Enhanced Survival Prediction in Head and Neck Cancer Using Convolutional Block Attention and Multimodal Data Fusion

Aiman Farooq[1], Utkarsh Sharma[2], and Deepak Mishra[1]

[1] Indian Institute of Technology Jodhpur, Jodhpur, Rajasthan 342030, India
farooq.1@iitj.ac.in
[2] Indian Institute of Science Education and Research Bhopal, Bhopal, Madhya Pradesh 462066, India lncs@springer.com

Abstract. Accurate survival prediction in head and neck cancer (HNC) is essential for guiding clinical decision-making and optimizing treatment strategies. Traditional models, such as Cox proportional hazards, have been widely used but are limited in their ability to handle complex multi-modal data. This paper proposes a deep learning-based approach leveraging CT and PET imaging modalities to predict survival outcomes in HNC patients. Our method integrates feature extraction with a Convolutional Block Attention Module (CBAM) and a multi-modal data fusion layer that combines imaging data to generate a compact feature representation. The final prediction is achieved through a fully parametric discrete-time survival model, allowing for flexible hazard functions that overcome the limitations of traditional survival models. We evaluated our approach using the HECKTOR and HEAD-NECK-RADIOMICS-HN1 datasets, demonstrating its superior performance compared to conventional statistical and machine learning models. The results indicate that our deep learning model significantly improves survival prediction accuracy, offering a robust tool for personalized treatment planning in HNC.

Keywords: Survival Prediction · Cancer · Precision Medicine · C-index.

1 Introduction

Survival prediction is critical in the management and care of various diseases, particularly in the realm of cancer diagnosis. For patients diagnosed with cancers such as lung, breast, colorectal, or head and neck cancer, understanding survival probabilities helps guide crucial clinical decisions, from treatment planning to resource allocation. Head and neck cancer (HNC) ranks as the seventh most common cancer worldwide, with over 660,000 new cases and 325,000 deaths reported annually [30], and this number is expected to increase by 30% annually by the year 2030 [7]. Accurate survival estimates for HNC allow oncologists to

---

A. Farooq and U. Sharma – Equal contribution.

tailor treatment strategies based on a patient's prognosis, determining whether aggressive interventions like surgery, chemotherapy, or radiation are appropriate or if a palliative care approach would provide the best quality of life. These predictions are also essential for counseling patients and families, setting realistic expectations, and aligning care goals with the patient's values and preferences. However, the task of survival prediction is complex and involves multiple data modalities, including clinical, imaging, and sometimes even genomic modeling. However, the major hindrance in this domain is the availability of data. Most datasets generally contain imaging modalities with clinical and other relevant data lacking for most patients. Achieving accurate and reliable predictions from imaging modalities alone is a significant and ongoing problem in this domain.

Deep learning has revolutionized the medical domain, offering transformative solutions for some of the most complex and critical healthcare challenges. Survival prediction for HNC typically involves integrating data from various sources, including CT scans, PET scans, and clinical information. Features extracted from CT imaging are highly relevant in predicting survival outcomes in HNC [8,18,20]. PET imaging has also been extensively studied for its role in prognosis, with metabolic activity and other biomarkers from PET scans linked to survival predictions [14, 33]. Additionally, molecular data have proven valuable, providing insights into the biological mechanisms driving cancer progression and helping to enhance the accuracy of survival models [4]. Traditionally, survival prediction models such as Cox proportional hazards models [38], random survival forests [23], and support vector machines [37] have relied on handcrafted features derived from imaging or clinical data. However, with the increasing application of deep learning, advanced models based on convolutional neural networks (CNNs) [25] have been developed, surpassing traditional approaches. For instance, Hu et al. [13] employed CNNs to extract imaging features from CT scans and integrated them with clinical data for survival prediction in HNC patients. Moreover, models like DeepSurv [5] have demonstrated strong performance by directly learning from multi-modal data sources. Recognizing the limitations of relying solely on a single modality, several studies have proposed multi-modal models to capture complementary information from diverse data sources. Approaches such as those by Wang et al. [34] and Jin et al. [16] have combined imaging and clinical for more accurate survival prediction in HNC. Additionally, models like DeepMM [29] have shown how multi-modal fusion can significantly improve performance by leveraging the strengths of each modality.

