http://bit.ly/dlcv2018

#DLUPC

Day 2 Lecture 6

Medical Imaging at DCU

Kevin McGuinness
kevin.mcguinness@dcu.ie

Assistant Professor
School of Electronic Engineering
Dublin City University

DCU

Insight
Centre for Data Analytics

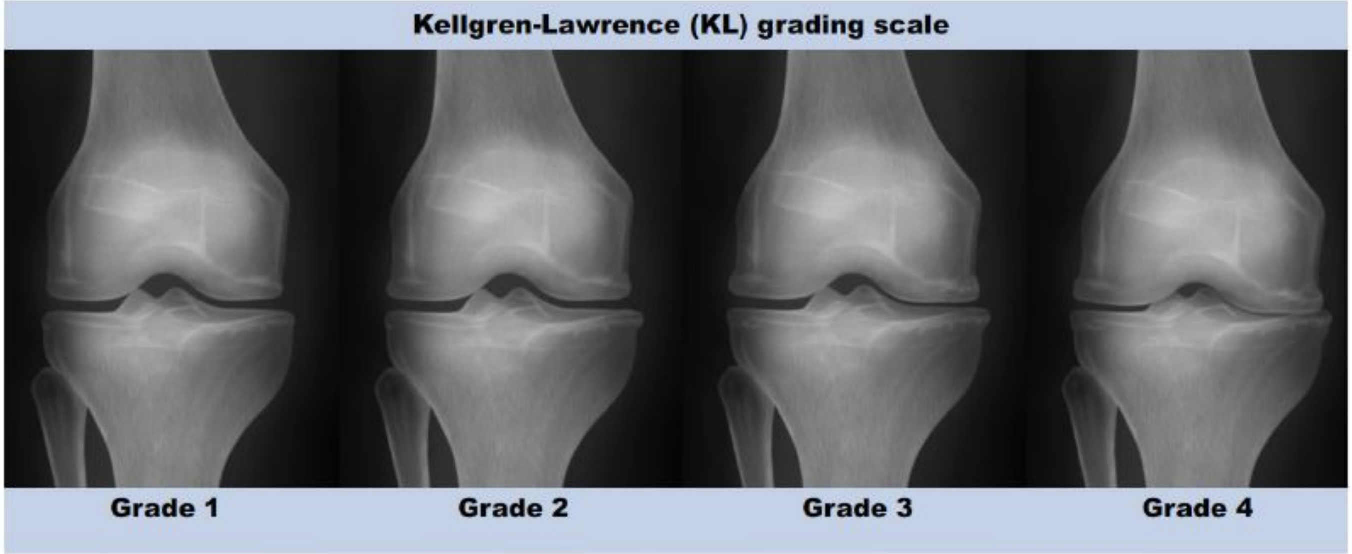
Overview

Computer aided diagnosis of knee-osteoarthritis

Neonatal brain image segmentation

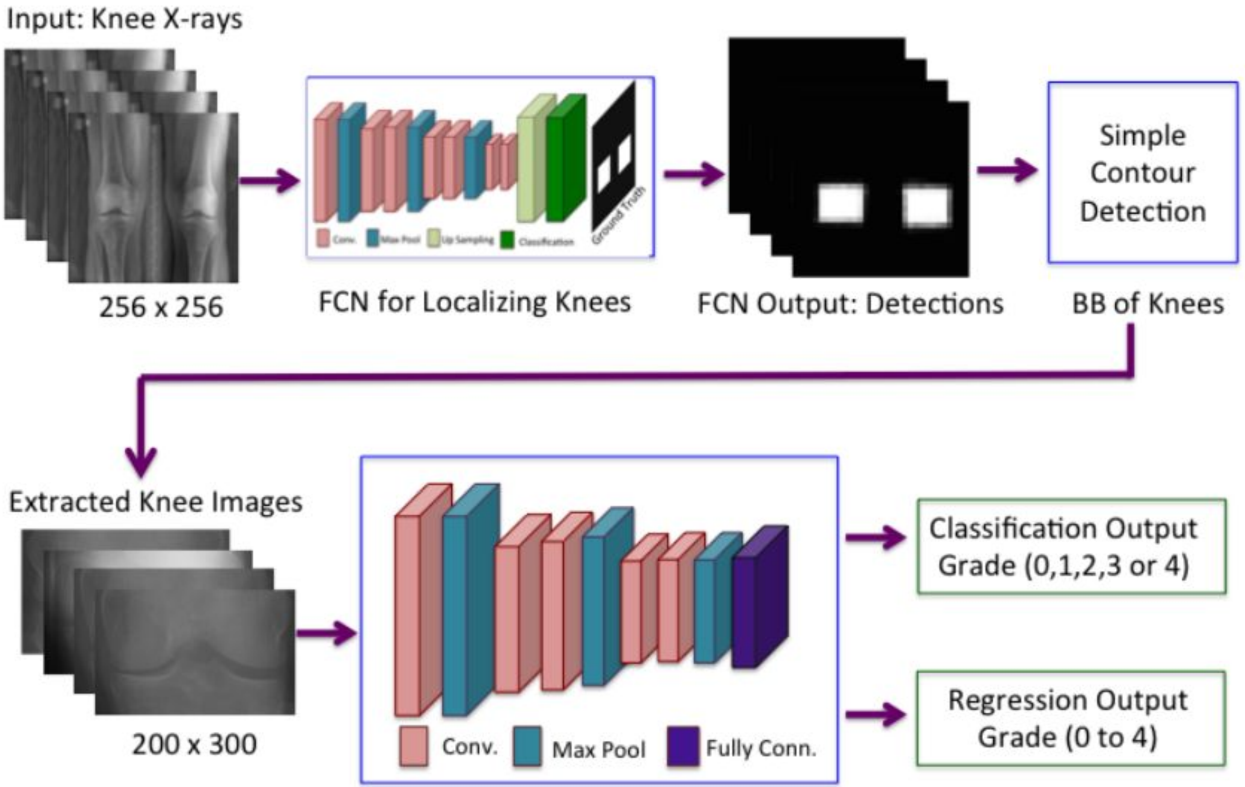
Detecting ischemic lesions in neonatal brains

