Physica Medica 62 (2019) 111–119

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Original paper

Neo-adjuvant chemoradiotherapy response prediction using MRI based ensemble learning method in rectal cancer patients

Sajad P. Shayesteh[a], Afsaneh Alikhassi[b], Armaghan Fard Esfahani[c], M. Miraie[d], Parham Geramifar[c], Ahmad Bitarafan-rajabi[e,f], Peiman Haddad[g,*]

[a] Department of Physiology, Pharmacology and Medical Physics, Faculty of Medicine, Alborz University of Medical Sciences, Karaj, Iran
[b] Department of Radiology, Cancer Institute of Iran, Tehran University of Medical Sciences, Tehran, Iran
[c] Research Center for Nuclear Medicine, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran
[d] Cancer Research Centre & Radiation Oncology Department, Cancer Institute, Tehran University of Medical Sciences, Tehran, Iran
[e] Cardiovascular Intervention Research Center, Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Sciences, Tehran, Iran
[f] Echocardiography Research Center, Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Sciences, Tehran, Iran
[g] Radiation Oncology Research Center, Cancer Institute, Tehran University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Keywords:
MRI
Radiomics
Machine learning
Rectal cancer
Chemoradiotherapy

ABSTRACT

Objectives: The aim of this study was to investigate and validate the performance of individual and ensemble machine learning models (EMLMs) based on magnetic resonance imaging (MRI) to predict neo-adjuvant chemoradiation therapy (nCRT) response in rectal cancer patients. We also aimed to study the effect of Laplacian of Gaussian (LOG) filter on EMLMs predictive performance.

Methods: 98 rectal cancer patients were divided into a training (n = 53) and a validation set (n = 45). All patients underwent MRI a week before nCRT. Several features from intensity, shape and texture feature sets were extracted from MR images. SVM, Bayesian network, neural network and KNN classifiers were used individually and together for response prediction. Predictive performance was evaluated using the area under the receiver operator characteristic (ROC) curve (AUC).

Results: Patients' nCRT responses included 17 patients with Grade 0, 28 with Grade 1, 34 with Grade 2, and 19 with Grade 3 according to AJCC/CAP pathologic grading. In without preprocessing MR Image the best result was for Bayesian network classifier with AUC and accuracy of 75.2% and 80.9% respectively, which was confirmed in the validation set with an AUC and accuracy of 74% and 79% respectively. In EMLMs the best result was for 4 (SVM.NN.BN.KNN) classifier EMLM with AUC and accuracy of 97.8% and 92.8% in testing and 95% and 90% in validation set respectively.

Conclusions: In conclusion, we observed that machine learning methods can used to predict nCRT response in patients with rectal cancer. Preprocessing LOG filters and EL models can improve the prediction process.

1. Introduction

Colorectal cancer is one of the most common malignancies and cause of cancer related death in developing countries [1]. The standard therapy approach for colorectal cancer is neo-adjuvant chemoradiation therapy (nCRT) followed by total mesorectal excision (TME). Previous studies have consistently demonstrated that TME is associated with challenging therapy response. On the other hand, concerns still exist regard the response assessment and prediction [2].

Response prediction in colorectal cancer therapy can be made using histopathological examination of surgically resected specimens. By advances in medical imaging and by introducing new imaging techniques, interests have been found to have a non-invasive method for cancer therapy response prediction. Several studies have identified the potential roles of imaging modalities including magnetic resonance imaging (MRI), positron emission tomography (PET) and computed

Abbreviations: EMLMs, ensemble machine learning models; LOG, Laplacian of Gaussian; nCRT, neo-adjuvant chemoradiation therapy; MRI, magnetic resonance imaging; PET, positron emission tomography; CT, computed tomography; TRG, tumor regression grade; AUC, area under the curve; ROC, receiver operator characteristic; ML, machine learning; TME, total mesorectal excision; LARC, locally advanced rectal cancer; AJCC, American Joint Committee on Cancer; CAP, College of American Pathologists; SVM, support vector machine; BN, Bayesian network; NN, neural network; KNN, K nearest neighbor; PCR, Pathologic Complete Response

* Corresponding author.
E-mail address: haddad@tums.ac.ir (P. Haddad).

https://doi.org/10.1016/j.ejmp.2019.03.013
Received 31 August 2018; Received in revised form 23 February 2019; Accepted 17 March 2019
Available online 15 May 2019

tomography (CT) for therapy response evaluating in different cancers. In regard to colorectal cancer, tumor regression grade (TRG) evaluation by MRI is based on the size of the tumor and qualitative assessment of change in signal intensity. Therefore, evaluation and detection of tumor response at the early stage of nCRT would be beneficial to patient treatment but change in size and qualitative features of tumor are not apparent in early stages of therapy [3,4].

Early prediction of tumor response to therapy is of great value in locally advanced rectal cancer patients to avoid ineffective treatment and to get a chance to receive another beneficial treatment. More accurate biomarkers for early prediction of Response to therapy can lead to alteration of the initial treatment plan. Then to reduce treatment morbidity and improve treatment performance qualified stratification can be used. So, by early detection and prediction of patient's response to therapy and based on a personalized treatment approach we can divide patients into different prognostic groups. In patients with locally advanced rectal cancer undergoing nCRT and surgery, 45% of patients will require permanent colostomy. In these patients, Identification of patients with a clinical complete response prior to surgery would enable optimization of the surgical approach with “organ-preserving” procedures, which would result in a reduction in surgical morbidity. As well as, early detection of poor responders to nCRT in LARC patients would provide the opportunity for these patients to proceed directly to surgery thereby avoiding morbidity of nCRT or for intensified treatment regimens [5–7]. In recent years, evidence showing that qualitative parameters obtained by medical images texture analysis, reflecting the spatial variation and heterogeneity of voxel intensities within a tumor, may led to additional predictive and prognostic information for better tissue characterization as well as personalization of cancer treatment in the clinic [8–10].

Recently, a considerable amount of data has been published on radiomics, which aims to quantify tumor heterogeneity by mineable data extracted from tumor images. Previous radiomics studies have indicated that image features extracted from different medical images could predict nCRT response in the rectal cancer patients [11–13]. Also the ability of radiomic features as the potential biomarkers, have been investigated in breast cancer, prostate cancer, cerebral gliomas, head and neck cancer and liver diseases [14–19].

Based on several radiomics studies, a wide range of machine learning (ML) algorithms have been applied on radiomic features for feature selection and classification. Studies have proved that the selected ML algorithm has a great impact on prognostic/predictive performance of developed radiomic models [20–22]. As new approach, ensemble learning (EL) has shown great potential in different radiomic studies such as skin, head and neck and colorectal cancer [23–26].