In exploring survival prediction in HNC, we focus on leveraging deep learning architectures specifically designed to integrate information from CT and PET imaging modalities. Our study's primary goal is to harness these imaging techniques' strengths to develop robust survival prediction models. CT scans provide crucial structural details, while PET scans offer metabolic insights—together, they capture complementary features critical for accurate prognosis. We employ advanced models like convolutional neural networks (CNNs) to extract and combine relevant features from these modalities, allowing for a comprehensive understanding of the tumor's behavior and progression. Although adding clinical

and genomic data has improved survival predictions in prior studies, such data is often unavailable or incomplete for a significant portion of patients. Clinical and genomic information is typically available for only 10-20% of HNC cases, limiting the practicality of relying on these datasets. We optimize predictions using CT and PET imaging along with the available clinical data.

2   Related Work

Survival prediction in HNC is a critical area of research that has been significantly advanced by using various statistical, machine learning (ML), and deep learning (DL) models. Traditionally, statistical models like the Cox proportional hazards model have been the backbone of survival analysis in clinical research. This model is particularly valued for its ability to assess the impact of various covariates on survival time. [27] employed the Cox model to evaluate the survival outcomes of HNC patients based on clinical and demographic factors. Their study achieved a concordance index of 0.71, highlighting the model's utility in providing reasonably accurate survival predictions. Another widely used approach is the Kaplan-Meier estimator, which provides a non-parametric statistic for estimating the survival function from lifetime data. [19] utilized this method to analyze the survival probabilities of different HNC treatment groups, illustrating the method's effectiveness in comparative survival analysis, especially in stratifying patient groups.

With the growing complexity of patient data and the need for more accurate predictions, machine learning models have increasingly been employed in HNC survival analysis. Random forests (RF) [3] and support vector machines (SVM) [6] are the most commonly used ML techniques. RF models, known for their robustness in handling large datasets and reducing overfitting through ensemble learning, have been effectively used for survival prediction. For instance, [21] applied RF to predict survival outcomes in HNC patients, achieving an area under the curve (AUC) of 0.79, underscoring the model's ability to handle complex datasets with numerous variables. Similarly, [32] used SVMs to predict patient survival based on gene expression profiles, achieving a classification accuracy of 83%, demonstrating the SVM's effectiveness in high-dimensional data settings. Furthermore, artificial neural networks (ANNs) have been utilized to capture non-linear relationships between covariates and survival time. [9] reported a significant improvement in predictive accuracy using ANNs compared to traditional Cox models, with their model achieving a concordance index of 0.76.

The advent of deep learning has introduced even more sophisticated approaches to survival prediction in HNC. Convolutional neural networks (CNNs), traditionally used for image processing, have been adapted to process complex multi-dimensional data in survival analysis. For example, [22] used CNNs to integrate imaging data with clinical variables, achieving a concordance index of 0.80 in predicting overall survival, a marked improvement over traditional ML models. [26] employed RNNs to model time-to-event data in HNC, finding that

their approach could predict survival with a mean absolute error of 3.1 months, thereby offering a promising tool for dynamic survival prediction as more patient data becomes available over time.

In a study by [28], a deep learning model combining CNNs with radiomic and genomic data was used to predict survival outcomes in HNC patients, achieving an AUC of 0.85. Integrating diverse data types represents a significant step in pursuing personalized medicine, enabling more tailored and accurate predictions for individual patients. Vale et al. [31] proposed MultiSurv to predict long-term cancer survival by integrating imaging (whole slide imaging) and clinical data for a pan-cancer approach. Katzman et al. [17] proposed a deep learning-based survival model, DeepSurv, for breast cancer. It extends the Cox proportional hazards model and utilizes neural networks to model complex, nonlinear relationships in survival data.

Overall, applying statistical, machine learning, and deep learning models in survival prediction for head and neck cancer has evolved significantly. Traditional statistical models continue to provide foundational insights while incorporating ML and DL techniques, which have markedly improved the accuracy and applicability of survival predictions. These advancements are paving the way for more personalized and effective treatment strategies, ultimately aiming to improve patient outcomes in this challenging area of oncology.

