Iranian Journal of Medical Physics

ijmp.mums.ac.ir

Diagnostic Value of Radiomics Based on T2- and T1 Fat-Saturated MRI Sequences for Differentiating Ovarian Endometrioma from Hemorrhagic Cyst

Farrokh Seilanian Toosi[1], Behzad Aminzadeh[1], Shahrokh Nasseri[2], Farzaneh Sohrabi[1] , Saeed Shirdel[1*]

1. Department of Radiology, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
2. Medical Physics Research Center, Basic Science Research Institute, Mashhad University of Medical Sciences, Mashhad, Iran

ARTICLE INFO

Article type:
Original Paper

Keywords:
Endometriosis
Hemorrhagic Ovarian Cysts
Magnetic Resonance Imaging
Texture Analysis
Consensus Feature Selection

ABSTRACT

Introduction: Radiomics is not yet widely adopted in clinical practice due to limited understanding of the biological significance of texture features and insufficient evidence of its superiority over conventional imaging. This study aimed to evaluate radiomics-based quantification of endometriomas and hemorrhagic cysts and assess their differentiation.

Material and Methods: This diagnostic study evaluated the performance of radiomics-based texture analysis in distinguishing ovarian endometriomas from hemorrhagic ovarian cysts using pelvic MRI data from 150 patients. Segmentation and feature extraction were performed using 3D Slicer, followed by statistical analysis of 107 radiomic features. A nested cross-validation strategy with consensus Lasso-based feature selection was applied to train Support Vector Classifier (SVC) and K-Nearest Neighbors (KNN) models. Model performance was assessed via receiver operator characteristic (ROC) analysis, yielding metrics such as area under the curve (AUC), accuracy, sensitivity, specificity, and F1-score.

Results: A total of 107 features per dataset were analyzed using SVC and KNN classifiers, with feature selection guided by Shapley additive explanation (SHAP) values and statistical testing. Shape-based features such as Sphericity, and first-order statistical features such as Skewness, were among the most important predictors. Sphericity, Skewness, and textural features such as Coarseness were frequently selected and showed statistically significant differences in T2 and T1 Fat-Sat +T2 datasets, while the T1 Fat-Sat dataset exhibited weaker discriminative power. The highest classification performance was achieved with the T1 Fat-Sat +T2 dataset, followed by T2, and then T1 Fat-Sat. Similar trends were observed with the KNN classifier.

Conclusion: The study shows that T2-weighted MRI performs well on its own, especially when data is limited, while T1 Fat-Sat alone is unreliable. Combining T1 Fat-Sat and T2 offers the best results for radiomics-based classification of some of ovarian lesions. Among classifiers, SVC had higher specificity and accuracy, whereas KNN showed better sensitivity.

► *Please cite this article as:*
Seilanian Toosi F, Aminzadeh B, Nasseri Sh, Sohrabi F, Shirdel S. Diagnostic Value of Radiomics Based on T2- and T1 Fat-Saturated MRI Sequences for Differentiating Ovarian Endometrioma from Hemorrhagic Cyst. Iran J Med Phys 2026; 23 (3): 219-233. 10.22038/ijmp.2026.95767.2694.

Introduction

Endometriomas are among the common causes of gynecological morbidity and represent the most prevalent type of ovarian cysts. These lesions may impair fertility and are associated with an increased risk of ovarian cancer [1, 2]. Due to their similar clinical presentation and imaging characteristics, endometriomas can be misdiagnosed as hemorrhagic ovarian cysts, which are benign and typically self-limiting conditions [3]. Accurate diagnosis and differentiation between these two entities remain challenging. The presence of endometriomas often indicates concomitant endometriotic lesions in the pelvic and intestinal regions, which may lead to a spectrum of complications [4]. Moreover, distinguishing between these lesions is crucial to avoid unnecessary surgical interventions [5].

Transvaginal ultrasonography (TVUS) is recognized as the primary modality for the diagnosis of endometriomas [6]. In specific circumstances—such as the presence of concomitant lesions, suspected complications, or atypical sonographic features—and depending on the severity of clinical symptoms, additional imaging techniques like magnetic resonance imaging (MRI) may be employed. Compared to TVUS, pelvic MRI offers a more comprehensive mapping of endometriotic lesions. It demonstrates superior diagnostic accuracy, particularly for anterior and notably posterior compartment lesions [6, 7].

Typically, T1- and T2-weighted sequences are utilized for the evaluation of ovarian cystic lesions [8]. Advanced imaging techniques, such as diffusion-

*Corresponding Author: Tel: +98 9139520854; Email: saeedshirdel1368@gmail.com

weighted imaging (DWI) and dynamic contrast-enhanced sequences, may also be employed in selected cases [9]. One of the hallmark MRI features for identifying endometriomas is the "T2 shading" phenomenon [8]. This refers to an adnexal cyst that appears hyperintense on T1-weighted images and hypointense on T2-weighted images [10].

However, the "T2 shading" phenomenon has also been observed in hemorrhagic ovarian cysts, as these lesions may contain substantial amounts of blood, thereby mimicking the imaging appearance of endometriomas [11]. Recent studies indicate that the diagnostic performance of the T2 shading sign is lower than previously estimated, with reported sensitivity ranging from 68% to 93% and specificity between 45% and 93% [12].

Another distinguishing imaging feature of endometriomas on MRI is the presence of "T2 dark spots." Although blood clots may also form within simple hemorrhagic ovarian cysts, the transient nature of these cysts limits the time available for free water reduction and hemosiderin accumulation. Consequently, such cysts typically do not exhibit focal signal loss on T2-weighted images. Findings have demonstrated that "T2 dark spots" exhibit high specificity (93%) but relatively low sensitivity (36%) in differentiating endometriomas from non-endometriotic hemorrhagic cystic lesions [11, 12].

Radiomics represents a rapidly evolving domain of research that enables the extraction of clinically meaningful data from radiological images. These analyses predominantly rely on anatomical imaging techniques. By uncovering nuanced correlations between image-derived features and disease characteristics, radiomics can be employed to forecast clinical outcomes [13].

Texture analysis is considered a fundamental component of radiomics. This approach reveals significant differences in the physical and biochemical properties between endometriomas and hemorrhagic ovarian cysts, based on the extraction and processing of specific imaging parameters through quantification of pixel intensity and distribution patterns [14]. Texture analysis enables an objective characterization of image content, allowing for the computation of histogram-based and morphological features from MRI scans [15].

Despite its potential, radiomics has not yet been widely adopted in clinical practice due to the limited understanding of the biological significance of the measured texture features and the lack of sufficient clinical evidence supporting its superiority over conventional imaging analysis methods. Furthermore, the high false-positive rate and the limited availability of validated radiomics models within diagnostic and therapeutic systems have contributed to its underutilization [16]. In light of these considerations, the present study was conducted to evaluate radiomics-based approaches for quantifying imaging features of endometriomas and hemorrhagic cysts, as well as to investigate their differentiation.

Materials and Methods

This was a retrospective study conducted between April 2018 and March 2025. This was a diagnostic study conducted to evaluate the validity (accuracy, sensitivity, and specificity) of radiomics methods aimed at differentiating ovarian endometriomas from hemorrhagic ovarian cysts. The validation process was performed through pelvic MRI data analysis at Qaem Educational, Research, and Treatment Center in Mashhad.

Inclusion and Exclusion Criteria

Inclusion criteria consisted of patients over 18 years of age with a confirmed diagnosis of ovarian endometrioma or hemorrhagic ovarian cyst. Exclusion criteria included lesions smaller than 15 mm, concurrent ovarian pathologies, and lack of confirmed pathological or clinical diagnosis. Initially, concurrent ovarian pathologies were considered exclusion criteria; however, several conditions that are commonly associated with endometrioma were subsequently permitted. These conditions included Deep Infiltrating Endometriosis (DIE), Hematosalpinx, Hydrosalpinx, and hemorrhagic cysts in patients with endometrioma. In cases where the endometrioma was consistently located contralateral to the hemorrhagic cyst, the patient was included; otherwise, the patient was excluded.

The analysis was performed at the lesion level, with only one lesion per patient included in the dataset to avoid intra-patient correlation. In patients with coexisting endometrioma and hemorrhagic cyst, only one representative lesion was included, and the second lesion was excluded. Accordingly, each patient contributed only a single sample to the dataset, ensuring that no patient appeared in more than one subset of the analysis. The selection process for both patient groups is illustrated in Figure 1.

For endometriomas, 73% of cases had histopathological and surgical confirmation. The remaining cases were diagnosed based on characteristic imaging findings on ultrasound and MRI and/or imaging follow-up. Only cases with typical and concordant imaging features were included, while equivocal cases were excluded. In nine cases with initially incomplete documentation, direct contact with patients was used to obtain prior medical records, and lesion stability over time was considered supportive of the diagnosis.
For hemorrhagic cysts, seven cases had histopathological confirmation, and the remainder were confirmed based on follow-up imaging demonstrating resolution. The minimum follow-up duration was 8 weeks. In cases with shorter interval imaging, a ≥50% reduction in lesion size, together with supportive features such as cyst wall collapse or wrinkling, was used as diagnostic criteria.