EL method is a machine learning process formed by fusing an array of individual classifiers to get better prediction performance. It works by merging multiple machine learning models into a single model which helps to avoid over fitting and encourages generalization of the predictor. In addition to improved prediction performance, models combination also remove the difficulty in choosing the models with the best prognostic/predictive performance. Each individual classifier used in the construction of the overall ensemble model will have its own performance strengths, resulting in powerful and more accurate predictions when these machine learning models are used in combination [27–30].

In the present study, we aimed to investigate and validate different individual and ensemble ML algorithms on radiomic features extracted from rectal MRI to find best predictive models. Also the effect of Laplacian of Gaussian (LOG) filter on EMLMs predictive performance was studied.

2. Materials and methods

The overall framework of this study was depicted in Fig. 1. In the below, different phases of study are outlined.

2.1. Patient data

This prospective study was approved by ethics committee of Iran and Tehran University of Medical Sciences and informed consent requirement was waived. 98 patients who met the inclusion criteria were divided into two groups including training dataset with 53 patients and the validation dataset with 45 patients from October 2016 to Jun 2018. To reduce the imaging protocol and treatment system effect on image features all patients of training and validation groups were selected randomly from same treatment and imaging center with same inclusion criteria. Inclusion criteria were location of tumor within 15 cm above anal verge, tumor penetration to perirectal fat (cT3–4) or lymph node involvement, age ≤ 80 years, WHO performance status of 0–2, normal CBC, liver and renal function tests and lack of any prior treatment for the disease. Varian Clinac 2100C system (Varian Medical Systems, Palo Alto, CA, USA) is used for radiation therapy. All patient received concurrent nCRT. They received 45–46 Gy external beam radiation in 23–25 fractions with 18 MV photons to the tumor and loco regional disease including pre-sacral and internal iliac lymph nodes with a boost to the tumor for a total of 50–50.4 Gy, concurrent capecitabine at 825 mg/m2 twice daily.

2.2. Image acquisition:

All images were acquired a week before nCRT on a 3.0-T MRI system (Tesla-Trio, Siemens Healthcare, Germany). A 32-channel pelvic phased array coil was used for signal reception. Transverse, coronal and sagittal T2W-weighted were acquired with same protocol as follows: turbo spin echo (TSE) sequence, repetition time/echo (msec) 4800/97, 2D acquisition, signal average (NSA) 3, acceleration factor 2, phase resolution 70, section thickness 3 mm, intersection gap 0.8 mm, field of view 35 cm, and matrix, 256 × 248.

2.3. Tumor segmentation

Segmentation of rectal tumor tissues on the axial T2w MR images was done by two readers: a ten years’ experience radiation oncologist, and a fifteen years’ experience radiologist using a designated multi-platform, free and open source software package for visualization and medical image computing (3D slicer, version 4.8.1; available at: http://slicer.org/). Tumor tissues were first reviewed in all slices and then were drawn on the T2W images. All the tumor sites were segmented on every MR image creating a volume of interest (VOI).

2.4. Preprocessing and feature extraction

Texture analysis used to quantify the spatial variations in gray level values within an image and thus can provide useful quantitative information of medical images for prognostic/predictive Targets. However, they are sensitive to data acquisition conditions due to the use of different protocols and to intra- and inter-scanner variations in the MR Imaging. Accordingly, Before feature extraction and in order to noise reduction, intensity normalization and discretization, all MR images were pre-processed by the method proposed by Collewet et al. and also discretization to 64 Gy levels. In Collewet et al. method, all image intensities are normalized between $\mu \pm 3\sigma$, where μ is the mean value of gray-levels inside the region of interest (ROI), and σ is the standard deviation [31].

Also, to test the filter effect on radiomic model performance, we applied our feature extraction on T2W MR Images with and without processing filters. The filters were including Laplacian-of-Gaussian (LOG) filter with sigma 0.5 and 1.5. For preprocessing with LOG filters and feature extraction we used the freely available radiomic software, imaging biomarker explorer (IBEX) that runs in Matlab platform.

Various radiomic features from different feature sets including intensity, shape and texture based features were extracted from processed

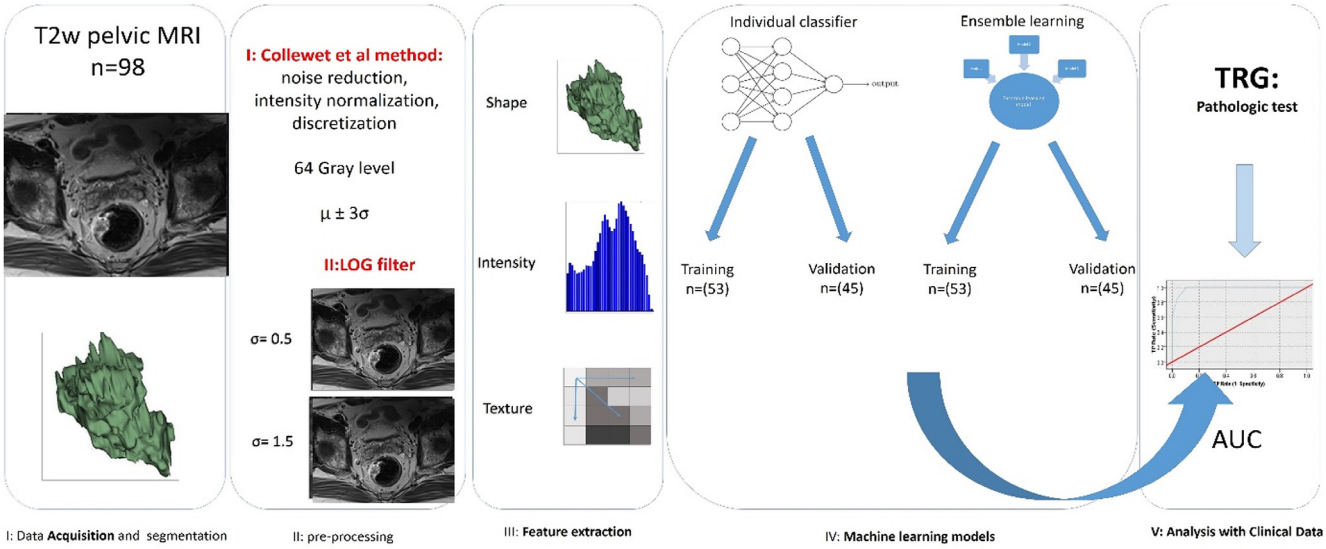

Fig. 1. Overall framework of study.

and un-processed T2W MR Images. Extracted features included shape features (n = 17) [32], intensity histogram features (n = 9), intensity direct (n = 19), neighbor intensity difference (n = 5), co-occurrence matrix features (n = 19), and gray level run-length matrix features (n = 11) [33–36].