Task: predict KL grade from X-Ray images



CLASSIFICATION	Normal	Doubtful	Mild	Moderate	Severe
DESCRIPTION	No features of OA	Minute osteophyte: doubtful significance	Definite osteophyte: normal joint space	Moderate joint space reduction	Joint space greatly reduced: subchondral sclerosis

Pipeline: locate and classify



Detection performance

FCN detection performance

Test Data	$J \geq 0.25$	$J \geq 0.5$	$J \geq 0.75$	Mean	Std.Dev
OAI	**100%**	**99.9%**	89.2%	0.83	0.06
MOST	99.5%	98.4%	85.0%	0.81	0.09
Combined OAI-MOST	99.9%	**99.9%**	**91.4%**	0.83	0.06

Template matching: (J > 0.5) 8.3%

SVM on handcrafted features: (J > 0.5): 38.6%

Multi-objective learning helps!

Same network used to regress on KL grade and predict a discrete KL grade

Jointly train on both objectives

Grade	Joint training for Clsf & Reg				Training for only Clsf			
	Precision	Recall	F_1	AUC	Precision	Recall	F_1	AUC
0	0.68	0.80	0.74	0.87	0.63	0.82	0.71	0.83
1	0.32	0.15	0.20	0.71	0.25	0.04	0.06	0.66
2	0.53	0.63	0.58	0.82	0.47	0.57	0.51	0.78
3	0.78	0.74	0.76	0.96	0.76	0.71	0.73	0.94
4	0.81	0.75	0.78	0.99	0.78	0.77	0.77	0.99
Mean	0.61	0.63	0.61	-	0.56	0.60	0.56	-

Comparison with the state-of-the-art

Method	Test Data	Accuracy	Mean-Squared Error
Wndchrm	OAI	29.3%	2.496
Wndchrm	MOST	34.8%	2.112
Fine-Tuned BVLC CaffeNet	OAI	57.6 %	0.836
Our CNN trained from Scratch	OAI & MOST	**60.3%**	0.898

How far are we from human-level accuracy?

Most errors are between grade 0 and 1 and grade 1 and 2.
Human experts have a hard time with these grades too.

Agreement among humans on OAI

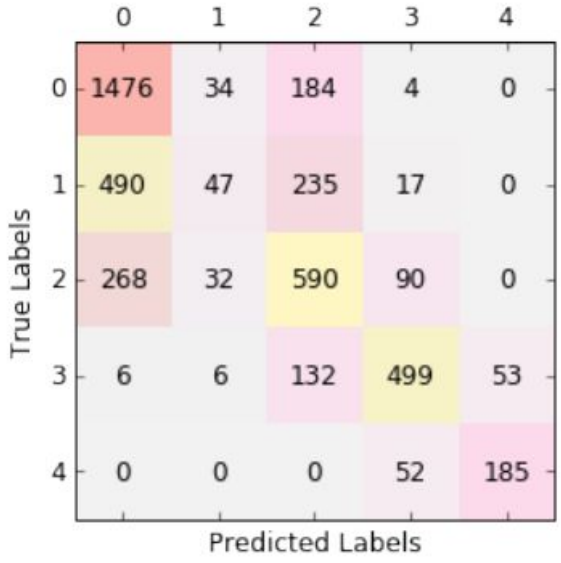
- weighted kappa of 0.70 [0.65-0.76]

Human machine agreement

- weighted kappa of 0.67 [0.65-0.68]

Predictions agree with the "gold standard" about as well as the "gold standard" agrees with itself.

Confusion matrix

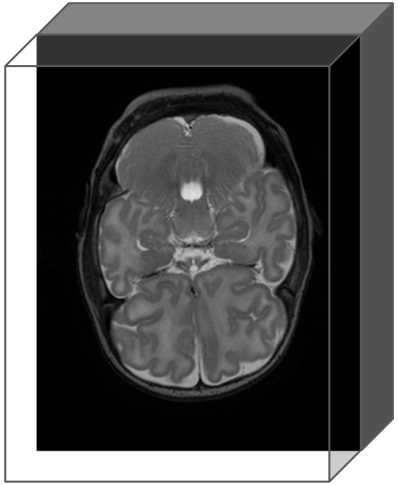


Neonatal brain image segmentation

Volumetric **semantic segmentation:** label each pixel with class of brain matter.

Applications:

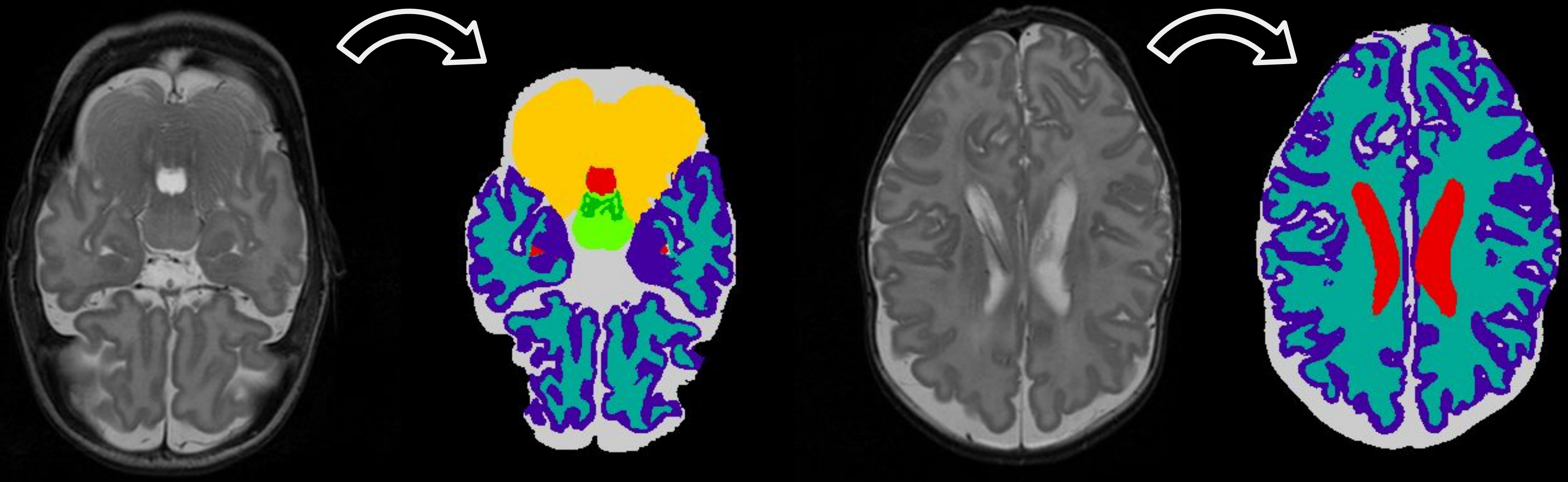
- Prerequisite for volumetric analysis
- Early identification of risk factors for impaired brain development



