# Convolutional neural networks for artifact free OCT retinal angiography

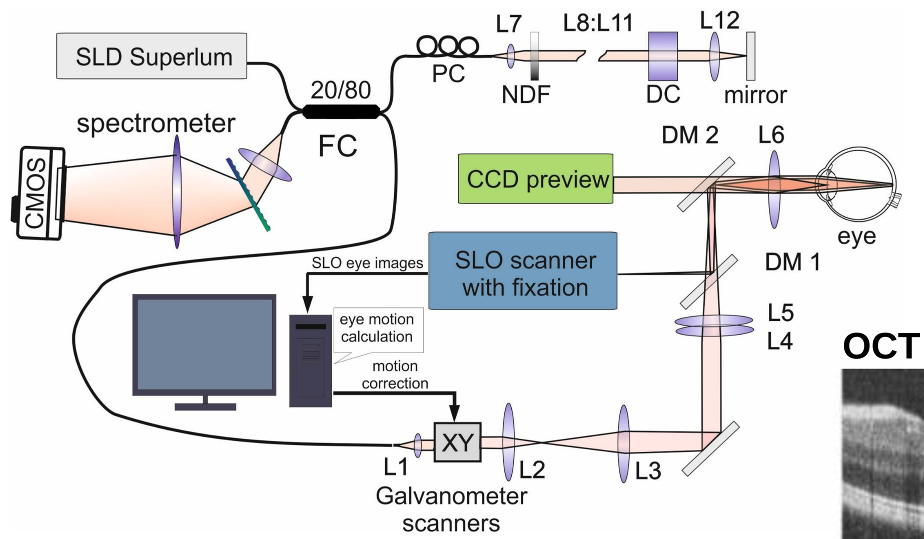
Maciej Szkulmowski,[1,*] Paweł Liskowski,[2] Bartosz Wieloch,[2] Krzysztof Krawiec,[2] Bartosz L. Sikorski[3]
[1]Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Toruń, Poland
[2] Laboratory of Intelligent Decision Support Systems, Poznan University of Technology, Poznań, Poland
[3] Department of Ophthalmology, Collegium Medicum, Nicolaus Copernicus University, Bydgoszcz, Poland
* E-mail: maciej.szkulmowski@fizyka.umk.pl

649 B0396

## Purpose

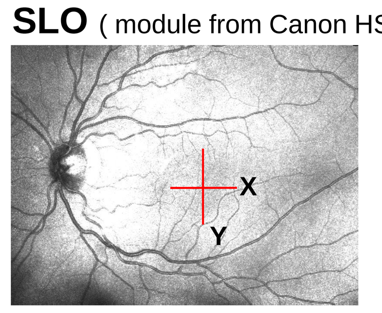
To demonstrate ability of deep convolutional neural network (CNN) to provide noninvasive visualization of retinal microcapillary network (RMN) with the use of data from a device combining Scanning Laser Ophthalmoscope (SLO) and Spectral Optical Coherence Tomography (SOCT). The approach allows RMN to be presented in form of 3D visualization as well as in forms of angiographic maps of different retinal layers **free of shadow artifacts** blurring standard RMS visualizations.

## OCT+SLO image acquisition



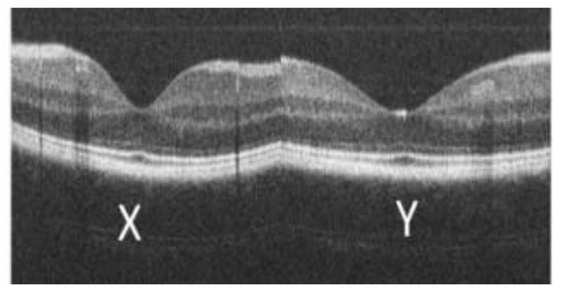

SLD - superluminescent diode; FC – fiber coupler; L1:L12 – lenses; PC - polarization controller; NDF - neutral density filter; DC - dispersion compensation; DM1&2 – dichroic mirrors.