2.5. Feature selection

In ML you need not use every feature at your access for creating a prediction algorithm. You can design your algorithm by using really relevant and robust features. In order to 1 – enable the ML model to train faster 2 – reduce the complexity of a model for easer interpretation 3 – improve the accuracy of a model with right subset selection and 4 – overcome overfitting issue, identify and remove unneeded, irrelevant and redundant features that do not contribute to the accuracy of a predictive model. To provide faster and more cost-effective models, we used Pearson's chi-square X2 algorithm. Feature selection was performed by SPSS Modeler 18 software (IBM Corp., Armonk, NY, USA).

2.6. Response assessment

For all patients surgery was done 4–8 weeks after nCRT. After inking, the specimens were fixed in formalin for 24 h. The whole tumor and mesorectum were serially sliced, axially, at 3 mm intervals, and treatment response was assessed according to the 4-category American Joint Committee on Cancer and College of American Pathologists (AJCC/CAP). All surgically resected specimens were histopathologically examined and analyzed by an experienced pathologist. The TRG according to AJCC/CAP was established as the following: grade 0 (Pathologic Complete Response (PCR)): is defined as no viable cancer cell, grade: 1 (moderate response): means single cells or small groups of cancer cells, grade 2 (minimal response): is residual cancer outgrown by fibrosis and grade 3 (poor response): is fibrosis outgrown by residual cancer [37–39]. Patients were divided into either the responder (Grade 0 or Grade 1) or non-responder (Grade 2 or Grade 3) group according to according to AJCC/CAP pathologic grading.

2.7. Univariate radiomic analysis

For univariate analysis, significant radiomic features were selected and a logistic regression classifier was used to find their predictive performance (based on AUC). Also, these features were compared between responder and non-responder groups. A paired *t*-test was performed to assess the significance of the differences between two groups. Statistical significance was assumed if $p < 0.05$ and all reported *p*-values are two-sided.

2.8. Multivariate radiomic analysis

For multivariable radiomic modelling, we separated responder from non-responder groups by using four individual classifiers separately, including support vector machine (SVM), Bayesian network (BN), neural network (NN) and K nearest neighbor (KNN). In addition, ensemble learning models are used to raise individual performance of multiple classifiers via information fusion in the ensemble learning methods (Table 1) [40,41]. All classifications were performed by SPSS Modeler 18 software (IBM Corp., Armonk, NY, USA). Predictive performance was evaluated using the area under the receiver operator characteristic (ROC) curve (AUC) and accuracy. Independent MR Images (n = 45) is used for external validation.

For the training and external validation sets, Student's *t*-test was used to assess differences in the imaging parameters between the two groups. A P value of less than 0.05 was considered statistically significant. Statistical analyses were performed using statistical software (R version 3.5.1, R Core Team, Vienna, Austria).

Table 1
Individual and boosted models tested in this study for nCRT response prediction in advanced rectal cancer patients.

N of combined classifiers	Model type	Used machine learning models
1	Individual	SVM
1	Individual	NN
1	Individual	BN
1	Individual	KNN
2	Ensemble	SVM.NN
2	Ensemble	SVM.BN
2	Ensemble	SVM.KNN
2	Ensemble	NN.BN
2	Ensemble	NN.KNN
2	Ensemble	BN.KNN
3	Ensemble	SVM.NN.BN
3	Ensemble	SVM.NN.KNN
3	Ensemble	SVM.BN.KNN
3	Ensemble	NN.BN.KNN
4	Ensemble	SVM.NN.BN.KNN

Table 2
Patient characteristics.

Demographics	Training dataset	Validation dataset
Gender (N [%])		
Male	33 [62.3]	30 [66.6]
Female	20 [37.7]	15 [33.4]
Total	53 [54.1]	45 [45.9]
Age (N [%])		
18–40	14 [26.4]	10 [22.2]
41–60	17 [32.1]	16 [35.6]
> 61	22 [41.5]	19 [42.2]
Response (N [%])		
Grade 0	9 [17]	8 [17.7]
Grade 1	15 [28.3]	13 [29]
Grade 2	19 [35.8]	15 [33.3]
Grade 3	10 [18.9]	9 [20]

3. Results

3.1. Patients and response

98 patients (63 men; mean age, 67.01 years; age range, 31–80 years; 35 women; mean age, 55.3 years; age range, 27–67 years) with local advanced rectal cancer were included in the study. 45 patients (17 patients with Grade 0, 28 patients with Grade 1) have been found as responder and 53 patients (34 patients with Grade 2, and 19 patients with Grade 3) have no response to nCRT according to AJCC/CAP pathologic grading. Patient data was detailed in table 2. There was no difference between the training dataset and the validation dataset in the clinicopathologic characteristics (p = 0.428–0.523).

3.2. Radiomic analysis

Several radiomic features were selected as highly correlated with therapy response. Our univariate analysis showed that eleven radiomic features have high correlation with nCRT response. We found that all eleven top radiomic features are from co-occurrence matrix (COM) feature set. The results on AUC logistic regression classifier for these feature are shown in Fig. 2. As was seen, the feature Dissimilarity has the highest performance (AUC, 0.65) followed by Sum Average (AUC, 0.64), Inter Quartile Range (AUC, 0.63), Cluster Tendency (AUC, 0.63), Variance (AUC, 0.63) and Cluster Prominence (AUC, 0.61).

Results on selected radiomic feature values between responder and non-responder groups were shown in Fig. 3. As was shown, although there are differences among features between two groups, but there are no significant differences between two groups in all features (p-value > 0.05).

For multivariate radiomic modelling, our results were shown in Table 3 for testing and external validation dataset. For each pat of multivariate analysis, feature selection algorithm is used to select most robust features from 11 top features of univariate analysis. Classification for all ML models have been done using 4 top ranked features. For individual ML classification, BN classifier with AUC 75.2% was found as high predictive model for un-processed T2W image features, followed by KNN (AUC, 73.40), NN (AUC, 71.30), and SVM (AUC, 68). These result were confirmed in validation dataset with AUC of BN (74%), KNN (73%), NN (70%), and SVM (72%) (Fig. 4). Our result on pre-processed images showed that LOG filter improves the classifiers performance. For images pre-processed with σ = 0.5, our results showed that KNN classifier is more predictive rather than other models (AUC, 86.60 training set, 81.7 validation set). In regard to σ = 1.5, BN classifier had the highest performance (AUC, 0.92 training set, 87.8 validation set) (Fig. 5).

