



Article

A Radiotherapy Dose Map-Guided Deep Learning Method for Predicting Pathological Complete Response in Esophageal Cancer Patients after Neoadjuvant Chemoradiotherapy Followed by Surgery

Wing-Keen Yap ^{1,*}, Ing-Tsung Hsiao ², Wing-Lake Yap ³, Tsung-You Tsai ⁴, Yi-An Lu ⁴, Chan-Keng Yang ⁵, Meng-Ting Peng ⁵, En-Lin Su ⁶ and Shih-Chun Cheng ^{7,*}

- ¹ Department of Radiation Oncology, Proton and Radiation Therapy Center, Chang Gung Memorial Hospital—Linkou Medical Center, 5 Fu-Shin Street, Kwei-Shan, Taoyuan 333, Taiwan
 - ² Department of Medical Imaging and Radiological Sciences, Healthy Aging Research Center, Chang Gung University, Taoyuan 333, Taiwan
 - ³ Department of Post Baccalaureate Medicine, Kaohsiung Medical University, Kaohsiung 807, Taiwan
 - ⁴ Department of Otolaryngology—Head and Neck Surgery, Chang Gung Memorial Hospital—Linkou Medical Center, 5 Fu-Shin Street, Kwei-Shan, Taoyuan 333, Taiwan
 - ⁵ Division of Hematology and Oncology, Department of Internal Medicine, Chang Gung Memorial Hospital at Linkou Branch, Chang Gung University College of Medicine, Taoyuan 333, Taiwan
 - ⁶ Department of School of Medicine, National Yang Ming Chiao Tung University, Taipei 112, Taiwan
 - ⁷ Taiwan AI Labs, Taipei 103, Taiwan
- * Correspondence: carsonyap2010@gmail.com (W.-K.Y.); andys0975@gmail.com (S.-C.C.)



Citation: Yap, W.-K.; Hsiao, I.-T.; Yap, W.-L.; Tsai, T.-Y.; Lu, Y.-A.; Yang, C.-K.; Peng, M.-T.; Su, E.-L.; Cheng, S.-C. A Radiotherapy Dose Map-Guided Deep Learning Method for Predicting Pathological Complete Response in Esophageal Cancer Patients after Neoadjuvant Chemoradiotherapy Followed by Surgery. *Biomedicines* **2023**, *11*, 3072. <https://doi.org/10.3390/biomedicines11113072>

Academic Editor: Ryszard Smolarczyk

Received: 13 September 2023
Revised: 28 October 2023
Accepted: 3 November 2023
Published: 16 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: The authors conducted this research to improve the prediction of how patients with esophageal cancer respond to a treatment called neoadjuvant chemoradiotherapy, which can enhance survival rates. However, doctors struggle to accurately predict how well a patient will respond to this treatment using existing imaging methods. To address this, the researchers developed a computer-based method called DCRNet, that not only analyzes medical images but also considers the distribution of radiation therapy doses on the radiotherapy treatment plans to make more accurate predictions. They tested this method on 80 patients with esophageal cancer and found that the HRNetV2p model with DCR performed the best, significantly improving prediction accuracy compared to other models. This breakthrough has the potential to help doctors better anticipate patient responses to treatment, which could lead to more personalized and effective care, and improving the treatment planning of radiotherapy.

Abstract: Esophageal cancer is a deadly disease, and neoadjuvant chemoradiotherapy can improve patient survival, particularly for patients achieving a pathological complete response (ypCR). However, existing imaging methods struggle to accurately predict ypCR. This study explores computer-aided detection methods, considering both imaging data and radiotherapy dose variations to enhance prediction accuracy. It involved patients with node-positive esophageal squamous cell carcinoma undergoing neoadjuvant chemoradiotherapy and surgery, with data collected from 2014 to 2017, randomly split into five subsets for 5-fold cross-validation. The algorithm DCRNet, an advanced version of OCRNet, integrates RT dose distribution into dose contextual representations (DCR), combining dose and pixel representation with ten soft regions. Among the 80 enrolled patients (mean age 55.68 years, primarily male, with stage III disease and middle-part lesions), the ypCR rate was 28.75%, showing no significant demographic or disease differences between the ypCR and non-ypCR groups. Among the three summarization methods, the maximum value across the CTV method produced the best results with an AUC of 0.928. The HRNetV2p model with DCR performed the best among the four backbone models tested, with an AUC of 0.928 (95% CI, 0.884–0.972) based on 5-fold cross-validation, showing significant improvement compared to other models. This underscores DCR-equipped models' superior AUC outcomes. The study highlights the potential of dose-guided deep learning in ypCR prediction, necessitating larger, multicenter studies to validate the results.

Keywords: deep learning; pathological complete response; esophageal cancer; neoadjuvant chemoradiotherapy; artificial intelligence; radiotherapy treatment planning; radiomics; dosiomics

1. Introduction

Esophageal cancer presents a grim prognosis with a 5-year survival rate below 25% [1]. Particularly prevalent in Asia, esophageal squamous cell carcinoma (ESCC) predominates over other forms, defying the Western trend where adenocarcinoma predominates [2]. Despite surgical advancements, disease recurrence remains a pressing concern, prompting exploration into neoadjuvant chemoradiotherapy (nCRT) as a means to bolster long-term survival in locally advanced ESCC patients [3,4]. Recent meta-analysis results indicate that at least 24–32% of patients attain a pathological complete response (ypCR) following nCRT, setting the stage for a “watch-and-wait” strategy as an alternative to surgery, aimed at reducing adverse events [5]. Regrettably, present imaging techniques, including computed tomography, positron emission tomography, endoscopic ultrasound, and magnetic resonance imaging, suffer from inadequate sensitivity (0.35, 0.62, 0.01, and 0.80, respectively) in detecting ypCR [5]. Consequently, there is a growing interest in computer-aided detection methods, like radiomics and machine learning, to facilitate ypCR prediction, aiding both patients and physicians in treatment planning.

