Robust Multi-Scale Anatomical Landmark Detection in Incomplete 3D-CT Data

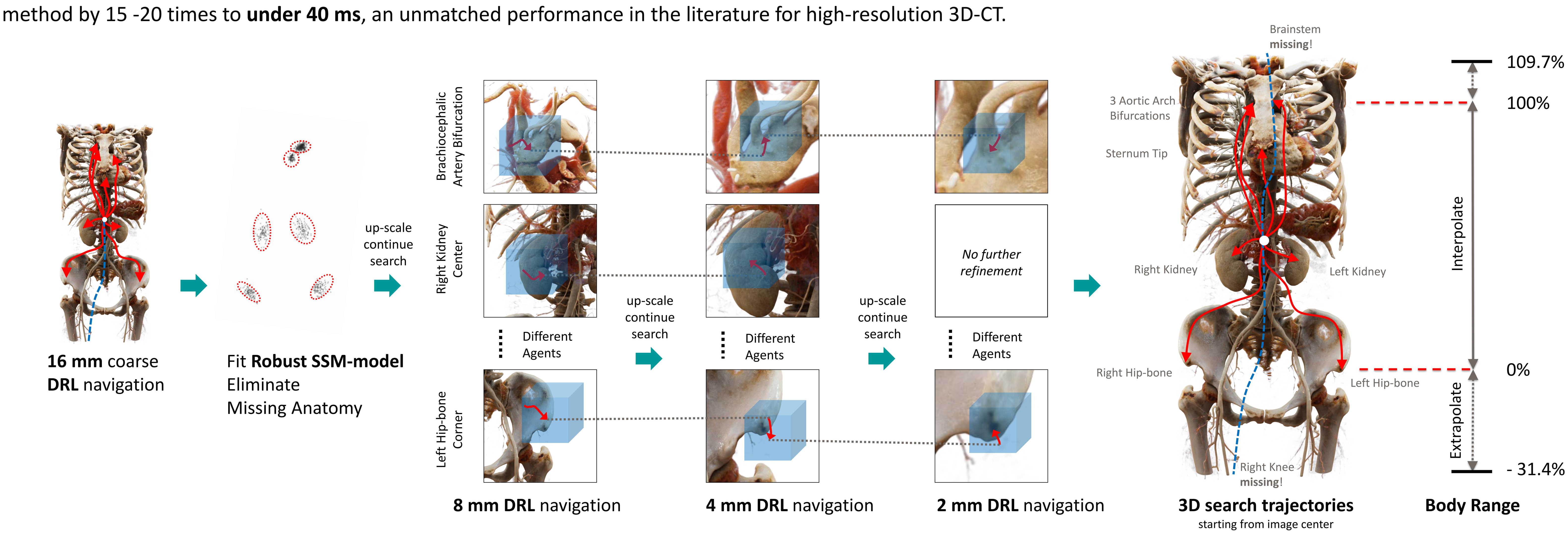
Florin C. Ghesu[1,2], Bogdan Georgescu[1], Sasa Grbic[1], Andreas Maier[2], Joachim Hornegger[2], and Dorin Comaniciu[1]

Medical Imaging Technologies, Siemens Healthineers, Princeton NJ, USA
Pattern Recognition Lab, Friedrich-Alexander-Universität, Erlangen, Nürnberg