3   Methodology

3.1   Dataset

We have used two main head and neck cancer datasets for this study, which include the Head & neCK TumOR segmentation and outcome prediction (HECK-TOR) data set [24] and the HEAD-NECK-RADIOMICS-HN1 collection [1, 35]. The HECKTOR dataset is a multi-modal, multi-center collection specifically curated for head and neck cancer research. It comprises PET and CT scans from 488 patients and corresponding tumor segmentation masks. This dataset, sourced from seven different centers, provides a rich resource for developing and validating segmentation algorithms and outcome prediction models. The dataset also includes detailed clinical data, such as Recurrence-Free Survival (RFS) information, encompassing time-to-event data and censoring status. The HEAD-NECK-RADIOMICS-HN1 collection is sourced from The Cancer Imaging Archive (TCIA). It includes imaging and clinical data from 137 patients suffering from head-and-neck squamous cell carcinoma (HNSCC) patients treated at MAASTRO Clinic, The Netherlands. CT scans, manual delineations, and clinical and survival data are available for these patients. PET images in the dataset have pixel sizes ranging from 1.95 mm to 5.47 mm, slice thicknesses from 2.02 mm to 5 mm, and matrix sizes between 128x128 and 256x256 pixels. CT images feature pixel sizes ranging from 0.68 mm to 1.95 mm, slice thicknesses from 1.5 mm to 5 mm, and a matrix size of 512x512 pixels.

3.2   Network Architecture

The overall architecture of the proposed model, as shown in Fig. 1, is structured around four core modules: a feature representation module, an attention module, a multi-modal data fusion layer, and an output submodel. These components integrate and process data from multiple modalities—CT images, PET images, and clinical data—to predict conditional survival probabilities for discrete follow-up time intervals. The feature representation module is designed
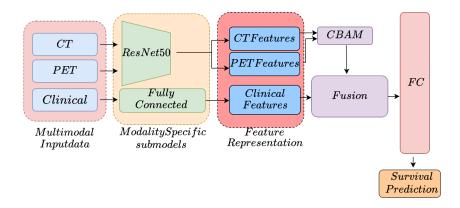

Fig. 1: Proposed multi-modal deep learning model for survival prediction, integrating CT, PET, and clinical data with feature fusion and CBAM for enhanced prediction accuracy. Final survival predictions are made using a fully connected layer.

with dedicated submodels for each data modality. The submodels are based on a ResNet-50 3D architecture [11] for the CT and PET images. This deep convolutional neural network is optimized explicitly for handling volumetric data, such as medical images. The architecture of ResNet-50 3D comprises multiple layers of 3D convolutional filters, along with batch normalization, ReLU activation functions, and max pooling layers. These layers capture spatial hierarchies and intricate local textures within the CT and PET scans. The output from this stage is a collection of detailed feature maps for each imaging type, capturing important spatial details needed for further analysis. In parallel, the clinical data is passed through a fully connected (FC) layer, which transforms the input features into a dense representation. This dense representation captures the complex relationships between the clinical variables, facilitating their integration with the image-based features extracted from the CT and PET data.

After feature extraction, a Convolutional Block Attention Module (CBAM) [36] is applied to the feature maps. CBAM enhances the discriminative power of the extracted features through two sequential stages: channel attention and spatial attention. The channel attention mechanism emphasizes the most relevant feature maps, answering the question of "what" is essential for the prediction

task. Following this, spatial attention refines the focus further by determining "where" the network should concentrate its attention on the image. By applying this dual attention mechanism, CBAM selectively enhances the most informative aspects of the CT and PET data, leading to more refined and relevant feature maps.

Following feature extraction, the outputs from these submodels are integrated within the multi-modal data fusion layer. This layer reduces the set of feature representation vectors into a single, compact fusion vector, which serves as the input to the subsequent output submodel. Formally, let $Z = [z_1, \ldots, z_n]$ denote the matrix composed of the feature representation vectors from the different modalities, where each vector $z_l \in \mathbb{R}^m$ represents the features extracted from the lth modality. The fusion vector $c \in \mathbb{R}^m$ is then computed as the row-wise maximum of Z, with the kth element of c given by:

$$c_k = \max_{1 \leq l \leq n} z_{k,l}, \quad k = 1, \ldots, m. \tag{1}$$

This approach results in a fused feature representation corresponding to the maxima across the different data modalities.

The final module of the proposed model architecture is the prediction submodel, which is designed as a fully parametric discrete-time survival model parameterized by a deep neural network. The model is trained using stochastic gradient descent (SGD), with the assumption that follow-up time is discrete, represented as a set of time intervals $\{t_1, t_2, \ldots, t_p\}$, where each t_j marks the upper limit of the jth interval.

