Weekly Meeting with Dr. Hannah

Hasan Shaikh

Quantitative Imaging Research and
Artificial Intelligence Lab (QIRAIL)

1. Backup of Orthanc Data

- Done with taking backup of Orthanc data till 1st December, 2025.

1. Pedro's Paper - Overview

Paper: Mateus et al. (2023) - *"Image based prognosis in head and neck cancer using convolutional neural networks: a case study in reproducibility and optimization"*

Key Points:

- Uses CT imaging to predict 3 clinical outcomes:
  - Distant Metastasis (DM)
  - Locoregional Recurrence (LRR)
  - Overall Survival (OS)
- Model: Convolutional Neural Network
- Training: Canadian dataset (298 patients)
- Testing: Maastro dataset (137 patients)

Goal: Validate if this model generalizes to our CMC dataset

1. Pedro's Paper - Overview

Paper: Mateus et al. (2023) - *"Image based prognosis in head and neck cancer using convolutional neural networks: a case study in reproducibility and optimization"*

Key Points:

- Uses CT imaging to predict 3 clinical outcomes:
  - Distant Metastasis (DM)
  - Locoregional Recurrence (LRR)
  - Overall Survival (OS)
- Model: Convolutional Neural Network
- Training: Canadian dataset (298 patients)
- Testing: Maastro dataset (137 patients)

Goal: Validate if this model generalizes to our CMC dataset

1. Pedro's Paper - Overview

2. Training (CANADIAN) & Testing (MAASTRO) Tumor Location Distribution

CANADA_Tumor Location Distribution

unknown
(n=9, 3.0%)

Hypopharynx
(n=13, 4.4%)

Nasopharynx
(n=28, 9.4%)

Larynx
(n=45, 15.1%)

Oropharynx
(n=203, 68.1%)

MAASTRO_Tumor Location Distribution

larynx
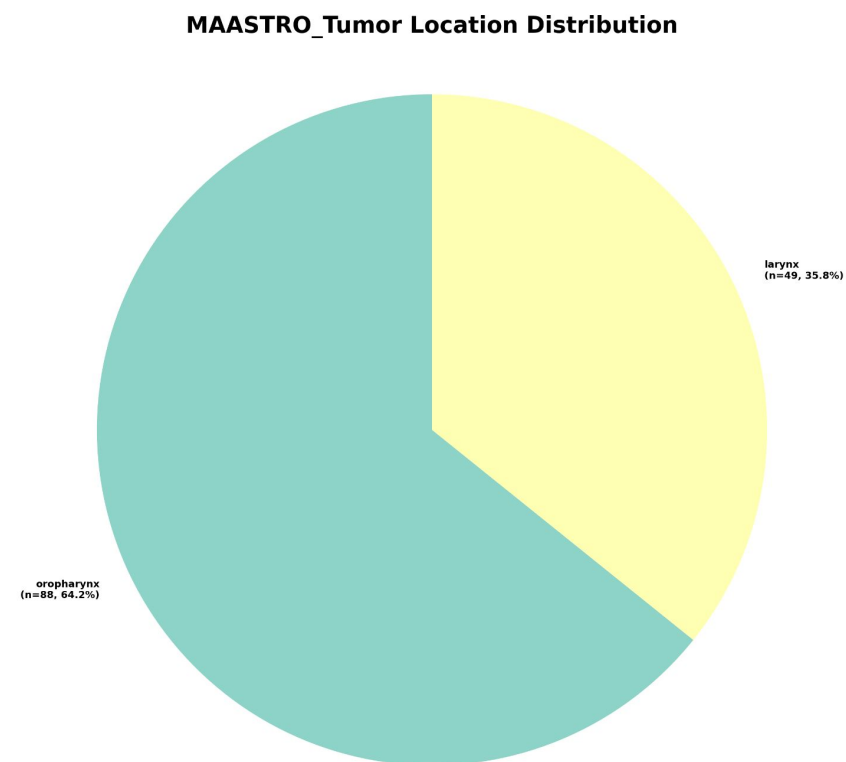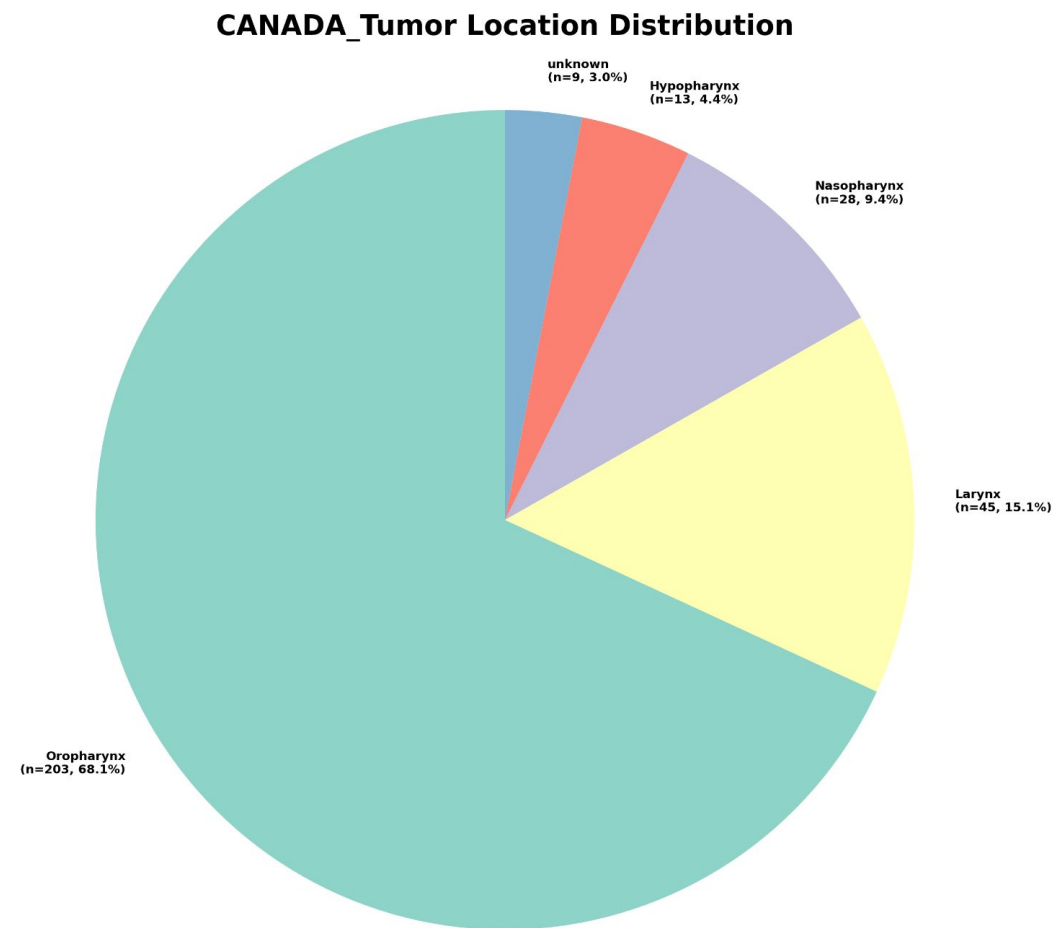(n=49, 35.8%)