While prior research has delved into various machine learning models to predict treatment responses in esophageal cancer centering on imaging features [6–12], these studies have chiefly focused on imaging data, overlooking the potential ramifications of variations in radiotherapy dose planning practices on treatment responses. Notably, previous published studies have demonstrated that a good response after chemoradiotherapy, particularly a ypCR is prognosticated for excellent overall survival and recurrence-free survival in ESCC patients [13–15]. Thus, the capability to predict a ypCR after nCRT in individual patients remains a crucial yet unmet need.

Our study introduces an innovative approach to predicting esophageal cancer patient responses to nCRT. We employ a deep learning methodology that transcends conventional reliance on imaging data alone, incorporating the variances in radiotherapy dose mapping. Specifically, we propose a radiotherapy dose map-guided deep learning model that predicts a ypCR after nCRT in patients with ESCC.

2. Materials and Methods

2.1. Patients

This study was conducted with the approval of the institutional review boards at Chang Gung Memorial Hospital. To safeguard patient privacy, the data used were de-identified, obviating the need for informed consent. The study adhered to the Standards for Reporting of Diagnostic Accuracy (STARD) guidelines.

The data employed in this research were derived from patients diagnosed with node-positive ESCC who underwent nCRT followed by surgery from January 2014 to December 2017. To ensure robustness, the patient cohort was randomly divided into five subsets for 5-fold cross-validation, with each subset maintaining an equivalent ratio of patients who achieved a ypCR and those who did not, as illustrated in Figure 1.

All patients in the study received platinum-based chemotherapy concurrently with radiotherapy, with the total radiation dose ranging from 41.4 to 50.4 Gy. The gross tumor volume (GTV) encompassed both the primary tumor and the involved regional nodes, as identified through endoscopy, CT, or PET/CT scans. The clinical target volume (CTV) extended 3–4 cm along the esophagogastric wall, both proximally and distally to the GTV, with an added 1 cm radial expansion to accommodate potential microscopic periesophageal spread.

Both planning CT scans and radiation dose maps were obtained for the development of the model. Following nCRT, enrolled patients underwent radical esophagectomy within

a window of 4 to 8 weeks. Specialized pathologists, with expertise in ESCC, assessed the pathological response. The criteria for a ypCR were met when no viable cancer cells were observed in the primary tumor and regional lymph nodes.

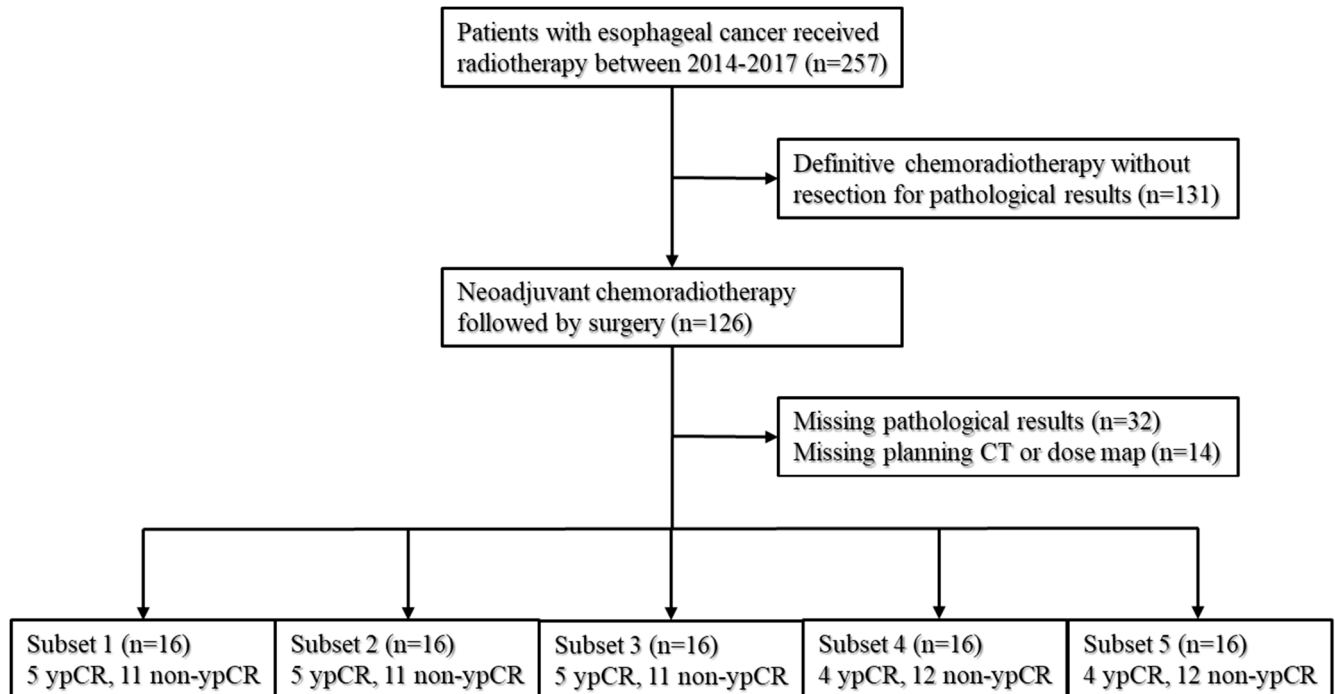


Figure 1. Patient inclusion and grouping.

2.2. Development of the Algorithm

The ypCR prediction algorithm represents an advancement of the state-of-the-art OCRNet (object contextual representations network) developed by Yuan et al. [16]. This algorithm is specifically designed for semantic segmentation tasks employing convolutional neural networks. To enhance its performance, the ypCR algorithm introduces the concept of soft regions, which prove particularly useful in approximate auto-planning scenarios. By integrating information related to radiation therapy dose distribution into the contextual representations of regions, the ypCR algorithm aims to refine the accuracy of semantic segmentation, as depicted in Figure 2. In this algorithm, object regions are defined based on varying levels of radiation isodose. To distinguish it from OCRNet, this variant is referred to as DCRNet (dose contextual representations network). The improved performance and distinctive attributes of the ypCR algorithm render it a valuable asset in the realms of semantic segmentation and radiation therapy planning.

