Neural Architecture Search (NAS) for Medical Image Segmentation

Dr. Susana Lai-Yuen

Associate Professor

Industrial & Management Systems Engineering

University of South Florida

Visiting Professor

Visual Computing and Machine Intelligence (VCMI)

INESC TEC, Porto

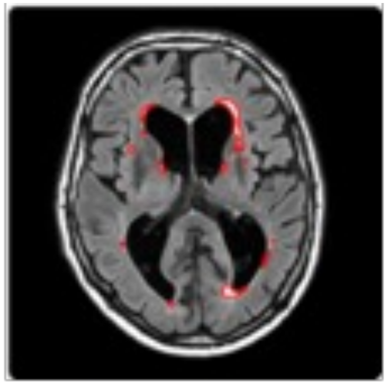
March 20, 2024

Outline

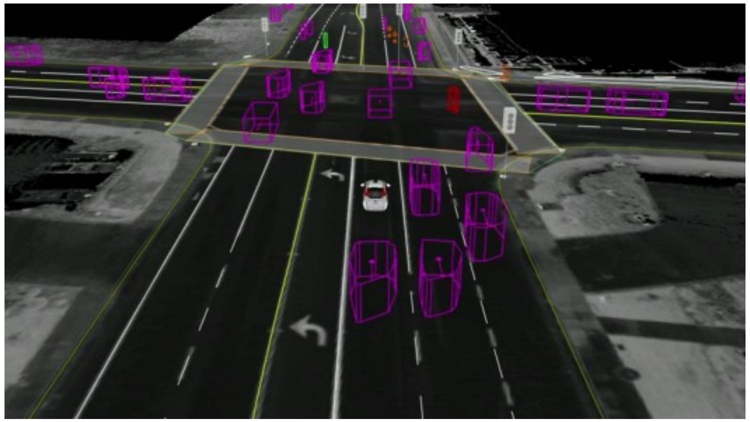
- Neural Architecture Search (NAS)
- Multi-Objective NAS
- **NAS for Medical Image Segmentation**
  - Medical Image Segmentation
  - Efficient Multi-objective NAS (EMONAS-Net)
  - Experiments and Results

Deep Neural Networks

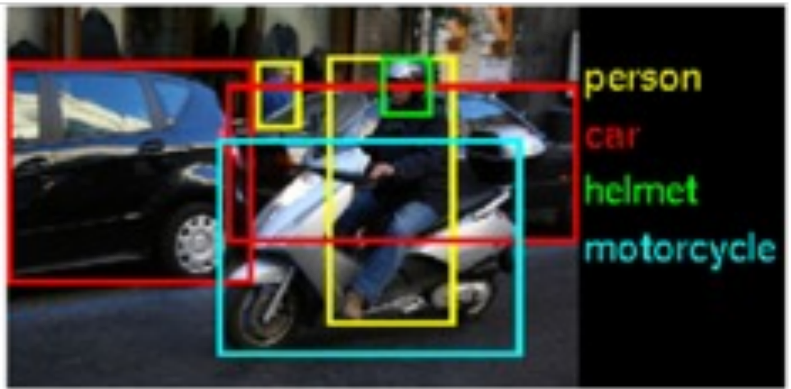
- Automatically learn features from data through multiple layers of abstraction
- Very successful in many applications:
  - Computer vision
  - Object recognition
  - Image segmentation



http://grand-challenge.org



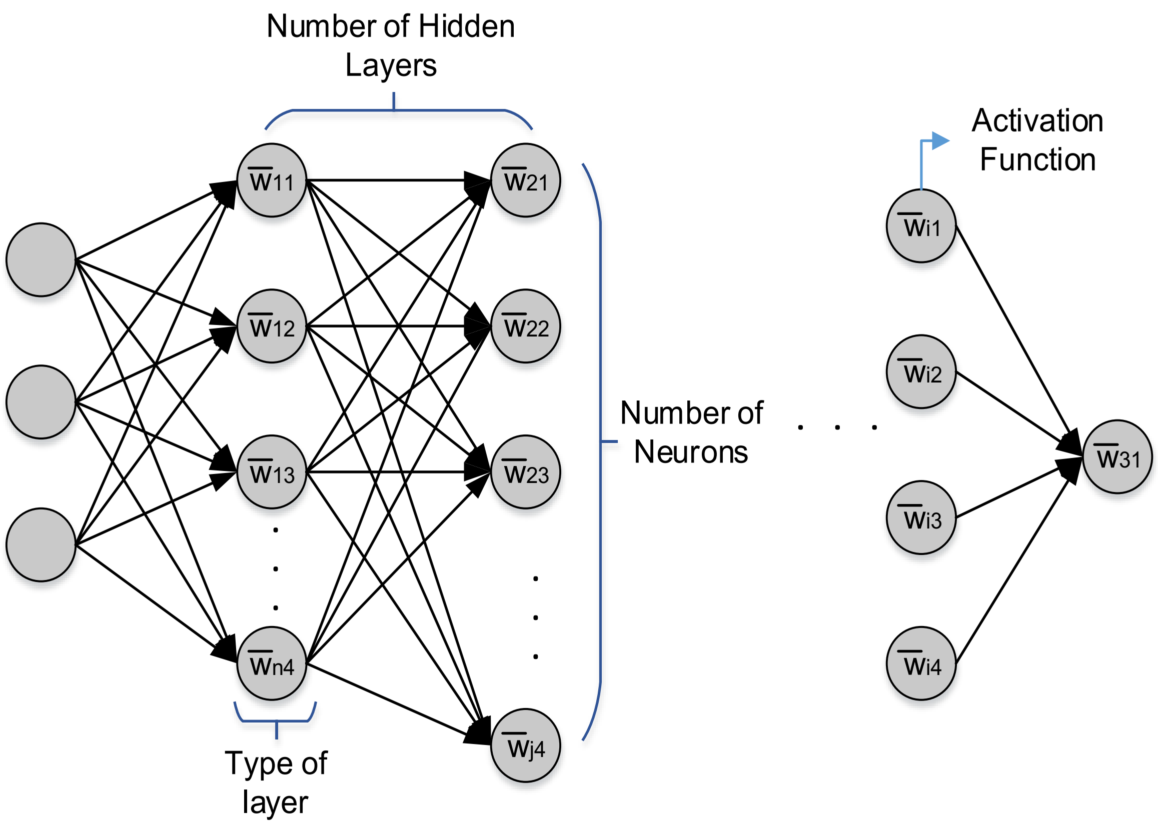
Source: https://www.google.com/selfdrivingcar



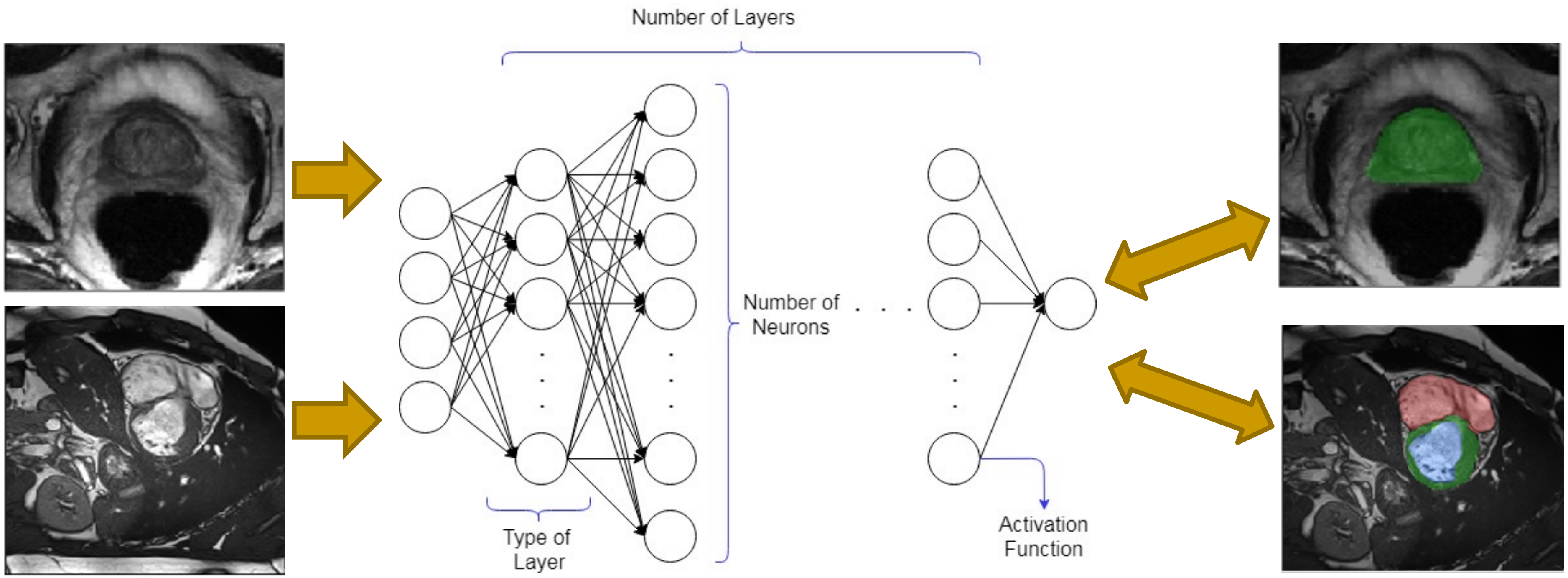
www.image-net.org

Neural Network Architecture

- Performance is highly dependent on architecture
- Extensive number of design parameters
  - Hyperparameters



Neural Network Architecture



- Time and expertise to design the best network architecture
- **Can we learn good architectures automatically?**