**SLO** ( module from Canon HS-100)



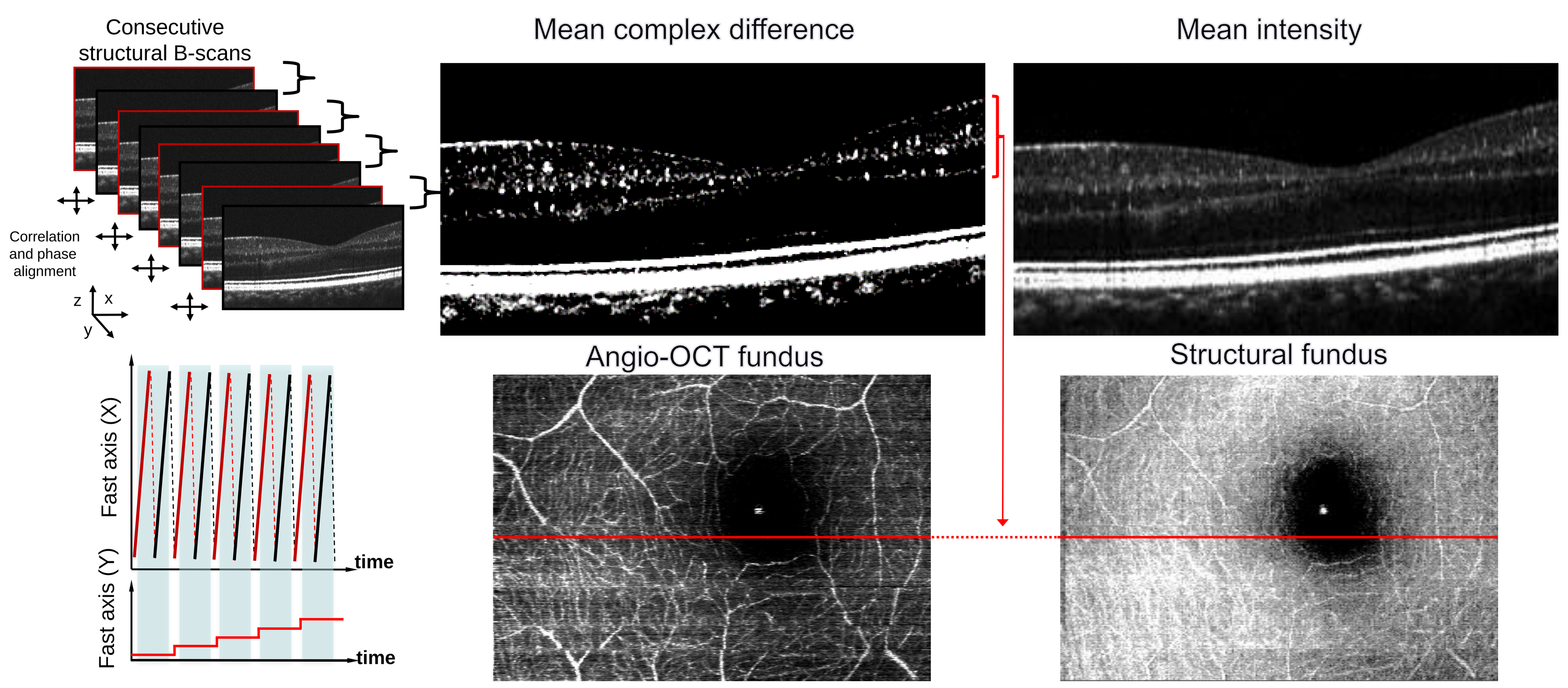

**Lateral resolution**
Δx, Δy = 30 µm
**Imaging speed**
27 Hz (800 x 300 Pixels)
*En face* preview
**Tracking accuracy:**
< 30 µm

