A lightweight model for CT segmentation of maxillary sinuses in children with chronic rhinosinusitis

Przemysław Olbratowski[1,2], Aldona Ząber-Kucharska[3], Józef Ginter[4], Natalia Gołuchowska[3], Bolesław Kalicki[3,4], Agata Tomaszewska[3]

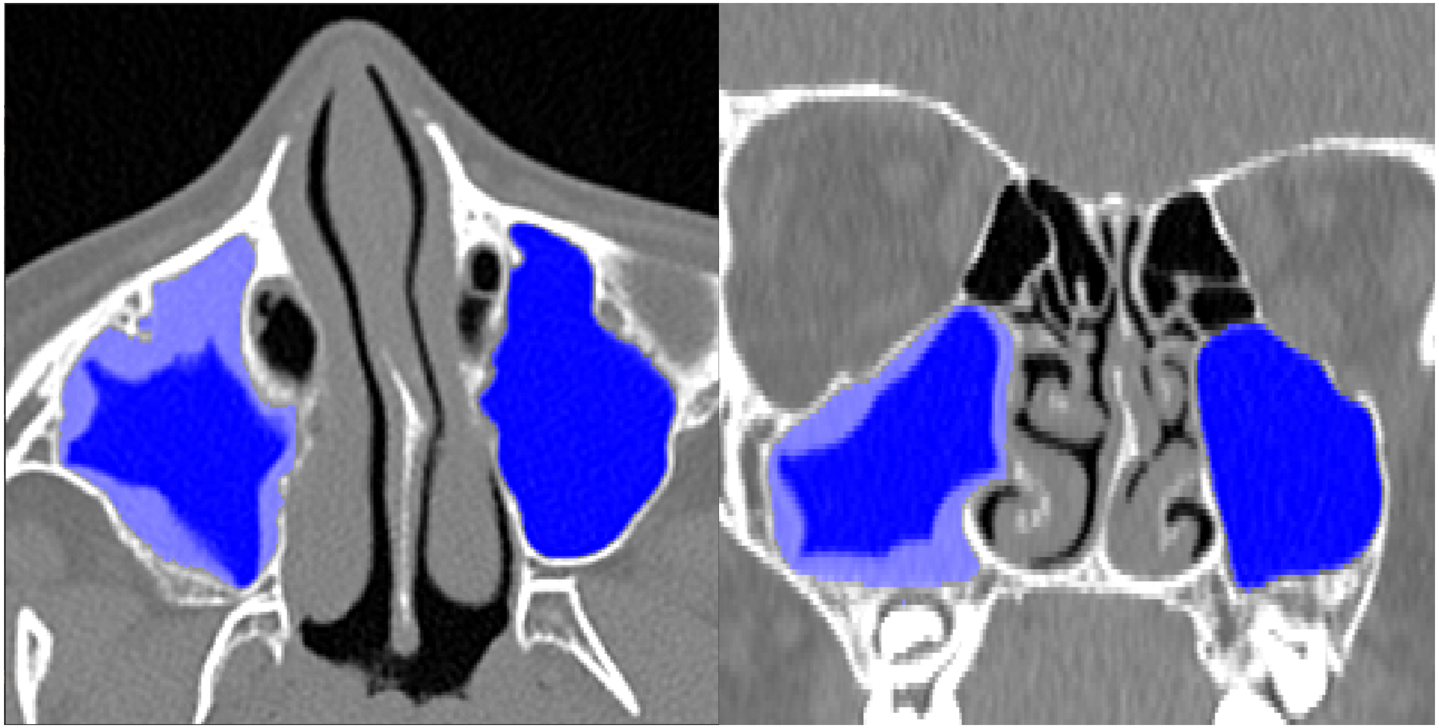
1) Faculty of Physics, University of Warsaw, Warsaw, Poland
2) Biological and Chemical Research Centre, University of Warsaw, Warsaw, Poland
3) Department of Pediatrics, Nephrology, and Allergology Military Institute of Medicine - National Research Institute, Warsaw, Poland
4) Faculty of Medicine, University of Warsaw, Warsaw, Poland

3rd International Workshop on Machine Learning and Quantum Computing Applications in Medicine and Physics, 7 September 2026, Warsaw, Poland

