EU-project „Rhinodiagnost“

The Segmentation of CT Scans using Artificial Intelligence

RHINO DIAGNOST
Morphological and functional precision diagnostics of the nose

Walter Koch, Matthias Lehner

Rhinodiagnost Final Symposium, Dec. 9, 2020

Angewandte Informationstechnik Forschungsgesellschaft mbH

Team: Odo Benda, Gerda Koch, Walter Koch , Matthias Lehner, Ramiro Ortiz, Jutta Stockklauser

Acknowledgments: Prim. Univ-Prof.Dr.med. Gerhard Ranner, Dr.med. Jochen Schachenreiter, Univ.-Prof.Dr.Dietmar Rafolt and Nikolaus Foidl† for providing CT data sets, their medical, biomedical engineering and medical physics expertise.

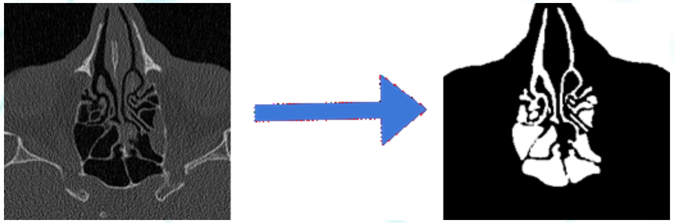
RHINO DIAGNOST

Automatic Reconstruction of the Nasal Geometry from CT Scans

Manual segmentation of head CT/MR scans of a single patient

- can take up to **24 hours**
- requires constant concentration > prone to errors even when experts are involved in the process.

BUT: in the everyday context, multiple sets of CT/MR segmentations of various patients have to be produced in short time.



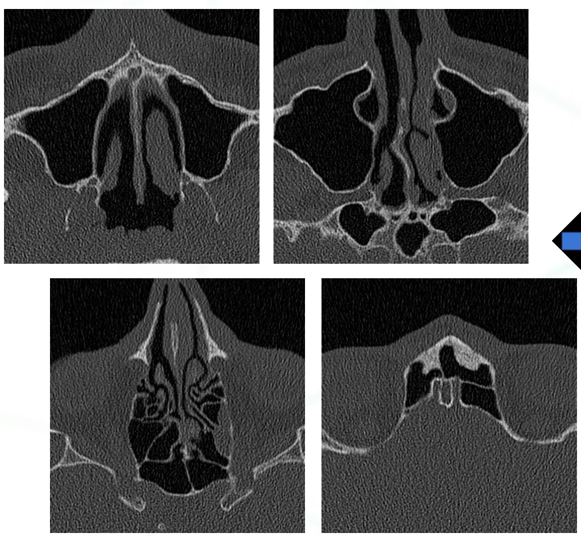
Furthermore:

- CT/MR scans are a **simplification** of the real situation (a two dimensional representation of tridimensional real models)
- Although they remain a valuable resource and a basis in every branch of medicine, an improvement of their representation power could yield **models closer to the reality**

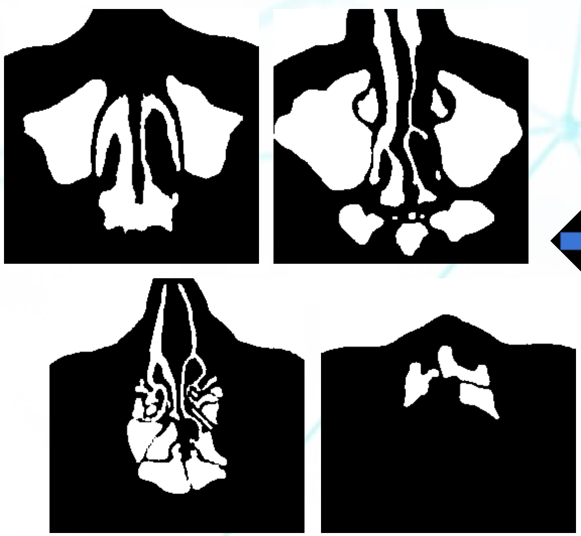
Automatic Reconstruction of the Nasal Geometry from CT Scans

Goal: Generate 3D models from CT scans automatically

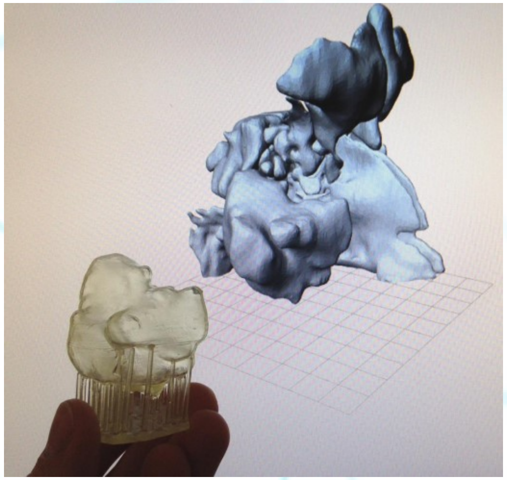
The 3D models can then be used for visualization of different anatomical regions, detection of pathologies, CFD simulations, virtual surgery, 3D printing, surgical training etc.



A set of CT images...

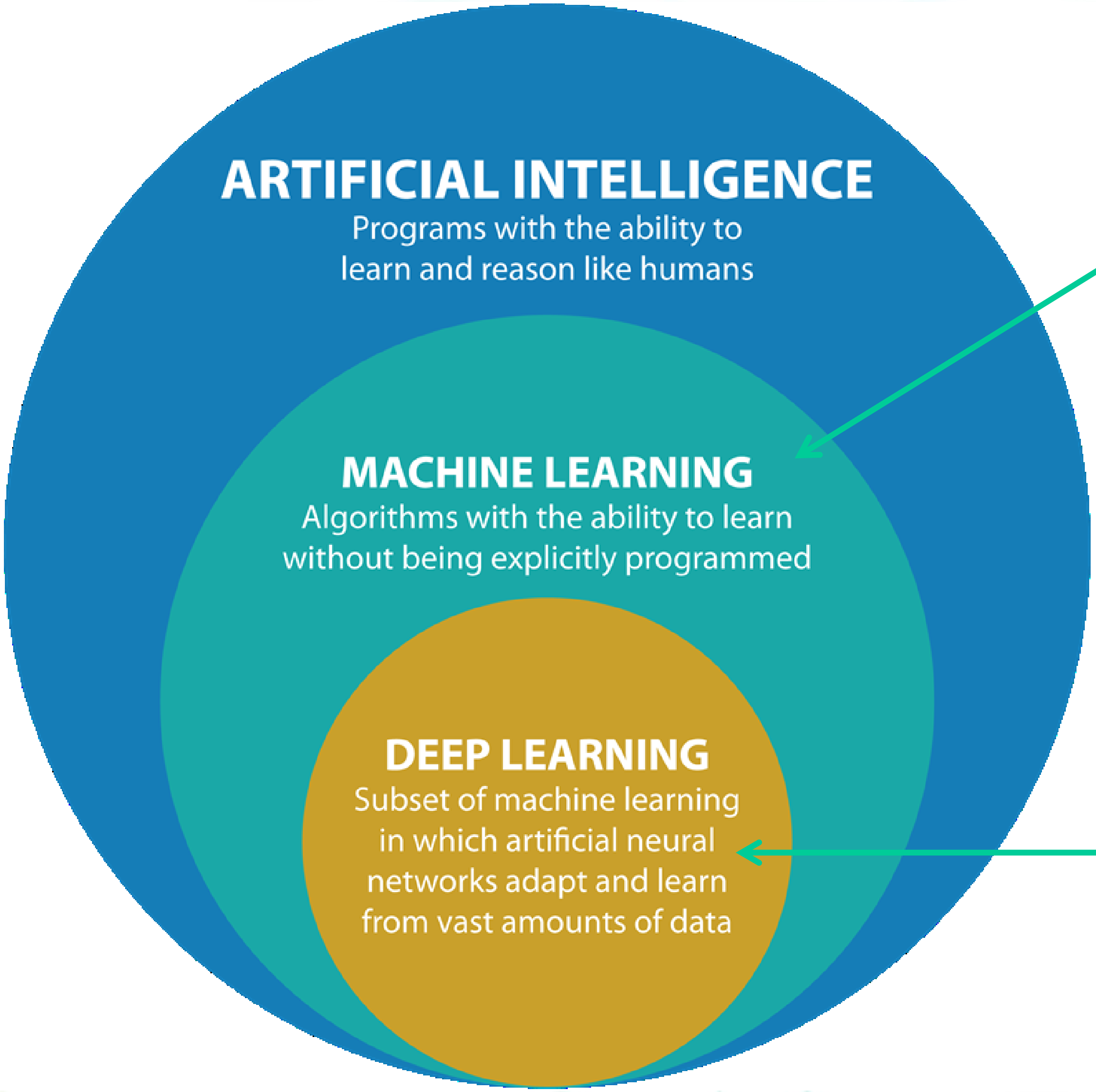


...can be used to train a Convolutional Neural Network which then segments new CT images automatically