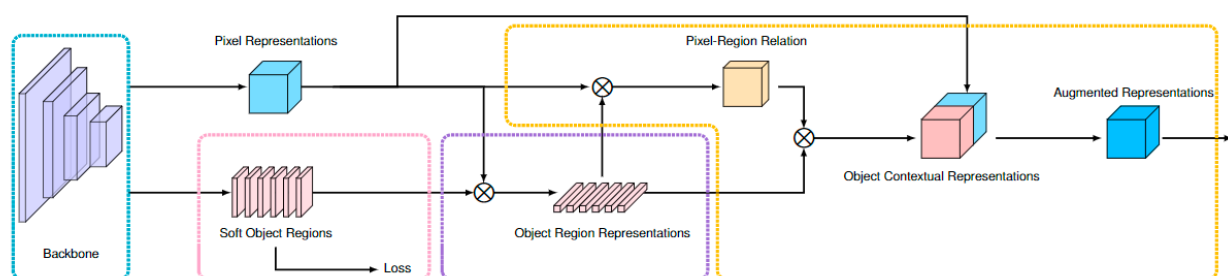


Figure 2. The architecture of DCRNet.

DCRNet is constructed upon HRNetV2p [17], which was initially pre-trained on the Cityscapes dataset, serving as the foundational framework for extracting feature maps from

planning CT images in a 2.5D fashion. The algorithm takes CT data as input, incorporating three consecutive slices covering the CTV from the radiotherapy (RT) structure set into the backbone model, as illustrated in Figure 3. Uniform ground-truth labels are assigned to all input units of the same patient, according to the patient's pathological response.

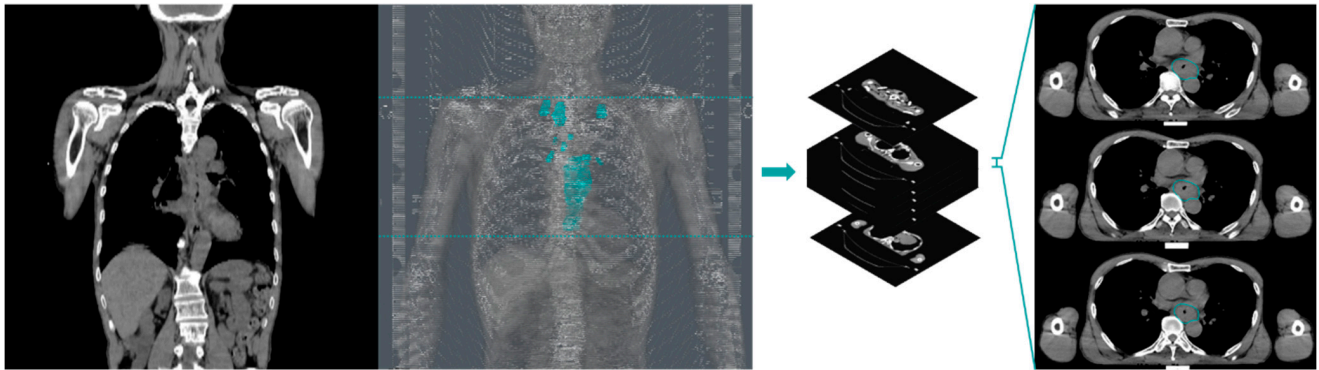


Figure 3. Image preprocessing.

To produce dose contextual representations, we employed channel softmax normalization for 10 soft regions, each corresponding to a 10% radiation isodose level (ranging from 0% to 100%), akin to dose-volume histogram analysis. In the model training phase, we utilized the dice loss function to supervise the model, simulating conditions akin to rough auto-planning. Subsequently, the resulting dose representation was amalgamated with the pixel representation to establish the dose contextual representations, following a methodology consistent with prior works [16].

To reflect the pathological outcomes, a fully connected layer was incorporated, consisting of two nodes. The output layer, comprising a single node, was supervised by a cross-entropy loss function designed to predict the likelihood of residual disease presence on each CT slice. The results derived from all input CT images for a given patient were summarized in three distinct manners: (1) the maximum value across the CTV, (2) the maximum between mean values obtained from every three consecutive slices, and (3) the mean value across the CTV (Figure 4). Performance comparisons were conducted across different models and summarization methods.

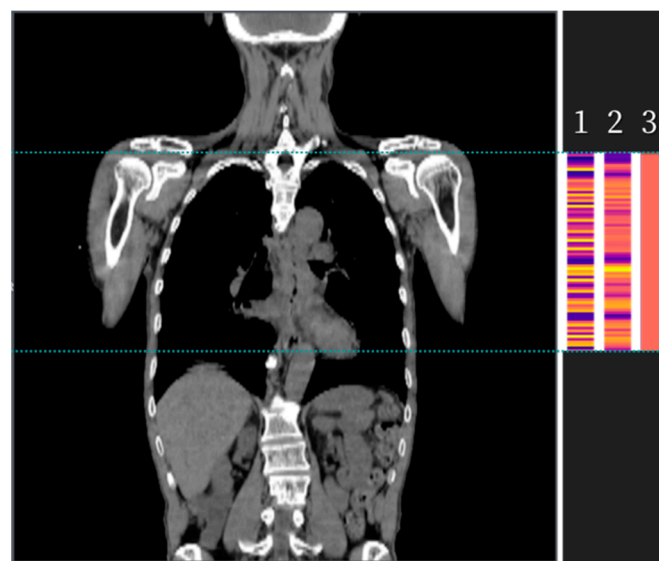


Figure 4. Three methods to summarize possibility in patient: (1) the maximum value across the CTV, (2) the maximum between mean values of every three continuous slices, (3) the mean value across the CTV.