In regard to ensemble learning for response prediction, our results for training and validation datasets are summarized in Table 4. As was shown, ensemble learning provided by combination of four and three classifiers are resulted in high predictive models in all un-processed and processed images. For un-processed images, SVM.NN.BN.KNN (AUC, 0.926, training set, 0.90 validation set) and SVM.BN.KNN (AUC, 0.926 training set, 0.884 validation set) were found as high predictive models, followed by NN.BN.KNN (AUC, 0.923, training set, 0.872 validation set), and SVM.NN.KNN (AUC, 0.89, training set, 0.864 validation set) (Figs. 6 and 7).

For processed images also, SVM.NN.BN.KNN had the highest performance. When we compared ensemble learning applied on processed and un-processed, it was found that, processing will result in more predictive models and processing using high filter value (σ = 1.5) will

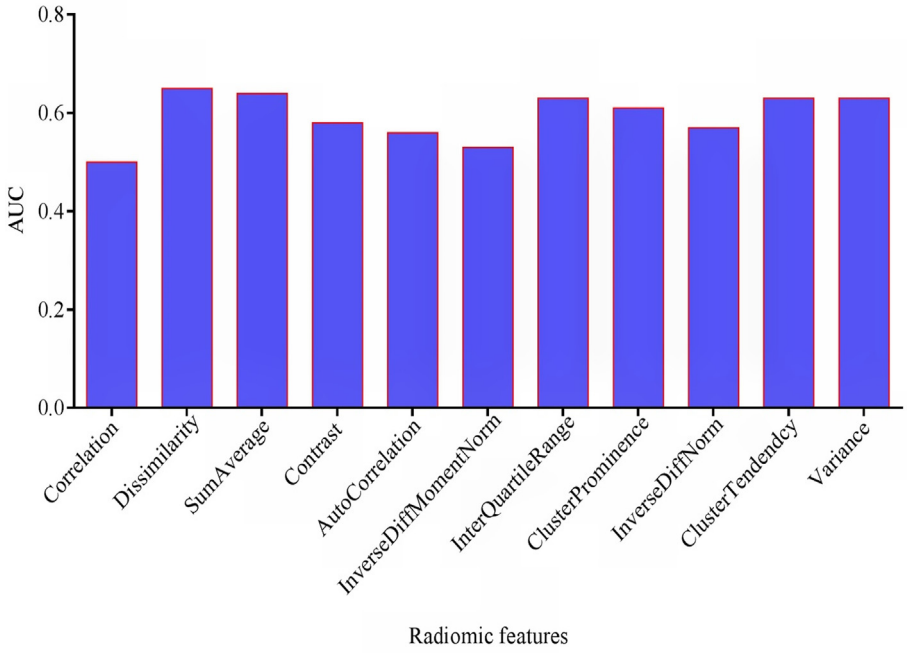

Fig. 2. Selected radiomic feature performances.

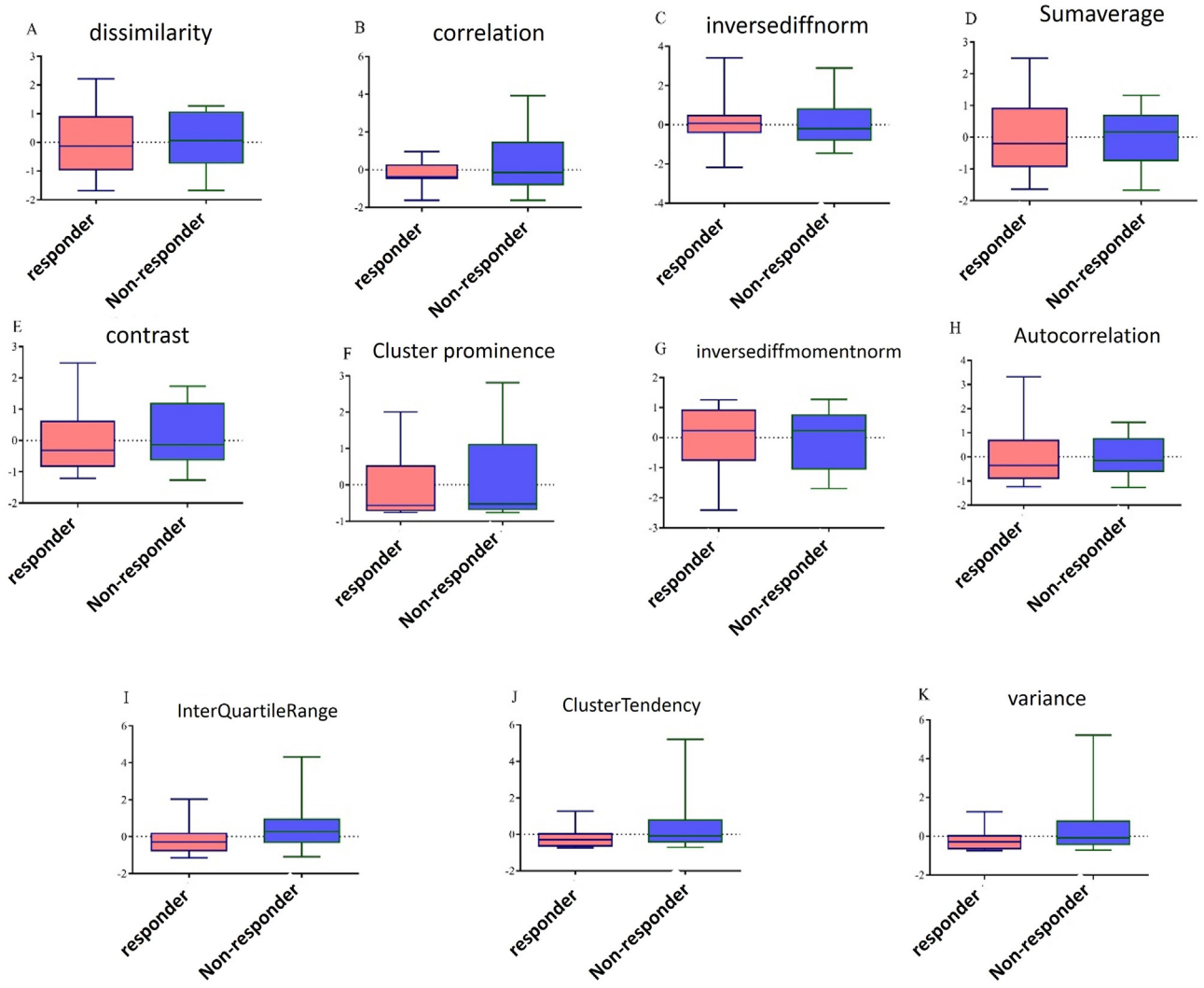

Fig. 3. Selected feature values between responder and non-responder groups.