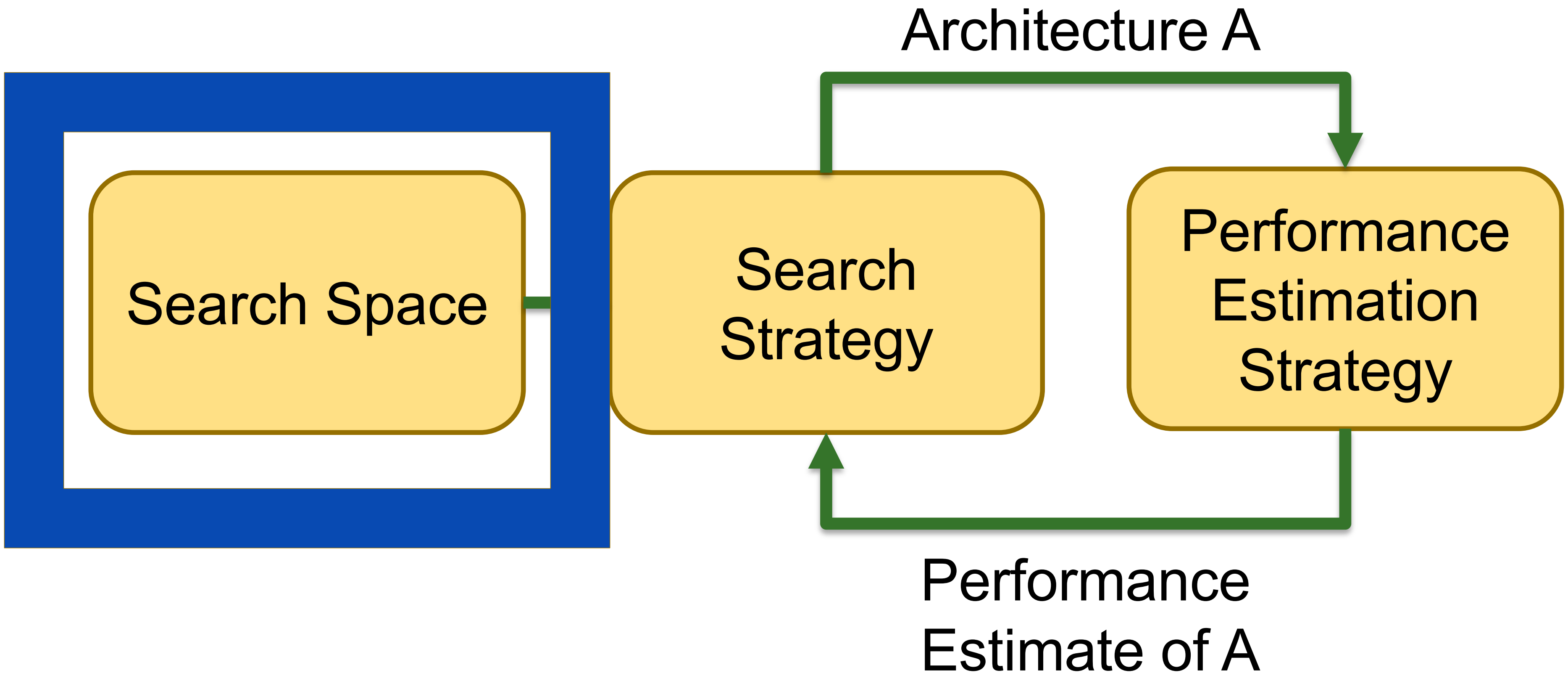
Neural Architecture Search (NAS)

- Automate the design of neural networks through optimization algorithms
  - **Network topology**: nodes connections, operators, etc.
  - **Metrics**: Accuracy, model size, inference time, etc.

- Applications where NAS methods have outperformed manually designed architectures:
  - Image classification [Real 2019]
  - Object detection [Zoph 2018]
  - Semantic segmentation [Chen 2018]

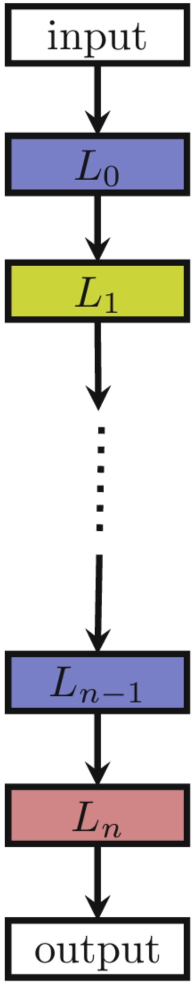
Source: https://www.automl.org/

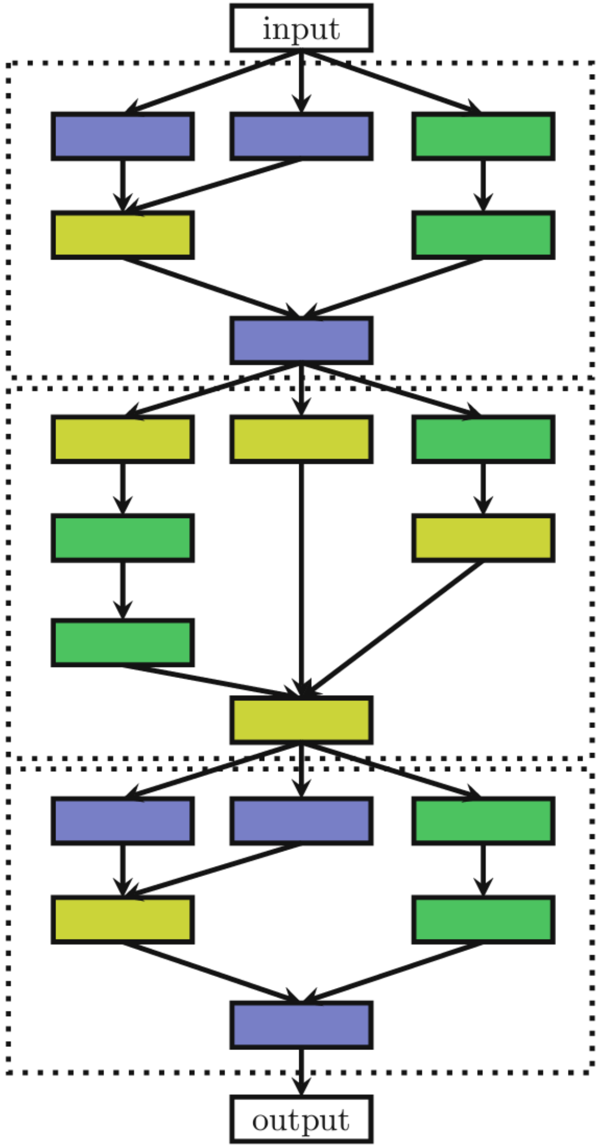
Neural Architecture Search (NAS)



NAS – Search Space

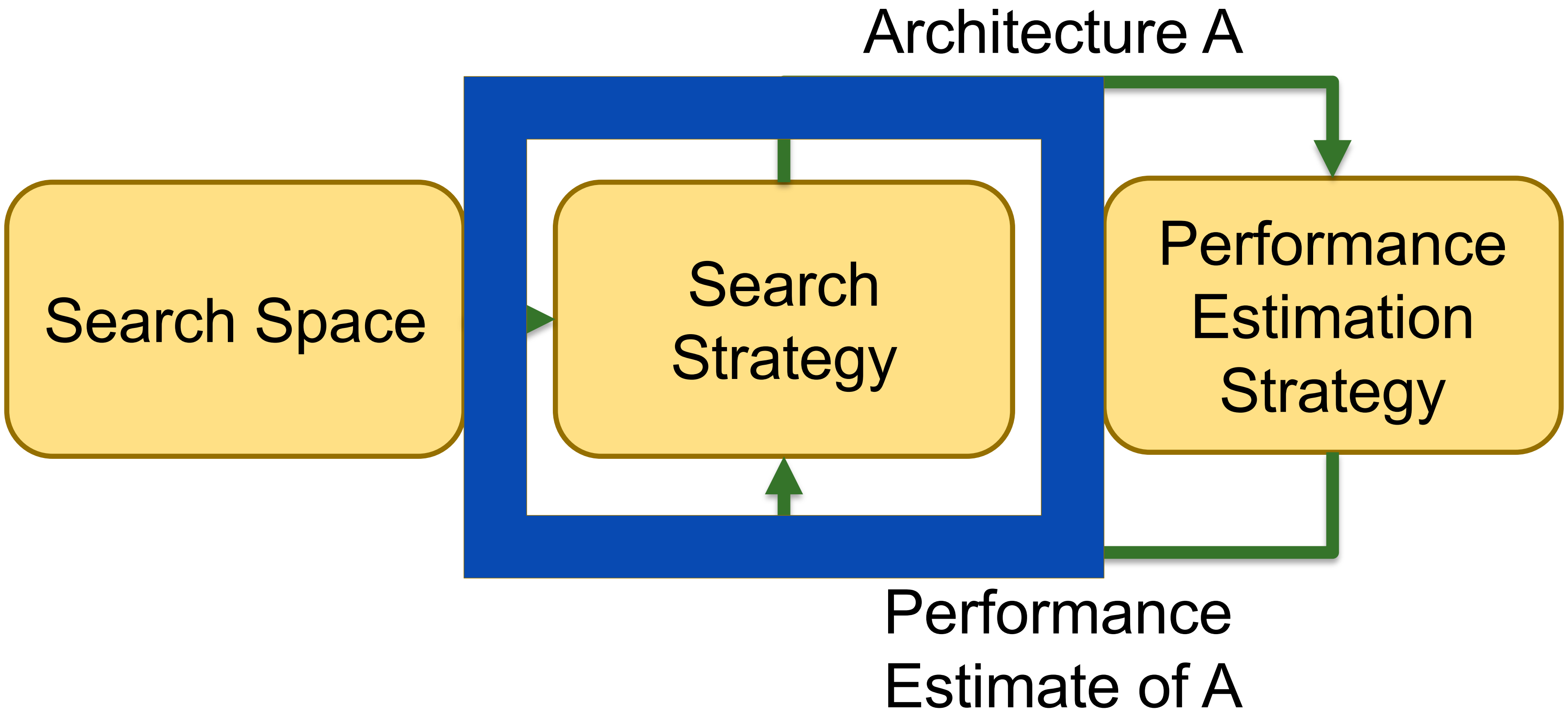
- Defines which architectures can be represented
- Incorporate prior knowledge about types of architectures suitable for the task
  - Reduces the size of search space
  - Can introduce human bias
- Examples:
  - Chain-structured
  - Cell-based





Source: Elsken, T., Metzen, J. H., and Hutter, F., "Neural Architecture Search: A Survey," *JMLR*, 20, 2019.

