2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

A Multi-Stage Multi-Level Attention-based Feature Learning Framework for Multi-Endpoint Survival Prediction in Head and Neck Cancer

Xiaoying Pan, PhD[1]; Xingzhao Pan[1]; X. Sharon Qi, PhD[2]
[1]Xi'an University of Posts and Telecommunications, [2]University of California Los Angeles

ABSTRACT

Accurate patient stratification is critical for personalized radiation therapy in head-and-neck cancer (HNC). However, high-dimensional feature spaces and limited sample sizes for specific endpoints present major challenges.

We present DeepCox_MTL, a multi-task attention-based deep learning (DL) framework for joint prediction of multiple survival endpoints.

The framework consists of three core modules:

(1) Multi-Modal Feature Extraction (MMFE module): Fuses radiomics, deep imaging embeddings, and clinical data.

(2) Multi-Stage Feature Selection (MSFS module): Uses an Enhanced Genetic Algorithm (EGA) to construct a compact, highly informative feature set.

(3) Multi-Endpoint Survival Prediction (MESP module): Uses a shared encoder with task-specific heads and an uncertainty-based adaptive loss weighting strategy.

Validated on 427 head and neck cancer patients, DeepCox_MTL achieved strong performance: C-indices of 0.7770 (OS), 0.8018 (LFFS), 0.7507 (RFFS), and 0.7179 (DFFS).

DeepCox_MTL captures inter-task relationships to improve prediction accuracy for both primary and sparse endpoints, demonstrating high potential for clinical translation.

CONTACTs

Xiaoying Pan, PhD
Xi'an University of Posts and Telecommunications
Email: panxiaoying@xupt.edu.cn

X. Sharon Qi, PhD
University of California Los Angeles
Email: xqi@mednet.ucla.edu

INTRODUCTION

Clinical Challenge: Individualized treatment in HNC is limited by disease heterogeneity and high feature dimensionality. Furthermore, certain endpoints (e.g., RFFS, DFFS) often suffer from limited sample sizes and training imbalance.

Our Solution: We propose a multi-task learning framework to jointly predict four endpoints: Overall Survival (OS), Local Failure-Free Survival (LFFS), Regional Failure-Free Survival (RFFS), and Distant Failure-Free Survival (DFFS).

Core Contributions:
An MSFS module that combines Spearman correlation, Lasso regression, and an Enhanced Genetic Algorithm (EGA) to select optimal feature subsets.
A multi-task learning architecture with an uncertainty-based adaptive loss weighting strategy to balance optimization across imbalanced clinical endpoints.

METHODS AND MATERIALS

I. Patient Cohort and Data

Source: A public OPC dataset from The Cancer Imaging Archive (TCIA), originating from Princess Margaret Cancer Centre (2005-2010).
Inclusion/Exclusion: From an initial cohort of 606 patients, we included 427 after excluding cases with no delineated Gross Tumor Volume (GTV), unknown HPVp16 status, or from scanner models with very few samples.
Dataset:
Evaluated on a public cohort of **427 Patients** with pre-treatment head-and-neck cancer.

II. Framework Components:

MMFE Module: Fuses low-level semantic radiomic features, high-level deep imaging representations (via an attention-guided global module), and clinical variables.
MSFS Module: Employs an Enhanced Genetic Algorithm (EGA) with a regularized fitness function to identify a compact, non-redundant predictor subset.
MESP Module: Features a shared backbone network to capture common representations across tasks, alongside separate output heads for each survival endpoint.

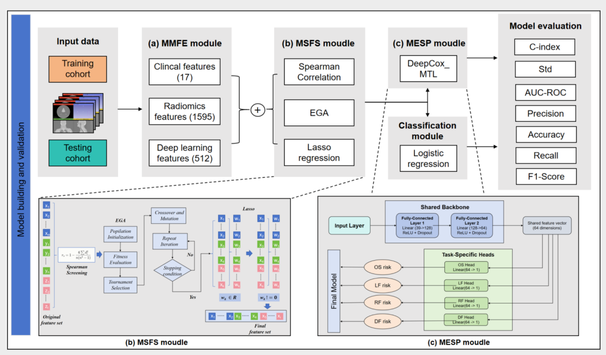

Figure 1. An illustrative diagram showing the overall framework.

RESULTS

Ablation Study: Table 1 demonstrates that the integration of all three proposed modules (MMFE, MSFS, MESP) yields the highest predictive performance (C-index) across all survival endpoints.

Optimization Analysis: Table 2 shows the performance under different training iteration counts. The early stopping strategy successfully balances training time and clinical endpoint prediction accuracy.