For training, we employed an Adam optimizer with a learning rate set at 1×10^{-4} , implementing gradient clipping. To enhance training focus on challenging instances, we utilized Online Hard Example Mining (OHEM) [18]. Model training took place on a system equipped with an Intel Xeon E5-2630 v4 2.20 GHz CPU, an NVIDIA Tesla T4 GPU boasting 16 GB of memory, and 128 GB of RAM. The algorithm was implemented in PyTorch, building upon the OCRNet codebase.

2.3. Statistical Analysis

Categorical variables underwent statistical comparison using the χ^2 or Fisher exact test, while numeric variables were assessed using the Kruskal–Wallis test. The performance of DCRNet models was gauged using the area under the ROC curve (AUC) through 5-fold cross-validation, with p values for differences determined through DeLong’s test [19]. Cutoff points for accuracy, sensitivity, and specificity were established using the Youden Index. Calibration performance and clinical utility were evaluated via calibration curves and decision curves. Statistical significance was denoted by a 2-tailed p value less than 0.05. Data visualization was accomplished using Python software version 3.8, while R software version 4.0.5 was employed for statistical analysis.

2.4. Five-Fold Cross-Validation

Five-fold cross-validation is a widely used technique in machine learning for model evaluation and performance estimation. It involves the division of a dataset into five equal subsets or “folds”. The model is then trained and tested five times, where in each iteration, one fold is designated as the test set, and the remaining four folds are used for training. This process is repeated five times, ensuring that each fold is used as a test set exactly once. The results from these iterations are then averaged to provide a comprehensive and reliable assessment of the model’s performance. Five-fold cross-validation helps reduce the impact of randomness and ensures that the model’s performance estimate is more representative of its generalization capabilities compared to a single train/test split, making it a valuable tool in model selection and performance assessment.

3. Results

3.1. Baseline Characteristics

In this study, 80 patients participated, with a mean age of 55.68 (± 9.5) years, consisting of 76 (95%) males and four (5%) females. Predominantly, the lesions were situated in the middle part (41.1%), and the majority (71.2%) were diagnosed with stage III disease. The observed ypCR rate across all patients stood at 28.75%, aligning with findings from previous research. Notably, there were no statistically significant differences in age, lesion location, clinical staging, radiation dosage, or overall survival outcomes between the ypCR and non-ypCR groups (Table 1).

Table 1. Patients’ characteristics.

Characteristics	Category	All Patients, <i>n</i> (%)	ypCR, <i>n</i> (%)	Non-ypCR, <i>n</i> (%)	<i>p</i>
Sex	Male	76 (95)	20 (87.0)	56 (98.2)	0.04
	Female	4 (5)	3 (13.0)	1 (1.8)	
Age (year)	Mean (SD)	55.68 $\pm$ 9.5	55.48 $\pm$ 7.3	55.76 $\pm$ 10.3	0.28
Location	Proximal	5 (8.9)	3 (13.1)	2 (6.1)	0.16
	Middle	23 (41.1)	12 (52.2)	11 (33.3)	
	Distal	28 (50)	8 (34.8)	20 (60.6)	
Clinical stage	II	5 (6.3)	2 (8.7)	3 (5.3)	0.70
	III	57 (71.2)	17 (73.9)	40 (70.1)	
	IVA	18 (22.5)	4 (17.4)	14 (24.6)	
Total radiation dose (Gy)		44.0 $\pm$ 2.1	44.5 $\pm$ 2.4	43.8 $\pm$ 2.8	0.81
OS (month)		24.1 $\pm$ 13.7	27.7 $\pm$ 16.2	22.6 $\pm$ 14.8	0.08

SD, standard deviation. OS, overall survival.

3.2. The Comparison of Three Methods in the Patient Summary

The model computed a potential non-ypCR value based on images. To consolidate this value effectively across all images for a given patient, we assessed three distinct summarization methods. Initially, we adopted the approach of selecting the maximum value within the CTV. Subsequently, through thorough evaluation, we identified an optimal cutoff for this potential value at 0.525. This particular method yielded an impressive AUC score of 0.928, accompanied by a sensitivity of 0.845 and specificity of 0.852, as illustrated in Figure 5.

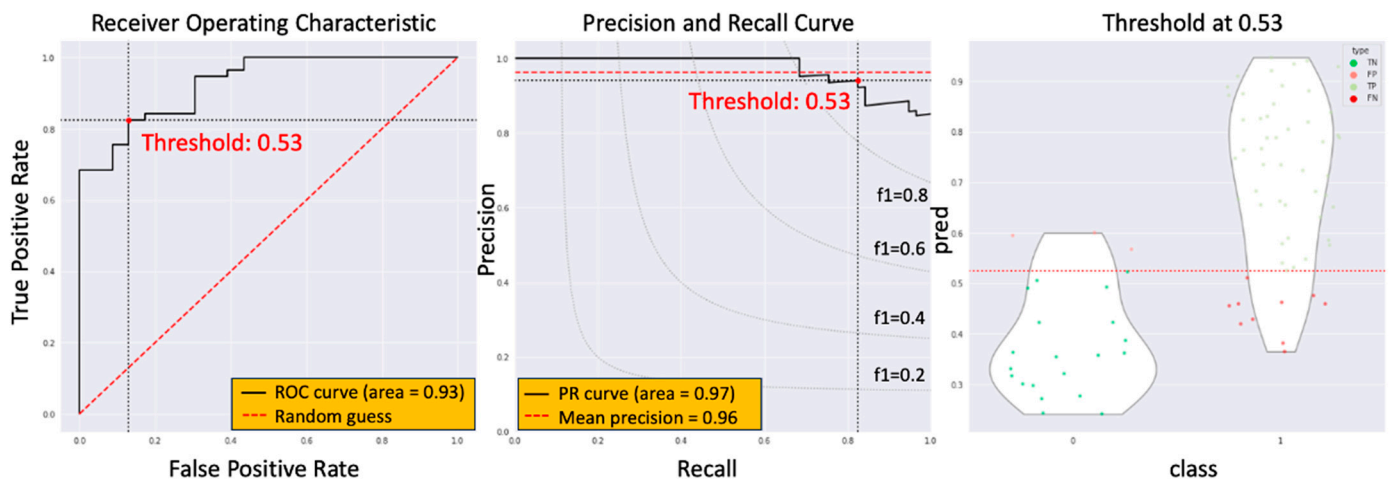


Figure 5. Model performance of the first summary method.