Neural Architecture Search (NAS)

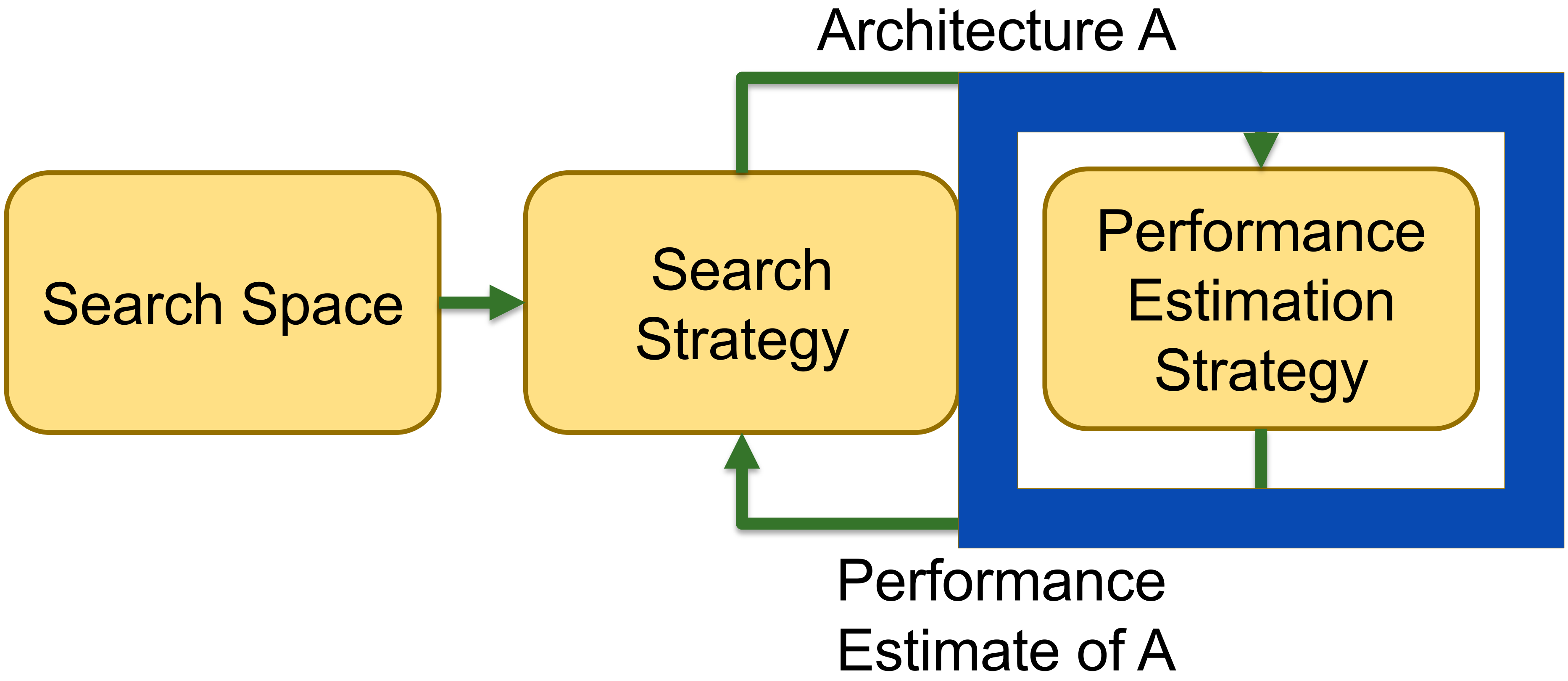


NAS – Search Strategy

- Defines how to explore the search space
  - Exploration vs. exploitation

- Some strategies:
  - Random search
  - Reinforcement learning
  - Gradient-based methods
  - Evolutionary methods

Neural Architecture Search (NAS)



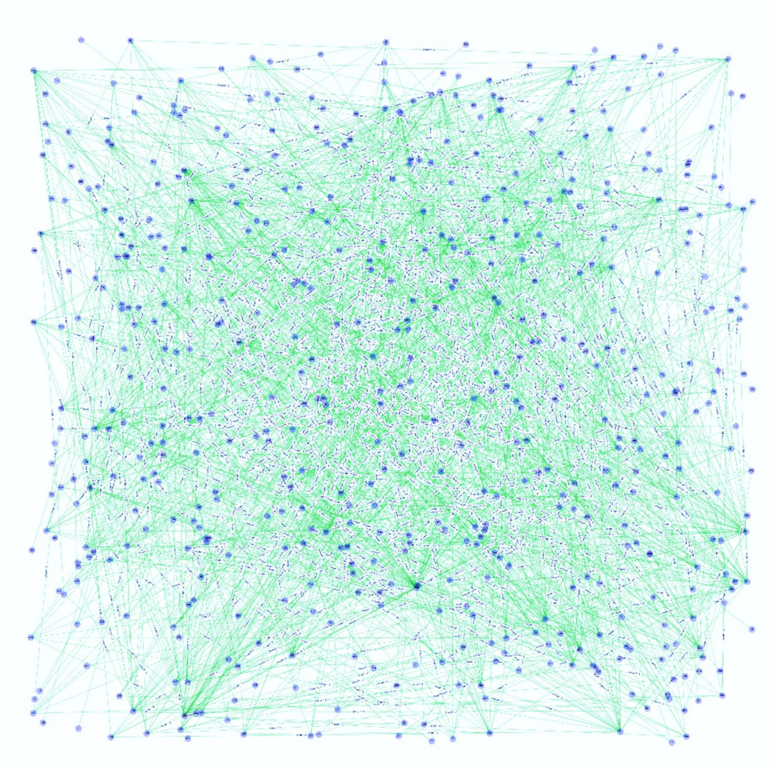
NAS – Performance Estimation Strategy

- Estimate the performance of the candidate architecture
  - Cannot train all the candidate architectures from scratch

- Some strategies:
  - Lower fidelity estimates
  - Network morphisms
  - Weight sharing
  - Surrogate models

Neural Network Efficiency

- Current architectures can have tens of millions of parameters
  - Over-parametrized
  - Highly inefficient
  - Overfitting



Multi-Objective NAS

- NAS is usually modeled as a single objective optimization problem
  - Maximize prediction accuracy

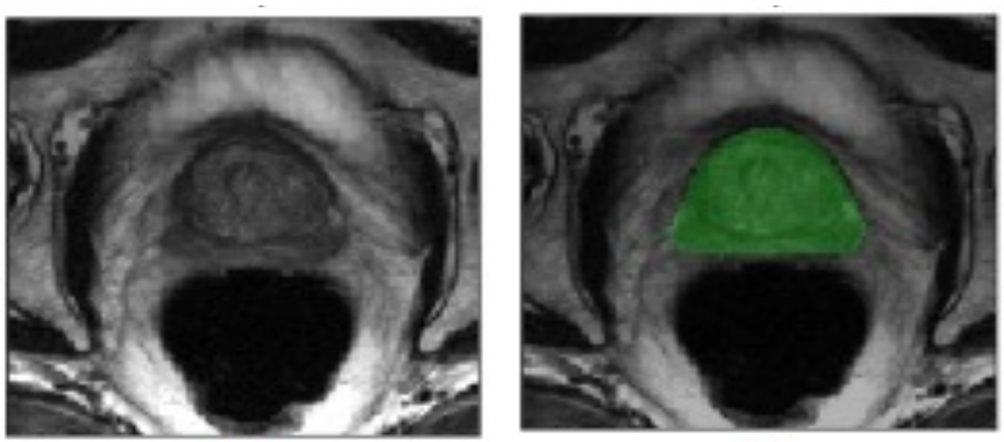
- **Multi-objective NAS** aims to identify architectures that optimize multiple objectives:
  - Maximize model’s performance
  - Minimize model’s size
  - Or other objectives

Efficient Multi-Objective NAS (EMONAS-Net) for Medical Image Segmentation

Baldeon-Calisto, M., and S. Lai-Yuen. “EMONAS-Net: Efficient Neural Architecture Search Framework using Surrogate-Assisted Evolutionary Algorithm." *Artificial Intelligence in Medicine,* 119, 2021.

Medical Image Segmentation

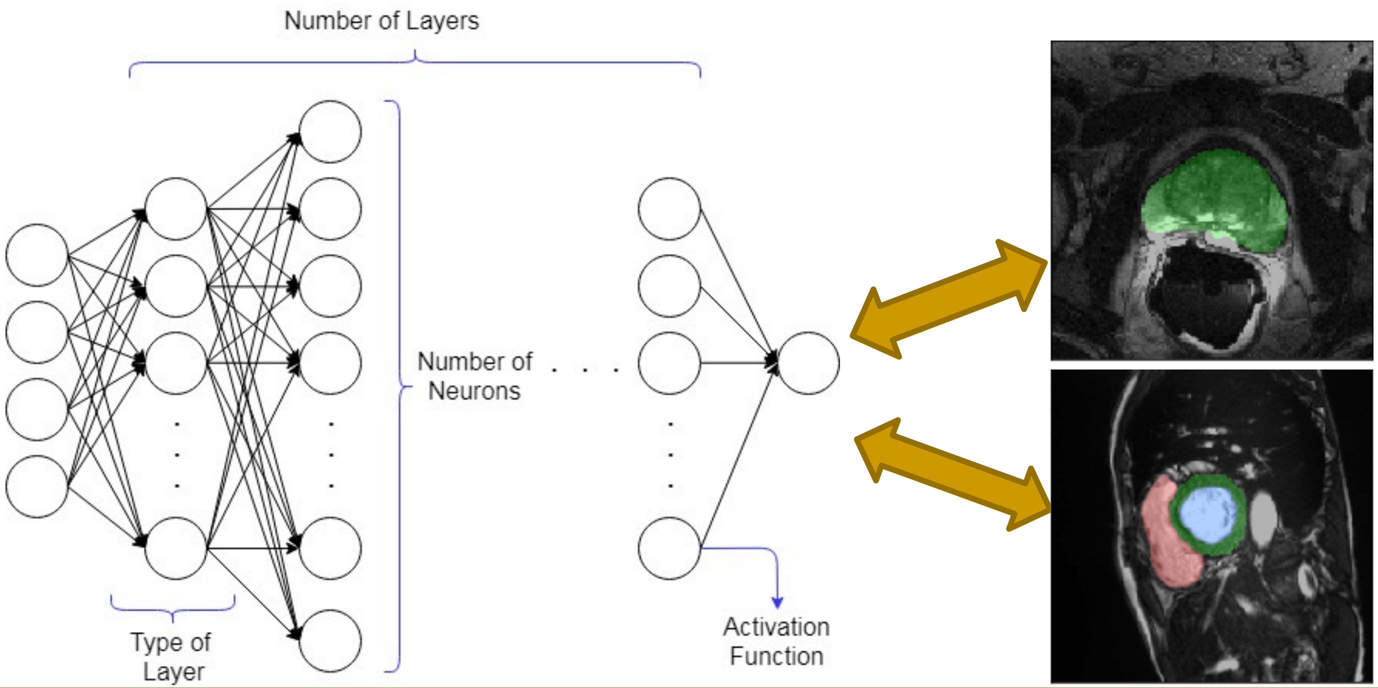
- Critical task for identifying structures in medical images
  - Studying anatomical structures
  - Identifying tumors and lesions
  - Assist in treatment planning before radiation therapy