oropharynx
(n=88, 64.2%)

Table 1. Comparative AUC performance for three clinical outcomes in head and neck cancer across the original published study (**"Paper Results"**), our reproduced models on the same dataset (**"Our Results"**), and our proposed CNN trained on our in-house dataset (**"Our Results with Our Dataset"**)

Events	Cohort	Paper Results (CNN)	Our Results (CNN)	Our Dataset Results (CNN)
		Cohort split (CI 95%) / 5-fold CV	Cohort split (CI 95%) / 5-fold CV	Cohort split (CI 95%) / 5-fold CV
Distant Metastasis (2y)	Train	0.91 [0.84, 0.96] / 0.87 (0.84–0.92)	**0.85 [0.75, 0.93] / 0.xx**	-
	Val	0.89 [0.81, 0.96] / 0.86 (0.77–0.96)	**0.87 [0.73, 0.98] / 0.xx**	-
	Test	0.89 [0.79, 0.98] / 0.83 (0.76–0.90)	**0.87 [0.67, 0.99] / 0.xx**	-
Locoregional failure (2y)	Train	0.76 [0.64, 0.88] / 0.77 (0.72–0.86)	**0.71 [0.57, 0.84] / 0.xx**	-
	Val	0.77 [0.58, 0.92] / 0.76 (0.72–0.84)	**0.72 [0.53, 0.88] / 0.xx**	-
	Test	0.45 [0.32, 0.57] / 0.53 (0.48–0.59)	**0.49 [0.36, 0.62] / 0.xx**	-
Overall survival (4y)	Train	0.84 [0.75, 0.92] / 0.82 (0.68–0.94)	**0.75 [0.61, 0.86] / 0.xx**	-
	Val	0.80 [0.66, 0.91] / 0.77 (0.62–0.96)	**0.77 [0.62, 0.90] / 0.xx**	-
	Test	0.67 [0.57, 0.77] / 0.63 (0.57–0.72)	**0.67 [0.56, 0.76] / 0.xx**	-