We employed the second method, which involved identifying the maximum among the mean values of every three consecutive slices. Through optimization, a possibility value cutoff of 0.457 was determined as optimal. This method yielded an AUC of 0.923, alongside a sensitivity of 0.840 and a specificity of 0.826, as depicted in Figure 6.

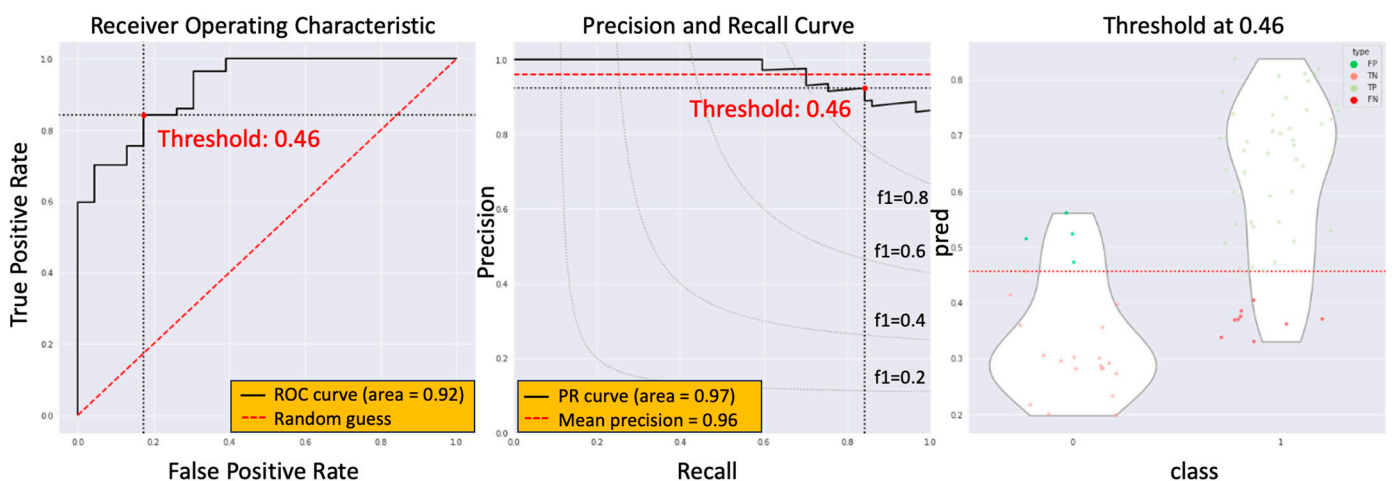


Figure 6. The model performance of the second summary method.

The third approach involved computing the mean value across the CTV, with an optimal cutoff set at 0.309. This method yielded an AUC of 0.924, a sensitivity of 0.825, and a specificity of 0.869 (Figure 7). These findings highlight the similarity in predictive performance among all three methods for identifying non-ypCR cases, albeit with minor discrepancies in optimal cutoff values and performance metrics.

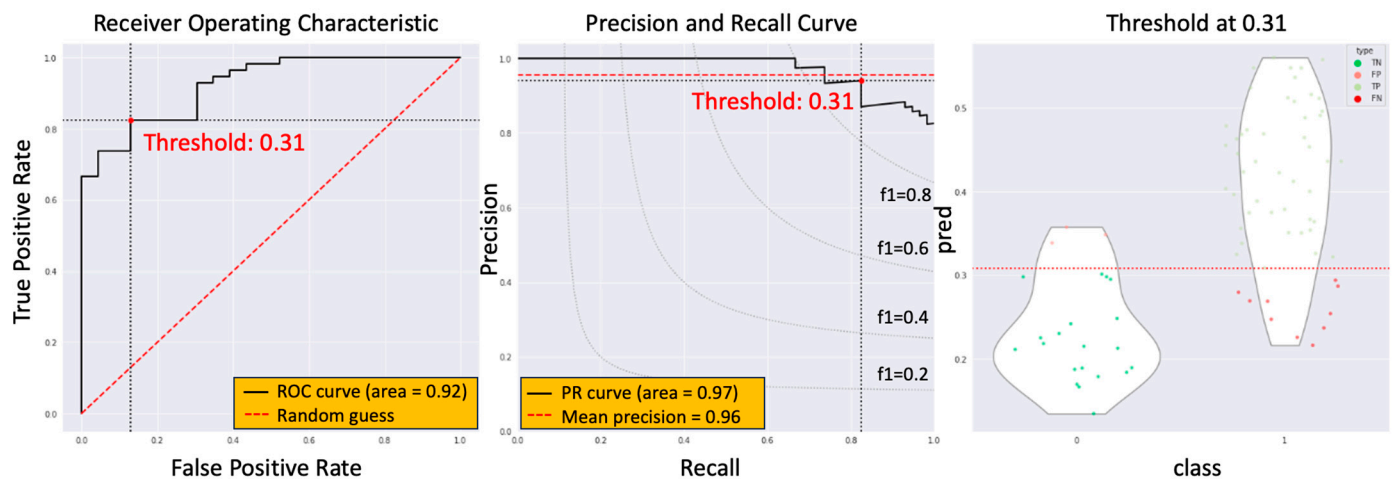


Figure 7. The model performance of the third summary method.