MICCAI Quebec City CONFERENCE 10.-14. September 2017

Abstract

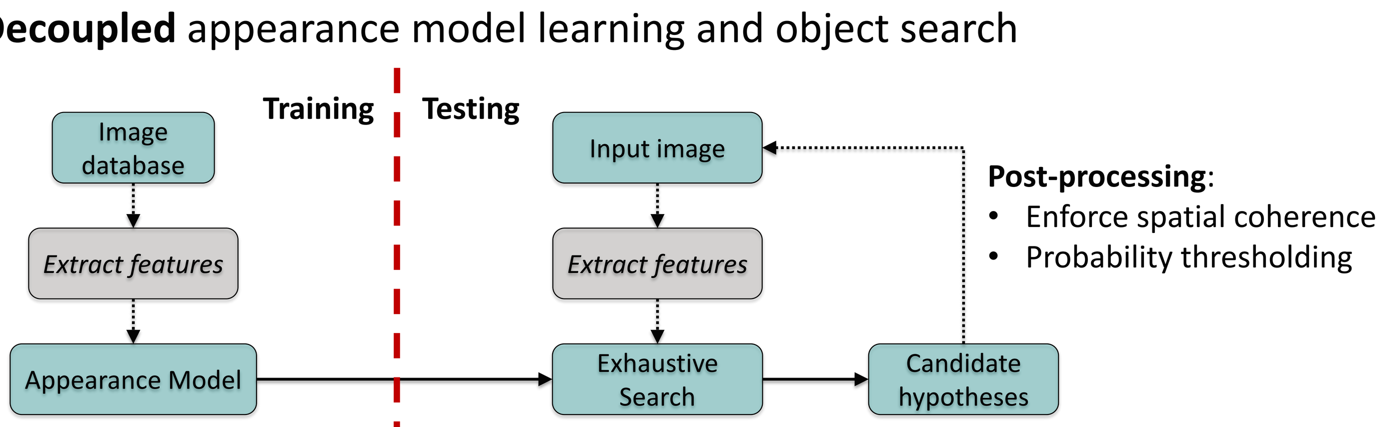
Robust and fast detection of anatomical structures is an essential prerequisite for the next-generation automated medical support tools. Previous solutions are typically driven by suboptimal and exhaustive strategies and do not effectively address cases of incomplete data, i.e., scans taken with a partial field-of-view. We address these limitations using the capabilities of **deep reinforcement learning** with **multi-scale image analysis** and **robust statistical shape modeling**. Artificial agents are taught optimal navigation paths in the image scale-space that can account for missing structures to ensure the **robust and spatially-coherent detection** of the observed anatomical landmarks. The identified landmarks are then used as robust guidance in estimating the extent of the body-region. Experiments show that our solution outperforms state-of-the-art deep learning in detecting different anatomical structures, without any failure, on a dataset of over **2300 3D-CT volumes**. In particular, we achieve **0% false-positive and 0% false-negative rates** at detecting the landmarks or recognizing their absence from the field-of-view of the scan. In terms of runtime, we reduce the detection-time of the reference method by 15 -20 times to **under 40 ms**, an unmatched performance in the literature for high-resolution 3D-CT.



Motivation

Typical Machine Learning for Anatomical Landmark Detection:

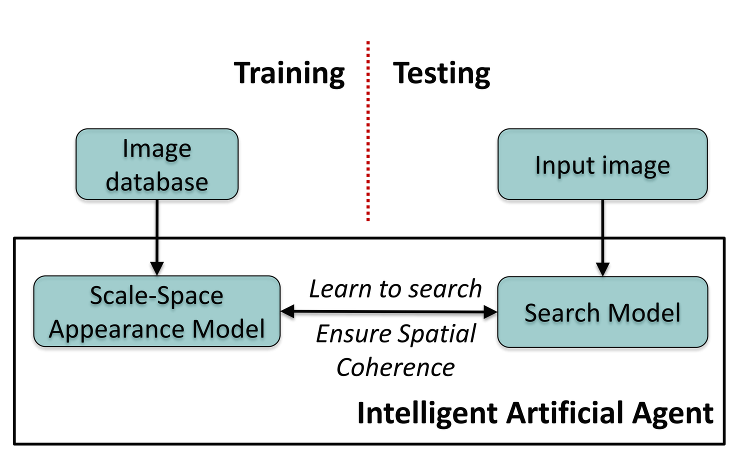
- Inefficient exhaustive scanning solutions subject to local false responses
- No understanding of underlying context
- **Heuristics** used for candidate aggregation and absence recognition
- **Decoupled** appearance model learning and object search



Proposed Method

Learning to search across image scales:

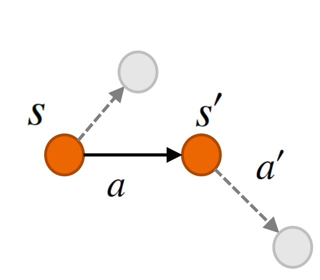
- Landmark detection – 3D navigation (multi-scale search-strategy)
- Focus from coarse to fine scale
- Discard large subspaces from the search
- Diverging trajectories signal ***absence***



