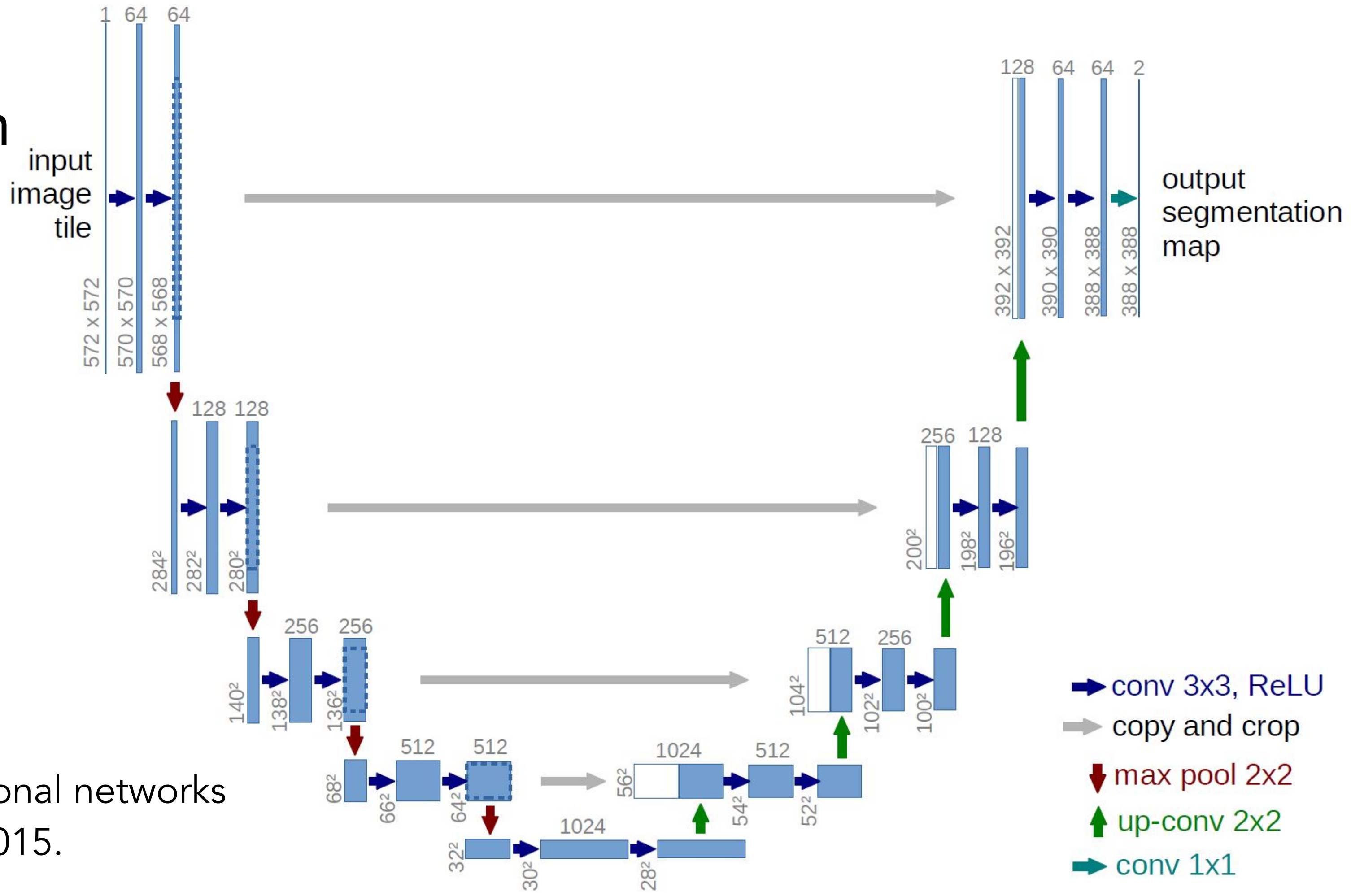


Liu, F., Zhou, Z., Jang, H., Samsonov, A., Zhao, G., Kijowski, R., 2018. Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging. Magnetic resonance in medicine.

U-Net: *a golden hammer*

Most commonly used DNN for biomedical image segmentation (>27,000 citations on GS)

- easy to implement and train
- requires few examples
- a very strong baseline
- translates well to 3D



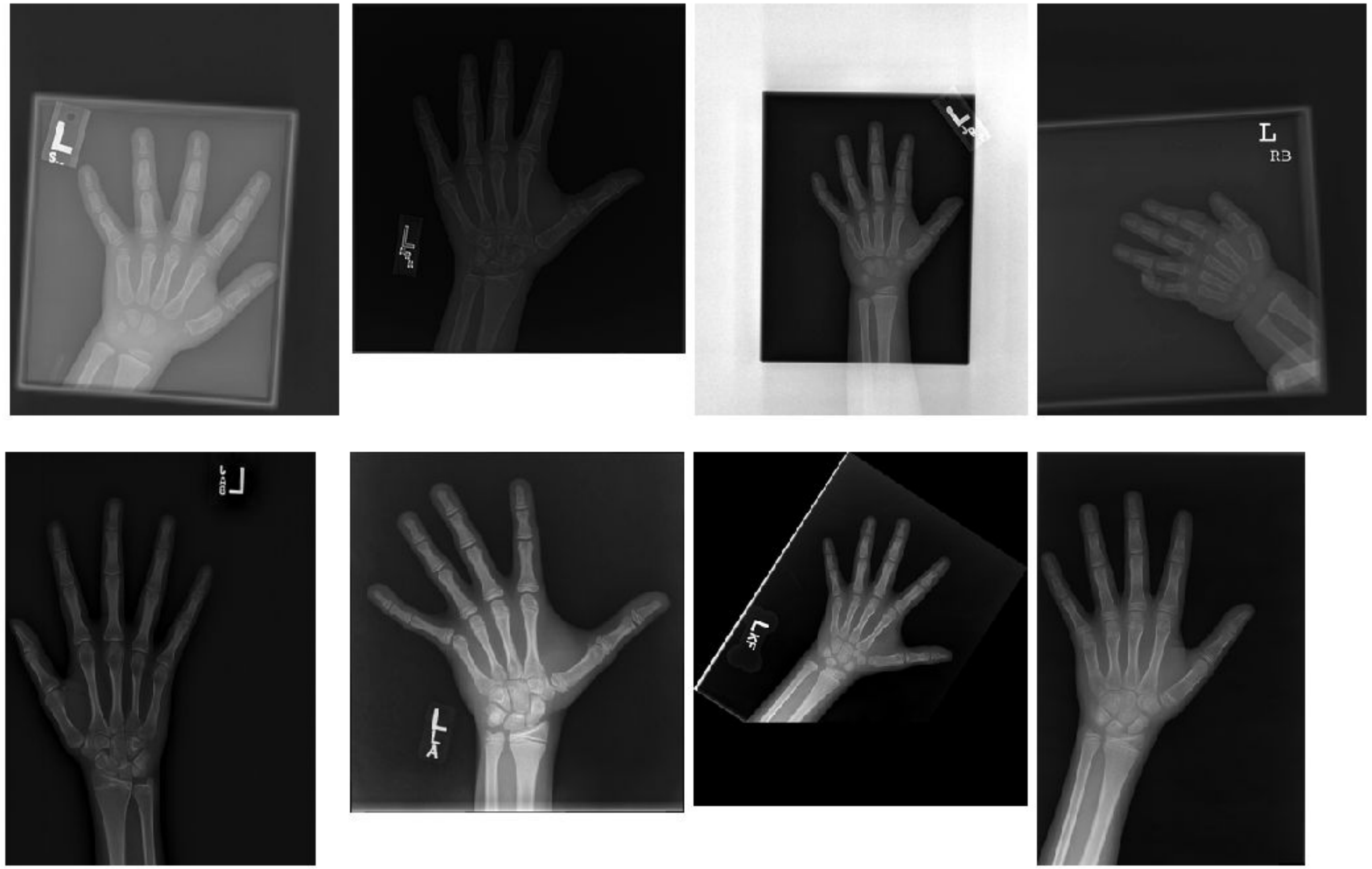
Ronneberger, O., et al., 2015. U-net: Convolutional networks for biomedical image segmentation. MICCAI 2015.