3.3. The Model Performance of DCRNet in Pathological Results

We observed no significant performance differences among the three summary methods and opted for simplicity by selecting the maximum value across the CTV. Subsequently, we conducted a comparative analysis of four backbone models (HRNetV2p, ResNet-101, EfficientNet, and DenseNet-121), both with and without dose contextual representations (Table 2). The HRNetV2p model integrated with dose contextual representations (DCR) outperformed the others, yielding an AUC of 0.928 (95% CI, 0.884–0.972) through 5-fold cross-validation, demonstrating a notable superiority over alternative models. Models incorporating DCR exhibited significantly better AUC outcomes in comparison to those without DCR.

Table 2. Comparison of predictive AUC in 5-fold cross validation.

AUC (95% CI)	ypCR	ypT0	ypN0
HRN + DCR	0.928 (0.884–0.972)	0.939 (0.928–0.950)	0.891 (0.881–0.901)
HRN	0.865 (0.856–0.873)	0.877 (0.865–0.889)	0.859 (0.846–0.872)
RN + DCR	0.829 (0.809–0.849)	0.845 (0.828–0.862)	0.813 (0.799–0.827)
RN	0.763 (0.751–0.775)	0.775 (0.759–0.791)	0.751 (0.733–0.769)
EN + DCR	0.836 (0.818–0.854)	0.847 (0.825–0.869)	0.825 (0.806–0.844)
EN	0.766 (0.749–0.783)	0.780 (0.756–0.804)	0.752 (0.736–0.768)
DN + DCR	0.832 (0.812–0.852)	0.844 (0.826–0.862)	0.820 (0.803–0.837)
DN	0.761 (0.742–0.780)	0.776 (0.756–0.796)	0.746 (0.731–0.761)

4. Discussion

Accurately detecting ypCR in esophageal cancer remains challenging with traditional imaging methods, such as computed tomography, positron emission tomography, endoscopic ultrasound, and magnetic resonance imaging, as their sensitivity ranges from 0.35 to 0.80 [5]. To overcome this limitation, computer-aided detection techniques, including radiomics and machine learning, have garnered interest as tools to enhance ypCR prediction. Prior research by Zhang et al. utilized a support vector machine model, achieving high accuracy by combining clinical, conventional PET, and radiomic PET features [6]. Beukinga et al. [7] and Yang et al. [8] developed multivariate logistic regression models incorporating textural features from PET and CT scans. Ypsilantis et al. demonstrated that a convolutional neural network outperformed traditional machine learning methods, such as support vector machine, logistic regression, random forest, and gradient boosting [9]. Hou et al. conducted two studies comparing an artificial neural network and support vector machine based on CT and MRI features, both yielding accurate treatment response predictions [10,11]. Notably, no significant difference in predictive performance emerged

between artificial neural network and support vector machine models, emphasizing the model choice may not be the primary factor [10,11]. Additionally, recent research delves into radiomic features of the peritumoral region, deemed reflective of the tumor microenvironment. Hu et al. [12] developed a support vector machine model with a radial basis function kernel, enhancing the AUC by integrating intratumoral and peritumoral features. While these studies underscore the efficacy of imaging features and machine learning models in predicting treatment response in esophageal cancer, wider clinical adoption awaits further validation in larger patient cohorts and the standardization of feature extraction methods.

The current study introduces an innovative approach, the radiotherapy dose map-guided deep learning method, to predict esophageal cancer patients' responses to nCRT. This method integrates imaging data with radiotherapy dose map variations, offering a more comprehensive assessment of patient responses. Notably, the study population's demographics and ypCR rate aligned with previous research findings [5]. There were no statistically significant differences observed in age, tumor location, clinical stage, or radiation dose between the ypCR and non-ypCR groups. Three methods were employed to summarize the potential value of non-ypCR predicted by the model. The first method, taking the maximum value across the CTV, yielded an AUC score of 0.928, a sensitivity of 0.845, and a specificity of 0.852, with an optimal cutoff at 0.525. The second method, based on the maximum between the mean values of three consecutive slices, produced an AUC of 0.923, a sensitivity of 0.840, and a specificity of 0.826, with an optimal cutoff at 0.457. The third method, utilizing the mean value across the CTV, resulted in an AUC of 0.924, a sensitivity of 0.825, and a specificity of 0.869 with an optimal cutoff at 0.309. These findings indicate that all three methods effectively predicted the presence of non-ypCR, albeit with variations in optimal cutoff values and performance metrics. To assess the impact of dose map information, this study compared four distinct deep learning models (HRNetV2p, ResNet-101, EfficientNet, and DenseNet-121) both with and without dose contextual representations. The simplest summary method, the maximum value across the CTV, was chosen. Results revealed that the HRNetV2p model, when enhanced with dose contextual representations (DCR), outperformed others, achieving an AUC of 0.928 through 5-fold cross-validation. This outcome marked a significant improvement over models lacking DCR. Notably, models without DCR demonstrated AUC results in line with prior research findings [6–12].

Given the small number size of this study, five-fold cross-validation method was performed for model evaluation to address the limitations of the traditional train/test split method. The comprehensive process of 5-fold cross-validation provides several key advantages compared to a single train/test split. First, it delivers a more reliable performance estimate by averaging results over five iterations, reducing the influence of random data splits and outliers. Second, it maximizes data utilization, as all data points are used for both training and testing, enhancing the model's ability to learn from the entire dataset. Third, it mitigates the risk of overfitting by testing the model on five different subsets, making it less likely to memorize peculiarities of a single split. Moreover, it assesses model robustness by revealing whether the model's performance is consistent across the folds, indicating its generalization capabilities. Lastly, five-fold cross-validation enables fair model comparisons, aiding in the selection of the best-performing model, making it a valuable tool in machine learning for accurate and reliable model assessment.