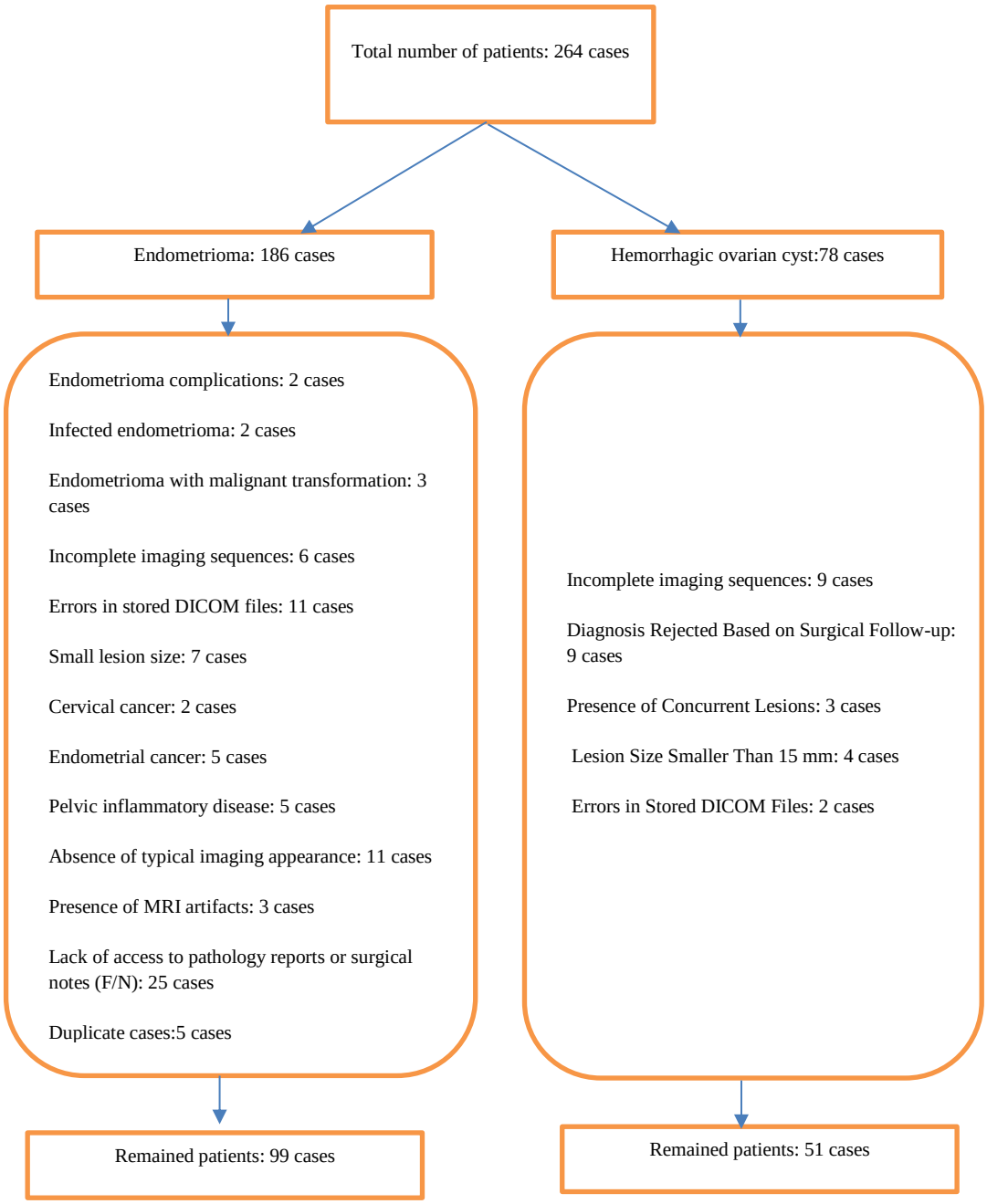


Figure 1. Patient Selection Process for Hemorrhagic Cyst and Endometrioma Groups

Study design

In total, pelvic MRI data from 150 female patients diagnosed with ovarian endometrioma or hemorrhagic ovarian cyst were collected. Initially, DICOM images and clinical reports were reviewed. For patients who had undergone surgery, the pathological findings were considered the diagnostic reference. In other cases, clinical follow-up and imaging results were used for diagnosis. For patients lacking both, direct contact was made. Given the extended duration of the study, lesions that showed size reduction over time were classified as hemorrhagic cysts, while stable lesions were considered endometriomas.

Imaging Protocol

MRI scans were performed with patients in the supine position and with a partially filled bladder. Patients were preferably asked to fast for 4 to 6 hours prior to scanning to minimize gas-related artifacts. Pelvic MRI was conducted using a 1.5 Tesla Avanto system (Siemens, Germany) in coronal, sagittal, and axial planes with T2, DWI, ADC, and T1 sequences.

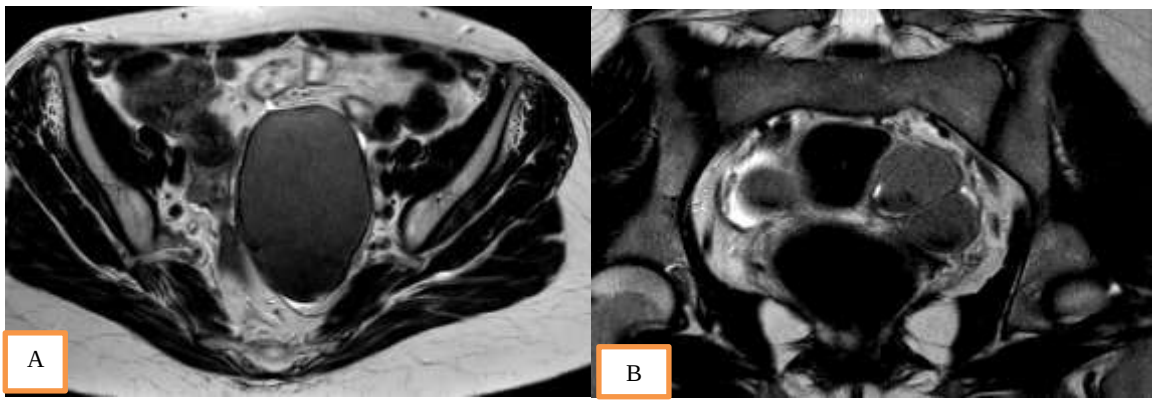


Figure 2. (A) T2 dark spot sign, this sign describes a focal T2 hypointense nodule against the wall. (B) The shading sign, this sign describes lesions that are hyperintense on T1 weighted sequences and hypointense on T2-weighted sequences affecting variable portions of the cyst

Non-contrast sequences included T2 sagittal, T1 sagittal, fat-saturated T1 turbo spin-echo, T2 axial turbo spin-echo, T1 axial, fat-s
aturated T1 transverse spin-echo, T2 turbo spin-echo coronal, DWI (B0/B400/B800), ADC, and T1 VIBE sagittal. Eligible patients first underwent pre-contrast MRI, followed by gadolinium injection (0.1 cc/kg body weight) at a rate of 2 cc/sec. Five scans were then acquired within the first 3 minutes post-injection (every 30 seconds), along with a delayed scan at 3 minutes. Although the standard imaging protocol at our tertiary center routinely includes all of the sequences listed above for evaluating adnexal lesions, only the T2-weighted and pre-contrast T1-weighted fat-saturated sequences were used for the purposes of the present study (Figure 2).

Radiomics Approach

To separate hemorrhagic cysts from ovarian endometriomas, radiomics analysis was applied in this investigation. The radiomics process was comprised of three primary phases: segmenting images through regions of interest (ROIs), extracting features, selecting features, and evaluating models. Intensity normalization was performed using z-score standardization, in which voxel intensities were normalized within each image based on the mean and standard deviation of the corresponding image volume. This step was applied to reduce minor variations in signal intensity and to improve comparability of radiomics features across patients. All MRI examinations were acquired at a single institution using a standardized imaging protocol, which minimized inter-scanner variable.

Segmentation: In this study, the 3D Slicer software was utilized for image segmentation. Axial views from both T1 fat-saturated and T2-weighted MRI sequences were used for segmentation. Patients lacking either of these sequences were excluded from the study. On T1-weighted images, both hemorrhagic cysts and endometriomas appeared hyperintense across slices. Endometriomas exhibited a characteristic "shading" appearance on T2-weighted images, whereas hemorrhagic cysts generally also demonstrated hyperintensity on T2 sequences. Radiomic features were extracted independently from T1-weighted and T2-weighted MR images. Lesions were initially segmented on the T1-weighted images. The corresponding lesion location was then visually assessed on the T2-weighted images. Whenever complete anatomical correspondence was not observed due to differences in slice positioning, image acquisition, or displacement caused by respiratory motion and patient movement, the lesion was manually re-segmented directly on the T2-weighted images. Thus, each MRI sequence had its own independently generated three-dimensional region of interest (ROI). No algorithmic image registration or interpolation was performed during this process. These segmented regions were saved as separate segmentation files. The analysis was performed at the lesion level, with only one lesion per patient included in the dataset to avoid intra-patient correlation. In cases of bilateral disease Lesion segmentation was manually performed in a three-dimensional manner using 3D Slicer by a radiologist and subsequently reviewed and approved by a senior radiologist with more than 30 years of experience in gynecologic imaging. The radiologist performing image interpretation and lesion segmentation was blinded to the histopathological results at the time of analysis to minimize interpretation bias.