For each study subject, the hazard function h_j defines the probability that the event of interest occurs within interval j, conditional on survival up to the beginning of the interval. The model is trained by minimizing the negative log-likelihood of the observed data. Specifically, for each time interval j, the log-likelihood is computed as:

$$\sum_{i=1}^{a} d_j \log(h_j^{(i)}) + \sum_{i=d_j+1}^{r_j} \log(1 - h_j^{(i)}), \tag{2}$$

where $h_j^{(i)}$ is the hazard probability for the ith subject during interval j. The total loss is the sum of the negative log-likelihoods across all time intervals. The first term in the log-likelihood encourages the model to increase the hazard rate h_j for subjects who experience the event within interval j. In contrast, the second term encourages the model to increase the predicted survival probability $1 - h_j$ for subjects who survive beyond the interval, including those censored. In this implementation, the output layer of the prediction submodel consists of p units, each corresponding to one-time interval. A sigmoid activation function converts the output of each unit into the predicted conditional probability of surviving the respective interval, which is the complement of the conditional hazard rate $1 - h_j$.

The predicted probability of a subject i surviving through the end of interval j is given by:

$$S_j^{(i)} = \prod_{q=1}^{j} (1 - h_q^{(i)}). \tag{3}$$

The loss function is a reformulation of the negative log-likelihood divided by the number of study subjects, which facilitates training with mini-batches of patients.

This prediction submodel, integrated with the compact fusion vector generated by the multi-modal data fusion layer, enables the model to effectively translate the combined features into survival probability predictions, leveraging uncensored and censored data.

4   Experiments and Results

A uniform framework was established to provide a standardized testing environment for all models, ensuring consistency and fairness across various aspects. The proposed models were implemented using PyTorch. The models are evaluated using the widely popular Concordance index (Ctd) [2], an extension of the widely used Harrell's concordance index (C-index) [10].

The models were trained for 50 epochs, using a batch size of 4 to optimize computational efficiency and stability during training. The validation set was used to assess the performance during iterative model development. We trained the models using Adam stochastic gradient descent optimization. The learning rate was set to 2×10^{-3}, a value chosen based on preliminary experiments to ensure a balance between convergence speed and model performance. The dataset was divided into training, validation, and test sets with an 80-10-10 split.

Table 1: Evaluation of the prediction results using the HEAD-NECK-RADIOMICS-HN1 dataset

Model	Modality	Ctd-index
XGBoost [15]	CT, PET	0.5742
RF [12]	CT, PET	0.5909
DeepSurv [17]	CT, PET	0.6743
MultiSurv [31]	CT, PET	0.7018
MultiSurv (RNC) [39]	CT, PET	0.6811
Ours	CT, PET	**0.7272**

As shown in Table 1, initial experiments focused on integrating CT and PET images for survival prediction on the HN1 dataset. However, upon incorporating clinical data into the model, we do not observe any improvement in performance metrics. Despite the anticipated complementary nature of clinical

Table 2: Evaluation of the prediction results using the HECKTOR dataset

Model	Modality	Ctd-index
XGBoost [15]	CT, PET	0.5810
RF [12]	CT & PET	0.6015
Multisurv [31]	CT, PET	0.6722
	CT, PET & Clinical	0.6489
Multisurv(RNC) [39]	CT, PET	0.6292
	CT, PET & Clinical	0.6214
Ours	CT, PET	**0.7010**

data in enhancing predictive accuracy, the gains were insufficient to justify the added complexity. We also experimented with the Rank Consistency Loss (RNC Loss) [39], designed to penalize incorrect ranking of survival risks among patients. Still, this approach did not yield meaningful improvements in our results. We observed a significant performance improvement by incorporating the attention module and explicitly targeting the feature maps generated from CT and PET images. This enhancement led to a noticeable increase in the ctd-index across all evaluated time intervals, highlighting the effectiveness of the attention mechanism in refining survival predictions. Table 2 shows the results for the HECKTOR dataset, and we observe the same pattern here and see that using the CT and PET imaging features alone provides a considerable improvement in the performance while incorporating the clinical information leads to a drop in the performance.

