# Computed tomography-based radiomics machine learning models for differentiating enchondroma and atypical cartilaginous tumor in long bones

## Computertomographie-basierte Radiomics Machine Learning Modelle zur Unterscheidung von Enchondromen und atypischen Knorpeltumoren in Röhrenknochen

Authors
Rui Hong[1], Qian Li[1], Jielin Ma[2], Chunmiao Lu[1], Zhiwei Zhong[1]

Affiliations
1 Radiology, The Third Hospital of Hebei Medical University, Shijiazhuang, China
2 Oncology, The Third Hospital of Hebei Medical University, Shijiazhuang, China

Keywords
Radiomics, Machine learning, Enchondroma, Atypical cartilaginous tumor, Long bone, Computed tomography

received 9.4.2024
accepted after revision 3.6.2024
published online 2024

Bibliography
Fortschr Röntgenstr
DOI 10.1055/a-2344-5398
ISSN 1438-9029

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Correspondence
Dr. Zhiwei Zhong
Radiology, The Third Hospital of Hebei Medical University, Ziqiang Road, 050000 Shijiazhuang, China
zhongzhw@sina.com

## ABSTRACT

**Purpose** To explore the value of CT-based radiomics machine learning models for differentiating enchondroma from atypical cartilaginous tumor (ACT) in long bones and methods to improve model performance.

**Materials and Methods** 59 enchondromas and 53 ACTs in long bones confirmed by pathology were collected retrospectively. The features were extracted from preoperative CT images of these patients, and least absolute shrinkage and selection operator (LASSO) regression was used for feature selection and dimensionality reduction. The selected features were used to construct classification models by thirteen machine learning algorithms. The data set was randomly divided into a training set and a test set at a proportion of 7:3 by ten-fold cross-validation to evaluate the performance of these models.

**Results** A total of 1199 features were extracted, 9 features were selected, and 13 radiomics machine learning models were constructed. The area under the curve (AUC) of 11 models was more than 0.8, and that of 3 models was more than 0.9. The Extremely Randomized Trees model achieved the best performance (AUC = 0.9375 ± 0.0884), followed by the Adaptive Boosting model (AUC = 0.9188 ± 0.1010) and the Linear Discriminant Analysis model (AUC = 0.9062 ± 0.1459).

**Conclusion** CT-based radiomics machine learning models had great ability to distinguish enchondroma and ACT in long bones. By using filters to deeply mine high-order features in the original image and selecting appropriate machine learning algorithms, the performance of the model can be improved.

**Key points**

- CT-based radiomics machine learning models can distinguish enchondroma and ACT in long bones.
- Using filters and selecting advanced machine learning algorithms can improve model performance.
- Clinical features have limited utility in distinguishing enchondroma and ACT in long bones.

**Citation Format**

- Hong R, Li Q, Ma J et al. Computed tomography-based radiomics machine learning models for differentiating enchondroma and atypical cartilaginous tumor in long bones. Fortschr Röntgenstr 2024; DOI 10.1055/a-2344-5398

## ZUSAMMENFASSUNG

**Ziel** Untersuchung des Nutzens maschineller Lernmodelle für CT-basierte Radiomics bei der Unterscheidung von Enchondromen und atypischen Knorpeltumoren (ACT) in langen Knochen und Methoden zur Verbesserung der Modellleistung.

**Materialien und Methoden** 59 Enchondrome und 53 ACTs in langen Knochen, die histopathologisch bestätigt wurden, wurden retrospektiv erhoben. Die Merkmale wurden aus präoperativen CT-Bildern dieser Patienten extrahiert, und die Least

Absolute Shrinkage and Selection Operator (LASSO)-Regression wurde zur Merkmalsauswahl und Dimensionalitätsreduktion verwendet. Die ausgewählten Funktionen wurden verwendet, um Klassifizierungsmodelle mit dreizehn Algorithmen des maschinellen Lernens zu konstruieren. Der Datensatz wurde zufällig in einen Trainingssatz und einen Testsatz mit einem Anteil von 7 × 3 durch zehnfache Kreuzvalidierung unterteilt, um die Leistung dieser Modelle zu bewerten.

**Ergebnisse** Insgesamt wurden 1199 Merkmale extrahiert, neun Merkmale ausgewählt und dreizehn Modelle für maschinelles Lernen in der Radiomik konstruiert. Die Fläche unter der Kurve (AUC) von elf Modellen betrug mehr als 0,8 und die von drei Modellen mehr als 0,9. Das Extremely Randomized Trees-Modell erzielte die beste Leistung (AUC = 0,9375 ± 0,0884), gefolgt von dem Adaptiven Boosting-Modell (AUC = 0,9188 ± 0,1010) und dem Linear Discriminant Analysis-Modell (AUC = 0,9062 ± 0,1459).

**Schlussfolgerung** CT-basierte Radiomics Machine Learning-Modelle hatten eine große Fähigkeit, ein Enchondrom und von einem ACT in langen Knochen zu unterscheiden. Durch die Verwendung von Filtern, um hochrangige Features im Originalbild tief abzubauen und geeignete Algorithmen für maschinelles Lernen auszuwählen, kann die Leistung des Modells verbessert werden.

**Kernaussagen**

- CT-basierte Radiomics Machine Learning-Modelle können Enchondrome und ACTs in langen Knochen unterscheiden.
- Die Verwendung von Filtern und die Auswahl fortschrittlicher Algorithmen für maschinelles Lernen können die Modellleistung verbessern.
- Klinische Merkmale haben begrenzten Nutzen bei der Unterscheidung eines Enchondroms von einem ACT in Röhrenknochen.