Feature Extraction: Following segmentation, feature extraction was performed using the Radiomics module within the 3D Slicer software. Approximately 107 features were extracted from the region of interest (ROI), which were subsequently used for statistical analysis. After completing the segmentation process and running the Radiomics module, a set of quantitative features was derived from the ROI. These features were categorized into several main groups, as presented in Figure 3.

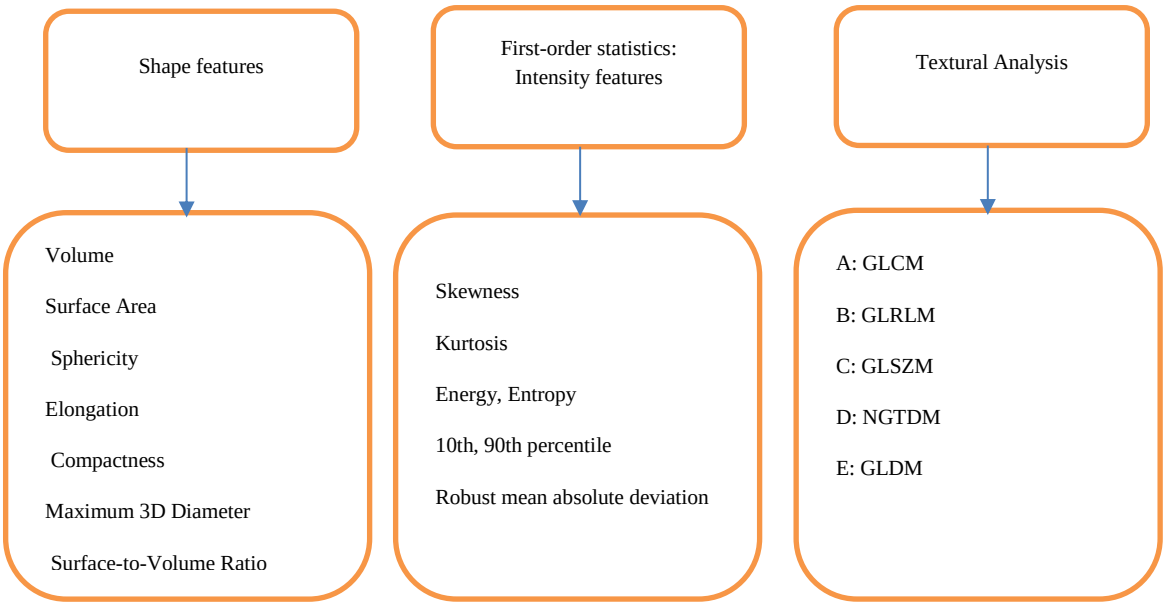


Figure 3. Feature Classification in Radiomics Analysis

Gray Level Co-occurrence Matrix (GLCM) Gray Level Run Length Matrix (GLRLM)
Gray Level Size Zone Matrix (GLSZM) Neighboring Gray Tone Difference Matrix (NGTDM)
Gray Level Dependence Matrix (GLDM)

Feature selection and Model Evaluation: A repeated nested k-fold stratified cross-validation (CV) strategy with consensus-based feature selection was implemented [17]. The outer CV loop was designated for evaluating model performance, while the inner loop handled feature selection and hyperparameter tuning. To balance bias and variance, both loops used k=5k = 5 folds per iteration [18]. At each step, feature vectors from the training set were passed into the inner loop. Within this loop, data were standardized and Lasso-based feature selection was applied using a logistic regression model [19]. Feature selection frequency was tracked to assess stability, and features selected in more than one fold of inner loop (threshold frequency of 0.2) were compiled into the final set.

All steps of feature selection (using LASSO), hyperparameter tuning, and data normalization were performed exclusively within the inner cross-validation loop to avoid data leakage. The outer cross-validation loop was used only for model evaluation. Feature selection was based on a consensus approach across inner folds, and features selected consistently across folds were retained for model construction in each outer iteration. All preprocessing and model optimization steps were confined to the training data within each inner fold, thereby minimizing the risk of data leakage.

The study dataset included 99 endometriomas and 51 hemorrhagic cysts, representing a moderate class imbalance. To preserve the original characteristics of the dataset, no synthetic oversampling methods, including the Synthetic Minority Oversampling Technique (SMOTE), were applied during model development. This decision was made because synthetic sample generation may alter the distribution of radiomics features, increase class overlap, and introduce artificial variability, particularly in relatively small datasets.These factors may result in unrealistic decision boundaries and reduced model generalizability.

Support Vector Classifier (SVC) and K-Nearest Neighbors (KNN) models were trained on the full outer training set using the selected features and optimized hyperparameters from the inner loop. Hyperparameter tuning—including the L1 regularization strength for Lasso, the C parameter for SVC, and the number of neighbors for KNN—was performed using the Tree Parzen Estimator algorithm in HyperOpt, with 200 evaluation runs focused on maximizing accuracy. Using classification accuracy as the primary optimization measure, hyperparameter tuning was conducted. Even though the area under the receiver operating characteristic curve (AUC) was computed, it was employed for model evaluation rather than parameter optimization. Classification accuracy is measured at the established decision threshold (0.5 in our study), however AUC is a threshold-independent measure that condenses model discrimination across all potential decision thresholds shown by the ROC curve. Because the final classifier was designed to operate with this constant threshold, it was believed that optimizing hyperparameters based on classification accuracy was more appropriate for the proposed clinical application. Despite the fact that balanced accuracy was taken into account, it was not chosen as the main optimization metric since our goal was to maximize overall classification performance at the predetermined operating threshold rather than to optimize a mean class-wise performance metric. To provide a thorough evaluation of model performance, sensitivity, specificity,

F1-score, and AUC were also disclosed. Feature importance was assessed using Shapley Additive Explanations (SHAP) values derived from the trained models [20]. To account for variability in CV splits, the entire process was repeated 50 times with different random seeds.

Statistical analysis

The normality of variables and selected features was assessed using the Shapiro-Wilk test. Clinical variables between the two groups were compared using the independent t-test and the chi-square test. Selected features were analyzed using the Mann-Whitney U test or independent t-test as appropriate to evaluate significant differences across labels within each sequence. A p-value less than 0.05 was considered statistically significant. statistical testing, p-values were computed only for the features most frequently selected across models to provide an additional descriptive interpretation of their overall association with the outcome. These values were not used in the feature selection or model training process.

Model performance was reported in terms of the area under the ROC curve (AUC), along with accuracy, sensitivity, specificity, and F1-score with 95% confidence interval. AUC values were categorized as follows: excellent (0.9–1), good (0.8–0.9), fair (0.7–0.8), poor (0.6–0.7), and non-discriminative (<0.6). Statistical analyses were conducted using Python (version 3.12.11) with the following libraries: SciPy (version 1.16.1) [21], scikit-learn (version 1.6.1) [22], and SHAP (version 0.48.0) [23].

Ethical Considerations

In this study, all information was kept strictly confidential and used solely for research purposes. The study was approved by the Institutional Ethics Committee of the Faculty of Medicine, Mashhad University of Medical Sciences, under the code IR.MUMS.MDICSAL.REC.1403.101.

Results

Demographic and clinical Findings

In total, 150 patients were enrolled in this study. Among them, 99 patients (66%) were diagnosed with endometrioma, and 51 patients (34%) were diagnosed with hemorrhagic ovarian cysts. The mean age of patients in the endometrioma group was 32.77 ± 7.16 years, while in the hemorrhagic cyst group it was 33.48 ± 9.51 years. There was no statistically significant difference in age between the two groups (t = 0.505, P = 0.614).

In the endometrioma group, the cyst was located on the right side in 25.5% of cases, on the left side in 30.6%, and bilaterally in 43.9%. In the hemorrhagic cyst group, the cyst was located on the right side in 48% of cases, on the left side in 48%, and bilaterally in 4%. Bilateral cysts were more common in the endometrioma group, whereas unilateral cysts were more prevalent in the hemorrhagic cyst group (P < 0.005). The types of concurrent pathologies are presented in Table 1. The most common concurrent pathology in patients with endometrioma was deep infiltrating endometriosis (56% among those with concurrent pathology), while in patients with hemorrhagic cysts, the most frequent concurrent pathology was contralateral endometrioma (70%).The higher rate of contralateral endometrioma observed in patients with hemorrhagic cysts is mainly explained by the fact that many of these patients had a prior history of endometrioma and underwent MRI due to worsening symptoms. In such cases, the hemorrhagic cyst was largely an incidental finding rather than the primary pathology.

