DEEP RECURRENT-CONVOLUTIONAL MODEL FOR AUTOMATED SEGMENTATION OF CRANIOMAXILLOFACIAL CT SCANS

F. Murabito, S. Palazzo, F. Proietto Salanitri, 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

- Craniomaxillofacial malformations due to abnormal development comprise over one-third of all congenital birth defects.
- CT and CBCT are the most common imaging modalities for diagnosis and treatment of CMF disorders.
- Segmentation and Landmarking are the upstream steps for any subsequent clinical evaluation

Motivation

- Challenge of automatically CMF CT segmentation:
  - Multiple structures
  - Irregular and complex shape patterns
  - Lack of contrast in joints
  - Significant morphological variations among different patients

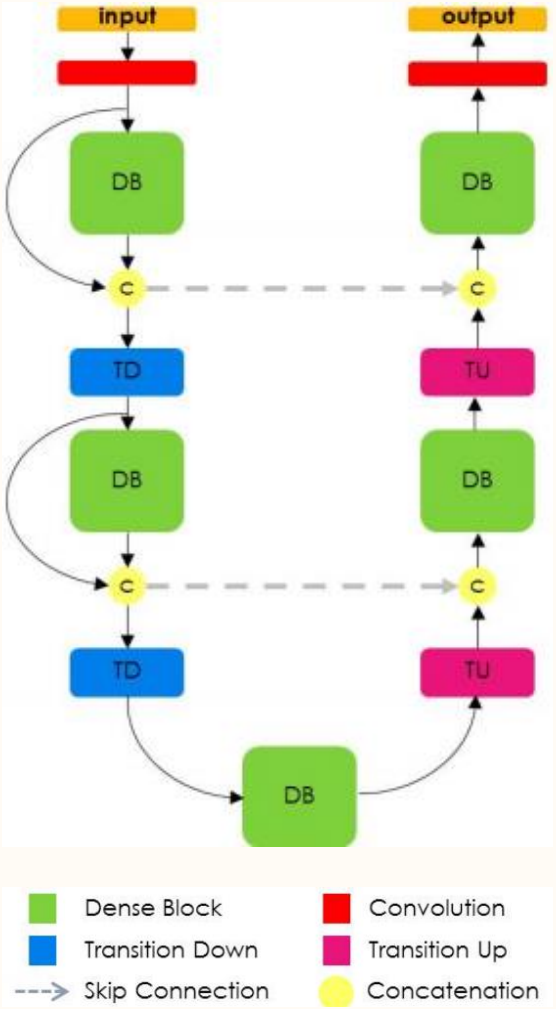
State of the art

- Before DL: atlas-guided methods, with a preliminary registration step
- After DL: Encoder-decoder fully convolutional networks (several variants)

Problems

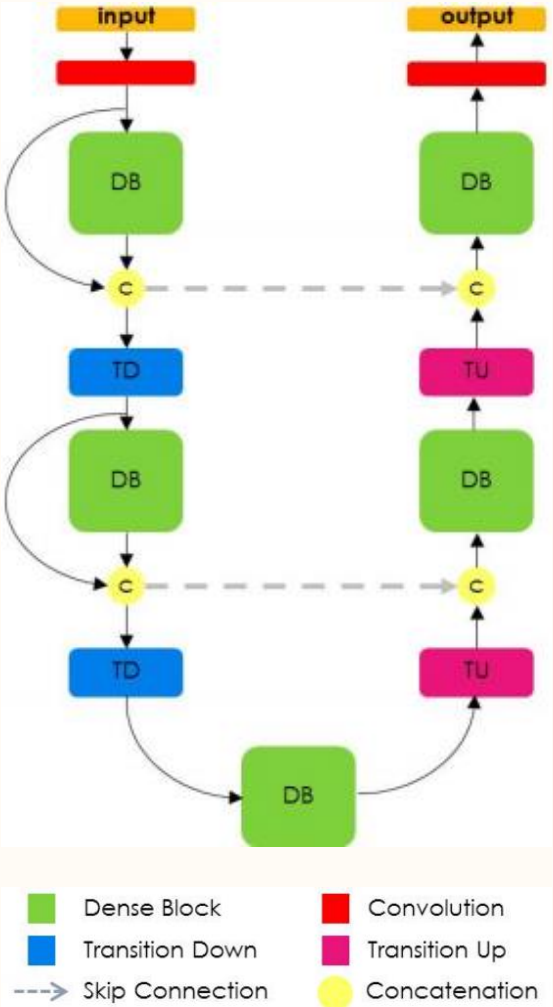
- Large appearance variability of CMF CT scans
- Limited generalization capabilities

