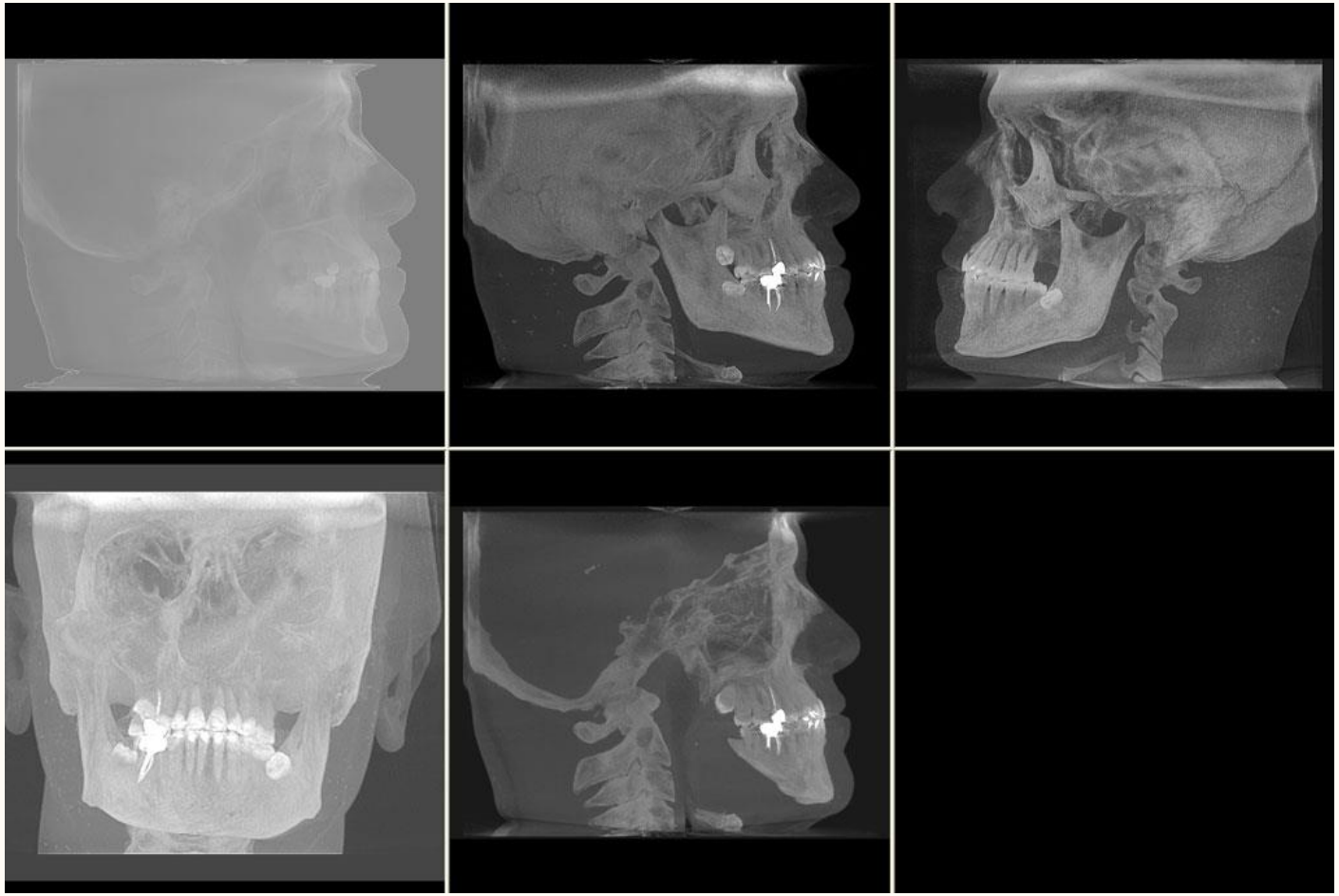
DEEP MULTI-STAGE MODEL FOR AUTOMATED LANDMARKING OF CRANIOMAXILLOFACIAL CT SCANS