Chronic rhinosinusitis in children

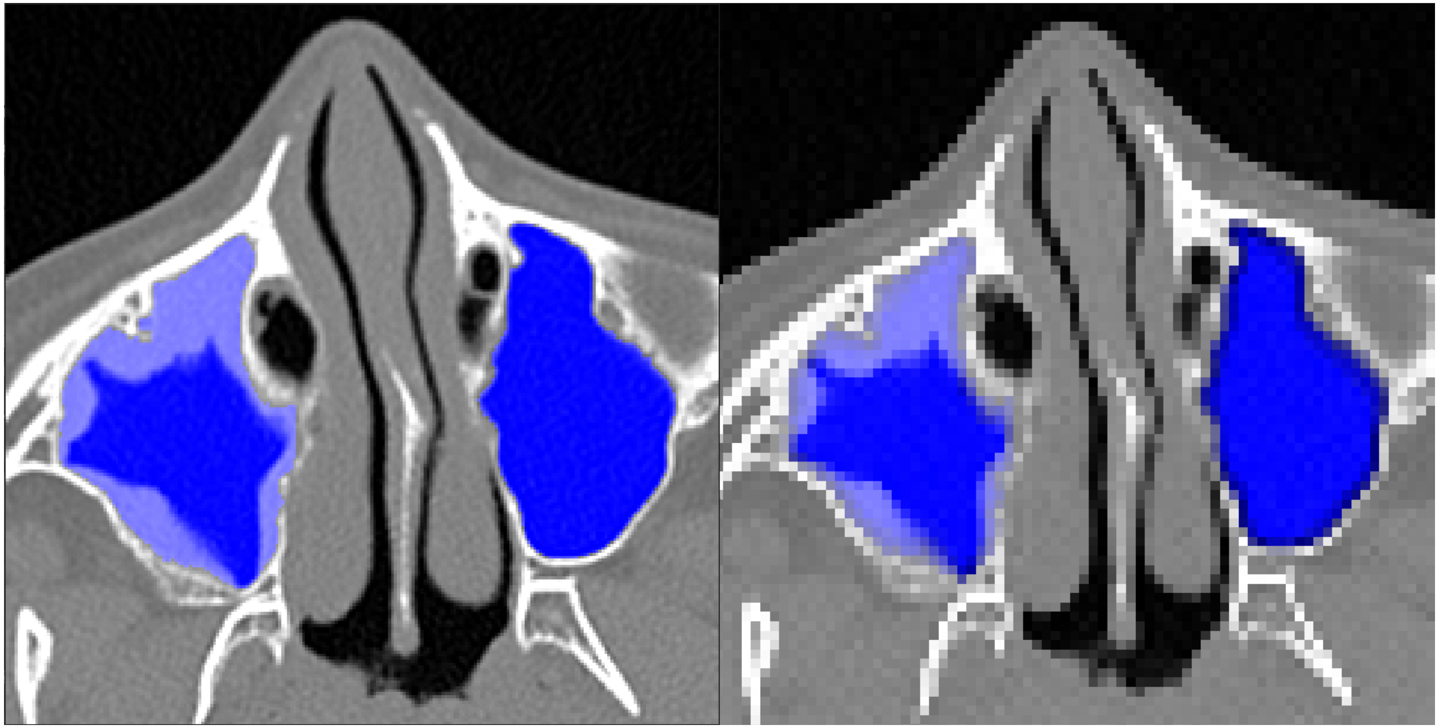
- In chronic rhinosinusitis sinuses are partially filled with mucus
- The goal is to compute the total volume of each maxillary sinus and its fraction filled with mucus
- Pediatric cases are more difficult but very important