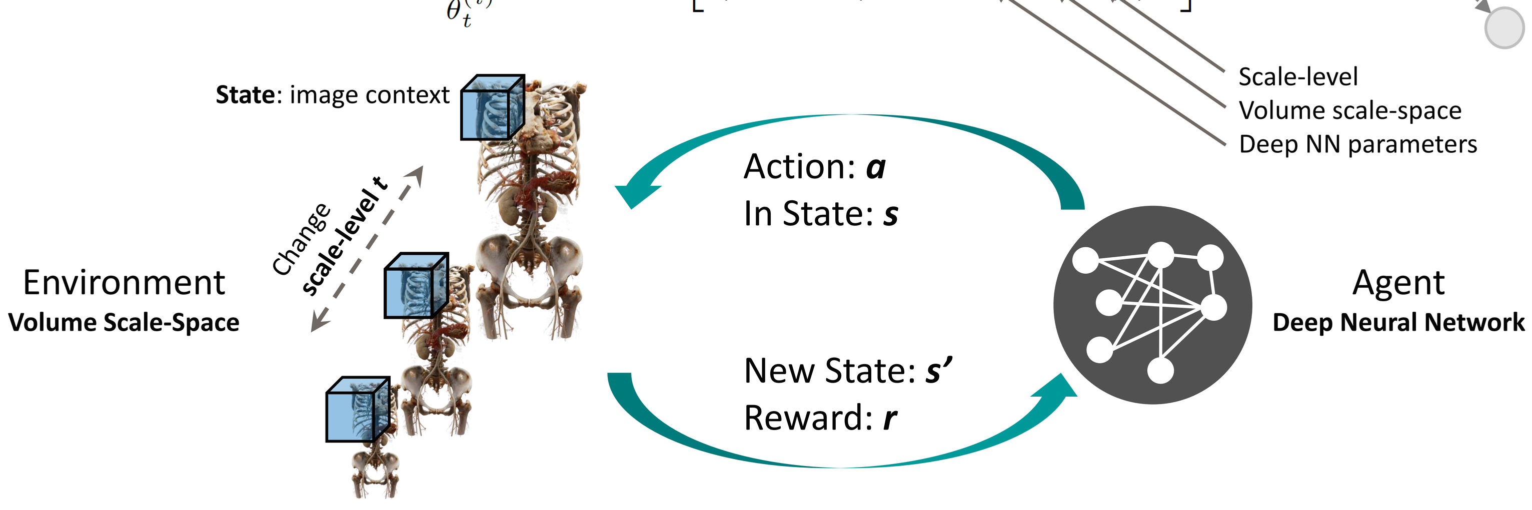
1. Multi-Scale Deep Reinforcement Learning

- Learn navigation policy with deep Q-Learning[1,2]:

$$\hat{\theta}_t^{(i)} = \arg\min_{\theta_t^{(i)}} \mathbb{E}_{s,a,r,s'}\left[\left(y - Q\left(s,a;\theta_t^{(i)} \mid L_d, t\right)\right)^2\right]$$



Scale-level
Volume scale-space
Deep NN parameters



2. Robust Statistical Shape Model (SSM) fitted with MSAC

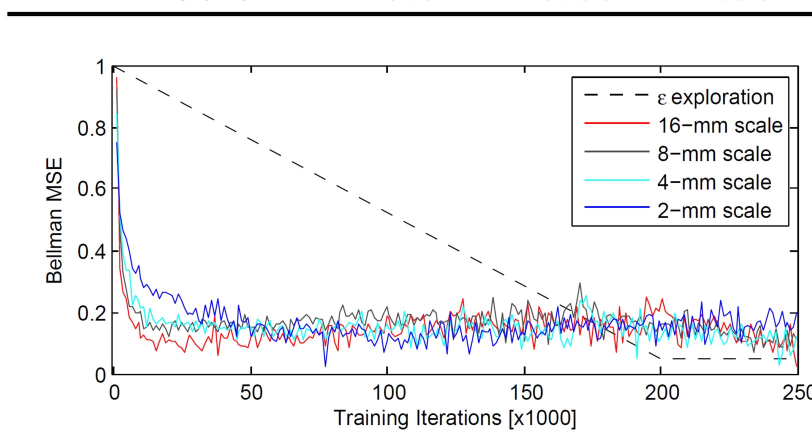
- Model landmark location as multi-variate normal distribution: $p_i \sim \mathcal{N}(\mu_i, \Sigma_i)$
- Robust model fitting with random 3-samples – yield largest consensus set:

$$\hat{S} \leftarrow \arg\min_{S\in I_3(\tilde{P})} \sum_{i=0}^{|\tilde{P}|} \min\left[\frac{1}{Z_i}\left(\phi(\tilde{p}_i)-\mu_i\right)^\top \Sigma_i^{-1}\left(\phi(\tilde{p}_i)-\mu_i\right), 1\right]$$