S. Palazzo, G. Bellitto, L. Prezzavento, F. Rundo,
U. Bagci, D. Giordano, R. Leonardi, C. Spampinato

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Motivation

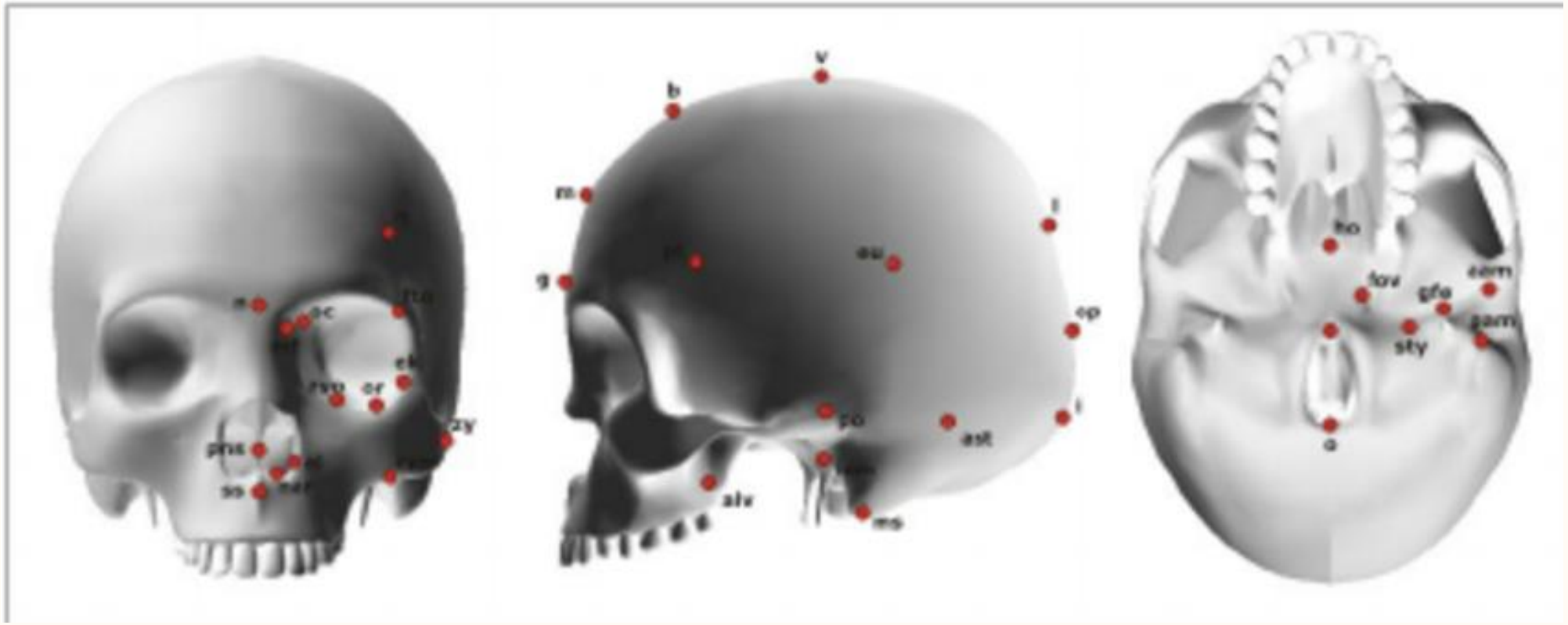
Computer tomography (CT) and **cone-beam computer tomography (CBCT)** are the most common imaging modalities for diagnosis and treatment of **craniomaxillofacial (CMF) disorders**.



Motivation

Limitations of automated landmarking:

- Landmarking requires to work at full resolution scale
- Existing models need upstream segmentation to control landmark identification.



Motivation

Limitations of automated landmarking:

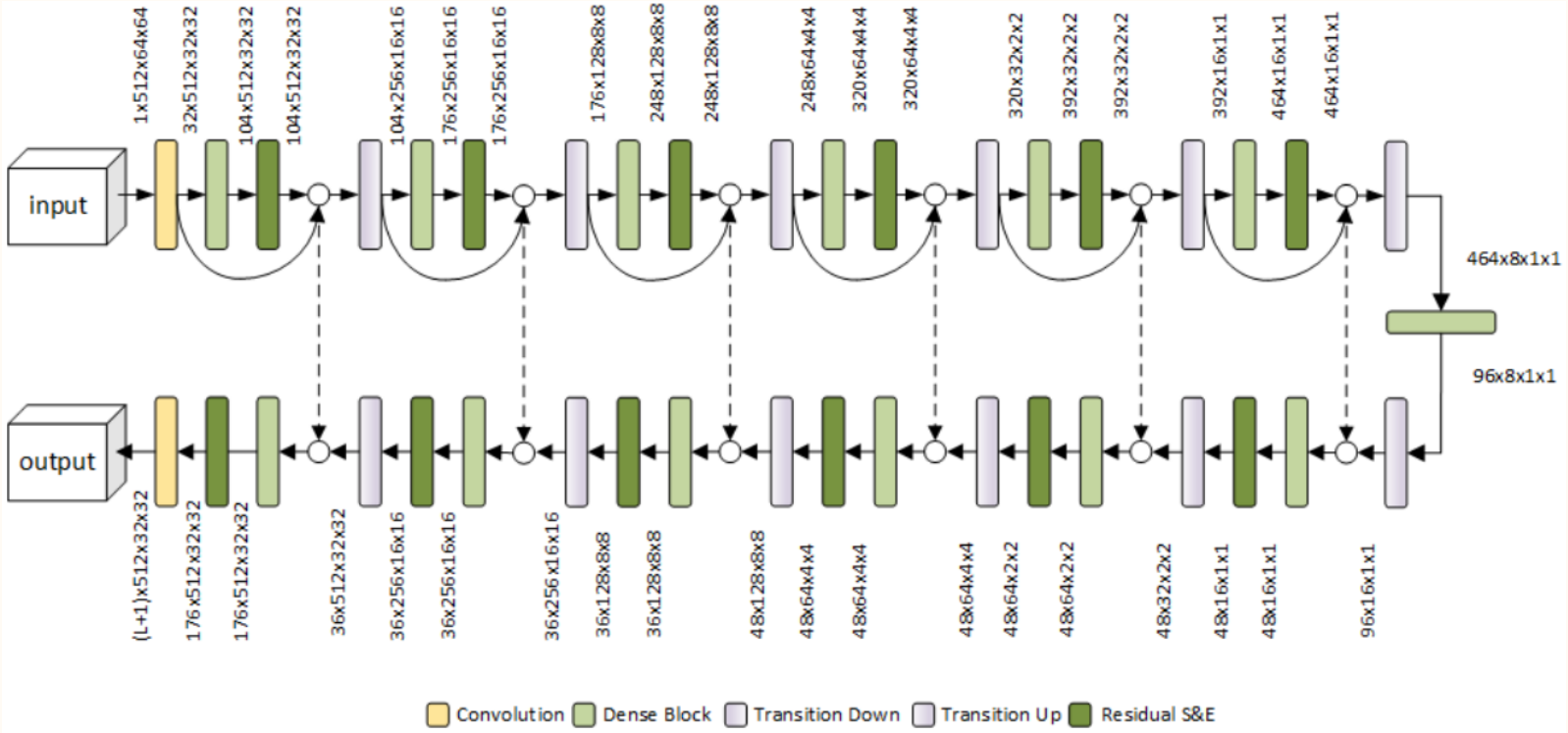
- Landmarking requires to work at full resolution scale
- Existing models need upstream segmentation to control landmark identification.

- We face this problem through a multi-stage pipeline that first **identifies areas** where landmarks may be found, on reduced-resolution images, and then **refines localization**, at full resolution scale.