Segmentation task



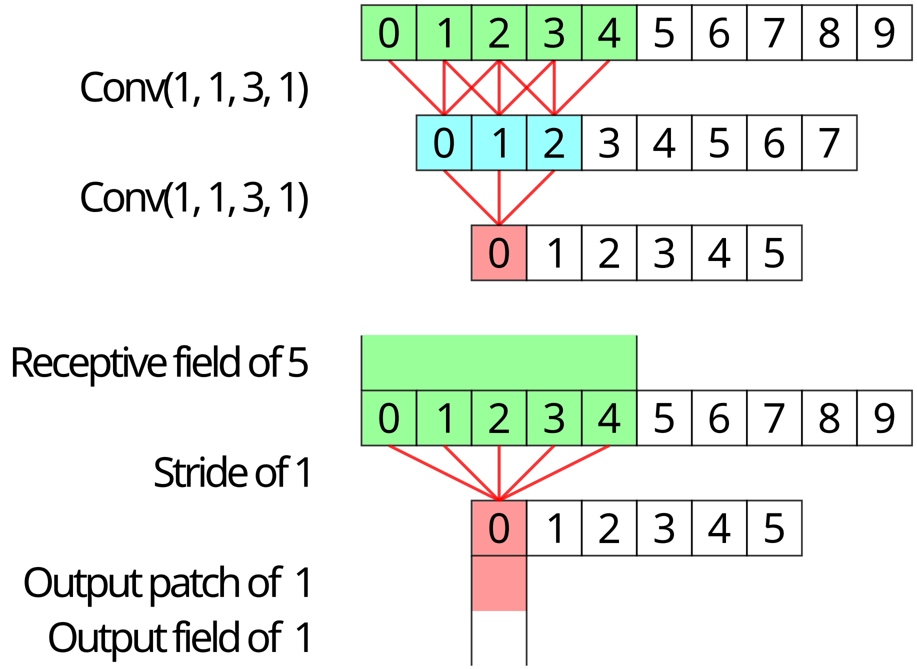
- A neural network segments maxillary sinuses as one class
- Left and right are distinguished by diagonalizing the inertia tensor
- Free and congested regions are identified based on HU values
- Five human experts segmented 92 scans using 3D Slicer

Fractional labels in reduced resolution



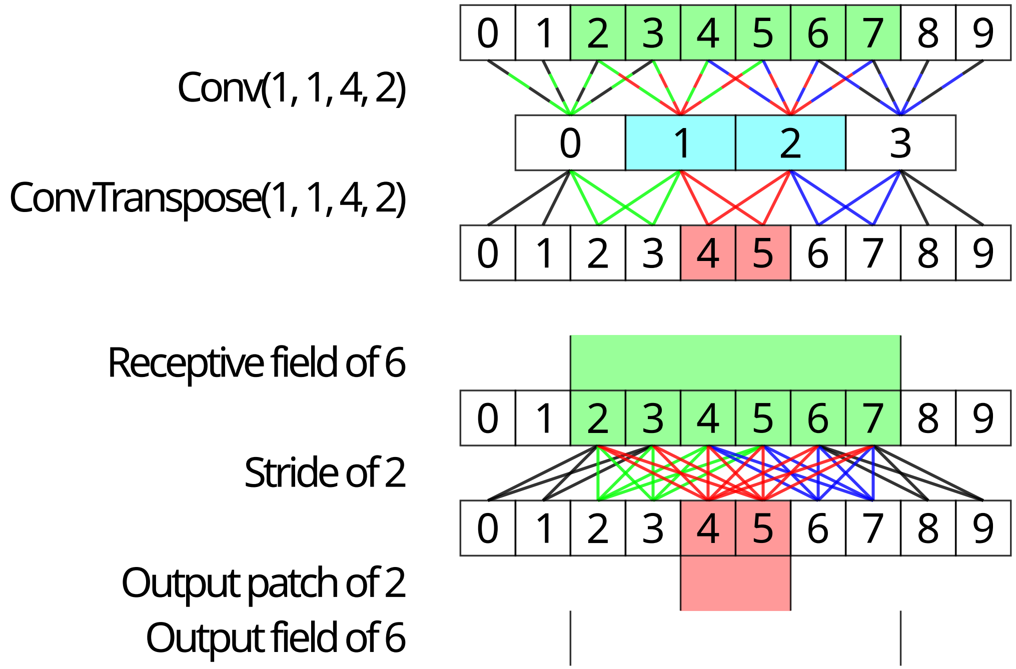
- CT images are about 512×512×448 with 0.3×0.3×0.6 mm³ voxels
- They are scaled down to 192×192×168 with 1×1×1 mm³ voxels
- Fractional labels account for the partial-volume effect

Receptive field



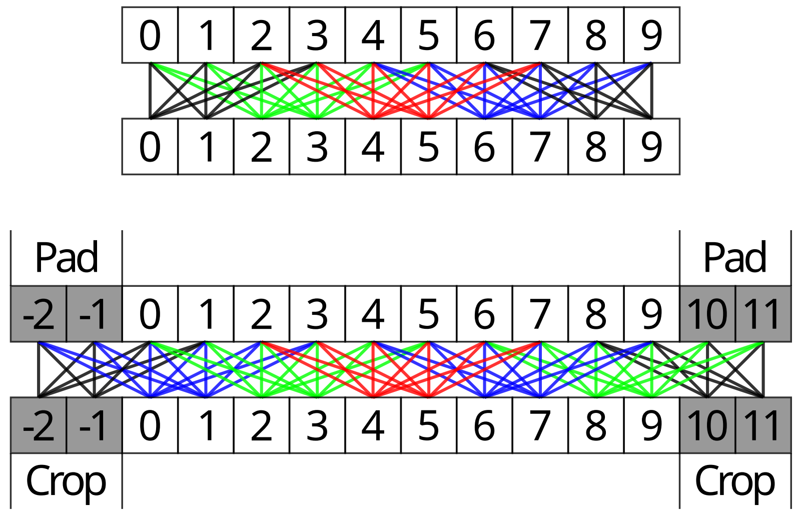
- Receptive field is a set of input voxels seen by one output voxel
- Large architectures often have different receptive fields for different output voxels
- We aim at equal receptive field for each output voxel