Challenge:

- Neonatal brains very different
- Sparse training data! Neobrains challenge has 2 training examples

The task



Cerebellum

Unmyelinated white matter

Cortical grey matter

Ventricles

Cerebrospinal fluid

Brainstem

Some background...

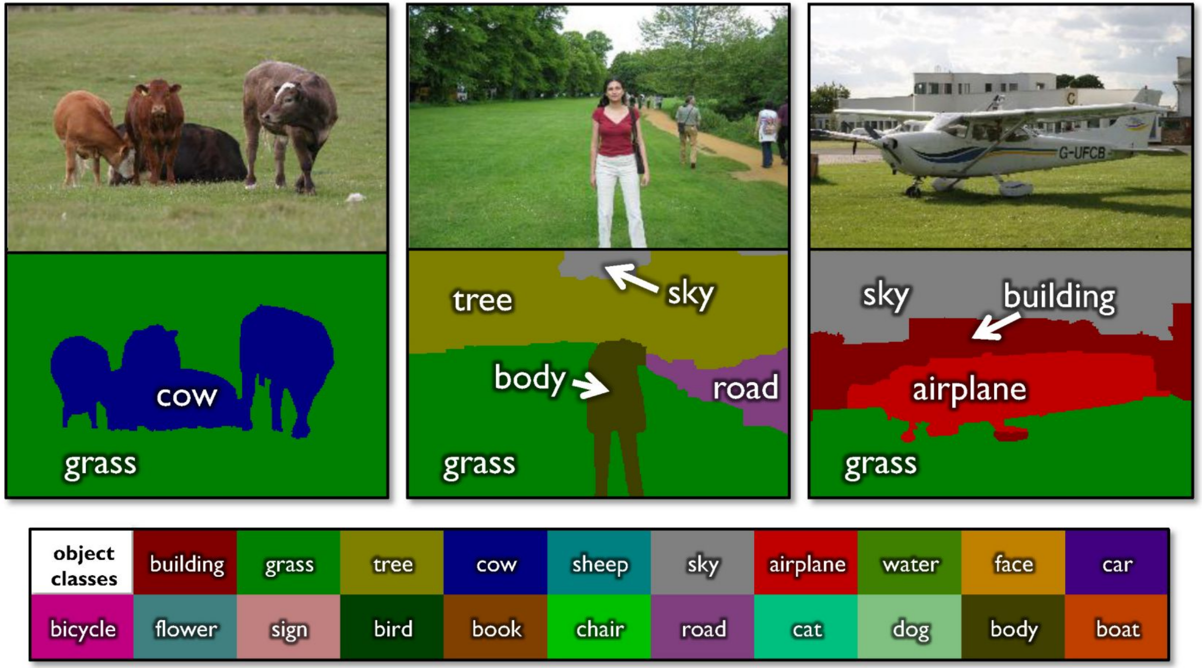
Semantic segmentation and fully convolutional networks

Semantic segmentation

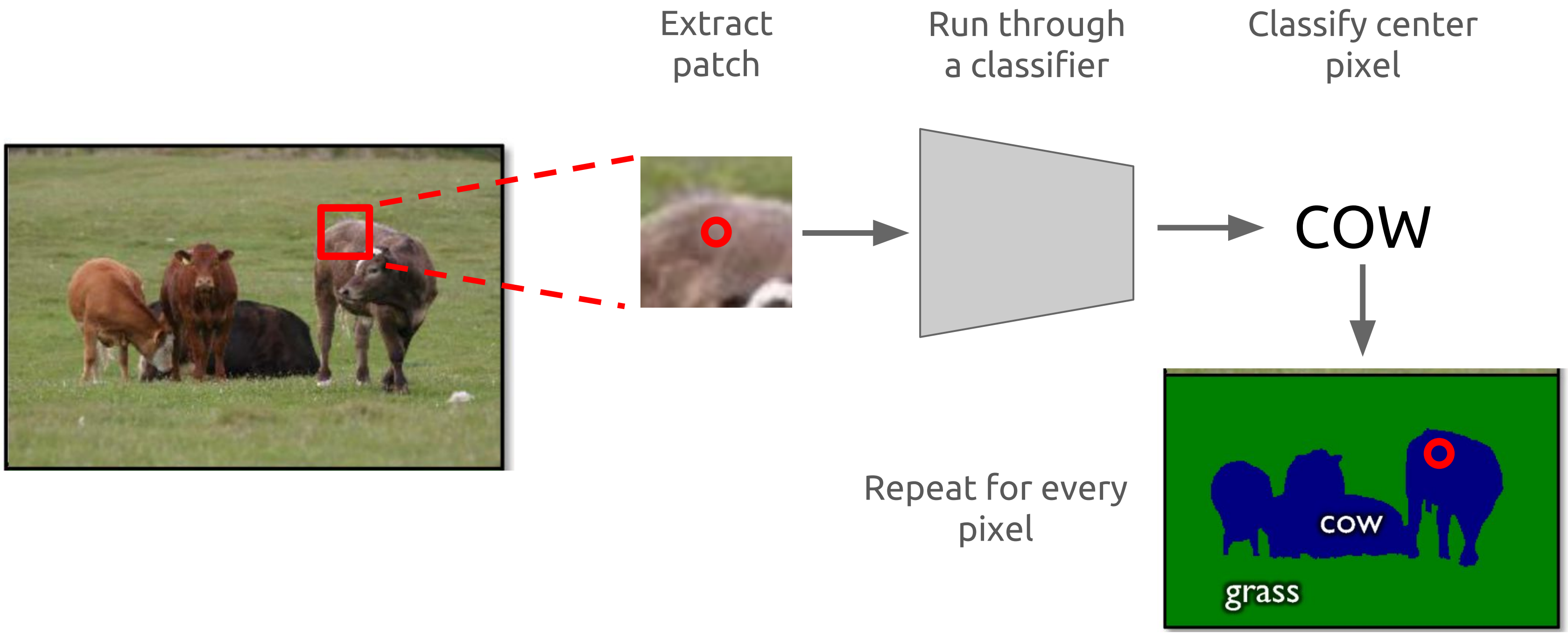
Label every pixel!

Don’t differentiate instances (cows)

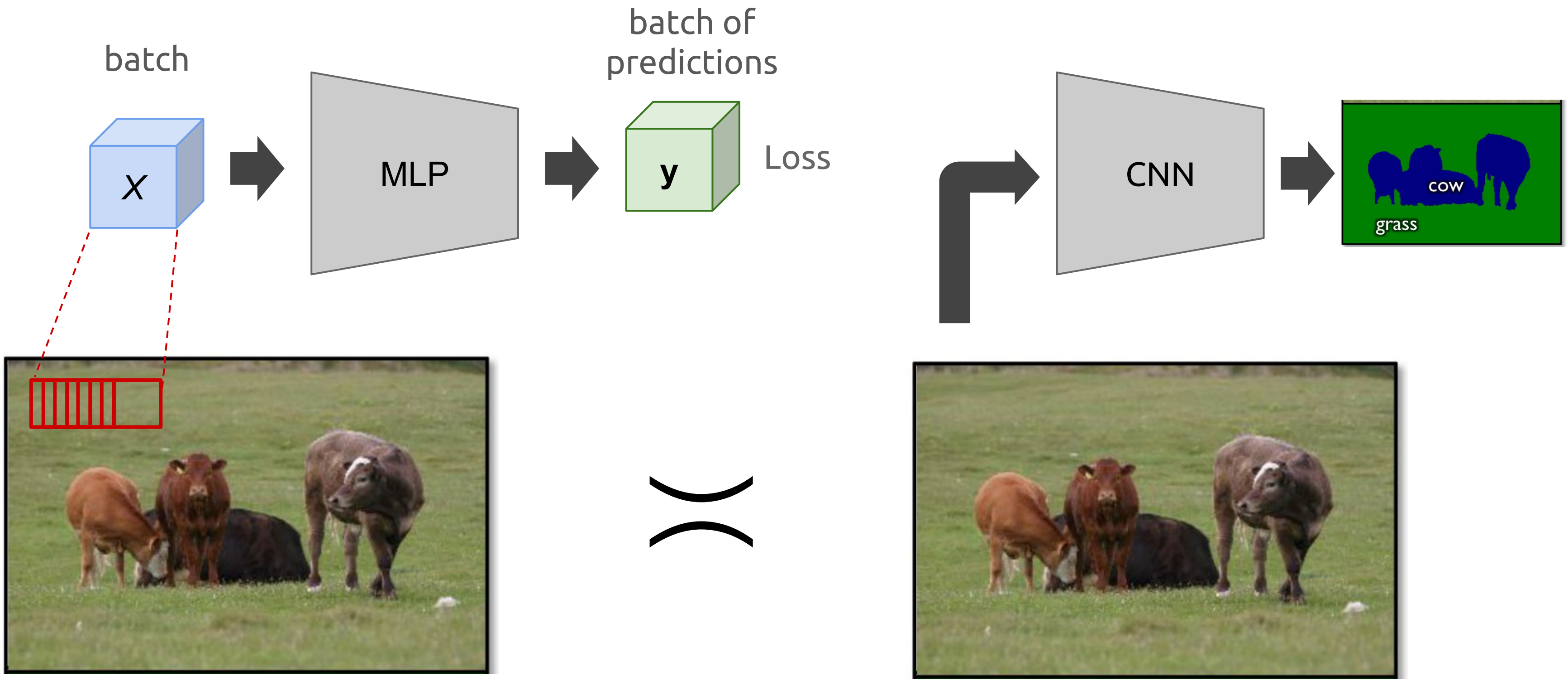
Classic computer vision problem



Semantic segmentation via pixelwise classification

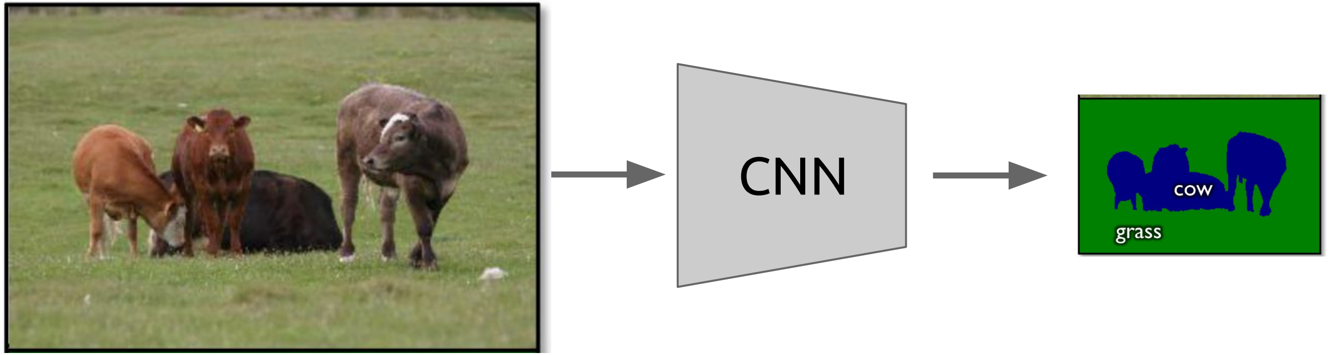


Convolutionalizing



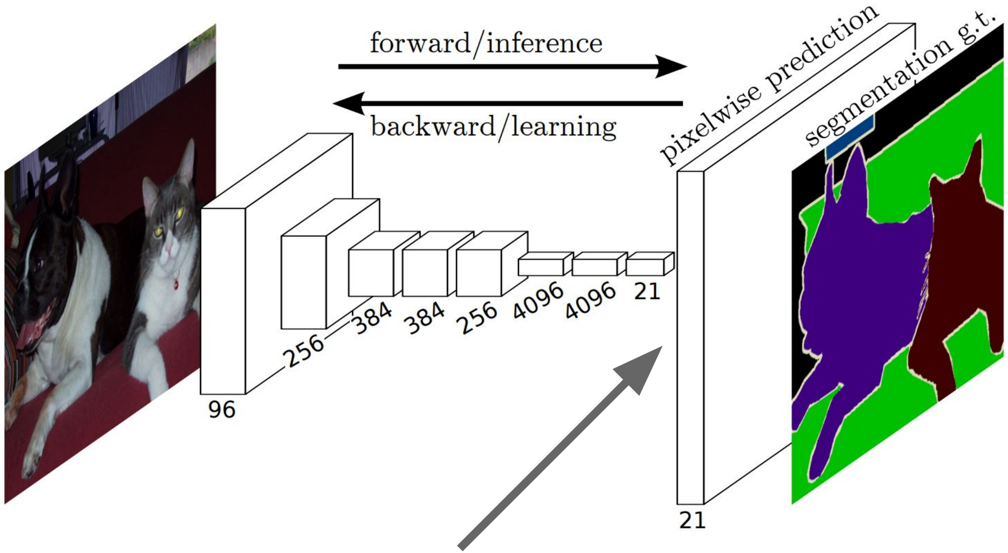
Semantic segmentation using a FCN

Run “fully convolutional” network to get
all pixels at once



Smaller output
due to pooling

Fully convolutional networks



Learnable upsampling!

Long et al. [Fully Convolutional Networks for Semantic Segmentation.](#) CVPR 2015