Table 3
AUC and accuracy of different individual classifiers for T2W MR Images without and with (σ = 0.5 and σ = 1.5) preprocessing LOG filters.

Filter value	Training dataset								Validation dataset							
	SVM		KNN		NN		BN		SVM		KNN		NN		BN	
	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy
No filtration	0.68	0.73	0.73	0.74	0.71	0.74	0.75	0.81	0.72	0.71	0.73	0.70	0.70	0.76	0.74	0.78
σ = 0.5	0.76	0.76	0.87	0.80	0.85	0.76	0.81	0.76	0.71	0.75	0.82	0.71	0.71	0.70	0.80	0.80
σ = 1.5	0.89	0.79	0.91	0.83	0.88	0.80	0.92	0.83	0.88	0.78	0.83	0.78	0.87	0.79	0.88	0.81

lead to develop models with higher performances.

4. Discussion

In the present study, we developed several MRI based ensemble and individual learning method to predict rectal cancer response to nCRT. We demonstrated the efficacy of these methods in therapy response prediction. We also found that ensemble learning classification will result in more predictive models rather than individual issues. In overall, our results showed the AUC more than 70% in all ML methods. We also found that among individual ML methods, BN seems to be the best model for response prediction (AUC, 75.2%) which is consistent with previous studies [15,29].

In this study the best performance for response prediction obtained using gray level co-occurrence matrix texture features. In univariate analysis, we used feature selection algorithm to select robust features with highest performance. In this stage, 11 features selected and all of them were from gray level co-occurrence matrix texture features. Previous studies have shown the feasibility of radiomic modelling in rectal cancer. Nie et al. found that radiomic features extracted from T1/T2W, diffusion-weighted MRI (DWI) and dynamic contrast-enhanced (DCE) MR images could enhance the predictive power of pathologic response after preoperative nCRT for locally advanced rectal cancer (LARC) [42]. They used artificial neural network for classification. In another study, Meng et al., used MRI texture analysis for nCRT response prediction and found several textures such as standard Deviation (SD),

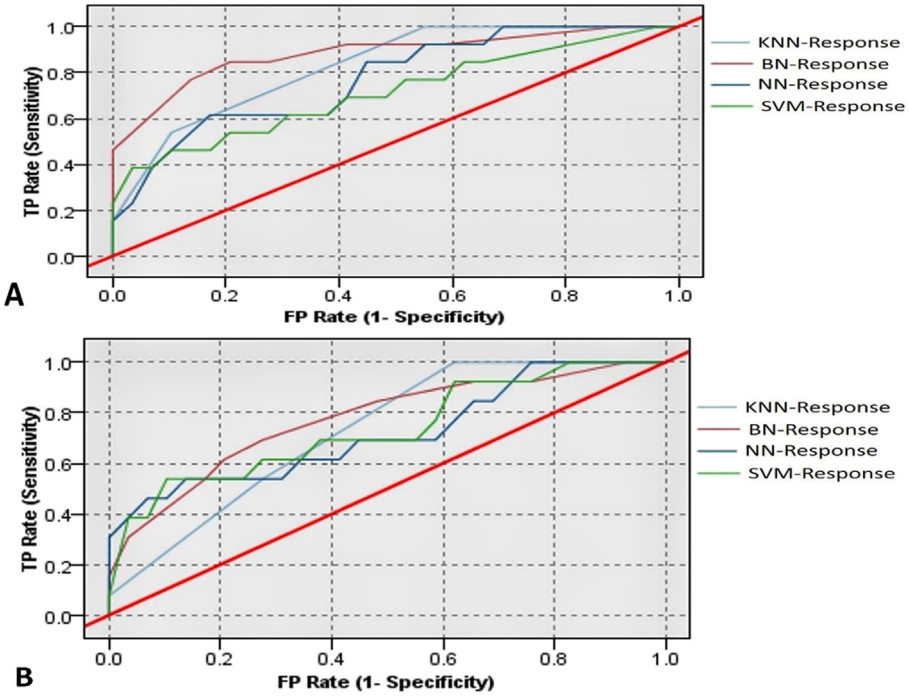

Fig. 4. AUC of KNN, BN, NN and SVM classifiers for T2WMR images without preprocessing LOG filter. A: Training dataset. B: validation dataset.

kurtosis, energy, and uniformity were statistically different between responder and non-responder groups [12].

Ensemble learning is statistical ML paradigm that combines results from multiple models to create a single and optimal prediction tool. Ensemble learning method raises the performance of individual classifiers via integration of their shared and complementary information and data fusion. In this study we used boosted classifier based on 2, 3 and 4 classifier ensemble learning method. All boosted methods improved the performance of models and the best AUC and accuracy was for 3 and 4 classifier ensemble learning methods with fine and coarse LOG filters. The application of ensemble learning in radiomics is new and is reported in some few previous studies. Lu et al. applied this