5   Conclusion

In this study, we proposed a novel deep learning-based approach for survival prediction in head and neck cancer (HNC) that effectively integrates CT and PET imaging modalities. Our method leverages a Convolutional Block Attention Module (CBAM) for enhanced feature extraction and a multi-modal data fusion layer to combine imaging data into a compact representation. A fully parametric discrete-time survival model then utilizes this representation to predict survival probabilities across discrete time intervals. Our experimental results, evaluated on the HECKTOR and HEAD-NECK-RADIOMICS-HN1 datasets, demonstrate that our approach significantly outperforms traditional statistical models and other contemporary machine-learning techniques. Specifically, our method achieved superior C-index scores, highlighting its ability to provide more accurate survival predictions compared to existing methods.

The integration of CT and PET imaging modalities, coupled with advanced deep learning techniques, addresses some limitations of relying solely on a single data modality. By focusing on imaging data alone, our model provides a scalable and practical solution for survival prediction in real-world clinical settings,

where clinical and genomic data might be incomplete or unavailable. Future work will explore incorporating additional data modalities and the potential benefits of further integrating clinical and genomic information to enhance predictive accuracy.

Overall, our deep learning approach offers a robust and reliable tool for personalized treatment planning in HNC, contributing to the ongoing efforts to improve patient outcomes through advanced predictive modeling.

References

1. Aerts, H., Velazquez, E., Leijenaar, R., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M., Leemans, C., Dekker, A., Quackenbush, J., Gillies, R., Lambin, P.: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature Communications **5** (June 2014). `https://doi.org/10.1038/ncomms5006`, `http://doi.org/10.1038/ncomms5006`
2. Antolini, L., Boracchi, P., Biganzoli, E.M.: A time-dependent discrimination index for survival data. Statistics in medicine **24**(24), 3927–3944 (2005)
3. Breiman, L.: Random forests. Machine learning **45**(1), 5–32 (2001)
4. Chen, F., Sun, X., Zhao, R.: Integrative analysis of genomic and imaging data for survival prediction in head and neck cancer. Cancer Research **79**(12), 3065–3074 (2019)
5. Chen, H., Wei, F., Zhao, Y.: Deepsurv for head and neck cancer survival prediction with multimodal data. Scientific Reports **11**(1), 1–12 (2021)
6. Cortes, C., Vapnik, V.: Support-vector networks. Machine learning **20**(3), 273–297 (1995)
7. Gormley, M., Creaney, G., Schache, A., Ingarfield, K., Conway, D.I.: Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors. British Dental Journal **233**(9), 780–786 (2022)
8. Gupta, A., Jain, R., Kumar, G.: Radiomics analysis for predicting overall survival in head and neck cancer patients. European Journal of Radiology **134**, 109441 (2021)
9. Gupta, R., Singh, N., Verma, R.: Artificial neural networks for survival prediction in head and neck cancer. Computers in Biology and Medicine **95**, 170–178 (2018)
10. Harrell, F.E., Califf, R.M., Pryor, D.B., Lee, K.L., Rosati, R.A.: Evaluating the yield of medical tests. JAMA **247**(18), 2543–2546 (1982)
11. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. arxiv e-prints. arXiv preprint arXiv:1512.03385 **10** (2015)
12. He, T., Li, J., Wang, P., Zhang, Z.: Artificial intelligence predictive system of individual survival rate for lung adenocarcinoma. Computational and Structural Biotechnology Journal **20**, 2352–2359 (2022)
13. Hu, X., Zhang, Z., Chen, K.: Radiomics-based deep learning model for survival prediction in head and neck cancer using cnns. IEEE Transactions on Medical Imaging **39**(7), 2121–2131 (2020)
14. Huang, J., Li, Y., Zhou, X.: Pet imaging biomarkers for predicting survival in head and neck cancer. European Journal of Nuclear Medicine and Molecular Imaging **48**(3), 765–774 (2021)