In our study, a noteworthy observation was the statistically significant difference in the male-to-female ratio between the ypCR group and the non-ypCR group, with the ypCR group having a higher proportion of female patients. This finding aligns with a prior study that identified female sex as an independent favorable prognostic factor in patients with ESCC who underwent definitive radiotherapy [20]. It is worth noting that our study cohort consisted of 95% male patients, reflecting the male-to-female ratio observed in the general population of individuals diagnosed with ESCC in Taiwan [21]. According to the Cancer Registry Annual Report 2020 published by the Ministry of Health

and Welfare of Taiwan in December 2022, out of the 2652 newly diagnosed ESCC cases in 2020, 93% were male patients [21]. However, it is essential to recognize that while ESCC is typically more prevalent in males, with a global male-to-female age-standardized incidence rate ratio (ASIR) of approximately 3.3:1, significant variations exist by country [22]. For instance, in the Republic of Korea and several Baltic/European countries, including Belarus, Ukraine, Slovakia, Lithuania, Latvia, and Estonia, studies have reported extremely high male-to-female ASIRs, exceeding 10:1 and even reaching as high as 21:1 for ESCC [22]. The male-to-female ASIR for ESCC in Taiwan is similar to these countries, highlighting the geographical diversity in the incidence of this disease.

This study highlights the potential of a radiotherapy dose map-guided deep learning approach, combining imaging data and radiotherapy dose map variables, to predict ypCR effectively. These results hold the promise of enhancing patient response prediction, enabling more personalized and effective care. This includes avoiding unnecessary surgeries and optimizing radiotherapy treatment planning by incorporating the algorithm into the treatment planning system, thereby potentially improving patient outcomes and treatment strategies. Nevertheless, several limitations in the experiment may affect the validity of the results. Firstly, the relatively small sample size, comprising only 80 patients, could potentially restrict the generalizability of the findings, as larger sample sizes are often required for more robust and accurate predictions. Secondly, the use of data from a single center introduces potential bias into the results, particularly if there are variations in patient populations or treatment protocols across centers. Finally, the study did not evaluate the model's impact on patient outcomes, including survival and quality of life, which should be addressed in future research to provide a more comprehensive understanding of its clinical significance.

5. Conclusions

This study introduced a novel approach to predict nCRT response in esophageal cancer patients, employing a deep learning method that considers both imaging data and radiotherapy dose map variations. The HRNetV2p with DCR model demonstrated superior performance, achieving an AUC of 0.928 through 5-fold cross-validation, signifying significant improvement over alternative models. While all three summary methods yielded similar results for non-ypCR prediction, the simplicity of selecting the maximum value across the CTV was preferred. This research highlights the promise of radiotherapy dose map-guided deep learning in enhancing ypCR prediction. However, further studies involving larger patient cohorts are essential to validate these findings and unlock the full potential of the proposed methodology.

Author Contributions: Conceptualization, W.-K.Y. and S.-C.C.; methodology, S.-C.C. and I.-T.H.; software, W.-L.Y.; validation, T.-Y.T., Y.-A.L. and C.-K.Y.; formal analysis, M.-T.P.; investigation, W.-L.Y.; resources, W.-K.Y.; data curation, E.-L.S.; writing—original draft preparation, W.-K.Y. and S.-C.C.; writing—review and editing, T.-Y.T., Y.-A.L., C.-K.Y. and M.-T.P.; visualization, T.-Y.T.; supervision, W.-K.Y.; project administration, I.-T.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Chang Gung Memorial Hospital (protocol code 202001098B0 approved on 29 June 2020).

Informed Consent Statement: All participants included in the study signed informed consent forms for treatment before the start of study treatments. Patient consent for enrollment into this study was waived due to the retrospective design.

Data Availability Statement: Data supporting reported results can be provided upon request.