Output patch and field



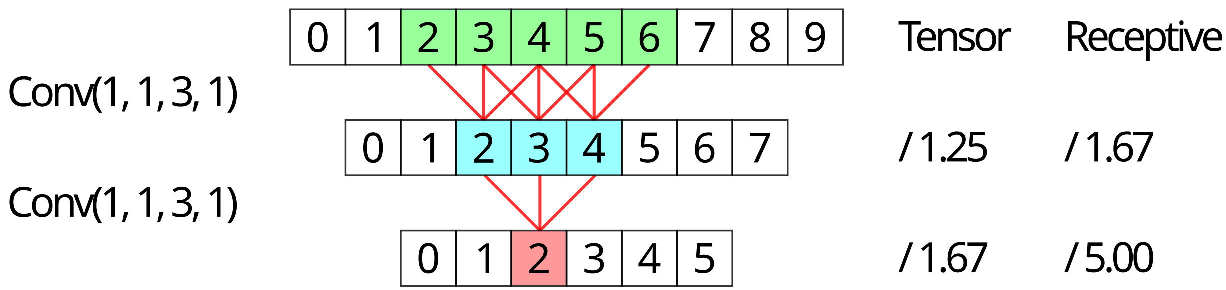
- Output patch is a set of output voxels that depend on the same receptive field and only on that field
- Output field is what the model outputs when fed with a single receptive field

Padding and cropping



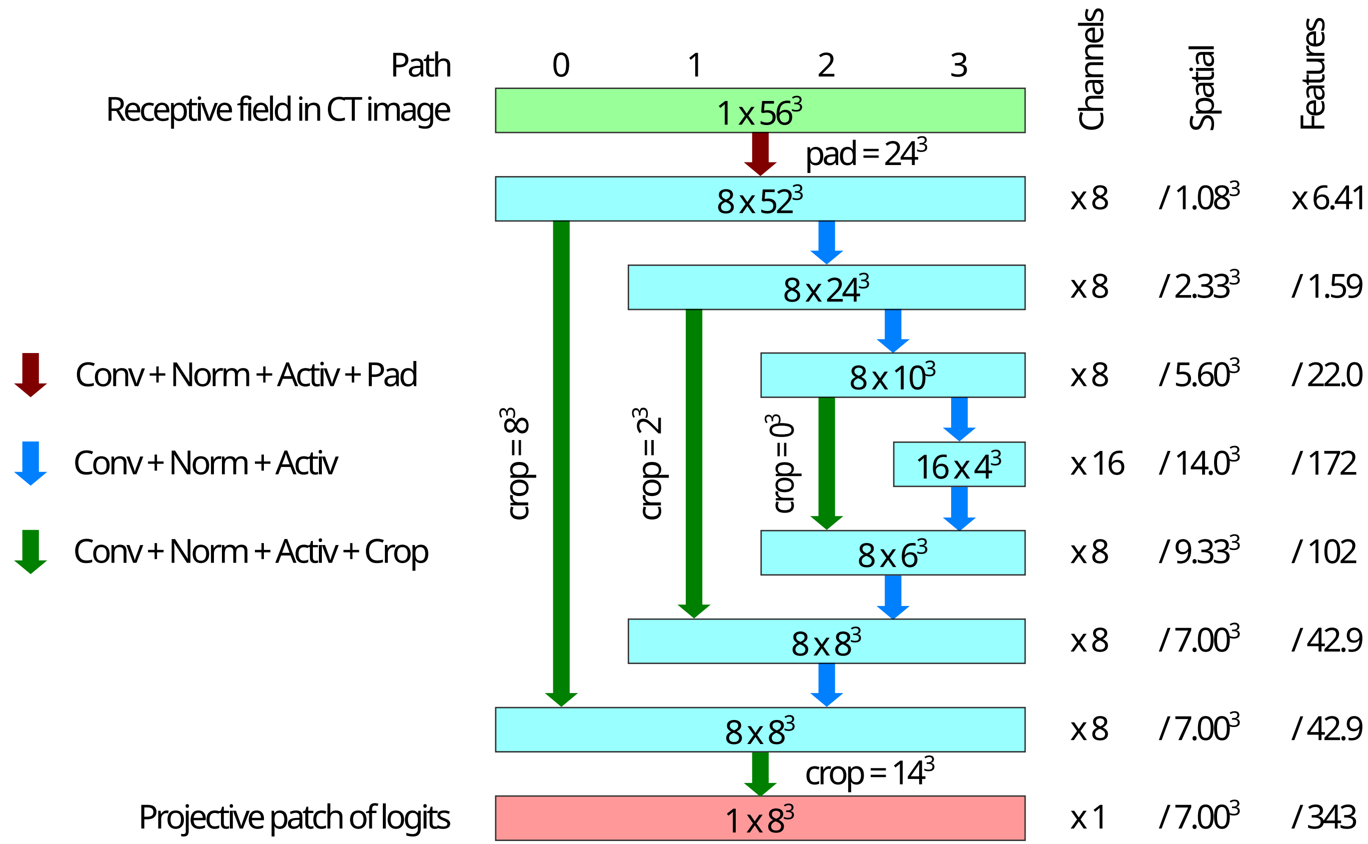
- Large architectures use zero-padding in every convolutional layer to maintain spatial dimensions, which inserts meaningless zeros
- Our model applies zero-padding only in the input layer where it is equivalent to the image background