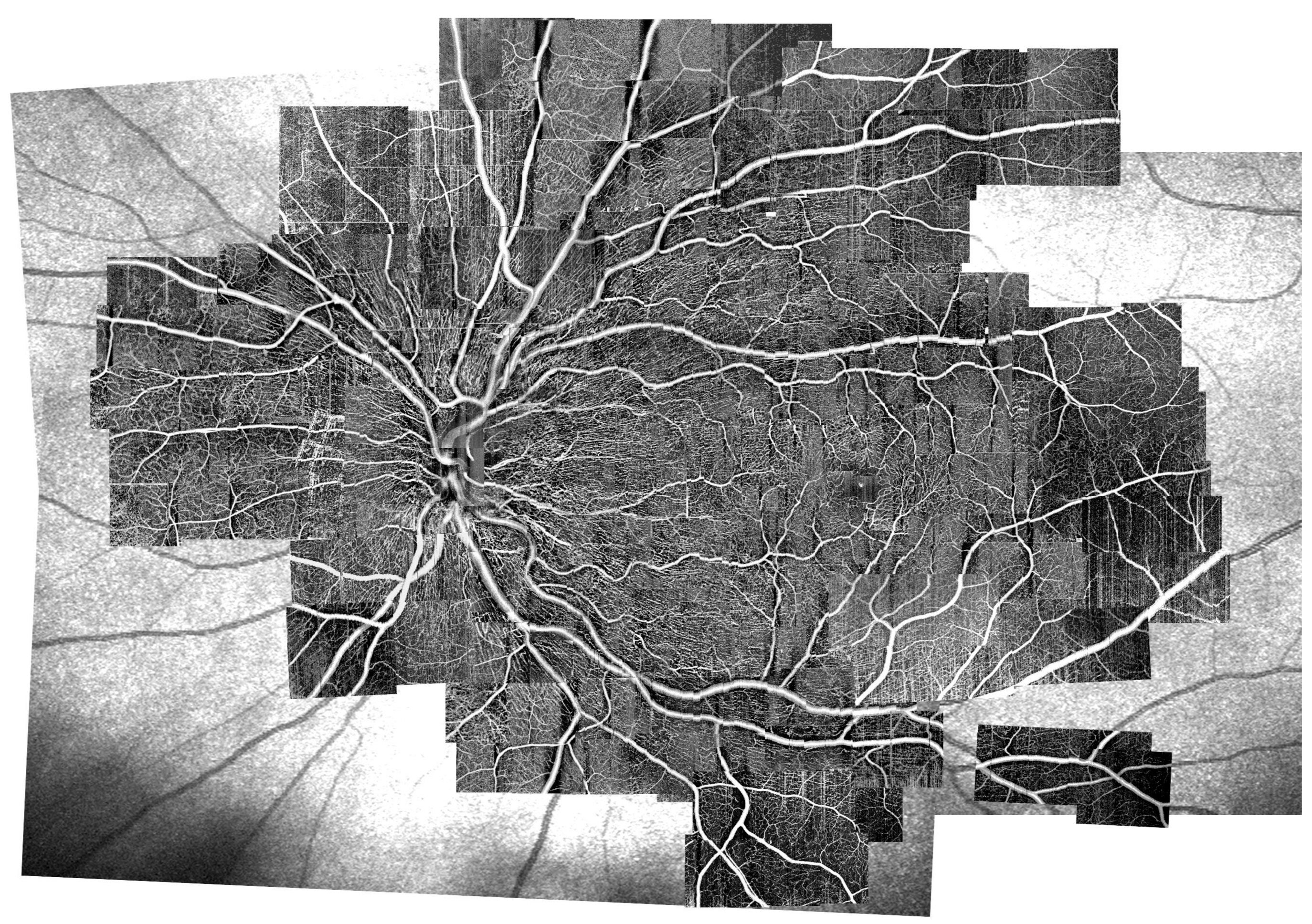
**OCT**




**Axial resolution**
Δz = 4.5 µm
**Lateral resolution**
Δx, Δy = 16 µm
**A-scan rate**
100 kHz
(max 350 kHz)

## Angio OCT data analysis



**Normal retinal microcapillary network with tile tracking and rescan mode.** Data collected from 29 years old healthy volunteer. Visual acuity was 20/20 in both eyes. Data overlaid on SLO image. SLO tracking mode: tiles positioning, rescan.

