Cross-Modality Domain Adaptation Challenge for Medical Image Segmentation

crossMoDA

MICCAI2021

Mon, Sept 27

Tom Vercauteren

Challenge task and dataset

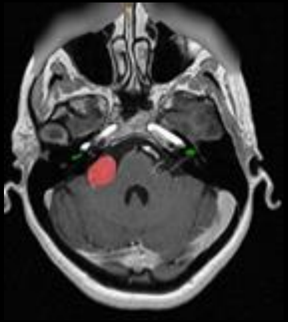
Task: treatment planning of vestibular schwannoma (VS) – segment of two key structures: **tumor and cochlea**

Dataset:

- All images were obtained on a 32-channel Siemens Avanto 1.5T scanner
- Image resolution: 0.5×0.5×1.0mm or 0.5×0.5×1.5mm
- Consecutive patients

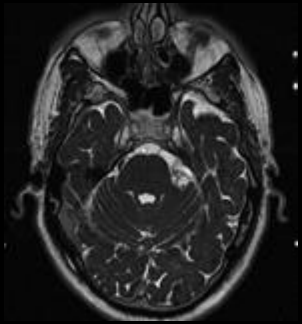
Training

Contrast-enhanced T1



105

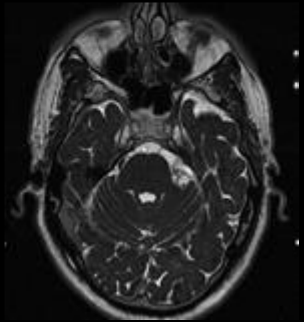
High resolution T2



105

Validation

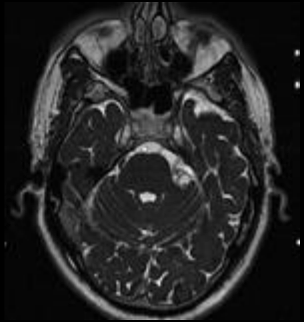
High resolution T2



32

Testing

High resolution T2



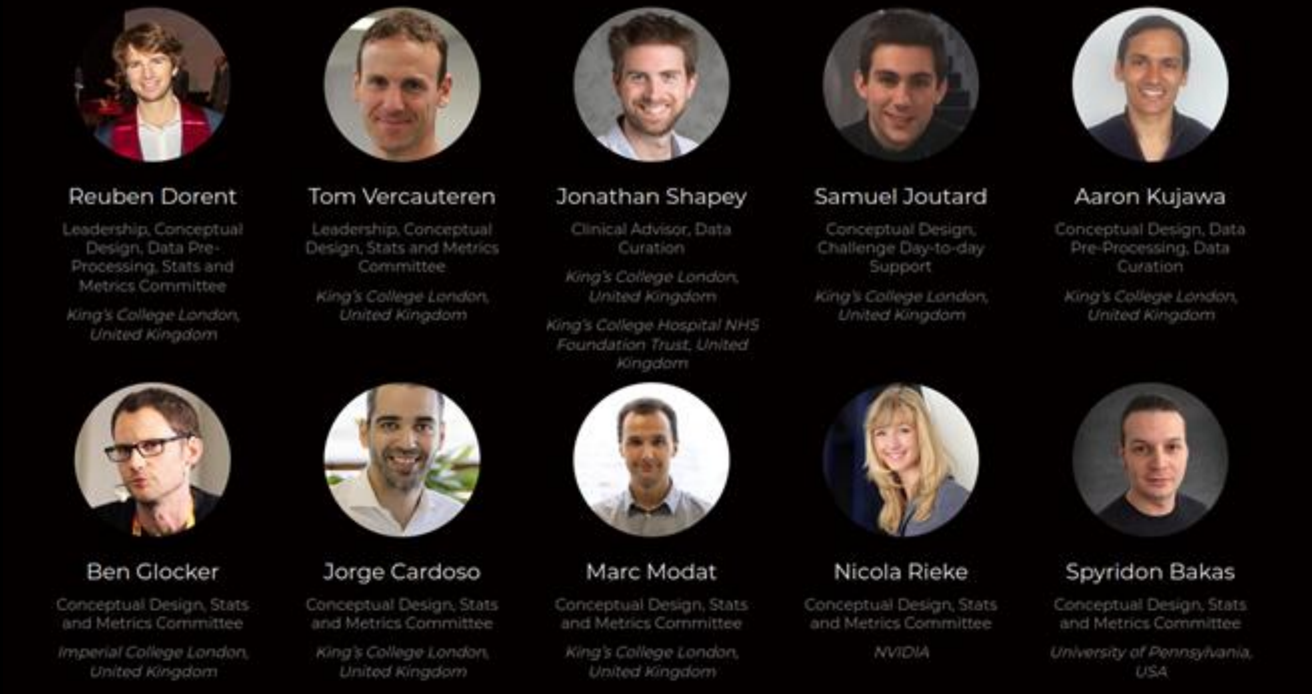
138

Open on TCIA

Program

- 11:10 UTC — Keynote (20 minutes)
  - Dr. Jonathan Shapey
  - Title: Artificial Intelligence Opportunities for Vestibular Schwannoma Management
- 11:30 UTC — Oral Session (40 minutes)
  - Unsupervised Domain Adaptation in Semantic Segmentation Based on Pixel Alignment and Self-Training (PAST)