Receptive features



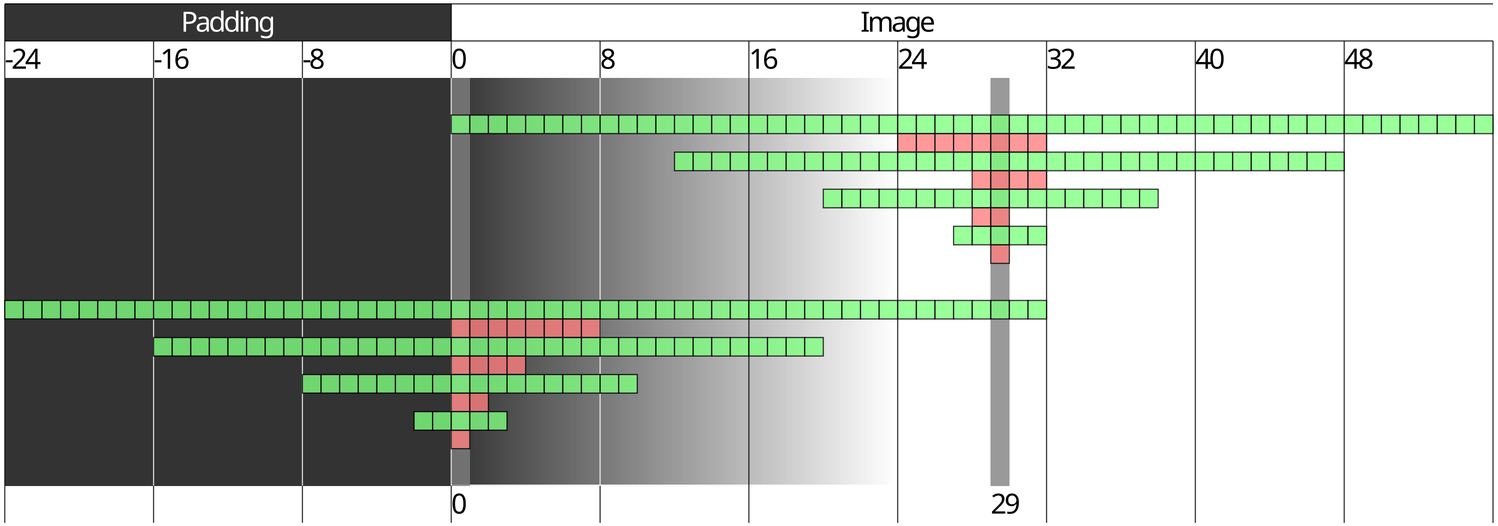
- Adjusting the number of channels involves counting relevant features in each layer
- People often consider just the tensor shapes
- We count receptive features affecting a single output patch
- We adjust the channel count to first inflate and then shrink the set of receptive features

Architecture



- 16 convolutional layers comprising 39,745 parameters
- 4 receptive fields and output patches corresponding to the 4 paths