Prostate segmentation

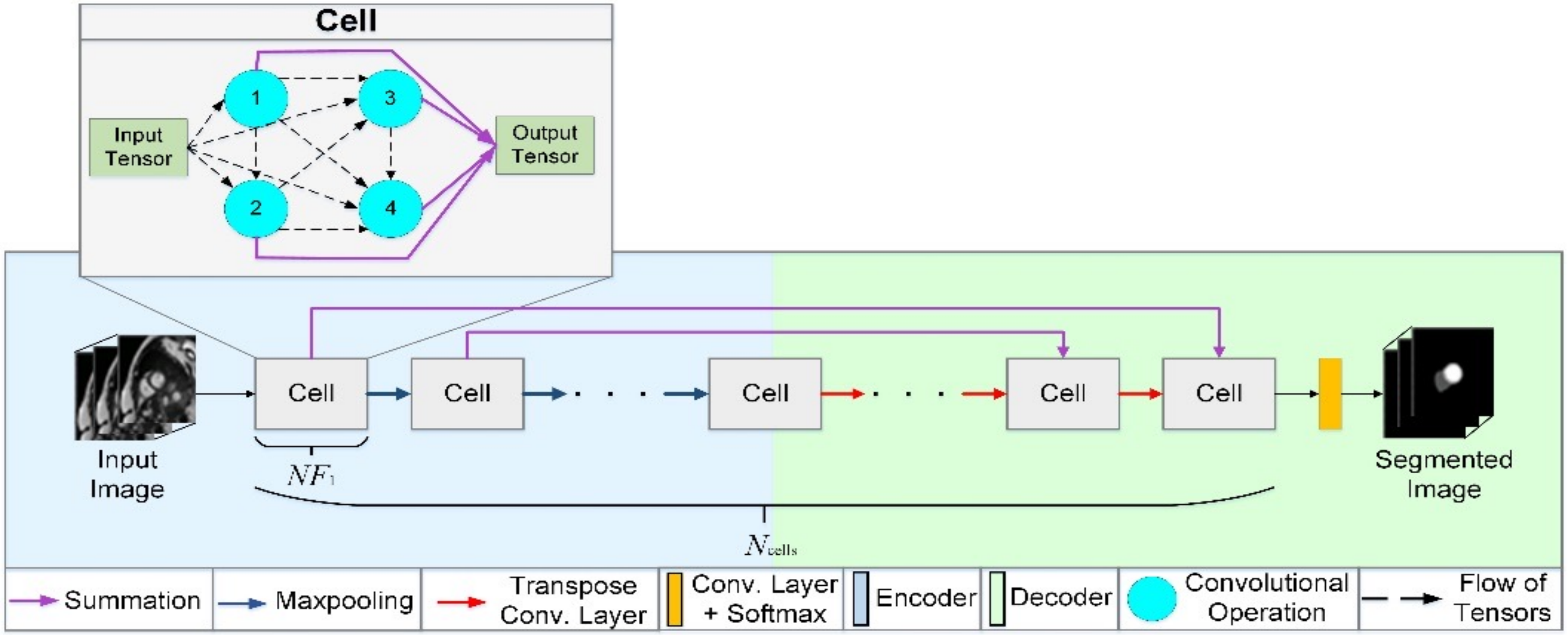
Efficient Multi-Objective NAS (EMONAS-Net)

- Design an efficient multi-objective NAS framework for medical image segmentation
  - Automatically identifies accurate and small networks for a particular dataset
  - Improves search efficiency



Efficient Multi-objective NAS (EMONAS-Net)

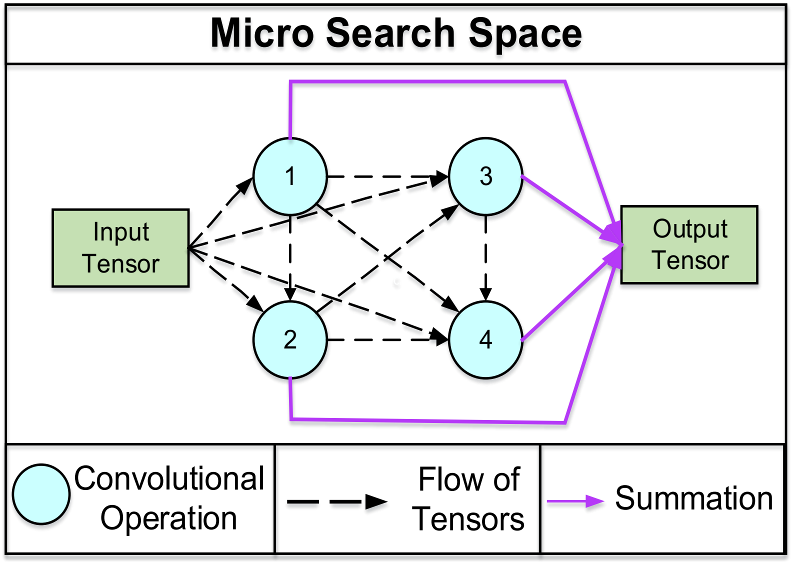
1. Novel search space
2. Surrogate-assisted Multi-objective Evolutionary based Algorithm (SaMEA)
   - Finds the best micro and macro hyperparameters

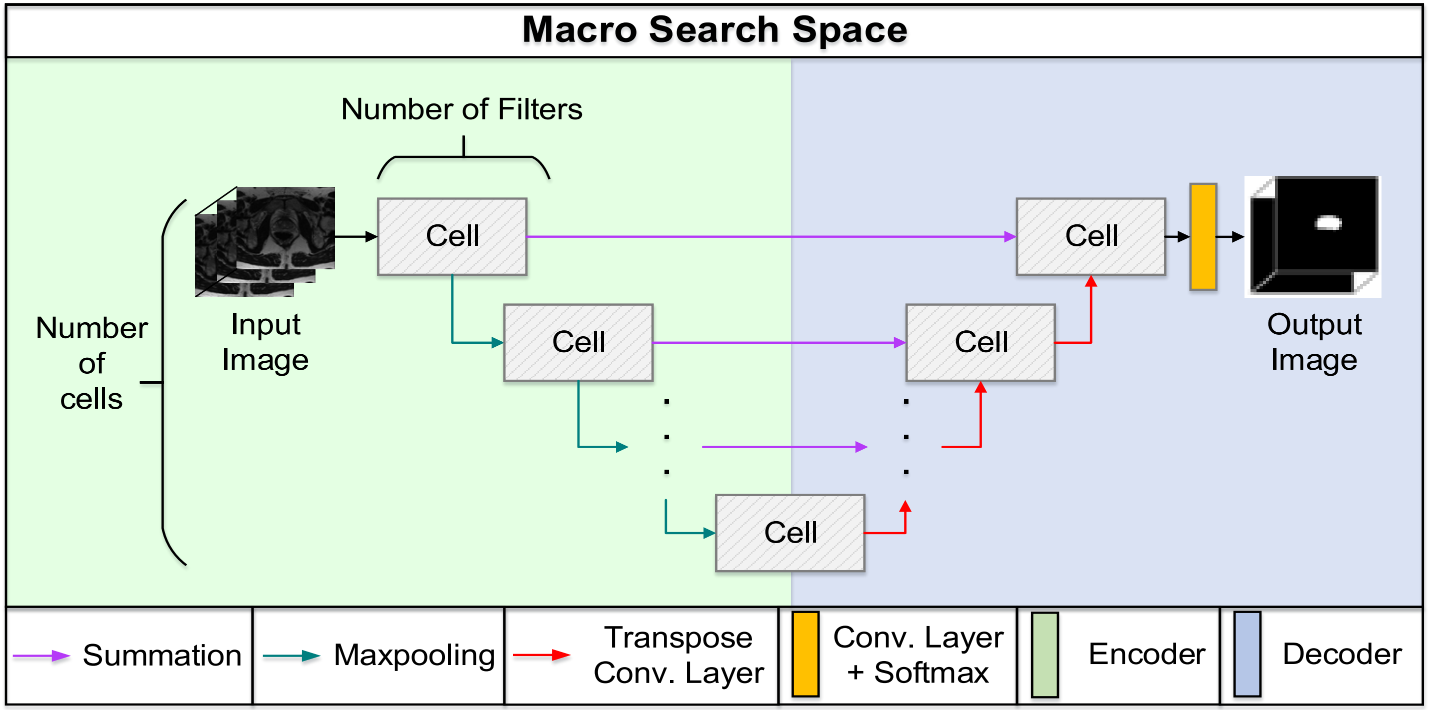


EMONAS-Net framework components

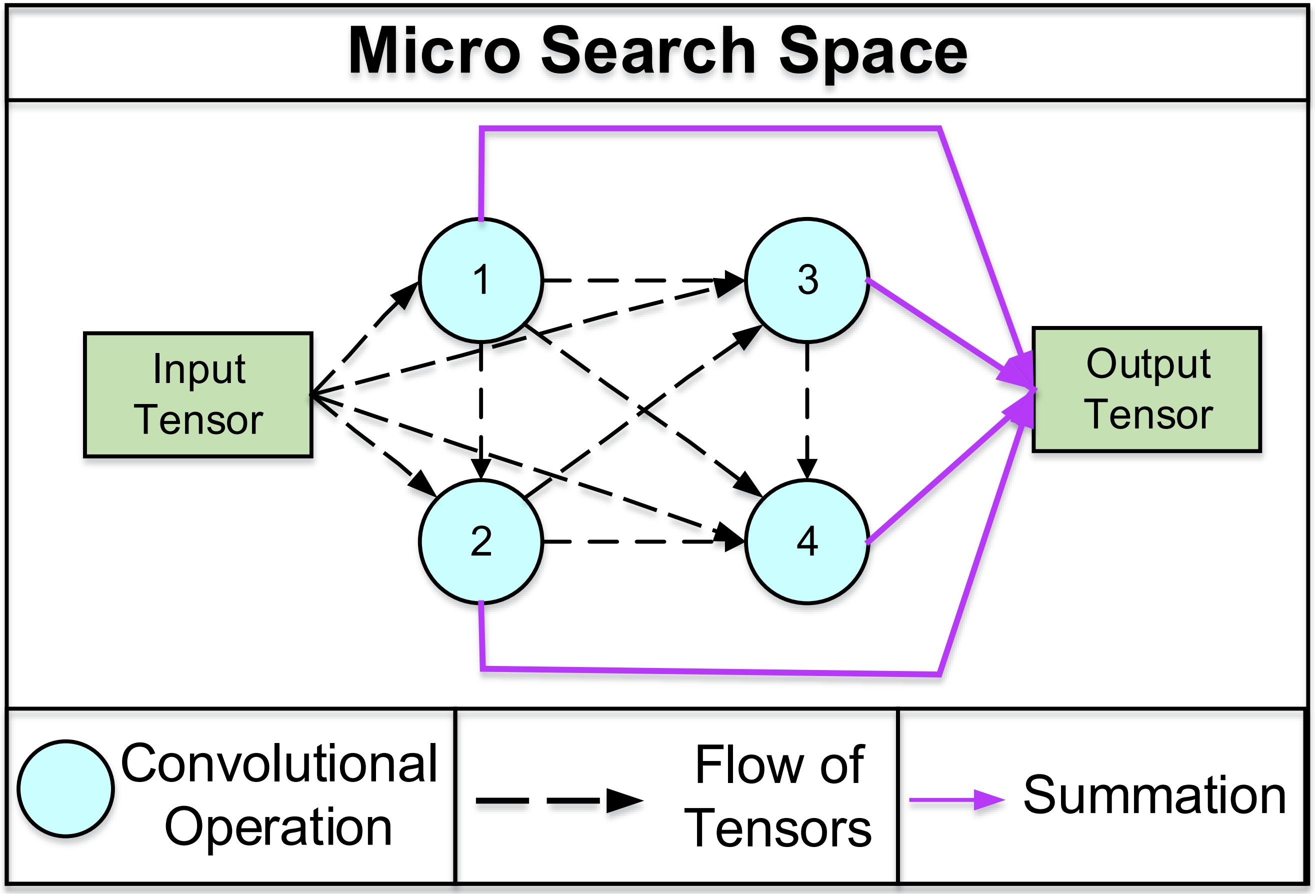
Search Space

- Macro and micro search space to be optimized





Micro Search Space



Cell represented as a directed acyclical graph.