Landmark Localization

Landmark plane localization

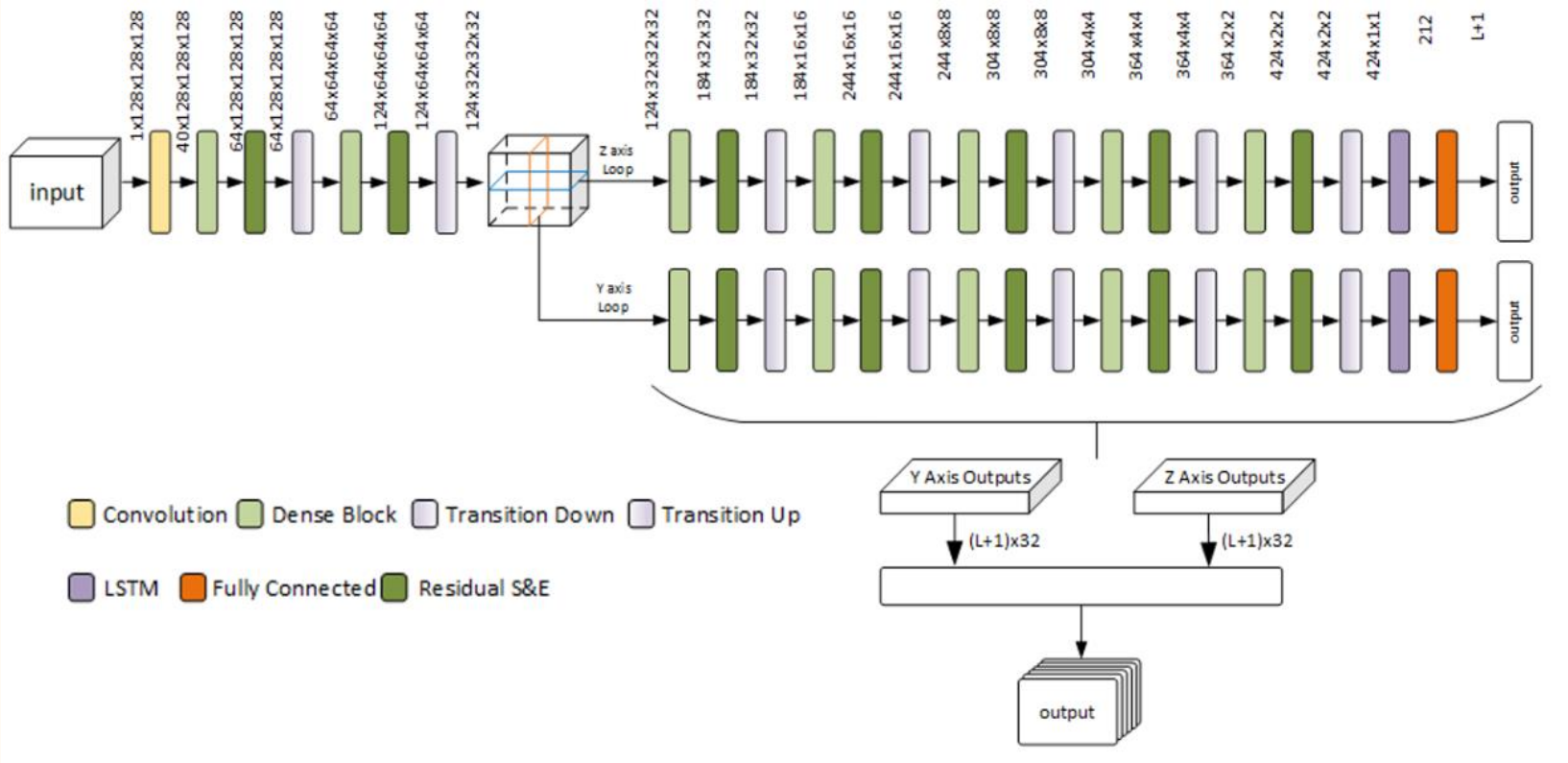
- Localization of the *x* coordinate for each target landmark
- 3D convolutions
- Output treated as classification volume (maximal activation is selected)
- Full-resolution for *x* axis (512 pixels)
- Shared features among landmarks



