Weekly Meeting with Dr. Hannah

Hasan Shaikh

Quantitative Imaging Research and
Artificial Intelligence Lab (QIRAIL)

1. 📝 Full Draft of AIHC 2025 Conference (NIT Calicut)

- Completed and shared the **initial full draft** of the paper with **Dr. Hannah** for feedback and revision, prepared for submission to the **AIHC 2025** conference hosted by NIT Calicut.
- Carefully design the **experimental setup** to be reported in the paper.

2. 📄 Review and Preparation of nnU-Net Segmentation Paper

- **Reviewed the nnU-Net manuscript** (written by Nabeel) to assess whether all results, metrics, and claims were correctly and consistently reported.
- Discussion with Nabeel to clarify key aspects of the methodology, performance, and correctness of the reported outcomes.
- Began **converting the manuscript** into the official **AIHC 2025 LaTeX template** for submission readiness.

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)

Table 3: *Performance comparison of machine learning models using radiomics features. ROC-AUC with 95% CI values is reported for ("Train/Validation/Test")*

Classifier	LASSO	SelectKBest	Particle Swarm Optimization	Whale Optimization Algorithm	Grey Wolf Optimizer	Genetic Algorithm	Simulated Annealing
Logistic Regression	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx
Naive Bayes	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx
Linear SVM	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx
RBF SVM	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx
Decision Tree	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx
Random Forest	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx	0.xx | 0.XX | 0.xx