Table 1. Presence of Concurrent Pathologies in the Hemorrhagic Cyst and Endometrioma Groups

Concurrent Pathologies	Endometrioma (n=99)	Hemorrhagic Cyst (n=51)
Hematosalpinx	4 (8)	1 (10)
Hydrosalpinx	2 (4)	--
DIE	28 (56)	--
Hemorrhagic Cyst	5 (10)	--
PIC and Hematosalpinx	2 (4)	--
PIC	2 (4)	--
DIE and PIC	2 (4)	--
Follicular Cyst	2 (4)	--
DIE and Hemorrhagic Cyst	2 (4)	--
Contralateral Endometrioma	--	7 (70)
Ovarian Torsion	--	1 (10)
PCO	--	1 (10)
PCO and Hydrosalpinx	1 (2)	--

DIE: Deep Infiltrating Endometriosis, PIC: Pelvic Inflammatory Condition, PCO: Polycystic Ovary

Percentage were calculated using the number of patients with concurrent pathologies as the denominator

Feature Selection

A total of 107 radiomic features were initially extracted from each lesion. After removing 23 features with zero variance, 107 features remained for further analysis. The findings of the present study were derived from 107 features in the T1 Fat-Sat dataset and 107 features in the T2 dataset. In the SVC classification, three datasets— T1 Fat-Sat, T2, and T1 Fat-Sat +T2—were evaluated, as illustrated in Figure 4. A total of 250 models were trained using 50 repetitions of 5-fold cross-validation,. features selection has been done during inner folds with lasso method and as the consensus feature selection features repeated in more than one-fold (threshold = 0.2) which mean feature that isn't selected totally or feature that just one time selected during inner fold validation (5-fold) isn't consider as a feature for corresponding outer fold. For final frequency analyze for each feature, frequencies from all repeats summed up and plotted. But as depicting all feature frequencies and their corresponding mean SHAP and graphs is confusing we just used features selected in more 40 percent of 250 models (5 fold * 50 repeat = 250 model train on outerfolds) that lead to 13 feature for T1 fat sat, 7 for T2 and 12 for combined for visually graphs. (Table 2)

All steps of feature selection (using LASSO), hyperparameter tuning, and data normalization were performed exclusively within the inner cross-validation loop to avoid data leakage. Feature selection was based on a

consensus approach across inner folds, and features selected consistently across folds were retained for model construction in each outer iteration. Regarding statistical testing, p-values were computed only for the features most frequently selected across models to provide an additional descriptive interpretation of their overall association with the outcome. These values were not used in the feature selection or model training process.

13 features were found in over 40% of the models for distinguishing between ovarian endometriomas and hemorrhagic ovarian cysts in T1 Fat-Sat dataset. Among these, Sphericity as shape feature and Small Area Emphasis were selected most frequently, indicating their relative stability across different model validation stages. The Independent t test revealed that these two features showed statistically significant differences between endometrioma and hemorrhagic cyst groups. However, in the T1 Fat-Sat sequence evaluated with the SVC classifier, several frequently selected features—including Kurtosis_ first order, Strength_ NGTDM, Large Area Emphasis _GLSZM, and Correlation _ glcm—were found to be statistically non-significant, suggesting that the T1 Fat-Sat sequence performs poorly in label classification. Therefore, despite their frequent selection, certain features from the T1 Fat-Sat dataset alone were not effective in distinguishing between the two lesion types.

For the T2 dataset, seven features were identified, while for the combined T1 Fat-Sat +T2 sequence, twelve features were selected for classifying lesions as ovarian endometrioma or hemorrhagic ovarian cyst. Using the SVC classifier, all selected features in the T2 and T1 Fat-Sat +T2 datasets showed statistically significant differences between the two lesion types ($P < 0.05$), indicating superior classification performance of these sequences. In the T2 sequence, Sphericity and Skewness were most frequently selected, whereas in the T1 Fat-Sat +T2 dataset, T1Fat-sat sequence showed the highest selection rates for coarseness, while the T2 sequence had the highest rates for skewness. On the other hand, for the T2 sequence, the GLCM features Inverse Difference Normalized (IDN) and Inverse Difference Moment Normalized (IDMN) indicate the greatest influence on model predictions based on SHAP values. For the T1+T2 dataset, the median from the T2 sequence and Size Zone Non Uniformity Normalized from the T1 sequence are identified as having the most significant impact on model predictions (Table 3) .

Table 2. Features with theirs frequency and mean SHAP value in SVC :SCV Classifier

	Feature	Frequency	Mean SHAP Value	P-value
T1 Fat-Sat	Sphericity_shape	250	0.297	<0.001
	SmallAreaEmphasis_glszm	243	0.290	0.001
	Coarseness_ngtdm	230	0.161	<0.001
	DependenceEntropy_gldm	192	0.241	<0.001
	GrayLevelNonUniformityNormalized_glrlm	189	0.131	0.002
	Correlation_glcm	177	0.150	0.678
	Maximum2DDiameterColumn_shape	150	0.184	<0.001
	LargeAreaEmphasis_glszm	134	0.027	0.845
	LargeAreaHighGrayLevelEmphasis_glszm	125	0.025	0.015
	MeshVolume_shape	117	0.051	0.002
	Kurtosis_firstorder	104	0.048	0.896
	VoxelVolume_shape	102	0.047	0.002
	Strength_ngtdm	101	0.15	0.665
T2	Sphericity_shape	248	0.376	<0.001
	Skewness_firstorder	237	0.315	<0.001
	LongRunHighGrayLevelEmphasis_glrlm	216	0.268	<0.001
	Median_firstorder	201	0.364	<0.001
	Idn_glcm	196	0.424	<0.001
	Idmn_glcm	190	0.534	<0.001
	10Percentile_firstorder	120	0.341	<0.001
T1 Fat-Sat + T2	Coarseness_ngtdm_T1	242	0.189	<0.001
	Skewness_firstorder_T2	232	0.295	<0.001
	Sphericity_shape_T1	216	0.282	<0.001
	Median_firstorder_T2	213	0.413	<0.001
	SmallAreaEmphasis_glszm_T1	205	0.337	0.001
	LongRunHighGrayLevelEmphasis_glrlm_T2	200	0.274	<0.001
	SizeZoneNonUniformityNormalized_glszm_T1	199	0.369	<0.001
	Idmn_glcm_T2	194	0.362	<0.001
	Sphericity_shape_T2	182	0.216	<0.001
	DependenceEntropy_gldm_T1	153	0.255	<0.001
	Contrast_ngtdm_T2	126	0.245	<0.001
	10Percentile_firstorder_T2	107	0.355	<0.001

Table 3. Comparison of T1 Fat-Sat - and T2-Weighted Sequences SVC Classifier

SCV Classifier		Endometrioma	Hemorrhagic cyst	P-value	Test
T1 Fat-Sat	Sphericity_shape	-0.26± 1.01	0.53±0.71	<0.001	Man-Whitney
	SmallAreaEmphasis_GLSZM	0.27±0.71	-0.54±1.22	<0.001	Man-Whitney
	Coarseness_NGTDM	-0.26±0.38	0.53±1.501	<0.001	Man-Whitney
	DependenceEntropy_GLDM	0.31±0.82	-0.62±1.03	<0.001	T-test
	GrayLevelNonUniformityNormalized_GLRLM	-0.25±0.48	0.503±1.45	<0.001	Man-Whitney
	Correlation_GLCM	-0.02 ±0.79	0.04±1.307	0.678	Man-Whitney
	LargeAreaEmphasis_GLSZM	-0.13±0.32	0.25±1.63	0.845	Man-Whitney
	Maximum2DDiameterColumn_SHAPE	0.22±1.01	-0.44±0.801	<0.001	Man-Whitney
	LargeAreaHighGrayLevelEmphasis_GLSZM	-0.05±0.502	0.108±1.56	0.015	Man-Whitney
	ZoneVariance_GLSZM	-0.12±0.32	0.25±1.63	0.823	Man-Whitney
	MeshVolume_SHAPE	0.03±0.83	-0.07±1.26	0.002	Man-Whitney
	Kurtosis_FIRSTORDER	0.00±1.05	0.00±0.88	0.896	Man-Whitney
	VoxelVolume_SHAPE	0.03±0.83	-0.07±1.26	0.002	Man-Whitney
	Strength_NGTDM	-0.14±0.72	0.28±1.35	0.665	Man-Whitney
T2	Sphericity_shape	-0.26±1.04	0.52± 0.64	<0.001	Man-Whitney
	Skewness_firstorder	0.32±0.98	-0.63 ±0.65	<0.001	Man-Whitney
	LongRunHighGrayLevelEmphasis_glrlm	-0.32± 0.55	0.64±1.32	<0.001	Man-Whitney
	Idmn_glcm	0.32±0.77	-0.64±1.08	<0.001	Man-Whitney
	Median_firstorder	-0.39±0.77	0.78±0.92	<0.001	Man-Whitney
	Idn_glcm	0.308±0.79	-0.61±1.08	<0.001	Man-Whitney
	10 Percentile_firstorder	-0.35±0.82	0.71±0.92	<0.001	Man-Whitney

The T1 Fat-Sat, T2, and T1 Fat-Sat +T2 datasets evaluated using the KNN classifier are also presented in Figure 5. In the T1 Fat-Sat dataset, 11 features were identified for lesion classification. When assessed with the KNN classifier, features such as Zone Variance _GLSZM , Large Area Emphasis _GLSZM, and Correlation_ GLCM were found to be statistically non-significant based on the Independent t test ($P > 0.05$), indicating weaker classification performance for the T1 Fat-Sat sequence. Again, Sphericity and Small Area Emphasis were frequently selected, reflecting their relative stability across model validation stages. Nevertheless, these features alone were insufficient for effective discrimination between the two lesion types.