U-Net example

Task: age regression from hand x-ray images

- 12,000 noisy images

- no segmentation masks



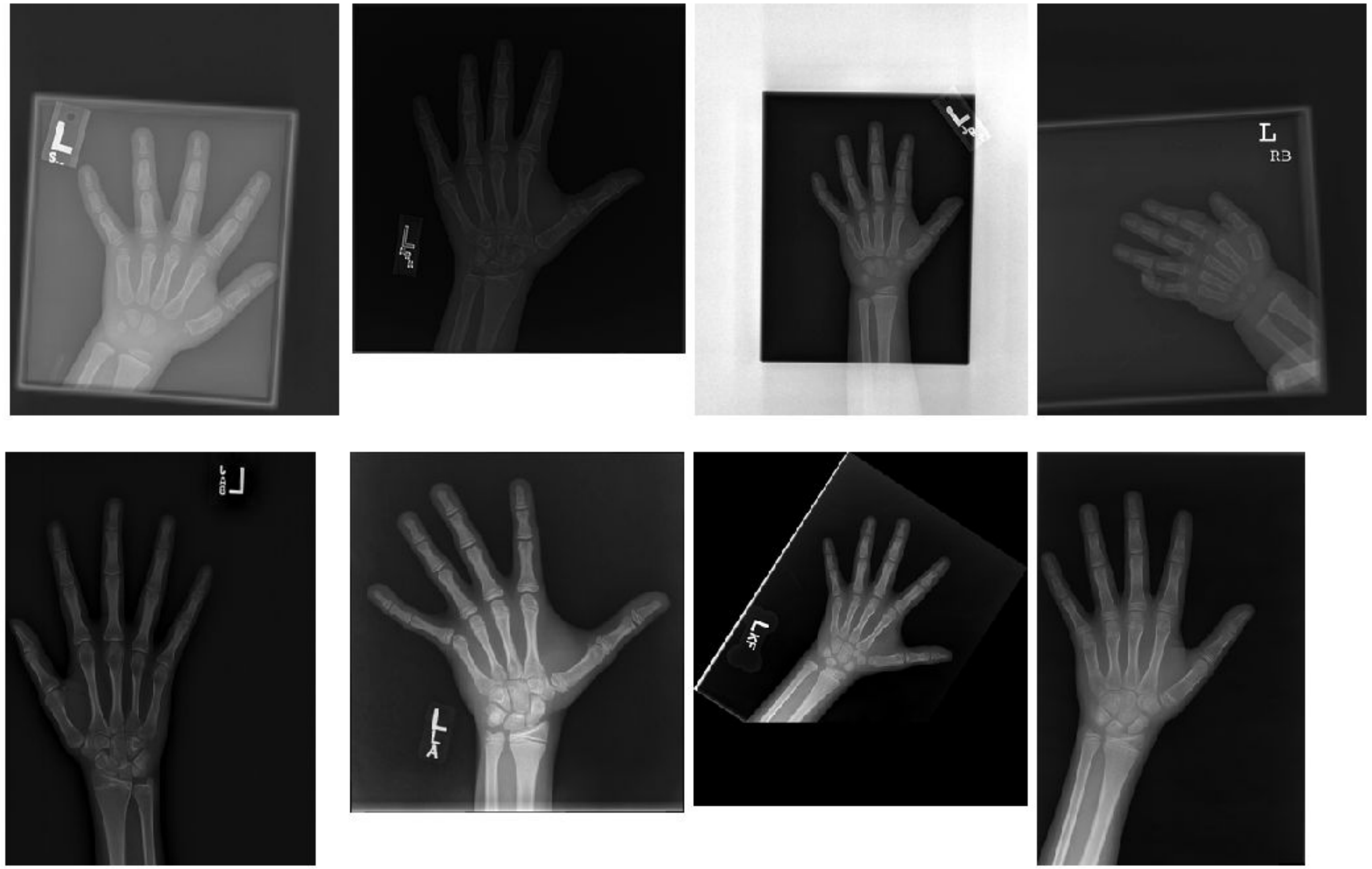
Iglovikov, V.I., et al., 2018. Paediatric bone age assessment using deep convolutional neural networks. MICCAIW 2018.

U-Net example

Task: age regression from hand x-ray images

- 12,000 noisy images
- no segmentation masks

Can we use U-Net for pre-processing to clean up the data?

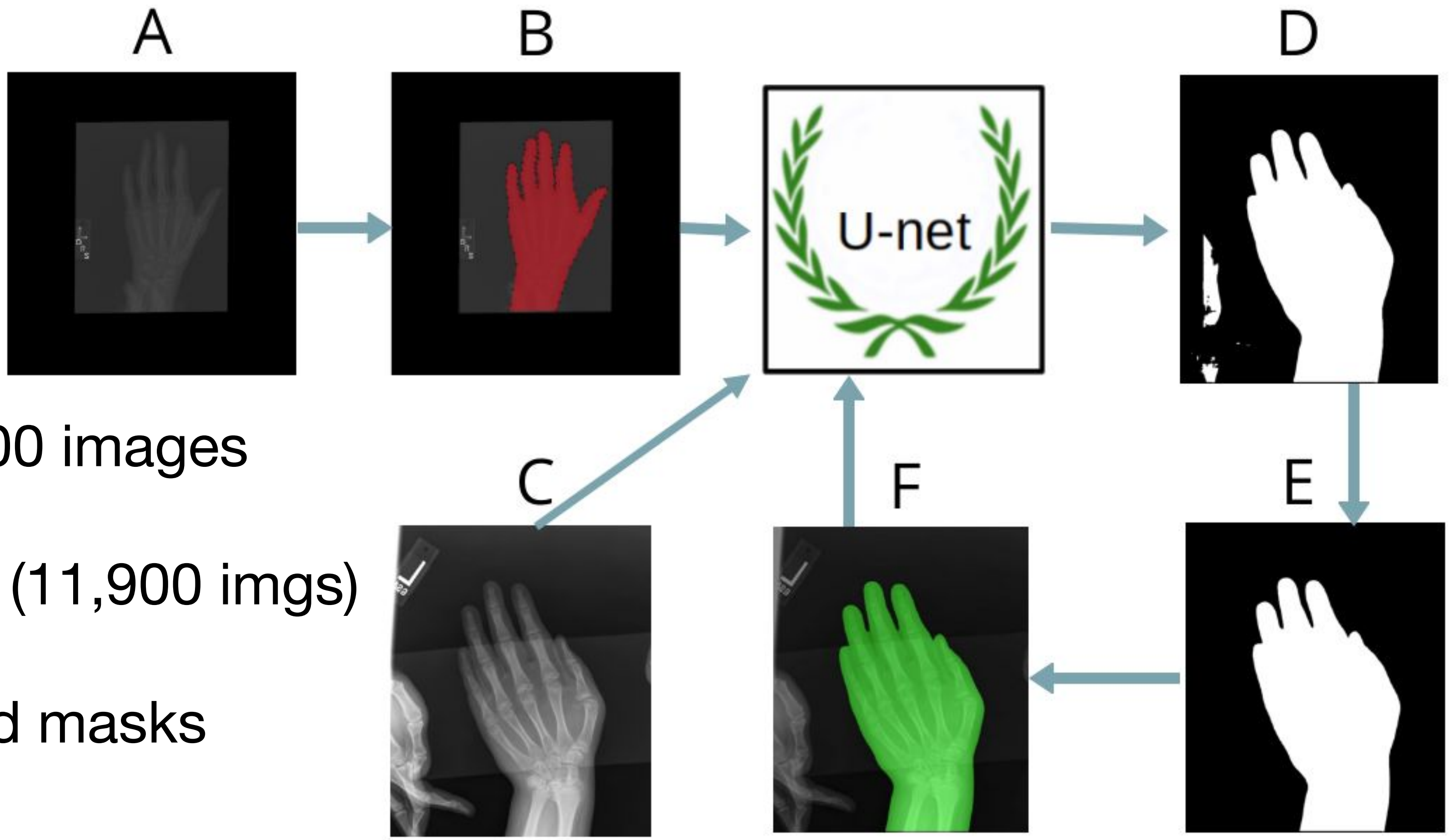


Iglovikov, V.I., et al., 2018. Paediatric bone age assessment using deep convolutional neural networks. MICCAIW 2018.