Original study   **Our reproduced study**   **Validate on our Dataset**
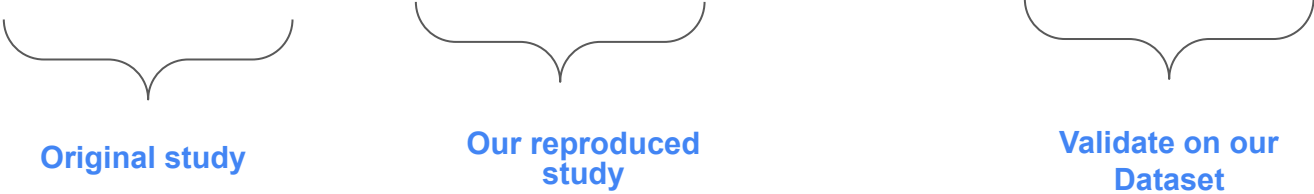
Table 2. Comparative performance (AUCs) including clinical data for predicting multiple clinical outcomes in head and neck cancer across the original study results (*"Paper Results"*), our reproduced CNN and ANN models (*"Our Results"*), and our proposed CNN trained on our in-house dataset (*"Our Results with Our Dataset"*)

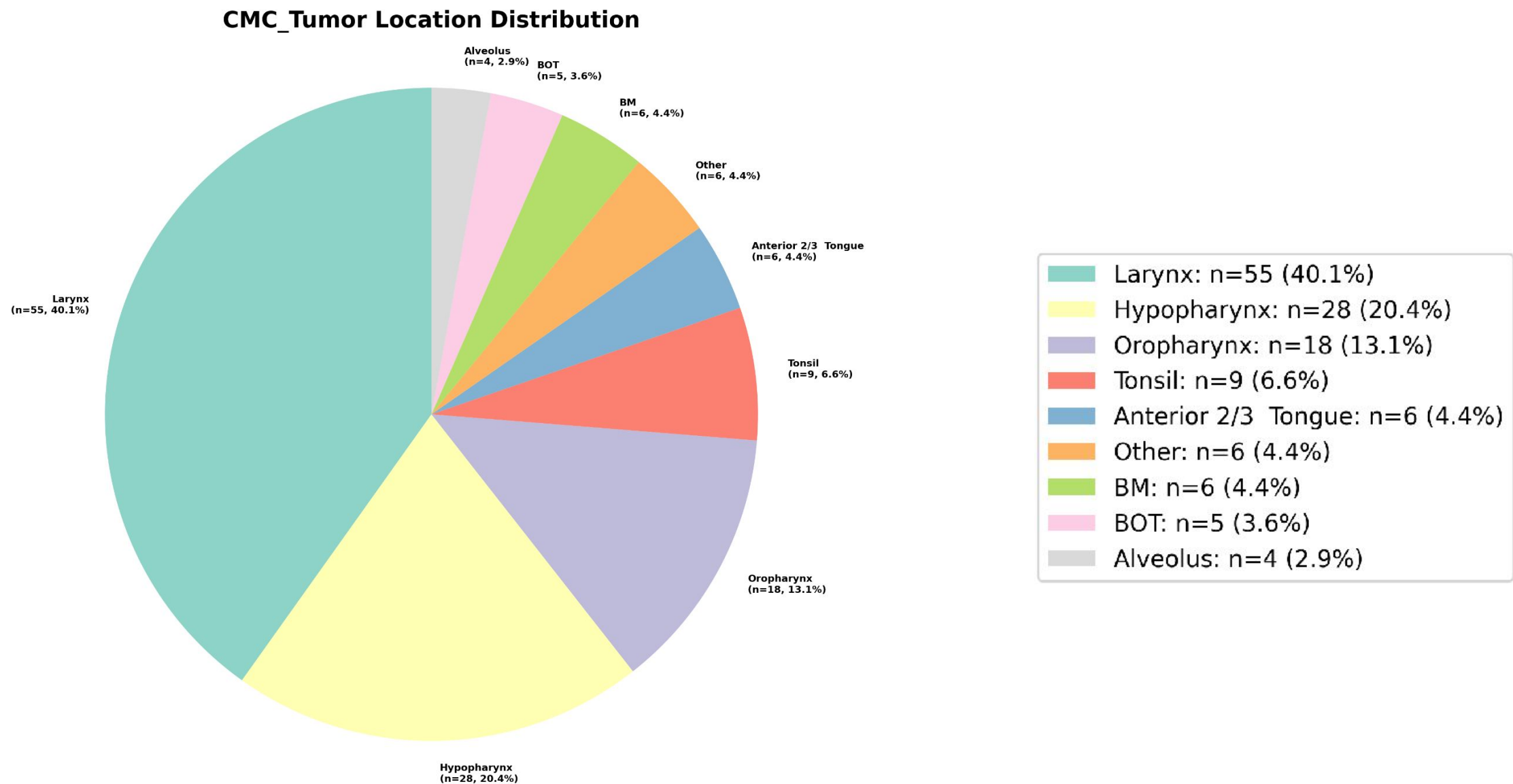| Events	Cohort	Paper Results (CNN)	Our Results (CNN)	Paper Results (ANN)	Our Results (ANN)	Our Dataset Results (CNN)
		Cohort split (CI 95%) / 5-fold CV	Cohort split (CI 95%) / 5-fold CV	Cohort split (CI 95%) / 5-fold CV	Cohort split (CI 95%) / 5-fold CV	Cohort split (CI 95%) / 5-fold CV
Distant Metastasis (2y)	Train	0.91 [0.86, 0.95] / 0.88 (0.81–0.93)	**0.90 [0.84, 0.95] / 0.86**	0.87 [0.78, 0.93] / 0.87 (0.81–0.92)	**0.91 [] / 0.76**	- / -
	Val	0.89 [0.79, 0.98] / 0.88 (0.81–0.93)	**0.86 [0.68, 0.98] / 0.89**	0.79 [0.65, 0.93] / 0.83 (0.79–0.88)	**0.76 [] / 0.80**	- / -
	Test	0.93 [0.86, 0.99] / 0.88 (0.86–0.90)	**0.92 [0.86, 0.98] / 0.88**	0.87 [0.78, 0.95] / 0.86 (0.81–0.89)	**0.87 [] / 0.82**	- / -
Locoregional failure (2y)	Train	0.84 [0.76, 0.93] / 0.77 (0.62–0.87)	**0.75 [0.63, 0.86] / 0.92**	0.71 [0.61, 0.80] / 0.74 (0.70–0.84)	**0.78 [] / 0.78**	- / -
	Val	0.70 [0.54, 0.84] / 0.72 (0.60–0.84)	**0.70 [0.52, 0.85] / 0.87**	0.66 [0.48, 0.82] / 0.71 (0.60, 0.81)	**0.54 [] / 0.69**	- / -