Decision Variables

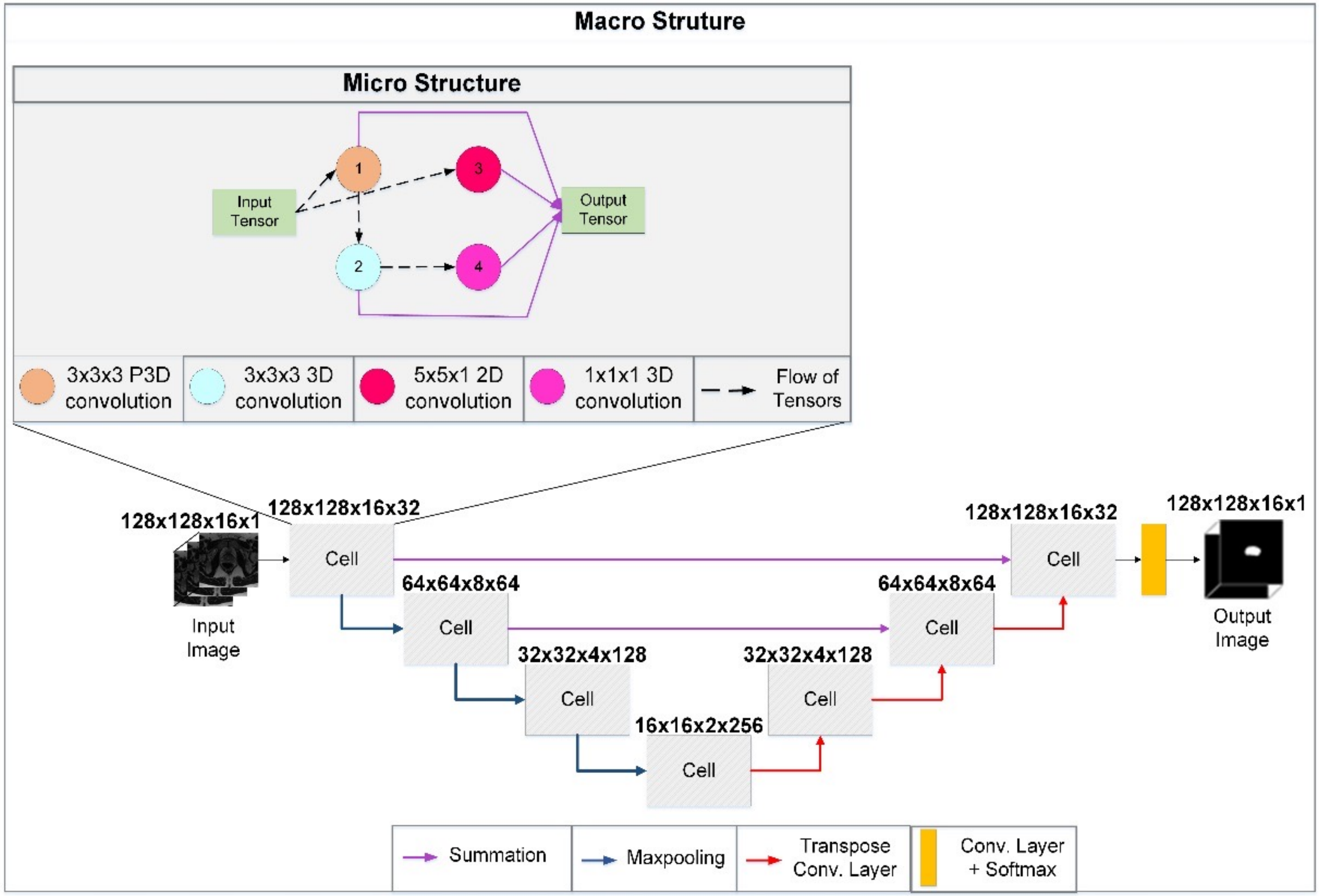
• 1×1×1 3D convolution
• 3×3×3 3D convolution
• 5×5×5 3D convolution
• 3×3×1 2D convolution
• 5×5×1 2D convolution
• 7×7×1 2D convolution
• 3×3×3 Pseudo-3D convolution
• 5×5×5 Pseudo-3D convolution
• 7×7×7 Pseudo-3D convolution
• identity

Table 1. Set of possible convolutional operations in a node.

Decision Variable	Formula	Search Range
Input to node 2 (I_2)	-	[Input tensor, node 1]
Input to node 3 (I_3)	-	[Input tensor, node 1, node 2]
Input to node 4 (I_4)	-	[Input tensor, node 1, node 2, node 3]
Type of Convolutional operation in node 1 (O_1)	-	Refer to Table 1.
Type of Convolutional operation in node 2 (O_2)	-	Refer to Table 1.
Type of Convolutional operation in node 3 (O_3)	-	Refer to Table 1.
Type of Convolutional operation in node 4 (O_4)	-	Refer to Table 1.
Number of cells (N_{cells})	$2n_c + 1$	$n_c \in [2,3,4]$
Number of filters for NF_1	2^{n_f}	$n_f \in [3,4,5]$
Learning rate	-	$[1 \times 10^{-5}, 9 \times 10^{-3}]$

Table 2. Decision variables constituting the hyperparameter search space and search range.

Decision Variables

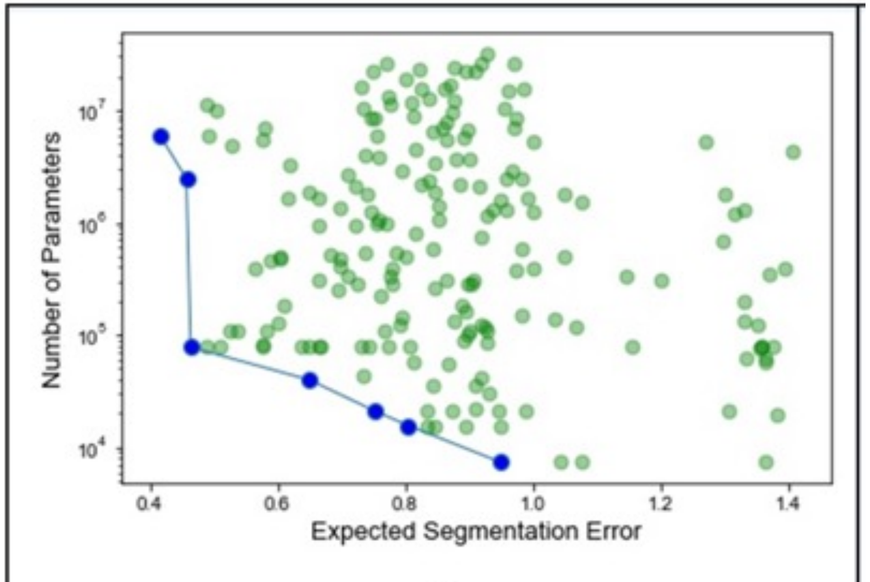


Hyperparameter Vector

Node 1	Input tensor	Node 2	3x3x3 P3D	3x3x3 3D	5x5x1 2D	1x1x1 3D	7	32	3×10^{-4}
Input node 2	Input node 3	Input node 4	Conv. operation node 1	Conv. operation node 2	Conv. operation node 3	Conv. operation node 4	Number of cells	Number of filters	Learning rate

Surrogate Assisted Multi-objective Evolutionary Algorithm (SaMEA)

- MOEA/D based algorithm with penalty-based boundary intersection (PBI) approach
  - Selection probabilities
  - Random forest surrogate



a)

UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING

SaMEA Algorithm

Expected Segmentation Error

$$Minimize\ f_1(x) = \alpha\left(C - MCDice_{Train}(\theta)\right) + \left(C - MCDice_{Val}(\theta)\right) + \beta\left(\frac{E - e_{max}}{E}\right)$$

Number of Trainable Parameters

$$Minimize\ f_2(x) = \log\left(|\theta|\right)$$

$$subject\ to\ x \in \Omega$$

Ω = hyperparameter search space
θ = network parameters
$\mathrm{MCDice}(\theta)$ = multiclass Dice similarity coefficient
α, β = weight parameters
E = maximum number of training epochs
e_{max} = epoch with the maximum validation $\mathrm{MCDice}(\theta)$