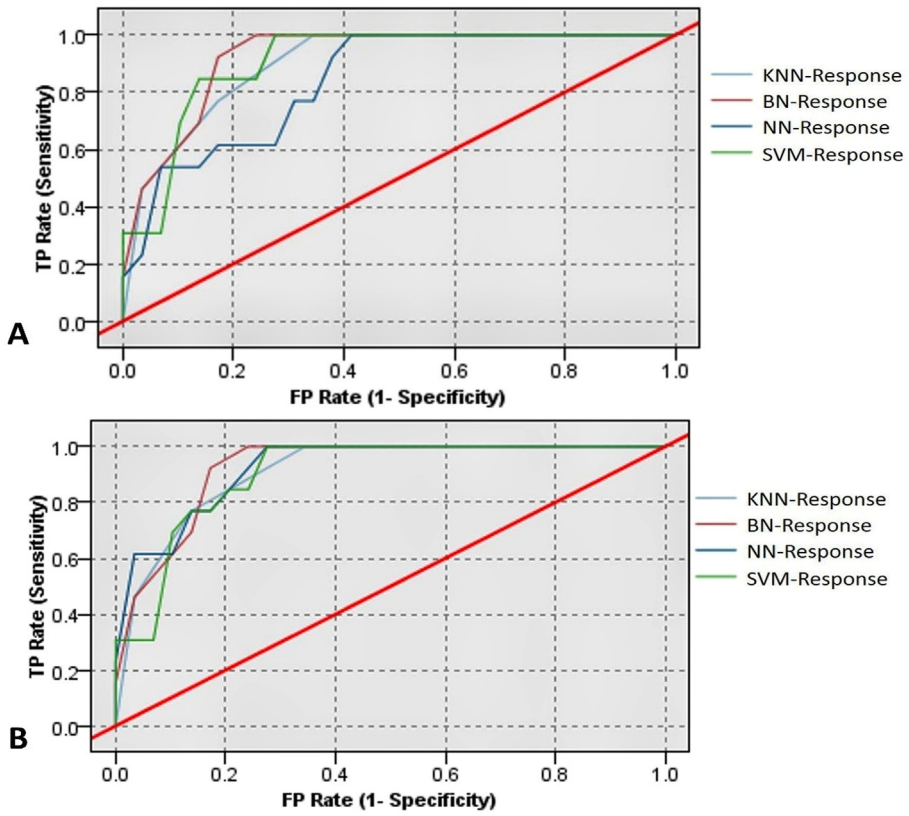

Fig. 5. AUC of KNN, BN, NN and SVM classifiers for T2WMR images with 1.5 value preprocessing LOG filter. A: Training dataset. B: validation dataset.

Table 4
AUC and accuracy of ensemble learning models for T2W MR Images without and with (σ = 0.5 and σ = 1.5) preprocessing LOG filters.

Boosted method	Training dataset						Validation dataset					
	No filtration		Fine (0.5)		Coarse (1.5)		No filtration		Fine (0.5)		Coarse (1.5)	
	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy
SVM.NN	0.70	0.76	0.78	0.78	0.91	0.81	0.73	0.73	0.77	0.77	0.90	0.78
SVM.BN	0.88	0.78	0.86	0.74	0.91	0.81	0.88	0.75	0.87	0.72	0.89	0.80
SVM.KNN	0.84	0.74	0.91	0.80	0.91	0.81	0.83	0.75	0.84	0.81	0.90	0.78
NN.BN	0.83	0.71	0.88	0.79	0.93	0.83	0.81	0.70	0.87	0.78	0.90	0.81
NN.KNN	0.83	0.74	0.89	0.85	0.93	0.86	0.81	0.72	0.89	0.82	0.91	0.79
BN.KNN	0.83	0.71	0.90	0.79	0.094	0.83	0.80	0.73	0.88	0.77	0.092	0.81
SVM.NN.BN	0.88	0.76	0.84	0.74	0.92	0.83	0.88	0.74	0.82	0.76	0.94	0.82
SVM.NN.KNN	0.89	0.73	0.90	0.76	0.90	0.83	0.86	0.75	0.89	0.77	0.88	0.83
SVM.BN.KNN	0.92	0.86	0.91	0.83	0.92	0.83	0.88	0.88	0.88	0.84	0.91	0.81
NN.BN.KNN	0.92	0.84	0.91	0.80	0.93	0.88	0.87	0.84	0.89	0.79	0.95	0.88
SVM.NN.BN.KNN	0.93	0.82	0.98	0.93	0.97	0.93	0.90	0.89	0.95	0.90	0.94	0.91

method on MR images to predict molecular subtyping in Glioma cancer patients and found that ensemble learning has better performance when single models fail. Also, Dorani et al. used ensemble learning method for detecting gene-gene interactions in colorectal cancer. They combined random forests and gradient machine learning classifiers to search for single-nucleotide polymorphisms that contribute to the disease risk through non-additive gene-gene interactions. They found that, ML algorithms have powerful abilities of modeling complex relationships between a large number of features. Also they found a novel design of an informatics framework of using two ensemble learning algorithms to search for interacting genetic factors associated with cancer witch may help us better understand the etiology of colorectal cancer [29]. Pearson et al used ensemble ML to predict treatment outcomes following an Internet intervention for depression. In this study an elastic net and random forest algorithms were used for treatment prediction in 283 patients from USA. Finally they found that, ensemble ML may be a promising statistical approach for identifying the cumulative contribution of many weak predictors to psychosocial depression treatment response [30]. In another study, ensemble learning and MR images was recruited for computed assisted diagnosis of Glioma [43–45].

As an interesting part of our study, we demonstrated the great impact of LOG filter on predictive model performance. We found that LOG filter with higher sigma value will lead to more predictive power models. The application of LOG filtering has reported by several radiomic studies. In a study by Chee et al. they demonstrated the feasibility of coarse LOG filter (sigma value, 2.5) for response prediction in advanced rectal cancer by CT texture radiomics. Also, De Cecco et al investigate the ability of filtered and unfiltered T2w MR Images for response prediction in rectal cancer patients. They found that, medium texture-scale quantified as kurtosis was significantly low in the pCR patients and Mid treatment kurtosis without filtration was significantly high in pCR patients and Finally, texture parameters can be used as response prediction biomarker in rectal cancer patients. Our results also were interesting, when we used LOG filter and ensemble learning as a combination rule to improve prediction performance [1,9,46–48].