U-Net example

Approach:

- manually label 100 images
- train U-Net on these 100 images
- predict the rest of train (11,900 imgs)
- take 100 best predicted masks
- re-train U-Net on 200 images

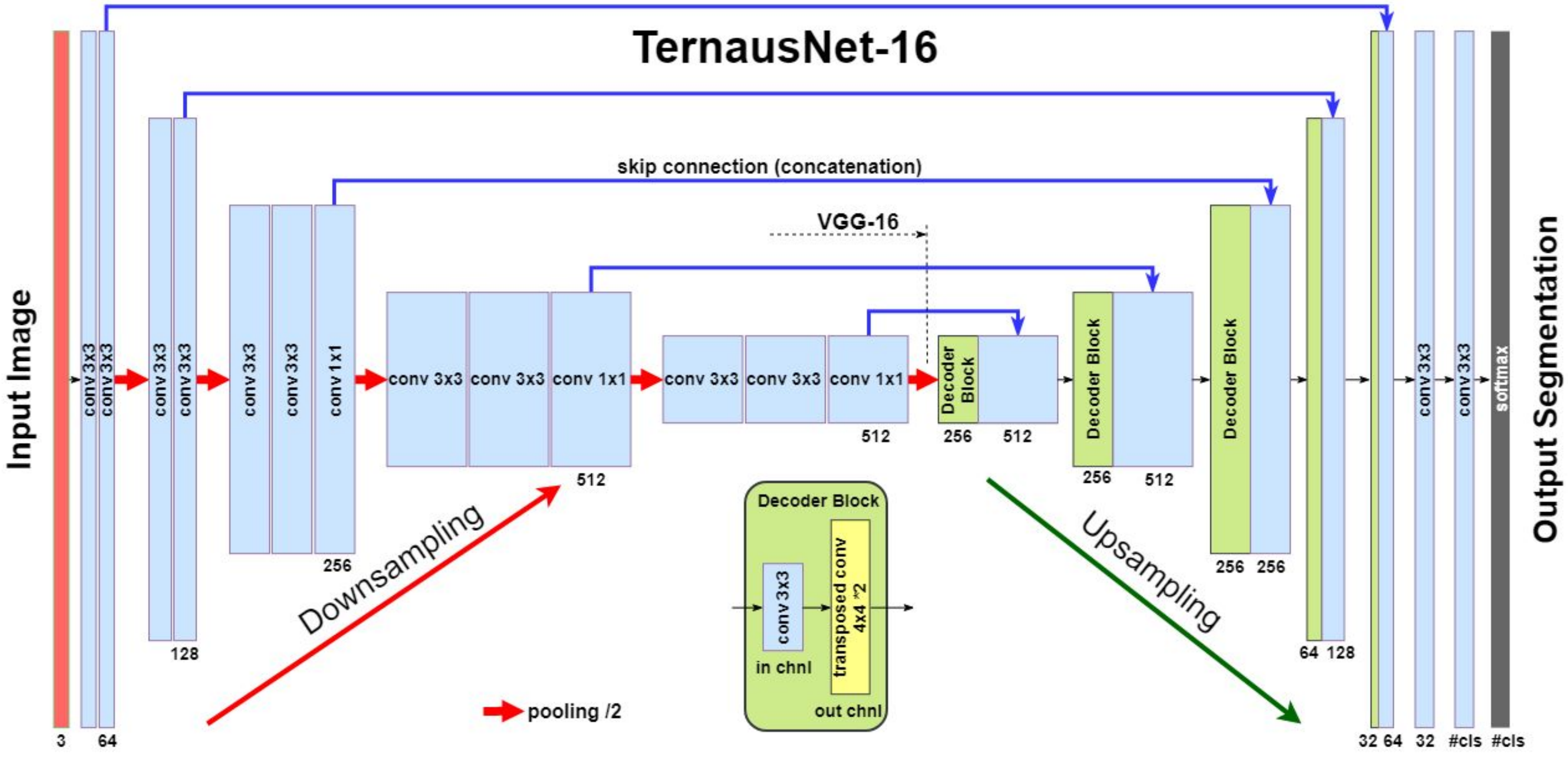


Iglovikov, V.I., et al., 2018. Paediatric bone age assessment using deep convolutional neural networks. MICCAIW 2018.

U-Net with pre-training: TernausNet

Improvement by:

- replacing encoder with pre-trained CNN
- transposed convolutions



Iglovikov, V., Shvets, A., 2018. Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. arXiv preprint arXiv:1801.05746.

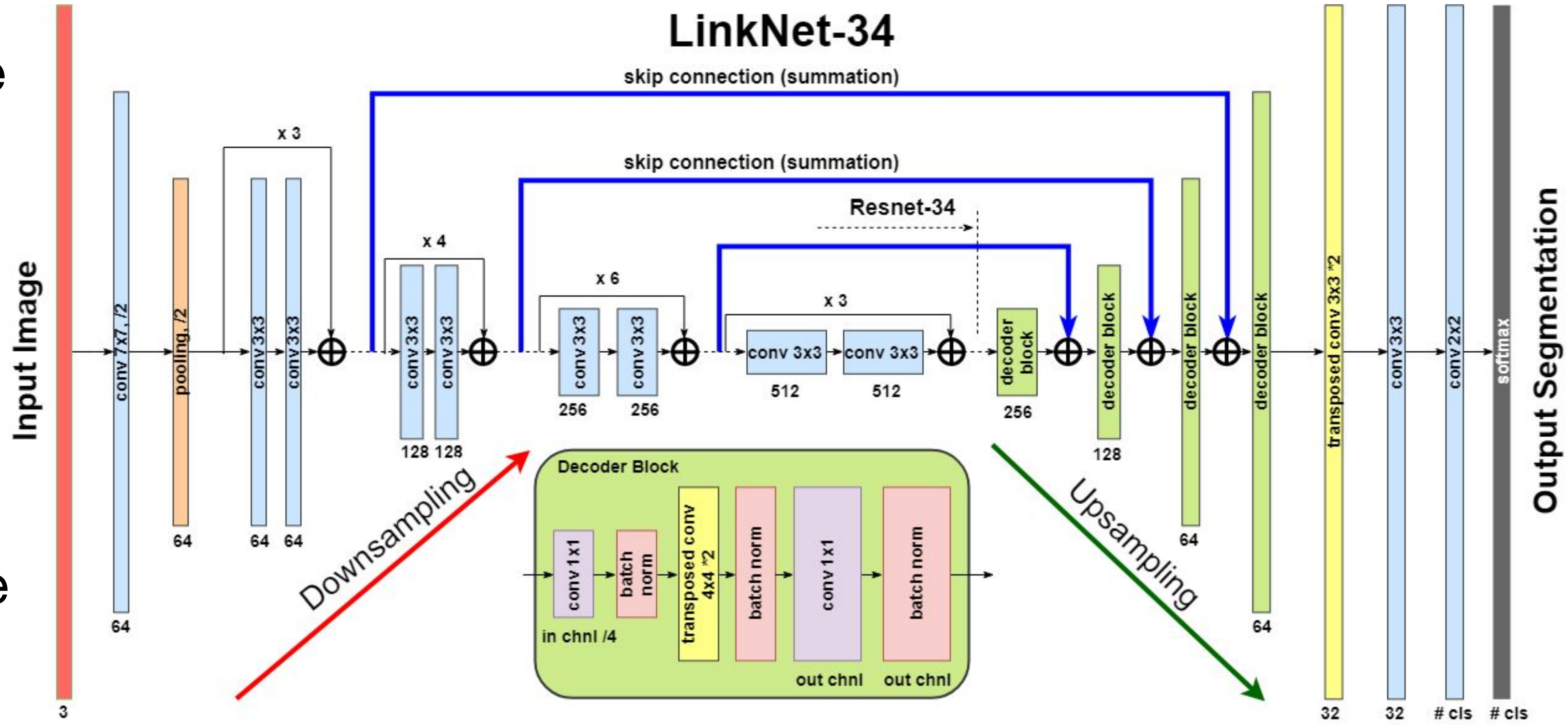
Shvets, A.A., Rakhlin, A., Kalinin, A.A., Iglovikov, V.I., 2018. Automatic instrument segmentation in robot-assisted surgery using deep learning. IEEE ICMLA 2018.

LinkNet-34

Chaurasia, A., & Culurciello, E., 2017. Linknet: Exploiting encoder representations for efficient semantic segmentation. VCIP 2017