In the T2 dataset, seven features were identified using the KNN classifier, and in the T1 Fat-Sat +T2 dataset, twelve features were selected for distinguishing ovarian endometrioma from hemorrhagic ovarian cyst. All selected features in the T2 and T1 Fat-Sat +T2 sequences showed statistically significant differences between the two lesion types ($P < 0.05$), confirming the strong classification performance of these sequences. Analysis of feature stability revealed that Sphericity and Skewness were frequently selected in the T2 sequence, while Coarseness from T1 Fat-Sat and Skewness from T2 were the most consistent features in the multi-modal dataset. These findings were complemented by SHAP-based importance rankings; specifically, IDMN and Median exerted the greatest influence on the T2-based model. For the combined T1 Fat-Sat + T2 model, the most significant predictive contributions were provided by T1 Fat-Sat features, specifically Size Zone Non Uniformity Normalized and Small Area Emphasis.

Model Performance Evaluation

In this study , Table 4 presents the values of AUC, accuracy, sensitivity, specificity, and F1-score for both the SVC and KNN classifiers. For the SVC classifier, the best performance was observed with the T1 Fat-Sat +T2 dataset, achieving an excellent mean AUC of 0.914 (95% confidence interval: 0.907 to 0.922). This was followed by the T2 dataset, with a good mean AUC of 0.892 (95% CI: 0.885 to 0.899), and then the T1 Fat-Sat dataset, which also showed a good mean AUC of 0.805 (95% CI: 0.796 to 0.814).

For the KNN classifier, the highest performance was again associated with the T1 Fat-Sat +T2 dataset, yielding an excellent mean AUC of 0.905 (95% CI: 0.898 to 0.913). The T2 dataset followed with a good mean AUC of 0.872 (95% CI: 0.865 to 0.88), while the T1 Fat-Sat dataset showed moderate performance with a mean AUC of 0.779 (95% CI: 0.768 to 0.79). The mean AUC trends based on repeated cross-validation (CV) iterations for both SVC and KNN classifiers are illustrated in Figure 6. The findings indicate that the T1 Fat-Sat +T2 dataset yielded the best classification performance, followed by T2, with T1 Fat-Sat showing the lowest performance.

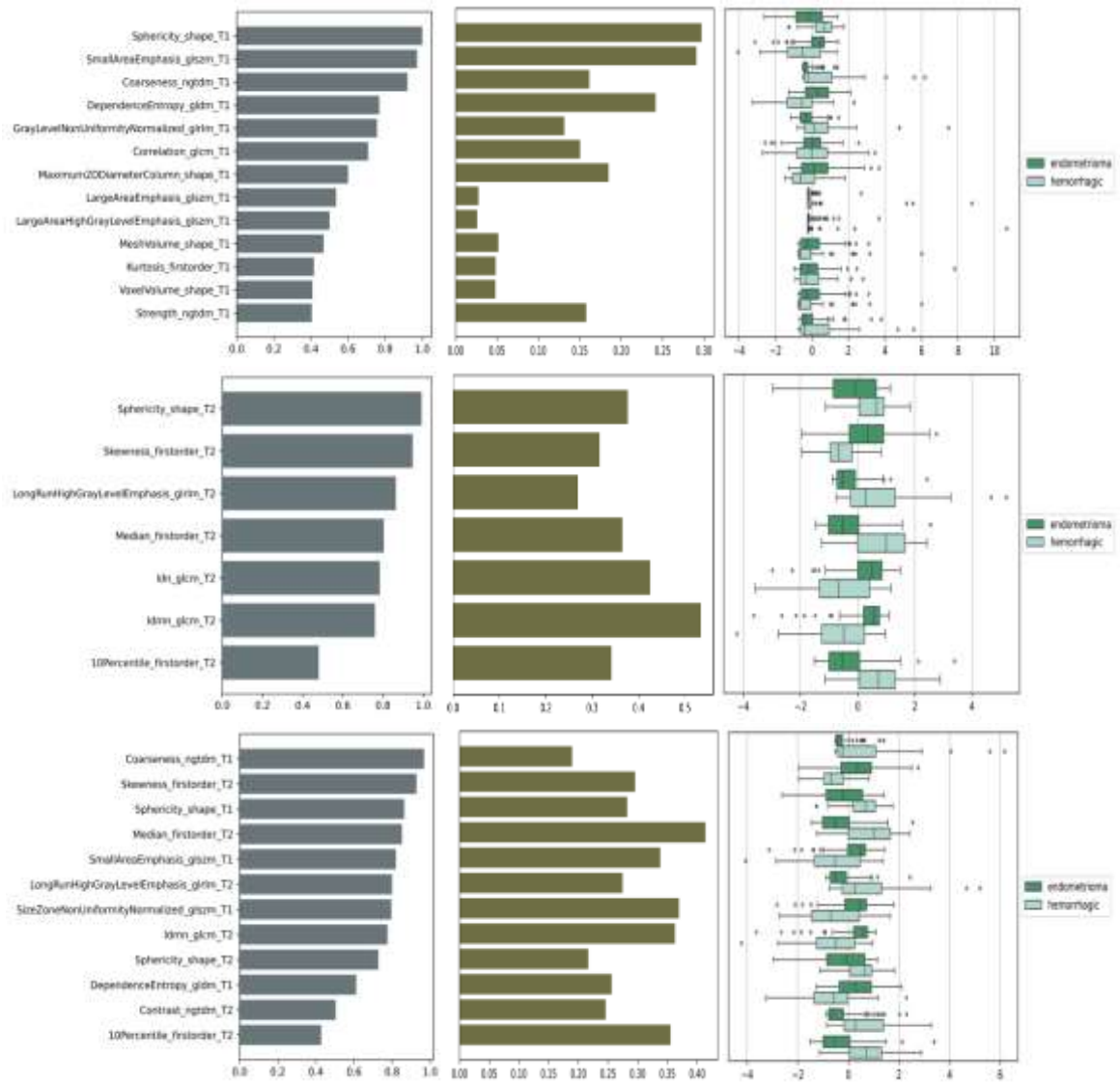


Figure 4. Selected features for subgroup prediction in T1 Fat-Sat, T2, and T1 Fat-Sat +T2 datasets using the SVC classifier. Bar charts on the left indicate the frequency of feature selection. Central bar charts show feature importance measured by the mean absolute SHAP values, reflecting each feature's impact on individual model predictions. Box plots on the right display standardized feature values grouped by primary category. Some sequence features in the T1 Fat-Sat dataset showed no significant difference between groups ($P > 0.05$), while selected features in the T2 and T1 Fat-Sat +T2 datasets exhibited statistically significant differences across all examined categories ($P < 0.05$).