## Convolutional neural network (CNN) for OCT angiography

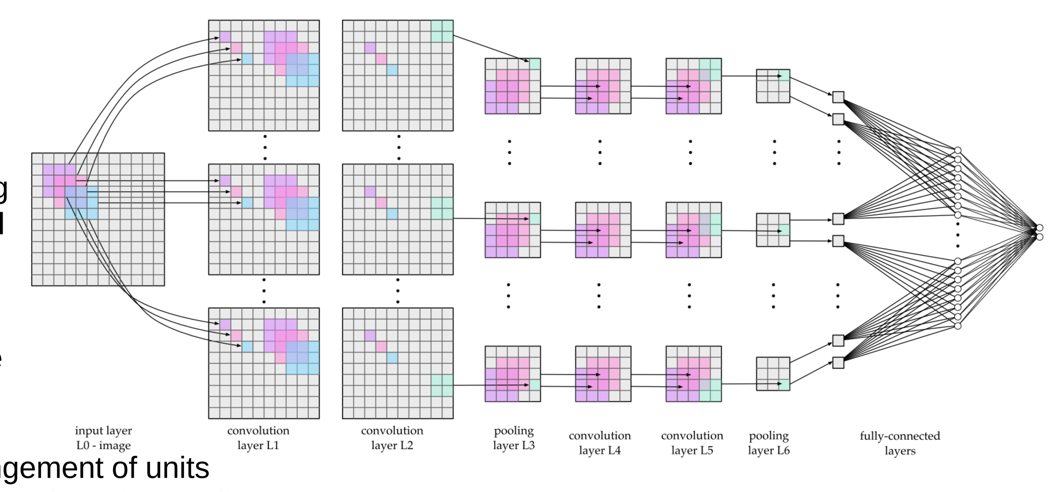
### Motivation

Convolutional Neural Networks (CNNs) :

1. are currently the **best known machine learners** for many problems in pattern recognition, computer vision, and beyond, very well suited for **large volumes of training data** provided by OCT,
2. prove superior for segmenting blood vessels in fundus imaging (best-to-date accuracy (Liskowski & Krawiec 2016)),
3. are an effective tool in **differentiating retinal angiographic signal from shadow artifacts**.

### CNN architecture

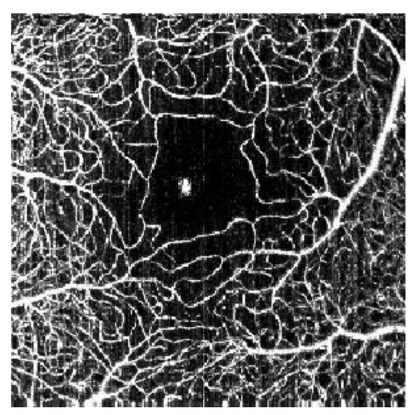
**Convolutional Neural Network (CNN)** – a large number of interconnected processing **units**, arranged in a series of **convolutional layers** followed by some **fully-connected layers** (feed-forwarded network).



**Convolutional layers** learn localized image features:

1. **Layer** = a set of feature maps
2. **Feature map** (FM) = a rectangular arrangement of units
3. **Unit** = weighted sum of input signals (convolution) + nonlinearity
4. Each unit in a FM corresponds to a location in the input image and has a spatially limited **receptive field (RF)** that weighs (convolves) input signals with a trainable vector of weights. Units in a FM **share weights,** which reduces the number of required connections.
5. FMs may be alternated with **max-pooling layers** that perform localized spatial aggregation of features and so reduce dimensions of subsequent layers.
6. In **3D network architectures** (**CNN 3D**), RFs and FMs are three-dimensional structures: each RF is a 11x11x11 cube, and each FM is a 3D cube of units, so the 3D input image (a cube of voxels taken from the OCT image) is fed directly into network.
7. In **2D architectures** (**CNN**), RFs and FMs are two-dimensional: each RF is a 11x11 square, and the 3D piece of OCT image is sliced along the Z axis into 11 2D images, which are fed into individual FMs. Thus, 2D architectures do not perform convolution along the Z axis, while 3D architectures do.
8. For CNN with **structural prediction** (**CNN SP**), networks have 27 output units that predict class for each of the 3 × 3 × 3 voxels centered at the input RF. As the input RF is shifted across the input OCT image, each voxel is classified 27 times, and the final decision is made via majority voting.