    - Hexin Dong, Fei Yu, Jie Zhao, Bin Dong and Li Zhang
  - Using Out-of-the-Box Frameworks for Unpaired Image Translation and Image Segmentation for the crossMoDA Challenge
    - Jae Won Choi
  - Self-Training Based Unsupervised Cross-Modality Domain Adaptation for Vestibular Schwannoma and Cochlea Segmentation
    - Hyungseob Shin, Hyeon Gyu Kim (Yonsei University), Taejoon Eo and Dosik Hwang
- 12:10 UTC — Sponsor presentation: NVIDIA (8 minutes)
  - Ahmadi Seyed-Ahmad
- 12:18 UTC — Challenge design and results announcement (20 minutes)
  - Reuben Dorent
- 12:40 UTC — Poster session (40 minutes)
  - The poster session on Gather Town: https://gather.town/app/tcaaFFi2mJYrrDGK/crossmoda-2021

Organizing team & sponsors



NVIDIA

BANA
British Acoustic Neuroma Association

Cross-Modality Domain Adaptation Challenge for Medical Image Segmentation

crossMoDA

MICCAI2021

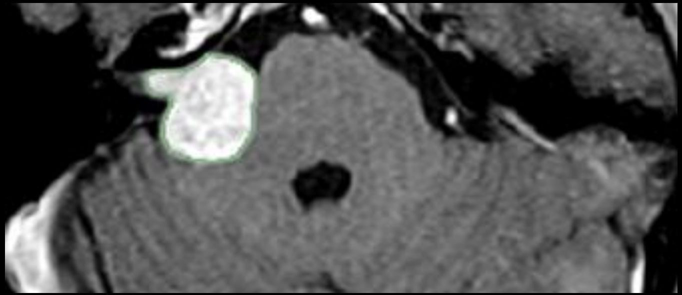
Mon, Sept 27
Reuben Dorent

Supervised learning

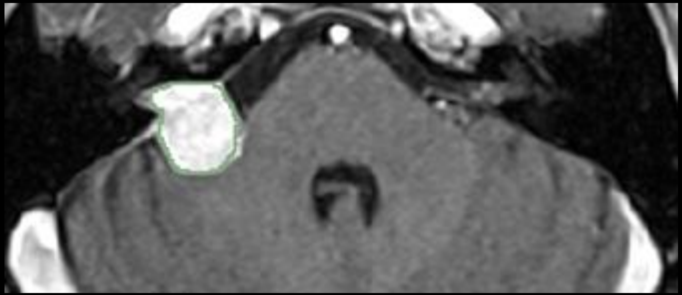
Underlying assumption of **supervised training** on data distributions:

Source (Training) = Target (Test)

Source



Target



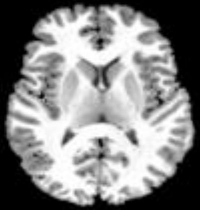
Domain shift in medical applications

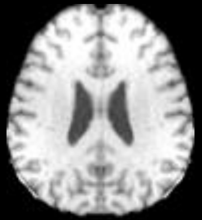
In practice:

Source (Training) ≠ Target (Test)

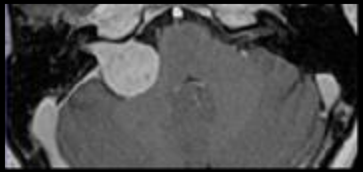
Source | Target

1 **Different acquisition protocols:**
- Scanner characteristic (manufacturer, strength)
- Sequence parameters
- Type of acquisition (axial, coronal, sagittal, isotropic - slice thickness)





2 **Different imaging modalities:**
CT vs MR
Contrast-enhanced T1 vs T2

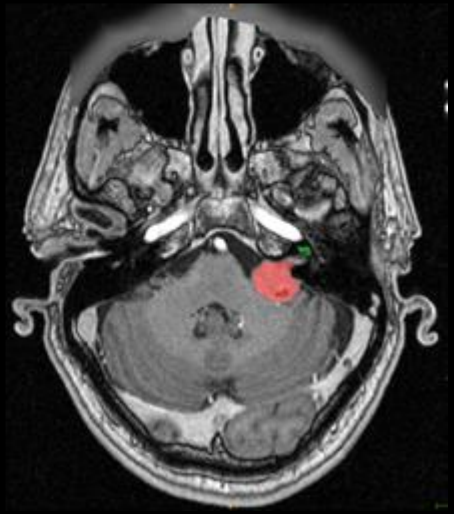


CNNs have been shown to have poor generalization capability

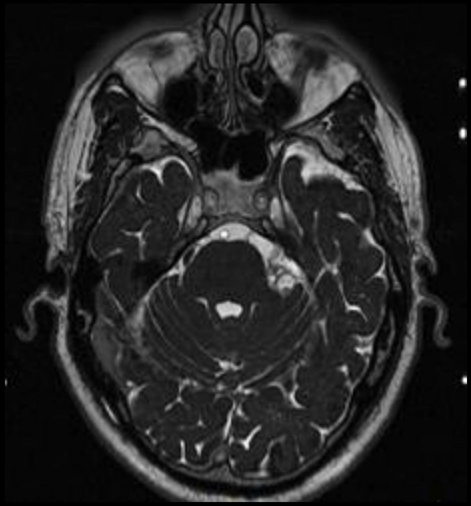
Unsupervised Domain Adaptation (UDA)

Goal: Learning a domain-invariant feature representation of the data without any target **labelled** data.

Source



Target



Various UDA approaches…

Transforming the source data in target-like data:

→ data augmentation

→ generative models (e.g., CycleGAN) [4,6]

Minimizing the discrepancy between the feature distributions:

→ distribution discrepancy loss

→ discriminative adversarial loss [1,2,3,4,6]

Self-training:

→ self-supervision via pretext tasks [5]

Large range of techniques can be used to enforce networks to be modality-invariant.

