平安科技
PINGAN TECHNOLOGY

CHANG GUNG MEDICAL FOUNDATION
Chang Gung Memorial Hospital, Linkou

CVPR SEATTLE WASHINGTON JUNE 16-18 2020

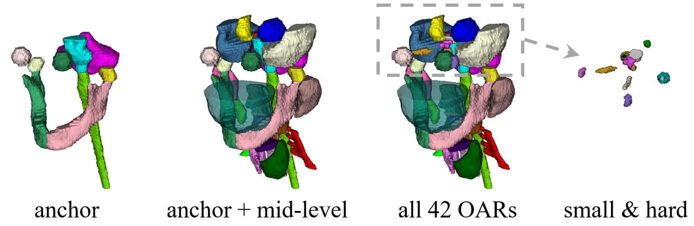
Organ at Risk for Head and Neck Cancer using Stratified Learning and Neural Architecture Search

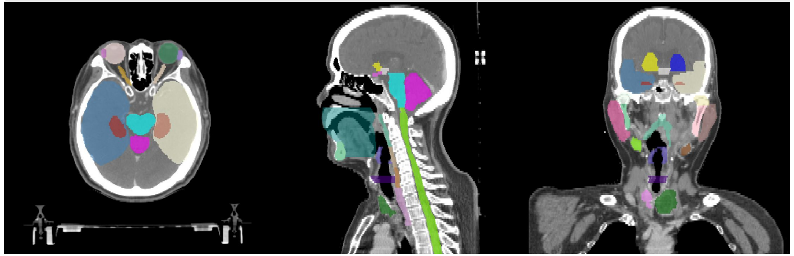
Dazhou Guo[1], Dakai Jin[1], Zhuotun Zhu[3], Tsung-Ying Ho[2], Adam P. Harrison[1], Chun-Hung Chao[4], Jing Xiao[5], Le Lu[1]

1. PAII Inc., Bethesda, MD, USA 2. Chang Gung Memorial Hospital, Linkou, Taiwan, ROC 3. The Johns Hopkins University
4. National Tsing Hua University, Hsinchu City, Taiwan, ROC 5. Ping An Technology, Shenzhen, China

Motivation

- **Organs at risk (OAR) at head & neck (H&N) are critical**
  - In charge of living quality of norm life
  - Responsible for many cancers, e.g. OPX, NPC, HPX
- **Manual delineation is extremely time and labor intensive**
  - H&N OARs are *complex* in anatomical shapes, *dense* in spatial distributions, *large* in size variations, and *low* in RTCT image contrast.
  - *> 2 hours (for 9 OARs)* manual work for each patient
- **Automated SOTA methods are usually hard to optimize**
  - Model optimization becomes increasingly difficult as greater number of OARs need to be segmented.
  - Different OARs likely require different network architectures for optimal performance