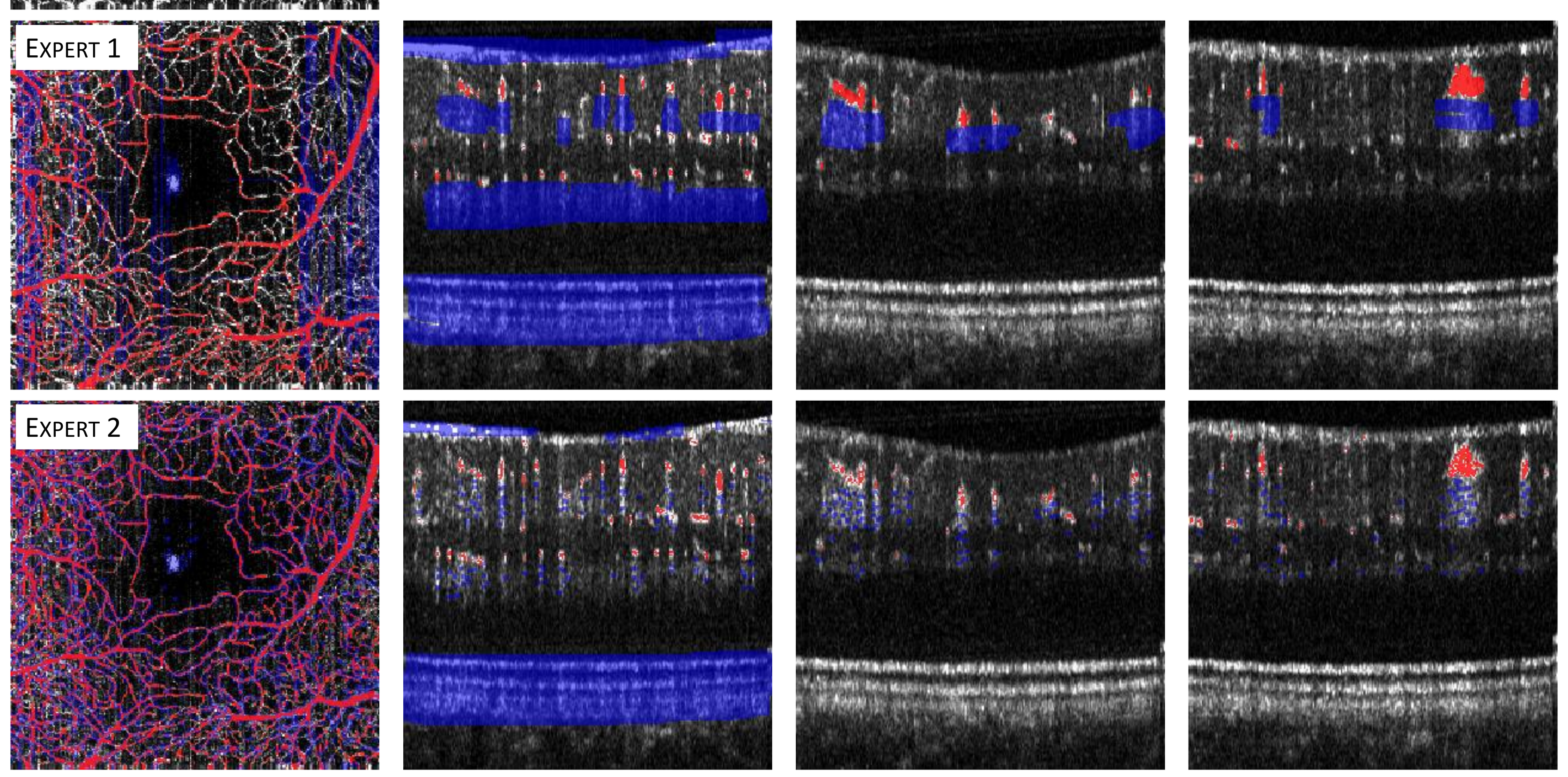
### CNN training



**Normal retinal microcapillary network.** Data collected from 32 years old volunteer. Field of view 1.5mm x 1.5mm. 240Ascans x 240Bscans x 400 pixels (~23 milion voxels).

Positive (**red – vessel**) and negative (**blue – not a vessel**) example voxels were labeled by two experts:

1. **Expert 1**: 51 063 positive and 358 699 negative examples
2. **Expert 2**: 38 909 positive and 144 667 negative examples




**Patch** – 11×11×11 cube surrounding a labeled voxel (1331 voxels)

- Central voxel determines class label (vessel, non-vessel)
- Voxel values standardized

Patches **partitioned into training set** and **test set**;
two variants:

- **random split**: 75% training set, 25% test set,
- data set labeled by different experts for training and testing sets

**Training –** stochastic gradient descent via backpropagation: patches forward-propagated through the network, output errors propagated backwards, units' weights updated (in batches) according to delta rule.

**Training details:**

- loss function: cross-entropy, weights initialized with N(0, 1),
- up to 50 epochs (full passes over the training set), nearly 26,000 iterations, total of 13,272,000 examples seen by the network,
- batch of 512 examples are simultaneously processed,
- stopping condition: validation error stops improving.

