Multi-Scale 3D Convolutional Neural Networks for Lesion Segmentation in Brain MRI

Konstantinos Kamnitsas, Liang Chen, Christian Ledig, Daniel Rueckert, Ben Glocker

Biomedical Image Analysis Group, Imperial College London, UK

Objective: Fully automatic, accurate segmentation

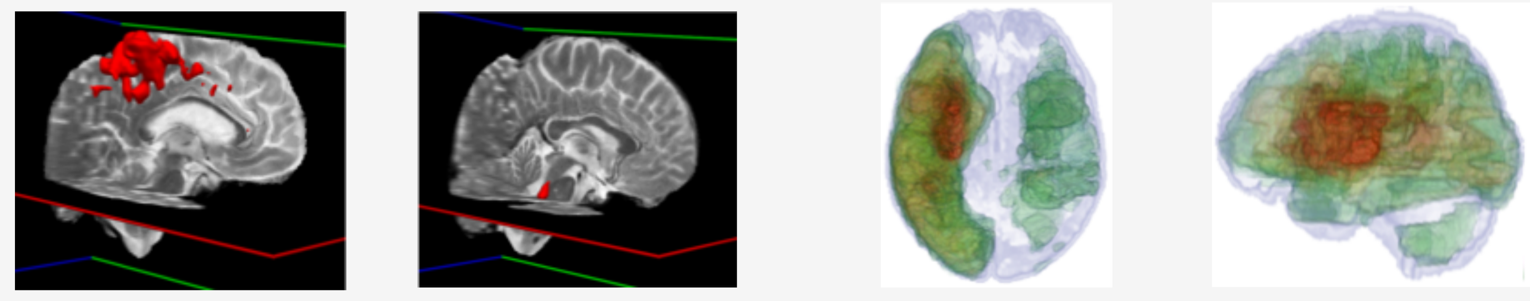
- Manual segmentation is time consuming
- Multi-modal 3D scans can be challenging to segment even for an expert
- An automatic system could enable studies on large cohorts of patients
- Accurate performance is critical. Small lesions may be of high importance
- Computational analysis may reveal characteristics of the lesions
- Investigate Deep Learning and CNNs, following advances in Computer Vision

Method: Convolutional Neural Networks

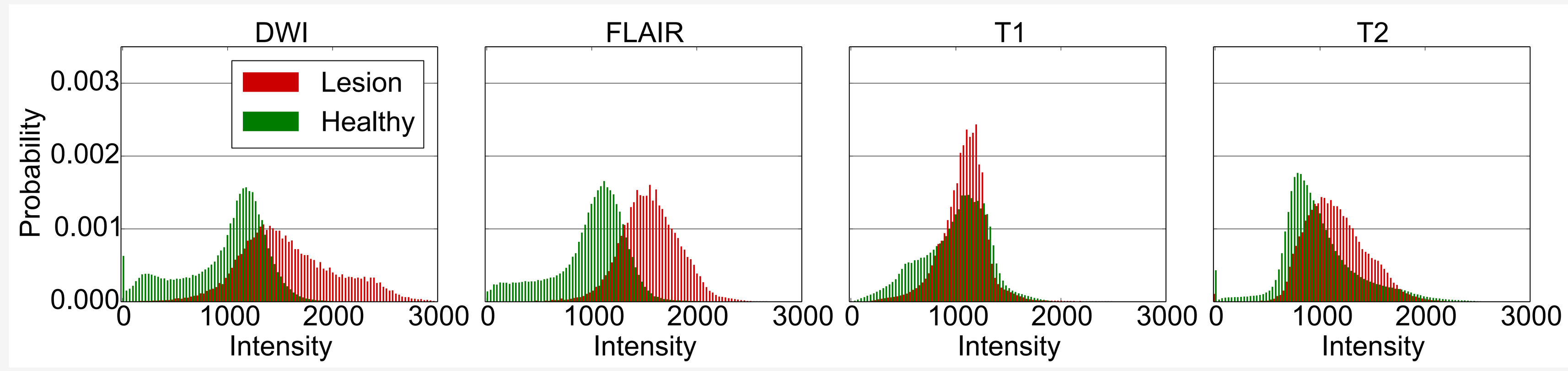
- Automatic learning of data-driven features for the segmentation
- Computationally efficient processing of multi-modal scans
- Training on small datasets possible with modern regularisation methods
- Scalable to large training datasets
- Generic technique, applicable to various segmentation tasks