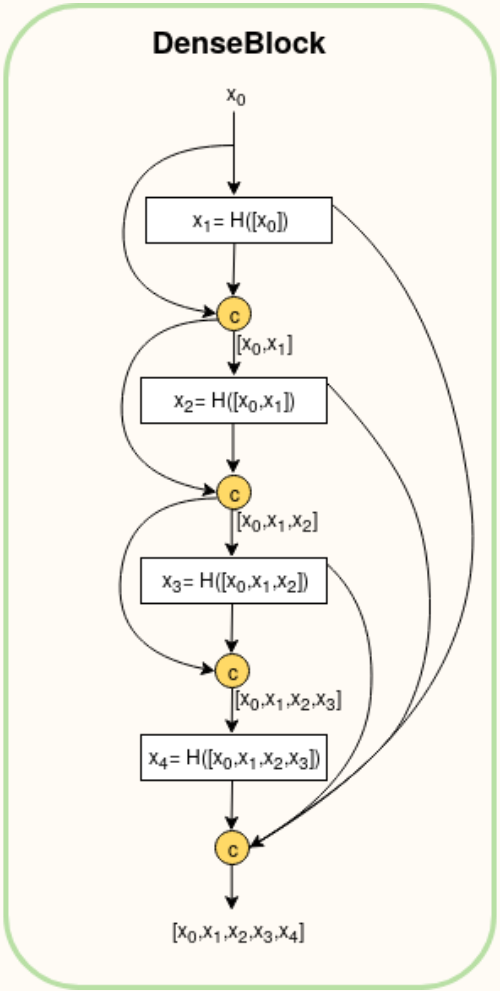
The Tiramisu Network [1]



[1] S. J´egou, M. Drozdzal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: Fully convolutional dense nets for semantic segmentation," in CVPRW 2017. IEEE, 2017, pp. 1175–1183.

The Tiramisu Network [1]