LinkNet architecture with pre-trained ResNet-34 encoder

- reduces the number of parameters
- faster and more accurate

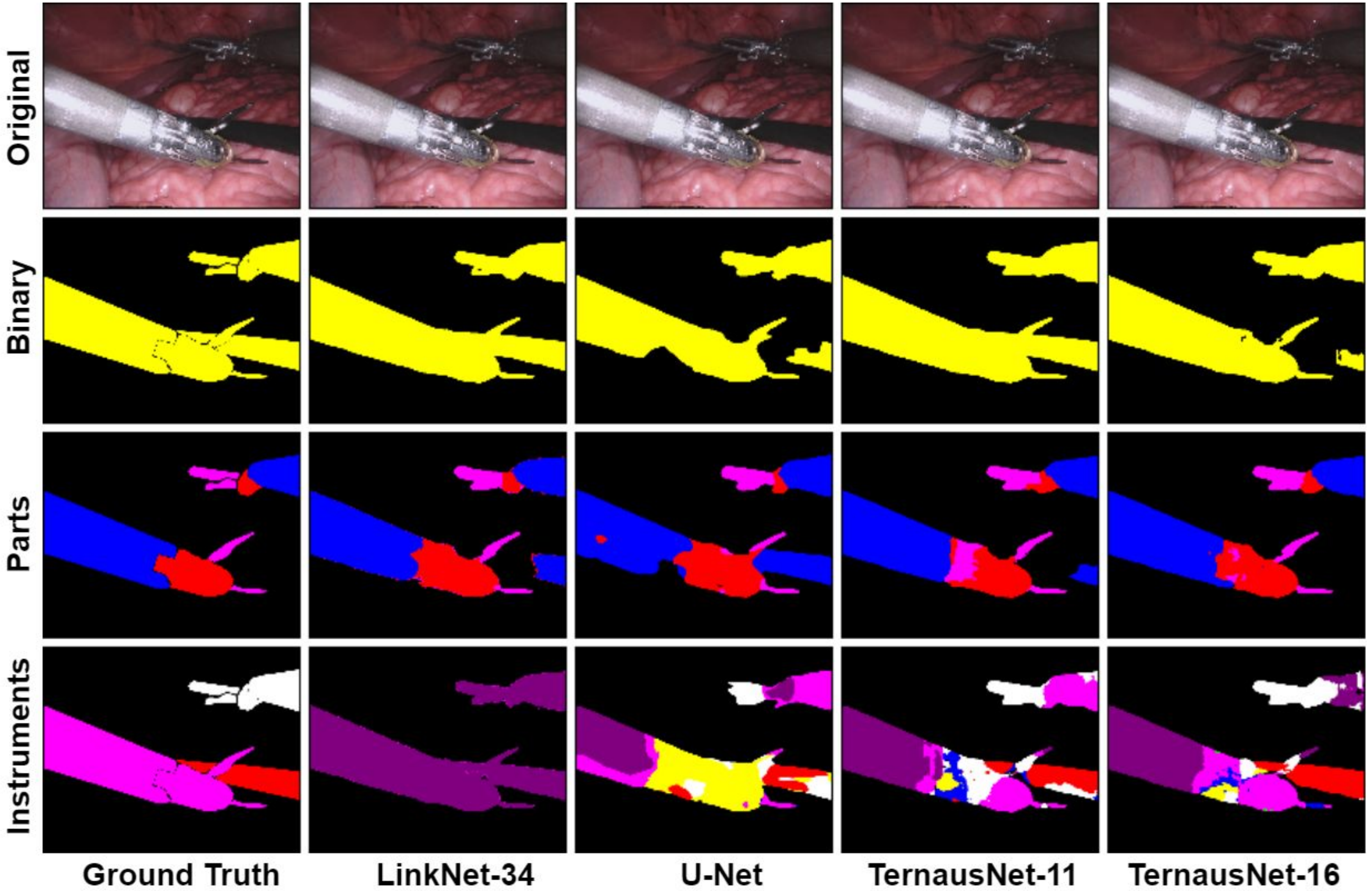


Shvets, A.A., et al, 2018. Automatic instrument segmentation in robot-assisted surgery using deep learning. IEEE ICMLA 2018.

@alxndrkalinin, 2021

Comparison of different encoders

- U-net provides a strong baseline
- best model depends on a task
- just “deeper” is not always better

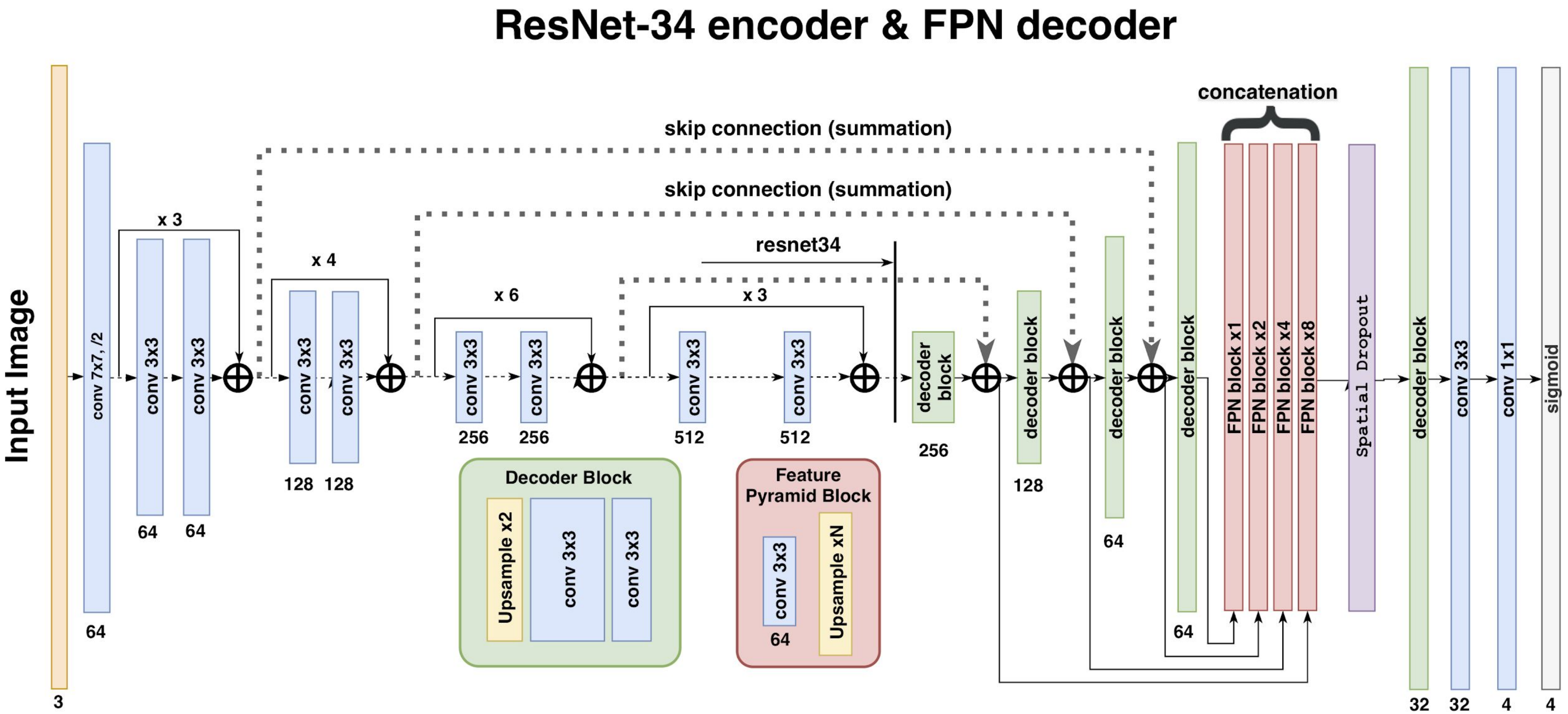


Kalinin, A.A., et al., 2020. Medical Image Segmentation Using Deep Neural Networks with Pre-trained Encoders. Deep Learning Applications.

LinkNet-34-FPN

Feature Pyramid Network

- feature reconstruction in the decoder at 4 different scales



Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S., 2017. Feature pyramid networks for object detection. CVPR 2017.

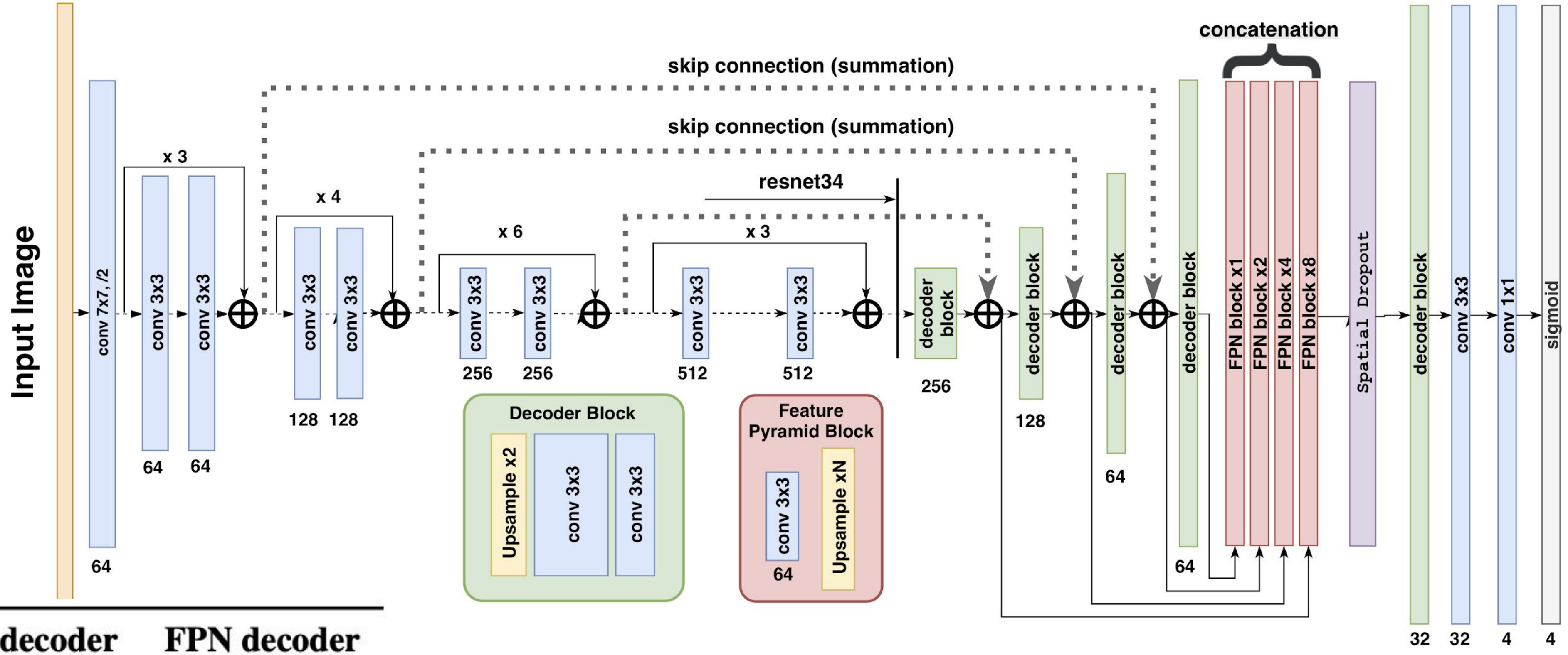
Rakhlin, A., et al, 2019. Breast Tumor Cellularity Assessment using Deep Neural Networks. ICCVW 2019.