Landmark Localization

Landmark region localization

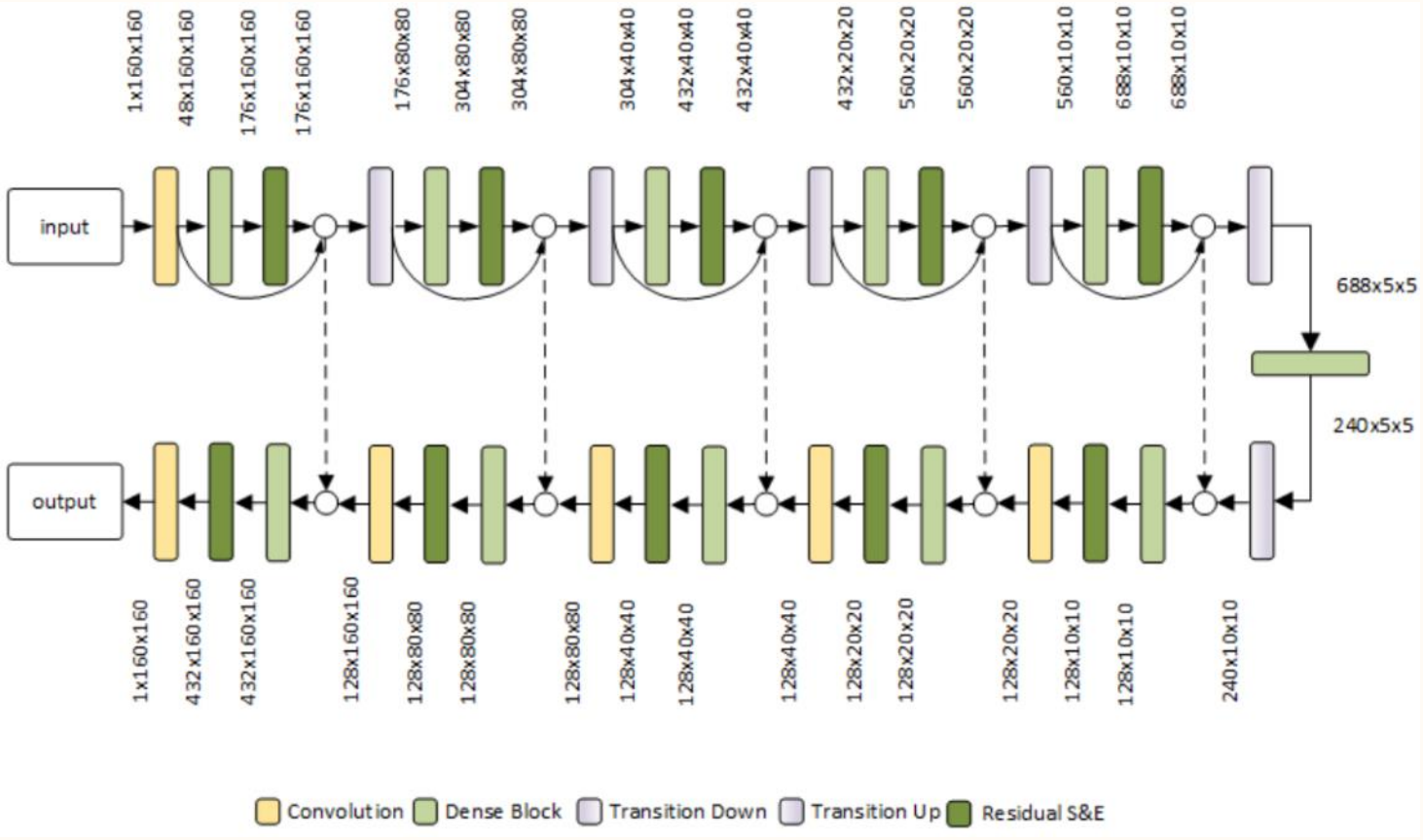
- Localization of a small area on the *yz* plane around each candidate landmark
- 3D convolutions
- Sequential scanning of slices through LSTMs along the two directions
- Maximal activation in output maps identifies center of target region
- Down-sampled input (128x128x128)
- Shared features



Landmark Localization

Landmark point localization

- Landmark-specific models
- 2D convolutions
- Full-resolution on a reduced area (160x160)
- Output treat as classification map (maximal likelihood pixel denotes *yz* coordinates)



Performance Evaluation

Datasets used:

- **AirwaysSet**
  - 19 CT scans of 19 patients.
  - Five landmarks for airways analysis (ANS, PNS, C2sp, C3ai, N), employed to study obstructed sleep apnea syndrome (OSAS)

- **Anomymized CBCT scans of 50 patients [1]**
  - Nine cephalometric landmarks (Me, Gn, Pg, B, Id, CdL, CdR, CorL, CorR) adopted in clinical routine for planning CMF surgery.

[1] N. Torosdagli, D. K. Liberton, P. Verma, M.Sincan, J. S. Lee, U. Bagci, "Deep geodesic learning for segmentation and anatomical landmarking", IEEE Transactions on Medical Imaging, 2018

Performance Evaluation - Results

Landmark Localization Error (in mm) on AirwaysSet

Baseline: 3D Tiramisu operating at largest possible resolution that fits in a Titan X GPU (128x128x128)