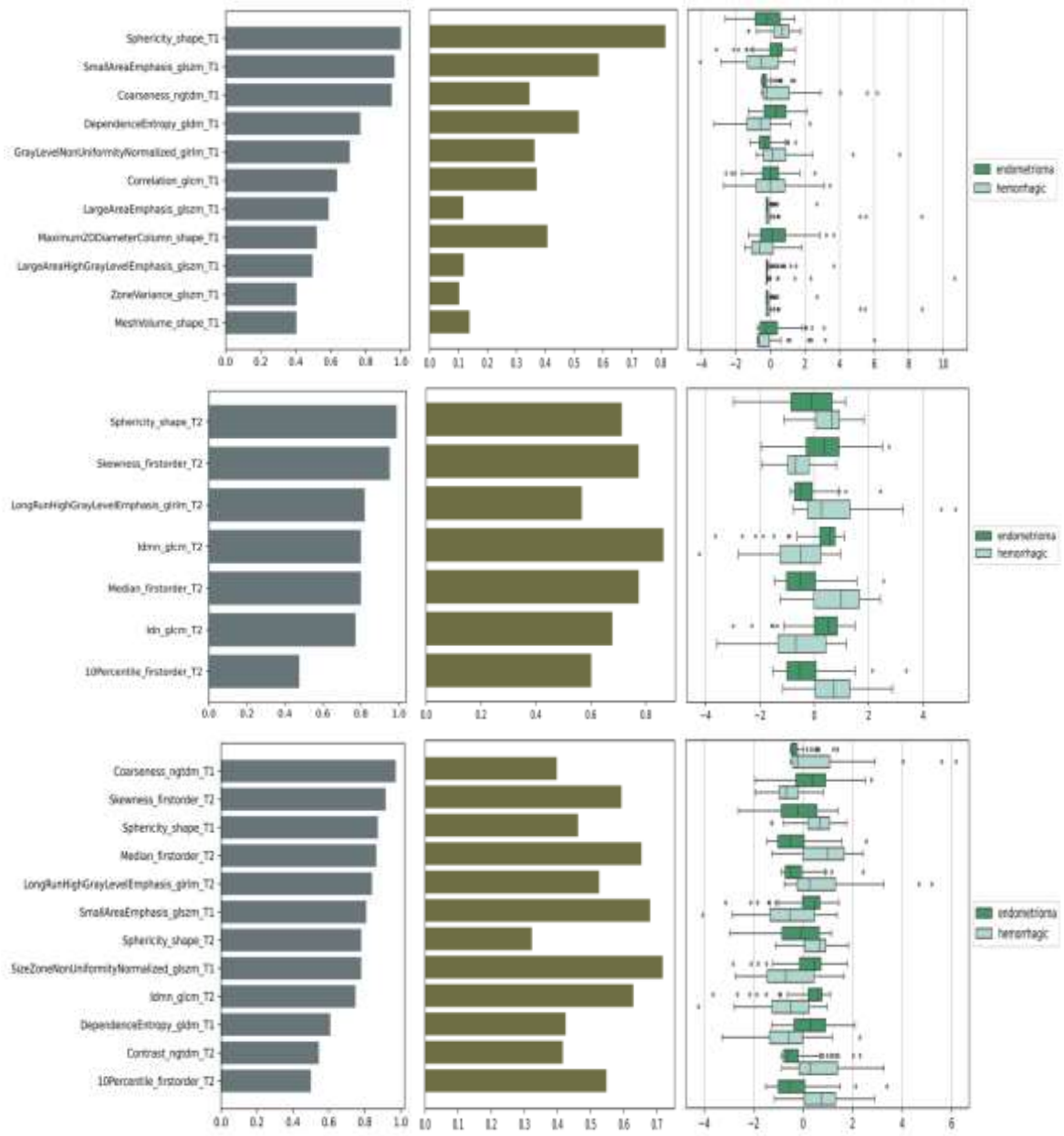


Figure 5. Feature selection results for T1 Fat-Sat, T2, and T1+T2 datasets using the KNN classifier. The bar plots on the left illustrate how frequently each feature was selected, indicating the consistency of selection. The middle bar plots represent feature importance based on the average absolute SHAP values, highlighting each feature's contribution to model predictions. The box plots on the right depict standardized feature values categorized by primary group. In the T1 Fat-Sat dataset, some selected features showed no statistically significant differences between groups ($P > 0.05$), whereas in the T2 and T1 Fat-Sat +T2 datasets, all selected features showed significant differences across categories ($P < 0.05$).

Table 4. Performance metrics of SVC and KNN classifier

	sequence	AUC(%[CI])	Accuracy(%[CI])	Sensitivity(%[CI])	Specificity(%[CI])	F1-score(%[CI])
SVC	T1 Fat-Sat	0.805[0.796-0.814]	73.1[72.2-74.1]	74.4[73-75.8]	70.7[68.9-72.5]	78.3[77.4-79.2]
	T2	0.892[0.885-0.899]	81.7[80.9-82.5]	79.7[78.6-80.8]	85[84.1-87]	85.1[84.3-85.8]
	T1 Fat-Sat + T2	0.914[0.907-0.922]	84[83.2-84.9]	83[81.9-84]	86.2[84.8-87.6]	87.2[86.5-87.9]
KNN	T1 Fat-Sat	0.779[0.768-0.79]	77.1[76.3-77.9]	89.8[89-90.6]	52[50-54]	83.9[83.3-84.4]
	T2	0.872[0.865-0.88]	80.4[79.7-81.2]	85.9[85-86.8]	69.6[67.7-71.5]	85.3[84.7-85.9]
	T1 Fat-Sat + T2	0.905[0.898-0.913]	85.1[84.3-85.8]	91.6[90.9-92.4]	72.2[70.3-74]	89.1[88.5-89.6]

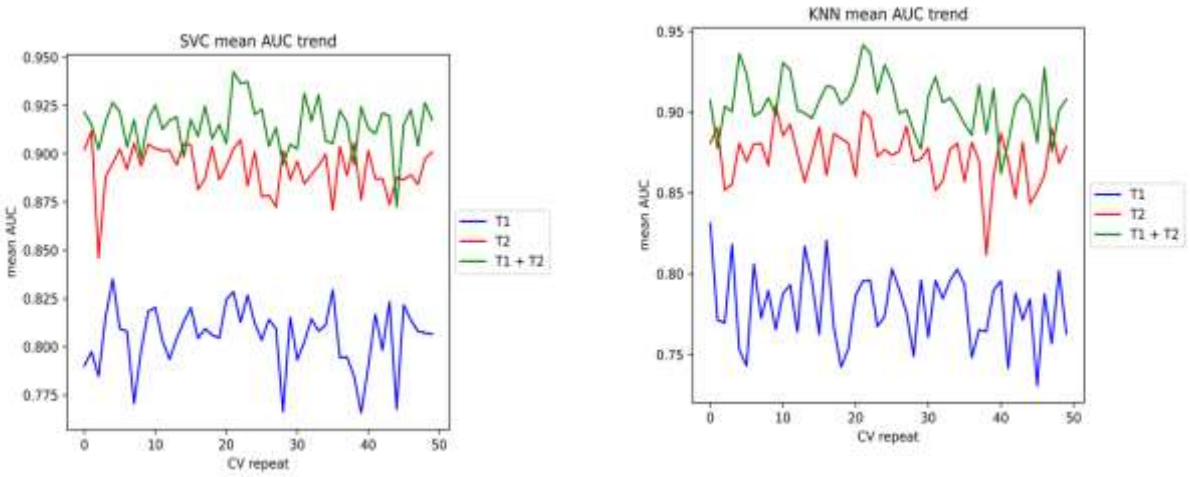


Figure 6. Mean AUC for SVC and KNN classifier across each repetition

Discussion

The findings of the present study revealed that the T1-weighted sequence demonstrated poor performance in label classification when evaluated using both KNN and SVC models. This limitation was evident in both classifiers, particularly in the KNN model, where specificity was reported at only 52%, indicating a high rate of false positives. In contrast, the T2-weighted sequence exhibited superior performance. Several extracted features demonstrated statistical significance in both KNN and SVC models, suggesting the potential discriminative capability of the T2 sequence in identifying adnexal lesions. Performance metrics such as AUC, sensitivity, and F1-score were notably higher. Specifically, in the SVC model, specificity reached 85% and the F1-score was 85%, underscoring the robust diagnostic potential of this sequence. When evaluating the combined T1 Fat-Sat + T2 sequence using both classifiers, all prominent features were statistically significant, indicating the effective selection of key attributes from both imaging modalities. Given that the intended clinical application relied on a specific decision threshold of 0.5, classification accuracy was selected as the primary standard for optimizing hyperparameters. To offer a comprehensive and threshold-independent assessment of model performance, AUC was stated in conjunction with sensitivity, specificity, F1-score, and balanced accuracy.

In this context, the study by Lupean et al. (2020) employed MRI-based texture analysis using the standard T2 FR-FSE HR sequence without the need for contrast agent administration to differentiate between endometriomas and hemorrhagic cysts. Although texture analysis on the standard T2 sequence was able to distinguish endometriomas from hemorrhagic cysts, several evaluated features—such as Sum Average, Percentile Histogram, and Correlation—did not reach statistical significance [24]. In contrast, the present study, through rigorous statistical testing and selection of frequently occurring features, identified a set of statistically significant parameters, particularly within the T2 sequence and the combined T1 Fat-Sat +T2 dataset.

The high diagnostic potential of MRI-based radiomic features in characterizing adnexal lesions has also been confirmed in other studies [26, 27]. Additional research has emphasized the importance of selecting appropriate MRI sequences to enhance diagnostic accuracy, suggesting that optimal sequence selection forms a critical foundation for the development of robust radiomic models and further supports the imaging design strategy adopted in the present study [28, 29].

Scientific evidence suggests that integrating clinical data with imaging features and radiomic parameters can enhance diagnostic accuracy [30-36]. However, the findings of the present study indicate that in specific ovarian lesions such as endometriomas and hemorrhagic cysts, MRI-derived imaging features alone possess strong discriminative power—particularly when multiple sequences are utilized. The T2-weighted sequence, due to its superior soft tissue contrast, demonstrated greater capability in extracting statistically significant features compared to T1 Fat-Sat. Even in the absence of clinical data, both the T2 sequence and the combined T1 Fat-Sat +T2 dataset demonstrated promising performance in classifying the two lesion types**.**This underscores the potential value of MRI in scenarios where clinical information is limited.

In line with this, the study by Ștefan et al. demonstrated that quantitative image-derived features—whether through radiomics in MRI or texture analysis in ultrasound—can substantially improve diagnostic precision. These findings further validate the use of quantitative features in the current study and highlight the potential of texture analysis as an effective complement to conventional visual assessment [37].

Features extracted from the T1 Fat-Sat sequence demonstrated lower statistical significance and limited discriminative power compared to those derived from T2-weighted imaging in the SVC model. In the KNN model, the combination of T1 Fat-Sat and T2 sequences resulted in improved sensitivity and F1-score, indicating better overall performance in lesion identification. Overall, these results highlight the potential of machine

learning models for the quantitative assessment of ovarian lesions.