LinkNet-34-FPN

Feature Pyramid Network

- feature reconstruction in the decoder at 4 different scales

ResNet-34 encoder & FPN decoder

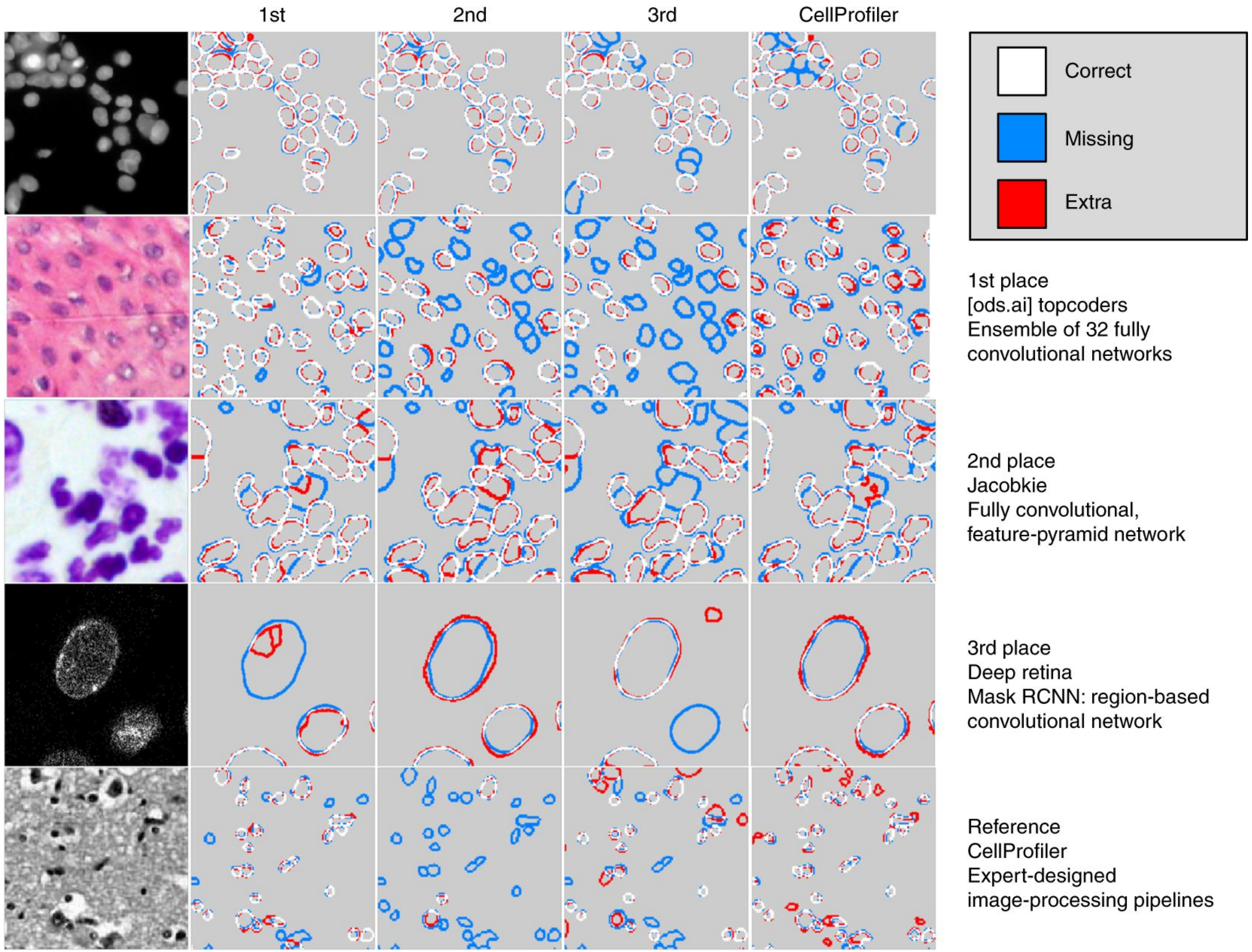


Initialization	Standard decoder	FPN decoder
Random	0.35	0.47
ImageNet	0.50	**0.53**

Rakhlin, A., et al, 2019. Breast Tumor Cellularity Assessment using Deep Neural Networks. ICCVW 2019.

Nucleus segmentation @ 2018 Data Science Bowl

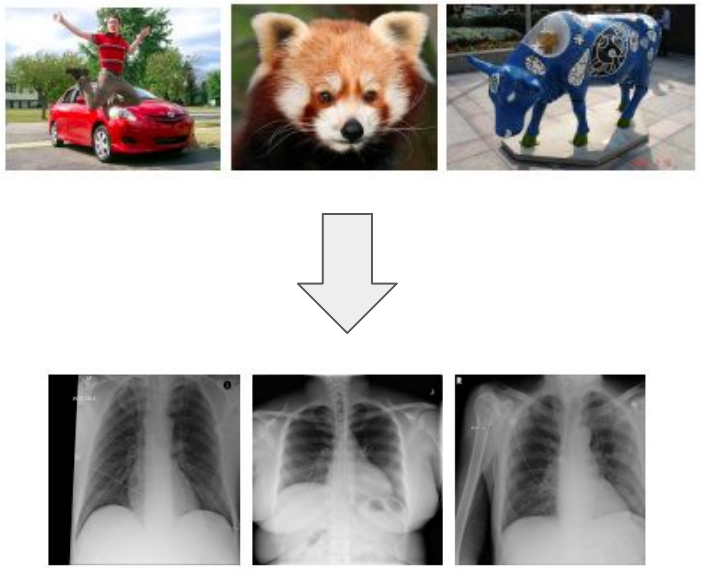
- U-Net-like FPN networks with ImageNet-pretrained Resnet-(34, 50, 101, 152) encoders
- additional targets (touching)
- heavy augmentations
- postprocessing
- even single models wins!



Caicedo, J. C., et al., 2019. Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. Nature methods.

Does transfer learning always work?

Medical images differ from natural domain, so why ImageNet pretraining works?