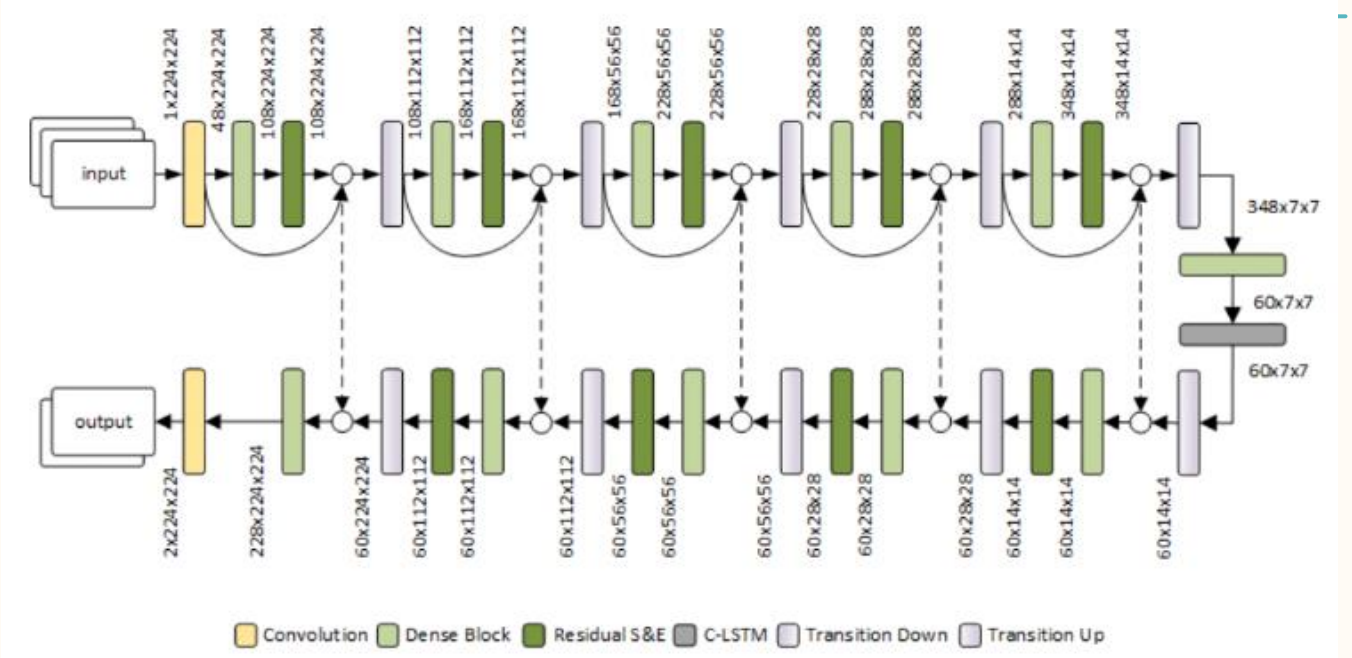




[1] S. J´egou, M. Drozdzal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: Fully convolutional dense nets for semantic segmentation," in CVPRW 2017. IEEE, 2017, pp. 1175–1183.

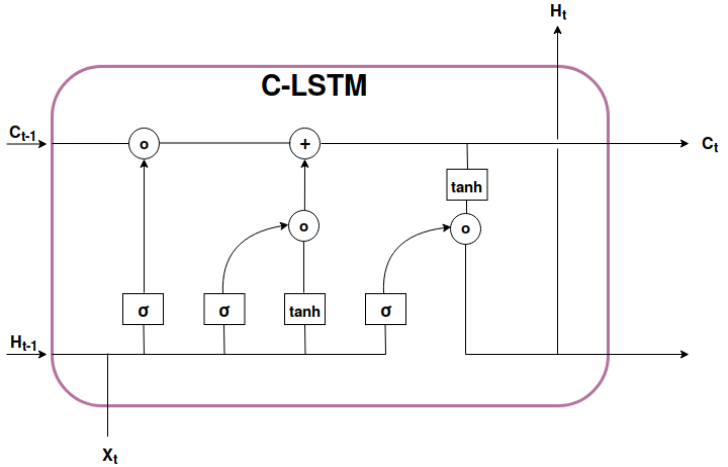
Method

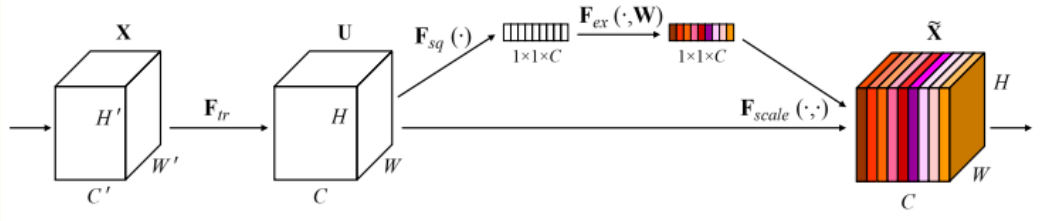
- Fully convolutional DenseNet following U-Net architecture
- C-LSTMs at bottleneck to exploit spatial axial correlation of consecutive scan slices
- Residual squeeze-and-excitation layers to emphasize relevant features and improve representational power