… tested on different problems

	Public	Large testing set (>20)	Multi-Class Problem	Cross-modality
Traumatic brain injuries [1]		✓		
Liver Segmentation [2]		✓		✓
White Matter Lesions [5]	✓			
Cardiac structure segmentation [3,4,6]	✓		✓	✓

Need for a benchmark on a large, publicly available, multi-class dataset

Challenge task and dataset

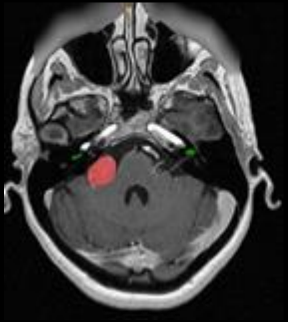
Task: treatment planning of vestibular schwannoma (VS) – segment of two key structures: **tumor and cochlea**

Dataset:

- All images were obtained on a 32-channel Siemens Avanto 1.5T scanner
- Image resolution: 0.5×0.5×1.0mm or 0.5×0.5×1.5mm
- Consecutive patients

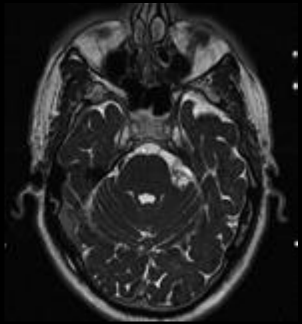
Training

Contrast-enhanced T1



105

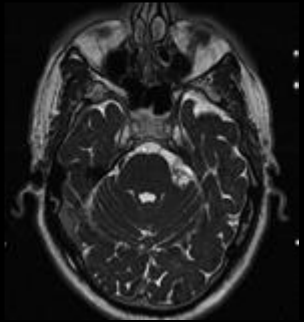
High resolution T2



105

Validation

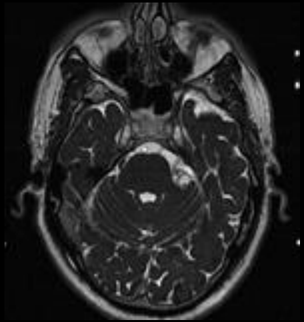
High resolution T2



32

Testing

High resolution T2



138

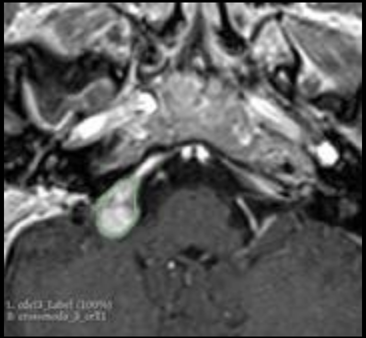
Open on TCIA

A challenging task

Vestibular Schwannoma

- Uniform on ceT1
- Borders may not be clear on hrT2

ceT1

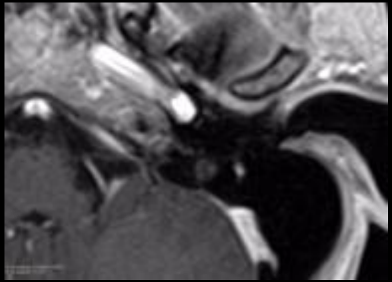


hrT2

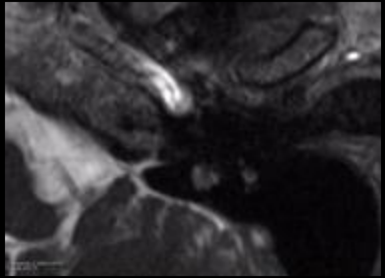
Cochlea

- Two sides
- Very small structure (92 ± 14 mm^3 - 0.002% voxels)
- Unclear borders on ceT1