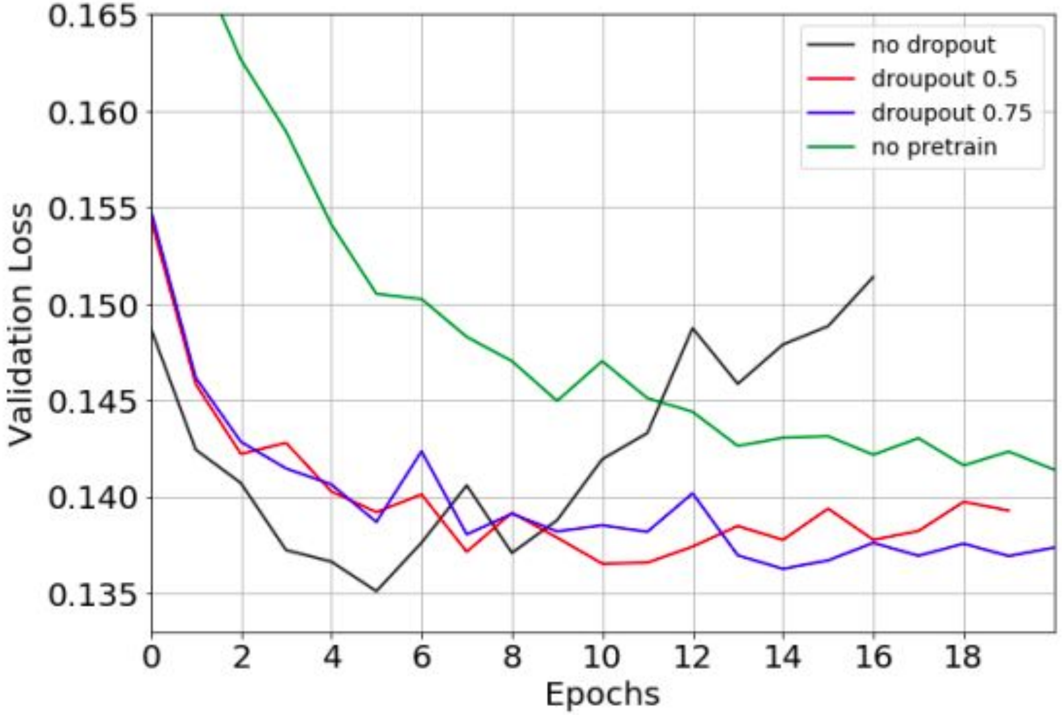
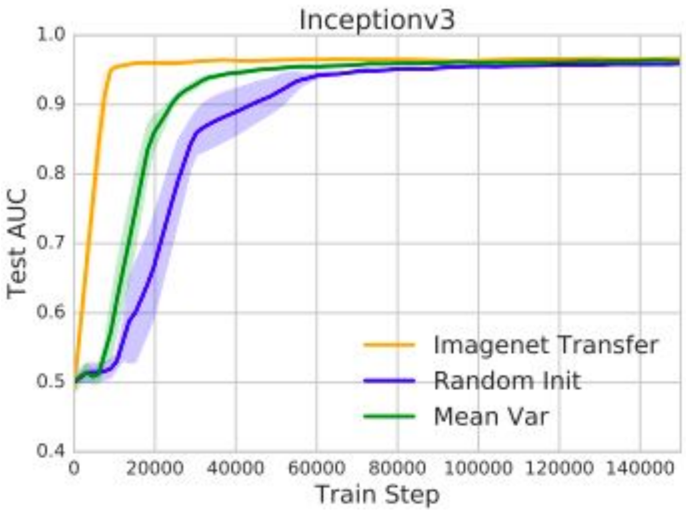
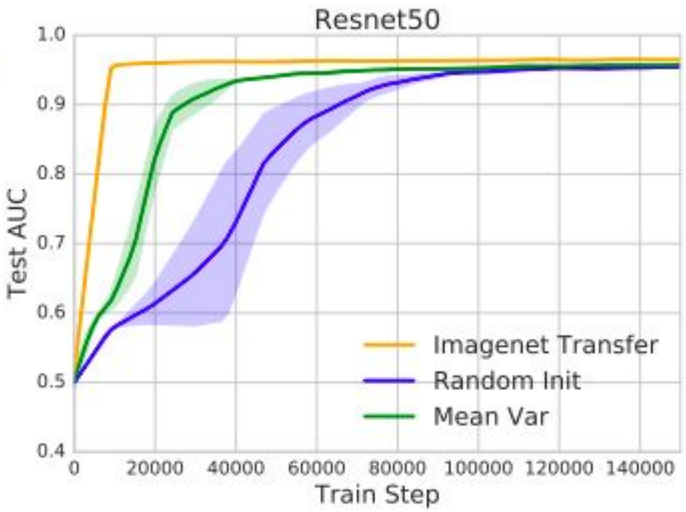


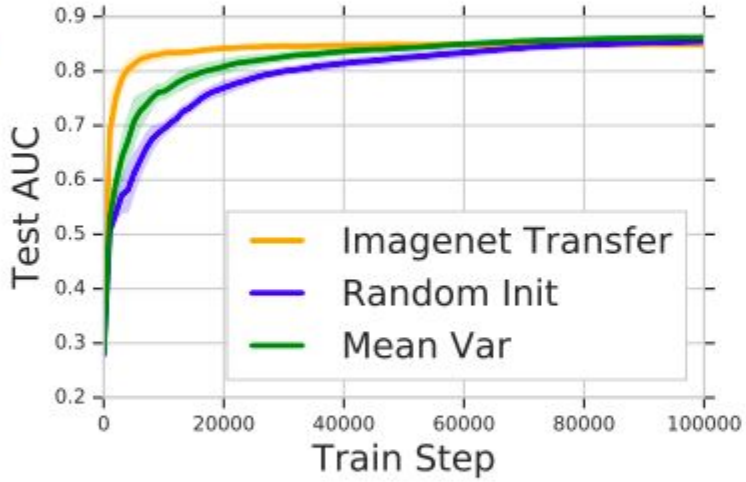
Figure 4. Evolution of the validation loss during training for different versions of RetinaNet with SE-ResNext-101 encoders.

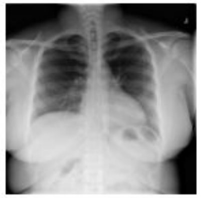
Gabruseva, T., et al, 2020. Deep Learning for Automatic Pneumonia Detection. CVPRW 2020.

ImageNet pre-training helps even on biomedical images







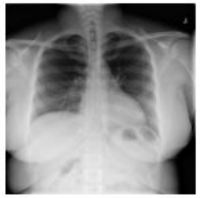



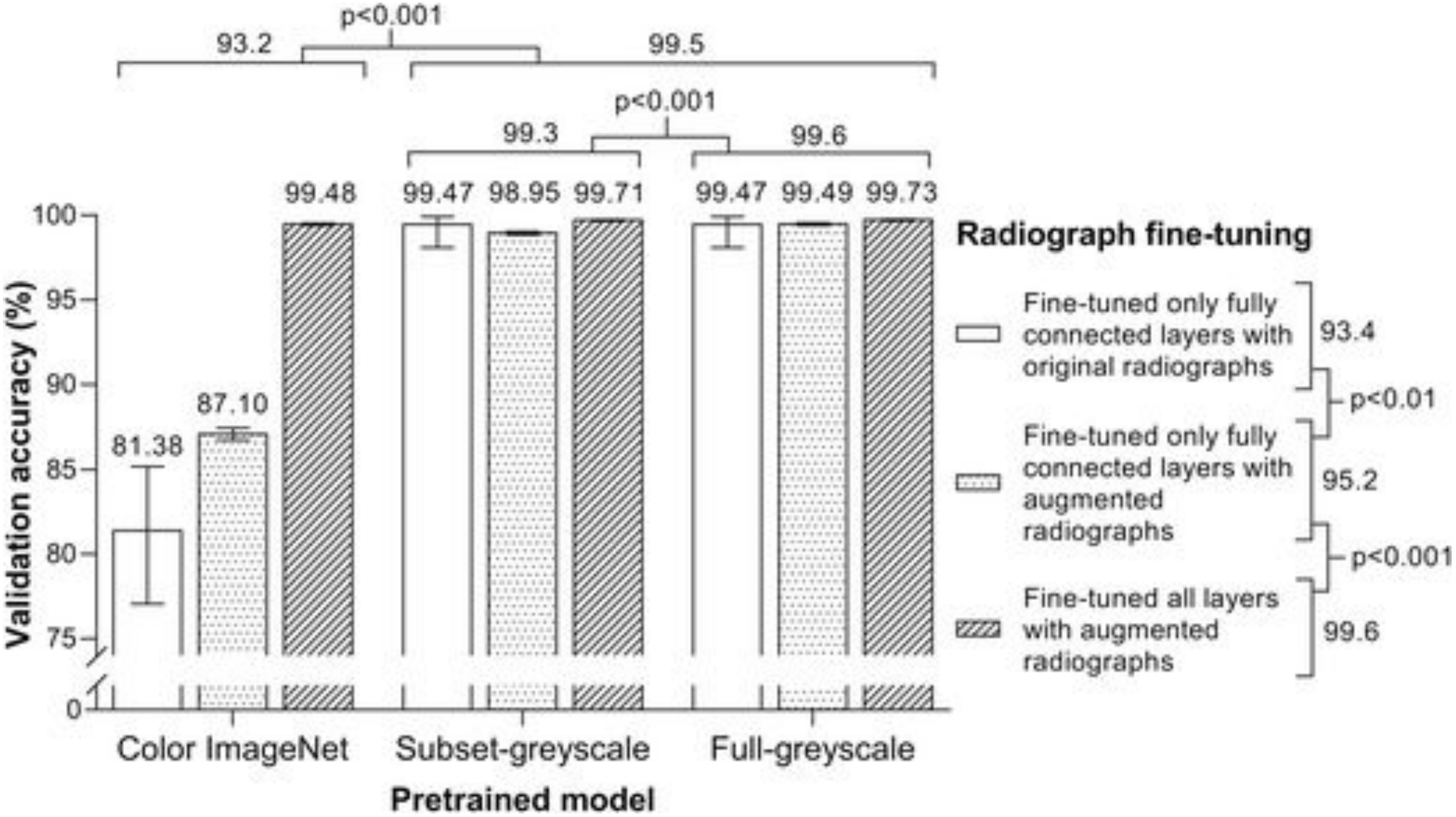

Often, it provides feature-independent convergence speed up due to better weight scaling

Raghu, M., et al, 2019. Transfusion: Understanding transfer learning for medical imaging. NeurIPS 2019.

Can ImageNet pre-training be improved?

- needs fine-tuning
- progressive unfreezing
- start with lower LR

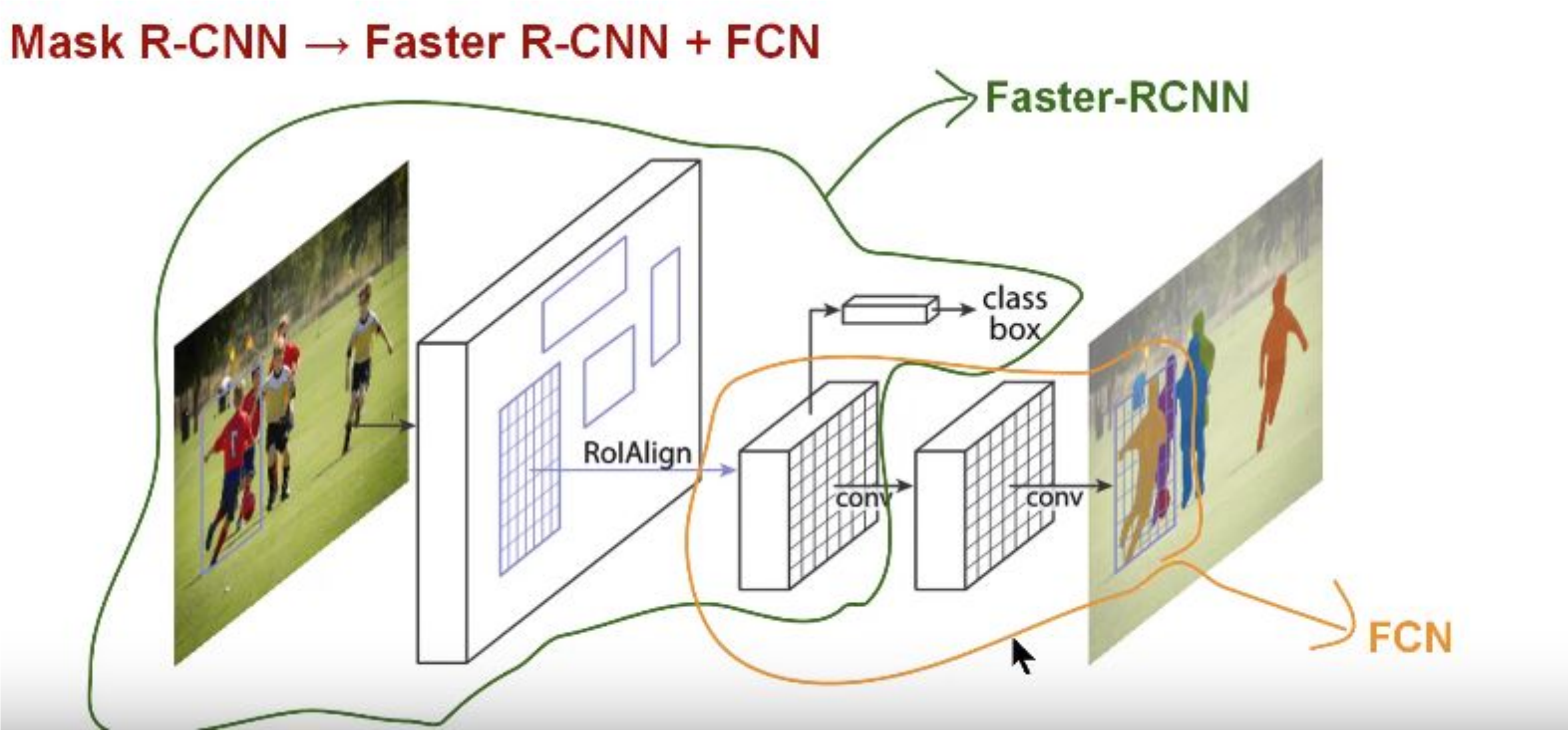
Rajkomar, A., et al, 2017. High-throughput classification of radiographs using deep convolutional neural networks. Journal of digital imaging.

Deep networks for segmentation: Mask R-CNN

Simultaneous detection and segmentation:

- Faster R-CNN detection network
- various encoders
- FCN for segmentation

He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask r-cnn. CVPR 2017.



Priya Dwivedi, 2018. Ultimate Guide: Building a Mask R-CNN Model for Detecting Car Damage (with Python codes) https://www.analyticsvidhya.com/blog/2018/07/building-mask-r-cnn-model-detecting-damage-cars-python/.

Further improvements

New developments are focused on improving some parts of existing models:

- more powerful encoders (ResNeXt, ResNeSt, EfficientNet, Transformers)
- more complex architectures: cascaded, attention-based, etc.
- sophisticated losses & train schedules: cosine, Lovasz-Softmax, etc
- combining with GANs & self-supervised learning

Useful libraries

Detectron2 by FAIR: https://github.com/facebookresearch/Detectron2

segmentation_models
- Keras: https://github.com/qubvel/segmentation_models
- PyTorch: https://github.com/qubvel/segmentation_models.pytorch

MMSegmentation: https://github.com/open-mmlab/mmsegmentation

DeepLab-PyTorch: https://github.com/kazuto1011/deeplab-pytorch

Image Segmentation Keras:
https://github.com/divamgupta/image-segmentation-keras

Improving performance: augmentations

Good image augmentations improve:

- generalization
- convergence
- robustness on out-of-distribution samples
- prediction accuracy

Buslaev, A., et al., 2020. Albumentations: fast and flexible image augmentations. Information.

Thank you!

Presentation contains materials from my joint work with:

- Vladimir Iglovikov
- Alexey Shvets
- Alexander Rakhlin
- Ivo Dinov’s group at UMich
- Brian Athey’s group at UMich

Influenced by: http://dlcourse.ai by Simon Kozlov

Useful discussions:

- ODS.ai – biggest (mostly Russian-speaking) Data Science community: English speakers welcome!

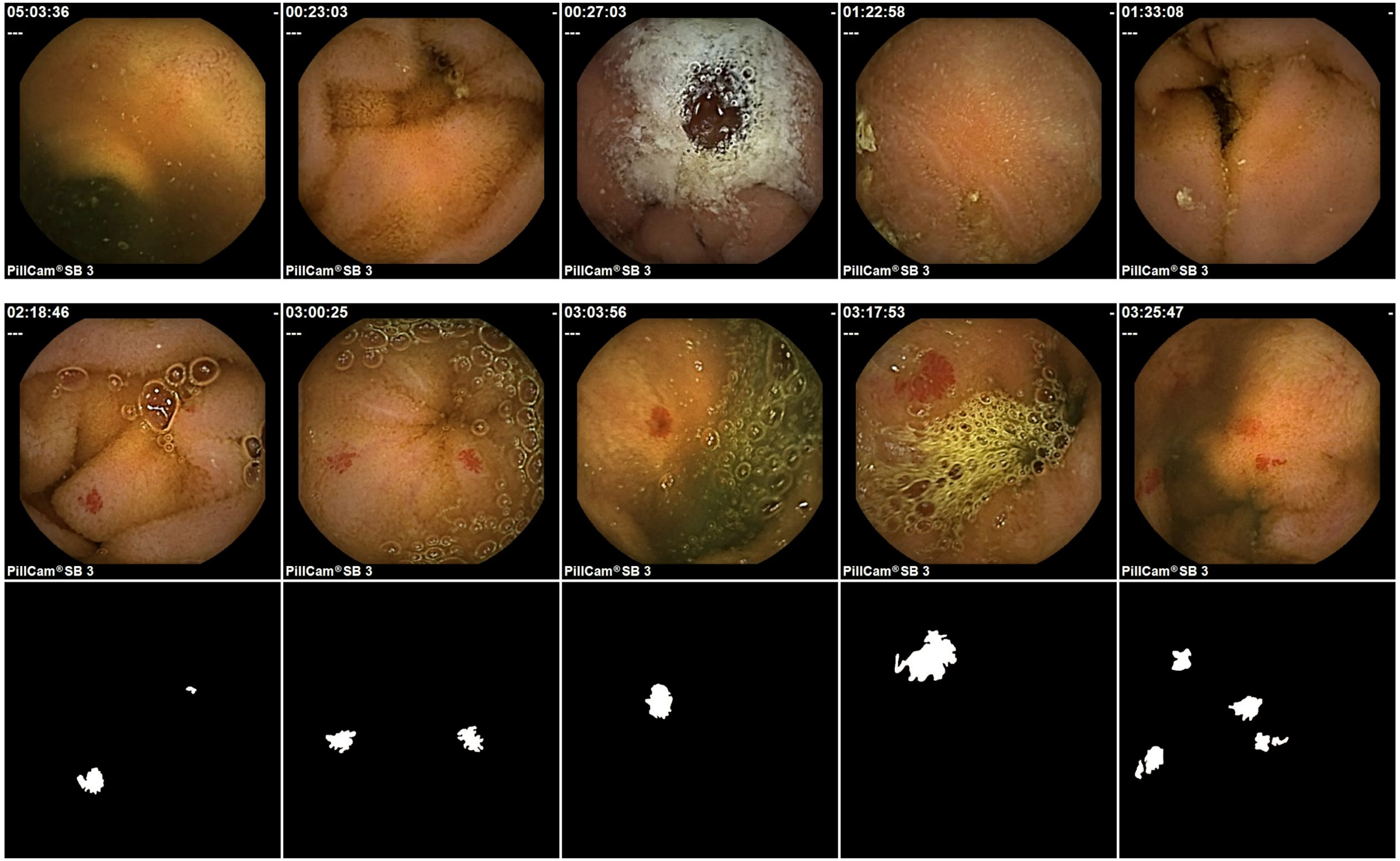
Angiodysplasia segmentation

Angiodysplasia – the most common vascular lesion of the gastrointestinal tract.

Endoscopic videos recorded with wireless capsule.

The goal is to segment lesions.

MICCAI Endoscopic Vision Challenge. 2017. https://endovissub2017-giana.grand-challenge.org/Angiodysplasia-ETISDB/

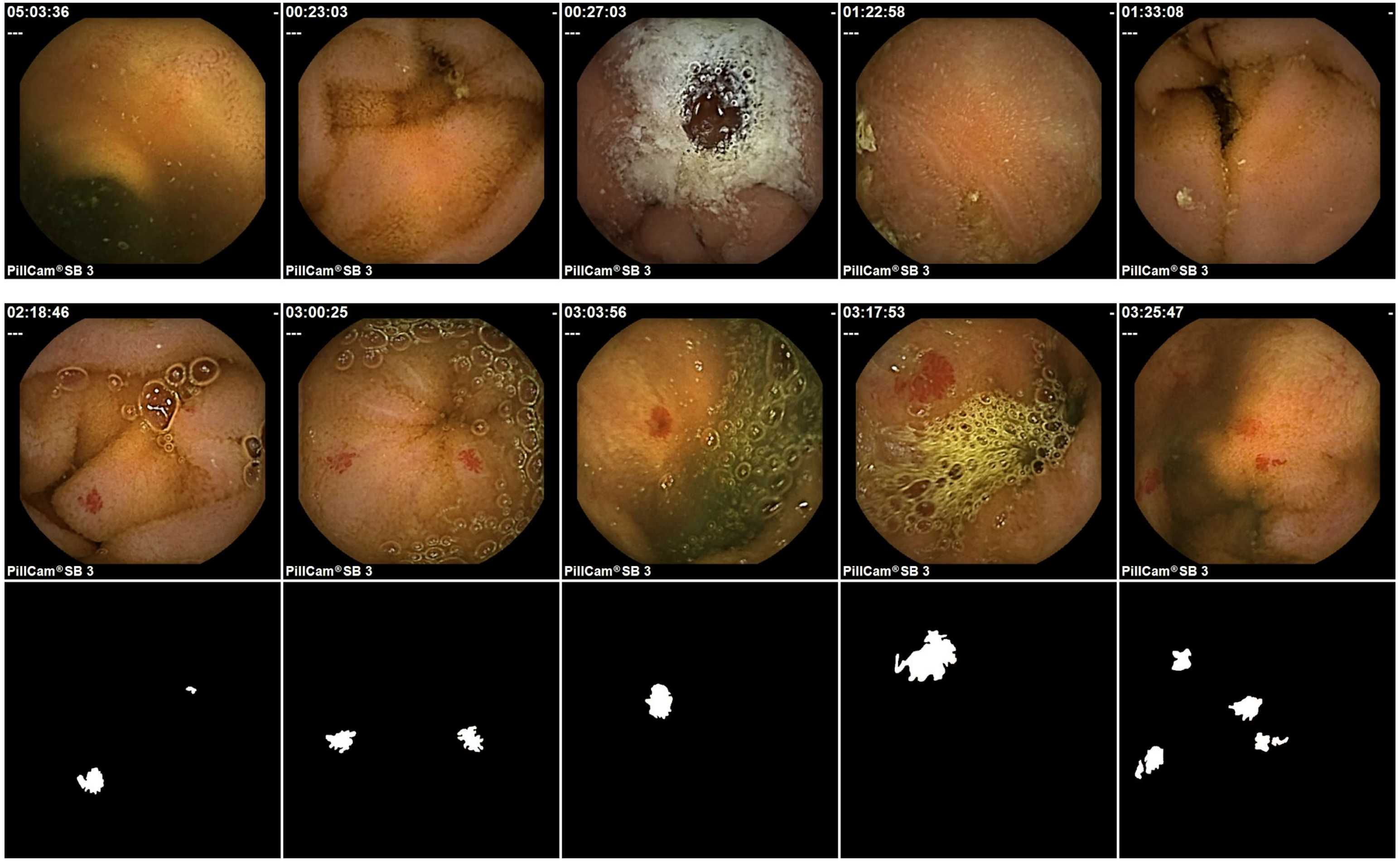


Shvets, A.A., Iglovikov, V.I., Rakhlin, A., Kalinin, A.A., 2018. Angiodysplasia detection and localization using deep convolutional neural networks. ICMLA

Angiodysplasia segmentation

600 images for training
600 for testing

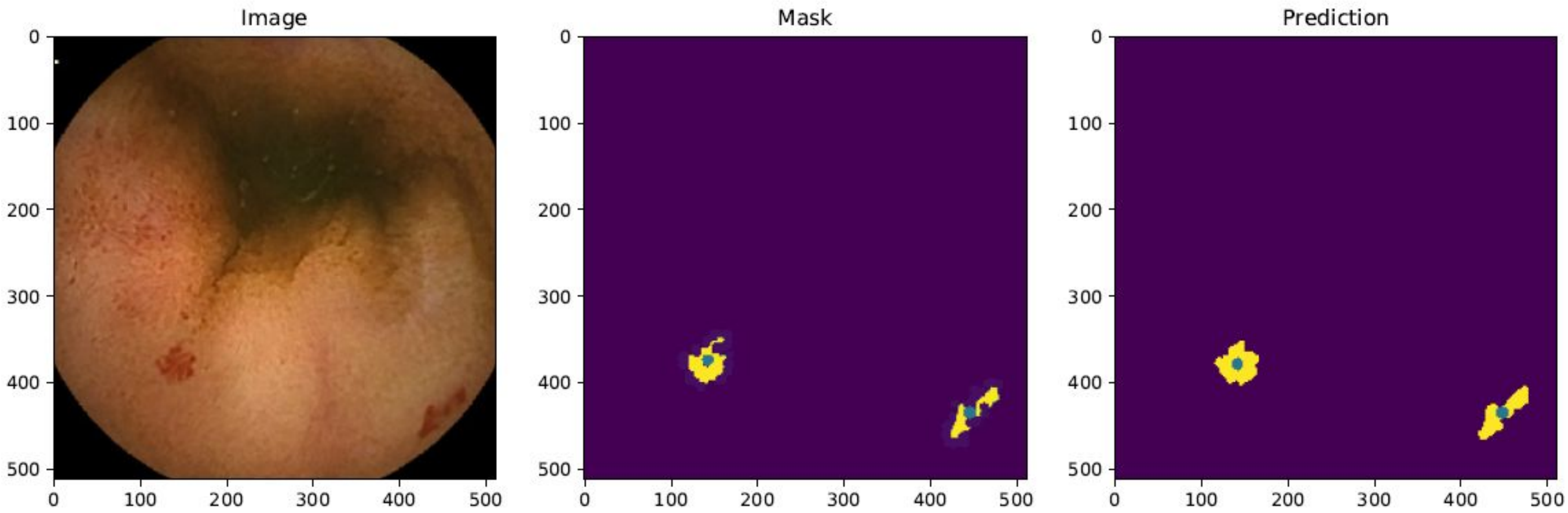
Each set: 300 images with angiodysplasia and 300 without the pathology



MICCAI Endoscopic Vision Challenge. 2017.
https://endovissub2017-giana.grand-challenge.org/Angiodysplasia-ETISDB/

Shvets, A.A., Iglovikov, V.I., Rakhlin, A., Kalinin, A.A., 2018. Angiodysplasia detection and localization using deep convolutional neural networks. ICMLA

Angiodysplasia segmentation



Model	IOU	Dice	Time
U-Net	73.18	83.06	30
TernausNet-11	74.94	84.43	51
TernausNet-16	73.83	83.05	60
AlbuNet-34	**75.35**	**84.98**	**21**

Shvets, A.A., Iglovikov, V.I., Rakhlin, A., Kalinin, A.A., 2018. Angiodysplasia detection and localization using deep convolutional neural networks. ICMLA

Coding tutorial

https://bit.ly/issondl2021_segment

https://github.com/alxndrkalinin/angiodysplasia-segmentation

Questions?