DISCUSSION

Multi-task learning effectively mitigates training imbalances in data-scarce endpoints (such as RFFS and DFFS) by transferring knowledge through the shared backbone encoder.

The MSFS module successfully screens high-dimensional radiomic and deep features, reducing overfitting while maintaining crucial prognostic markers.

Table 1. Results of Traditional Methods

A. MMFE module	B. MSFS module	C. MESP module	C-index (OS)	C-index (LFFS)	C-index (RFFS)	C-index (DFFS)
×	×	× (Linear Predict)	0.5593	0.4074	0.5996	0.5121
√	×	× (Linear Predict)	0.6014	0.4484	0.4276	0.6164
√	×	√	0.6044	0.6194	0.7397	0.5775
×	√	× (Linear Predict)	0.7411	0.5197	0.4301	0.3685
×	×	√	0.6456	0.6164	0.6508	0.6149
√	√	× (Linear Predict)	0.7595	0.4611	0.2885	0.4667
×	√	√	0.7417	0.7243	0.6073	0.6574
√	√	√	**0.7702**	**0.8018**	**0.7507**	**0.7179**

Table 2. Performance of MESP module under Different Iteration Counts and with Early Stopping Strategy

Epoch	Task	Mean C- index	Std
100	Overall Survival(OS)	**0.7770**	**0.0312**
100	Local Failure(LFFS)	0.7950	0.0757
100	Regional Failure(RFFS)	√0.7495	0.1257
100	Distant Failure(DFFS)	0.6978	0.1047
300	Overall Survival(OS)	0.7660	0.0289
300	Local Failure(LFFS)	**0.8147**	**0.0841**
300	Regional Failure(RFFS)	**0.7507**	**0.1053**
300	Distant Failure(DFFS)	0.6816	0.1131
Early Stopping	Overall Survival(OS)	0.7702	0.0368
Early Stopping	Local Failure(LFFS)	**0.8018**	**0.0551**
Early Stopping	Regional Failure(RFFS)	0.6793	0.2261
Early Stopping	Distant Failure(DFFS)	**0.7179**	**0.1118**

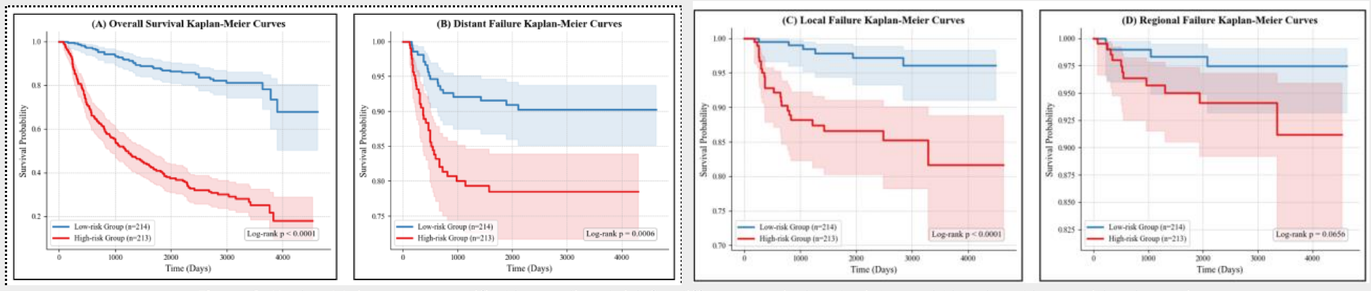

Figure 2. Kaplan-Meier curves stratified by predicted risk for different survival endpoints would be placed here, with sub-captions:
(A) Overall Survival; (B) Distant Failure-Free Survival; (C) Local Failure-Free Survival; (D) Regional Failure-Free Survival

Conclusion

We successfully established a deep learning framework, DeepCox_MTL, which achieves joint and simultaneous prediction of four distinct survival endpoints (OS, LFFS, RFFS, DFFS) in HNC patients, bypassing the need for separate single-task models.

The integration of multi-modal features (radiomics, deep imaging embeddings, and clinical data) combined with the Multi-Stage Feature Selection (MSFS) module successfully reduces high-dimensional noise and prevents overfitting, ensuring highly robust representations.

By employing an uncertainty-based adaptive loss weighting strategy within the MESP module, the framework successfully overcomes training imbalances. This significantly improves the prediction accuracy of sparse endpoints (such as RFFS and DFFS) through shared representation learning.

Validated on a large clinical cohort with high C-indices (such as 0.8018 for LFFS), this framework demonstrates strong clinical utility. It offers a reliable decision-support tool for patient stratification, potentially guiding personalized radiotherapy and treatment planning in oncology.