**Performance:**

- classifying a batch of 512 examples: 250 ms,
- segmenting whole 3D image: **up to 30 min** depending on the architecture (not optimized in this research)

## Results

### Quality of segmentation

CNNs trained on voxels labeled by:

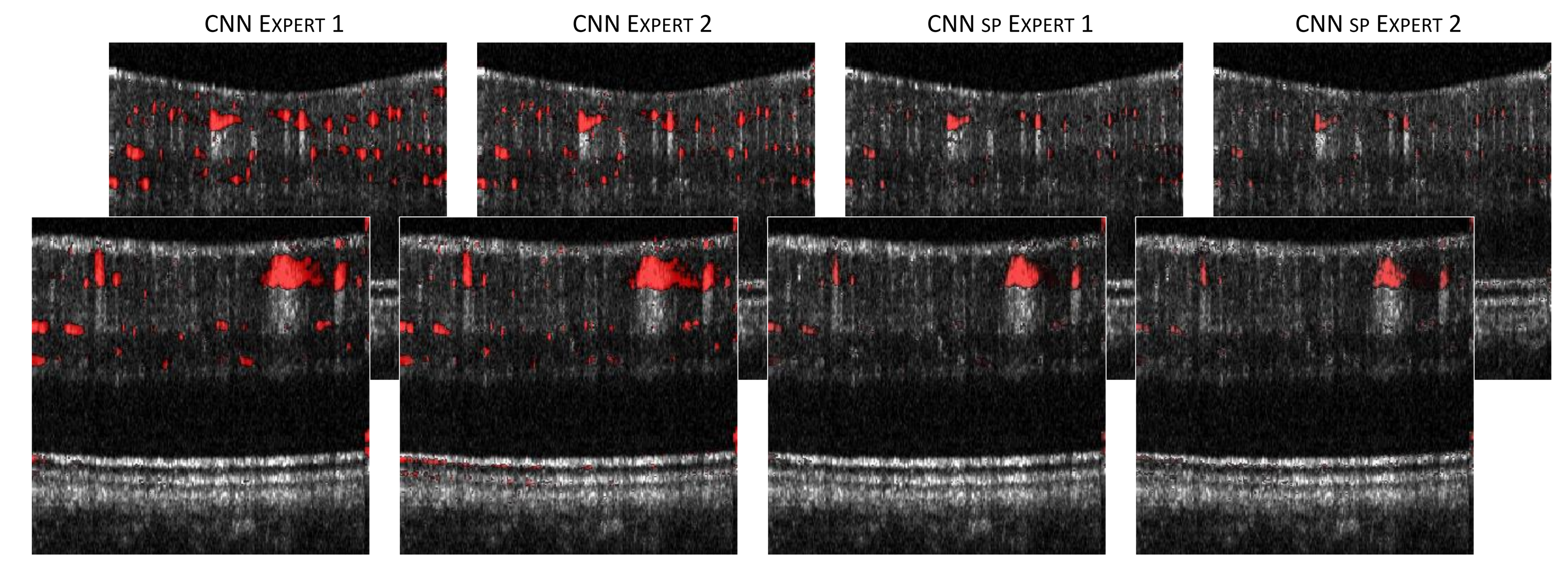
1. **Expert 1** (training part) were tested on voxels labeled Expert 1 (testing part).
2. **Expert 2** (training part) were tested on all voxels labeled by Expert 1.

Test set is always disjoint from the training set.