In comparison, the study by Lupean et al. (2021) employed MRI signal intensity analysis and the T2 shading index to differentiate between endometriomas and hemorrhagic cysts. Their findings showed that two quantitative methods for assessing T2 shading achieved acceptable performance, with sensitivity of 100% and specificity of 81.82%, whereas classical subjective evaluation of this index yielded sensitivity of 92.86% and specificity of 66.67%. However, DWI and ADC parameters were found to be ineffective in distinguishing between the two lesion types [38].

The present study suggests that multi-sequence MRI-based radiomics—without the need for advanced sequences such as DWI or contrast agent administration—may offer improved diagnostic performance compared to traditional signal intensity-based methods. Moreover, machine learning models like SVC and KNN are capable of extracting complex patterns from imaging data that may not be discernible through conventional approaches.

In database searches, most prior studies have focused on CT-based radiomics for differentiating ovarian cystic lesions, particularly distinguishing dermoid cysts from endometriomas [39, 40]. In contrast, the present study aimed to differentiate endometriomas from hemorrhagic cysts using MRI-based radiomics—two lesions that, unlike dermoid cysts, exhibit highly similar imaging signal characteristics, making their differentiation more challenging. Despite this complexity, the combination of T1 Fat-Sat - and T2-weighted sequences in the SVC model yielded high AUC and specificity values, indicating satisfactory performance under more demanding diagnostic conditions. The superiority reported in previous studies should be interpreted in light of lesion type, imaging modality, and the complexity of differentiation [40].

The limited soft tissue contrast in the T1 Fat-Sat sequence led to statistical overlap between classes, and features inherently dependent on pixel intensity differences or correlated structures failed to demonstrate sufficient discriminative power. The exclusion of these features during the final selection phase was not due to algorithmic limitations, but rather reflected the inadequacy of the T1 Fat-Sat sequence in capturing relevant textural differences between endometriomas and hemorrhagic cysts. This finding underscores the importance of selecting appropriate imaging sequences when designing radiomics models and highlights that even frequently occurring features may lack classification value if not supported by adequate imaging contrast [41].

The T2-weighted sequence, by providing superior soft tissue contrast, offered a robust foundation for extracting meaningful texture features that played a decisive role in differentiating adnexal lesions. Among these, features such as sphericity, skewness, and textural measures including Inverse Difference Normalized (IDN) and Inverse Difference Moment Normalized (IDMN)—which reflect textural homogeneity—showed strong performance based on SHAP values. Additionally, sphericity, coarseness, and skewness demonstrated consistently high predictive value across the models. Due to their high discriminative power and statistical significance, these features emerged as key indicators in radiomics-based classification models, reinforcing the diagnostic value of the T2 sequence in MRI texture analysis [24, 42].

These features were not only statistically significant but also clinically aligned with established patterns of tissue behavior in the two lesion types. This outcome can be attributed to the physical and imaging characteristics of the T2-weighted sequence, which—due to its high sensitivity to soft tissue contrast—can capture subtle structural variations within lesions with greater clarity. Endometriomas typically exhibit more organized architecture, homogeneous texture, and complex cellular composition, appearing on T2 images as regions with uniform intensity and well-defined geometric shapes. In contrast, hemorrhagic cysts, due to the presence of fresh or degraded blood, often display heterogeneous signals, irregular shapes, and unstable intensity distributions. These differences are well captured by the T2 sequence, making shape features such **as** sphericity and homogeneity-related features as IDM and IDMN statistically significant and effective for classification. Therefore, the strong performance of machine learning models in T2 sequence can largely be attributed to the discriminative power of geometric and textural features which accurately reflect the underlying structural differences between lesions. Importantly, these findings address the limited biological interpretability of radiomic features highlighted in the Introduction.

The use of radiomics in MRI image analysis—particularly for differentiating adnexal lesions—can substantially enhance diagnostic validation. Unlike traditional methods that rely on subjective interpretation by radiologists, radiomics enables precise, reproducible, and data-driven evaluation through quantitative feature extraction. This high level of validation may help reduce the need for multiple imaging sessions or additional sequences in future clinical workflows, as radiomics-based machine learning models may provide useful diagnostic support even with limited data. Moreover, when models achieve high accuracy using core sequences such as T1 Fat-Sat and T2, the need for advanced or time-consuming sequences like DWI or contrast-enhanced imaging diminishes. This may potentially translate into shorter scan times, reduced use of contrast agents, and lower patient costs in future clinical applications. Additionally, reducing patient time in the MRI scanner may improve imaging system efficiency and potentially allow more patients to be accommodated in less time [43].

The present study is among the few investigations that have employed radiomics in MRI image analysis to differentiate endometriomas and hemorrhagic cysts. However, it has several limitations. First, due to its

retrospective design, there is a potential for selection bias, highlighting the need for prospective validation to confirm the findings. Second, the number of endometriomas was nearly twice that of hemorrhagic cysts, which may have influenced the study outcomes. Our dataset exhibited moderate class imbalance (99 endometriomas versus 51 hemorrhagic cysts). Synthetic oversampling techniques such as SMOTE were intentionally not applied because they may amplify existing biases, increase the risk of class overlap by generating synthetic samples near class boundaries, and introduce additional noise, particularly in relatively small datasets. These factors may result in unrealistic decision boundaries and reduced model generalizability. Instead, the original class distribution was preserved to better reflect the underlying clinical data. The relatively better performance of the T2-based KNN model may indicate greater robustness of T2-derived features in the presence of class imbalance. Third, the single-center nature and limited sample size represent additional weaknesses that restrict the generalizability of the results. In this study, we did not apply filters to extract features in order to improve the interpretability of the model , aiming for a more clinically oriented and less complex framework. While this approach may have influenced overall model performance and slightly altered its diagnostic value, it supports practical applicability and clinical relevance. Another limitation of this study is that inter-reader and intra-reader reproducibility analyses of radiomic features using metrics such as the intraclass correlation coefficient (ICC) were not performed. Future studies with independent segmentations by multiple readers and formal reproducibility assessment are warranted to further validate the robustness of the extracted radiomic features.

Given the relatively limited number of extracted radiomics features, a permissive consensus threshold was intentionally adopted to avoid excessive feature elimination.

multiple radiomic features were tested, which increases the risk of false-positive findings in the absence of formal correction for multiple comparisons. Therefore, the reported p-values should be interpreted with caution as hypothesis-generating rather than confirmatory evidence.

Conclusion

Overall, the findings of this study indicate that the T2-weighted sequence alone demonstrates acceptable performance and may be useful in scenarios with limited data availability and warrants further validation. In contrast, the T1-weighted sequence alone does not provide reliable classification capability. The combination of T1 Fat-Sat and T2 sequences yields the best performance and is therefore recommended for the development of radiomics-based models aimed at differentiating ovarian lesions. Among the classification algorithms, the SVC model outperformed KNN in terms of specificity and overall accuracy, whereas KNN demonstrated superior sensitivity.

Acknowledgment

The authors gratefully acknowledge the Research and Technology Deputy of Mashhad University of Medical Sciences for supporting this study.

References

1. Park BY, Yao R, Rossi J, Lee AW. Severe maternal morbidity associated with endometriosis: a population-based, retrospective cohort study. Fertility and Sterility. 2023 Aug 1;120(2):360-8.
2. Kalaitzopoulos DR, Mitsopoulou A, Iliopoulou SM, Daniilidis A, Samartzis EP, Economopoulos KP. Association between endometriosis and gynecological cancers: a critical review of the literature. Archives of gynecology and obstetrics. 2020 Feb;301(2):355-67.
3. Arafah M, Rashid S, Akhtar M. Endometriosis: a comprehensive review. Advances in anatomic pathology. 2021 Jan 1;28(1):30-43.
4. Kim HJ, Lee SY, Shin YR, Park CS, Kim K. The value of diffusion-weighted imaging in the differential diagnosis of ovarian lesions: A meta-analysis. PloS one. 2016 Feb 23;11(2):e0149465.
5. Tanase Y, Kawaguchi R, Takahama J, Kobayashi H. Factors that differentiate between endometriosis-associated ovarian cancer and benign ovarian endometriosis with mural nodules. Magnetic Resonance in Medical Sciences. 2018;17(3):231-7.
6. Bausic A, Coroleuca C, Coroleuca C, Comandasu D, Matasariu R, Manu A, Frincu F, Mehedintu C, Bratila E. Transvaginal ultrasound vs. magnetic resonance imaging (MRI) value in endometriosis diagnosis. Diagnostics. 2022 Jul 1;12(7).
7. Bauşic AI, Matasariu DR, Manu A, Brătilă E. Transvaginal ultrasound vs. Magnetic resonance imaging: what is the optimal imaging modality for the diagnosis of endometriosis?. Biomedicines. 2023 Sep 23;11(10):2609.
8. Franco PN, Annibali S, Viganò S, Cazzella C, Marra C, Smedile A, Bonaffini PA, Marra P, Otero García MM, Reinhold C, Sironi S. T2*-weighted imaging performance in the detection of deep endometriosis among readers with different experience: comparison with conventional MRI sequences. Diagnostics. 2022 Jun 24;12(7):1545.
9. Wang LJ, Tseng YJ, Wee NK, Low JJ, Tan CH. Diffusion-weighted imaging versus dynamic contrast-enhanced imaging for pre-operative diagnosis of deep myometrial invasion in endometrial cancer: A meta-analysis. Clinical Imaging. 2021 Dec 1;80:36-42.
10. Khan I, Gupta A, Ilyas M. T2 Shading and T2 Dark Spot Sign in the Diagnosis of Endometriotic Cysts: A Prospective Study. CHRISMED Journal of Health and Research. 2022 Jan 1;9(1):31-4.
11. Nakai Y, Maeda E. Differential Diagnosis of Ovarian Endometriotic Cysts and Hemorrhagic Functional Cysts. InMRI and CT for Decision-Making in Obstetrics and Gynecology Practice 2025 Apr 27 (pp. 105-113). Singapore: Springer Nature Singapore.
12. Corwin MT, Gerscovich EO, Lamba R, Wilson M, McGahan JP. Differentiation of ovarian

endometriomas from hemorrhagic cysts at MR imaging: utility of the T2 dark spot sign. Radiology. 2014 Apr;271(1):126-32.
13. Tomaszewski MR, Gillies RJ. The biological meaning of radiomic features. Radiology. 2021 Mar;298(3):505-16.
14. Bibbo M, Draganova-Tacheva R, Fitzpatrick BT. Peritoneal washings and ovary. InComprehensive cytopathology 2015 Jan 1 (pp. 237-245). Content Repository Only!.
15. Varghese BA, Cen SY, Hwang DH, Duddalwar VA. Texture analysis of imaging: what radiologists need to know. American Journal of Roentgenology. 2019 Mar;212(3):520-8.
16. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, Pellot-Barakat C, Soussan M, Frouin F, Buvat I. LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. Cancer research. 2018 Aug 15;78(16):4786-9.
17. Blüthgen C, Patella M, Euler A, Baessler B, Martini K, von Spiczak J, Schneiter D, Opitz I, Frauenfelder T. Computed tomography radiomics for the prediction of thymic epithelial tumor histology, TNM stage and myasthenia gravis. PloS one. 2021 Dec 20;16(12):e0261401.
18. Arlot S, Lerasle M. Choice of V for V-fold cross-validation in least-squares density estimation. Journal of Machine Learning Research. 2016;17(208):1-50.
19. Perniciano A, Loddo A, Di Ruberto C, Pes B. Insights into radiomics: impact of feature selection and classification. Multimedia Tools and Applications. 2025 Aug;84(26):31695-721.
20. Napa KK, Govindarajan R, Sathya S, Murugan JS, Vijayammal BK. Comparative analysis of explainable machine learning models for cardiovascular risk stratification using clinical data and shapley additive explanations. Intelligence-Based Medicine. 2025 Aug 13:100286.
21. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, Van Der Walt SJ. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature methods. 2020 Mar 2;17(3):261-72.
22. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J. Scikit-learn: Machine learning in Python. the Journal of machine Learning research. 2011 Nov 1;12:2825-30.
23. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Advances in neural information processing systems. 2017;30.
24. Lupean RA, Ștefan PA, Csutak C, Lebovici A, Măluțan AM, Buiga R, Melincovici CS, Mihu CM. Differentiation of endometriomas from ovarian hemorrhagic cysts at magnetic resonance: the role of texture analysis. Medicina. 2020 Sep 23;56(10):487.
25. Mayerhoefer ME, Breitenseher M, Amann G, Dominkus M. Are signal intensity and homogeneity useful parameters for distinguishing between benign and malignant soft tissue masses on MR images?: Objective evaluation by means of texture analysis. Magnetic resonance imaging. 2008 Nov 1;26(9):1316-22.
26. Takeyama N, Sasaki Y, Ueda Y, Tashiro Y, Tanaka E, Nagai K, Morioka M, Ogawa T, Tate G, Hashimoto T, Ohgiya Y. Magnetic resonance imaging-based radiomics analysis of the differential diagnosis of ovarian clear cell carcinoma and endometrioid carcinoma: a retrospective study. Japanese Journal of Radiology. 2024 Jul;42(7):731-43.
27. Balaban M, Idilman IS, Toprak H, Unal O, Ipek A, Kocakoc E. The utility of diffusion-weighted magnetic resonance imaging in differentiation of endometriomas from hemorrhagic ovarian cysts. Clinical imaging. 2015 Sep 1;39(5):830-3.
28. Lee NK, Kim S, Kim KH, Suh DS, Kim TU, Han GJ, Lee JW, Kim JY. Diffusion-weighted magnetic resonance imaging in the differentiation of endometriomas from hemorrhagic cysts in the ovary. Acta radiologica. 2016 Aug;57(8):998-1005.
29. Bulut E, Peker M, Kupeli A, Danisan G, Bulut AC. The efficiency of susceptibility-weighted MRI in the differentiation of endometriomas from haemorrhagic ovarian cysts. Abdominal radiology. 2021 Nov;46(11):5337-43.
30. Bendifallah S, Puchar A, Suisse S, Delbos L, Poilblanc M, Descamps P, Golfier F, Touboul C, Dabi Y, Daraï E. Machine learning algorithms as new screening approach for patients with endometriosis. Scientific reports. 2022 Jan 12;12(1):639.
31. Chen T, Wei JL, Leng T, Gao F, Hou SY. The diagnostic value of the combination of hemoglobin, CA199, CA125, and HE4 in endometriosis. Journal of clinical laboratory analysis. 2021 Sep;35(9):e23947.
32. Chiappa V, Bogani G, Interlenghi M, Salvatore C, Bertolina F, Sarpietro G, Signorelli M, Castiglioni I, Raspagliesi F. The Adoption of Radiomics and machine learning improves the diagnostic processes of women with Ovarian MAsses (the AROMA pilot study). Journal of Ultrasound. 2021 Dec;24(4):429-37.
33. Hu Y, Weng Q, Xia H, Chen T, Kong C, Chen W, Pang P, Xu M, Lu C, Ji J. A radiomic nomogram based on arterial phase of CT for differential diagnosis of ovarian cancer. Abdominal Radiology. 2021 Jun;46(6):2384-92.
34. Li X, Ma Q, Tao C, Liu J, Nie P, Dong C. A CT-based radiomics nomogram for differentiation of small masses (< 4 cm) of renal oncocytoma from clear cell renal cell carcinoma. Abdominal Radiology. 2021 Nov;46(11):5240-9.
35. Qi L, Chen D, Li C, Li J, Wang J, Zhang C, Li X, Qiao G, Wu H, Zhang X, Ma W. Diagnosis of ovarian neoplasms using nomogram in combination with ultrasound image-based radiomics signature and clinical factors. Frontiers in genetics. 2021 Sep 28;12:753948.
36. Zheng Y, Wang H, Li Q, Sun H, Guo L. Discriminating between benign and malignant solid ovarian tumors based on clinical and radiomic features of MRI. Academic radiology. 2023 May 1;30(5):814-22.
37. Ștefan RA, Ștefan PA, Mihu CM, Csutak C, Melincovici CS, Crivii CB, Maluțan AM, Hîțu L, Lebovici A. Ultrasonography in the differentiation of endometriomas from hemorrhagic ovarian cysts:

the role of texture analysis. Journal of Personalized Medicine. 2021 Jun 28;11(7):611.
38. Lupean RA, Ştefan PA, Lebovici A, Csutak C, Rusu GM, Mihu CM. Differentiation of endometriomas from hemorrhagic cysts at magnetic resonance: The role of quantitative signal intensity measurements. Current Medical Imaging. 2021 Apr 1;17(4):524-31.
39. Liu L, Cai W, Zhou C, Tian H, Wu B, Zhang J, Yue G, Hao Y. Ultrasound radiomics-based artificial intelligence model to assist in the differential diagnosis of ovarian endometrioma and ovarian dermoid cyst. Frontiers in Medicine. 2024 Mar 8;11:1362588.
40. Li J, Wang F, Ma J, Zhang Z, Zhang N, Cui S, Ye Z. A CT-based radiomics nomogram for differentiating ovarian cystadenomas and endometriotic cysts. Clinical Radiology. 2023 Sep 1;78(9):e635-43.
41. Ghosh A, Serai SD. Radiomics and Texture Analysis. InAdvanced Clinical MRI of the Kidney: Methods and Protocols 2023 Nov 18 (pp. 407-418). Cham: Springer International Publishing.
42. Cansu A, Bulut E, Dinc G, Bekircavusoglu S, Eyuboglu I, Guven ES, Ahmetoglu A. Diagnostic efficacy of T2 Dark Spot, T2 dark rim signs, and T2 shading on magnetic resonance imaging in differentiating endometriomas from hemorrhagic cysts. Journal of computer assisted tomography. 2019 Jul 1;43(4):619-22.
43. Guo SW, Ding D, Shen M, Liu X. Dating endometriotic ovarian cysts based on the content of cyst fluid and its potential clinical implications. Reproductive Sciences. 2015 Jul;22(7):873-83.