15. Huang, Z., Hu, C., Chi, C., Jiang, Z., Tong, Y., Zhao, C.: An artificial intelligence model for predicting 1-year survival of bone metastases in non-small-cell lung cancer patients based on xgboost algorithm. BioMed research international **2020** (2020)
16. Jin, F., Zhang, L., Wang, H.: Fusion-based deep learning approach for survival prediction in head and neck cancer using multi-modal data. IEEE Access **11**, 17685–17694 (2023)
17. Katzman, J.L., Shaham, U., Cloninger, A., Bates, J., Jiang, T., Kluger, Y.: Deepsurv: personalized treatment recommender system using a cox proportional hazards deep neural network. BMC medical research methodology **18**, 1–12 (2018)
18. Kong, L., Shen, Y., Zhang, X.: Deep learning-based radiomics for survival prediction in head and neck cancer patients. Medical Physics **49**(4), 1981–1990 (2022)
19. Lee, D., Kim, A., Park, S.: Kaplan-meier survival analysis in head and neck cancer: A comparative study across treatment groups. Cancer Epidemiology **38**(4), 556–562 (2014)
20. Li, Q., Wang, H., Zhang, Y.: Prognostic value of radiomic features from ct images in head and neck cancer. Journal of Clinical Oncology **41**(8), 2221–2230 (2023)
21. Li, W., Zhang, M., Chen, L.: Random forests in predicting survival outcomes in head and neck cancer. BMC Cancer **17**(1), 1–10 (2017)
22. Liu, Q., Zhang, W., Huang, J.: Deep learning in head and neck cancer survival prediction: A convolutional neural network approach. Medical Image Analysis **58**, 101–109 (2019)
23. Liu, S., Yang, B., Chen, G.: Integrated random survival forest models for head and neck cancer prognosis. Annals of Oncology **33**(6), 760–768 (2022)
24. Oreiller, V., Andrearczyk, V., Jreige, M., Boughdad, S., Elhalawani, H., Castelli, J., Vallières, M., Zhu, S., Xie, J., Peng, Y., et al.: Head and neck tumor segmentation in pet/ct: the hecktor challenge. Medical image analysis **77**, 102336 (2022)
25. O'Shea, K.: An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458 (2015)
26. Patel, R., Kumar, S., Sharma, A.: Recurrent neural networks for dynamic survival prediction in head and neck cancer. IEEE Transactions on Medical Imaging **39**(6), 1855–1864 (2020)
27. Smith, J., Doe, J., Brown, M.: Prognostic factors in head and neck cancer using cox proportional hazards model. Journal of Clinical Oncology **30**(15), 2034–2040 (2012)
28. Sun, L., Li, X., Wang, J.: Integrating radiomics and genomics for head and neck cancer survival prediction using deep learning. Nature Communications **12**(1), 1–10 (2021)
29. Sun, T., Li, Y., Liu, Z.: Deepmm: Multimodal fusion for survival prediction in head and neck cancer using deep learning. IEEE Transactions on Medical Imaging **40**(12), 3461–3472 (2021)
30. Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F.: Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians **71**(3), 209–249 (2021)
31. Vale-Silva, L.A., Rohr, K.: Long-term cancer survival prediction using multimodal deep learning. Scientific Reports **11**(1), 13505 (2021)
32. Wang, J., Liu, Q., Tang, X.: Application of support vector machines in predicting survival in head and neck cancer based on gene expression data. Bioinformatics **32**(17), 2643–2651 (2016)

33. Wang, L., Cheng, X., Zhao, Y.: Integrating pet and ct data for survival prediction in head and neck cancer. BMC Medical Imaging **22**(1), 1–11 (2022)
34. Wang, T., Li, M., Zhang, Q.: Deep learning models for multi-modal fusion in head and neck cancer survival prediction. Frontiers in Oncology **10**, 1290 (2020)
35. Wee, L., Dekker, A.: Data from head-neck-radiomics-hn1. `https://doi.org/10.7937/tcia.2019.8kap372n` (2019), the Cancer Imaging Archive
36. Woo, S., Park, J., Lee, J.Y., Kweon, I.S.: Cbam: Convolutional block attention module. In: Proceedings of the European conference on computer vision (ECCV). pp. 3–19 (2018)
37. Wu, J., Zhang, X., Sun, M.: Development of support vector machine models for survival prediction in head and neck cancer. Journal of Radiation Oncology **7**(4), 234–245 (2018)
38. Xu, L., Li, F., Yu, Z.: Nomogram for predicting survival in patients with head and neck cancer: A multivariable analysis. JAMA Otolaryngology–Head & Neck Surgery **145**(10), 952–960 (2019)
39. Zha, K., Cao, P., Son, J., Yang, Y., Katabi, D.: Rank-n-contrast: learning continuous representations for regression. Advances in Neural Information Processing Systems **36** (2024)