Dilated convolutions

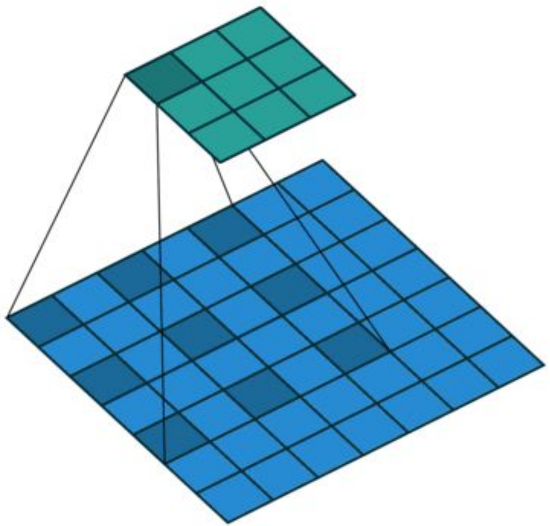
We usually use striding and pooling to reduce the resolution so subsequent filters can take into account a larger spatial region

- 2x, 3x effective aperture size
- Reduce computation
- Spatial invariance

Cost: loss of spatial resolution

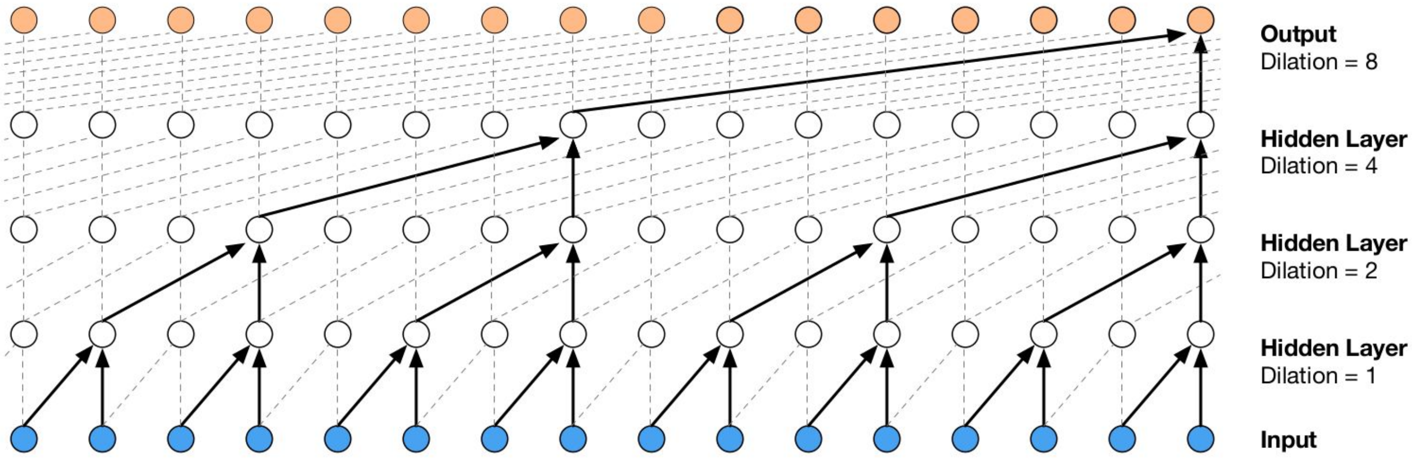
Dilated convolutions can be used to increase effective aperture size without sacrificing spatial resolution.

AKA. **Atrous convolutions**, convolution with holes.