Challenge: Vastly heterogeneous lesions

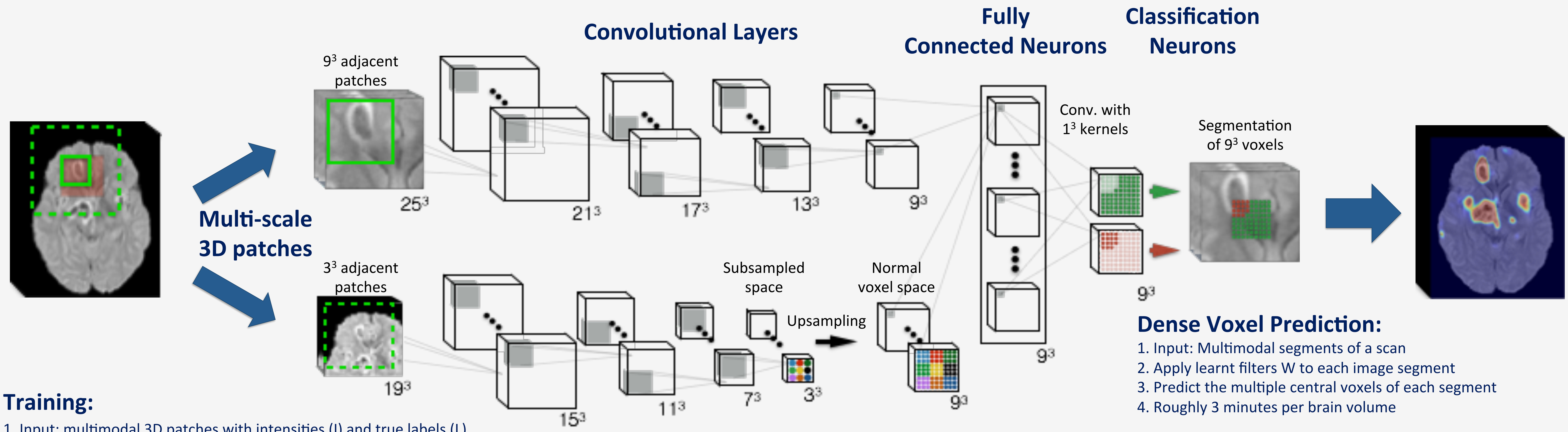
› Lesions vary in size and shape. They may also occur in various locations



› The intensity profile of the lesions is hard to model



› Hand-engineering feature detectors may be problematic

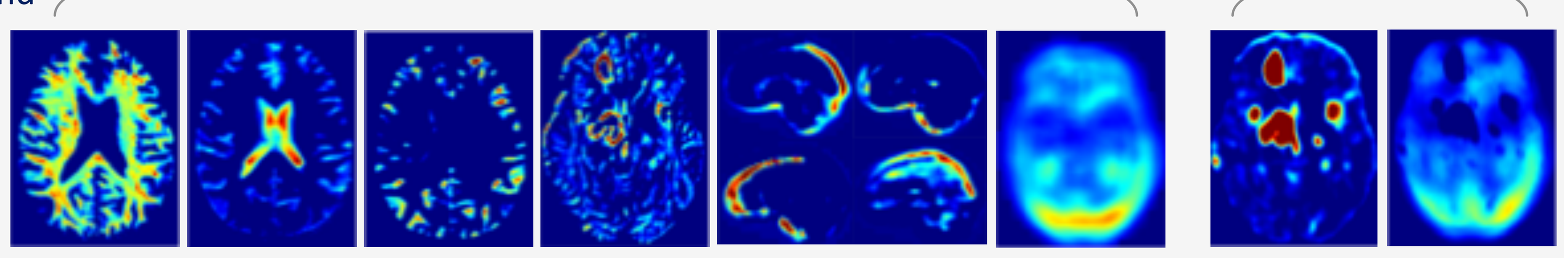


Dense Voxel Prediction:

1. Input: Multimodal segments of a scan
2. Apply learnt filters W to each image segment
3. Predict the multiple central voxels of each segment
4. Roughly 3 minutes per brain volume

Training:

1. Input: multimodal 3D patches with intensities (I) and true labels (L)
2. Learn filters (W) to minimize the error between prediction (C) and provided true labels (L): $\hat{W} = \max_W P(C = L | I, W)$
3. Dense voxel prediction is exploited in training as in [1]
4. Takes 1-2 days on a modern GPU

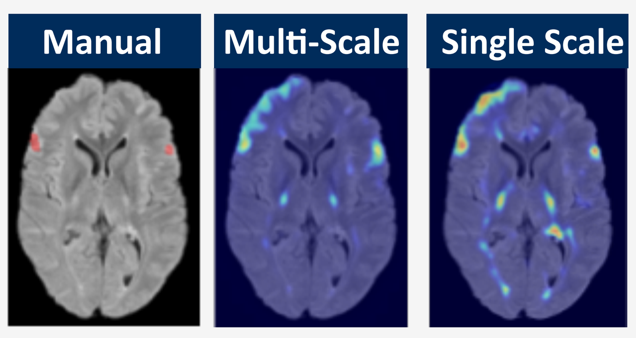


Deep Learning the Lesions:

The neurons of the network act as feature detectors. Neurons of deeper layers learn to detect more complicated features. This gradually leads to the final filters that perform the classification of the patch and the voxel's segmentation.

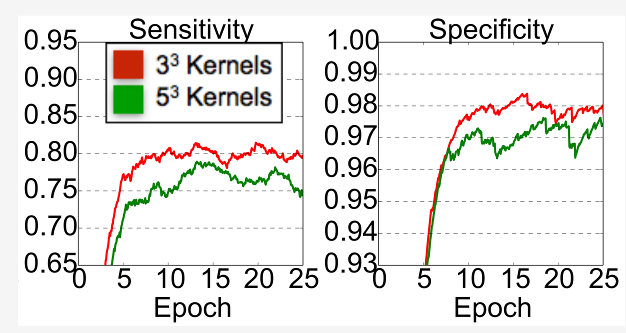
Multi-Scale Processing:

- Additional content should lead to improved classification
- Increasing CNN's receptive field is computationally prohibited
- We propose the parallel processing of the image at multiple scales with the use of parallel convolutional pathways



Going 11-Layers Deep:

- Deep architectures can learn to model highly non-linear functions
- The complexity of the task suggests this would be beneficial
- The final 11-layer deep 3D CNN is built by replacing 5^3 with 3^3 kernels
- Performance is improved with lower number of total parameters thanks to the smaller kernels