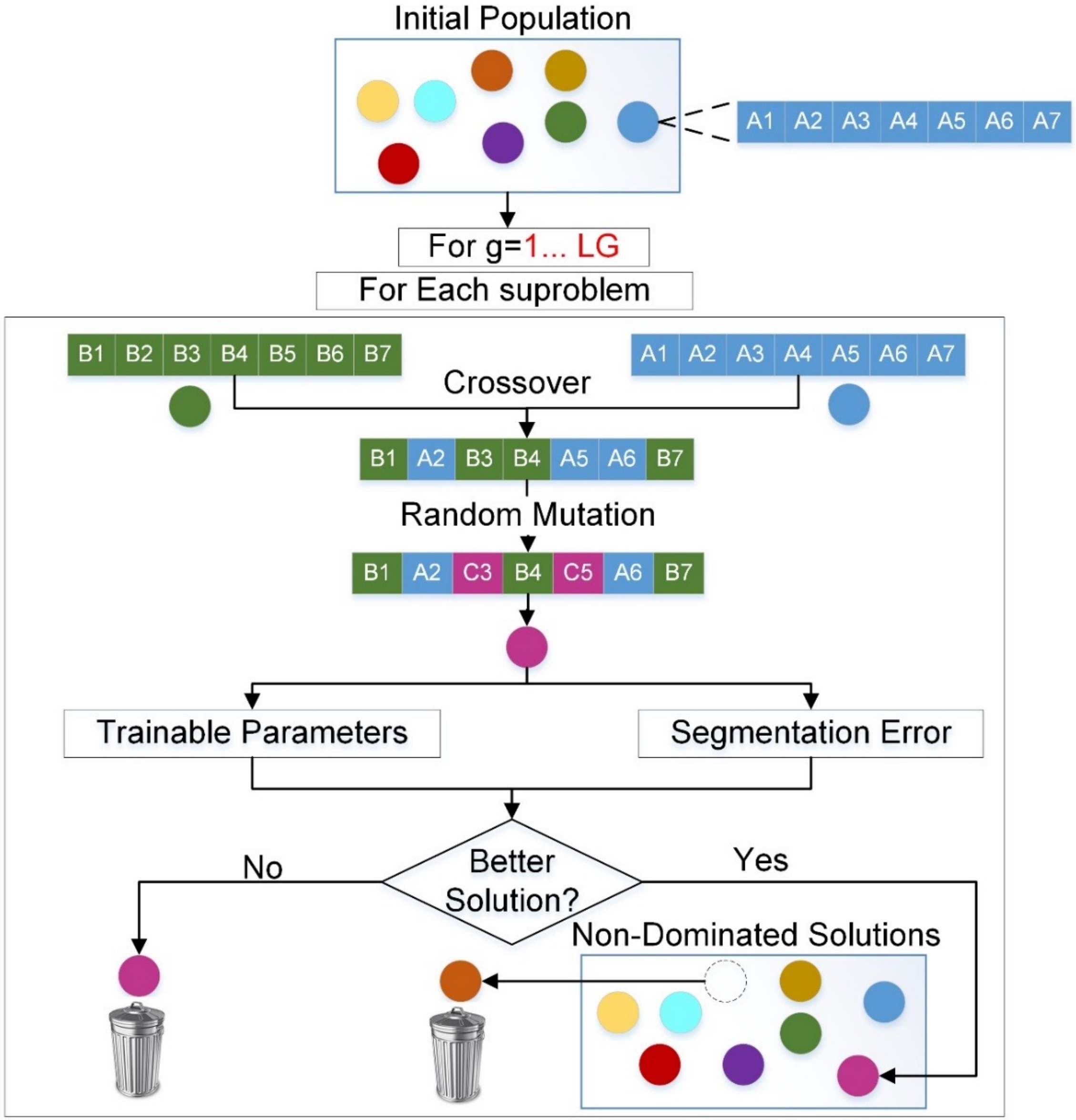
SaMEA Algorithm

- Three phases:
  1. Initialization phase
  2. Learning phase
  3. Exploitation phase

Exploring the search space

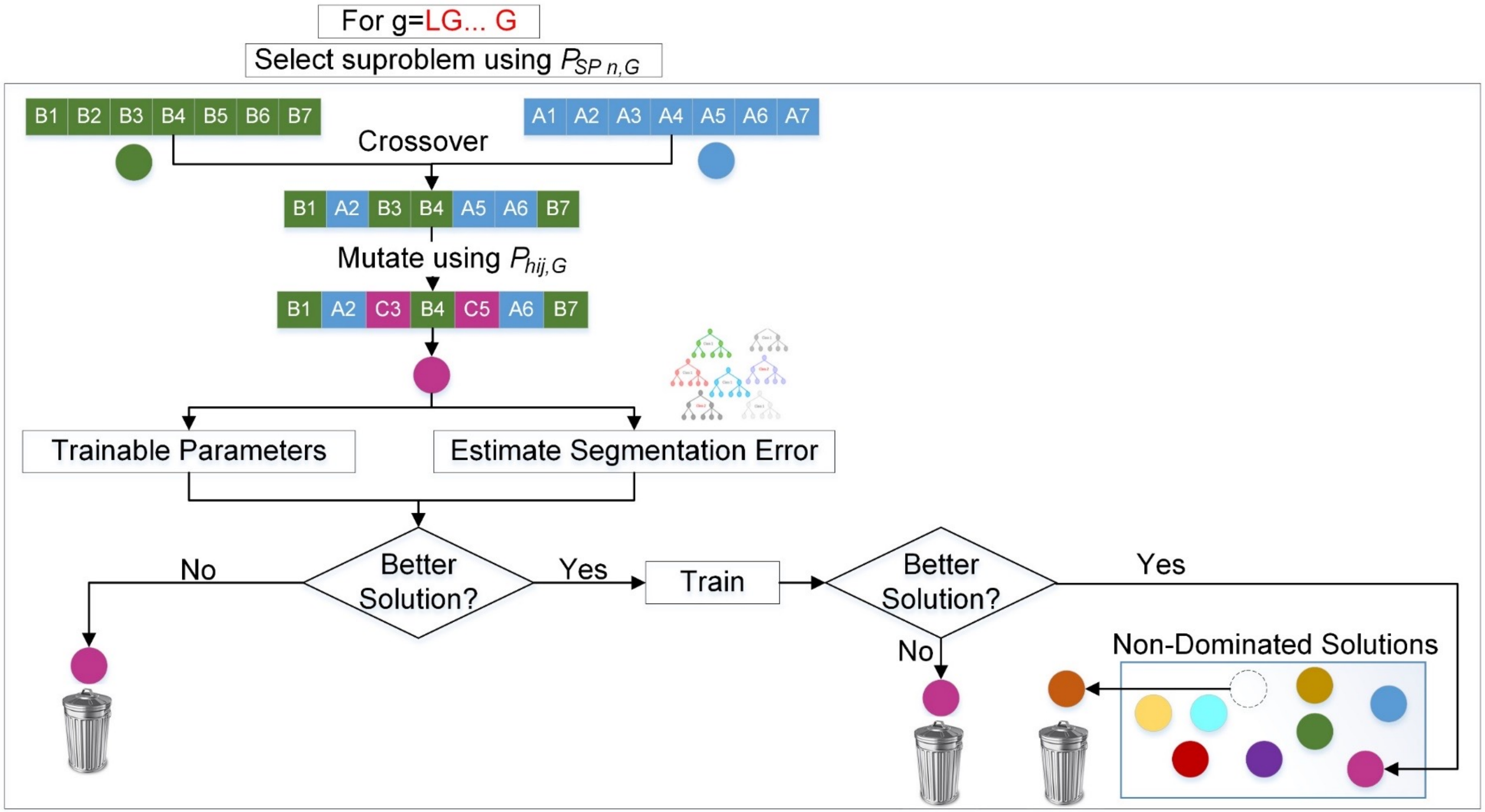
Exploiting the search space (Implement selection probabilities, Random Forest surrogate)

Initialization and Learning Phase



Phase I and II of the SaMEA Algorithm

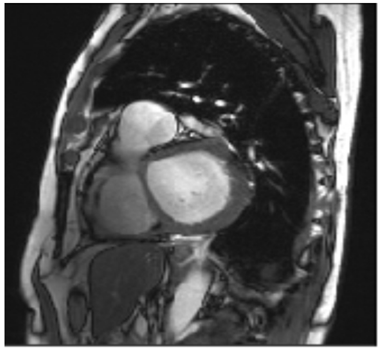
Exploitation Phase



Phase III of the SaMEA Algorithm

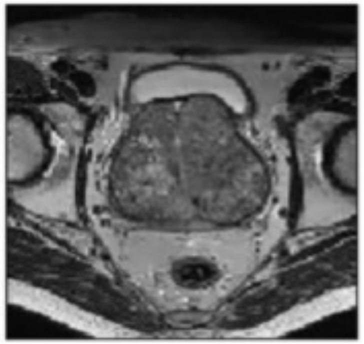
Experiments

Dataset		
Cardiac	**Hippocampus**	**Prostate**
• ACDC Grand Challenge • Left ventricle cavity • Left ventricle myocardium • Right ventricle cavity	• Medical Segmentation Decathlon • Posterior part • Anterior part	• PROMISE12 challenge • Prostate



(Bernard, 2018)

(Simpson, 2019)



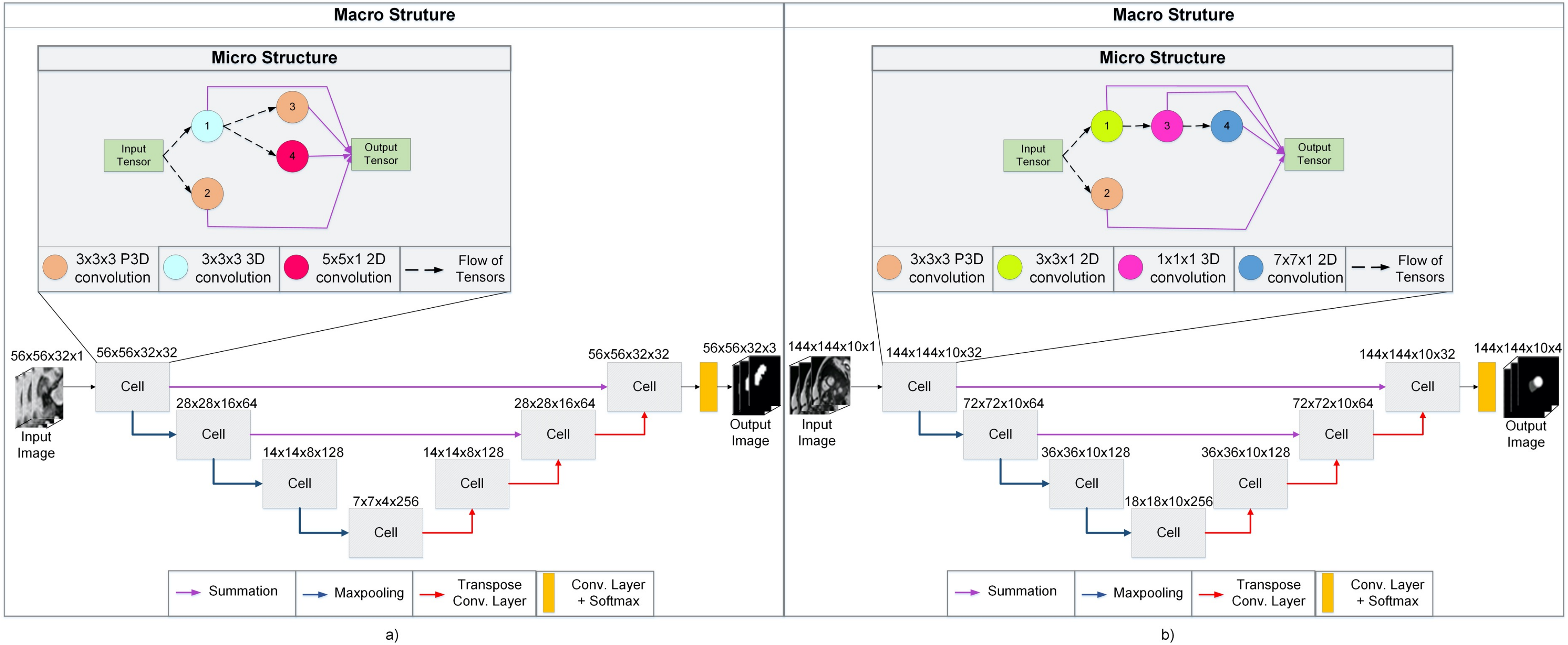
(Litjens, 2014)

Prostate Dataset: https://promise12.grand-challenge.org/Home/
Cardiac Dataset: https://acdc.creatis.insa-lyon.fr/#challenge/
Hippocampus Dataset: https://decathlon-10.grand-challenge.org/