Acknowledgments: The authors thank all members of the Esophageal Cancer Team at the Linkou Chang Gung Memorial Hospital for their dedication to patient care.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Rustgi, A.K.; El-Serag, H.B. Esophageal carcinoma. *N. Engl. J. Med.* **2014**, *371*, 2499–2509. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Shapiro, J.; van Lanschot, J.J.B.; Hulshof, M.C.C.M.; van Hagen, P.; van Berge Henegouwen, M.I.; Wijnhoven, B.P.L.; van Laarhoven, H.W.M.; Nieuwenhuijzen, G.A.P.; Hospers, G.A.P.; Bonenkamp, J.J.; et al. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): Long-term results of a randomised controlled trial. *Lancet Oncol.* **2015**, *16*, 1090–1098. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Yang, H.; Liu, H.; Chen, Y.; Zhu, C.; Fang, W.; Yu, Z.; Mao, W.; Xiang, J.; Han, Y.; Chen, Z.; et al. Neoadjuvant chemoradiotherapy followed by surgery versus surgery alone for locally advanced squamous cell carcinoma of the esophagus (NEOCRTEC5010): A phase III multicenter, randomized, open-label clinical trial. *J. Clin. Oncol.* **2018**, *36*, 2796–2803. [\[CrossRef\]](#) [\[PubMed\]](#)
5. de Gouw, D.J.; Klarenbeek, B.R.; Driessen, M.; Bouwense, S.A.; van Workum, F.; Fütterer, J.J.; Rovers, M.M.; Broek, R.P.T.; Rosman, C. Detecting Pathological Complete Response in Esophageal Cancer after Neoadjuvant Therapy Based on Imaging Techniques: A Diagnostic Systematic Review and Meta-Analysis. *J. Thorac. Oncol.* **2019**, *14*, 1156–1171. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Zhang, H.; Tan, S.; Chen, W.; Kligerman, S.; Kim, G.; D'Souza, W.D.; Suntharalingam, M.; Lu, W. Modeling pathologic response of esophageal cancer to chemoradiation therapy using spatial-temporal 18F-FDG PET features, clinical parameters, and demographics. *Int. J. Radiat. Oncol. Biol. Phys.* **2014**, *88*, 195–203. [\[CrossRef\]](#)
7. Beukinga, R.J.; Hulshoff, J.B.; van Dijk, L.V.; Muijs, C.T.; Burgerhof, J.G.; Kats-Ugurlu, G.; Slart, R.H.; Slump, C.H.; Mul, V.E.; Plukker, J.T. Predicting Response to Neoadjuvant Chemoradiotherapy in Esophageal Cancer with Textural Features Derived from Pretreatment 18F-FDG PET/CT Imaging. *J. Nucl. Med.* **2017**, *58*, 723–729. [\[CrossRef\]](#)
8. Yang, Z.; He, B.; Zhuang, X.; Gao, X.; Wang, D.; Li, M.; Lin, Z.; Luo, R. CT-based radiomic signatures for prediction of pathologic complete response in esophageal squamous cell carcinoma after neoadjuvant chemoradiotherapy. *J. Radiat. Res.* **2019**, *60*, 538–545. [\[CrossRef\]](#)
9. Ypsilantis, P.-P.; Siddique, M.; Sohn, H.-M.; Davies, A.; Cook, G.; Goh, V.; Montana, G. Predicting Response to Neoadjuvant Chemotherapy with PET Imaging Using Convolutional Neural Networks. *PLoS ONE* **2015**, *10*, e0137036. [\[CrossRef\]](#)
10. Hou, Z.; Li, S.; Ren, W.; Liu, J.; Yan, J.; Wan, S. Radiomic analysis in T2W and SPAIR T2W MRI: Predict treatment response to chemoradiotherapy in esophageal squamous cell carcinoma. *J. Thorac. Dis.* **2018**, *10*, 2256–2267. [\[CrossRef\]](#)
11. Hou, Z.; Ren, W.; Li, S.; Liu, J.; Sun, Y.; Yan, J.; Wan, S. Radiomic analysis in contrast-enhanced CT: Predict treatment response to chemoradiotherapy in esophageal carcinoma. *Oncotarget* **2017**, *8*, 104444–104454. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Hu, Y.; Xie, C.; Yang, H.; Ho, J.W.K.; Wen, J.; Han, L.; Chiu, K.W.H.; Fu, J.; Vardhanabhuti, V. Assessment of Intratumoral and Peritumoral Computed Tomography Radiomics for Predicting Pathological Complete Response to Neoadjuvant Chemoradiation in Patients With Esophageal Squamous Cell Carcinoma. *JAMA Netw. Open* **2020**, *3*, e2015927. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Murphy, M.B.; Xiao, L.; Patel, V.R.; Maru, D.M.; Correa, A.M.; Amlashi, F.G.; Liao, Z.; Komaki, R.; Lin, S.H.; Skinner, H.D.; et al. Pathological complete response in patients with esophageal cancer after the trimodality approach: The association with baseline variables and survival-The University of Texas MD Anderson Cancer Center experience. *Cancer* **2017**, *123*, 4106–4113. [\[CrossRef\]](#)
14. Shen, J.; AME Thoracic Surgery Collaborative Group; Kong, M.; Yang, H.; Jin, K.; Chen, Y.; Fang, W.; Yu, Z.; Mao, W.; Xiang, J.; et al. Pathological complete response after neoadjuvant treatment determines survival in esophageal squamous cell carcinoma patients (NEOCRTEC5010). *Ann. Transl. Med.* **2021**, *9*, 1516. [\[CrossRef\]](#)
15. Yap, W.-K.; Chang, Y.-C.; Hsieh, C.-H.; Chao, Y.-K.; Chen, C.-C.; Shih, M.-C.; Hung, T.-M. Favorable versus unfavorable prognostic groups by post-chemoradiation FDG-PET imaging in node-positive esophageal squamous cell carcinoma patients treated with definitive chemoradiotherapy. *Eur. J. Nucl. Med. Mol. Imaging* **2018**, *45*, 689–698. [\[CrossRef\]](#)
16. Yuan, Y.; Chen, X.; Chen, X.; Wang, J. Segmentation Transformer: Object-Contextual Representations for Semantic Segmentation. *arXiv* **2019**, arXiv:1909.11065.
17. Sun, K.; Zhao, Y.; Jiang, B.; Cheng, T.; Xiao, B.; Liu, D.; Mu, Y.; Wang, X.; Liu, W.; Wang, J. High-Resolution Representations for Labeling Pixels and Regions. *arXiv* **2019**, arXiv:1904.04514.
18. Shrivastava, A.; Gupta, A.; Girshick, R. Training Region-based Object Detectors with Online Hard Example Mining. *arXiv* **2016**, arXiv:1604.03540.
19. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* **1988**, *44*, 837–845. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Luo, H.-S.; Xu, H.-Y.; Du, Z.-S.; Li, X.-Y.; Wu, S.-X.; Huang, H.-C.; Lin, L.-X. Impact of sex on the prognosis of patients with esophageal squamous cell cancer underwent definitive radiotherapy: A propensity score-matched analysis. *Radiat. Oncol.* **2019**, *14*, 74. [\[CrossRef\]](#)

21. Ministry of Health and Welfare Taiwan. Cancer Registry Annual Report, 2020. 2022. Available online: <https://www.hpa.gov.tw/Pages/Detail.aspx?nodeid=269&pid=16434> (accessed on 15 October 2023).
22. Wang, S.; Zheng, R.; Arnold, M.; Abnet, C.; Zeng, H.; Zhang, S.; Chen, R.; Sun, K.; Li, L.; An, L.; et al. Global and national trends in the age-specific sex ratio of esophageal cancer and gastric cancer by subtype. *Int. J. Cancer* **2022**, *151*, 1447–1461. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.