Evaluation

Dataset: ISLES'15 SISS Training Data

› Multimodal brain-scans of 28 patients with Sub-acute ischemic stroke lesions

› Modalities: DWI, FLAIR, T1, T2

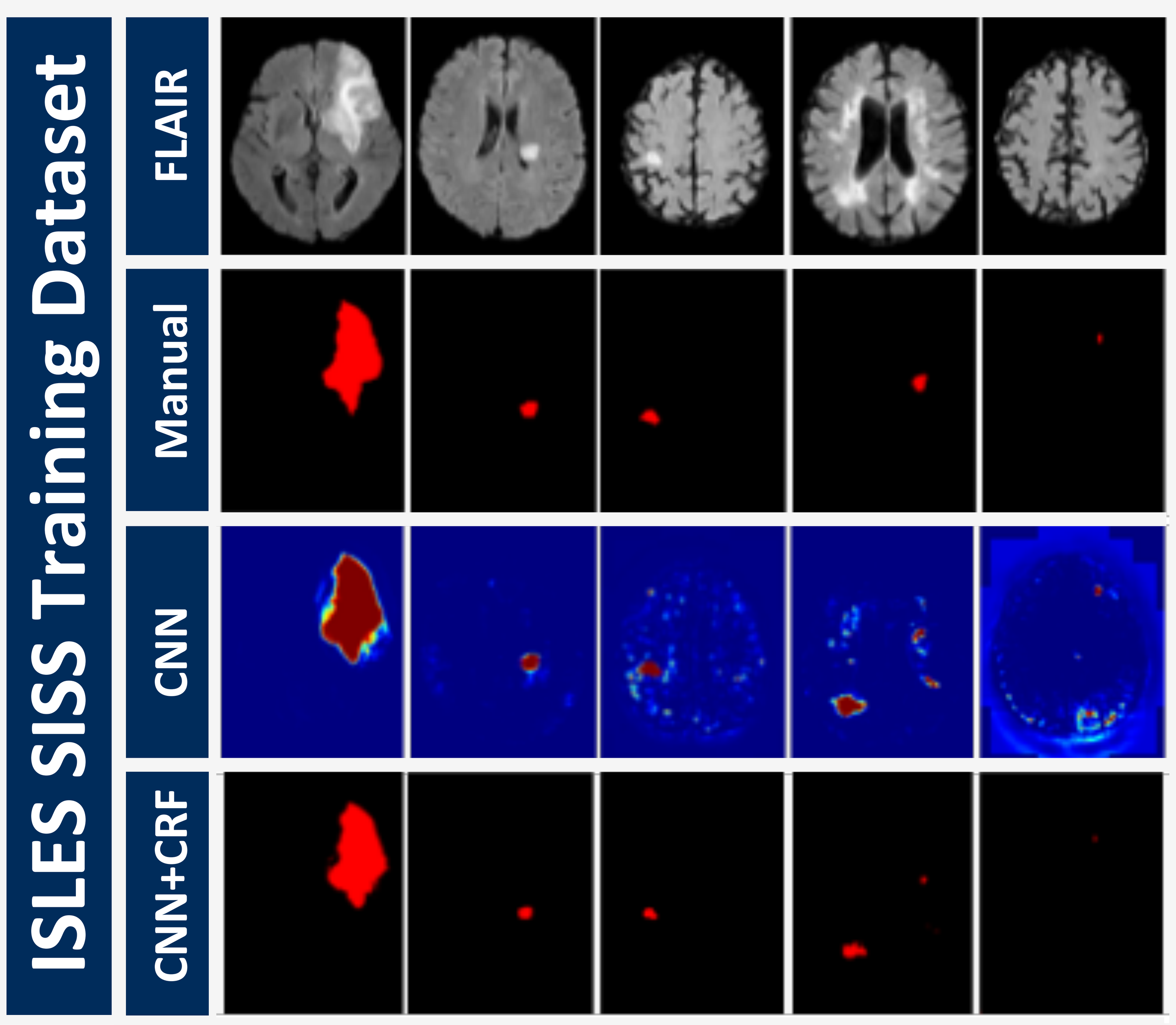
Pre-processing:

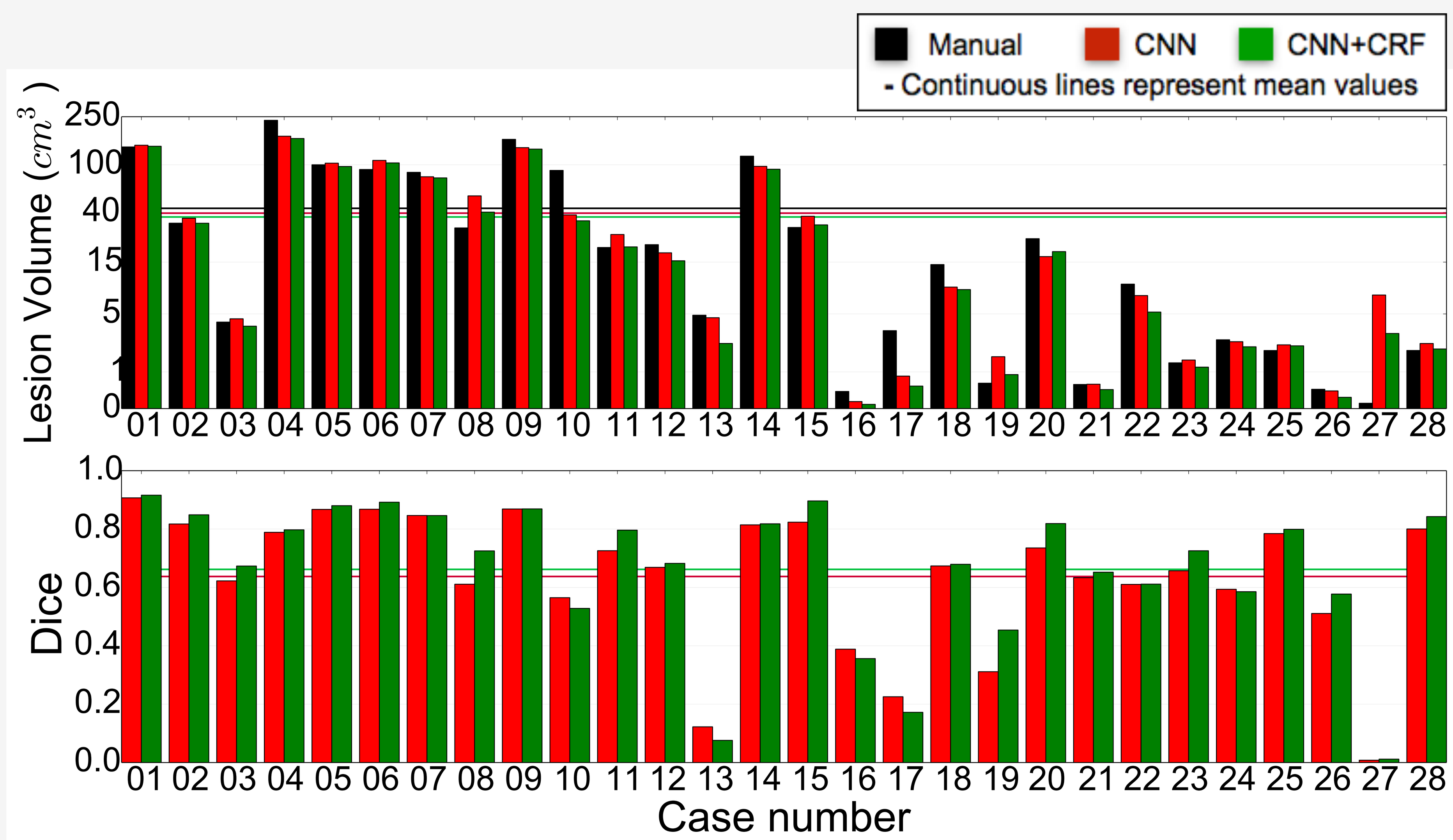
› The scans of each patient were normalized to zero mean and unity standard deviation

Post-processing:

› Our 3D version of the CRF presented in [2] was employed for reducing over-segmentation

› Connected components smaller than 20 voxels were eliminated





References:

[1] Urban, G., Bendszus, M., Hamprecht, F., Kleesiek, J.: *"Multi-modal brain tumor segmentation using deep convolutional neural networks"*, in proc of BRATS-MICCAI (2014)

[2] Krähenbühl, P., Koltun, V.: *"Efficient inference in fully connected CRFs with gaussian edge potentials"*, arXiv: 1210.5644 (2012) 1-9

Contact: konstantinos.kamnitsas12@imperial.ac.uk

Imperial College London