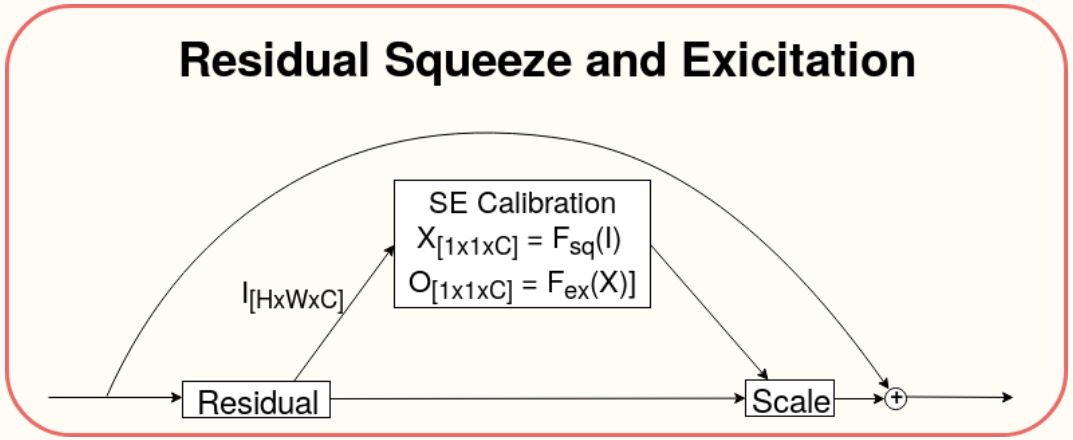


Method

Unidirectional Convolutional LSTM







Squeeze-and-Excitation block [2]

[2] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," arXiv preprint arXiv:1709.01507, vol. 7, 2017

Performance Evaluation - Datasets

- **MandibleSet**
  - 32 CT scans and 20 CBCT scans
  - Mandible segmentation
  - Patients with no clinically obvious facial asymmetries and no prior surgery in the head and neck
- **MICCAI Head and Neck 2015 Challenge dataset**
  - CT scans of 48 patients with manual segmentations of several anatomical regions
  - Challenging because of image artifacts and different disease
- **AirwaysSet (proposed in this paper)**
  - Subset of MandibleSets
  - 19 CT scans (one per subject)
  - Nasal cavity and pharyngeal areas segmentation

Performance Evaluation

Metrics

- Dice Similarity coefficient

$$DSC = \frac{2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2}$$

N number of pixels

p_i and g_i output of the model and ground truth at pixel i

Training procedure

- 3 consecutive slices as input processed individually, normalized to 0 mean and unitary standard deviation
- Training for 50 epochs on the MandibleSet and optimized with RMSProp
- Dice Score Coefficient computed on the central slice was used for backpropagation

Performance Evaluation

Model	Pure DL	DSC (%)
3D-UNet [24]	Yes	87.34
Tiramisu [7]	Yes	91.20
Ours	Yes	93.41
Robust segmentation [26]	No, post-processing	91.00
AnatomyNet [3]	No, post-processing	92.50
Best MICCAI 2015 [27]	No, landmark-guided	93.90
Hierarchical Vertex [25]	No, shape prior	94.00

- Performance on MICCAI Head and Neck 2015 Dataset

- DSC performance when fine-tuned(second columns) and trained from scratch on the CBCT part of MandibleSet and on the AirwaysSet

Model	Fine-Tuned (%)	Trained from scratch (%)
	CBCT - Mandible set	
3D-UNet [24]	80.32	73.36
Tiramisu [7]	83.74	76.78
Ours	89.25	80.31
	Airways	
3D-UNet [24]	81.82	77.30
Tiramisu [7]	84.02	80.21
Ours	93.31	85.12

