BMC Cancer

RESEARCH

Open Access

Identification of testicular cancer with T2-weighted MRI-based radiomics and automatic machine learning

Liang Wang[1†], PeiPei Zhang[2†], Yanhui Feng[1,3], Wenzhi Lv[2,4], Xiangde Min[2], Zhiyong Liu[3], Jin Li[1*] and Zhaoyan Feng[2*]

Abstract

Background Distinguishing between benign and malignant testicular lesions on clinical magnetic resonance imaging (MRI) is crucial for guiding treatment planning. However, conventional MRI-based radiomics to identify testicular cancer requires expert machine learning knowledge. This study aims to investigate the potential of utilizing automatic machine learning (AutoML) based on MRI to diagnose testicular lesions without the need for expert algorithm optimization.

Methods Retrospective preoperative MRI scans from 115 patients diagnosed with testicular disease through pathology were obtained. A total of 1781 radiomics features were extracted from each lesion on the T2-weighted images. Intraclass and interclass correlation coefficients were used to evaluate the intra-observer and interobserver agreements for each radiomics feature. We developed an AutoML method based on the tree-based pipeline optimization tool (TPOT) algorithm to construct a discriminant model. The best pipeline was determined through 100 repeated operations using a 5-fold cross-validation algorithm in TPOT. The model was evaluated for accuracy, sensitivity, and specificity using the area under the curve (AUC) value of the receiver operating characteristic (ROC) curve. Shapley Additive exPlanations were used to illustrate the optimization results.

Results Utilizing the TPOT method, 100 diagnostic models were developed to identify testicular lesions. The best model was determined based on the highest AUC in the training cohort. The prediction model yielded AUC values of 0.989 (95% confidence interval [CI]: 0.985–0.993) and 0.909 (95% CI: 0.893–0.923) in the training and testing cohorts, respectively.

Conclusions AutoML, based on the TPOT algorithm, holds potential as a noninvasive method for effectively discriminating between benign and malignant testicular lesions.

Keywords Automatic machine learning, Magnetic resonance imaging, Radiomics, Testicular cancer

†Liang Wang and PeiPei Zhang contributed equally to this work.

*Correspondence:
Jin Li
lijin@tjh.tjmu.edu.cn
Zhaoyan Feng
fzm198822@126.com
[1]Computer Center, Tongji Hospital, Tongji Medical College, Uazhong University of Science and Technology, Wuhan, China
[2]Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[3]School of Medicine and Health Management, Huazhong University of Science and Technology, Wuhan, China
[4]Britton Chance Center, MoE Key Laboratory for Biomedical Photonics, Wuhan National Laboratory for Optoelectronics-Huazhong University of Science and Technology, Wuhan, China

Background

Testicular tumors are relatively uncommon in humans, constituting approximately 1% of all male malignancies [1]. Despite its relatively low overall incidence, testicular cancer is the most common solid tumor in men aged 20 to 34 years [2].

Due to the potential for tumor metastasis and infiltration, biopsy has not been widely employed in patients with suspected malignant tumors [3, 4]. Orchiectomy is the first choice for patients with suspected tumors, with diagnosis relying mainly on postoperative pathological examinations [5]. Some testicular masses are clinically misdiagnosed as malignant, leading to unnecessary radical orchiectomy, and some patients with benign masses may opt for surgery due to concerns about potential malignancy [6]. In such patients, testis-sparing surgery or clinical examination rather than radical orchiectomy is the ideal method [7, 8]. Therefore, accurate preoperative identification of benign tumors and cancer is crucial to minimize unnecessary tumor resections and surgeries, significantly impacting the clinical curative outcome in patients [9].

Conventional ultrasonography (US), such as grayscale and color Doppler flow imaging, is the most widely used method for scrotal examination. Despite reports suggesting the utility of US as an important indicator for distinguishing between benign and malignant testicular masses, varying conclusions exist across different studies [10–13]. In instances where traditional US methods have limitations, multiparameter ultrasound (mpUS) and magnetic resonance imaging (MRI) are effective auxiliary examination methods for tumor diagnosis [7]. Due to its robust soft tissue imaging capability and versatility, MRI is increasingly utilized as an adjunct method to address uncertain or inadequate scrotal US [14]. MRI can reveal the shape, location, and tissue structure (such as fat, blood, granulation tissue, and fibrosis) of lesions and provide an important basis for formulating surgical treatment plans and improving the overall medical quality of patient care.

Radiomics, as an effective quantitative image evaluation approach, aims to extract quantitative features such as shape features, first-order features, texture features, and filter features from images to characterize lesion features effectively [15]. Radiomics involves high-throughput quantitative analysis of image data using modern imaging technologies and has found extensive application in the diagnosis and treatment of neck cancer, breast cancer, and brain glioma [16–18]. While radiomics is predominantly a research tool, its translational potential is evident in ongoing clinical trials for breast cancer risk stratification [19] and glioma prognosis prediction [20].

However, there have been few reports on the use of radiomics to evaluate testicular lesions. Previous studies have indicated that radiomics models perform better than the statistical results of conventional methods in diagnosing testicular lesions [21, 22].

Nevertheless, the development of a conventional radiomics model involves preprocessing method, feature selection, classifiers, and optimal parameters, all of which require manual expert optimization [23]. Automatic machine learning (AutoML) is a new technology that has gained popularity in recent years and provides a solution to this issue. The tree-based pipeline optimization tool (TPOT) is an AutoML framework facilitating automatic optimization [24]. The fundamental principle of TPOT is to automatically optimize feature preprocessing, feature selection, and classifiers using genetic algorithms without manual design. Generations are generated through heredity or mutation based on the predicted performance and reliability of each generation to maximize predictive model performance and reduce complexity.

While TPOT has been applied to diagnose several diseases [25–27], its use for the treatment of testicular diseases has not been reported. In this study, our primary goal was to develop an AutoML (TPOT) model based on MRI to investigate the potential of differentiating between benign and malignant testicular masses without manual expert optimization.

Methods

Patients

This retrospective study underwent ethical review of the Tongji Hospital, Tongji Medical College, Huazhong University of Science & Technology. In total, all patients ($n = 424$) who underwent clinical testicular MRI and pathological examination between February 2014 and April 2023 were enrolled in this study. Exclusion criteria were applied as illustrated in Fig. 1: (1) MRI indicating no significant testicular lesion ($n = 198$); (2) absence of a testicular lesion confirmed by pathology ($n = 21$); (3) patients who had undergone biopsies, surgery, or treatment before MRI examination ($n = 81$); or (4) poor image quality ($n = 9$). After exclusion, 123 lesions in 115 patients were included in the study (Table 1). There were 49 benign and 74 malignant tumors. Testicular lesions were diagnosed through histopathological examination of testicular tissue sections obtained during surgical procedures or biopsy.

MRI protocol

All cases were imaged using a 3.0-T MRI scanner (MAGNETOM Skyra, Siemens Healthcare) using an 18-element matrix and a 32-channel spine coil. The detailed MRI parameters can be found in Supplemental Table 1. In view of the small sample size, diffusion-weighted imaging and dynamic contrast-enhanced MRI images were not used in the study.

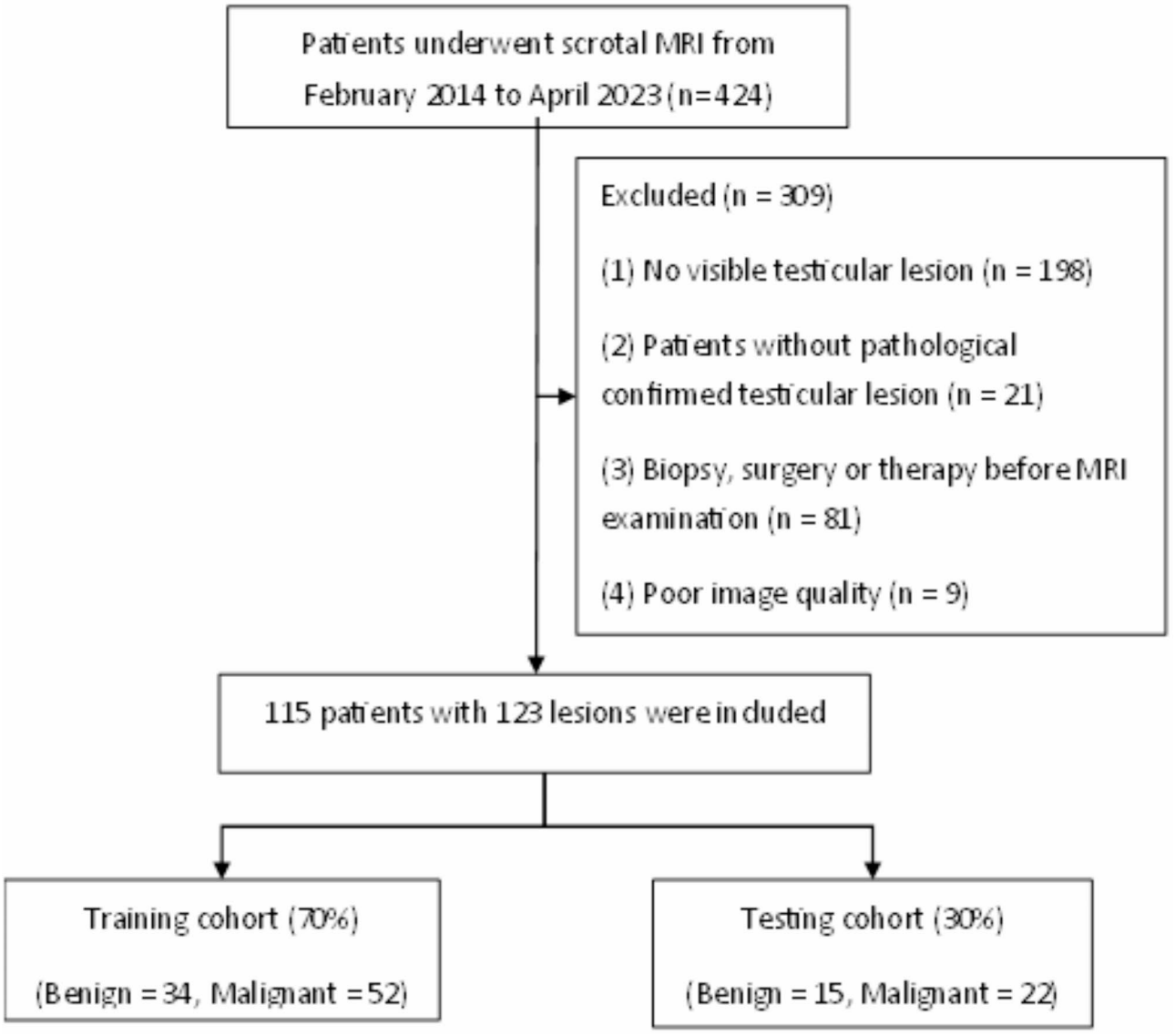

Fig. 1 Patient inclusion and exclusion criteria. A chart illustrating the process of patient inclusion and exclusion until the final selection of 115 patients with 123 lesions. MRI, magnetic resonance imaging

MRI segmentation and feature extraction

Axial T2-weighted images (T2WI) were utilized for manual segmentation of lesions and subsequent analyses. Three-dimensional volume of interest (VOI) manual segmentation was performed using ITK-SNAP (version 3.4.0). The VOI was delineated slice-by-slice along the contour of the mass on axial T2WI by a radiologist (Reader 1, with four years of experience) blinded to the histopathological information.

Radiomics features were extracted using the open-source IBSI-compliant PyRadiomics package (version 3.0.1) in Python [28]. Before feature extraction, the T2WI were resampled with the same voxel size of $0.46875 \times 0.46875 \times 3$ mm (x, y, z axis), with gray-level discretization set to 20 bin counts. Pixel intensities were normalized using the z-sore method by PyRadiomics. The largest connected component analysis was applied to remove small disconnected regions in the VOI, as it would lead to errors in feature extraction. A total of 1781 features were extracted from each VOI in the original and differently processed T2WI (Supplemental Table 2).

Inter-observer and intra-observer reproducibility of features

30% of patients (34 patients) were randomly selected to assess the reproducibility of radiomics features. Reader 1 manually segmented these 34 patients again one month later to evaluate intra-observer reproducibility. Another radiologist (Reader 2, with 5 years of experience) delineated the testicular lesions in these 34 patients to assess

Table 1 The distribution and characteristics of the included cohort

Benign			**Malignant**		
Age (years)	36.55 ± 2.63		**Age (years)**	37.55 ± 1.78	
Pathology	**Patient Number**	**Lesion Number**	**Pathology**	**Patient Number**	**Lesion Number**
Epidermoid cyst	12	12	Seminoma	27	27
Sex cord-mesenchymal tumor	8	8	Embryonal carcinoma	9	9
Infection	16	19	Mixed germinoma	10	10
Cyst	2	2	Lymphoma	11	14
Infarction	1	1	Teratoma	6	6
Contusion	3	3	Testicular metastases	3	4
Prepubertal teratoma	1	1	Neuroendocrine carcinoma	1	1
Prepubertal teratoma and epidermoid cyst	1	1	Yolk sac tumor	2	2
Testicular adrenal rest Tumor	1	2	Borderline serous tumor	1	1
Summation	45	49		70	74

interobserver reproducibility. The intraclass and interclass correlation coefficients (ICCs) were calculated between the values of radiomics features from the two Readers. Features with an ICC value of at least 0.8 were regarded as reliable and incorporated into the subsequent analysis.

TPOT analysis

In this study, TPOT was used as an AutoML algorithm to automatically generate optimal machine learning pipelines for the identification of testicular lesions (Fig. 2). A total of 123 testicular lesions were randomly divided into training and testing cohorts at a ratio of 7:3. Patient-wise splitting was applied to ensure all lesions from a single subject were assigned exclusively to either the training or testing cohort. The training cohort was utilized to develop an optimal model, which was evaluated in the testing cohort. The TPOT automatically optimized the machine learning pipeline for maximum predictive performance using tree-based genetic programming. The performance of each pipeline was evaluated using equilibrium accuracy with 5-fold cross-validation. Equilibrium accuracy is a measurement method for unbalanced datasets that prevents excessive performance by calculating the weighted average accuracy for each category. The optimal pipeline obtained through training utilized all training data to create the final model.

The hyperparameters for TPOT model training were set as follows: the generation was 50, the population size was 50, and a 5-fold cross-validation was utilized (Table 2). Acknowledging the inherent randomness of the evolutionary algorithm, which leads to model instability, 100 random seeds were used to build prediction models using TPOT during the development process. Among the 100 prediction models in the training cohort, the model with the highest area under the curve (AUC) was selected as the best. Finally, the Shapley Additive exPlanations (SHAP) values were utilized to evaluate the impact of each feature on the optimization model and visually display how these values affected the model output for a single case.

Statistical analysis

In this study, statistical analysis were performed utilizing the Python (version 3.7.13), with the packages used as illustrated in Supplemental Table 3. Continuous variables are presented as mean ± standard deviation. The independent t-test and Mann–Whitney U test were used to analyze variables with normal and non-normal distributions, respectively, with the Shapiro–Wilk test used to evaluate normality. The chi-square test or Fisher's exact test was utilized to compare differences in the classification variables. The diagnostic performance of the predictive model was evaluated by receiver operating characteristic (ROC) curves, and the maximum Youden index was used as the optimal cutoff value to calculate the accuracy, sensitivity, and specificity. A p-value < 0.05 on both sides indicated statistical significance. To enhance transparency, a the CLAIM (Checklist for Artificial Intelligence in Medical Imaging) [29, 30] is provided in Supplemental Table 4.

Results

Patients

In total, 115 patients with 123 testicular masses (mean age, 36.24 years ± 18 [SD]) were included in this study. The testicular masses were pathologically confirmed as 49 benign lesions (benign group) and 74 malignant lesions (malignant group), as listed in Table 1. No statistically significant difference in age was observed between the two groups ($p = 0.753$). All lesions were randomly allocated into two cohorts: a training cohort consisting of 86 lesions and a testing cohort comprising 37 lesions, following a split ratio of 7:3.

Radiomics feature extraction

A total of 1781 features were extracted from each lesion using various types of images (Supplemental Table 1), including 14 shape features, 18 first-order intensity

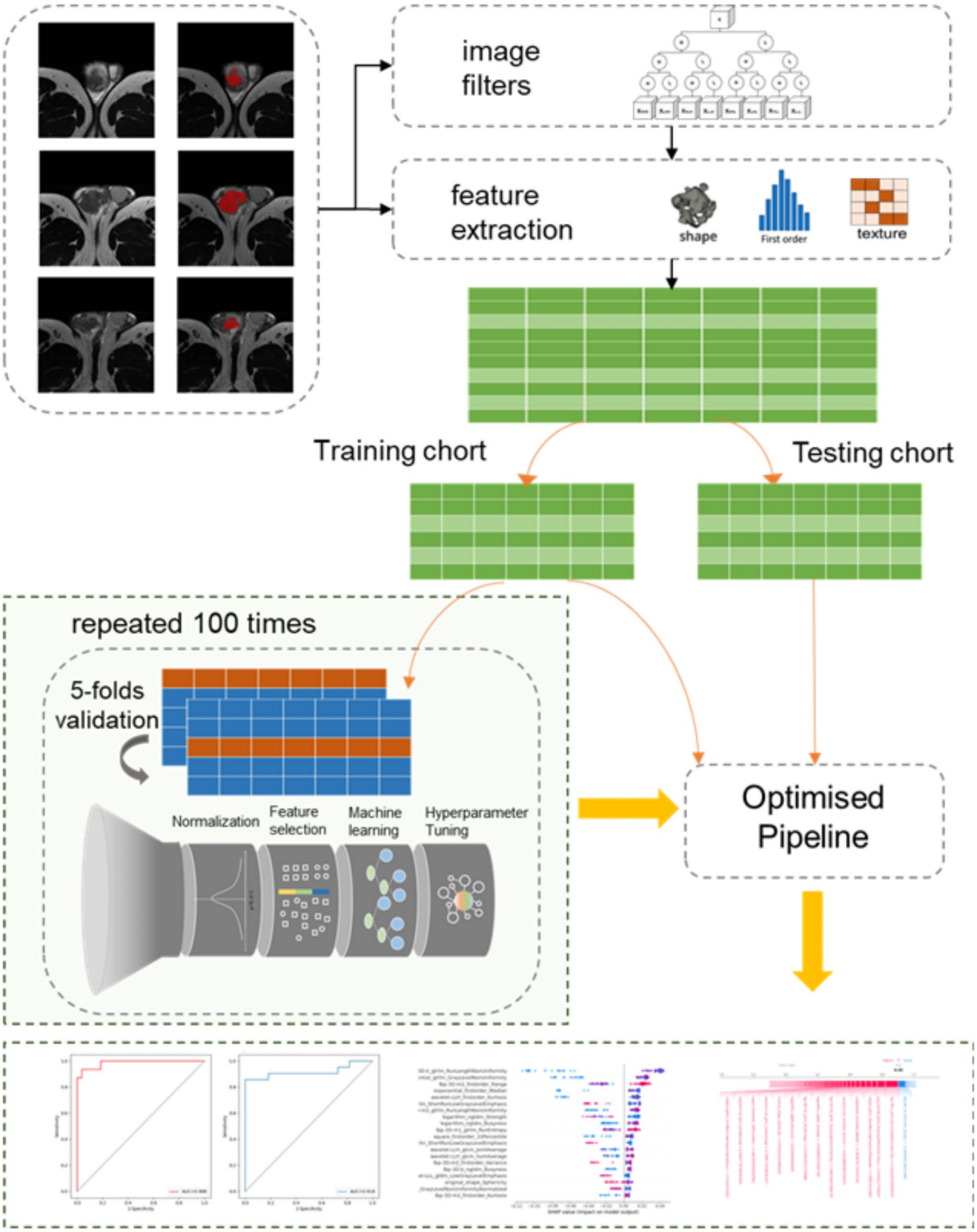

Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Study design illustration. Radiomics features are extracted from different images and processed by several filtering methods. The dataset was divided into two independent sets: a training cohort and an evaluation cohort. The training cohort was subjected to TPOT, which combined normalization, feature selection, classifier, and hyperparameter tuning, and the performance was evaluated. Five-fold cross-validation was carried out in each generation to select the optimal genetic model for that generation, followed by replacement or mutation, and then passed on to the offspring. The TPOT process was repeated 100 times, and based on this, the improved pipeline was retrained on a training cohort and tested on a test cohort. Finally, the analysis of consequences was conducted using different methods. TPOT, Tree-based pipeline optimization tool

statistical features, 75 texture features, 744 wavelet features, 186 Laplacian of Gaussian features, 93 square features, 93 square root features, 93 logarithmic features, 93 exponential features, 93 magnitude of local gradient features, and 279 local binary pattern features. Out of these, 1182 features were considered highly reproducible, with both intra- and inter-observer ICCs exceeding 0.8.

Classification model derived from TPOT

One hundred different models were constructed based on the training cohort to identify the optimal model for diagnosing testicular diseases (Fig. 2). The AUC values of the models generated by the TPOT under 100 different random seeds are displayed in Supplemental Fig. 1. The lowest AUC was 0.900, the overall performance was 0.977 ± 0.020, and the generated pipeline with the highest AUC was selected as the optimal model. The optimal pipeline comprises two operators, namely, recursive feature elimination and random forest, with the corresponding hyperparameters listed in Supplemental Table 5. Subsequently, ROC and AUC values were used to evaluate the performance of the optimal model. The ROC curve and AUC of the classification results for benign and malignant lesions are illustrated in Fig. 3a. In the testing cohort, the model achieved good discriminability with an accuracy of 0.784 for diagnosing testicular lesions (AUC = 0.909) (Fig. 3a; Table 3). The beeswarm plot in Fig. 3b and the chord plot in Fig. 3c show the classification results for benign and malignant lesions using our model.

After stability analysis and pipeline processing of the features extracted from different types of images, the probability of each lesion being benign or malignant was obtained using the prediction model. Subsequently, the importance of the features was interpretatively analyzed using SHAP. The global importance of each feature was considered the average absolute value of the feature for all given samples (Supplemental Fig. 2).

Figures 5 and 4 show the features ranked highest according to the SHAP values and waterfall plots of decisions for one malignant lesion and one benign lesion. The value at the intersection of red and blue represents the probability of the patient suffering from malignant tumors. The base value is the average probability of malignant tumors obtained from the training data. The red and blue arrows below the axes correspond to the factors. The wider the arrows, the more important they are. The red color represents factors that increase the risk of malignant tumors, while the blue color represents factors that decrease the risk. The figure only shows a few of the most important factors, and the number of factors shown is controlled by a threshold.

Discussion

Accurate preoperative diagnosis is crucial for devising a rational surgical plan, preventing unnecessary orchiectomy, and improving patient care [31]. AutoML, based on radiomics features, demonstrated strong performance in MRI image diagnosis without manual optimization design, effectively distinguishing between benign and malignant testicular lesions. In the testing cohort, this approach achieved an AUC value of 0.909. The output of our optimal model demonstrated the potential of TPOT as an automatic and robust approach for the preoperative characterization of benign and malignant testicular lesions.

Conventional radiomics methods have been employed to distinguish between benign and malignant lesions [32–35]. For instance, Xu et al. evaluated the application of a radiomics signature in distinguishing benign and malignant prostate lesions based on MRI and achieved an AUC of 0.92 in the validation cohort [36]. Based on the analysis of 44 testicular lesions in 42 patients, Feliciani et al. [37] utilized T2WI to evaluate the diagnosis of testicular germinoma and seminoma and approximately 500 features were extracted and selected using the Least Absolute Shrinkage and Selection Operator (Lasso) to build classifiers. The accuracies of the models with three and four features were 89% and 86%, respectively. In a previous study, Zhang et al. [38] established a T2WI-based radiomics signature for seminomas and non-seminomas. While deep learning models demonstrate remarkable capabilities in automated feature extraction from raw data, their performance is significantly constrained by dataset scale. When applied to limited datasets, these models frequently encounter overfitting issues due to their inherent architectural complexity. This limitation becomes particularly relevant in the medical imaging studies referenced, where the adoption of radiomics-based machine learning approaches offers distinct advantages through dimensionality reduction and simplified model architectures, thereby effectively mitigating overfitting risks.

Table 2 Genetic programming parameters for the TPOT

Parameter	Value
Population size	50
Generations	50
Selection	10% elitism, rest 3-way tournament (2-way parsimony)
Per-individual mutation rate	90%
Fitness function	Equilibrium accuracy
Per-individual crossover rate	5%
Replicate	30

Traditional machine learning methodologies present additional implementation challenges, requiring substantial domain expertise for manual optimization of critical components including feature engineering, algorithm selection, and hyperparameter configuration. These human-dependent processes not only introduce subjective biases but also create significant barriers to knowledge transfer across different clinical scenarios. By contrast, AutoML frameworks establish systematic pipelines that integrate automated feature processing, model selection, and parameter optimization. This algorithmic approach enhances reproducibility while maintaining model efficacy, substantially lowering the technical threshold for clinical implementation and facilitating broader adoption of AI-assisted diagnostic solutions.TPOT can be applied to genetic programming to automatically generate optimized machine learning pipelines and has been used in other radiomics research [39]. Further, some studies have shown that random searches are more likely to produce optimal models [40]. TPOT, based on genetic programming, continually updates and iterates populations in the process of "evolution" and selects models with improved efficiency and lower complexity based on balanced accuracy. The optimal model was produced in a larger population by generating the initial population multiple times, and the generated pipelines achieved considerable classification efficiency on the training set. Consequently, the TPOT-based model established in this study can noninvasively and quantitatively discriminate tumors, providing a foundation for patients to make informed treatment plans. Additionally, the interpretability of the model was enhanced by SHAP values, illustrating how features influenced the predicted results.

The higher proportion of benign lesions in our cohort (49/123, 39.8%) compared to general orchiectomy cohorts primarily reflects differences in patient inclusion criteria. Unlike studies restricted to radical orchiectomy, which inherently favor malignant diagnoses, our cohort incorporated patients undergoing testis-sparing surgery (TSS) and diagnostic biopsy for lesions with ambiguous imaging findings. These conservative procedures are selectively applied to cases with lower clinical suspicion of malignancy, thereby increasing the representation of benign pathologies. This aligns with prior studies. For instance, Isidori et al. (2014) reported a comparable benign rate (42/86, 49%) in a cohort evaluated via contrast-enhanced ultrasound, despite methodological differences in imaging modality and diagnostic workflow [10]. Such parallels suggest that higher benign proportions are inherent to studies evaluating diagnostically challenging lesions, regardless of the imaging technique employed. While this inclusion strategy enhances the clinical relevance of our model for ambiguous cases, it may limit direct comparability to cohorts dominated by radical orchiectomy.

This study has several limitations. First, although we selected more patients than most previous studies, our sample size may still be considered small due to the rarity of testicular cancer. Second, this study was retrospective, thereby potentially introducing certain selection biases. Third, manual VOI segmentation is subject to time consumption and objective instability.

In the absence of sufficient external validation data, we utilized various methods to ensure the reproducibility of the model. First, multiple readers were involved to achieve the division of the range of interest in the images. Second, the ICCs were computed to ensure the robustness of the selected features. Third, a standardized process of feature normalization, feature selection, parameter optimization, and classification based on the TPOT was implemented to realize feature processing and model recognition. Further, the inner cross-validation of each pipeline was performed. However, our results are limited to the respondents. Therefore, we intend to further test the universality of the model through external validation.

Conclusions

In summary, AutoML, based on the TPOT algorithm, can effectively establish a relationship between MRI features and testicular disease by automatically optimizing the machine learning process. This technique proves valuable in discriminating between benign and malignant testicular lesions without the subjective biases associated with the manual design model-building process.

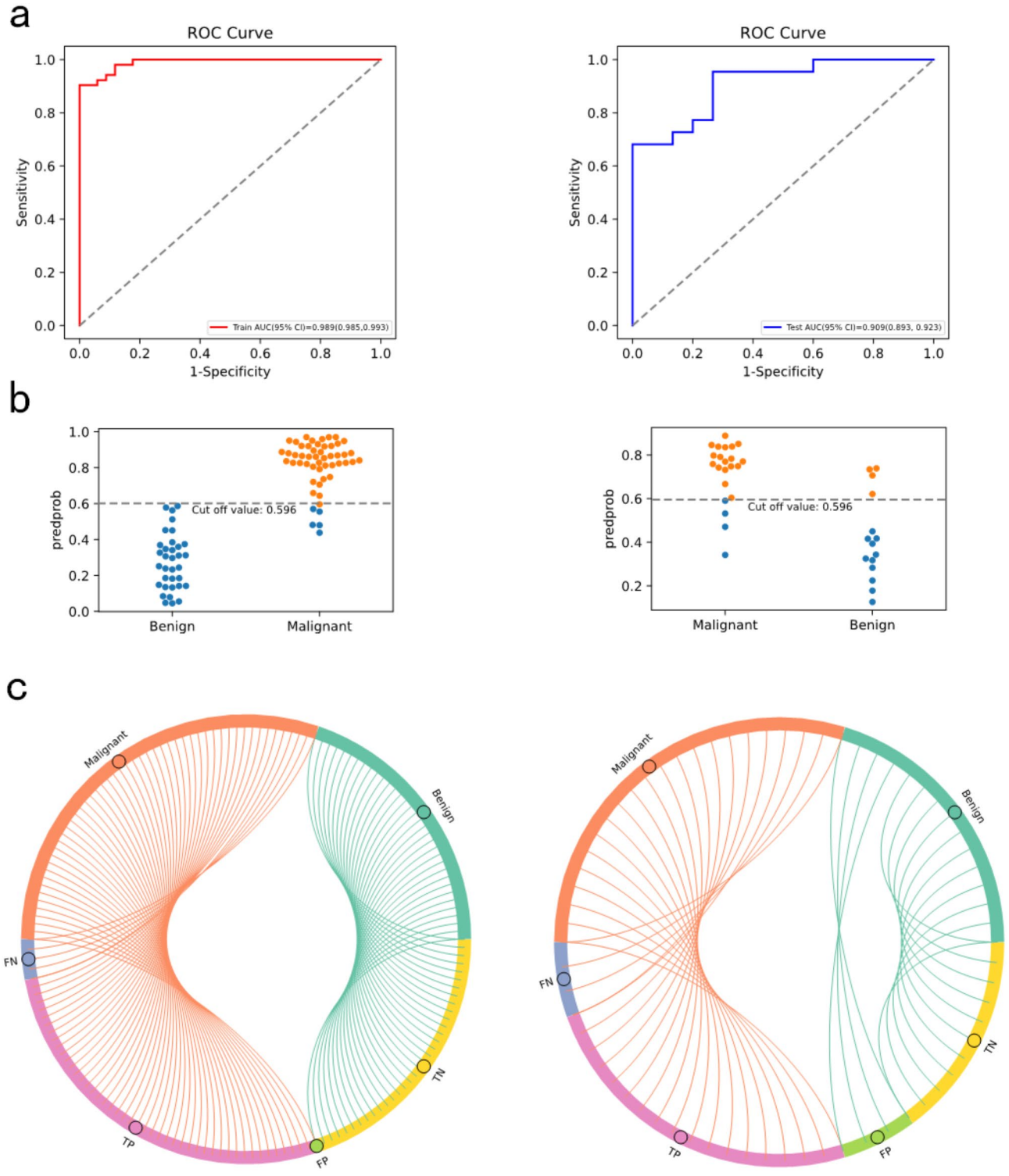

Fig. 3 ROC curves and predictions of optimal pipeline in training cohort (left) and testing cohort (right). (**a**) The ROC curve of the training and testing cohorts. (**b**) The beeswarm plots of predictions in the training and testing cohorts. (**c**) The chord diagrams of predictions in the training and testing cohorts. The upper half of the ring represents the actual pathological diagnosis, and the lower half represents the type of model prediction results. ROC curve, receiver operating characteristic curve; AUC, area under the curve; FN, false negative; TP, true positive; FP, false positive; TN, true negative

Table 3 Performance of TPOT in the diagnosis of testicular lesions

Group	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)	NPV (%)	PPV (%)
Training cohort	0.989 (0.985,0.993)	90.4 (88.4,92.4)	100 (100,100)	94.2 (93.0,95.4)	87.2 (84.9,89.6)	100 (100,100)
Testing cohort	0.909 (0.893,0.923)	81.8 (78.7,84.6)	73.3 (68.8, 77.4)	78.4 (75.9,80.7)	73.3 (69.7,76.5)	81.8 (79.4,84.0)

Note: Data in parentheses are 95% confidence intervals. AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value

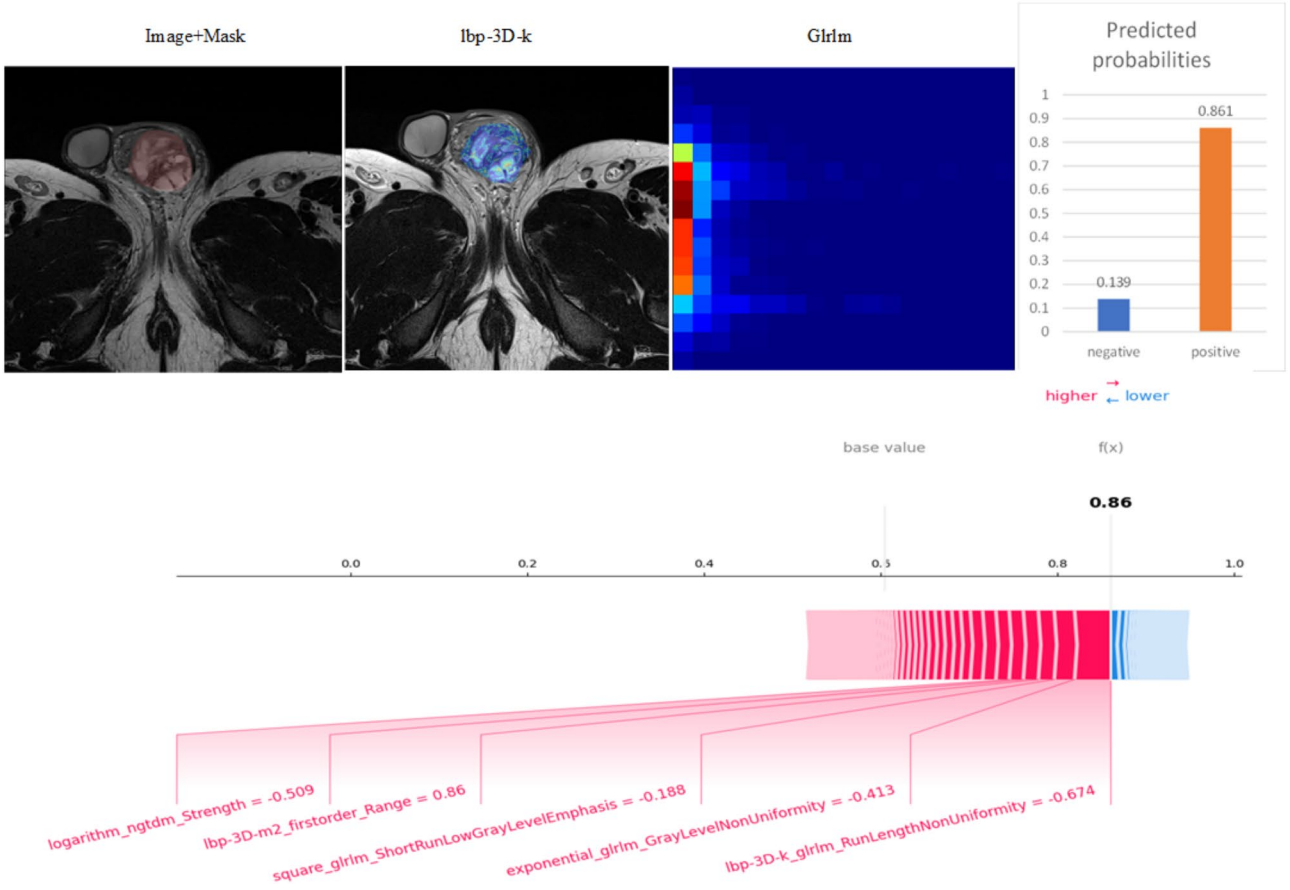

Fig. 4 Results of individual-level interpretation for one malignant lesion (embryonal carcinoma). The features are ranked in the first two places according to the SHAP values and waterfall plots of decisions. SHAP, Shapley additive exPlanations

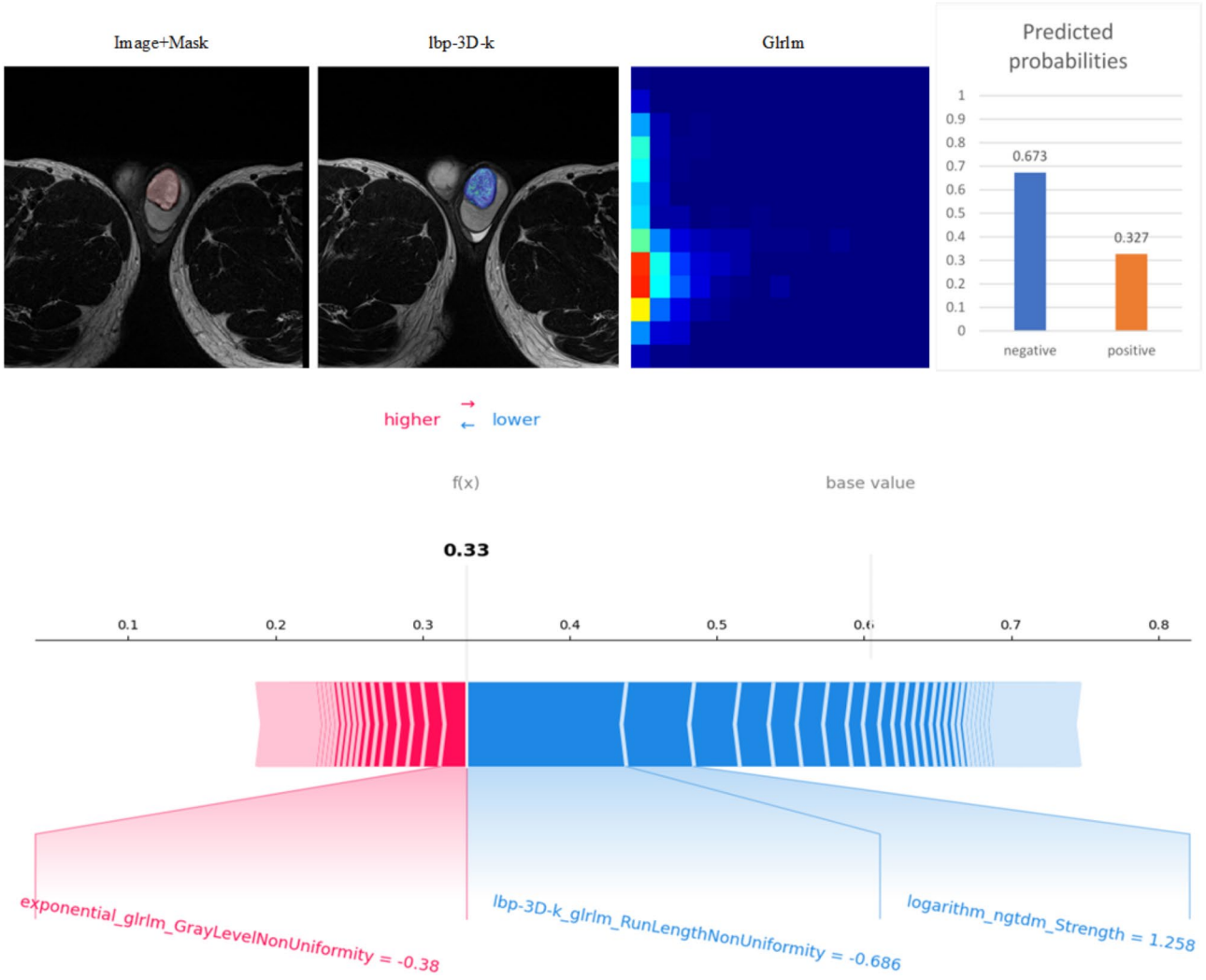

Fig. 5 Results of individual-level interpretation for one benign lesion (epidermoid cyst). The features are ranked in the first two places according to the SHAP values and waterfall plots of decisions. SHAP, Shapley additive exPlanations

Abbreviations

AUC	Area under the curve
AutoML	Automatic machine learning
ICCs	Intraclass and interclass correlation coefficients
mpUS	Multiparameter ultrasound
MRI	Magnetic resonance imaging
ROC	Receiver operating characteristic
SHAP	Shapley additive exPlanations
TPOT	Tree-based pipeline optimization tool
T2WI	T2-weighted Images
US	Ultrasonography
VOI	Volume of interest

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-13844-3.

Supplementary Material 1

Author contributions

LW, PP, ZY and LJ contributed to the conception and design of this study. PP and ZY contributed to data acquisition. LW, YH and WZ contributed to data interpretation and analysis. LW, PP and YH contributed to manuscript editing. XD and ZY contributed to analysis validation. LJ and ZY contributed to study supervision. All authors read and approved the manuscript.

Funding

This work was supported by the Knowledge Innovation Program of Wuhan-Shuguang Project; the National Natural Science Foundation of China (Nos. 81801668, 62072349, and 82001787).

Data availability

The raw datasets used and analyzed in this article are available from the corresponding author (fzm198822@126.com or lijin@tjh.tjmu.edu.cn) on reasonable request.

Declarations

Ethical approval and consent to participate

This retrospective study was performed in accordance with the Declaration of Helsinki. The Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology Institutional Review Board approved the study and waived informed consent, as it was a retrospective study and all procedures performed were part of clinical examinations.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 7 April 2024 / Accepted: 28 February 2025
Published online: 28 March 2025

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.
2. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cancer J Clin. 2023;73(1):17–48.
3. Ramanathan S, Dogra V. Current status of percutaneous testicular biopsy for focal lesions. Abdom Radiol (NY). 2018;43(11):3125–31.
4. Dieckmann KP, Kulejewski M, Heinemann V, Loy V. Testicular biopsy for early cancer detection–objectives, technique and controversies. Int J Androl. 2011;34(4 Pt 2):e7–13.
5. Chovanec M, Hanna N, Cary KC, Einhorn L, Albany C. Management of stage I testicular germ cell tumours. Nat Rev Urol. 2016;13(11):663–73.
6. Song G, Xiong GY, Fan Y, Huang C, Kang YM, Ji GJ, et al. The role of tumor size, ultrasonographic findings, and serum tumor markers in predicting the likelihood of malignant testicular histology. Asian J Androl. 2019;21(2):196–200.
7. Tsili AC, Bertolotto M, Turgut AT, Dogra V, Freeman S, Rocher L, et al. MRI of the scrotum: recommendations of the ESUR scrotal and penile imaging working group. Eur Radiol. 2018;28(1):31–43.
8. Lagabrielle S, Durand X, Droupy S, Izard V, Marcelli F, Huyghe E, et al. Testicular tumours discovered during infertility workup are predominantly benign and could initially be managed by sparing surgery. J Surg Oncol. 2018;118(4):630–5.
9. Tsili AC, Argyropoulou MI, Dolciami M, Ercolani G, Catalano C, Manganaro L. When to ask for an MRI of the scrotum. Andrology. 2021;9(5):1395–409.
10. Isidori AM, Pozza C, Gianfrilli D, Giannetta E, Lemma A, Pofi R, et al. Differential diagnosis of nonpalpable testicular lesions: qualitative and quantitative contrast-enhanced US of benign and malignant testicular tumors. Radiology. 2014;273(2):606–18.
11. Yu J, Jiang XH, Du LF, Bai M, Li ZJ, Shi QS, et al. Contrast-enhanced ultrasound as a valuable imaging modality for characterizing testicular lesions. Asian J Androl. 2022;24(2):201–6.
12. Luzurier A, Maxwell F, Correas JM, Benoit G, Izard V, Ferlicot S, et al. Qualitative and quantitative contrast-enhanced ultrasonography for the characterisation of non-palpable testicular tumours. Clin Radiol. 2018;73(3):322. e1-.e9.
13. Belfield J, Findlay-Line C. Testicular germ cell Tumours-The role of conventional ultrasound. Cancers (Basel). 2022;14(16).
14. Mittal PK, Abdalla AS, Chatterjee A, Baumgarten DA, Harri PA, Patel J, et al. Spectrum of extratesticular and testicular pathologic conditions at scrotal MR imaging. Radiographics. 2018;38(3):806–30.
15. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.
16. Gonçalves M, Gsaxner C, Ferreira A, Li J, Puladi B, Kleesiek J et al. Radiomics in head and neck Cancer outcome predictions. Diagnostics (Basel). 2022;12(11).
17. Wang X, Xie T, Luo J, Zhou Z, Yu X, Guo X. Radiomics predicts the prognosis of patients with locally advanced breast cancer by reflecting the heterogeneity of tumor cells and the tumor microenvironment. Breast Cancer Res. 2022;24(1):20.
18. Li G, Li L, Li Y, Qian Z, Wu F, He Y, et al. An MRI radiomics approach to predict survival and tumour-infiltrating macrophages in gliomas. Brain. 2022;145(3):1151–61.
19. Gu J, Jiang T. Ultrasound radiomics in personalized breast management: current status and future prospects. Front Oncol. 2022;12:963612.
20. Li Y, Qin Q, Zhang Y, Cao Y. Noninvasive determination of the IDH status of gliomas using MRI and MRI-Based radiomics: impact on diagnosis and prognosis. Curr Oncol. 2022;29:6893–907.
21. Fan C, Sun K, Min X, Cai W, Lv W, Ma X, et al. Discriminating malignant from benign testicular masses using machine-learning based radiomics signature of appearance diffusion coefficient maps: comparing with conventional mean and minimum ADC values. Eur J Radiol. 2022;148:110158.
22. Baessler B, Nestler T, Pinto Dos Santos D, Paffenholz P, Zeuch V, Pfister D, et al. Radiomics allows for detection of benign and malignant histopathology in patients with metastatic testicular germ cell tumors prior to post-chemotherapy retroperitoneal lymph node dissection. Eur Radiol. 2020;30(4):2334–45.
23. Dafflon J, Pinaya WHL, Turkheimer F, Cole JH, Leech R, Harris MA, et al. An automated machine learning approach to predict brain age from cortical anatomical measures. Hum Brain Mapp. 2020;41(13):3555–66.
24. Le TT, Fu W, Moore JH. Scaling tree-based automated machine learning to biomedical big data with a feature set selector. Bioinformatics. 2020;36(1):250–6.
25. Zhou H, Hu R, Tang O, Hu C, Tang L, Chang K, et al. Automatic machine learning to differentiate pediatric posterior Fossa tumors on routine MR imaging. AJNR Am J Neuroradiol. 2020;41(7):1279–85.
26. Radzi SFM, Karim MKA, Saripan MI, Rahman MAA, Isa INC, Ibahim MJ. Hyper-parameter tuning and pipeline optimization via grid search method and Tree-Based automl in breast Cancer prediction. J Pers Med. 2021;11(10).
27. Meng J, Wang G, Zhou L, Jiang S, Qian S, Chen L, et al. Mapping variation of extracellular matrix in human keloid Scar by label-free multiphoton imaging and machine learning. J Biomed Opt. 2023;28(4):045001.
28. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to Decode the radiographic phenotype. Cancer Res. 2017;77(21):e104–7.
29. Mongan J, Moy L, Charles E, Kahn J. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. Radiology: Artificial Intelligence. 2020;2(2):e200029.
30. Tejani AS, Klontzas ME, Gatti AA, Mongan J, Moy L, Park SH, et al. Updating the checklist for artificial intelligence in medical imaging (CLAIM) for reporting AI research. Nat Mach Intell. 2023;5(9):950–1.
31. Parenti GC, Feletti F, Carnevale A, Uccelli L, Giganti M. Imaging of the scrotum: beyond sonography. Insights Imaging. 2018;9(2):137–48.
32. Bian T, Wu Z, Lin Q, Mao Y, Wang H, Chen J, et al. Evaluating Tumor-Infiltrating lymphocytes in breast Cancer using preoperative MRI-Based radiomics. J Magn Reson Imaging. 2022;55(3):772–84.
33. Pan J, Zhang K, Le H, Jiang Y, Li W, Geng Y, et al. Radiomics nomograms based on Non-enhanced MRI and clinical risk factors for the differentiation of chondrosarcoma from enchondroma. J Magn Reson Imaging. 2021;54(4):1314–23.
34. Khanna M, Abualruz AR, Yadav SK, Mafraji M, Al-Rumaihi K, Al-Bozom I, et al. Diagnostic performance of multi-parametric MRI to differentiate benign sex cord stromal tumors from malignant (non-stromal and stromal) testicular neoplasms. Abdom Radiol (NY). 2021;46(1):319–30.
35. Manganaro L, Saldari M, Pozza C, Vinci V, Gianfrilli D, Greco E, et al. Dynamic contrast-enhanced and diffusion-weighted MR imaging in the characterisation of small, non-palpable solid testicular tumours. Eur Radiol. 2018;28(2):554–64.
36. Xu M, Fang M, Zou J, Yang S, Yu D, Zhong L, et al. Using biparametric MRI radiomics signature to differentiate between benign and malignant prostate lesions. Eur J Radiol. 2019;114:38–44.
37. Feliciani G, Mellini L, Carnevale A, Sarnelli A, Menghi E, Piccinini F, et al. The potential role of MR based radiomic biomarkers in the characterization of focal testicular lesions. Sci Rep. 2021;11(1):3456.
38. Zhang P, Feng Z, Cai W, You H, Fan C, Lv W, et al. T2-Weighted Image-Based radiomics signature for discriminating between seminomas and nonseminoma. Front Oncol. 2019;9:1330.
39. Olson RS, Bartley N, Urbanowicz RJ, Moore JH. Evaluation of a Tree-based Pipeline Optimization Tool for Automating Data Science2016 March 01, 2016:[arXiv:1603.06212 p.]. Available from: https://ui.adsabs.harvard.edu/abs/2016arXiv160306212O
40. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Mach Learn Res. 2012;13(1):281–305.

Publisher's note