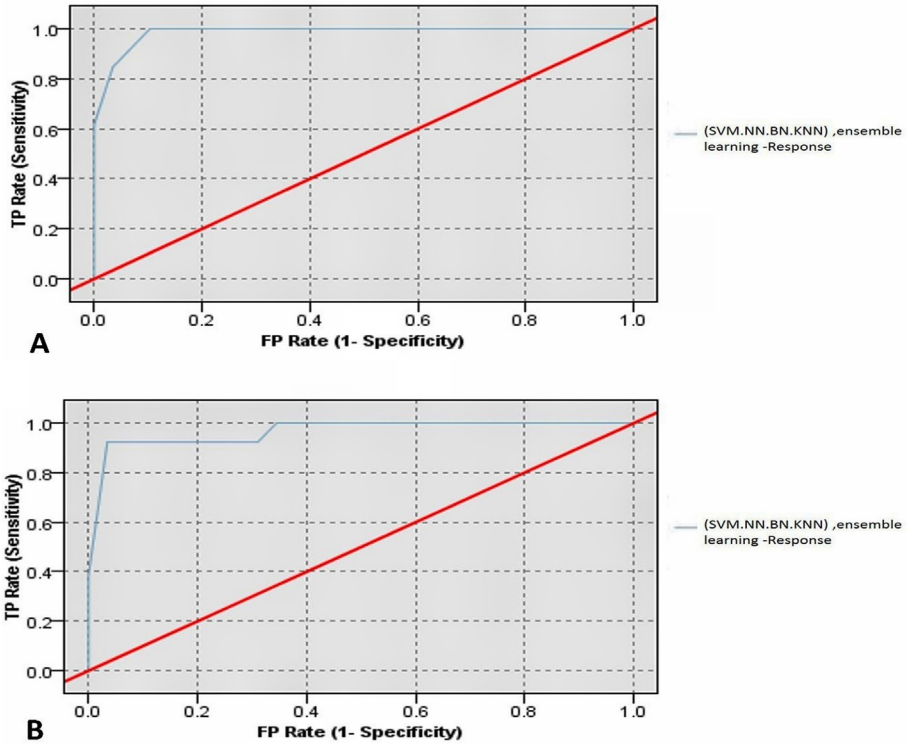

Fig. 6. AUC for 4 classifier ensemble learning method (SVM, NN, BN and KNN) with 0.5 value LOG filter in T2W MR Images. *A: Training dataset. B: validation dataset.*

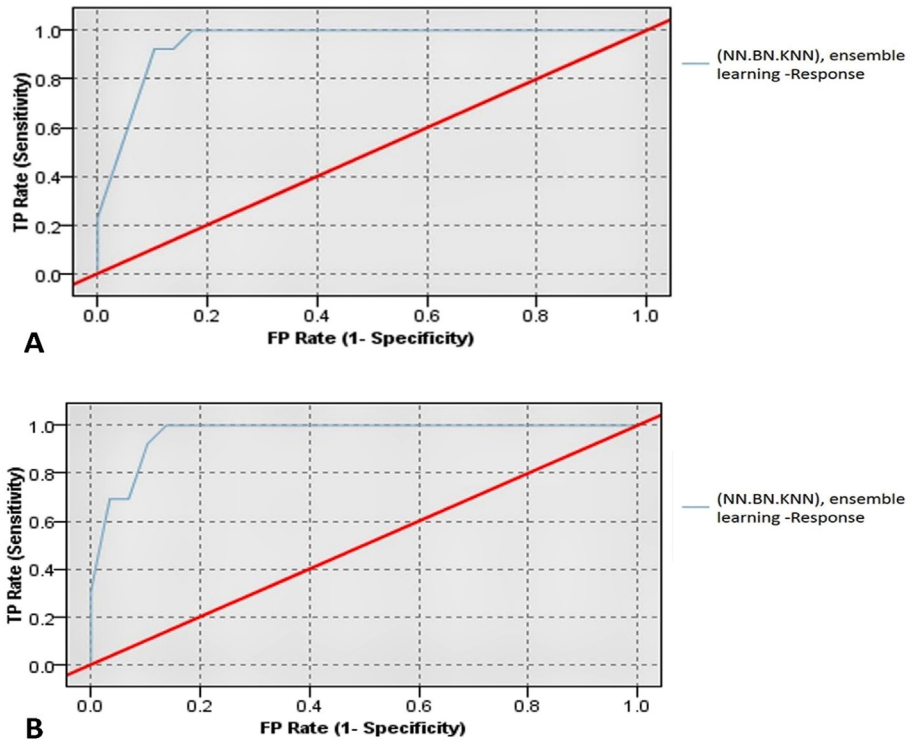

Fig. 7. AUC for 3 classifier ensemble learning method (NN, BN and KNN) with 1.5 value LOG filter in T2W MR Images. A: Training dataset. B: validation dataset.

Although our results are significant, but this study suffers from some limitation. First, the small sample size of 98 patients. Further studies with a large patient data is warranted to verify our results. Also, for multivariate analysis only 4 top features were used for classification. With 4 features and 53 patients the probability of over fitting is very low. In some part of our study, result for validation data set with 45 patient is better than tanning data set with 53 patients but in all sections difference between them in not significant and our result is approved in validation data set. Second, feature robustness and reproducibility. Based on several studies, radiomic feature are vulnerable against some challenges including image acquisition, reconstruction, segmentation and processing. Another limitation of our study is intensity Non-uniformity and with B1 + and B0 maps for each slice that are used to perform a proper spatial intensity uniformity correction can be reduced. Even when giving a dose of MRI contrast one should account for variations in spatial non-uniformity caused by scanner calibration [49–52].

5. Conclusion

In conclusion, the results presented here, showed that T2W MR imaging features can be used for response predicting in rectal cancer patients undergoing nCRT. We observed that preprocessing LOG filters and ensemble learning models can improve the prediction performances particularly with preprocessing LOG filter with higher value.