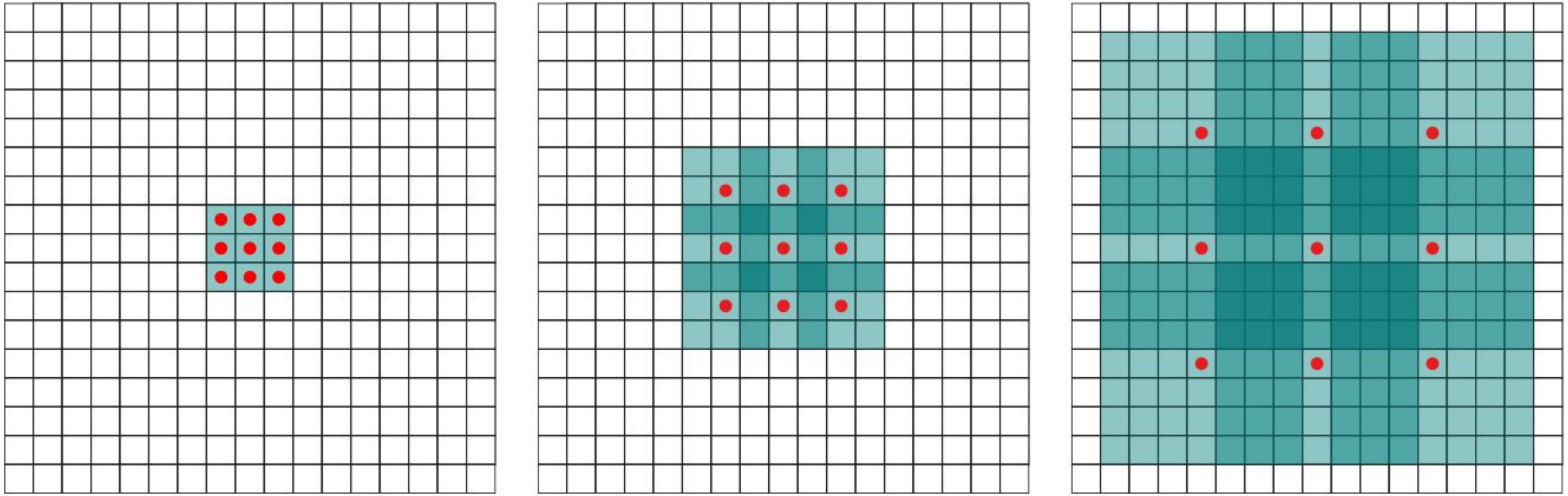


Animation from: http://nicolovaligi.com/deep-learning-models-semantic-segmentation.html

Dilated convolutions

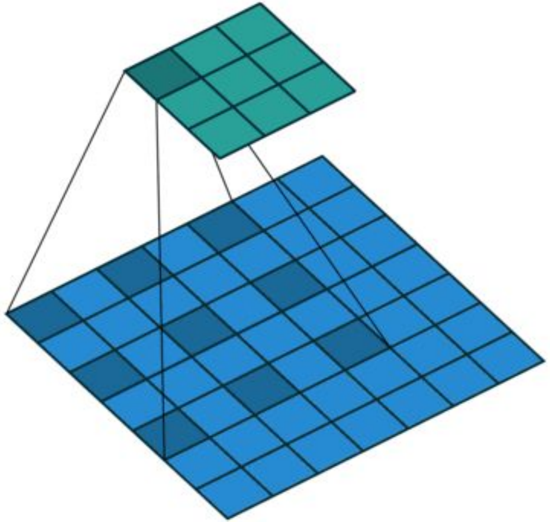


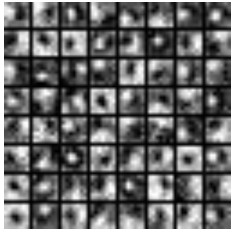
Dilated convolutions



Our model

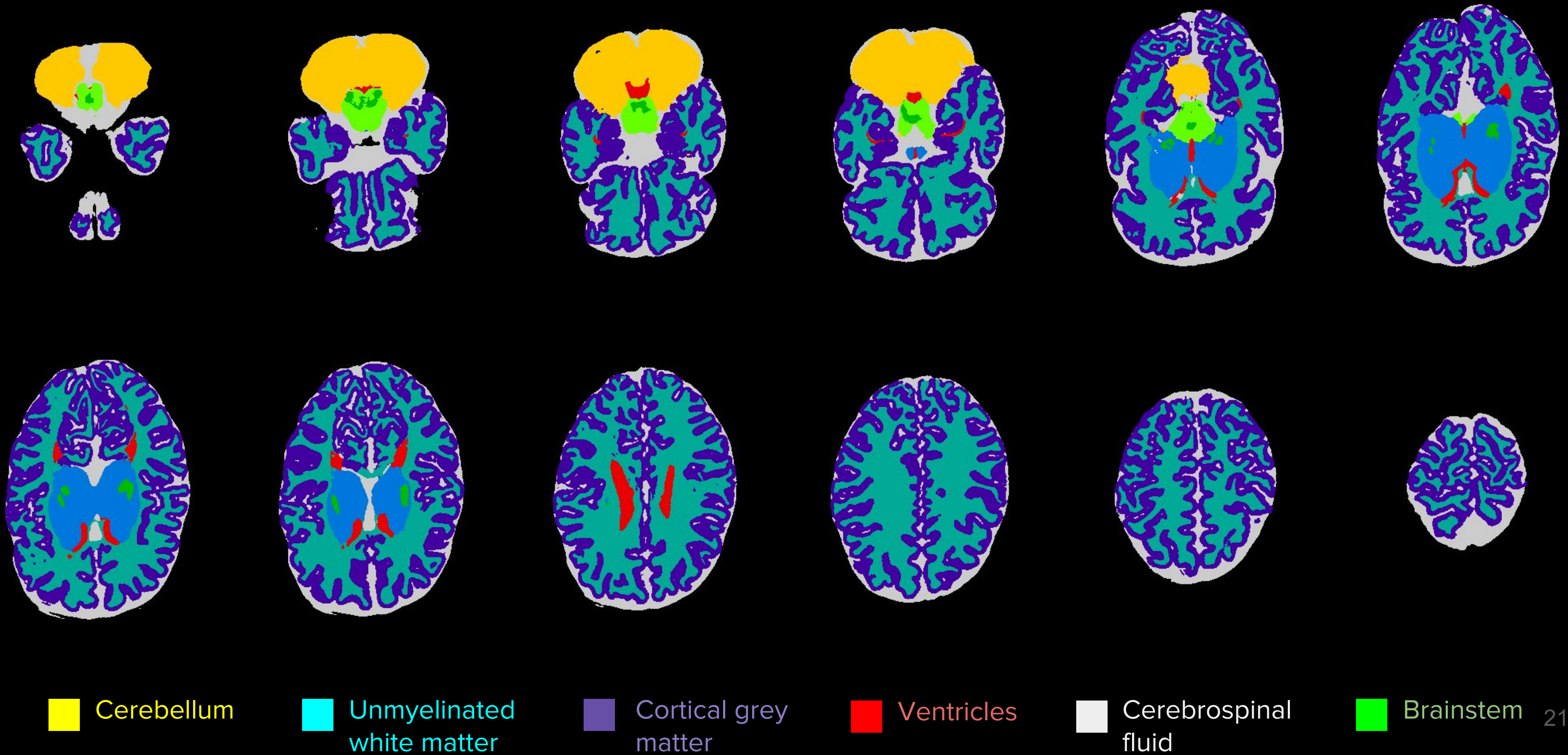
- 8 layer FCN
- 64/96 convolution filters per layer
- **Atrous (dilated) convolution** to increase receptive field without sacrificing prediction resolution
- 9D per pixel softmax over classes
- Binary cross entropy loss
- L^2 regularization
- **Aggressive data augmentation**: scale, crop, rotate, flip, gamma
- Train on 2 axial volumes (~50 slides per volume) for 500 epochs using Adam optimizer