ceT1



hrT2



Challenge evaluation

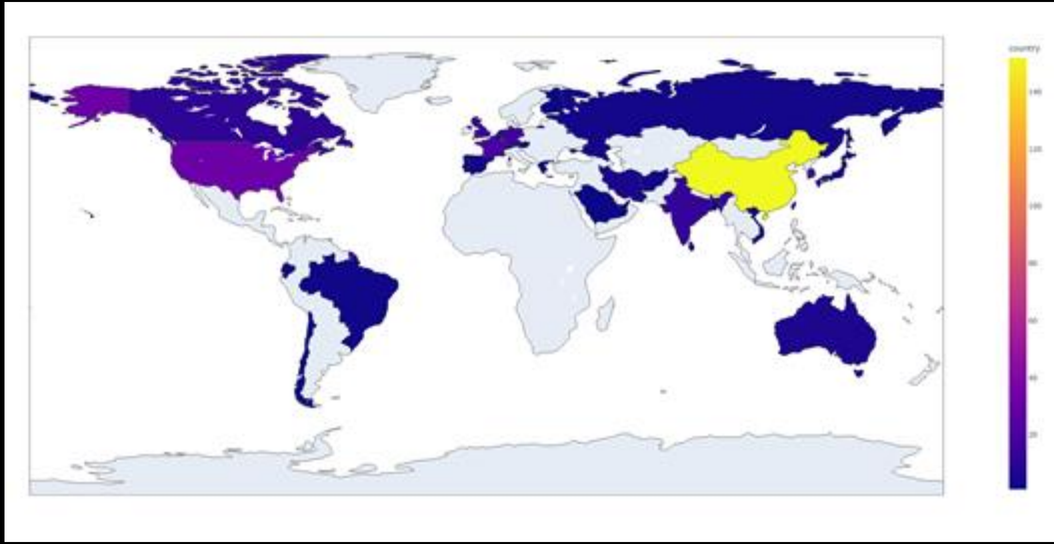
- **Metrics:**
  - Dice Score Coefficient (DSC)
  - Average Symmetric Surface Distance (ASSD)
- **Ranking method:**
  - Based on BraTS challenge methodology
  - Participating teams are ranked for each testing subjects, for each evaluated region (i.e., VS and cochlea), and for each measure (i.e., DSC and ASSD)
  - The final ranking score for each team is then calculated by firstly averaging across all these individual rankings for each patient, and then averaging these cumulative ranks across all patients for each participating team
- **Validation set submission process:**
  - Predictions submitted via *grand-challenge.org*
  - 1 submission allowed per day
- **Testing set submission process:**
  - 1 submission via a Docker container

Participation

Registration:
Number teams: **341**
Number countries: 34

Validation:
Number teams: **55**
Number countries: 16

Testing:
Number teams: **16**
Number countries: 9

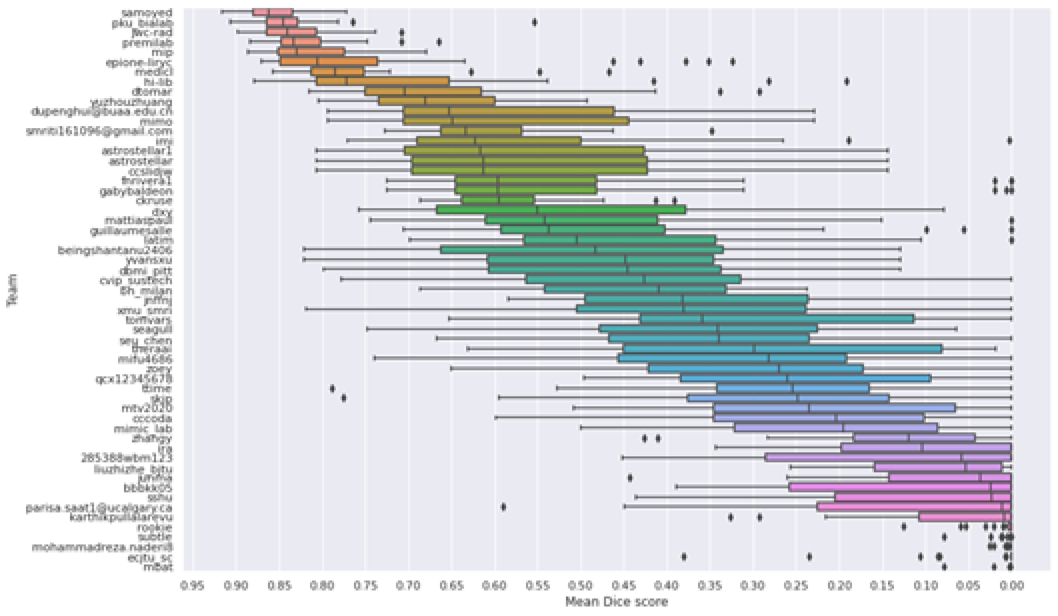


High level observations - validation (1)