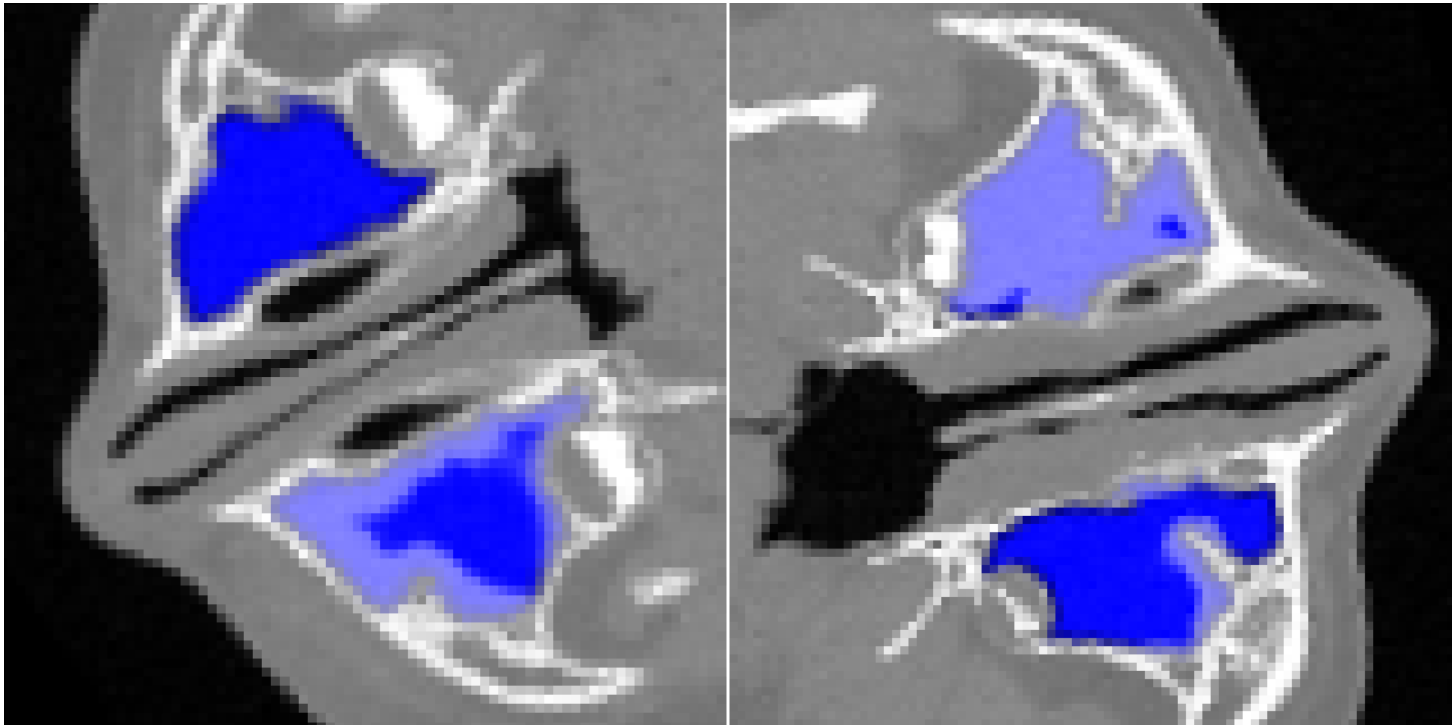
Paths



Path	0	1	2	3
Receptive field	5^3	18^3	36^3	56^3
Output patch	1^3	2^3	4^3	8^3

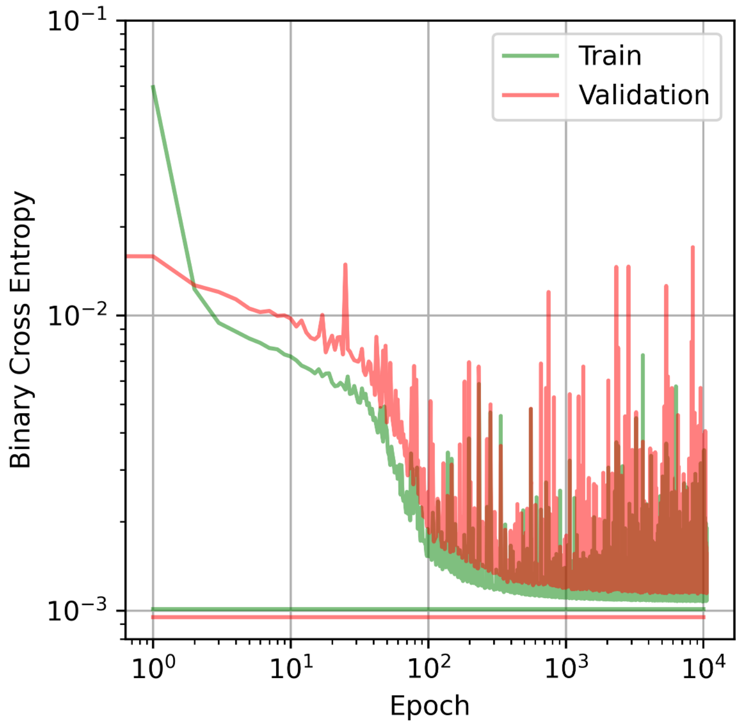
- Each output voxel receives information at 4 spatial scales
- 5^3 ensures voxel-level precision while 56^3 provides broad context
- Each receptive field is as centered at the output voxel as possible

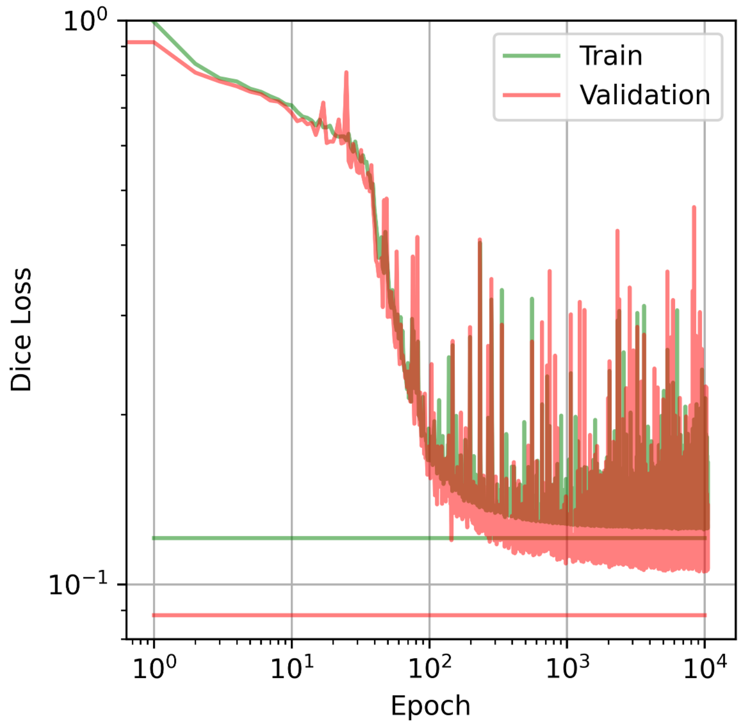
Data augmentation



- Random translations and arbitrary rotations about the body axis
- The model is trained to process scans at any axial orientation
- With fuzzy labels cross entropy and metrics have Shannon limits
- These limits are different for original and augmented images