Animation from: http://nicolovaligi.com/deep-learning-models-semantic-segmentation.html

Sample results



Neobrains challenge

New state of the art on Neobrains infant brain segmentation challenge for axial volume segmentation

Deep learning with only 2 training examples!

No ensembling yet. Best competing approach is a large ensemble.

Second best is also a deep net.

Tissue	Ours	LRDE_LTCI	UPF_SIMBioSys
Cerebellum	0.92	0.94	0.94
Myelinated white matter	0.51	0.06	0.54
Basal ganglia and thalami	0.91	0.91	0.93
Ventricles	0.89	0.87	0.83
Unmyelinated white matter	0.93	0.93	0.91
Brainstem	0.82	0.85	0.85
Cortical grey matter	0.88	0.87	0.85
Cerebrospinal fluid	0.83	0.83	0.79
UWM+MWM	0.93	0.93	0.90
CSF+Ven	0.84	0.84	0.79
	0.85	**0.80**	**0.83**

Detecting ischemic lesions

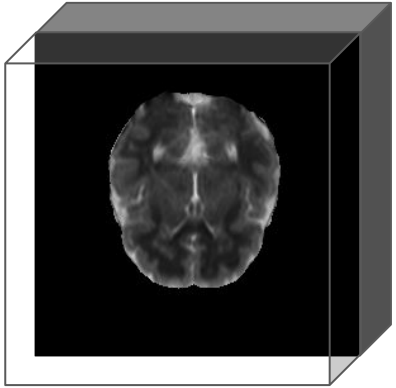
Another neonatal brain MRI analysis task

Pixelwize binary classification task: {lesion, no lesion}

Application:

- Automatic quantification of ischemic injury

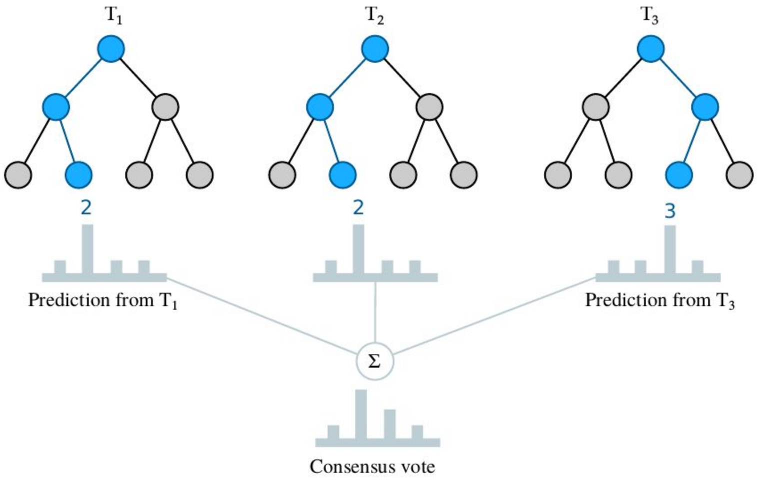
Same challenges as before: sparse labelled training data.



Previous state of the art

Random forest classifier trained on hand-crafted features

- Manually created brain mask
- Superpixel segmentation
- 9 hand crafted features based on:
    - Superpixel volume,
    - intensity,
    - spatial location,
    - distance to brain mask, etc.