Large range of performance

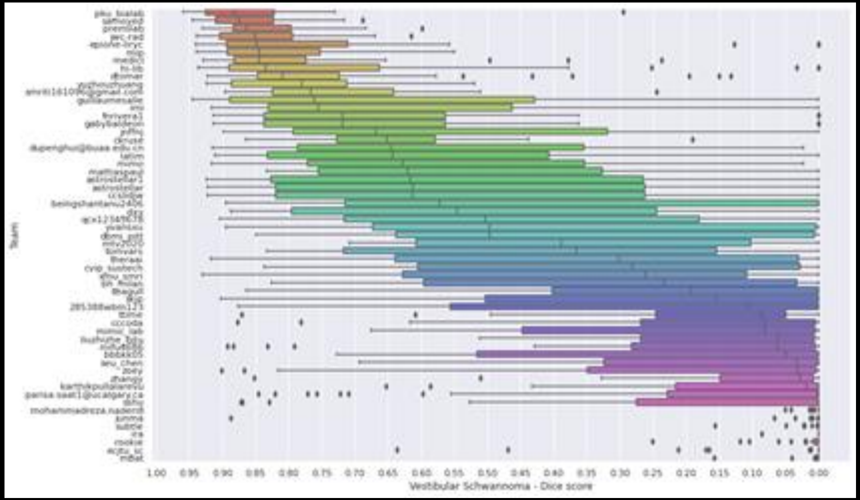
Challenging problem:

- 5 teams (10%) reached a high performance (>80% mean Dice Score).
- 47 teams (85%) obtained a relatively poor performance (<60% mean Dice Score).

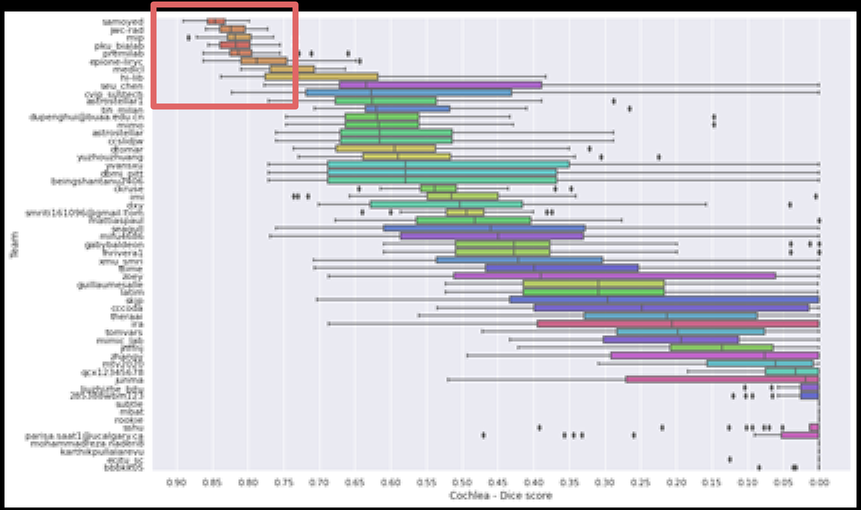


High level observations - validation (2)

Ranking depends on the structure
Larger performance variability for proposed approaches on the VS task
Clear performance gap for the top 5 teams and others on the cochlea task