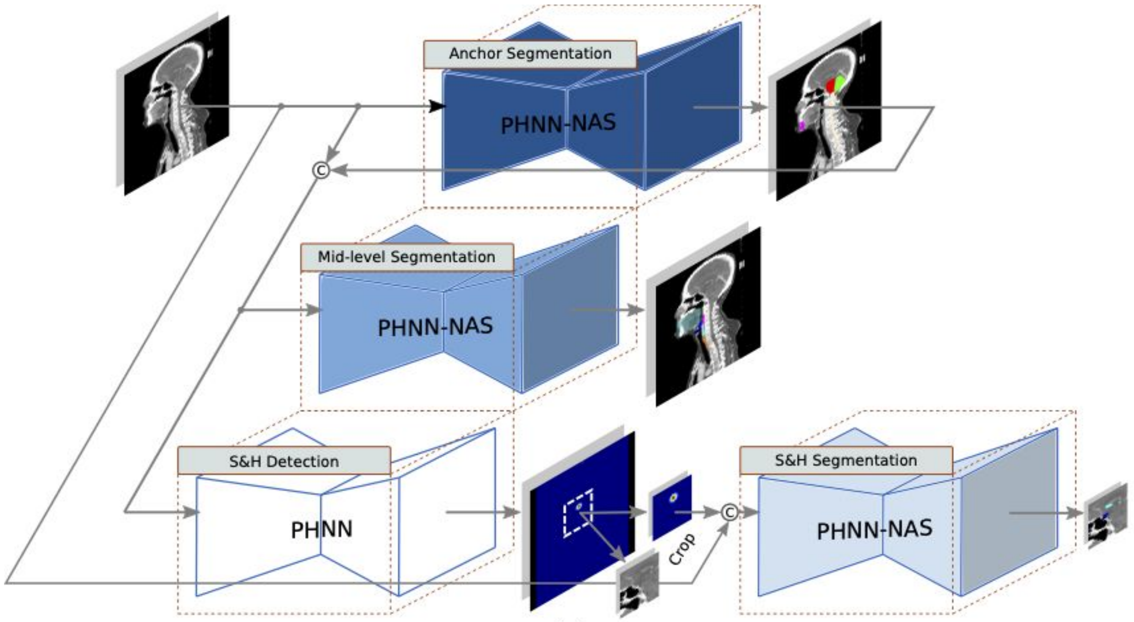




- **MICCAI 2019 challenge[1]**
  - 22 OARs
- **Nature Machine Intelligence paper[2]**
  - 28 OARs
- **Our most comprehensive dataset**
  - **42 OARs**

[1] "Automatic Structure Segmentation for Radiotherapy Planning Challenge 2019" (https://structseg2019.grand-challenge.org)
[2] H. Tang, et al., "Clinically applicable deep learning framework for organs at risk delineation in CT images". Nature Machine Intelligence, 2019

Method - Processing Stratification

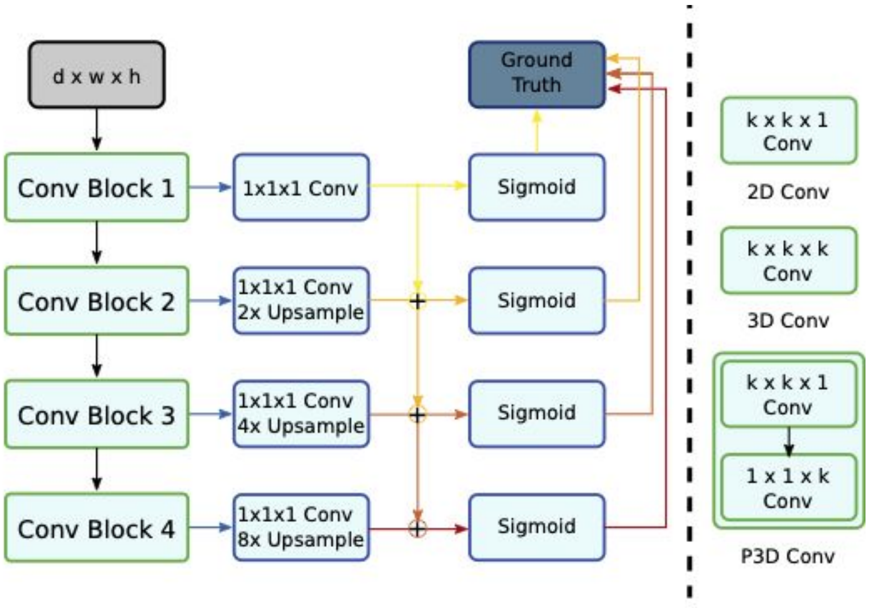


- **We propose a Stratified Organ At Risk Segmentation (SOARS) framework**
  - Stratifying different organs into **anchor**, **mid-level**, and **small & hard (S&H)** categories of OARs
  - **Anchor guided** mid-level and S&H OARs segmentation.
  - A **detection-segmentation** design for S&H OARs segmentation.