Best Hyperparameter Values for Datasets

Decision Variable	Prostate Dataset	Hippocampus Dataset	Cardiac Dataset
Input to node 2 (I_2)	Node 1	Input Tensor	Input Tensor
Input to node 3 (I_3)	Input Tensor	Node 1	Node 1
Input to node 4 (I_4)	Node 2	Node 1	Node 3
Type of Convolutional operation in node 1 (O_1)	3×3×3 P3D	3×3×3 3D Conv.	3×3×1 2D Conv.
Type of Convolutional operation in node 2 (O_2)	3×3×3 3D Conv.	3×3×3 P3D	3×3×3 P3D
Type of Convolutional operation in node 3 (O_3)	5×5×1 2D Conv.	3×3×3 P3D	1×1×1 3D Conv.
Type of Convolutional operation in node 4 (O_4)	1×1×1 3D Conv.	5×5×1 2D Conv.	7×7×1 2D Conv.
Number of Cells (N_{cells})	7	7	7
Number of filters for NF_1	32	32	32
Learning Rate	3×10^{-4}	4×10^{-4}	4×10^{-4}
Number of Parameters	5.9×10^{6}	7.0×10^{6}	7.1×10^{6}

Best Architectures for Hippocampus and Cardiac Datasets



Benchmark Results: Cardiac Dataset

Group	Method	Right Ventricle Cavity				Left Ventricle Cavity				Left Ventricle Myocardium				Size of the Network
		HD [mm]	Rank	DSC [%]	Rank	HD [mm]	Rank	DSC [%]	Rank	HD [mm]	Rank	DSC [%]	Rank	
EMONAS-1	Automatic	11.34	4	91.24	2	7.81	7	92.59	11	8.72	3	87.59	11	7.1×10^6
EMONAS-5	Automatic	11.18	2	91.20	3	7.35	5	93.10	7	8.92	4	88.00	10	35.5×10^6
Isensee et al.	Automatic	9.93	1	92.75	1	6.20	1	94.75	1	7.17	1	91.35	1	235.5×10^6
Baldeon et al.	Automatic	11.21	3	90.99	5	7.12	2	93.03	9	8.26	2	88.40	9	24.0×10^6
Zotti et al.	Manual	12.19	5	91.15	6	7.28	4	93.80	4	9.44	7	89.40	3	-
Zotti et al.	Manual	11.85	6	90.95	4	7.67	6	93.10	7	8.99	5	89.00	5	-
Painchaud et al.	Manual	13.52	8	90.85	7	7.22	3	93.60	6	9.12	6	88.90	7	-
Baumgartner et al.	Manual	13.68	9	90.75	8	785	8	93.70	5	9.67	8	89.65	2	-

Table 3. Evaluation metrics for the top competing methods in the ACDC challenge.

DSC: Dice Similarity Coefficient
HD: Hausdorff Distance

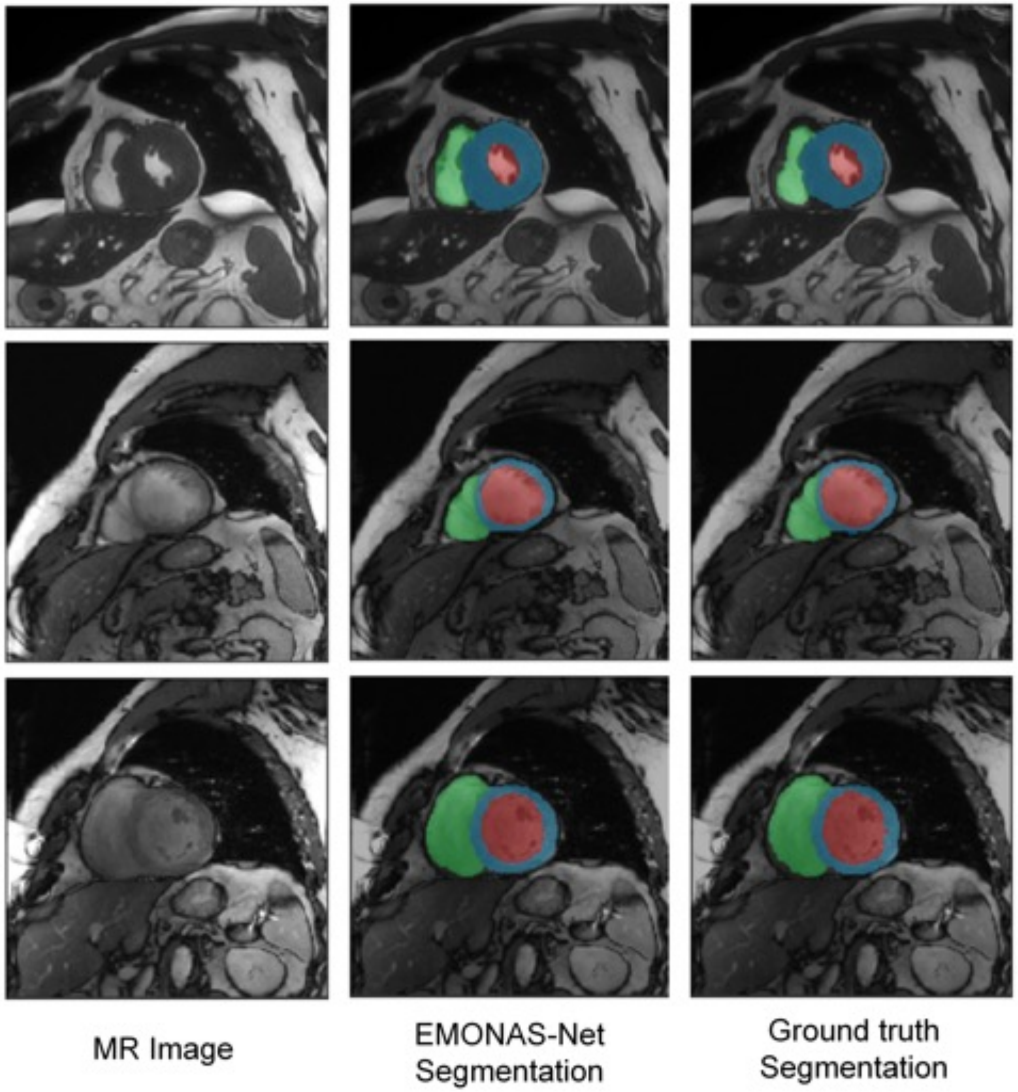
Efficiency Evaluation: Cardiac Dataset

NAS Method	Right Ventricle Cavity		Left Ventricle Cavity		Left Ventricle Myocardium		Size of the Network	GPU days	GPU Specification
	HD [mm]	DSC [%]	HD [mm]	DSC [%]	HD [mm]	DSC [%]			
EMONAS-1	11.3	91.2	7.8	92.6	8.7	87.6	7.1×10^6	11.4	1 GTX- 1070 Ti
Baldeon et al.	11.6	91.0	7.7	92.7	9.4	87.5	7.0×10^6	23.8	1 GTX- 1070 Ti
Mortazi et al.	14.3	86.8	8.9	92.8	10.7	84.9	30.0×10^6	10	**15** GTX Titan X

Table 4. Efficiency and performance metrics for competing NAS methods in the ACDC challenge.

DSC: Dice Similarity Coefficient
HD: Hausdorff Distance

Results: Cardiac Segmentation



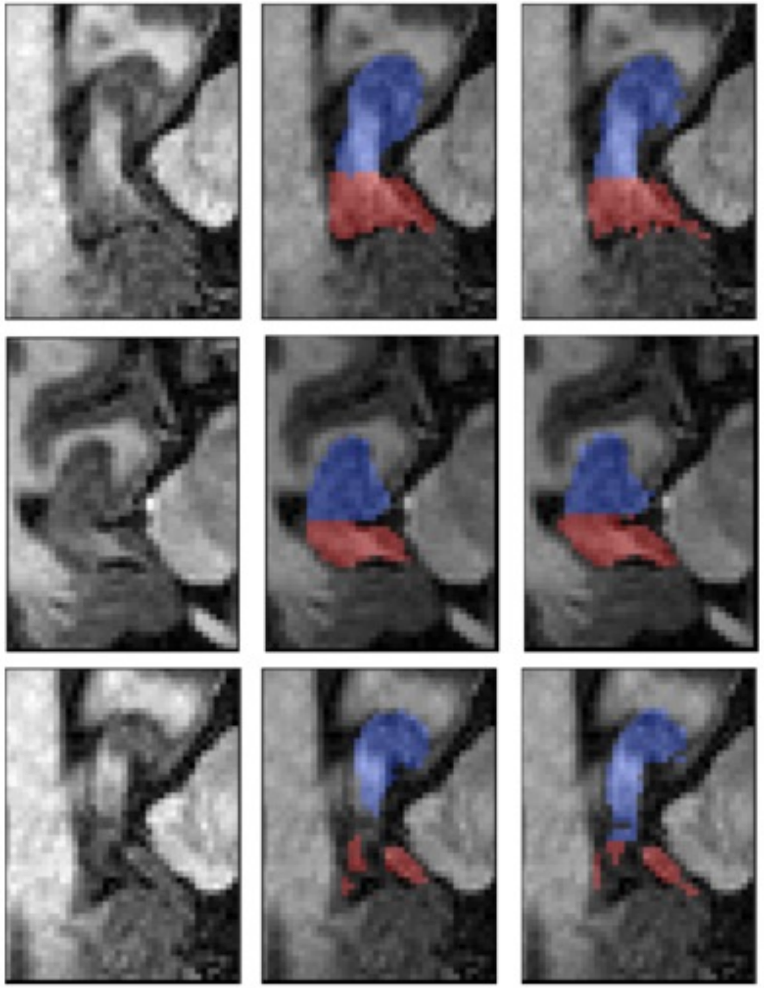
Benchmark Results: Hippocampus Dataset

Model	Method	DSC [%]		Size of the Network	GPU days	GPU Specification
		ANT	POS			
EMONAS-Net	Automatic	89.08±0.39	87.54±0.34	7.0×10^6	4.22	1 GTX- 1070 Ti
Baldeon et al.	Automatic	88.53±0.36	87.02±0.34	5.6×10^6	8.77	1 GTX- 1070 Ti
Yu et al.	Automatic	83.09	82.97	17.0×10^6	5	64 NVIDIA V100
Isensee et al.	Automatic	89.87	88.20	-	-	-
Wang et al.	Manual	87.93	88.69	-	NA	NA
Wang et al.	Manual	87.48	84.92	-	NA	NA
Wang et al.	Manual	85.75	88.29	-	NA	NA

Table 5. Evaluation metrics for competing methods in the Medical Decathlon Challenge dataset.