Vestibular Schwannoma



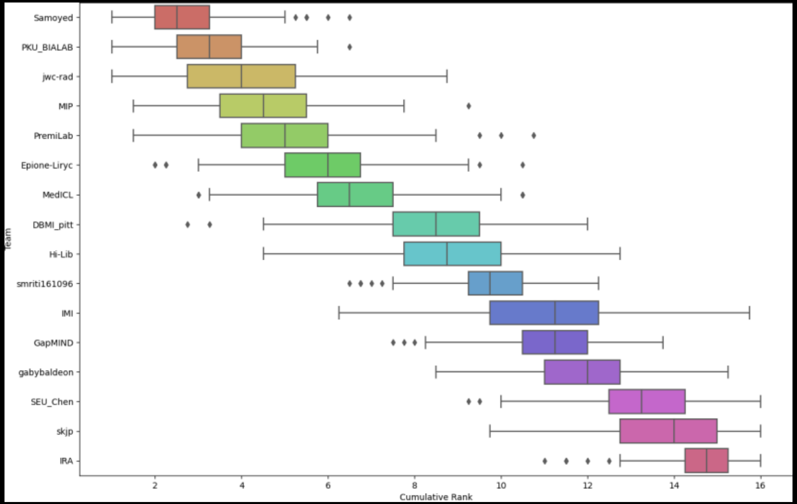
Cochlea

Results

1st - Samoyed - ranking score: **2.7**
Hyungseob Shin, Hyeon Gyu Kim, Taejoon Eo and Dosik Hwang (Yonsei University, Seoul, Korea)
Prize: NVIDIA RTX 3090

2nd - PKU_BIALAB - ranking score: **3.4**
Hexin Dong, Fei Yu, Jie Zhao, Bin Dong and Li Zhang (Peking University, Beijing, China)
Prize: £60

3rd - jwc-rad - ranking score: **4.1**
Jae Won Choi (Seoul National University, Seoul, Korea)
Prize: £40



Same ranking as in the validation set

Variability of the proposed approaches

	Ranking	Feature alignment	MIND features	Content-Style disentanglement	CycleGAN	nnUnet	Self-Supervision
Samoyed	1						
PKU_BIALAB	2						
jwc-rad	3						
MIP	4						
PremiLab	5						
epione-liryc	6						
MedICL	7						
DBMI_pitt	8						
Hi-Lib	9						
smriti161096	10						
IMI	11						
GapMIND	12						
gabybaldeon	13						
skjp	14						
SEU_Chen	15						
IRA	16						

Successful approach: CycleGAN + nnUnet + Self-Supervision

Comparison with full supervision

	Global			Vestibular Schwannoma		Cochlea	
Team	Ranking	DSC (%)	ASSD (mm)	DSC (%)	ASSD (mm)	DSC (%)	ASSD (mm)
Samoyed	1	83.9	0.43	83.0	0.52	84.9	0.34
PKU_BIALAB	2	83.4	0.33	87.0	0.37	79.8	0.30
jwc-rad	3	82.5	0.66	82.9	1.04	82.2	0.29
MIP	4	81.2	0.74	79.9	1.29	82.5	0.18
PremiLab	5	78.5	1.53	77.3	2.78	79.6	0.29
Full supervision (nnUnet)		**88.4**	**0.25**	**89.9**	**0.28**	**86.9**	**0.22**

Problem not solved → A new leaderboard will be created

Publication plans

- **Long paper** (deadline: 30th November 2021) published as part of the BrainLes workshop proceedings distributed by LNCS.
- Submission of a **joint manuscript** summarizing the results of the challenge to a high-impact journal in the field

Next edition: ideas

- **Koos classification** is used in the follow-up and treatment planning of VS:
  - → Adding structures to segment (cerebellum, brainstem)
  - → Unsupervised Domain Adaptation classification task

- The intra-domain data was **homogeneous**:
  - → Using T2 scans from different institutes
    - → T2 to T2 problem (non cross-modality)
    - → Adding T2 scans from other institutes in the testing set

- **Weakly-supervised Domain Adaptation** task (e.g., scribbles, points)

Thank you all!

Poster session: