Heliyon 8 (2022) e09311

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Classification of retinoblastoma-1 gene mutation with machine learning-based models in bladder cancer

Okan İnce [a,*], Hülya Yıldız [a], Tanju Kisbet [a], Şükrü Mehmet Ertürk [b], Hakan Önder [a]

[a] Health Sciences University Prof. Dr. Cemil Tascioglu City Hospital, Department of Radiology, Turkey
[b] Istanbul University Istanbul Medical Faculty, Department of Radiology, Turkey

ARTICLE INFO

Keywords:
Bladder
Computer applications-detection
Diagnosis
CT

ABSTRACT

Purpose: This study aims to evaluate the potential of machine learning algorithms built with radiomics features from computed tomography urography (CTU) images that classify RB1 gene mutation status in bladder cancer.
Method: The study enrolled CTU images of 18 patients with and 54 without RB1 mutation from a public database. Image and data preprocessing were performed after data augmentation. Feature selection steps were consisted of filter and wrapper methods. Pearson's correlation analysis was the filter, and a wrapper-based sequential feature selection algorithm was the wrapper. Models with XGBoost, Random Forest (RF), and k-Nearest Neighbors (kNN) algorithms were developed. Performance metrics of the models were calculated. Models' performances were compared by using Friedman's test.
Results: 8 features were selected from 851 total extracted features. Accuracy, sensitivity, specificity, precision, recall, F1 measure and AUC were 84%, 80%, 88%, 86%, 80%, 0.83 and 0.84, for XGBoost; 72%, 80%, 65%, 67%, 80%, 0.73 and 0.72 for RF; 66%, 53%, 76%, 67%, 53%, 0.60 and 0.65 for kNN, respectively. XGBoost model had outperformed kNN model in Friedman's test ($p = 0.006$).
Conclusions: Machine learning algorithms with radiomics features from CTU images show promising results in classifying bladder cancer by RB1 mutation status non-invasively.

1. Introduction

Bladder cancer is the most common cancer in the urinary system. It is in ninth place among all cancers [1]. Clinical staging and pathological grading systems have been widely used to evaluate patients and decide the best treatment protocol. Muscular invasion is one of the most critical clinical criteria for staging [2]. Tumors with muscular invasion are considered advanced cancers prone to recurrence and metastasize. Advanced bladder cancers are recently treated with neoadjuvant chemotherapy protocols before surgery [3]. Therefore, it is crucial to determine advanced bladder cancers preoperatively.

Multiple gene mutations from different signal pathways were reported in bladder cancers [4]. However, retinoblastoma-1 (RB1) gene mutations cause alterations in the activity of the members of the retinoblastoma (RB) protein family (e.g., Rb, p107, and p130), leading to advanced-stage bladder cancers with high recurrence and low survival rates [5]. That oncogenic alteration leads to the failure of dephosphorylation of RB protein, leading to failure of arrest in the mitotic cycle. Consequently, uncontrolled cell proliferation proceeds rapidly in an aggressive process [6].

Radiomics is a rapidly emerging field in radiology that makes it possible to analyze minimal differences between pixels and their relation to each other, which are unseeable to the human eye [7]. With radiomics, quantitative texture analysis of the tumors is being done by radiologists [8]. Moreover, mining the radiomic data with machine learning algorithms allows building various models to classify high-low-grade tumors or accurately predict response to treatment.

This study evaluates the potential of machine learning-based models that classify RB1 gene mutation status, with radiomics from computed tomography urography (CTU) images in bladder cancer.

2. Material and methods

2.1. Ethics and data source

No ethical approval was needed for this study because the patients' data were obtained from a freely available public dataset.

* Corresponding author.
E-mail address: okan_ince@yahoo.com (O. İnce).

https://doi.org/10.1016/j.heliyon.2022.e09311
Received 26 November 2021; Received in revised form 19 January 2022; Accepted 20 April 2022

From the Cancer Genome Atlas – Bladder Carcinoma (TGCA-BLCA) database [9], cases with and without RB1 mutations were selected, and their imaging data were downloaded from The Cancer Imaging Archive (TCIA) for scientific purposes [10]. For better visualization of the tumor in segmentation, CTU images of the dataset were enrolled. Cases without CTU images, non-diagnostic images due to various artifacts, and patients with multiple tumors were excluded. Radiomics workflow is summarized in Figure 1. The methodological quality of the workflow based on the Radiomics Quality Score [8] is assessed in Table 1.

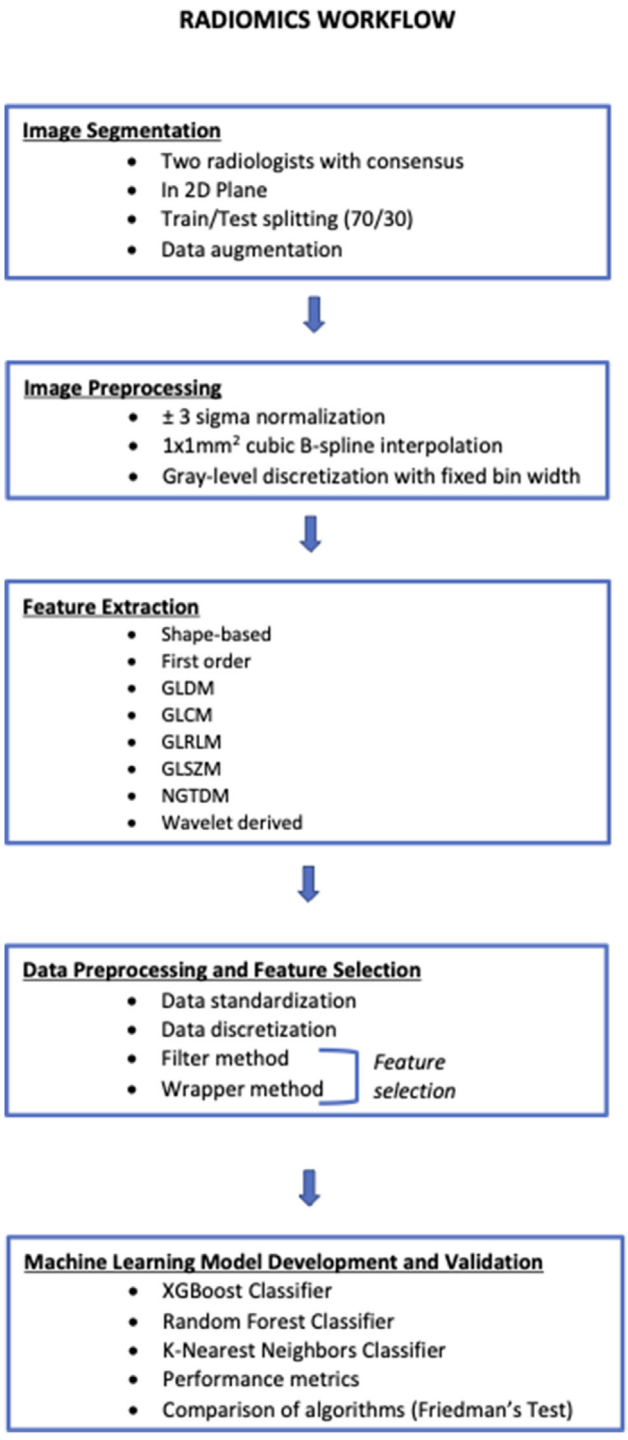

Figure 1. Radiomics workflow is presented. *GLDM: Gray-level dependence matrix, GLCM: Gray-level co-occurrence matrix, GLRLM: Gray-level run-length matrix, GLSZM: Gray-level size zone matrix, NGTDM: Neighboring gray-tone difference matrix.*

Table 1. According to the Radiomics Quality Score, the strengths of this study are as follows: (21 points).

Radiomics Quality Score
• Well documented image quality protocols. (Criteria-1, 1 point)
• Implementing a two-step feature reduction to reduce the risk overfitting. Using 5-fold cross validation technic in the wrapper method also played individual role in reducing the risk of overfitting. (Criteria-5, 3 points)
• Discussing the correlation of the radiomics models with a biological gene. (Criteria-7, 1 point)
• Reporting discrimination analysis with using cross-validation/bootstrapping (used within some of the algorithms) technic as well. (Criteria-9, 2 points)
• Reporting calibration statistics with using cross-validation/bootstrapping (used within some of the algorithms) technic as well. (Criteria-10, 2 points)
• The models' performances were validated using an independent test set that was created by splitting the main data from multiple centers. (Criteria- 12, 5 points)
• The gold standard of the study was histopathological and genomic examinations which were included in TGCA-BLCA database. The models' performances were evaluated according to the results of the gold standard. (Criteria-13, 2 points).
• Discussing the clinical implementations of models in the future. (Criteria-14, 2 points)
• Using an open-source data. Since the data was freely accessible, medical images can be downloaded from The Cancer Imaging Archive. The project IDs of the patients from each group can be found in the TCGA-BLCA project, are also shareable upon request to the corresponding author. The segmentation labels and radiomics features dataset are also shareable upon request to the corresponding author. (Criteria-16, 3 points)

2.2. CT acquisition parameters

Since the TGCA-BLCA project consisted of multicenter data, scanning parameters were mean slice thickness: 3.31mm (IQR 2.37–5), mean mAs: 155 mAs(IQR 84–216), and mean kV:120 kV(IQR 120-120). The images consisted of 11 and 14 different scanners from four vendors as RB1-mutated and non-mutated patients, respectively. To minimize the differences, the images have undergone multiple image preprocessing steps. The range between contrast delay time was 10 min-35min.

2.3. Segmentations

Two radiologists, one who has more than twenty years of experience in abdominal radiology the other one is a fourth-year radiology resident, have manually segmented the largest cross-sectional area of the tumor in CTU images in the 2D plane by using freely accessible 3D Slicer software (v.4.10.2) with consensus. To avoid high-density contrast material in the region of interest (ROI), 2mm shrinkage was applied in every segmentation label (Figure 2). For data augmentation to remove the imbalance between groups, one slice above and below of the tumor's largest cross-sectional area was also segmented in the RB1 mutation group. Segmented labels were shared in the corresponding author's github repository (https://github.com/okanince/HELIYON-OkanInce).

2.4. Image preprocessing and feature extraction

Images were normalized with the ±3 sigma technique for preprocessing steps [11]. Then, pixels were rescaled to $1 \times 1\text{mm}^2$ with cubic B-spline interpolation method, and gray-levels were discretized with a fixed bin width of 3 [12]. An optimal bin-width value was selected to keep total bins between 10-100.

Six separate feature subgroups were extracted from the original and wavelet filtered images using PyRadiomics [13], a built-in extension pack in 3D Slicer software. A detailed list of the extracted features was included in Table 2.

2.5. Data preprocessing and feature selection

Data preprocessing steps were observed for the stability and reliability of machine learning algorithms. All the data were standardized and discretized to 18 bins with a uniform bin width. The dataset was split

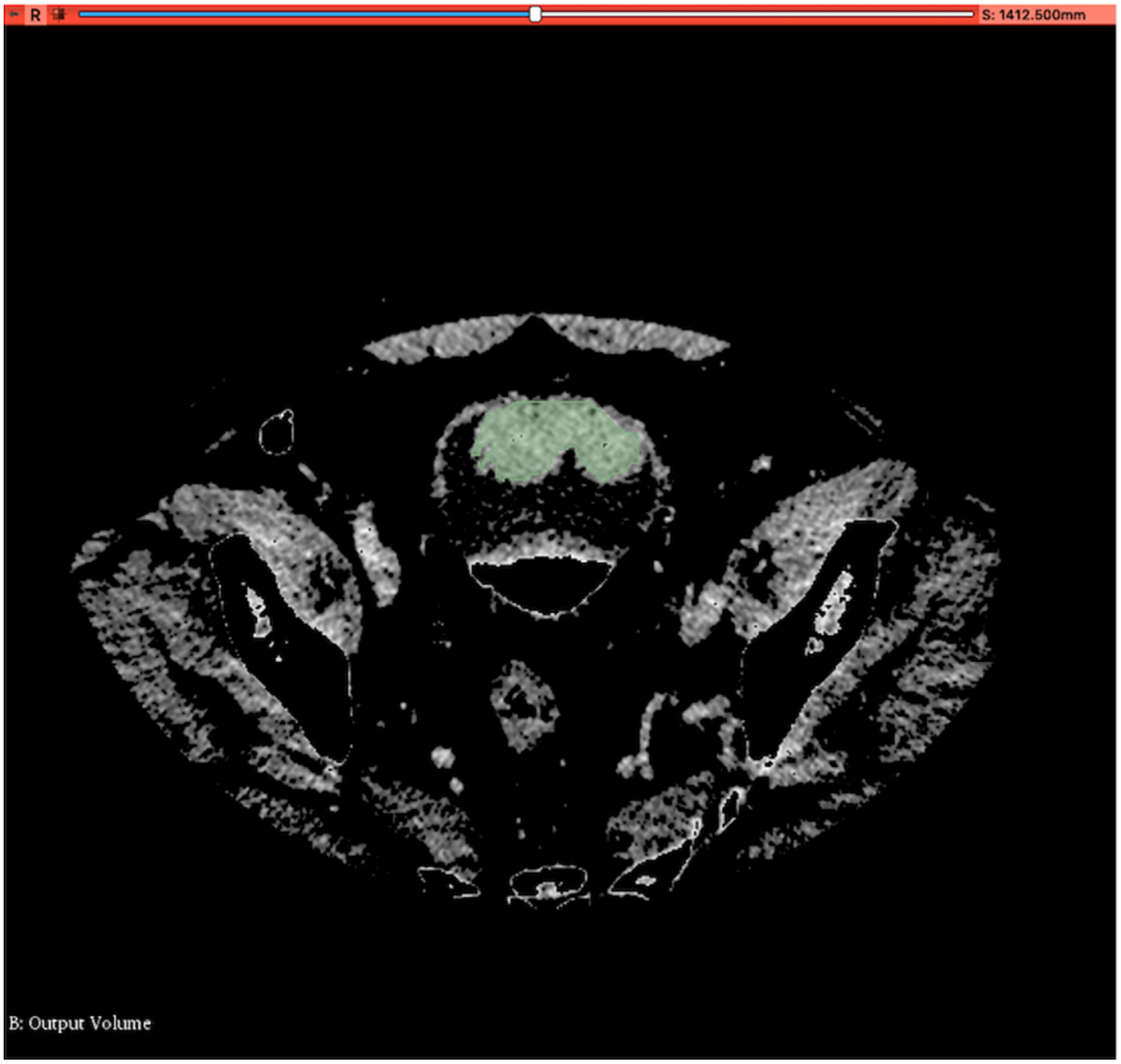

Figure 2. 2D segmentation of tumor from anterior bladder wall in the preprocessed image is presented.

to train and test sets with a ratio of 70/30. To avoid the injection of the train set to the test set, the data splitting process was performed before data augmentation. Consequently, a patient-based dataset splitting was ensured.

Feature selection is an essential step for building machine learning algorithms. Hence, exceeding the number of features in the model can cause overfitting bias [14]. To avoid that and reduce the multidimensional model input, we have followed two feature selection steps, the filter method, and the wrapper method, respectively.

In the filter method, features having high collinearity in Pearson's correlation analysis were excluded. The r threshold was selected as 0.7 [15]. The remaining features were the input of the second step. A wrapper-based sequential feature selection algorithm was built with backward propagated 5-fold cross-validation. The learning classifier was chosen as XGBoost with default hyperparameters [16, 17, 18, 19]. The wrapper method evaluates multiple models by including or excluding features in the wrapper-based sequential feature selection to reach the best feature combination. In the backward propagation, which the current study uses, the initial model is built with all features and validated by 5-fold cross-validation. Subsequent models were evaluated by excluding one of each feature respectively. The most relevant features were determined after multiple evaluations. Using cross-validation technique important features were selected utilizing only training folds. Thus, the "double-dipping" phenomenon was avoided [20]. Moreover, since the data were split to train and test sets before, as mentioned above, the test set was not used in any part of the feature selection.

2.6. Machine-learning algorithms based classification

The selected final features were used in the machine learning models. Three models were built by coding in python language (3.7.11). The first model's classifier was selected as XGBoost with hyperparameters of maximum estimators, learning rate, gamma, subsample ratio of columns by tree, and maximum depth as 200, 0.03, 0.3, 1, and 7, respectively. Second, the Random Forest (RF) classifier was selected with the hyperparameters of a number of estimators, criterion, maximum depth, minimum samples in leaf, minimum samples to split, and maximum features as 200, "entropy,"," 6, 2, 3 and "none", respectively. Third, a k-Nearest Neighbors (kNN) classifier was selected with the hyperparameters of the number of neighbors, weights, algorithm, power parameter, and distance metric as 6, "uniform,"," "auto," 3 and "Minkowski." The grid search algorithm tuned the hyperparameters of the models with 10-fold cross-validation by using the train set. The models were trained with the train set, and the test set evaluated their performances. Accuracy, sensitivity, specificity, precision, recall, F1 measure (a harmonic calculation of precision and recall), and area under the receiver operating characteristics curve (AUC) were calculated. The models' performances

Table 2. Total features extracted from original and wavelet filtered images. (GLDM: Gray-level dependence matrix, GLCM: Gray-level co-occurrence matrix, GLRLM: Gray-level run length matrix, GLSZM: Gray-level size-zone matrix, NGTDM: Neighboring gray-tone difference matrix).

Shape	First Order	GLDM	GLCM	GLRLM	GLSZM	NGTDM
Voxel Volume	Interquartile Range	Gray Level Variance	Joint Average	Short Run Low Gray Level Emphasis	Gray Level Variance	Coarseness
Maximum 3D Diameter	Skewness	High Gray Level Emphasis	Sum Average	Gray Level Variance	Zone Variance	Complexity
Mesh Volume	Uniformity	Dependence Entropy	Joint Entropy	Low Gray Level Run Emphasis	Gray Level Non Uniformity Normalized	Strength
Major Axis Length	Median	Dependence Non Uniformity	Cluster Shade	Gray Level Non Uniformity Normalized	Size Zone Non Uniformity Normalized	Contrast
Sphericity	Energy	Gray Level Non Uniformity	Maximum Probability	Run Variance	Size Zone Non Uniformity	Busyness
Least Axis Length	Robust Mean Absolute Deviation	Small Dependence Emphasis	Idmn	Gray Level Non Uniformity	Gray Level Non Uniformity	
Elongation	Mean Absolute Deviation	Small Dependence High Gray Level Emphasis	Joint Energy	Long Run Emphasis	Large Area Emphasis	
Surface Volume Ratio	Total Energy	Dependence Non Uniformity Normalized	Contrast	Short Run High Gray Level Emphasis	Small Area High Gray Level Emphasis	
Maximum 2D Diameter Slice	Maximum	Large Dependence Emphasis	Difference Entropy	Run Length Non Uniformity	Zone Percentage	
Flatness	Root Mean Squared	Large Dependence Low Gray Level Emphasis	Inverse Variance	Short Run Emphasis	Large Area Low Gray Level Emphasis	
Surface Area	90 Percentile	Dependence Variance	Difference Variance	Long Run High Gray Level Emphasis	Large Area High Gray Level Emphasis	
Minor Axis Length	Minimum	Large Dependence High Gray Level Emphasis	Idn	Run Percentage	High Gray Level Zone Emphasis	
Maximum 2D Diameter Column	Entropy	Small Dependence Low Gray Level Emphasis	Idm	Long Run Low Gray Level Emphasis	Small Area Emphasis	
Maximum 2D Diameter Row	Range Variance 10 Percentile Kurtosis Mean	Low Gray Level Emphasis	Correlation Autocorrelation Sum Entropy MCC Sum Squares Cluster Prominence Imc2 Imc1 Difference Average Id Cluster Tendency	Run Entropy High Gray Level Run Emphasis Run Length Non Uniformity Normalized	Low Gray Level Zone Emphasis Zone Entropy Small Area Low Gray Level Emphasis	

were compared with Friedman's test in SPSS v.23 (IBM Corp, Armonk, NY, USA) [21]. Post-hoc pairwise analysis was performed if a significant difference was found. The threshold of the statistical significance was set as 0.05.

3. Results

3.1. Data source

There were 78 patients with RB1 mutation and 334 without RB1 mutation of 412 in the TCGA-BLCA dataset. Images from 28 patients with RB1 mutation and 82 patients without RB1 mutation were available in TCIA. After exclusions, CTU images from 18 and 54 patients with and without RB1 mutation, were enrolled in the study. After data augmentation, 54 labeled data from both groups were in the study. Patients' demographics are described in Table 3.

3.2. Feature extraction and selection

In total, 851 features were extracted (14 shape-based, 18 first order, 14 gray-level-dependence matrix (GLDM), 24 gray-level co-occurrence matrix (GLCM), 16 gray-level run-length matrix (GLRLM), 16 gray-level size-zone-matrix (GLSZM), five neighboring gray-tone difference matrix (NGTDM) and 744 wavelet derived texture features).

In Pearson's correlation analysis, 95 non-redundant features were selected. After the wrapper-based sequential feature selection step, the selected features were reduced to 12. Details of the selected features are included in Table 4 and Figure 3.

3.3. Machine-learning algorithms based classification

XGBoost, RF, and kNN models classified RB mutation status with accuracy rates of 84%, 72%, and 66%, respectively. Sensitivity, specificity, and AUC were 80%, 88% and 0.84 for XGBoost; 80%, 65% and 0.72 for RF; 53%, 76% and 0.65 for kNN. Detailed performance metrics and confusion matrices are shown in Table 5. In Friedman's test, the XGBoost model showed better performance score than kNN model ($p = 0.006$). There was no statistically significant difference between performances of RF - kNN and RF – XGBoost models in posthoc pairwise analysis ($p = 0.54$, 0.25, respectively). The calibration plot of the models is shown in Figure 4.

4. Discussion

This study has evaluated machine learning-based models in classifying RB1 mutation presence from CTU images of bladder cancer. Three different models were built, and each of them achieved discriminative AUC scores. Achieving those scores from three different models indicates the feasibility of the machine learning models in RB1 mutation classification. According to these results, machine learning-based models would be utilizable for detecting RB1 mutation preoperatively in bladder cancer.

Table 3. Demographic characteristics of included patients by their RB1 mutation status.

	with RB1 mutation	without RB1 mutation
Age (mean ± SD)	69.1 ± 7.5	69.1 ± 11
Sex (female/male)	4/14	14/40
Scanner (Vendor/Model)	4/11	4/14

RB1: Retinoblastoma – 1 gene.

Table 4. Selected features after both filter and wrapper methods.

Feature Label	Image Type	Feature Class	Feature Name
TexF1	wavelet-HLL	GLDM	Dependence Entropy
TexF2	wavelet-LHL	FIRST ORDER	Maximum
TexF3	wavelet-LHH	GLRLM	Low Gray Level Run Emphasis
TexF4	wavelet-LHH	GLSZM	Gray Level Variance
TexF5	wavelet-LHH	GLSZM	High Gray Level Zone Emphasis
TexF6	wavelet-LLH	FIRST ORDER	Energy
TexF7	wavelet-LLH	GLSZM	Gray Level Variance
TexF8	wavelet-LLH	NGTDM	Strength
TexF9	wavelet-HHH	FIRST ORDER	Skewness
TexF10	wavelet-HHH	GLSZM	Gray Level Non Uniformity
TexF11	wavelet-LLL	GLCM	Correlation
TexF12	original	GLCM	Correlation

L: Low, H: High, GLDM: Gray-level dependence matrix, GLRLM: Gray-level run-length matrix, GLSZM: Gray-level size-zone-matrix, NGTDM: Neighboring Gray-tone difference matrix, GLCM: Gray-level co-occurrence matrix.

In the literature, radiomics with texture analysis has been studied in bladder cancer on differentiating low-grade and high-grade tumors [22, 23, 24] and tumors with the status of perivesical infiltration [25], prediction of disease-free survival [26], recurrences [27, 28], and response to the treatment [29]. Any gene mutation dedicated study could not be seen in the literature. To our knowledge, this study is the first that evaluates machine learning models' performances in the classification of bladder cancer by RB1 gene mutation status.

Histopathological examination of cystoscopic punch biopsy is the gold standard in the staging of bladder cancers. However, this procedure is invasive, requires hospitalization, and has challenges such asbiopsy specimen inadequacy and sampling error [30, 31]. Our study shows that machine learning-based models may contribute to the diagnosis preoperatively and non-invasively. Furthermore, it would be possible for patients to preoperatively benefit from rapidly developing precise medicine applications [32], which target the cyclin-dependent kinase (CDK) 4/6 pathway. There are currently promising new immunotherapy and vaccine trials targeting different molecular oncogenesis pathways in bladder cancer in the literature [33, 34, 35]. Considering these novel applications, determining the RB1 gene mutation directly at low cost and non-invasively would be pivotal in clinical practice.

Our study has several limitations. Firstly, this is a retrospective study, and all the data were obtained from previous recordings that could lead to a selection bias. Secondly, the patient population was small, and there was an imbalance between groups. Class imbalance is an issue for most machine learning algorithms that run around the assumption that all the classes are distributed equally. In such a case, models tend to predict in favor of the majority class. Data augmentation is a proven technique that can be used for preventing the imbalance between the classes or reducing the risk of overfitting where the number of samples is small [36]. Using various synthetic oversampling methods like Synthetic Minority Over-sampling (SMOTE) and Adaptive Synthetic (ADASYN) oversampling algorithm would be more timesaving [37, 38]. However, that would have caused the majority of the data to be synthetic.

For this reason, the authors performed multiple segmentations from the actual images of the minority class. Thirdly, two radiologists segmented the tumors in consensus to increase the segmentation accuracy. Therefore, interobserver analysis in segmentation could not be done. Fourthly, segmentations were done in the largest slice of the tumor in a 2D plane. Volumetric segmentation could represent the tumor better but requires an exceeding amount of time to perform [39]. Also, most of the studies in the literature are based on 2D segmentation. Several automated segmentation software is being used in the studies in the literature [40]. The authors acknowledge that they are currently in the development process of automated segmentation software for the bladder. Using such software, volumetric segmentations could have been performed more effortlessly and faster in the future. Fifthly, the TGCA-BLCA dataset consisted of various scanners from different centers with different protocols. Although it is essential since that represents clinical practice, this multiplicity may be challenging for machine learning algorithms. However, similar results from different models show the importance and necessity of image and data preprocessing steps. Lastly, we split the dataset to train and test sets that caused the models to

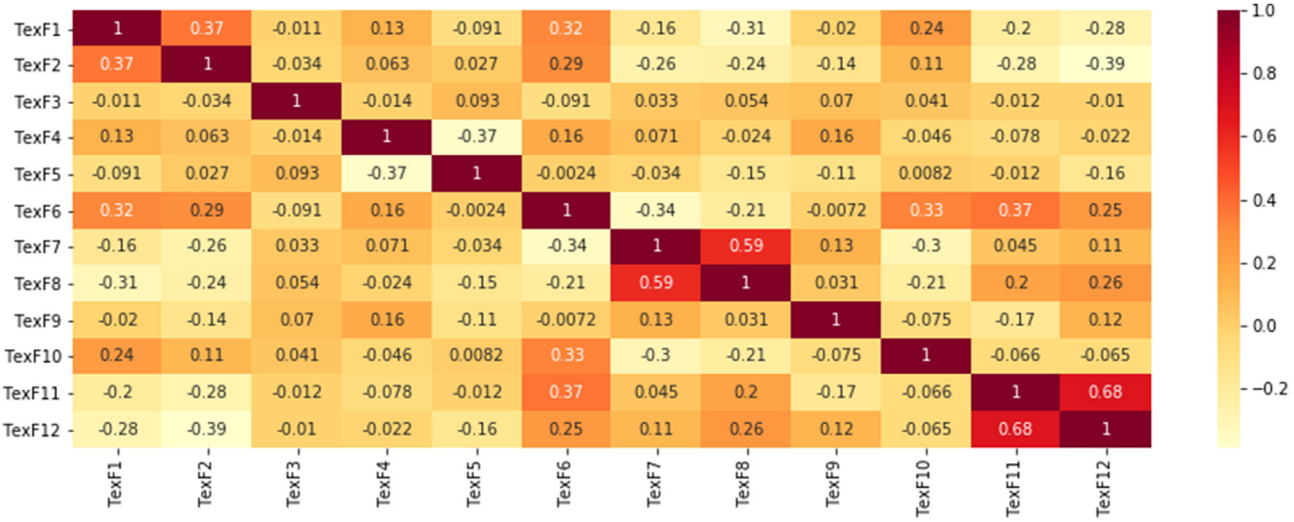

Figure 3. Correlation matrix of selected features are shown in the heatmap. None of the features are correlated to each other ($r < 0.7$).

Table 5. Performance metrics and confusion matrices of three algorithms.

	Accuracy	Sensitivity	Specificity	Precision	Recall	F1	AUC	TP/FP	TN/FN
XGBoost	84%	80%	88%	86%	80%	0.83	0.84	12/2	15/3
RF	72%	80%	65%	67%	80%	0.73	0.72	12/6	11/3
kNN	66%	53%	76%	67%	53%	0.60	0.65	8/4	13/7

RF: Random forest, kNN: k-Nearest Neighbors, AUC: Area under receiver operator characteristics curve, TP: True positive, FP: False positive, TN: True negative, FN: False-negative.

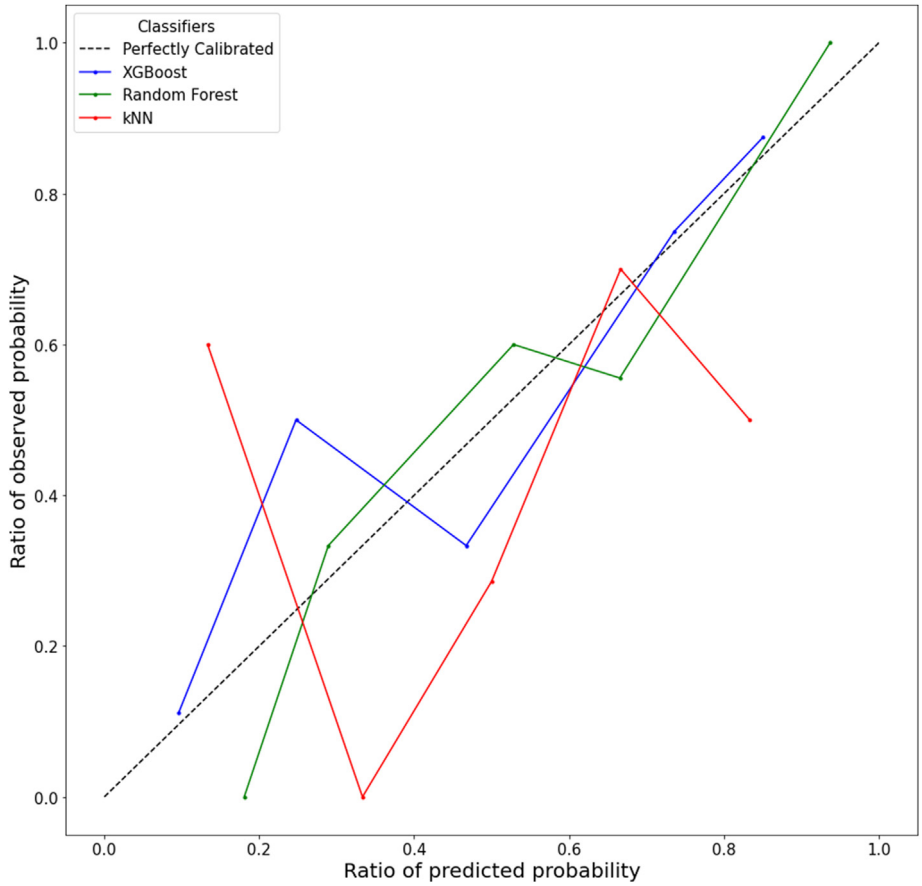

Figure 4. Calibration plot of three algorithms is shown. Each of the algorithms has better calibration, especially in higher probabilities.

test with smaller data sets. It is needed to test the models with larger datasets from external centers.

In conclusion, machine learning-based models with radiomics from CTU images show promising results in classifying bladder cancer by their RB1 gene mutation status non-invasively. Nevertheless, further studies with larger datasets are needed to test the models from external centers before their clinical use. Radiomics will have great potential when combined with artificial intelligence techniques like machine learning in the future.

Declarations

Author contribution statement

Okan İnce: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Hülya Yıldız: Conceived and designed the experiments; Performed the experiments.

Tanju Kisbet: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Şükrü Mehmet Ertürk: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Hakan Önder: Analyzed and interpreted the data.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data included in article/supplementary material/referenced in article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

[1] S. Antoni, J. Ferlay, I. Soerjomataram, A. Znaor, A. Jemal, F. Bray, Bladder cancer incidence and mortality: a global overview and recent trends, Eur. Urol. (2017) 96–108.

[2] American Joint Committee on Cancer, AJCC Cancer Staging Handbook, 2010.

[3] Advanced Bladder Cancer Meta-analysis Collaboration, Neoadjuvant chemotherapy in invasive bladder cancer: a systematic review and meta-analysis, Lancet (London, England) 361 (2003) 1927–1934.

[4] A.P. Mitra, R.J. Cote, Molecular pathogenesis and diagnostics of bladder cancer, Annu. Rev. Pathol. 4 (2009) 251–285.

[5] Q. Pan, A. Sathe, P.C. Black, P.J. Goebell, A.M. Kamat, B. Schmitz-Draeger, R. Nawroth, CDK4/6 inhibitors in cancer therapy: a novel treatement strategy for bladder cancer, Bladder Cancer 3 (2017) 79–88.

[6] A.P. Mitra, M. Birkhahn, R.J. Cote, p53 and retinoblastoma pathways in bladder cancer, World J. Urol. 25 (2007) 563–571.

[7] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, they are data, Radiology 278 (2016) 563–577.

[8] P. Lambin, R.T.H. Leijenaar, T.M. Deist, J. Peerlings, E.E.C. de Jong, J. van Timmeren, S. Sanduleanu, R.T.H.M. Larue, A.J.G. Even, A. Jochems, Y. van Wijk, H. Woodruff, J. van Soest, T. Lustberg, E. Roelofs, W. van Elmpt, A. Dekker, F.M. Mottaghy, J.E. Wildberger, S. Walsh, Radiomics: the bridge between medical imaging and personalized medicine, Nat. Rev. Clin. Oncol. 14 (2017) 749–762.

[9] S. Kirk, Y. Lee, F.R. Lucchesi, N.D. Aredes, N. Gruszauskas, J. Catto, J. Lemmerman, Radiology Data from the Cancer Genome Atlas Urothelial Bladder Carcinoma [TCGA-BLCA] Collection, Cancer Imaging Arch, 2016, pp. 96–108.

[10] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox, F. Prior, The cancer imaging archive (TCIA): maintaining and operating a public information repository, J. Digit. Imag. 26 (2013) 1045–1057.

[11] G. Collewet, M. Strzelecki, F. Mariette, Influence of MRI acquisition protocols and image intensity normalization methods on texture classification, Magn. Reson. Imaging 22 (2004) 81–91.

[12] M. Shafiq-Ul-Hassan, K. Latifi, G. Zhang, G. Ullah, R. Gillies, E. Moros, Voxel size and gray level normalization of CT radiomic features in lung cancer, Sci. Rep. 8 (2018) 1–9.

[13] J.J.M. Van Griethuysen, D.M.J. Lambregts, S. Trebeschi, M.J. Lahaye, F.C.H. Bakers, R.F.A. Vliegen, G.L. Beets, H.J.W.L. Aerts, R.G.H. Beets-Tan, Radiomics performs comparable to morphologic assessment by expert radiologists for prediction of response to neoadjuvant chemoradiotherapy on baseline staging MRI in rectal cancer, Abdom. Radiol. (2019) 632–643.

[14] B. Kocak, E.S. Durmaz, E. Ates, O. Kilickesmez, Radiomics with artificial intelligence: a practical guide for beginners, Diagnostic Interv, Radiol. 25 (2019) 485–495.

[15] C.F. Dormann, J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J.R.G. Marquéz, B. Gruber, B. Lafourcade, P.J. Leitão, T. Münkemüller, C. Mcclean, P.E. Osborne, B. Reineking, B. Schröder, A.K. Skidmore, D. Zurell, S. Lautenbach, Collinearity: a review of methods to deal with it and a simulation study evaluating their performance, Ecography (2013) 27–46.

[16] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2016, pp. 651–662.

[17] M. Nishio, M. Nishizawa, O. Sugiyama, R. Kojima, M. Yakami, T. Kuroda, K. Togashi, Computer-aided diagnosis of lung nodule using gradient tree boosting and Bayesian optimization, PLoS One (2018), e0212364.

[18] I.-J. Tsai, W.-C. Shen, C.-L. Lee, H.-D. Wang, C.-Y. Lin, Machine learning in prediction of bladder cancer on clinical laboratory data, Diagnostics 12 (2022).

[19] S. Otani, Y. Himoto, M. Nishio, K. Fujimoto, Y. Moribata, M. Yakami, Y. Kurata, J. Hamanishi, A. Ueda, S. Minamiguchi, M. Mandai, A. Kido, Radiomic machine learning for pretreatment assessment of prognostic risk factors for endometrial cancer and its effects on radiologists' decisions of deep myometrial invasion, Magn. Reson. Imaging (2022) 161–167.

[20] T.M. Ball, L.M. Squeglia, S.F. Tapert, M.P. Paulus, Double dipping in machine learning: problems and solutions, Biol. Psychiatry Cogn. Neurosci. Neuroimaging. 5 (2020) 261–263.

[21] IBM Corp. Released 2015, IBM SPSS Statistics for Macintosh, Version 23.0. Armonk, NY: IBM Corp., 2015.

[22] G. Zhang, L. Xu, L. Zhao, L. Mao, X. Li, Z. Jin, H. Sun, CT-based radiomics to predict the pathological grade of bladder cancer, Eur. Radiol. 30 (2020) 6749–6756.

[23] T. wei Fan, H. Malhi, B. Varghese, S. Cen, D. Hwang, M. Aron, N. Rajarubendra, M. Desai, V. Duddalwar, Computed tomography-based texture analysis of bladder cancer: differentiating urothelial carcinoma from micropapillary carcinoma, Abdom. Radiol. 44 (2019) 201–208.

[24] X. Zhang, X. Xu, Q. Tian, B. Li, Y. Wu, Z. Yang, Z. Liang, Y. Liu, G. Cui, H. Lu, Radiomics assessment of bladder cancer grade using texture features from diffusion-weighted imaging, J. Magn. Reson. Imag. 46 (2017) 1281–1288.

[25] C.S. Lim, S. Tirumani, C.B. van der Pol, F. Alessandrino, G.P. Sonpavde, S.G. Silverman, A.B. Shinagare, Use of quantitative T2-weighted and apparent diffusion coefficient texture features of bladder cancer and extravesical fat for local tumor staging after transurethral resection, AJR Am. J. Roentgenol. (2019) 1–10.

[26] P. Lin, D. yue Wen, L. Chen, X. Li, S. hua Li, H. biao Yan, R. quan He, G. Chen, Y. He, H. Yang, A radiogenomics signature for predicting the clinical outcome of bladder urothelial carcinoma, Eur. Radiol. 30 (2020) 547–557.

[27] G. Bartsch, A.P. Mitra, S.A. Mitra, A.A. Almal, K.E. Steven, D.G. Skinner, D.W. Fry, P.F. Lenehan, W.P. Worzel, R.J. Cote, Use of artificial intelligence and machine learning algorithms with gene expression profiling to predict recurrent nonmuscle invasive urothelial carcinoma of the bladder, J. Urol. 195 (2016) 493–498.

[28] X. Xu, H. Wang, P. Du, F. Zhang, S. Li, Z. Zhang, J. Yuan, Z. Liang, X. Zhang, Y. Guo, Y. Liu, H. Lu, A predictive nomogram for individualized recurrence stratification of bladder cancer using multiparametric MRI and clinical risk factors, J. Magn. Reson. Imag. (2019) 1893–1904.

[29] K.H. Cha, L. Hadjiiski, H.P. Chan, A.Z. Weizer, A. Alva, R.H. Cohan, E.M. Caoili, C. Paramagul, R.K. Samala, Bladder cancer treatment response assessment in CT using radiomics with deep-learning, Sci. Rep. (2017) 1–2.

[30] H.W. Herr, S.M. Donat, Quality control in transurethral resection of bladder tumours, BJU Int. 102 (2008) 1242–1246.

[31] D.E. Hansel, M.B. Amin, E. Comperat, R.J. Cote, R. Knüchel, R. Montironi, V.E. Reuter, M.S. Soloway, S.A. Umar, T.H. Van der Kwast, A contemporary update on pathology standards for bladder cancer: transurethral resection and radical cystectomy specimens, Eur. Urol. 63 (2013) 321–332.

[32] Q. Long, A.-H. Ma, H. Zhang, Z. Cao, R. Xia, T.-Y. Lin, G.P. Sonpavde, R. de Vere White, J. Guo, C.-X. Pan, Combination of cyclin-dependent kinase and immune checkpoint inhibitors for the treatment of bladder cancer, Cancer Immunol. Immunother. 69 (2020) 2305–2317.

[33] J.L. Pfail, A.B. Katims, P. Alerasool, J.P. Sfakianos, Immunotherapy in non-muscle-invasive bladder cancer: current status and future directions, World J. Urol. (2020) 1319–1329.

[34] P.A. Ott, S. Hu-Lieskovan, B. Chmielowski, R. Govindan, A. Naing, N. Bhardwaj, K. Margolin, M.M. Awad, M.D. Hellmann, J.J. Lin, T. Friedlander, M.E. Bushway, K.N. Balogh, T.E. Sciuto, V. Kohler, S.J. Turnbull, R. Besada, R.R. Curran, B. Trapp, J. Scherer, A. Poran, D. Harjanto, D. Barthelme, Y.S. Ting, J.Z. Dong, Y. Ware, Y. Huang, Z. Huang, A. Wanamaker, L.D. Cleary, M.A. Moles, K. Manson, J. Greshock, Z.S. Khondker, E. Fritsch, M.S. Rooney, M. DeMario, R.B. Gaynor, L. Srinivasan, A phase ib trial of personalized neoantigen therapy plus anti-PD-1 in patients with advanced melanoma, non-small cell lung cancer, or bladder cancer, Cell 183 (2020) 347–362.e24.

[35] W.M. Stadler, S.P. Lerner, S. Groshen, J.P. Stein, S.-R. Shi, D. Raghavan, D. Esrig, G. Steinberg, D. Wood, L. Klotz, C. Hall, D.G. Skinner, R.J. Cote, Phase III study of molecularly targeted adjuvant therapy in locally advanced urothelial cancer of the bladder based on p53 status, J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. 29 (2011) 3443–3449.

[36] B. Koçak, E.Ş. Durmaz, E. Ateş, Ö. Kılıçkesmez, Radiomics with artificial intelligence: a practical guide for beginners, Diagn. Interv. Radiol. 25 (2019) 485–495.

[37] N. V Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: synthetic minority over-sampling technique, J. Artif. Intell. Res. 16 (2002) 321–357.

[38] H. He, Y. Bai, E.A. Garcia, S. Li, ADASYN: Adaptive synthetic sampling approach for imbalanced learning, in: Proc. Int. Jt. Conf. Neural Networks, 2008.

[39] F. Ng, R. Kozarski, B. Ganeshan, V. Goh, Assessment of tumor heterogeneity by CT texture analysis: can the largest cross-sectional area be used as an alternative to whole tumor analysis? Eur. J. Radiol. 82 (2013) 342–348.

[40] S.S. Garapati, L. Hadjiiski, K.H. Cha, H.-P. Chan, E.M. Caoili, R.H. Cohan, A. Weizer, A. Alva, C. Paramagul, J. Wei, C. Zhou, Urinary bladder cancer staging in CT urography using machine learning, Med. Phys. 44 (2017) 5814–5823.