## Introduction

The most common intermediate (locally aggressive) chondrogenic tumor is the atypical cartilaginous tumor (ACT), which has been specifically found to occur in the appendicular skeletons (long and short tubular bones) according to the 2020 World Health Organization (WHO) classification of bone tumors [1, 2]. Enchondroma is the most common benign chondrogenic tumor [2]. Most of them are accidentally discovered without obvious symptoms, and with the widespread use of MRI, the incidence of them being accidentally found in long bones was higher than in short bones [3]. The typical imaging features are round osteolysis with popcorn-like calcifications in the medullary cavity of the metaphysis of long bones [2]. However, most enchondroma patients can choose regular surveillance over surgery, while the main treatments of ACT are surgical intralesional curettage and filling of the tumor cavity. The probability of local recurrence is about 7.5–11 % with only few metastases [1]. Therefore, it is important to improve the accuracy of identifying enchondroma and ACT.

Radiomics is a relatively objective method that can identify tumor heterogeneity and reflect potential structural and functional information by extracting quantitative features with high throughput from standard images and utilizing machine learning algorithms for mathematical operations [4]. MRI-based radiomics models have achieved great results regarding the differentiation of chondrogenic tumors in long bones [5]. When it comes to demonstrating osteolysis and calcification of chondrogenic tumors, CT is more advantageous than MRI. Nevertheless, there aren't many studies using CT-based radiomics models to identify chondrogenic tumors in long bones. The CT-based radiomics machine learning model developed by Gitto et al. [6] performed admirably with regard to distinguishing ACT from high-grade chondrosarcoma in long bones, but the validation set only included CT images from PET-CT examination. Deng et al. [7] developed the CT-based texture analysis model to classify enchondroma and low-grade chondrosarcoma in long bones due to the limited number of included patients and extracted features, resulting in low accuracy in the model.

The aim of this study is to explore the value of CT-based radiomics machine learning models for distinguishing enchondroma from ACT in long bones and methods to improve model performance.

## Materials and Methods

### Patient selection

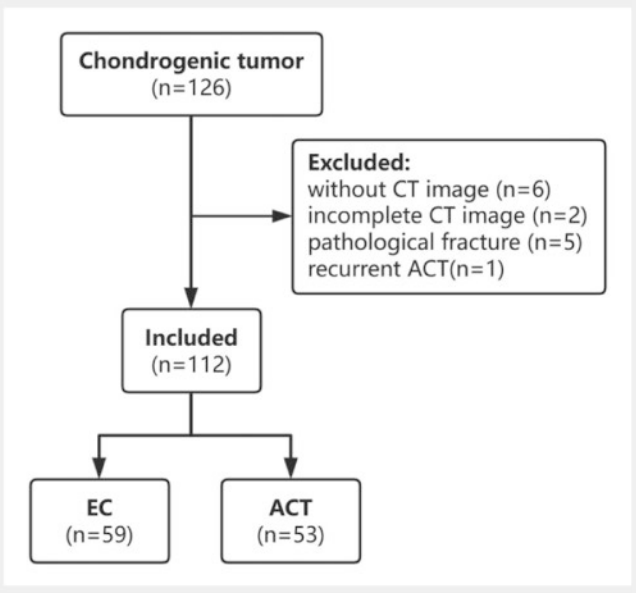
Approval from the Institutional Review Board was obtained, and in keeping with the policies for a retrospective review, informed consent was not required. Inclusion criteria: 1) enchondroma and ACT in long bones confirmed by pathology; 2) CT performed within 1 month before pathology. Exclusion criteria: 1) complicated with pathological fracture; 2) no first CT scans of recurrent ACT; 3) radiotherapy or chemotherapy before CT scans; 4) secondary ACT (▶ **Fig. 1**).

### Image acquisition

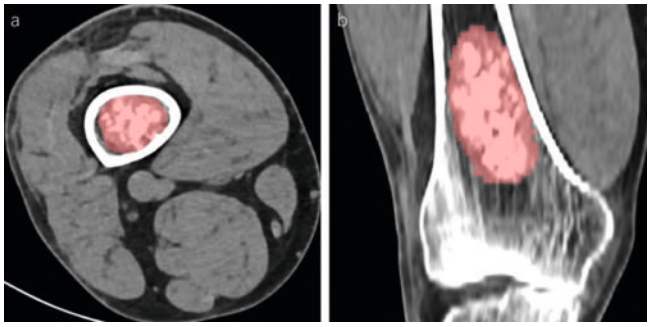
All enrolled patients underwent a multidetector row CT examination (Philips IQon Spectral CT; Siemens Somatom Definition AS 128; Siemens Somatom Sensatim 64). The CT scan parameters were: voltage: 120 kV; variable tube current; slice thickness: 0.65 to 1 mm; matrix: 512 × 512; field of view: from 120 × 120 mm to 400 × 400 mm.

### Image segmentation

Two musculoskeletal radiologists, observer 1 with three years of experience and observer 2 with fifteen years of experience, used the Research Oncology Suite of IntelliSpace Discovery (ISD, Version 3.0, Philips Healthcare, The Netherlands) to perform semi-automatic 3 D volume of interest (VOI) segmentation based on threshold intensity (▶ **Fig. 2**). The Interclass Correlation Coeffi-




▶ **Fig. 1** Flowchart of patient selection. EC: enchondroma; ACT: atypical cartilaginous tumor.