	Test	0.59 [0.47, 0.70] / 0.57 (0.53–0.60)	**0.57 [0.44, 0.68] / 0.54**	0.41 [0.29, 0.54] / 0.53 (0.50, 0.54)	**0.33 [] / 0.40**	- / -
Overall survival (4y)	Train	0.74 [0.64, 0.84] / 0.83 (0.74–0.94)	**0.74 [0.64, 0.84] / 0.88**	0.83 [0.74, 0.90] / 0.83 (0.77–0.85)	**0.78 [] / 0.84**	- / -
	Val	0.74 [0.58, 0.86] / 0.81 (0.73–0.93)	**0.72 [0.58, 0.86] / 0.92**	0.75 [0.62, 0.87] / 0.76 (0.71–0.78)	**0.74 [] / 0.79**	- / -
	Test	0.69 [0.59, 0.79] / 0.68 (0.63–0.71)	**0.69 [0.59, 0.79] / 0.68**	0.63 [0.52, 0.73] / 0.63 (0.61, 0.64)	**0.65 [] / 0.56**	- / -

Original study (CNN) Our reproduced study (CNN) Original study (ANN) Our reproduced study (ANN) Validate on our Dataset (CNN)

CMC_Tumor Location Distribution

Legend:
- Larynx: n=55 (40.1%)
- Hypopharynx: n=28 (20.4%)
- Oropharynx: n=18 (13.1%)
- Tonsil: n=9 (6.6%)
- Anterior 2/3 Tongue: n=6 (4.4%)
- Other: n=6 (4.4%)
- BM: n=6 (4.4%)
- BOT: n=5 (3.6%)
- Alveolus: n=4 (2.9%)

4. Preprocessing Pipeline - Overview

Step-by-step Processing:

1. **DICOM → NIFTI Conversion** (dcmrtstruct2nii)

2. **FSL Reorientation** (fslreorient2std)

3. **Slice Selection** (largest tumor area in HU window)

4. **Tumor-Centered Cropping** (180×180 pixels)

5. **HU Windowing & Normalization** (-50 to 300 HU)

6. **PNG Export** (Final 180×180 images)

5. Next Steps - Code Modifications for Training

Clinical Sheet Requirements:

- Match author's naming convention:
  - Column names
  - locoregional_recurrence_in_days