From the 2D segmentations, digital and physical 3D models can be generated

Neural Networks for Image Segmentation



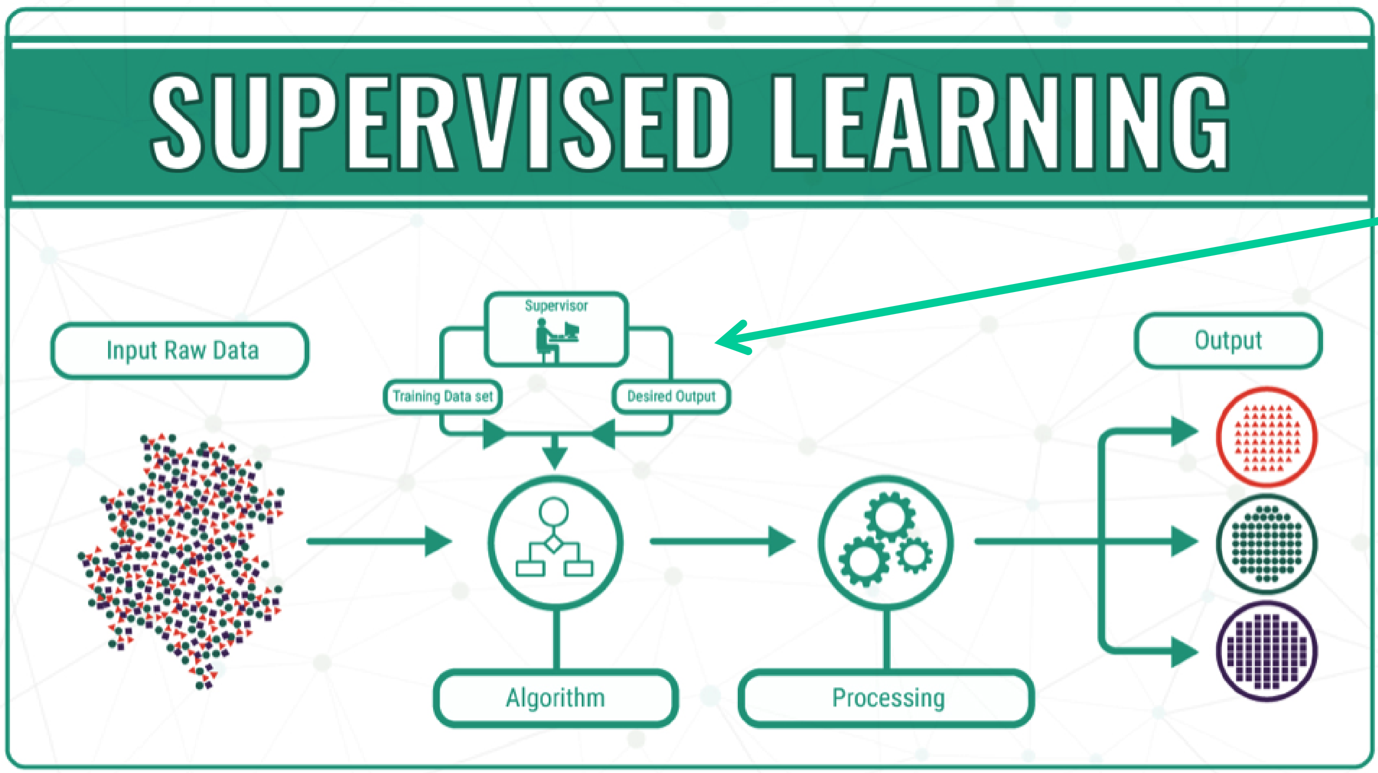
Machine Learning is a field of computer science that uses statistical techniques to give computers the ability to "learn" (i.e. progressively improve performance on a specific task) from data without being explicitly programmed.

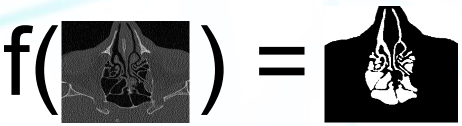
A **Convolutional Neural Network (CNN)** is a particular type of **Artificial Neural Network (ANN)** which was initially developed for **Computer Vision**, i.e. with the goal to **understand visual data** (images and videos).

What is a Convolutional Neural Network?

A Convolutional Neural Network is a Deep Learning algorithm which can take in an input image, assign importance (**learnable** weights and biases) to various aspects/objects in the image and is able to differentiate one from the other.

CNNs can be trained using **supervised learning**.
This means that in order to learn, a CNN needs to receive not only **input data** but also the **expected results**, i.e. the **desired output**.





1 = white (air)

0 = black (not air, i.e. bone & tissues)

Machine Learning Pipeline

1) Collect and preprocess the data needed to train the model

CT images of the nasal cavities and paranasal sinuses (anonymized DICOM files) are obtained from patients from partnering clinics. A researcher creates segmentations for a subset of the images manually.

2) Choose and implement an appropriate algorithm / mathematical model

A CNN with the previously depicted architecture is used

3) Split the data into training, validation and test sets

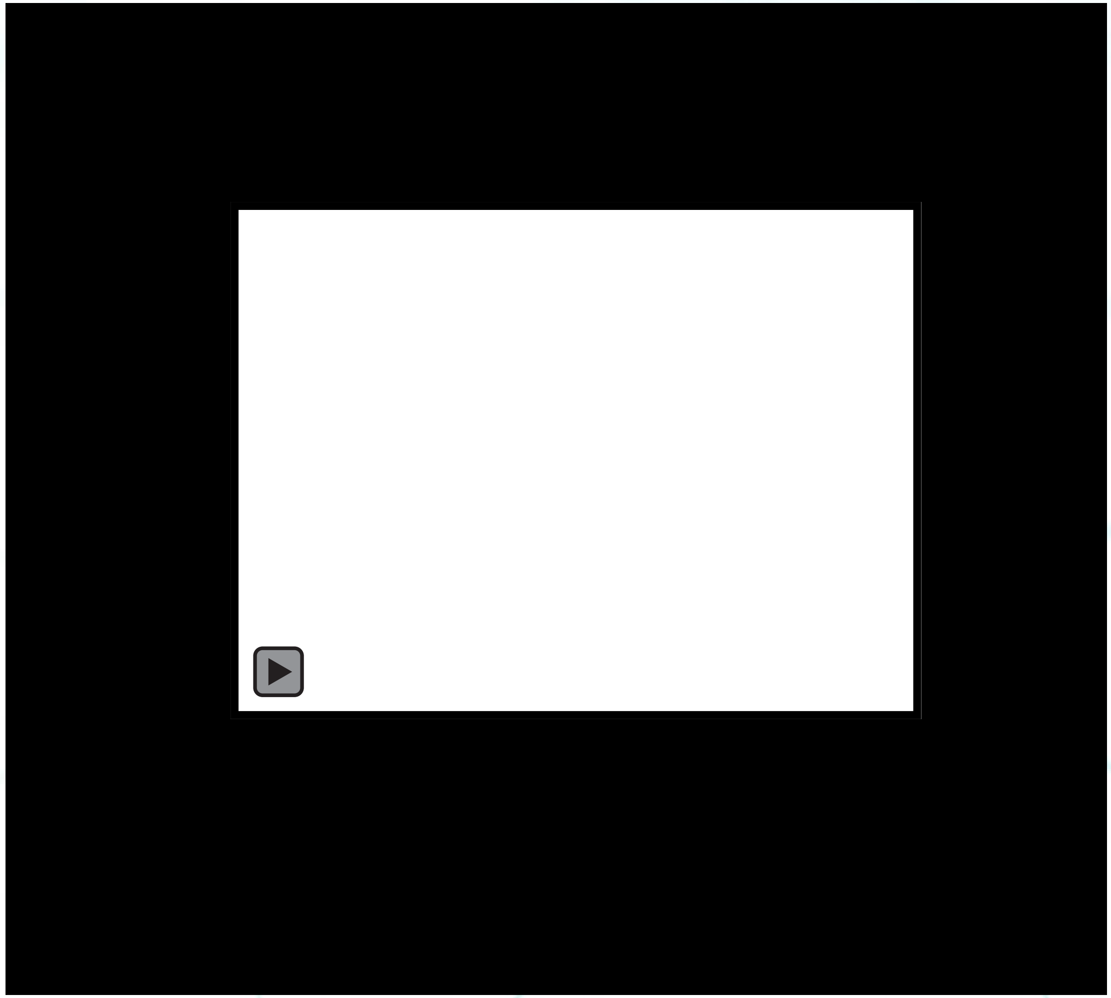
The **training set** is used to train several CNNs. The **validation set** is then used to select the best of these. The **test set** (CT images only) is used to segment images the selected CNN has never seen before.

4) Train models on the training set, select the best one using the validation set

We evaluate the performance using accuracy as a metrics. In our case, **accuracy** is calculated as the percentage of number of correctly classified pixels divided by the total amount of pixel in the image. Therefore, it ranges between 0% (all pixels wrongly predicted) to 1% (perfect prediction)

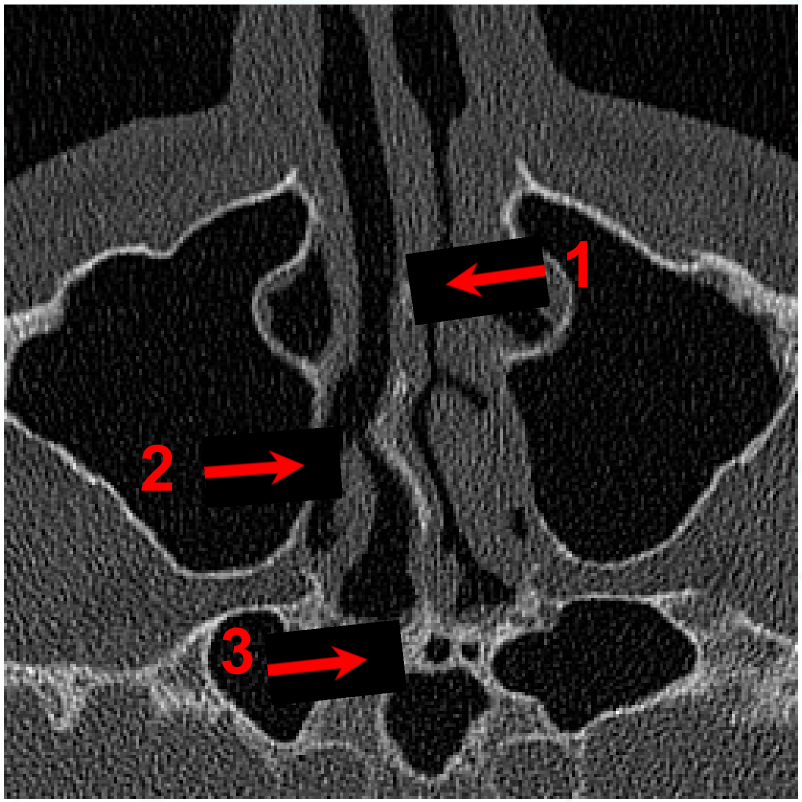
5) Use the test set to confirm the actual predictive power of the selected model

Video of Automated Air Segmentation

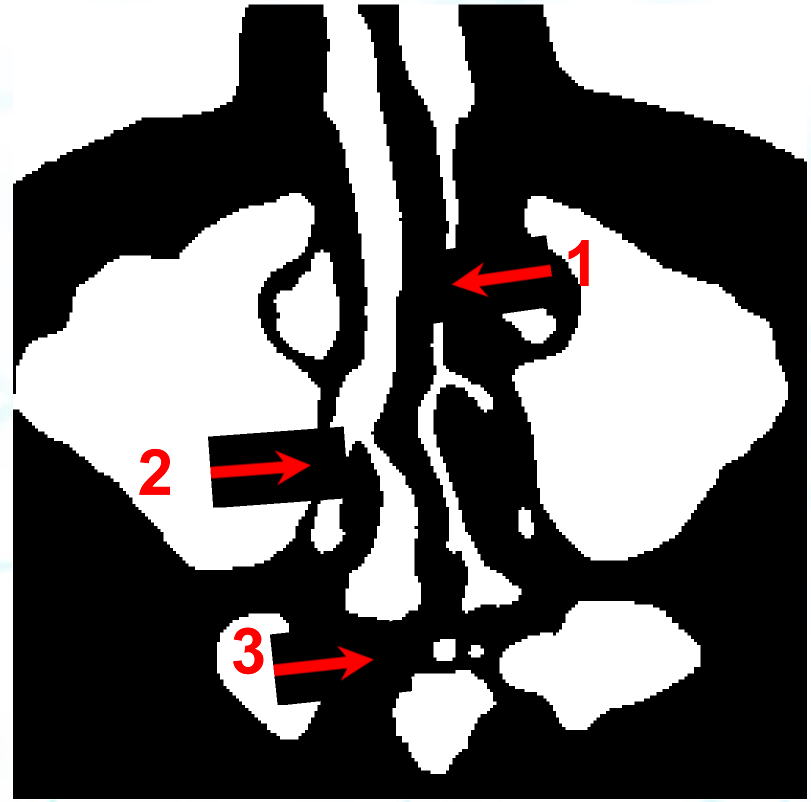


This video shows how the CNN learns to segment an input CT image better over time.
The video shows the first ~100 iterations of the learning process.

Quality of Air Segmentation



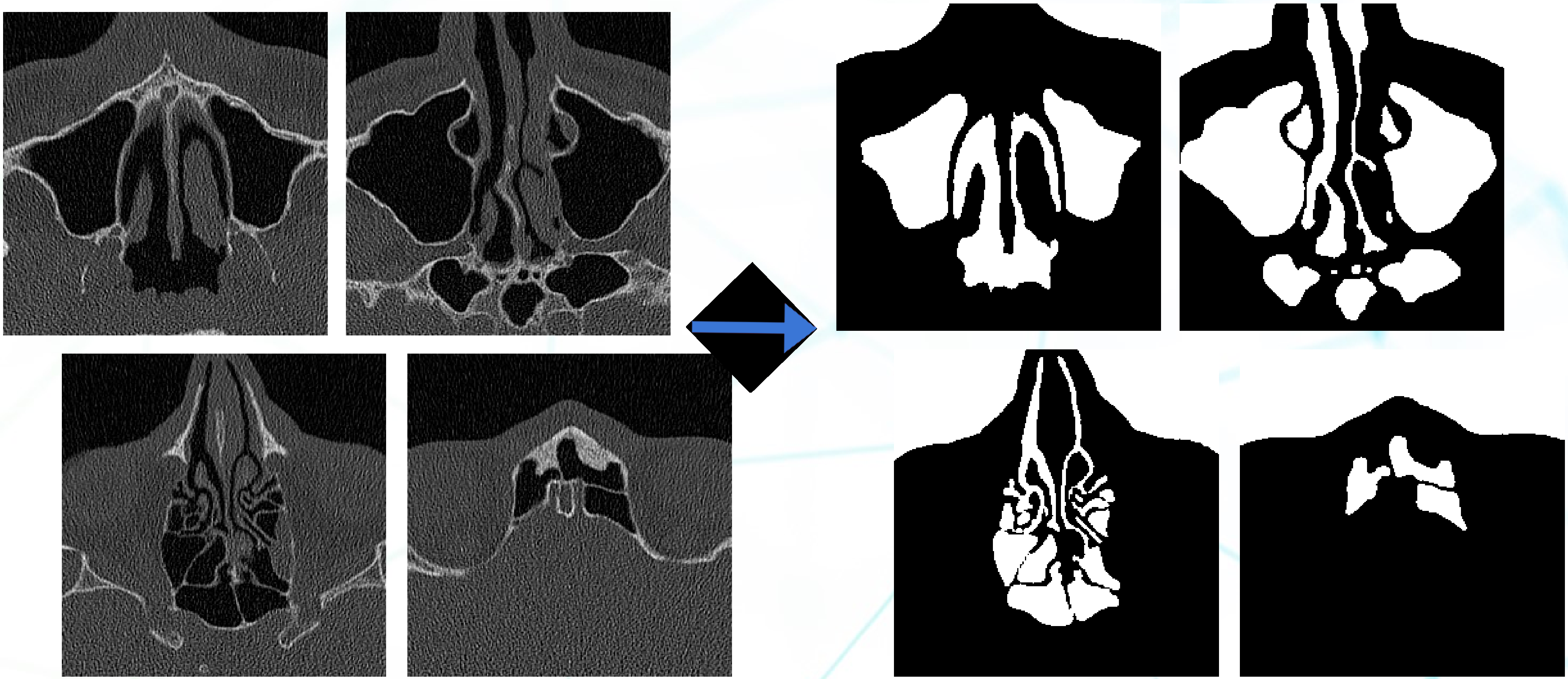
Input: **CT Image**



Output: **Binary Segmentation**

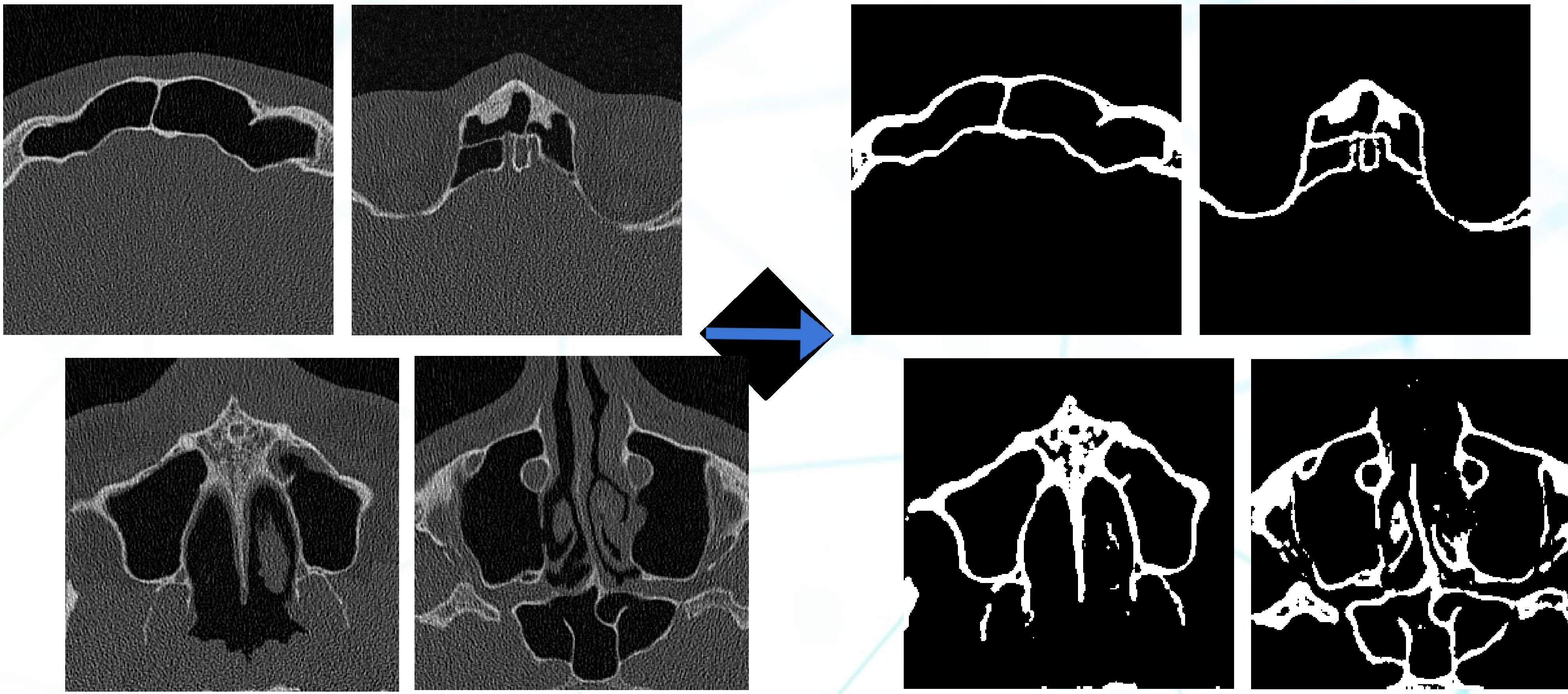
Notice how the CNN segments the thin, delicate pathways correctly (1)
It also keeps sinus and meatus separate despite the very thin wall between them (2)
Further notice that it detects the small black dots in the CT image on the left as well (3)

Quality of Air Segmentation



The results are still quite accurate even for the complex structures in the bottom left image.
Compared to a human annotator with quality control by an experienced ENT surgeon,
the CNN achieves an accuracy of 99+% for air segmentation.
This means that a trained person and the CNN classify almost all pixels equally.

Quality of Bone Segmentation



Bone segmentation is much more difficult as bone is much harder to see in CT images.

The intensity values of thin bones is highly similar to the values of surrounding tissues. Even for a trained human, it is hard and often impossible to know where exactly there is bone or isn't.

For all regions, the accuracy is **95-96%**. For the frontal sinuses, bone is more clearly visible in the CTs, so the accuracy there is about **97%**. Note: Every percentage point can make quite a difference!

Comparison: Human vs Neural Network

CT Image

Human Segmentation

CNN Segmentation

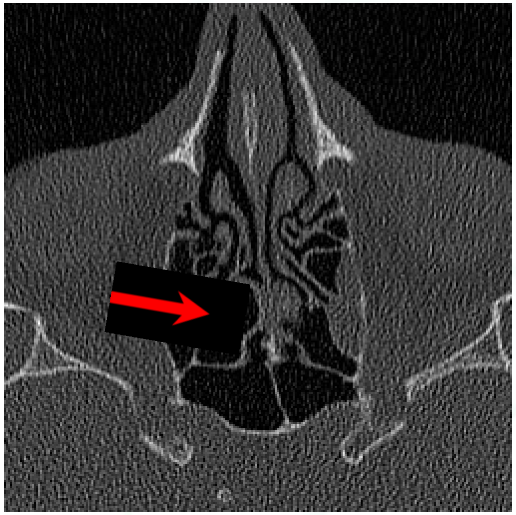
Difference

99.43% accuracy with the CNN used

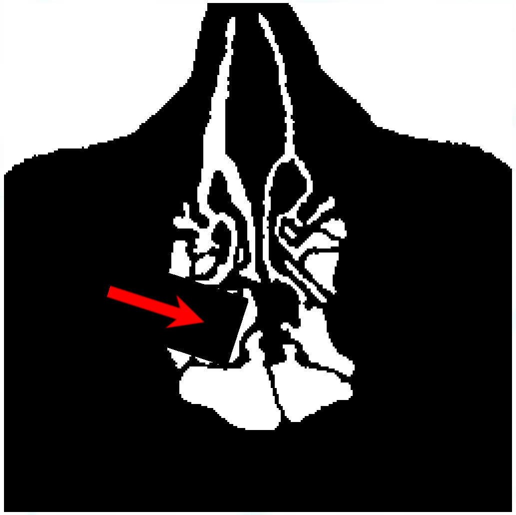
Accuracy = percentage of pixels classified correctly (as compared to the manual segmentation as reference)

We use other metrics as well (not shown).

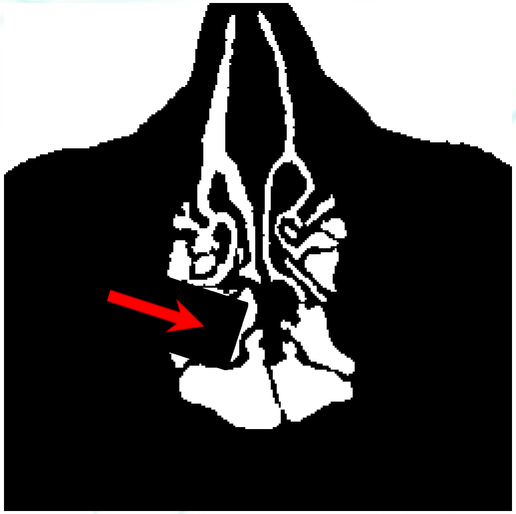
Comparison: Human vs Neural Network



CT Image



Human Segmentation



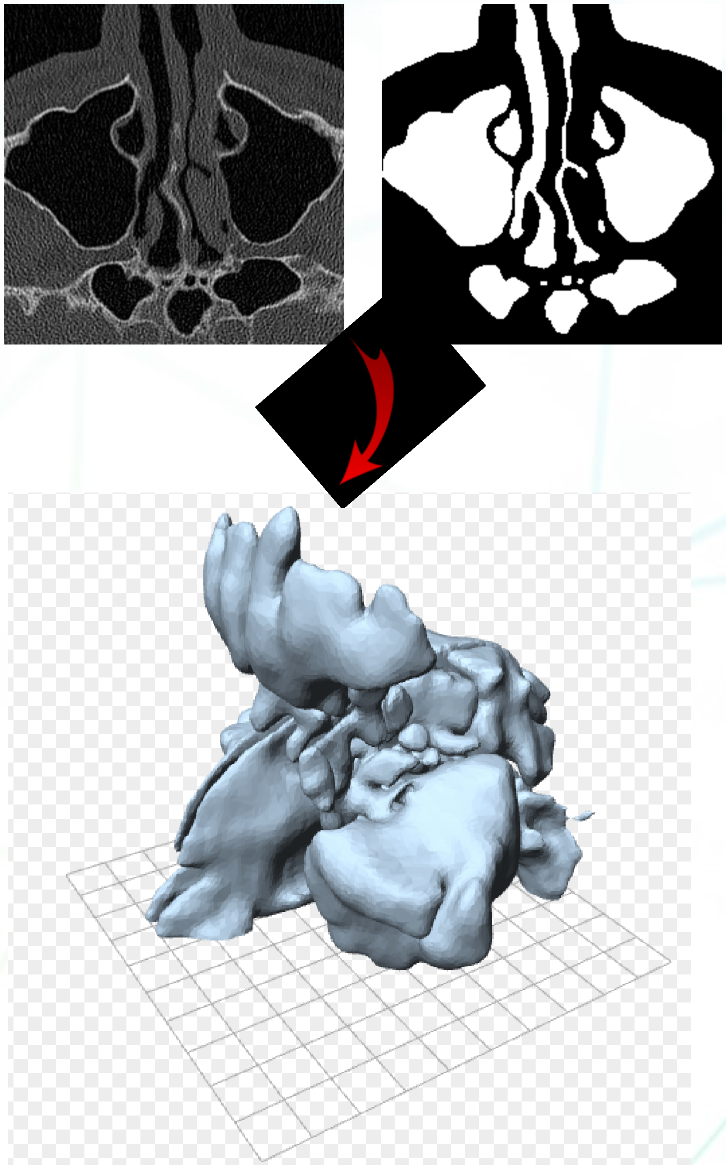
CNN Segmentation

Difference

99.43% accuracy with the CNN used

For some pixels it is hard to decide whether connections between regions exist or not. CT images are noisy and thin structures are more difficult to segment correctly.

The 3D-Model: Marching Cubes Algorithm



The **Marching Cubes** algorithm is used to extracts polygonal mesh from an input isosurface. Using this algorithm, one can generate 3D surface triangulations, from the 2D segmentations of air and bone.

The algorithm is accurate and well suited for surface reconstruction. The total time needed increases linearly with the increase of area.

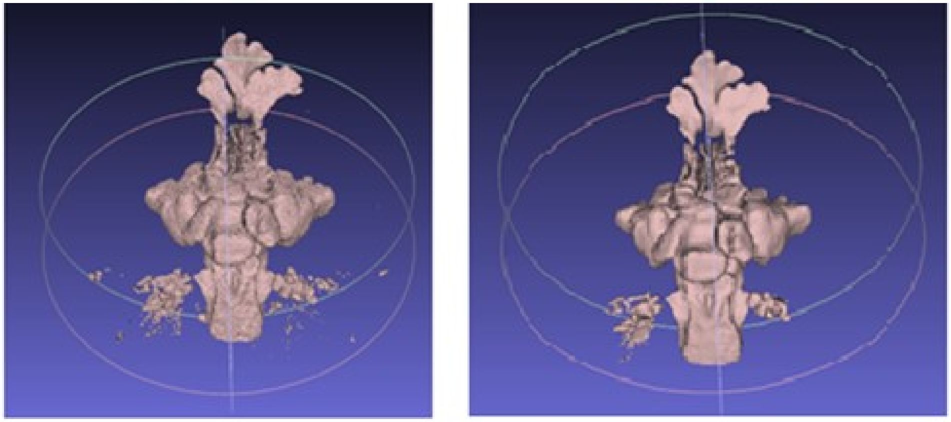
Further application:

- 3D printing
- CFD simulations
- fluid dynamics measurements with an artificial lung
- classify and distinguish different anatomical structures (i.e. different sinuses, left and right), etc.

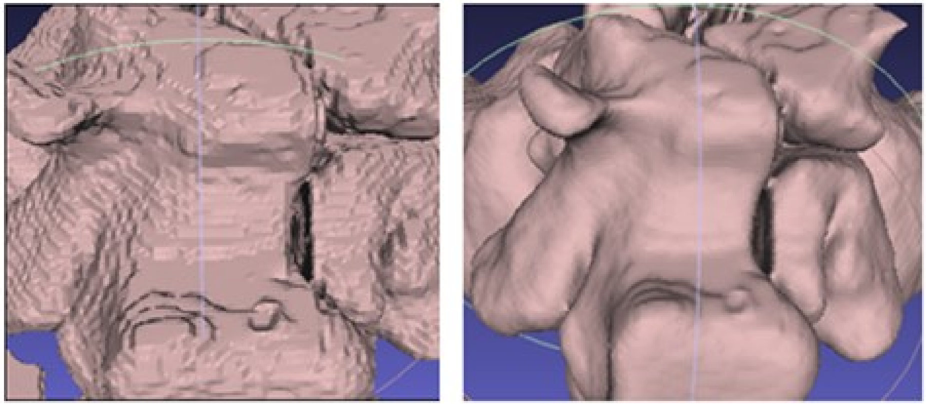
3D Models Postprocessing Steps

The 3D models are further refined by application a series of filters:

- automatically remove unwanted parts (ears, face) in order to isolate nasal cavity and sinuses.
- apply the MinComponentSize filter to remove connected components which are too small.
- apply a Laplacian smoothing filter several times. This filter implements the Laplacian smoothing algorithm which is widely used in computer graphics to smooth polygonal meshes.



After face and ear removal (left), after small connected component removal (right)



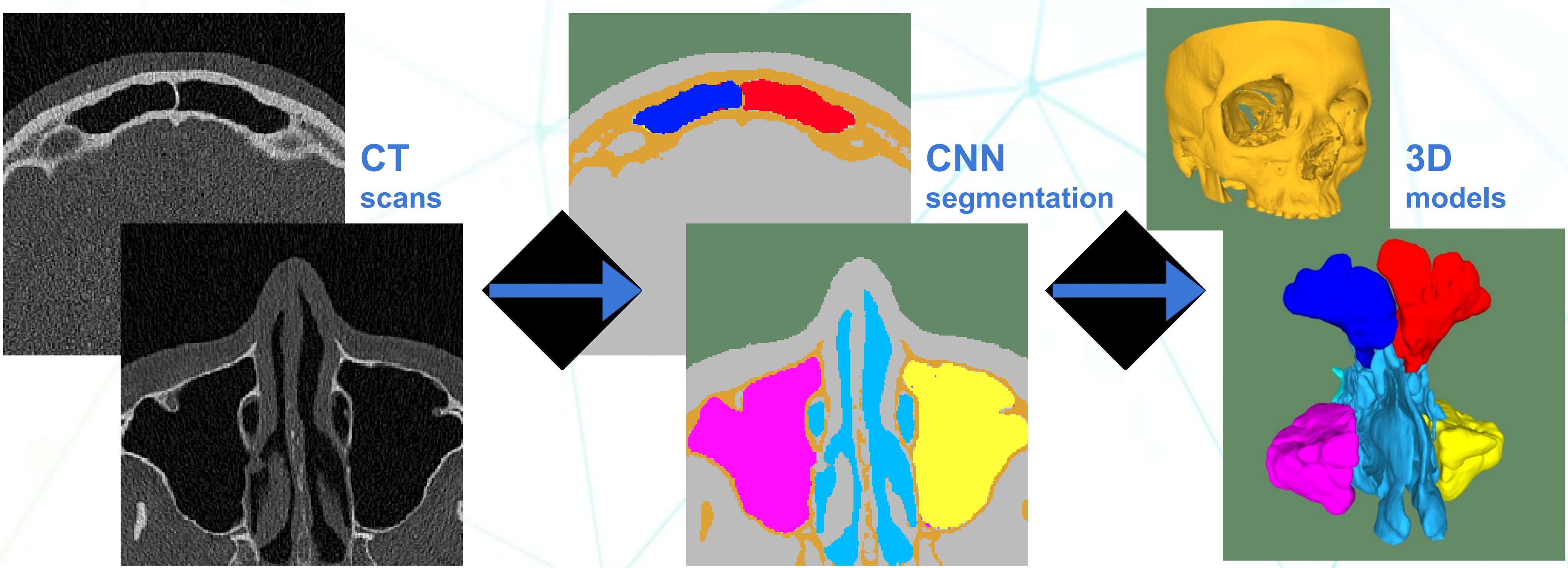
No smoothing (left), after 3 rounds of Laplacian smoothing (right)

All these steps, from the creation of the 3D model to its refinement using filters, do not require human interaction and are performed automatically after the 2D segmentation has been generated by the CNN.

Multi-Class Segmentation

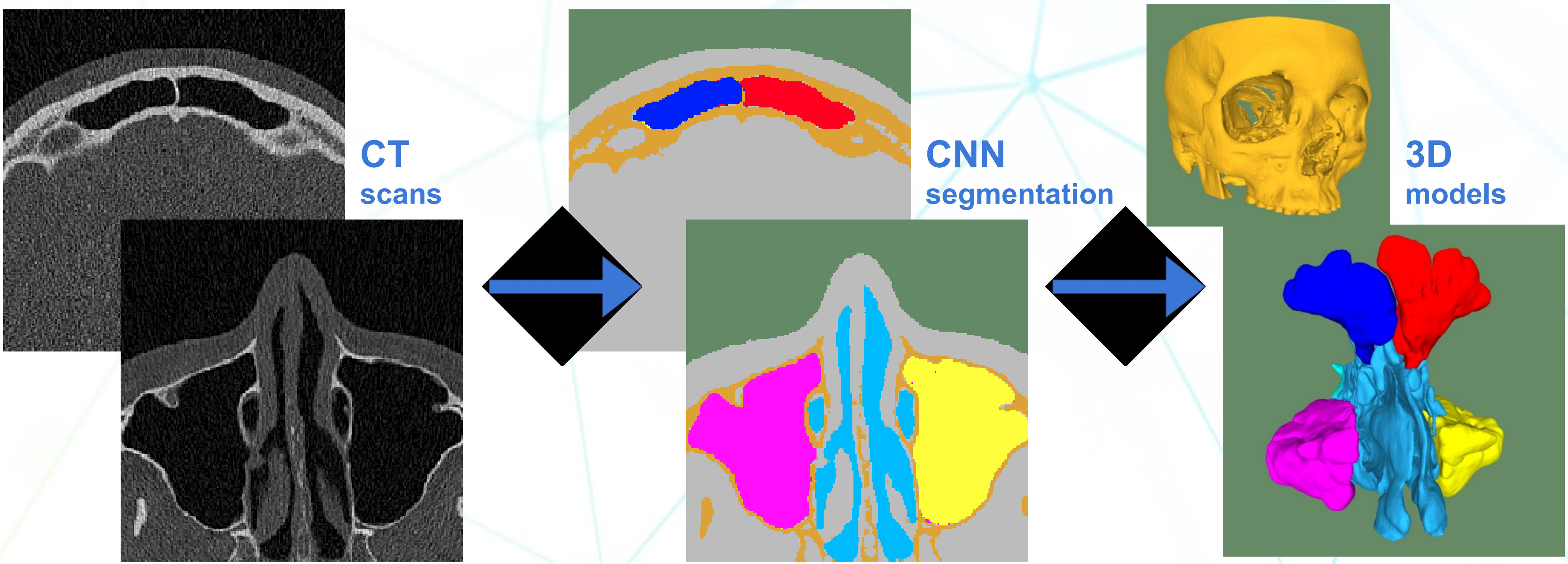
- goal: obtain different segments for different tissues / anatomical structures

- different linear combinations of the same feature map volume are used to obtain different segments

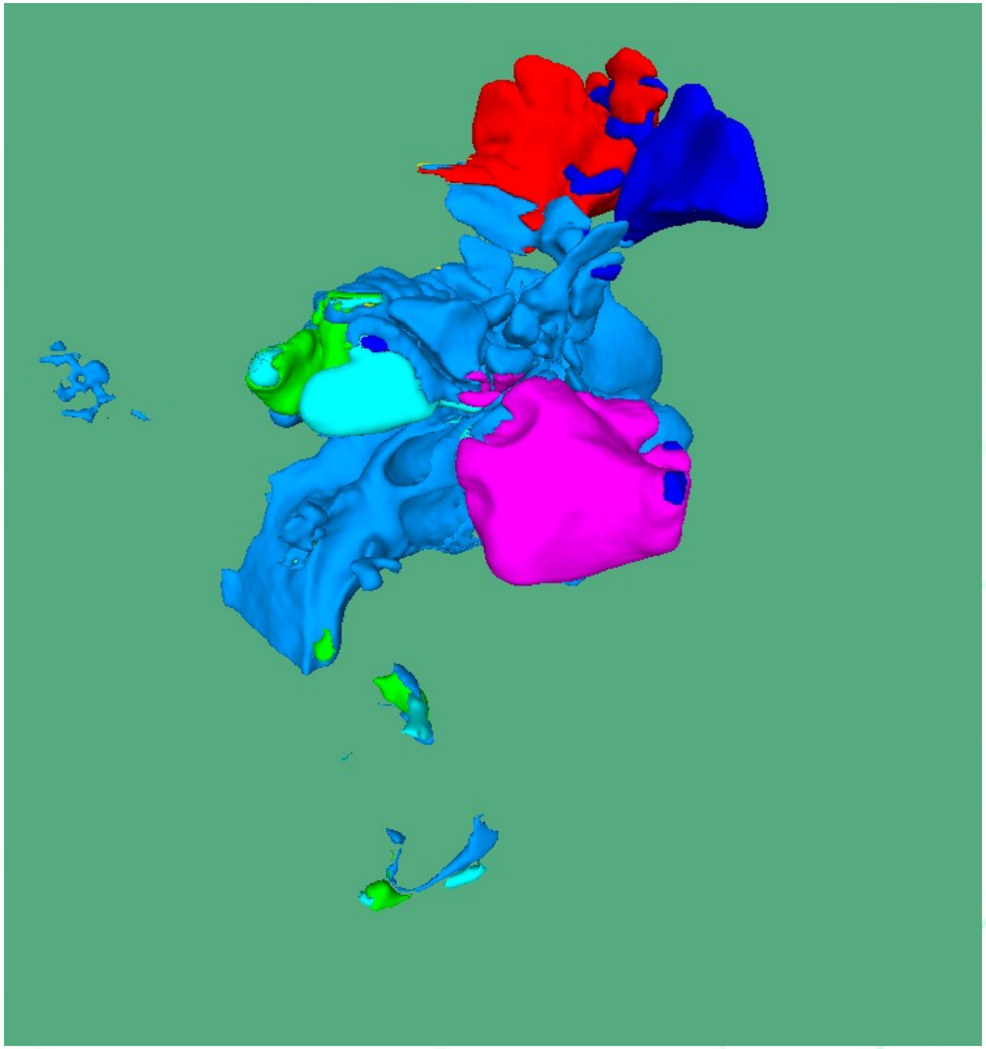


Goal: To obtain separate segmentations...

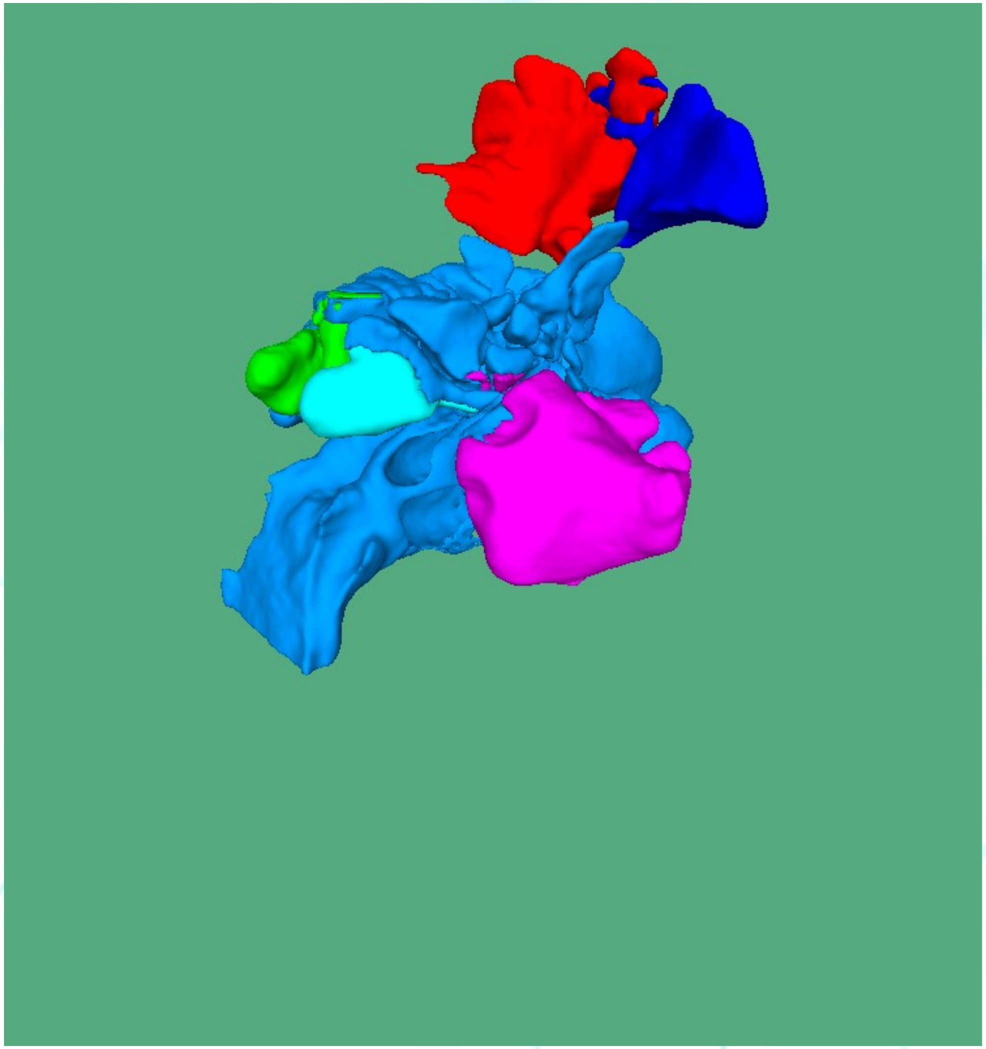
...of bone (one segment)
...of tissues (one segment)
..and different air-filled regions (multiple segments,
e.g. nasal cavity, each sinus, air surrounding the patient's head)



Postprocessing to improve CNN output



Patient Riga 0003 **before** postprocessing

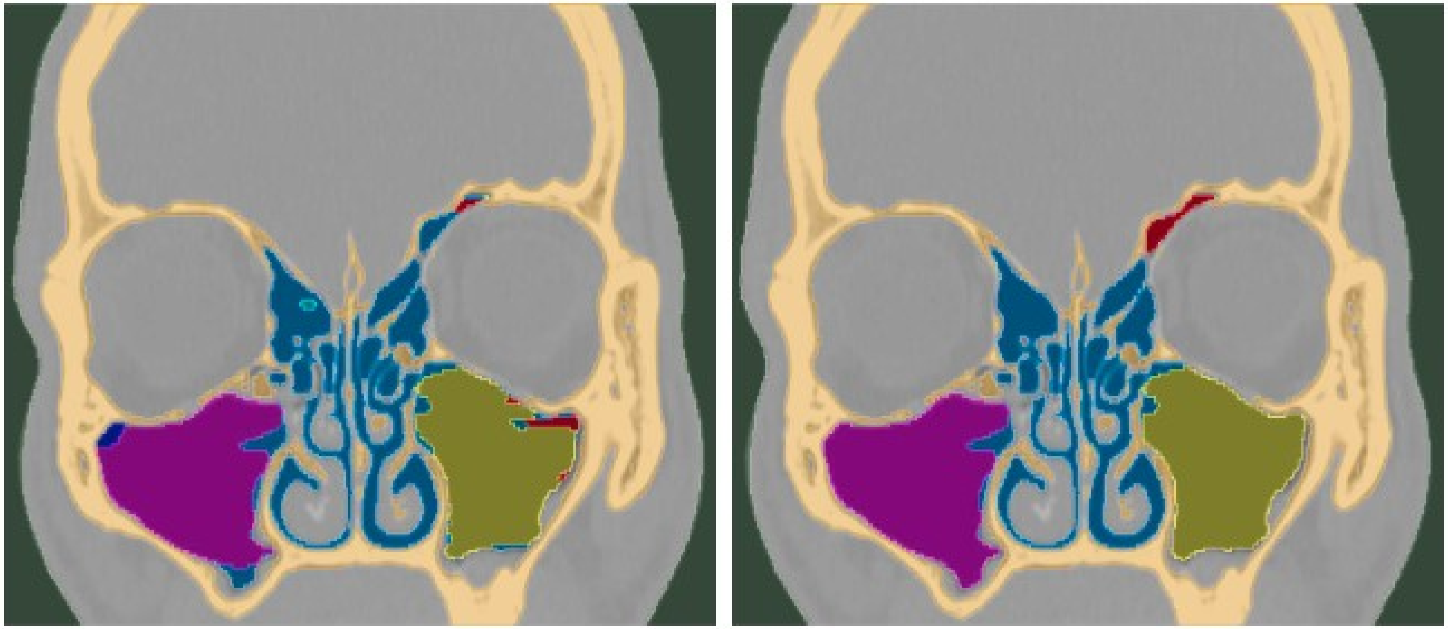


Patient Riga 0003 **after** postprocessing

Postprocessing to improve CNN output

Postprocessing steps (performed automatically):

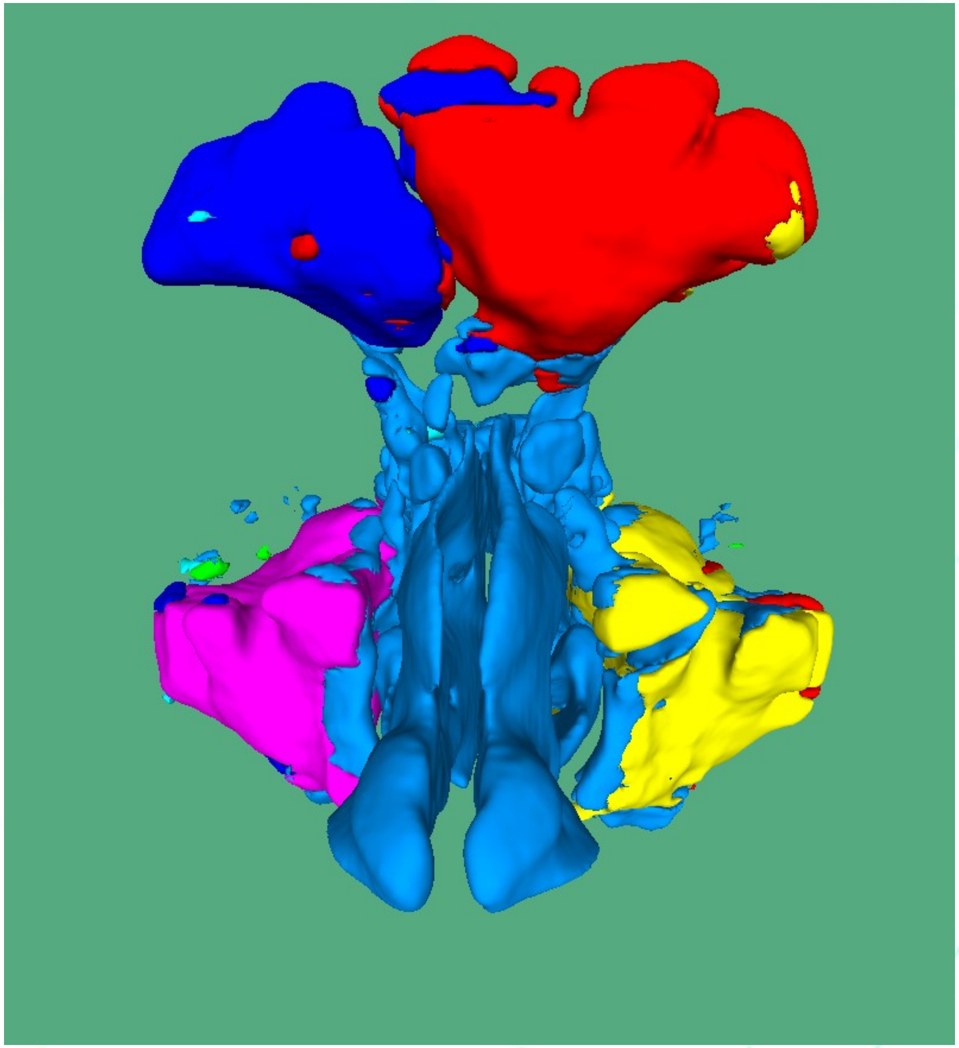
- removal of oral cavity (sometimes visible in the CT volume)
- removal of mastoid air cells (irrelevant for our purposes)
- correction of anatomically implausible components



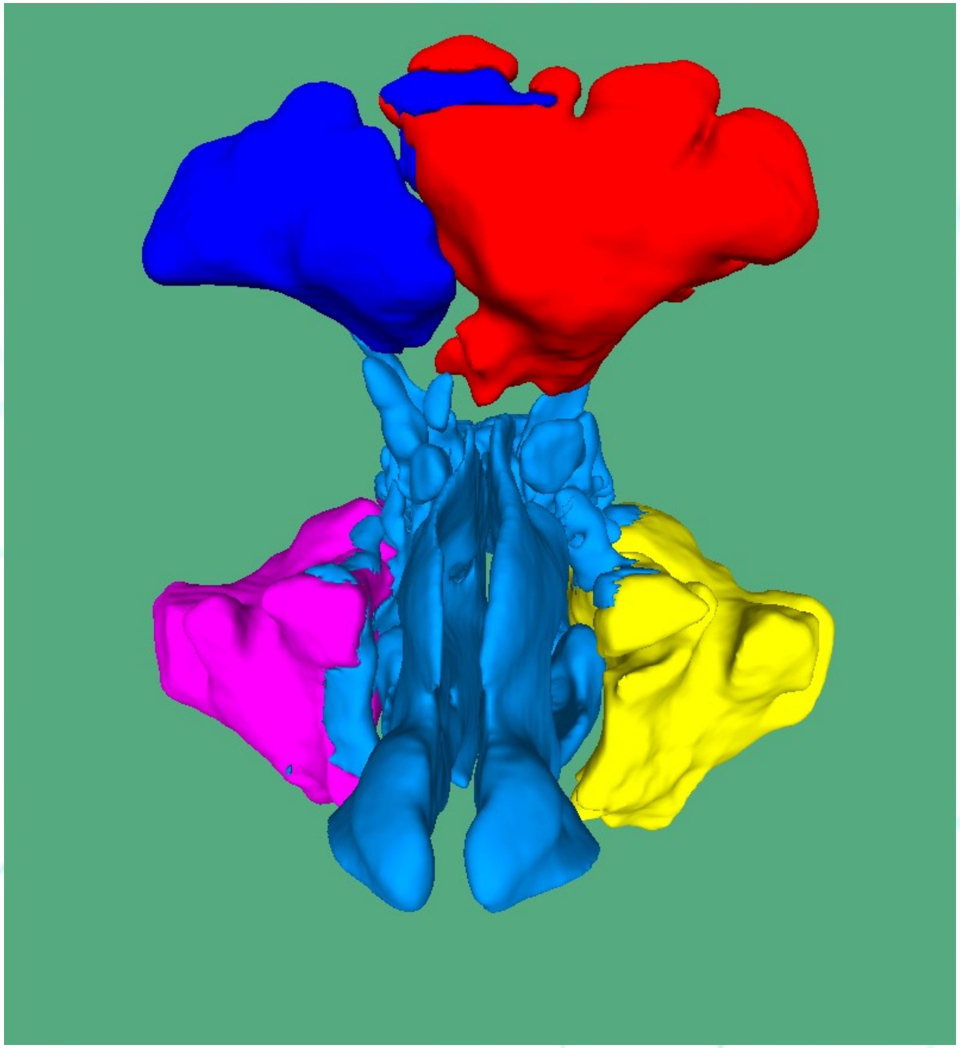
(a) Without postprocessing (b) With postproccesing

Fig 6. CNN segmentation with and without postprocessing for a single CT slice

Postprocessing to improve CNN output

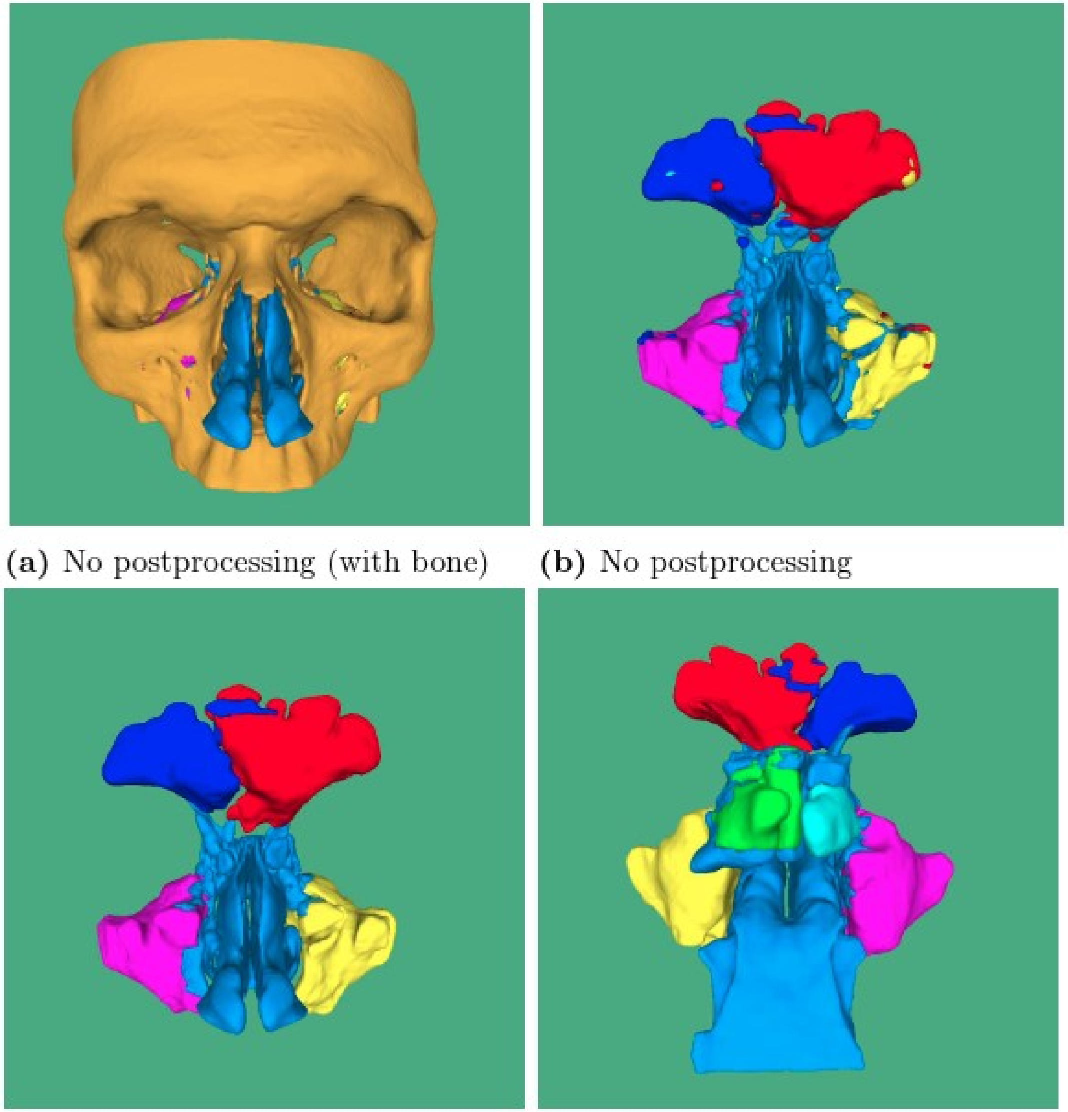


Patient Riga 0003 **before** postprocessing



Patient Riga 0003 **after** postprocessing

Postprocessing to improve CNN output



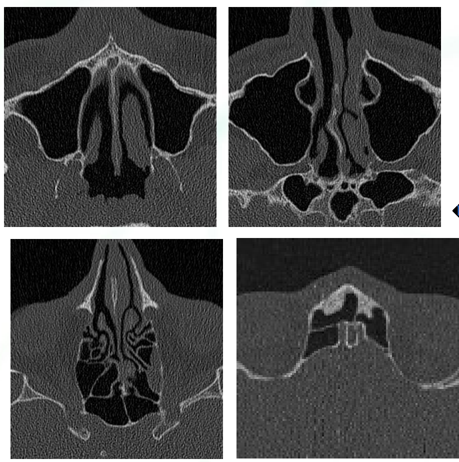
(a) No postprocessing (with bone) (b) No postprocessing

(c) After postprocessing (view 1) (d) After postprocessing (view 2)

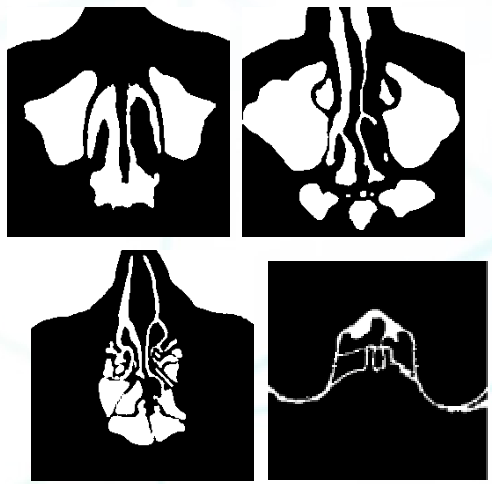
Fig 7. STL geometries generated from CNN segmentation

Outlook: Pathology Detection

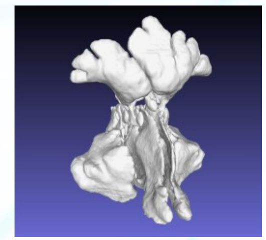
Overall aim: To extract 3D meshes of the paranasal sinuses from CT images, for both air and bone, and use the images and meshes to detect pathologies.



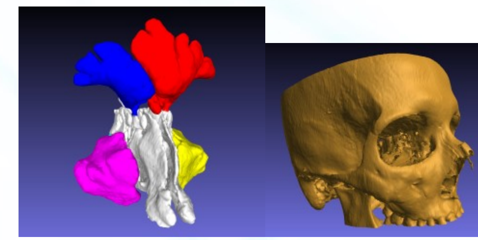
Obtain CT images of a patient



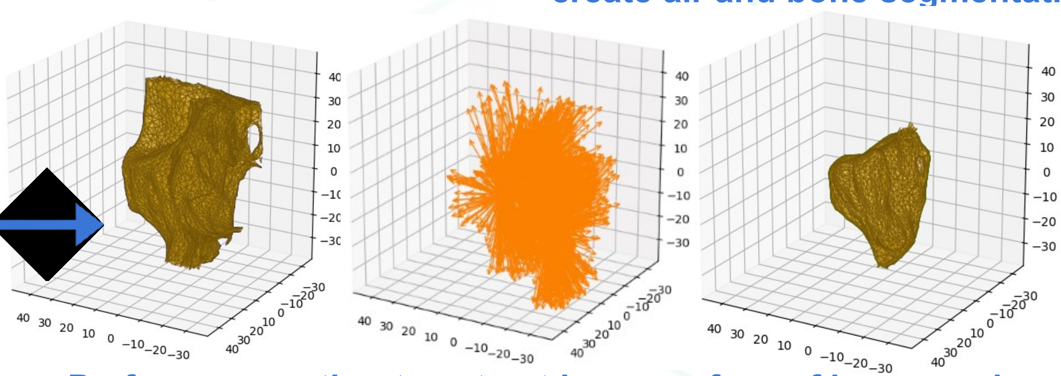
Use Convolutional Neural Networks to create air and bone segmentations



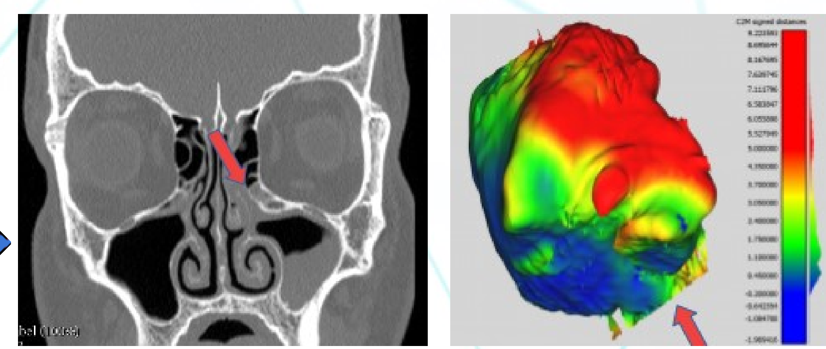
Create air and bone meshes from binary 2D segmentations and color air mesh, or generate colored mesh directly from CTs



Result: Air mesh with colored sinuses and bone mesh



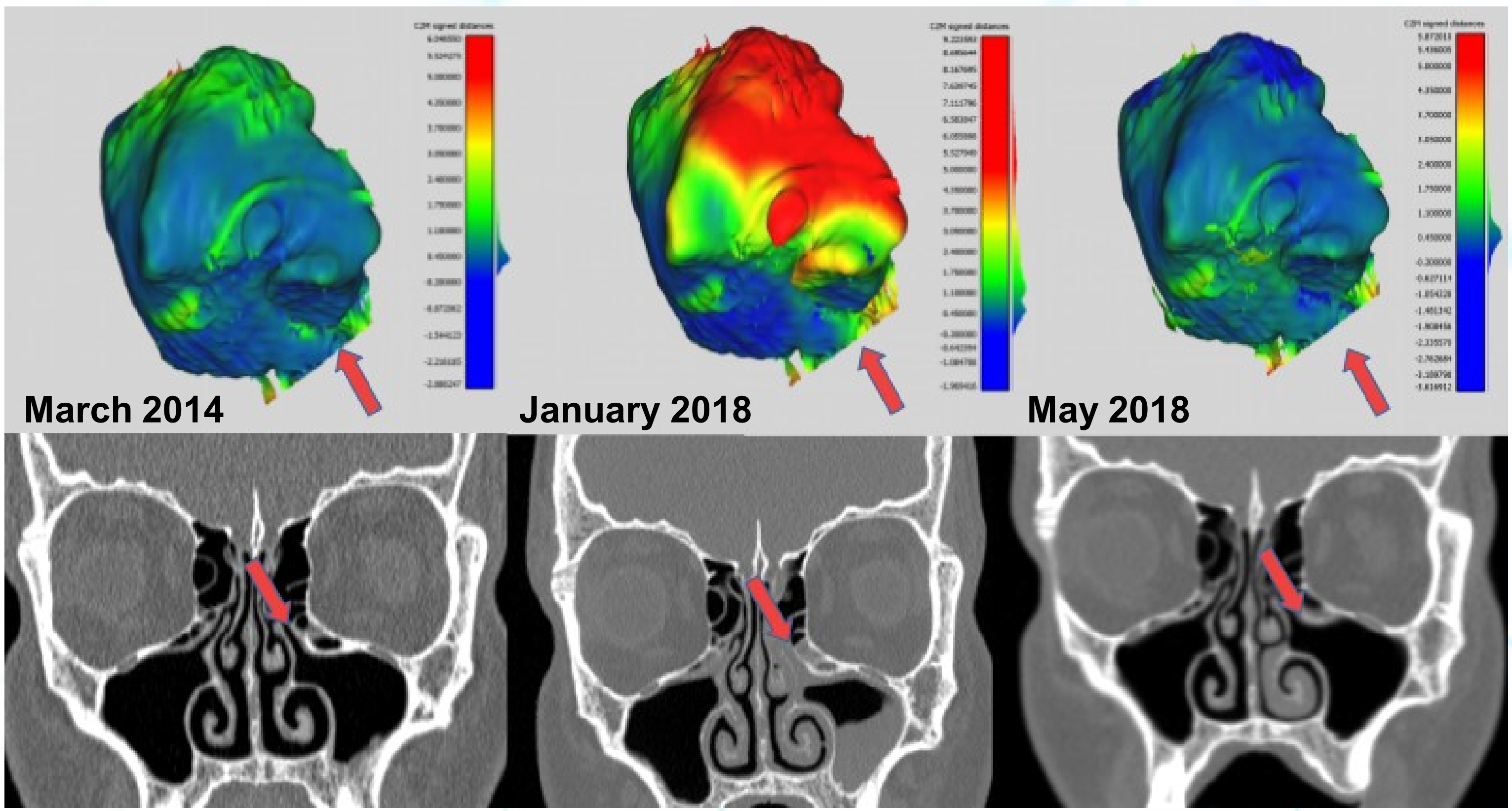
Perform raycasting to extract inner surface of bone mesh for each sinus (using ray origins inside air mesh)



Perform point cloud comparison to measure distance between air and bone surface mesh to detect obstructions

Outlook: Pathology Detection

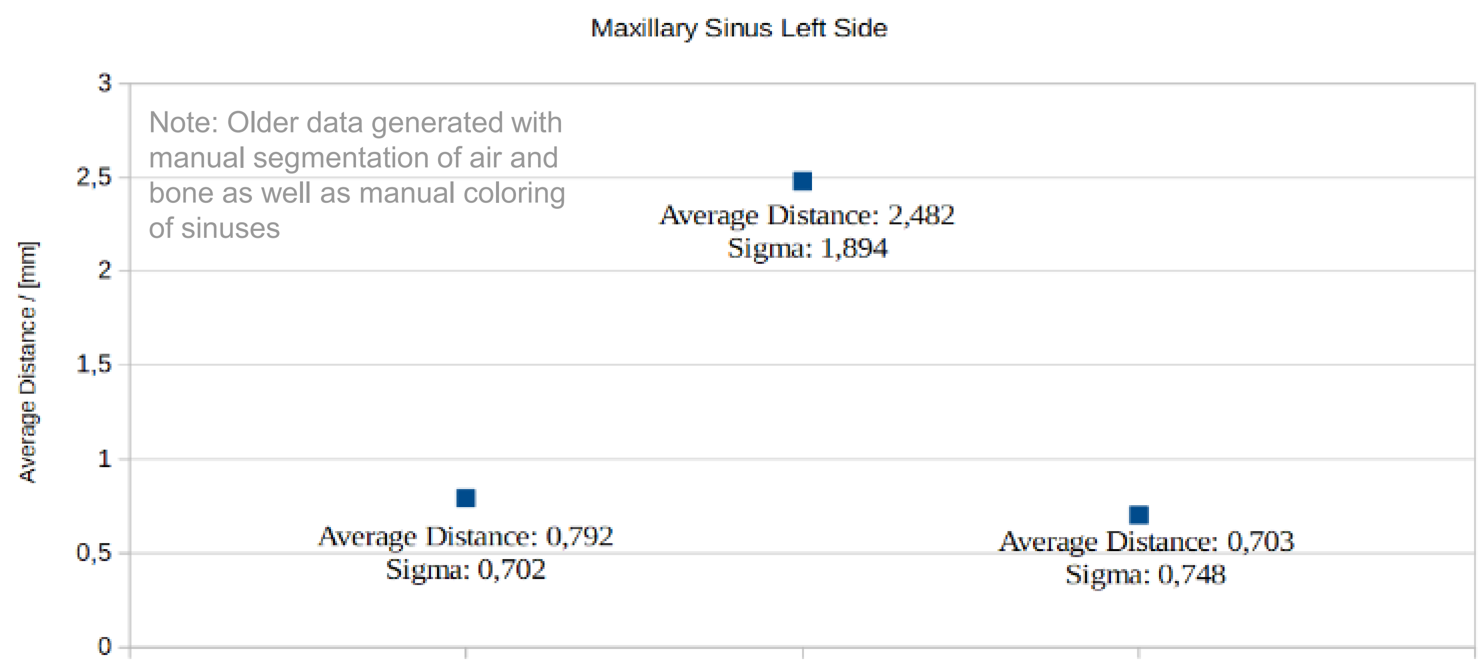
Aim: To obtain clinically relevant information about the physiology and pathophysiology of nasal cavity and paranasal sinuses, to automate pathology detection and provide actionable information for ENT specialists.



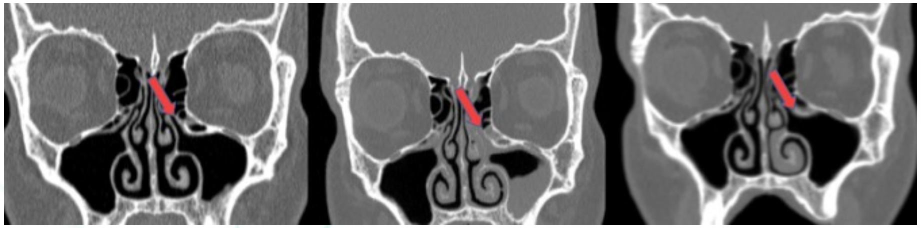
The distance between the inner surface of the bone around a cavity and the surface of a cavity itself (surface of the air mesh) can yield information about pathological obstructions. Furthermore, volumes, changes in volume and differences between volumes can be computed.

Outlook: Pathology Detection

Aim: To obtain clinically relevant information about the physiology and pathophysiology of nasal cavity and paranasal sinuses, to automate pathology detection and provide actionable information for ENT specialists.



March 2014 **January 2018** **May 2018**



Increase in average distance between bone and cavity of sinus suggests pathological change.

Thank you!

RHINO
Morphological and
DIAGNOST
functional precision diagnostics of the nose

Walter Koch

kochw@ait.co.at

Angewandte Informationstechnik Forschungsgesellschaft mbH

AUSTRIA

ANGEWANDTE INFORMATIONSTECHNIK
Forschungsgesellschaft mbH

IraSME

ZIM Zentrales Innovationsprogramm Mittelstand

FFG

bmwfw Bundesministerium für Wissenschaft, Forschung und Wirtschaft