Performance Evaluation

Ablation studies

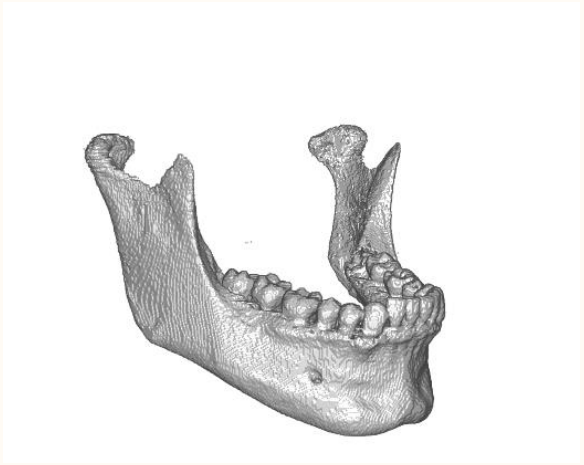
Model	DSC (%)	
[7]	92.81	Tiramisu Network
Res-SE	95.32	Tiramisu Network + SE
C-LSTM	95.04	Tiramisu Network + C-LSTM
Bi-C-LSTM	94.81	Tiramisu Network + Bi C-LSTM
Res-SE + C-LSTM	96.52	Tiramisu Network + SE + C-LSTM
Res-SE + Bi-C-LSTM	95.03	Tiramisu Network + SE + Bi C-LSTM

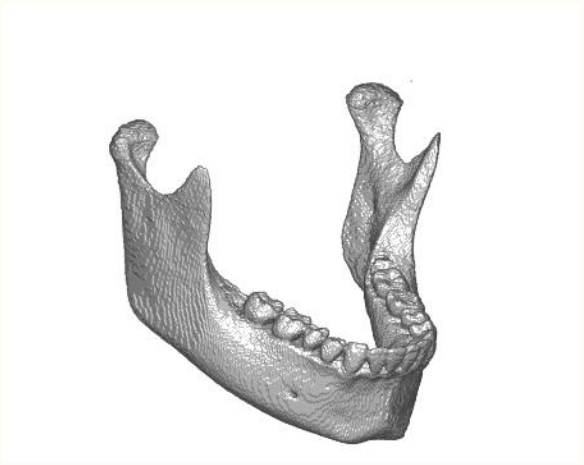
Performance Evaluation

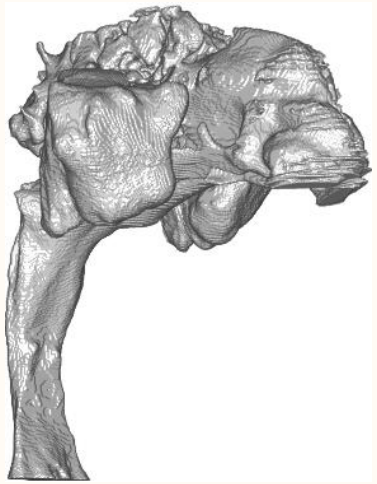
Ablation studies

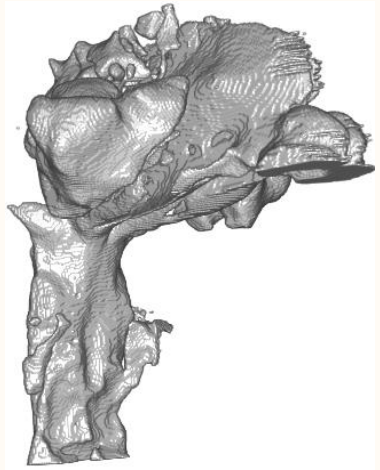
Model	DSC (%)	
[7]	92.81	Tiramisu Network
Res-SE	95.32	Tiramisu Network + SE
C-LSTM	95.04	Tiramisu Network + C-LSTM
Bi-C-LSTM	94.81	Tiramisu Network + Bi C-LSTM
Res-SE + C-LSTM	96.52	Tiramisu Network + SE + C-LSTM
Res-SE + Bi-C-LSTM	95.03	Tiramisu Network + SE + Bi C-LSTM

Segmentation examples









DATE

Conclusion

We propose:

- A Deep encoder-decoder network for automated segmentation of CMF structures
  - Improved segmentation performance
  - Enhanced generalization capabilities to multiple CMF structures
- Two new Datasets
  - MandibleSet containing multiple mandible segmentations
  - AirwaysSet containing multiple nasal cavity and pharynges segmentations