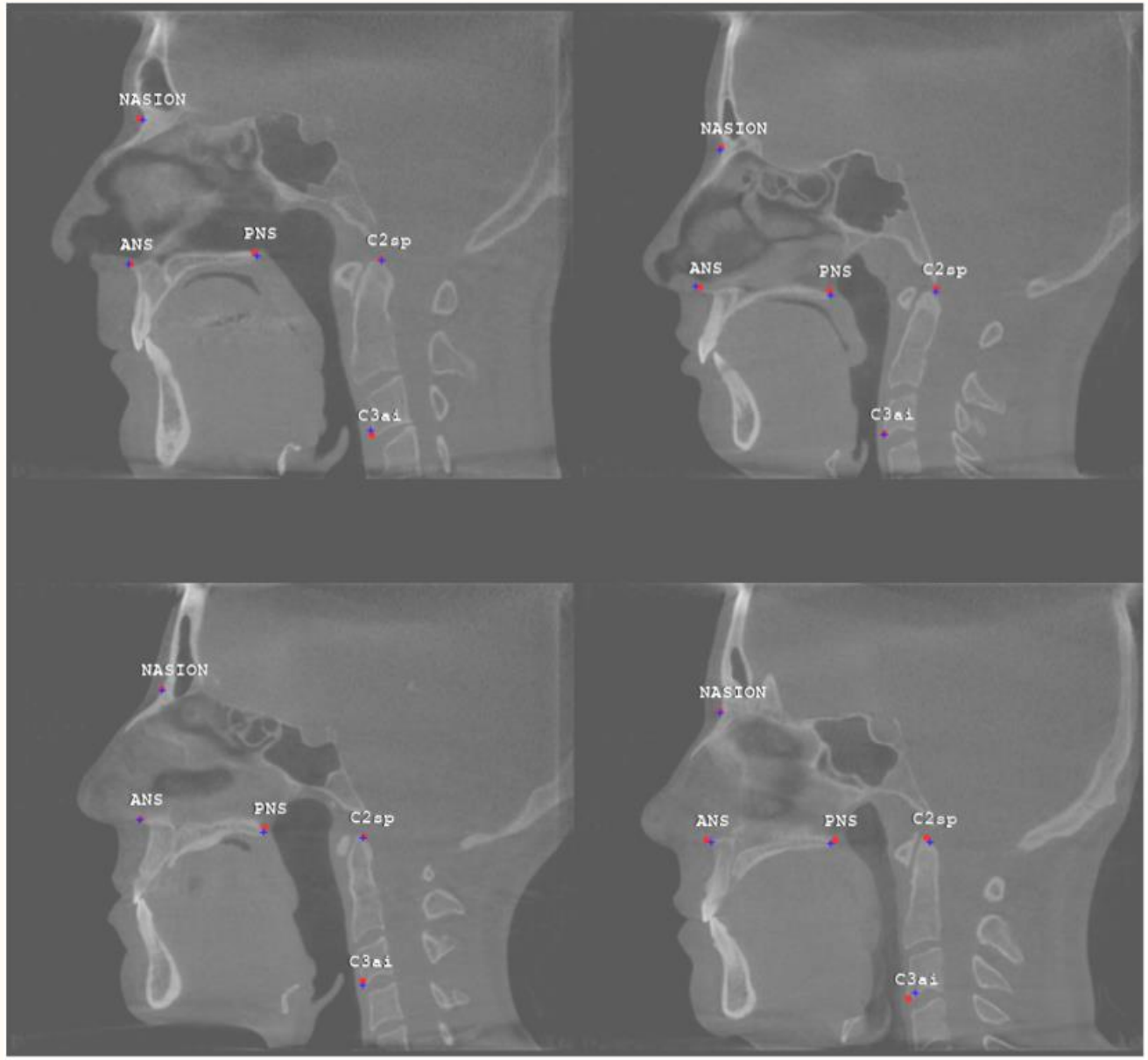
	Our model	Baseline
ANS	0.88 ± 0.64	1.49 ± 1.11
C2sp	0.74 ± 0.50	1.01 ± 1.01
C3ai	1.06 ± 0.83	5.80 ± 4.91
N	0.57 ± 0.53	1.23 ± 1.48
PNS	1.01 ± 0.56	1.09 ± 0.77
Average	0.85	2.12

Comparison with state of the art [1] in terms of localization error (in mm).

Note that [1] requires prior segmentation (hard to apply to airways, since landmarks are located in different structures)

	Our model	Torosdagli *et al* [1]
Me	0.84	0.33
Cor_L	0.71	0.27
Cor_R	0.52	0.03
Cd_L	1.08	1.01
Cd_R	0.47	0
Gn	0.91	0.41
Pg	0.93	1.36
B	0.73	0.68
Id	0.85	0.35
Average	0.78	0.55

Performance Evaluation - Results



Conclusion

- We propose a learning-based approach for automated landmarking of CT images that is able to work directly on the input data at its full resolution, employing a multi-stage pipeline
- This leads to an average landmarking error of less of 1 mm on two different CMF structures (mandibles and airways)
- Performance is comparable to the state of the art, without the need of performing a prior segmentation step