References

[1] Chee CG, et al. CT texture analysis in patients with locally advanced rectal cancer treated with neoadjuvant chemoradiotherapy: a potential imaging biomarker for treatment response and prognosis. PLoS One 2017;12(8):e0182883.
[2] Kim JC, et al. Preoperative concurrent radiotherapy with capecitabine before total mesorectal excision in locally advanced rectal cancer. Int J Rad Oncol Biol Phys 2005;63(2):346–53.
[3] Sathyakumar K, et al. Best MRI predictors of complete response to neoadjuvant chemoradiation in locally advanced rectal cancer. Br J Radiol 2016;89(1060):20150328.
[4] Sun Y-S, et al. Locally advanced rectal carcinoma treated with preoperative chemotherapy and radiation therapy: preliminary analysis of diffusion-weighted MR imaging for early detection of tumor histopathologic downstaging. Radiology 2009;254(1):170–8.
[5] Pham TT, et al. Functional MRI for quantitative treatment response prediction in locally advanced rectal cancer. Br J Radiol 2017;90(1072):20151078.
[6] Pham TT, et al. Study protocol: multi-parametric magnetic resonance imaging for therapeutic response prediction in rectal cancer. BMC Cancer 2017;17(1):465.
[7] Martens MH, et al. Tumor response to treatment: prediction and assessment. Curr Radiol Rep 2014;2(9):62.
[8] Chicklore S, et al. Quantifying tumour heterogeneity in 18 F-FDG PET/CT imaging by texture analysis. Eur J Nucl Med Mol Imaging 2013;40(1):133–40.
[9] De Cecco CN, et al. Texture analysis as imaging biomarker of tumoral response to neoadjuvant chemoradiotherapy in rectal cancer patients studied with 3-T magnetic resonance. Invest Radiol 2015;50(4):239–45.
[10] Zhang H, et al. MR texture analysis: potential imaging biomarker for predicting the chemotherapeutic response of patients with colorectal liver metastases. Abdominal Radiol 2018:1–7.
[11] Liu M, et al. Locally advanced rectal cancer: predicting non-responders to neoadjuvant chemoradiotherapy using apparent diffusion coefficient textures. Int J Colorectal Dis 2017;32(7):1009–12.
[12] Meng Y, et al. MRI texture analysis in predicting treatment response to neoadjuvant chemoradiotherapy in rectal cancer. Oncotarget 2018;9(15):11999.
[13] Dinapoli N, et al. Radiomics for rectal cancer. Transl Cancer Res 2016;5(4):424–31.
[14] Amano Y, et al. MRI texture analysis of background parenchymal enhancement of the breast. BioMed Res Int 2017;2017.
[15] Wibmer A, et al. Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. Eur Radiol 2015;25(10):2840–50.
[16] Skogen K, et al. Diagnostic performance of texture analysis on MRI in grading cerebral gliomas. Eur J Radiol 2016;85(4):824–9.
[17] Panek R, et al. Noninvasive imaging of cycling hypoxia in head and neck cancer using intrinsic susceptibility MRI. Clin Cancer Res 2017.
[18] House MJ, et al. Texture-based classification of liver fibrosis using MRI. J Magn Reson Imaging 2015;41(2):322–8.
[19] M.A.C.C. Head and N.Q.I.W. Group. Investigation of radiomic signatures for local recurrence using primary tumor texture analysis in oropharyngeal head and neck cancer patients. Sci Rep 2018:8.
[20] Abdollahi H, et al. Cochlea CT radiomics predicts chemoradiotherapy induced sensorineural hearing loss in head and neck cancer patients: a machine learning and multi-variable modelling study. Phys Med 2018;45:192–7.
[21] Liu L, et al. Application of texture analysis based on apparent diffusion coefficient maps in discriminating different stages of rectal cancer. J Magn Reson Imaging 2017;45(6):1798–808.
[22] Juntu J, et al. Machine learning study of several classifiers trained with texture analysis features to differentiate benign from malignant soft-tissue tumors in T1-MRI images. J Magn Reson Imaging 2010;31(3):680–9.
[23] Parmar C, et al. Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. Front Oncol 2015;5:272.

[24] Singh R, Shah P, Bagade J. Skin texture analysis using machine learning. Advances in Signal Processing (CASP), Conference on. IEEE; 2016.
[25] Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. Cancer Inf 2006;2. 117693510600200030.
[26] Zhao Q, et al. Ensemble learning for the detection of facial dysmorphology. Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE. IEEE; 2014.
[27] Somasundaram KS, Alli P. A machine learning ensemble classifier for early prediction of diabetic retinopathy. J MedSyst 2017;41(12):201.
[28] Papini S, et al. Ensemble machine learning prediction of posttraumatic stress disorder screening status after emergency room hospitalization. J Anxiety Disord 2018;60:35–42.
[29] Dorani F, et al. Ensemble learning for detecting gene-gene interactions in colorectal cancer. PeerJ 2018;6:e5854.
[30] Pearson R, et al. A machine learning ensemble to predict treatment outcomes following an Internet intervention for depression. Psychol Med 2018:1–12.
[31] Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. Magn Reson Imaging 2004;22(1):81–91.
[32] Bayanati H, et al. Quantitative CT texture and shape analysis: can it differentiate benign and malignant mediastinal lymph nodes in patients with primary lung cancer? Eur Radiol 2015;25(2):480–7.
[33] Fave X, et al. Delta-radiomics features for the prediction of patient outcomes in non–small cell lung cancer. Sci Rep 2017;7(1):588.
[34] Vallières M, et al. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol 2015;60(14):5471.
[35] Sutton EJ, et al. Breast cancer subtype intertumor heterogeneity: MRI-based features predict results of a genomic assay. J Magn Reson Imaging 2015;42(5):1398–406.
[36] Yang D, et al. Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma. Med Phys 2015;42(11):6725–35.
[37] Haddad P, et al. Addition of oxaliplatin to neoadjuvant radiochemotherapy in MRI-defined T3, T4 or N+ rectal cancer: a randomized clinical trial. Asia-Pac J Clin Oncol 2017;13(6):416–22.
[38] Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 2010;17(6):1471–4.
[39] Nougaret S, et al. Neoadjuvant chemotherapy evaluation by MRI volumetry in rectal cancer followed by chemoradiation and total mesorectal excision: initial experience. J Magn Reson Imaging 2013;38(3):726–32.
[40] Kotsiantis SB, Zaharakis I, Pintelas P. Supervised machine learning: a review of classification techniques. Emerging Artif Intell Appl Comput Eng 2007;160:3–24.
[41] Lin Y-Y, Liu T-L, Fuh C-S. Local ensemble kernel learning for object category recognition. Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on. IEEE; 2007.
[42] Nie K, et al. Rectal cancer: assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI. Clin Cancer Res 2016;22(21):5256–64.
[43] Lu CF, et al. Machine learning-based radiomics for molecular subtyping of gliomas. Clin Cancer Res 2018.
[44] Hu LS, et al. Multi-parametric MRI and texture analysis to visualize spatial histologic heterogeneity and tumor extent in glioblastoma. PLoS One 2015;10(11):e0141506.
[45] Emblem KE, et al. Machine learning in preoperative glioma MRI: survival associations by perfusion-based support vector machine outperforms traditional MRI. J Magn Reson Imaging 2014;40(1):47–54.
[46] Dinapoli N, et al. Magnetic resonance, vendor-independent, intensity histogram analysis predicting pathologic complete response after radiochemotherapy of rectal. Cancer Int J Radiat Oncol Biol Phys 2018.
[47] Chaddad A, et al. Prediction of survival with multi-scale radiomic analysis in glioblastoma patients. Med Biol Eng Compu 2018.
[48] Hu P, et al. Reproducibility with repeat CT in radiomics study for rectal cancer. Oncotarget 2016;7(44):71440–6.
[49] Shiri I, et al. The impact of image reconstruction settings on 18F-FDG PET radiomic features: multi-scanner phantom and patient studies. Eur Radiol 2017;27(11):4498–509.
[50] Yan J, et al. Impact of image reconstruction settings on texture features in 18F-FDG PET. J Nucl Med 2015;56(11):1667–73.
[51] Vallieres M, et al. Enhancement of multimodality texture-based prediction models via optimization of PET and MR image acquisition protocols: a proof of concept. Phys Med Biol 2017;62(22):8536–65.
[52] Watanabe H, Takaya N, Mitsumori F. Non-uniformity correction of human brain imaging at high field by RF field mapping of B1+ and B1. J Magn Reson 2011;212(2):426–30.