▶ **Fig. 2** Volume of interest (VOI) segmentation of a 51-year-old man with an atypical cartilaginous tumor in the distal femur and red areas of **a** (axial) and **b** (coronal) are masked.

cient (ICC) was used to evaluate the repeatability of VOI segmentation between observers. The CT images of 30 patients (enchondroma = 16, ACT = 14) were randomly selected from all patients with chondrogenic tumors by observer 2 for repeatability verification. When the ICC value of the feature was ≥ 0.90, it was considered to be a stable feature with great repeatability, and the follow-up steps were continued.

### Image Preprocessing and Feature extraction

The PyRadiomics plugin of ISD was used for image preprocessing and feature extraction. CT images were resampled at a spatial resolution of voxels with a size of 1 mm × 1 mm × 3 mm and discretized with a fixed bin width of 25HU [8]. All radiomic features were extracted with high throughput from the original and the filtered image, including first-order statistics, shape-based (3D), gray level co-occurrence matrix (GLCM), gray level size zone matrix (GLSZM), and gray level run length matrix (GLRLM). Filters included logarithm, exponential, square, square root, wavelet, and Laplacian of Gaussian (LoG).

### Feature selection and model development

R software (Version 2023.12.0 + 369) was used for feature selection and dimensionality reduction, and the IntelliSpace Medicina Scientia Research Platform (ISMS, Version 3.0, Philips Healthcare, The Netherlands) was used for model development. After standardizing the features, ICC, t-tests, and least absolute shrinkage and selection operator (LASSO) regression were used for feature selection and dimensionality reduction in R software. The selected features were imported into ISMS, utilizing 13 machine learning algorithms to develop the radiomics model. To evaluate the performance of these models using ten-fold cross-validation, the data set was randomly divided into a training set (n = 78) and a test set (n = 34) at a 7:3 proportion in each cross-validation.

### Statistical analysis

All statistical analyses were performed with SPSS (Version 27.0), R software, and ISMS. The statistically significant level was set to a two-sided P-value < 0.05. The T-test was used to analyze continuous variables, and the chi-square test was used to compare categorical variables. Indicators for evaluating the performance of the model included the area under the curve (AUC), accuracy (ACC), recall, precision, F1 score, Kappa, and Matthews correlation coefficient (MCC). The receiver operating characteristic curve (ROC), precision-recall curve (P-R), confusion matrix, and feature importance plot were drawn.

## Results

Finally, 112 patients with enchondroma (n = 59) or ACT (n = 53) in long bones met the inclusion and exclusion criteria. The age of patients with enchondroma (42.19 ± 18.30) was less than ACT (50.42 ± 11.94) (p = 0.005), and there was no significant difference between sex (p = 0.480) and tumor location (p = 0.909) (▶ **Table 1**).

A total of 1199 radiomic features were extracted, including 1172 stable features with an ICC ≥ 0.90, then 388 features with statistical significance (P < 0.05) by t-test, and finally nine most valuable features by LASSO regression for selection and dimensionality reduction (▶ **Fig. 3** and ▶ **Table 2**).

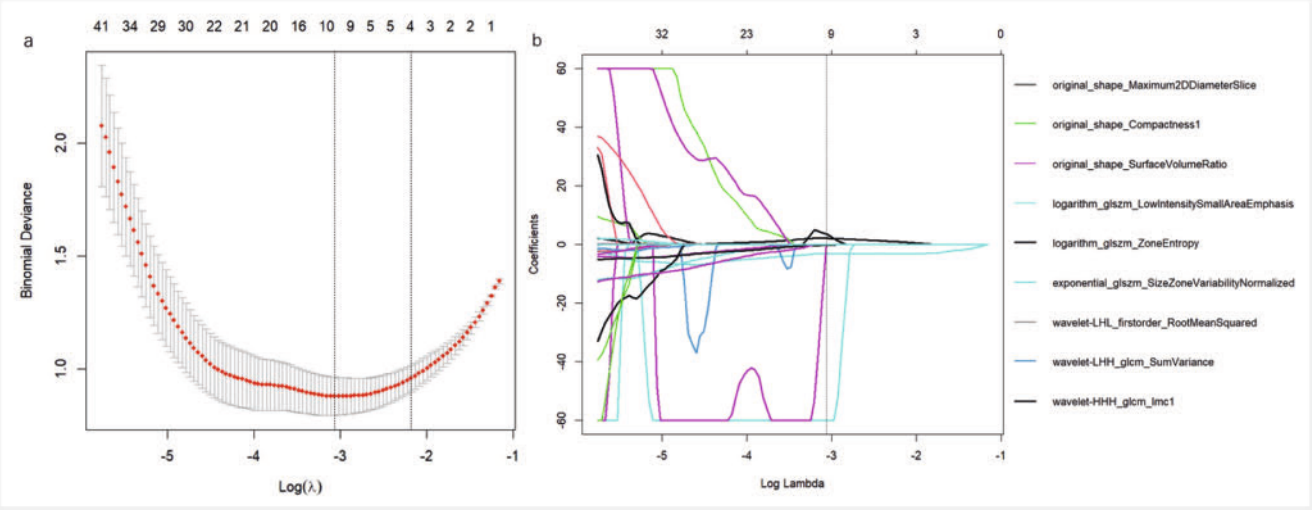
Among the 13 models constructed, eleven models had AUC values above 0.8 and three models above 0.9 (▶ **Table 3**). The Extremely Randomized Trees (ERT) model had the best performance (AUC = 0.9375 ± 0.0884, ACC = 0.8500 ± 0.1225), followed by the Adaptive Boosting (ADA) model (AUC = 0.9188 ± 0.1010, ACC = 0.8732 ± 0.0970), and the Linear Discriminant Analysis (IDA) model (AUC = 0.9062 ± 0.1459, ACC = 0.8500 ± 0.1346). The ROC curve, P-R curve, confusion matrix, and feature importance plot were drawn according to the analysis results of the ERT model, ADA model, and IDA model shown in ▶ **Fig. 4**, ▶ **Fig. 5** and ▶ **Fig. 6**.

According to the feature importance plot, the most important feature in the ERT model was the zone entropy (ZE) of the GLSZM

▶ **Table 1** Patient clinical data.

| | | Enchondroma (n = 59) | ACT (n = 53) | P-value | t/$\chi^2$ value |
|---|---|---|---|---|---|
| Age (years) | | | | 0.005 | −2.845* |
| | Mean ± SD | 42.19 ± 18.30 | 50.42 ± 11.94 | | |
| Sex | | | | 0.480 | 0.498** |
| | Male | 25 (42.37%) | 19 (35.85%) | | |
| | Female | 34 (57.62%) | 34 (64.15%) | | |
| Location | | | | 0.909 | 1.536** |
| | Proximal humerus | 15 (25.42%) | 14 (26.42%) | | |
| | Proximal femur | 10 (16.95%) | 10 (18.87%) | | |
| | Distal femur | 22 (37.29%) | 22 (41.51%) | | |
| | Proximal tibia | 7 (11.86%) | 4 (7.55%) | | |
| | Distal tibia | 2 (3.39%) | 2 (3.77%) | | |
| | Proximal fibula | 3 (5.08%) | 1 (1.89%) | | |

Note: In the column of the t/$\chi^2$ value, * is the t value and ** is the $\chi$ value. SD: standard deviation; ACT: atypical cartilaginous tumor




▶ **Fig. 3** Least absolute shrinkage and selection operator (LASSO) regression results. **a** Cross-validation curve, the left dotted line represents the minimum binomial deviance corresponding to the logarithm of the penalty coefficient (λ) (Log Lambda.min); the right dotted line symbolizes the Log Lambda.min plus one standard error corresponding to Log ambda.1se. **b** Coefficients path diagram, each line denotes a feature, and the coefficients of features tend to be sparse (0) as the logarithm of the penalty coefficient (λ) increases; the dotted line corresponds to Log Lambda.min, and legends are the nine features selected at Log Lambda.min.

feature from the logarithm-filtered image. By t-test again, the mean of the ZE in ACT (7.4503 ± 0.1799) was higher than in enchondroma (7.2035 ± 0.2375) ($P < 0.05$). In the ADA model, the most important feature was the surface area to volume ratio (SA/V) of the shape-based (3D) feature from the original image. The mean of the SA/V in ACT (0.2869 ± 0.0855) was discovered to be lower by t-test than in enchondroma (0.4782 ± 0.1421) ($P < 0.05$). The small area low gray level emphasis (SALGLE) of the GLSZM feature from the logarithm-filtered image was the most significant feature in the IDA model. The mean of the SALGLE in ACT (0.0014 ± 0.0010) was higher than in enchondroma (0.006 ± 0.003) by t-test once again ($P < 0.05$).

## Discussion

This study used a variety of filters to extract high-order radiomic features from CT images and constructed radiomics models utilizing 13 machine learning algorithms to identify the enchondroma and ACT in long bones. The results revealed that the AUC value of

▶ **Table 2** The most valuable feature.

| Feature name | Feature class | Source image |
|---|---|---|
| Maximum 2 D diameter (slice) | Shape (3 D) | Original |
| Compactness 1 | Shape (3 D) | Original |
| Surface area to volume ratio | Shape (3 D) | Original |
| Small area low gray level emphasis | GLSZM | Logarithm |
| Zone entropy | GLSZM | Logarithm |
| Size-zone non-uniformity normalized | GLSZM | Exponential |
| Root mean squared | First order | Wavelet (low-high-low pass filter) |
| Sum variance | GLCM | Wavelet (low-high-high pass filter) |
| Informational measure of correlation 1 | GLCM | Wavelet (high-high-high pass filter) |

Note: GLSZM: gray level size zone matrix; GLCM, gray level co-occurrence matrix

▶ **Table 3** Performance of radiomics machine learning models.

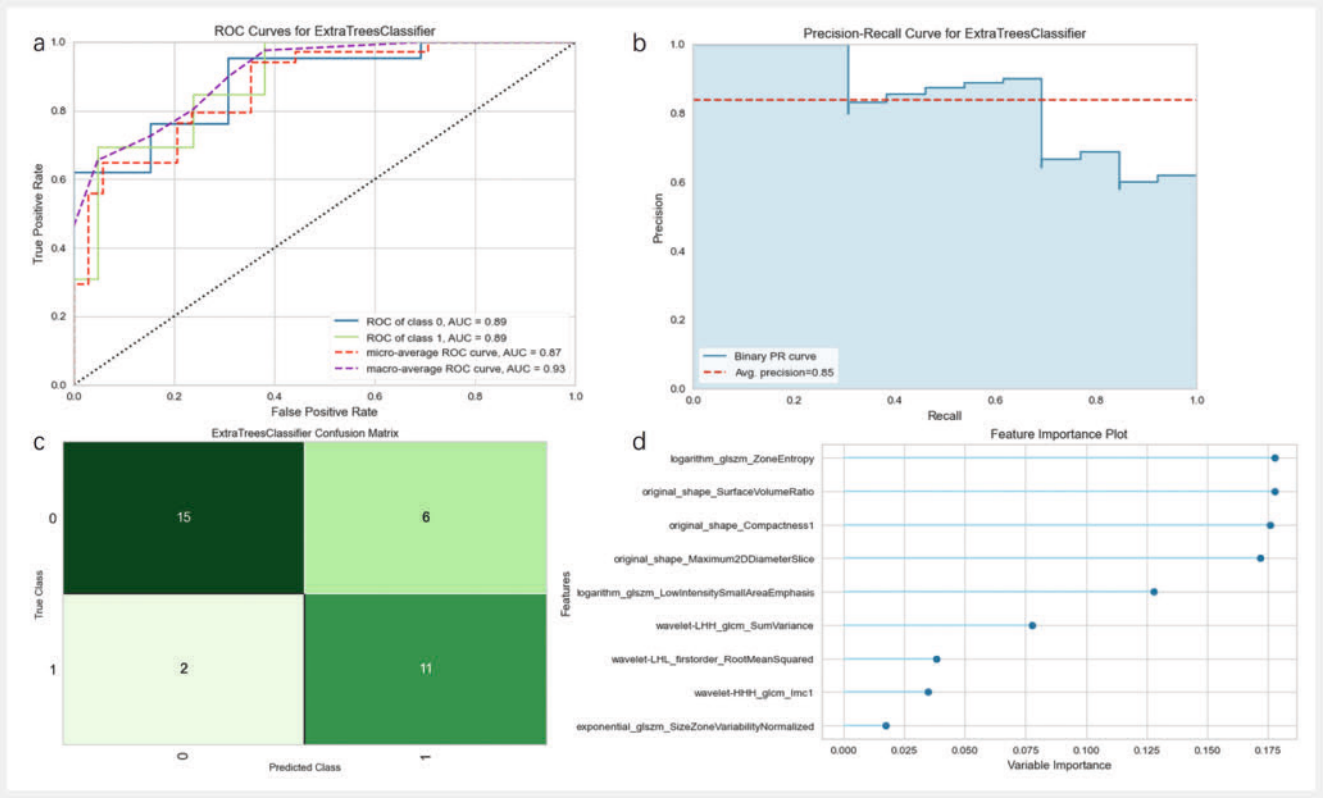
| Model | AUC | ACC | Recall | Precision | F1 score | Kappa | MCC |
|---|---|---|---|---|---|---|---|
| Extremely Randomized Trees | 0.9375 | 0.8500 | 0.8750 | 0.8433 | 0.8506 | 0.7000 | 0.7159 |
| Adaptive Boosting | 0.9188 | 0.8732 | 0.9000 | 0.8867 | 0.8689 | 0.7446 | 0.7786 |
| Linear Discriminant Analysis | 0.9062 | 0.8500 | 0.8750 | 0.8400 | 0.8467 | 0.7000 | 0.7192 |
| Random Forest | 0.8938 | 0.8357 | 0.8500 | 0.8350 | 0.8328 | 0.6720 | 0.6893 |
| Gradient Boosting Classifier | 0.8875 | 0.8607 | 0.8750 | 0.8767 | 0.8602 | 0.7220 | 0.7454 |
| Naive Bayes | 0.8854 | 0.8482 | 0.8750 | 0.8433 | 0.8506 | 0.6946 | 0.7115 |
| Logistic Regression | 0.8854 | 0.8357 | 0.8500 | 0.8367 | 0.8373 | 0.6696 | 0.6796 |
| Light Gradient Boosting Machine | 0.8792 | 0.8446 | 0.8500 | 0.8633 | 0.8475 | 0.6887 | 0.7051 |
| Quadratic Discriminant Analysis | 0.8521 | 0.8089 | 0.8250 | 0.8283 | 0.8090 | 0.6141 | 0.6406 |
| Decision Tree | 0.8510 | 0.8357 | 0.8250 | 0.8617 | 0.8290 | 0.6696 | 0.6873 |
| Support Vector Machine | 0.8152 | 0.8089 | 0.8000 | 0.8367 | 0.8087 | 0.6166 | 0.6320 |
| Multilayer perceptron | 0.7875 | 0.7714 | 0.7250 | 0.8350 | 0.7394 | 0.5391 | 0.5784 |
| K-Nearest Neighbor | 0.6948 | 0.6536 | 0.4750 | 0.7583 | 0.5626 | 0.3077 | 0.3374 |

Note: Sorted by AUC value in descending order. All values are the mean of the 10-fold cross-validation results. AUC: area under the curve; ACC: accuracy; MCC: Matthews correlation coefficient

eleven models was more than 0.8 and that of three models was more than 0.9.

In the clinical data, the age of patients with enchondroma is lower than in ACT. Although some researchers have pointed out that most chondrosarcoma patients are older than enchondroma patients [9], in this study the difference in mean age between enchondroma and ACT patients is just eight years with a sizable standard deviation. Consequently, we propose that age has limited utility in the differential diagnosis of ACT and enchondroma in long bones. There were no significant differences in terms of sex or tumor location between ACT and enchondroma in long bones in this study. Pan et al. [10] found that the location of chondrogenic tumors was the most significant clinical risk factor. In their study, there were notably more enchondroma patients in short bones than chondrosarcoma, and the number of chondrosarcoma patients in the pelvis was much higher than enchondroma, while there was no significant difference in the long bones. According to the 2020 WHO classification of bone tumors [1], ACT has been specifically referred to as an intermediate chondrogenic tumor that occurs in the long and short tubular bones. Therefore, we think that the long bone is the most valuable location for differential diagnosis between enchondroma and ACT.

Among the nine most valuable features selected, three shape-based (3 D) features were extracted from the original image, and




▶ **Fig. 4** The ability of the Extremely Randomized Trees model to correctly classify an enchondroma (label = 0, in figure) and atypical cartilaginous tumor (ACT, label = 1, in figure) in long bones in the test set. **a** Receiver operating characteristic curve (ROC), the blue solid line represents the ROC curve and corresponding area under the curve (AUC) for correctly classified enchondroma, the green solid line represents the ROC curve and corresponding AUC for correctly classified ACT, the red dotted line represents the micro-average ROC curve and corresponding AUC, and the purple dotted line represents the macro-average ROC curve and corresponding AUC; **b** Precision-recall curve (P-R), the blue solid line represents the binary P-R curve and AUC, and the red dotted lines represent the average precision; **c** Confusion matrix, the rows represent the enchondroma and ACT of the actual classification, and the columns represent enchondroma and ACT of model prediction classification; **d** Feature importance plot, the importance weight ranking of each feature for correct classification of enchondroma and ACT in this model.

six high-order features were extracted from the filtered image. Because shape features are independent of the gray value of the voxel, they can only be extracted from the original image, while other features are calculated based on the gray value of the voxel and can be extracted from both the original image and the filtered image [4]. By using a variety of filters, plenty of high-order features were extracted, and more valuable features were used to construct models than in previous research [7]. Compared with the logistic regression (LR) model in their research, it had a higher AUC value and accuracy in this study, and there were other models that performed better than the LR model. The performance of the radiomics model can be improved by increasing the sample size, using filters to deeply mine high-order features from the original image, and selecting appropriate machine learning algorithms.

In the feature importance plot of the ERT model, the most important feature was the ZE:

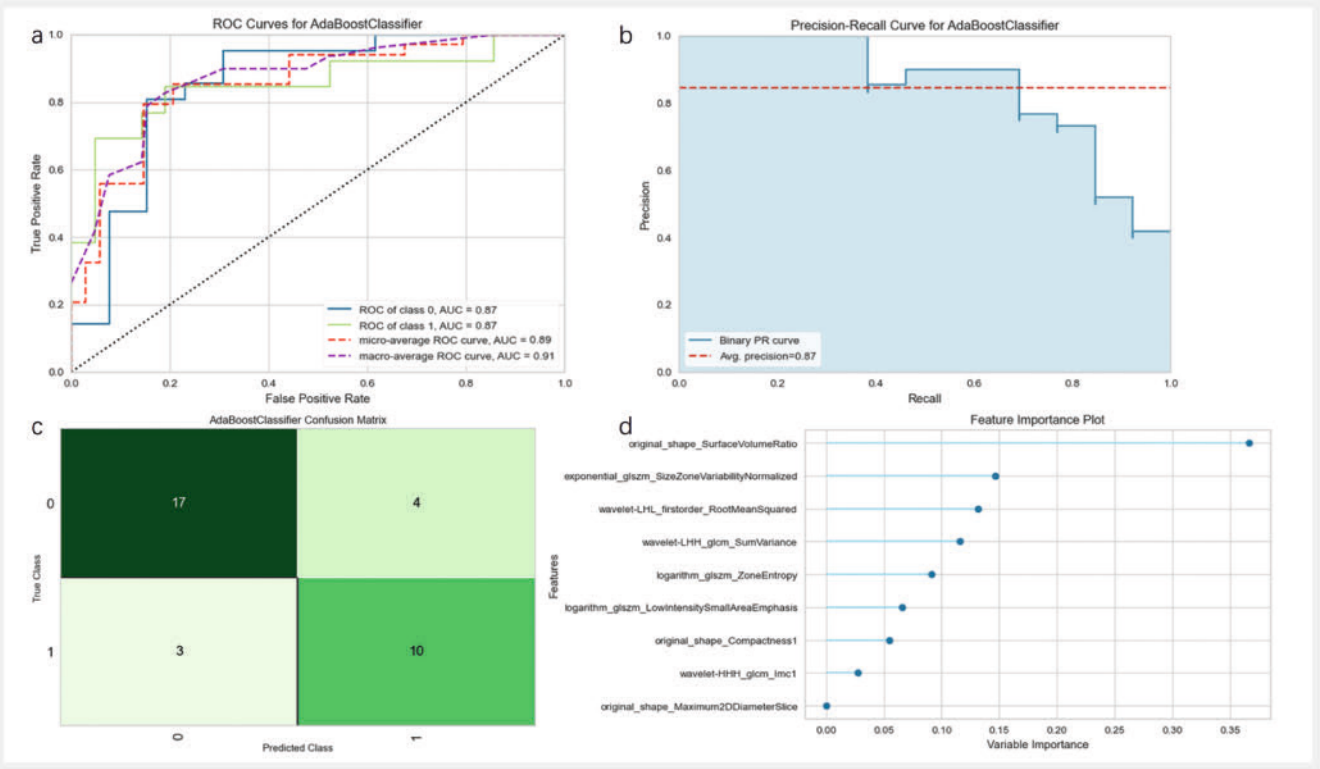
$$\textit{Zone Entropy} = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_s} P(i,j)\log_2(P(i,j)+\epsilon)$$

ZE measures the uncertainty or randomness in the distribution of zone sizes and gray levels, and the higher the value, the higher the heterogeneity in the texture patterns [4]. By t-test again, the mean of the ZE in ACT was higher than in enchondroma, indicating that the heterogeneity of ACT in long bones is higher than that of enchondroma. The most important feature in the ADA model was the SA/V:

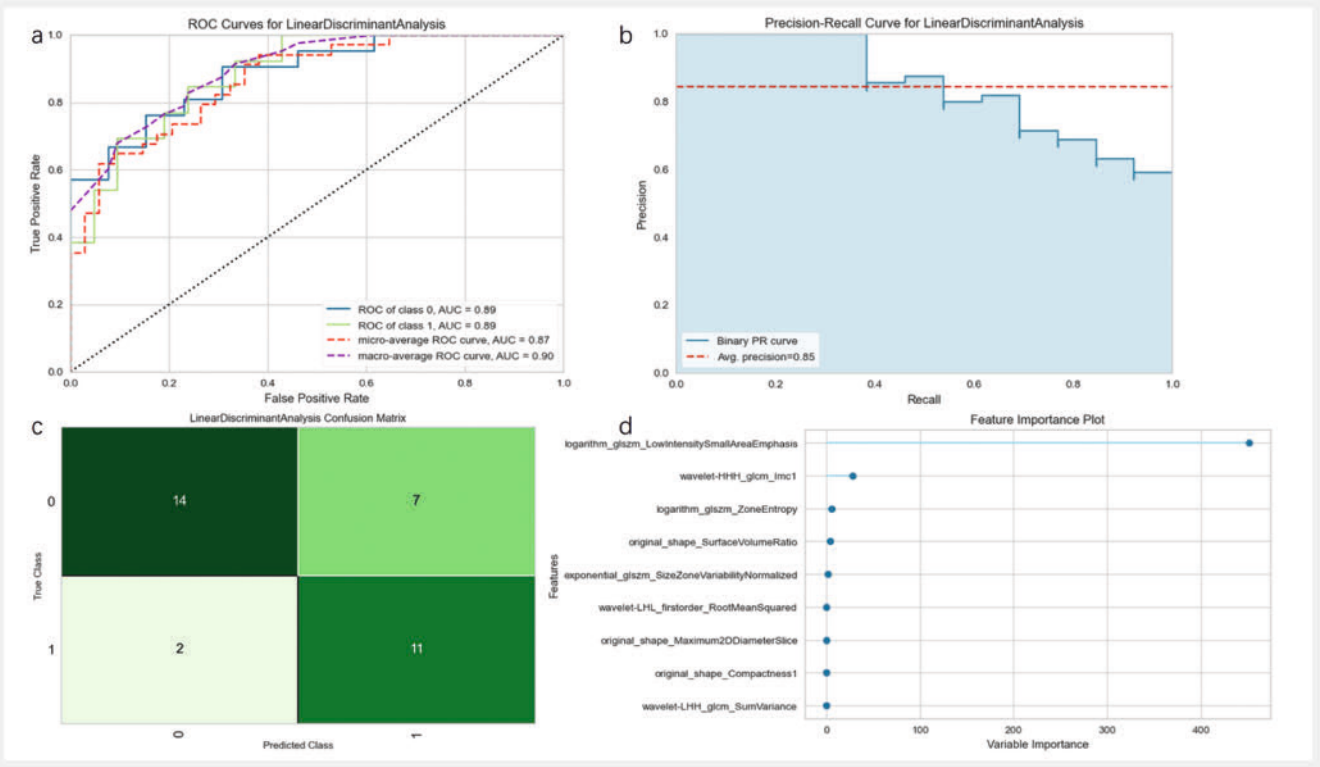
$$\textit{Surface Area to Volume ratio} = \frac{A}{V}$$

This feature is dependent on the volume of the segmented VOI, and a lower value indicates that the shape of the segmented VOI is closer to a sphere [8]. The mean of the SA/V in ACT was discovered to be lower by t-test than in enchondroma, meaning that ACT has a more spherical-shaped volume in long bones than enchondroma, which may be related to the former's more locally aggressive growth tendency [11]. The SALGLE was the most significant feature in the IDA model:

$$\textit{Small Area Low Gray Level Emphasis} = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_s}\frac{P(i,j)}{i^2j^2}}{N_z}$$




▶ **Fig. 5** The ability of the Adaptive Boosting model to correctly classify an enchondroma (label = 0, in figure) and atypical cartilaginous tumor (label = 1, in figure) in long bones in the test set. **a** Receiver operating characteristic curve; **b** Precision-recall curve; **c** Confusion matrix; **d** Feature importance plot.




▶ **Fig. 6** The ability of the Linear Discriminant Analysis model to correctly classify an enchondroma (label = 0, in figure) and atypical cartilaginous tumor (label = 1, in figure) in long bones in the test set. **a** Receiver operating characteristic curve; **b** Precision-recall curve; **c** Confusion matrix; **d** Feature importance plot.

SALGLE measures the proportion of joint distribution of smaller size zones with lower gray values in the image. The larger its value, the greater the prevalence of these areas in the image and the more uneven the distribution of these areas [8]. The mean of the SALGLE in ACT was higher than in enchondroma by t-test once again, which indicates that ACT in long bones may be more prone to small zone necrosis, leading to an uneven density of the image compared to enchondroma.

The ERT model, ADA model, IDA model, random forest (RF) model, and gradient boosting classifier (GBC) model are the first few models that performed well in this study. They all utilize ensemble learning algorithms. Ensemble learning algorithms integrate many different machine learning algorithms in order to construct multiple models to enhance prediction accuracy and reduce generalization errors [12]. Ishaq et al. [13] used nine machine learning algorithms to construct machine learning models for predicting the survival of 299 patients with heart failure. The results showed that the ERT model achieved the best accuracy (ACC = 0.9262), followed by the RF model (ACC = 0.9188), the ADA model (ACC = 0.8852), and the GBC model (ACC = 0.8852), which performed better than the decision tree model (ACC = 0.8778) and the LR model (ACC = 0.8442). Erdem et al. [14] asked two researchers to extract features from MRI images and utilize seven machine learning algorithms to construct radiomics models to classify enchondroma (n = 57) and chondrosarcoma (n = 31). The results showed that when using all features to construct models, both researchers found the best model to be the neural network (NN) model (AUC = 0.979, AUC = 0.984, respectively). When using selected features to construct models, the best model for the two researchers was the GBC (AUC = 0.990) model and the NN model (AUC = 0.979). The NN algorithm is the fundamental algorithm of deep learning that can automatically identify specific structures when combined with machine learning, but it generally requires a large sample size [15]. Combined with the findings of these studies, it is found that utilizing advanced machine learning algorithms such as ensemble learning algorithms and deep learning algorithms can improve the prediction performance of models.

There are some limitations: First, the number of patients is relatively small due to the low incidence rate of enchondroma and ACT, the absence of treatment for most enchondroma patients, and the limiting location and grading of chondrogenic tumors. Second, this study did not employ MRI images of chondrogenic tumor patients to construct models concurrently, making it impossible to directly assess the performance differences between CT-based and MRI-based radiomics machine learning models.

## Conclusion

This study found that CT-based radiomics machine learning models have great ability to distinguish enchondroma from ACT in long bones. The prediction performance of the model can be improved by using filters to deeply mine high-order features from the original image and selecting advanced machine learning algorithms.

## Conflict of Interest

The authors declare that they have no conflict of interest.

## References


[1] Anderson WJ, Doyle LA. Updates from the 2020 World Health Organization Classification of Soft Tissue and Bone Tumours. Histopathology 2021; 78: 644–657. doi:10.1111/his.14265

[2] Engel H, Herget GW, Füllgraf H et al. Chondrogenic Bone Tumors: The Importance of Imaging Characteristics. Fortschr Röntgenstr 2021; 193: 262–274. doi:10.1055/a-1288-1209

[3] Davies AM, Shah A, Shah R et al. Are the tubular bones of the hand really the commonest site for an enchondroma? Clinical Radiology 2020; 75: 533–537. doi:10.1016/j.crad.2020.02.004

[4] Griethuysen JJv, Fedorov A, Parmar C et al. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Research 2017; 77: e104–e107. doi:10.1158/0008-5472.CAN-17-0339

[5] Gitto S, Cuocolo R, Langevelde Kv et al. MRI radiomics-based machine learning classification of atypical cartilaginous tumour and grade II chondrosarcoma of long bones. EBioMedicine 2022; 75: 103757. doi:10.1016/j.ebiom.2021.103757

[6] Gitto S, Cuocolo R, Annovazzi A et al. CT radiomics-based machine learning classification of atypical cartilaginous tumours and appendicular chondrosarcomas. EBioMedicine 2021; 68: 103407. doi:10.1016/j.ebiom.2021.103407

[7] Deng XY, Chen HY, Yu JN et al. Diagnostic Value of CT- and MRI-Based Texture Analysis and Imaging Findings for Grading Cartilaginous Tumors in Long Bones. Frontiers in Oncology 2021; 11: 700204. doi:10.3389/fonc.2021.700204

[8] Zwanenburg A, Leger S, Vallières M et al. Image biomarker standardisation initiative. arXiv preprint arXiv 2020. doi:10.1148/radiol.2020191145

[9] Choi BB, Jee WH, Sunwoo HJ et al. MR differentiation of low-grade chondrosarcoma from enchondroma. Clinical Imaging 2013; 37: 542–547. doi:10.1016/j.clinimag.2012.08.006

[10] Pan J, Zhang K, Le H et al. Radiomics Nomograms Based on Non-enhanced MRI and Clinical Risk Factors for the Differentiation of Chondrosarcoma from Enchondroma. Journal of Magnetic Resonance Imaging 2021; 54: 1314–1323. doi:10.1002/jmri.27690

[11] Gassert FG, Breden S, Neumann J et al. Differentiating Enchondromas and Atypical Cartilaginous Tumors in Long Bones with Computed Tomography and Magnetic Resonance Imaging. Diagnostics 2022; 12: 2186. doi:10.3390/diagnostics12092186

[12] Mahajan P, Uddin S, Hajati F et al. Ensemble Learning for Disease Prediction: A Review. Healthcare (Basel) 2023; 11: 1808. doi:10.3390/healthcare11121808

[13] Ishaq A, Sadiq S, Umer M et al. Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques. IEEE Access 2021; 9: 39707–39716. doi:10.1109/access.2021.3064084

[14] Erdem F, Tamsel İ, Demirpolat G. The use of radiomics and machine learning for the differentiation of chondrosarcoma from enchondroma. Journal of Clinical Ultrasound 2023; 51: 1027–1035. doi:10.1002/jcu.23461

[15] Fritz B, Yi PH, Kijowski R et al. Radiomics and Deep Learning for Disease Detection in Musculoskeletal Radiology: An Overview of Novel MRI- and CT-Based Approaches. Investigative Radiology 2023; 58: 3–13. doi:10.1097/RLI.0000000000000907