Keelin Murphy et al, **Automatic quantification of ischemic injury on diffusion-weighted MRI of neonatal hypoxic ischemic encephalopathy**, NeuroImage: Clinical 14, 2017 https://doi.org/10.1016/j.nicl.2017.01.005

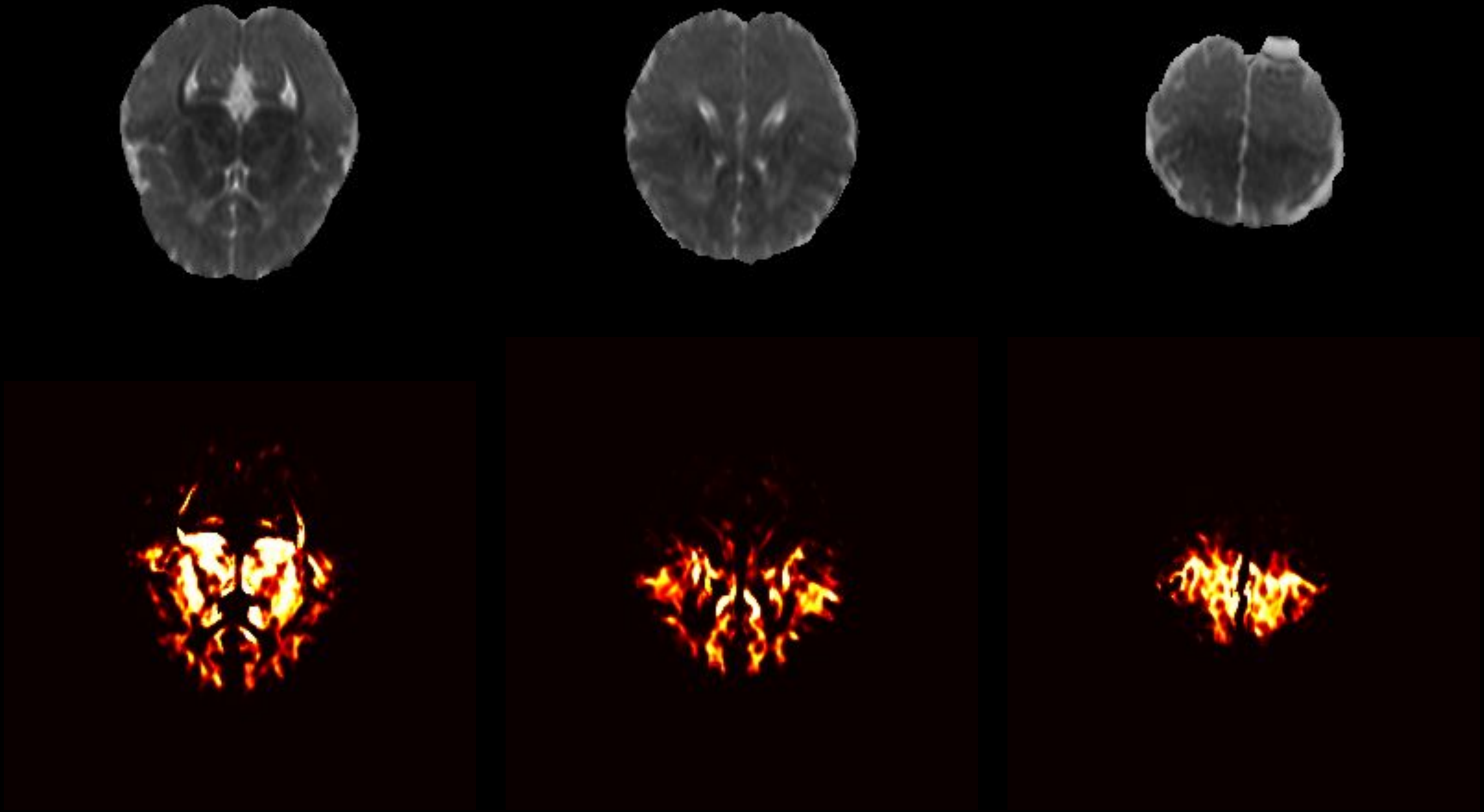
Our model

Almost identical to approach to brain segmentation!

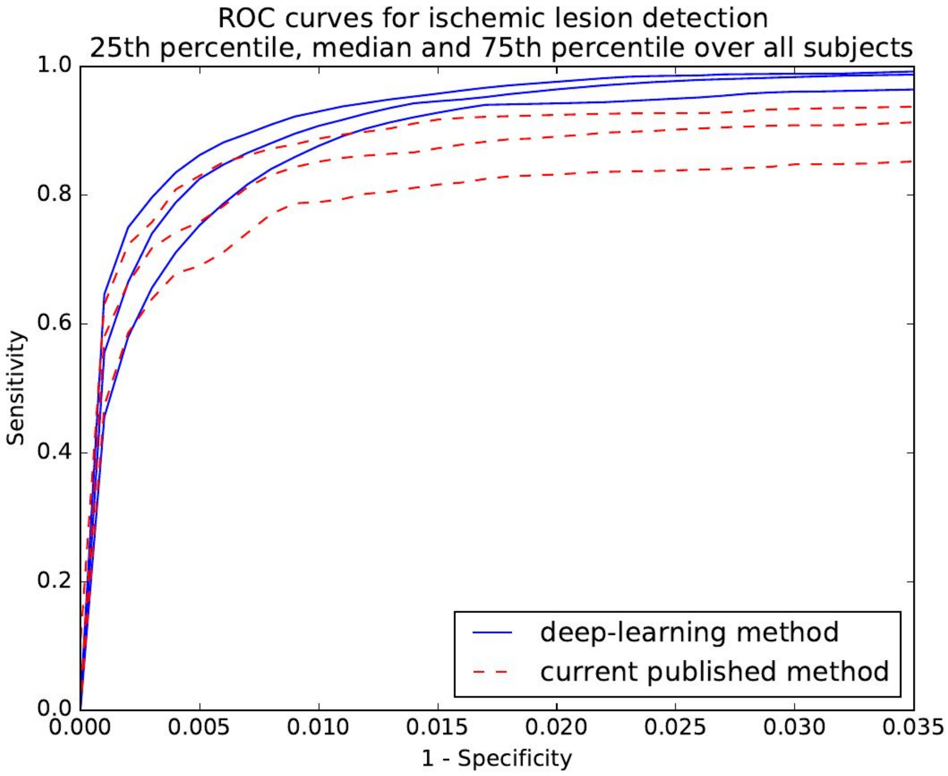
- Fully convolutional network (4 layer)
- Atrous convolutions
- Data augmentation
- 3x3 filters, 64-96-96-1 channels
- Adam optimizer

This is only an initial attempt: more improvements expected with improved model.

Sample results



Comparison with state of the art



Questions?