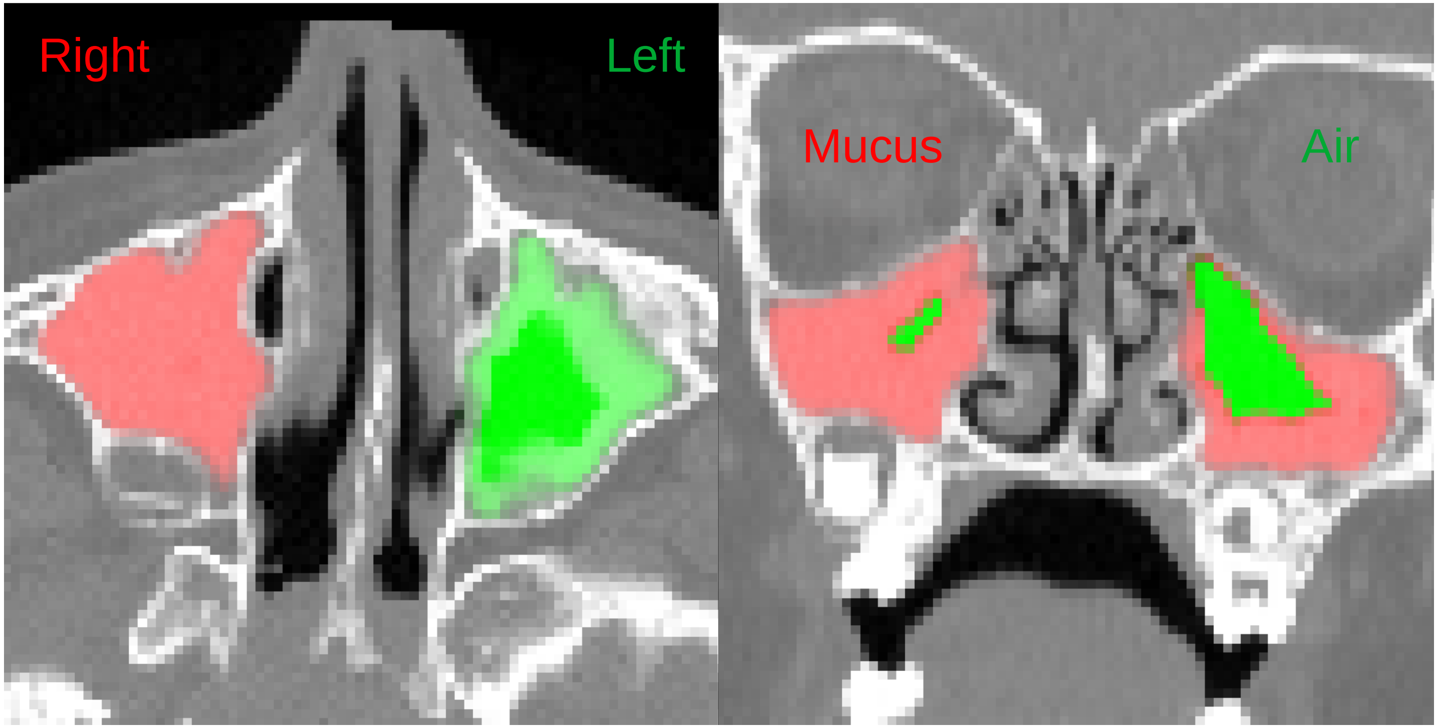
Training





- Pure cross-entropy because Dice tends to yield hard labels
- About 15 hours per 1000 epochs on a single gaming-class GPU
- About 10.000 epochs or 6.25 days needed for best results

Performance



- Sinus volume RMSE is 0.5 cm^3 with an average volume of 12 cm^3
- Involvement fraction RMSE is 1.5 percentage points
- Prediction takes a few seconds on CPU

Publication of the previous version

Aldona Ząber-Kucharska[1], Przemysław Olbratowski[2–4], Józef Ginter[5], Natalia Gołuchowska[1], Bolesław Kalicki[1,5], Agata Tomaszewska[1,5]

Received: 22.03.2026
Accepted: 19.05.2026
Published: 09.07.2026

A machine-learning model for quantitative assessment of maxillary sinus volume and involvement in children based on computed tomography

Algorytm uczenia maszynowego do automatycznej oceny objętości i stopnia zajęcia zatok szczękowych u dzieci na podstawie badań tomografii komputerowej

[1] Department of Paediatrics, Paediatric Nephrology and Allergology, Military Institute of Medicine – National Research Institute, Warsaw, Poland
[2] Faculty of Physics, University of Warsaw, Warsaw, Poland
[3] Biological and Chemical Research Centre, University of Warsaw, Warsaw, Poland
[4] WIT Academy, Warsaw, Poland
[5] Faculty of Medicine, University of Warsaw, Warsaw, Poland
Correspondence: Aldona Ząber-Kucharska, Department of Paediatrics, Paediatric Nephrology and Allergology, Military Institute of Medicine – National Research Institute, Szaserów 128, 04-141 Warsaw, Poland, e-mail: aldonazaber@gmail.com

https://doi.org/10.15557/PiMR.2026.0012

- Ząber-Kucharska A, Olbratowski P, Ginter J, Gołuchowska N, Kalicki B, Tomaszewska A: *A machine-learning model for quantitative assessment of maxillary sinus volume and involvement in children based on computed tomography*
  Pediatr Med Rodz 2026; 22 (2): 76–82

Thank you for your attention!

This project was financed by the Medical Research Agency under contract 2023/ABM/02/00014

Establishment of the Digital Medicine Centre

of the Military Institute of Medicine - National Research Institute and the University of Warsaw

MEDICAL RESEARCH AGENCY

UNIVERSITY OF WARSAW

MILITARY INSTITUTE OF MEDICINE
NATIONAL RESEARCH INSTITUTE

Photo of the conference venue from the conference website