- Our idea of stratifying the 42 OARs into three categories comes from the combination of *emulation of oncologists manual OAR contouring knowledge* and *the OAR's size distributions.*

Anchor OARs are reasonably large in object size with clear boundaries, are easy to segmentation
Mid-level OARs are reasonable in object size and are easy to locate but low in intensity contrast
S&H OARs are small in size, and are hard to locate, complex in shape, and poor in intensity contrast

Method - Architectural Stratification

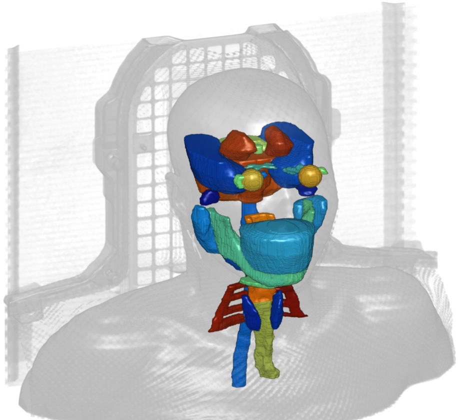


- **We use PHNN[3] as our backbone network**
- **We use NAS to search the network architecture for each convolution block**
  - Search between **2D, 3D & Pseudo 3D (P3D)** conv kernels
  - Search between **different kernel sizes**, e.g. 3x3, 5x5
- **Searched anchor branch block architecture:**
  - 2D-3x3, 2D-5x5, 2D-3x3, 3D-5x5
- **Searched mid-level branch block architecture:**
  - 2D-3x3, P3D-5x5, 2D-3x3, P3D-5x5
- **Searched S&H branch block architecture:**
  - 2D-3x3, 3D-5x5, 2D-3x3, 3D-5x5

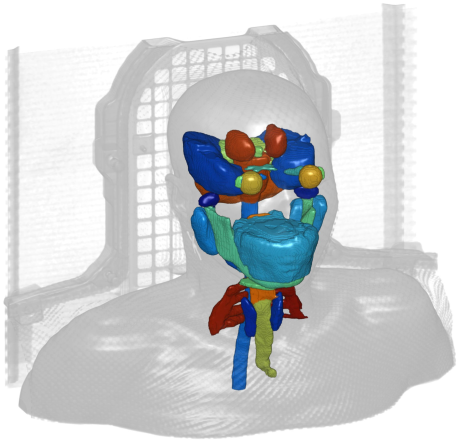
This is an interesting result, as it indicates that 3D kernels may not always be the best choice for segmenting objects with reasonable size, as mixed 2D or P3D kernels dominate both branches. As for the S&H oars, more 3D kernels are used, indicating small objects with low contrast rely more on the 3D spatial information for better segmentation.

[3] Harrison, A. P., Xu, Z., George, K., Lu, L., Summers, R. M., & Mollura, D. J. (2017, September). Progressive and multi-path holistically nested neural networks for pathological lung segmentation from CT images. In International conference on medical image computing and computer-assisted intervention (pp. 621-629). Springer, Cham.

Result - Ablation Study



Ground truth



Our prediction

- **CT Only vs. Baseline Dice score improvement:**
  - S&H OARs: **1.1%**
  - Mid-level OARs: **3.6%**
  - Anchor OARs: **0.1%**

- **SOARS w.o. NAS vs. Baseline Dice score improvement:**
  - S&H OARs : **10.8%**
  - Mid-level OARs: **7.1%**

Anchor OARs

	DSC	HD	ASD
Baseline	84.02	5.98	0.82
CT Only	84.14	5.25	0.79
CT+NAS	85.73	4.77	0.77

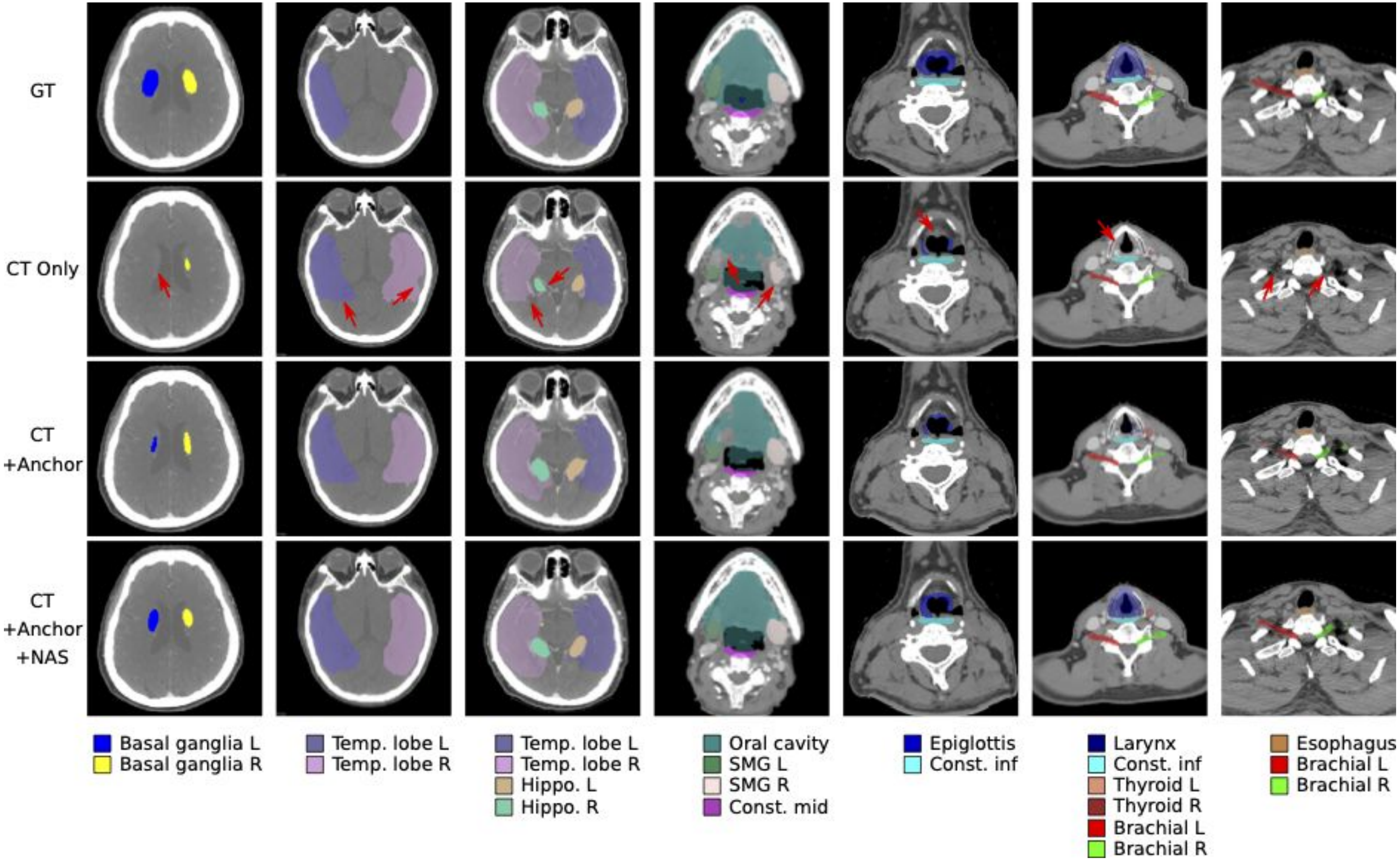
Mid-level OARs

	DSC	HD	ASD
Baseline	63.68	12.97	3.48
CT Only	67.31	12.03	3.97
CT+Anchor	70.73	10.34	1.67
CT+Anchor+NAS	72.55	9.05	1.31

S&H OARs

	DSC	HD	ASD
Baseline	60.97	4.86	0.98
CT Only	62.09	4.19	1.06
CT+Heat map	71.75	2.93	0.52
CT+Heat map +NAS	72.57	2.94	0.49

Result - Ablation Study

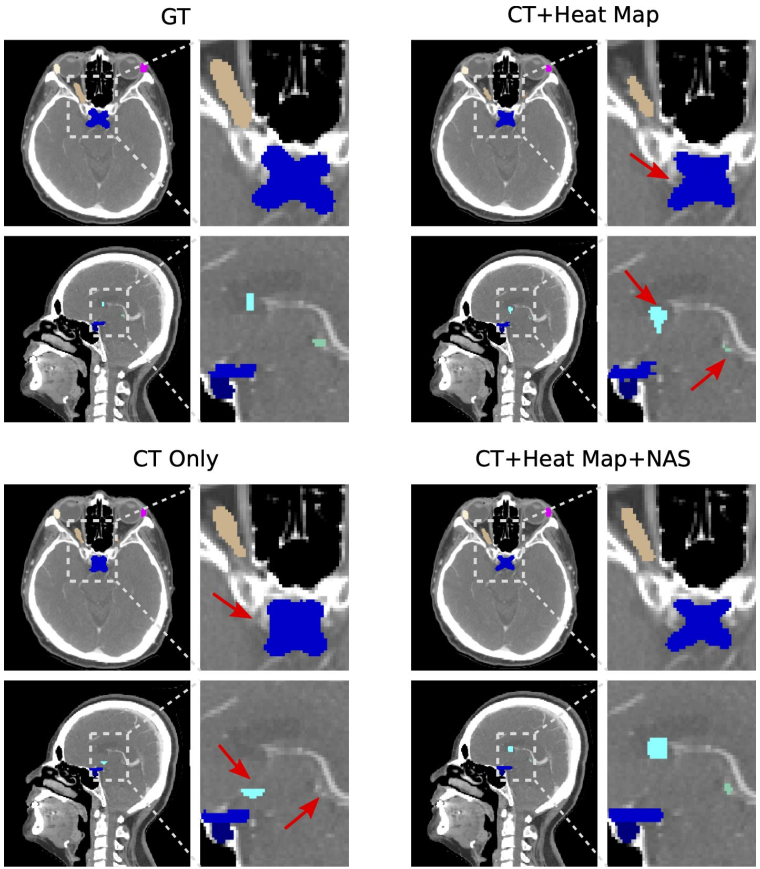


Result - Ablation Study

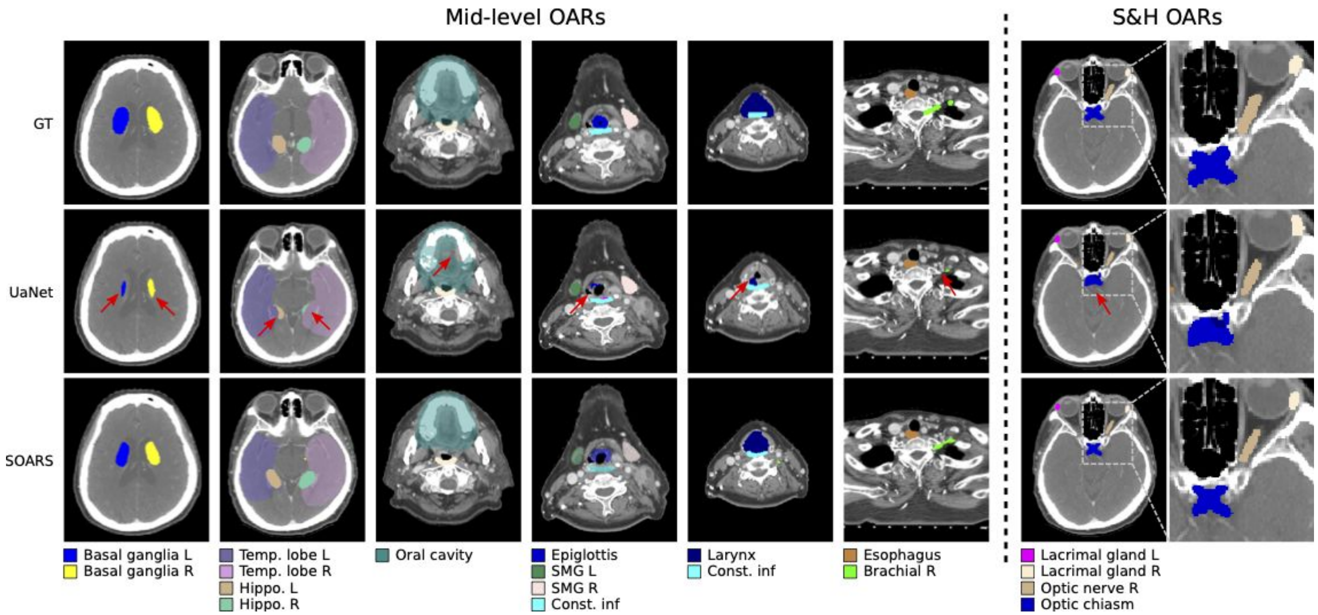
- **CT only vs. CT + anchor OAR S&H detection improvement:**
  - Center-to-center dist: **0.4mm**
  - Hausdorff dist: **12.7mm**

- **SOARS w. NAS vs. Baseline Dice score improvement:**
  - S&H OARs: **11.6%**
  - Mid-level OARs: **8.9%**
  - Anchor OARs: **1.7%**

	Dist (mm)		HD (mm)	
	CT Only	CT+Anchor	CT Only	CT+Anchor
Ear Lt	**3.9±2.5**	3.9±2.6	6.7±3.3	**5.7±2.1**
Ear Rt	1.9±1.4	**1.6±1.0**	4.4±1.8	**3.4±1.3**
Hypothalamus	2.6±1.7	**2.3±1.5**	4.0±2.0	**3.6±1.5**
Lacrimal Gland Lt	5.6±5.7	**4.6±3.1**	28.0±76.8	**14.7±20.7**
Lacrimal Gland Rt	3.3±1.9	**3.0±1.7**	47.4±112.0	**4.7±1.4**
Optic Chiasm	3.9±2.5	**3.4±1.9**	26.6±71.8	**10.6±25.6**
Optic Nerve Lt	**2.5±1.6**	2.6±1.5	4.6±1.8	**4.5±1.2**
Optic Nerve Rt	**3.0±1.2**	3.1±1.6	21.9±61.0	**4.9±1.6**
Pineal Gland	2.5±2.5	**1.8±0.7**	27.7±72.2	**3.9±1.3**
Average	3.3	**2.9**	18.9	**6.2**



Result - Comparison to SOTA



4-fold CV on 42 OAR segmentation results:

- Dice score: **75.1%**
- Hausdorff distance: **6.98mm**
- Average surface distance (ASD): **1.12mm**

Dice score improvement over SOTA -- UaNet

- Overall: **4.7%**
- S&H OARs: **10.1%**
- Mid-level OARs: **3.3%**
- Anchor OARs: **0.7%**

	Anchor OARs			Mid-level OARs			S & H OARs			All OARs		
	DSC	HD	ASD	DSC	HD	ASD	DSC	HD	ASD	DSC	HD	ASD
UNet [4]	82.97	8.90	1.06	63.61	11.06	1.92	59.64	6.38	1.31	66.62	9.26	1.86
P-HNN [12]	84.26	6.12	1.18	65.19	13.15	2.97	59.42	5.23	0.82	67.62	9.39	2.23
UaNet [37]	84.30	8.89	1.72	69.40	11.57	2.06	61.85	5.28	1.53	70.44	9.20	1.83
SOARS	**85.04**	**5.08**	**0.98**	**72.75**	**10.10**	**1.66**	**71.90**	**2.93**	**0.53**	**75.14**	**6.98**	**1.12**

Conclusion

- **Clinical Importance:**
  - Segmenting a comprehensive set of OARs is essential and critical for radiotherapy treatment planning in head and neck cancer. We work on the most clinically complete and desirable set of 42 OARs as compared to previous state-of-the-art work.

- **Methodological contribution:**
  - Our main methodological contribution is the proposed whole framework on stratifying different organs into different categories of OARs which to be dealt respectively with tailored segmentors (achieved by NAS). Our method is a well-calibrated framework of integrating organ stratification, multi-stage segmentation and NAS in a synergy.

- **Result:**
  - We demonstrate that our proposed SOARS can outperform all state-of-the-art baseline networks, including the most recent representative work UaNet, by margins as high as 4.70% in DSC. Thus, our work represents an important step forward toward reliable and automated H&N OAR segmentation.

Thank you!
Please feel free to discuss!