DSC: Dice Similarity Coefficient
ANT: Anterior Part of the Hippocampus
POS: Posterior Part of the Hippocampus

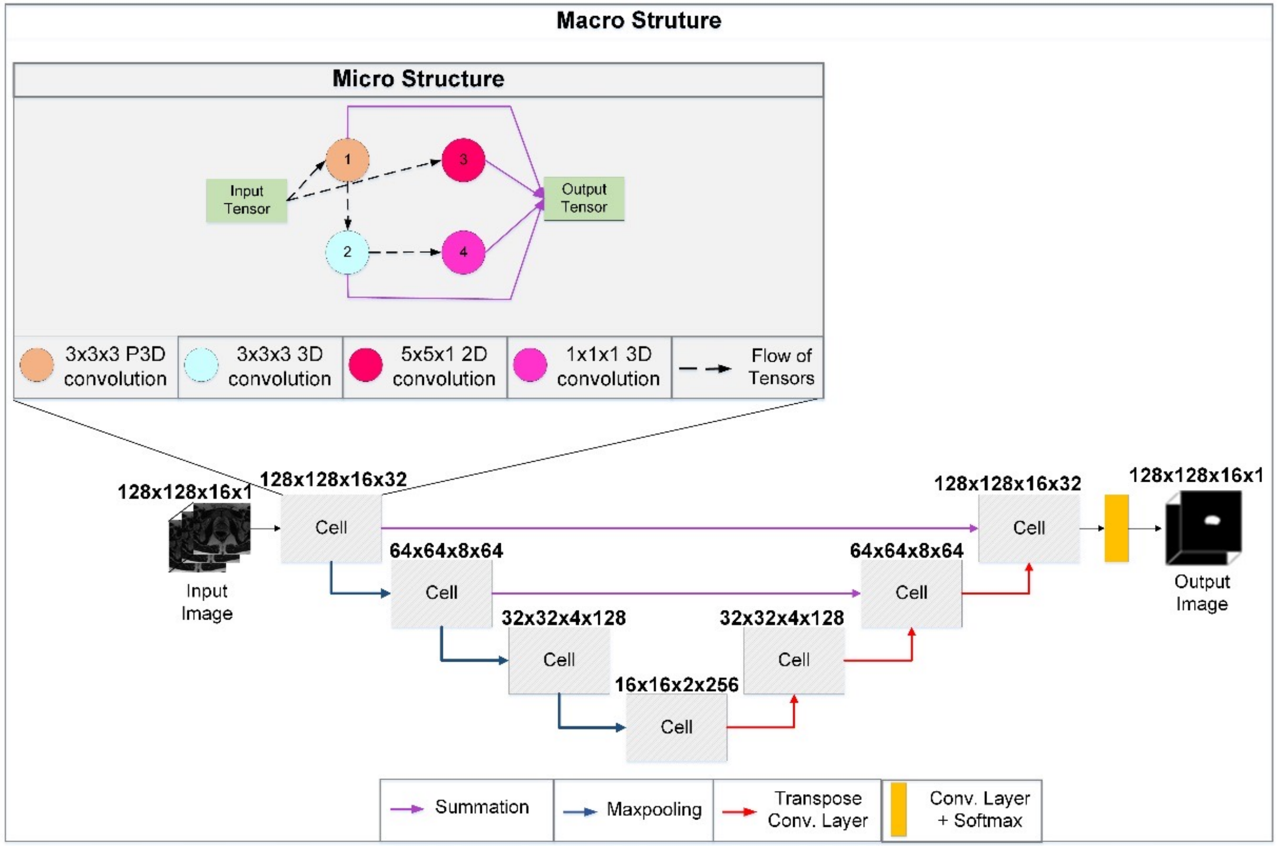
Results: Hippocampus Segmentation



MR Image | EMONAS-Net Segmentation | Ground truth Segmentation

UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING

Best Architecture for Prostate Dataset



Hyperparameter Vector

Node 1	Input tensor	Node 2	3x3x3 P3D	3x3x3 3D	5x5x1 2D	1x1x1 3D	7	32	3×10^{-4}
Input node 2	Input node 3	Input node 4	Conv. operation node 1	Conv. operation node 2	Conv. operation node 3	Conv. operation node 4	Number of cells	Number of filters	Learning rate

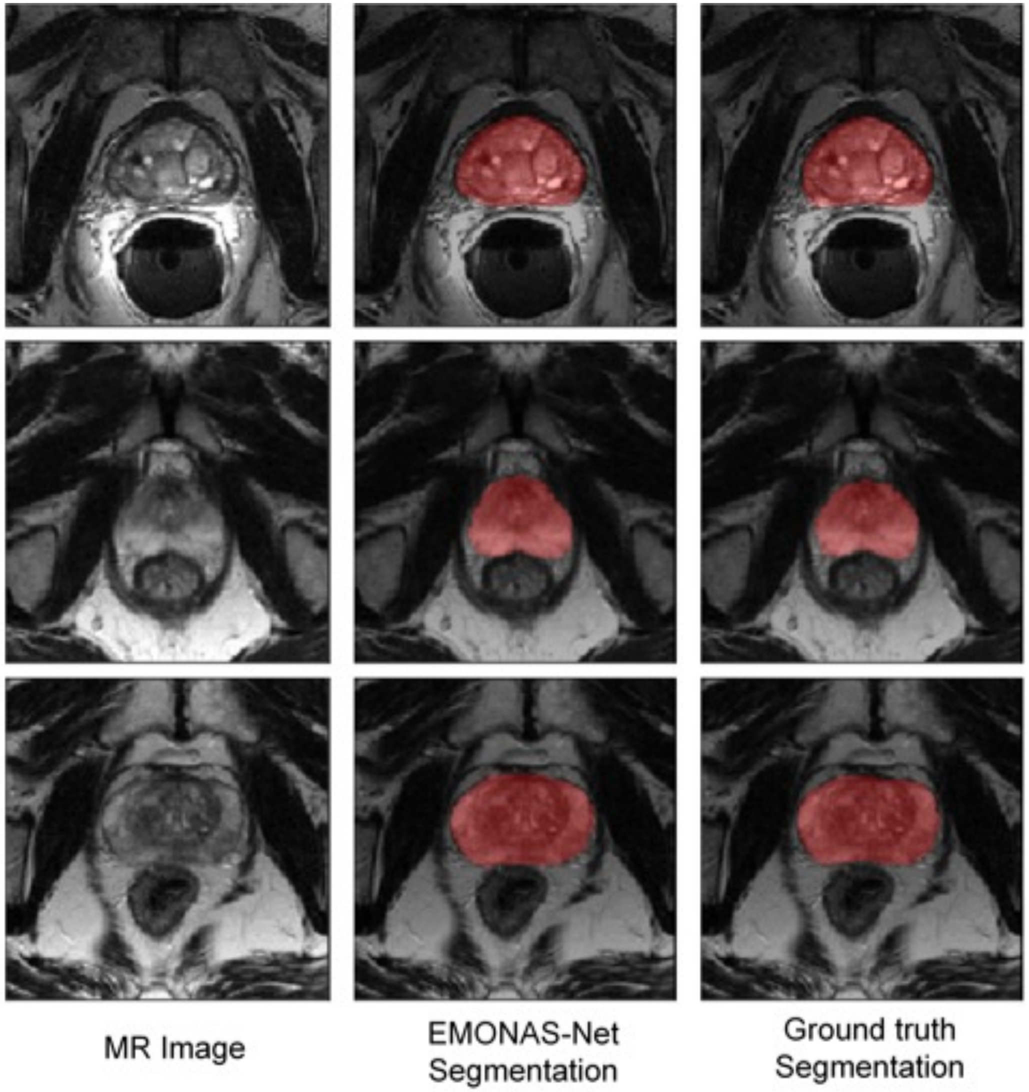
Benchmark Results: Prostate Dataset

Model	Method	DSC[%]			95 HD [mm]			Size of the Network	GPU days
		Whole	Apex	Base	Whole	Apex	Base		
EMONAS-Net	Automatic	91.01	88.16	89.50	4.40	3.69	4.37	5.9×10^6	3.02
Baldeon et al.	Automatic	91.45	88.10	89.54	4.08	3.58	4.43	27.5×10^6	7.68
Isensee et al. (I)	Automatic	91.61	91.61	90.29	4.00	3.79	4.05	365.5×10^6	-
Isensee et al. (II)	Automatic	91.56	91.56	89.59	4.17	3.77	4.42	365.5×10^6	-
Isensee et al. (III)	Automatic	91.93	88.82	91.93	3.95	3.69	4.21	365.5×10^6	-

Table 6. Evaluation metrics for the top competing NAS methods in the Promise12 challenge.

DSC: Dice Similarity Coefficient
95 HD: 95% Hausdorff distance

Results: Prostate Segmentation



Summary

- Presented an efficient NAS framework for medical image segmentation
  - Novel search space for micro and macro structure of architecture
  - SaMEA algorithm to improve exploration of search space and convergence
  - Random forest surrogate to reduce search time
- Experiments demonstrate that the presented framework identifies accurate and small networks while improving search efficiency

Some Additional Resources

- NAS Survey [Elsken 2019]
- Reinforcement Learning NAS [Jaafra 2019]
- Evolutionary Learning NAS [Liu 2021]
- Hardware-aware NAS [Benmeziane 2021]
- NAS for Transformers [Chitty-Venkata 2022]

References

1. E. Real, A. Aggarwal, Y. Huang, and Q. V. Le, “Aging Evolution for Image Classifier Architecture Search,” AAAI, 2019.
2. B. Zoph, V. Vasudevan, J. Shlens and L. Quoc V, "Learning transferable architectures for scalable image recognition.," in *IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 2018
3. L.-C. Chen, M. Collins, Y. Zhu, et al., “Searching for Efficient Multi-scale Architectures for Dense Image Prediction,” In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, 31, 2018.
4. T. Elsken, J. H. Metzen, and F. Hutter, ‘‘Neural architecture search: A survey,’’ J. Mach. Learn. Res., vol. 20, no. 1, pp. 1997–2017, 2019.
5. Y. Jaafra, J. L. Laurent, A. Deruyver, and M. S. Naceur, ‘‘Reinforcement learning for neural architecture search: A review,’’ Image Vis. Comput., vol. 89, pp. 57–66, Sep. 2019.
6. Y. Liu, Y. Sun, B. Xue, M. Zhang, G. G. Yen, and K.C.Tan, ‘‘A survey on evolutionary neural architecture search,’’ IEEE Trans. Neural Netw. Learn. Syst., early access, Aug. 6, 2021.
7. H. Benmeziane, K. El Maghraoui, H. Ouarnoughi, S. Niar, M. Wistuba, and N.Wang,‘‘Acomprehensive survey on hardware-aware neural architecture search,’’ 2021.
8. K. T. Chitty-Venkata, M. Emani, V. Vishwanath, and A. K. Somani, “Neural Architecture Search for Transformers: A Survey,” IEEE Access, 2022.

Thank You!

Questions?