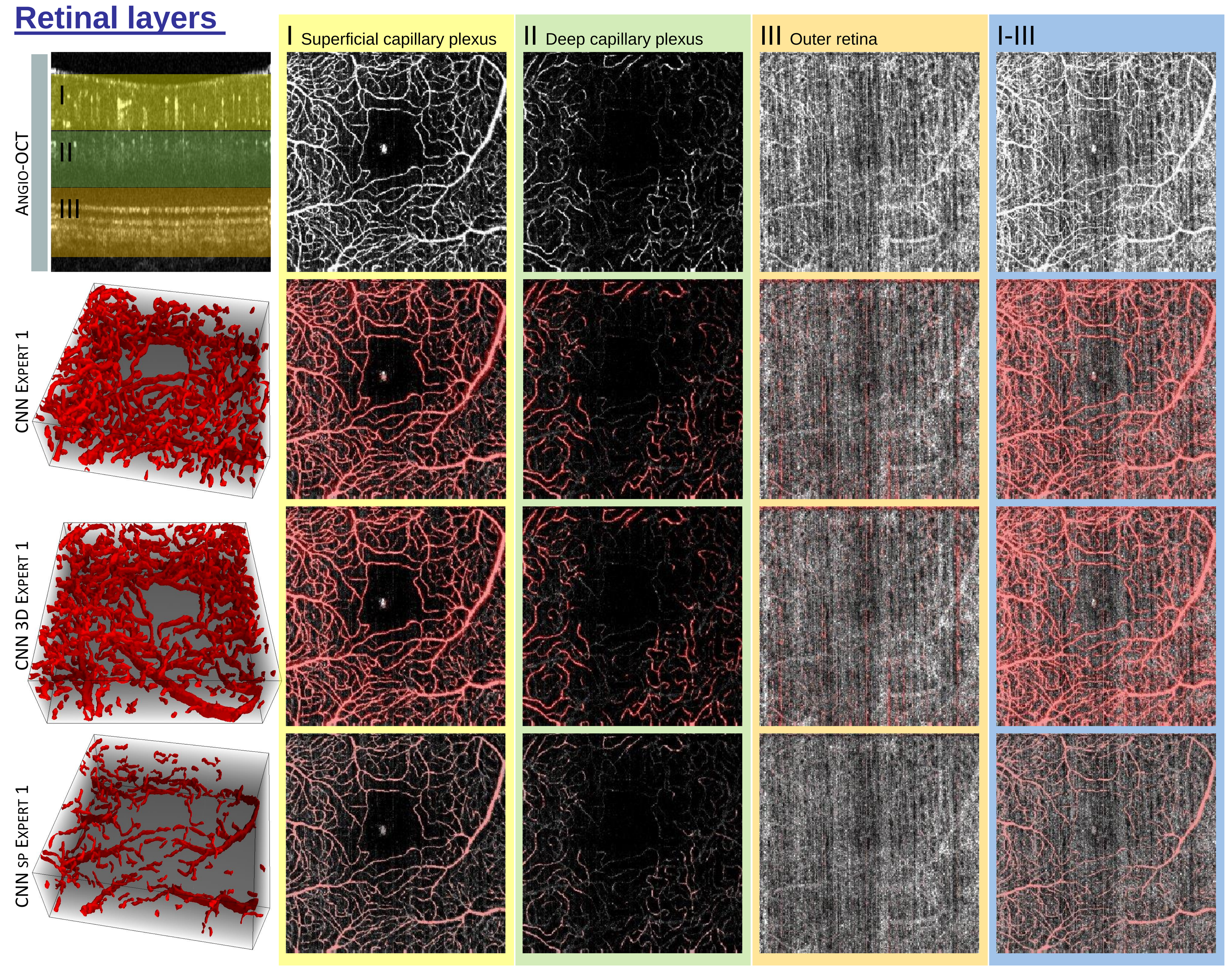
| Model | | AUC | Acc | Sens | Spec |
|---|---|---|---|---|---|
| OCT angio (baseline) | | 82.53% | 73.24% | 73.28% | 73.20% |
| CNN | Expert 1 | 99.99% | 99.82% | 99.11% | 99.93% |
| CNN | Expert 2 | 99.74% | 98.92% | 94.95% | 99.53% |
| CNN 3D | Expert 1 | 99.99% | 99.86% | 99.28% | 99.95% |
| CNN 3D | Expert 2 | 99.72% | 99.03% | 95.78% | 99.53% |
| CNN SP | Expert 1 | 99.97% | 99.77% | 99.00% | 99.89% |
| CNN SP | Expert 2 | 99.38% | 98.40% | 93.69% | 99.12% |

AUC – area under ROC curve; Acc – accuracy of classification; Sens – sensitivity; Spec – specificity. For Acc, Sens and Spec, network output thresholded at 0.5.

### Inter-expert repeatability




### Retinal layers




## Conclusions

Our results shows that CNN approach to RMN visualization provides accurate microcapillary detection and segmentation in 3 dimensions, incorporating a priori knowledge of skilled specialists, and allows for increased sensitivity and specificity of SOCT based angiography. It also is able to differentiate between vessels and the angiographic shadow below the vessels leading to artifact-free Angio-OCT images.

## Acknowledgements

This work was supported by National Centre for Research and Development Grant No. PBS1/A9/20/2013 and by the Foundation for Polish Science (grant TEAM TECH/2016-2/13).
None of the authors has commercial interest.

European Funds Smart Growth
FNP Foundation for Polish Science
European Union European Regional Development Fund