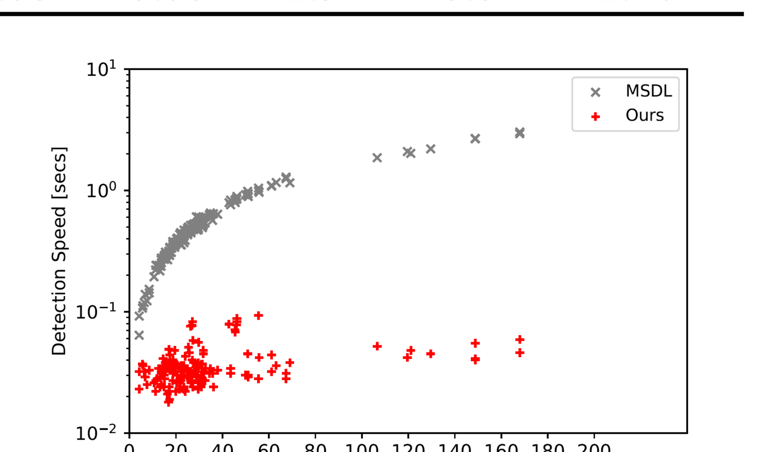
Experiments and Results

- Comprehensive evaluation on 8 landmarks from different anatomical structures (from left to right in table: left/right kidney centers, front corner left/right hip-bones, bifurcation of left common carotid artery, brachiocephalic artery and left subclavian artery, and the bronchial bifurcation)
- Comparison with **Marginal Space Deep Learning**[3]
- FP-rate / FN-rate quantify the accuracy in detecting the presence / absence of landmarks
- MSDL detects the absence of objects using a probability threshold (set to allow 1.5% FP)
- **Significant improvement** – both FP-/FN-rate and accuracy
- **Real-time detection-speed** which scales sub-linearly to volume size

		Anatomical Structures							
		LK	RK	LHB	RHB	LCCA	BA	LSA	BB
FP-rate	MSDL [1]	1.5%	1.5%	1.1%	1.2%	1.0%	1.0%	1.1%	1.0%
	Ours	0%	0%	0%	0%	0%	0%	0%	0%
FN-rate	MSDL [1]	13.9%	9.4%	1.2%	0.4%	10.8%	11.3%	7.2%	4.9%
	Ours	0%	0%	0%	0%	0%	0%	0%	0%
Mean	MSDL [1]	6.17	6.36	4.92	3.66	4.78	5.05	5.25	5.10
	Ours	6.83	6.98	3.61	2.63	4.02	4.26	4.23	4.07
Median	MSDL [1]	5.64	5.80	4.70	3.44	4.17	4.54	4.62	4.53
	Ours	6.32	6.63	2.83	2.49	2.86	3.46	3.21	3.77
STD	MSDL [1]	3.32	3.06	2.09	1.83	3.30	3.02	3.51	2.82
	Ours	3.52	3.83	2.08	1.53	3.33	2.97	3.37	2.16



Bellman Error during training for each scale-level



Correlation between image size and speed

Conclusion

- New learning paradigm for intelligent and robust multi-scale image parsing
- Learn multi-scale navigation paths instead of exhaustive scanning or regression
- Recognize if anatomy is missing and estimate body range based on present anatomy
- Significant improvement over state-of-the-art both in terms of accuracy and speed

References

[1] Mnih, V. et al.: Human-level control through deep reinforcement learning. Nature 518(7540), 2015
[2] Ghesu, F.C. et al.: An artificial agent for anatomical landmark detection in medical images. MICCAI, 2016
[3] Ghesu, F.C. et al.: Marginal Space Deep Learning: Efficient architecture for volumetric image parsing. IEEE TMI 35(5), 2016

usa.siemens.com/abouthealthineers

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG TECHNISCHE FAKULTÄT

SIEMENS Healthineers