

Diagnostic model based on CT and plantar pressure features for moderate-to-severe knee osteoarthritis

JiaCheng Wang

Yantaishan Hospital

Lin Xu

The Second Clinical Medical School of Shandong Medical and Pharmaceutical University

ShengJie Dong

Yantaishan Hospital

ZhengFeng Liu

The Second Clinical Medical School of Shandong Medical and Pharmaceutical University

JiangKun Qu

Yantaishan Hospital

HaiYang Tang

Yantaishan Hospital

YongQi Wang

Yantaishan Hospital

JunJie Jiang

JunjieJiang235@yeah.net

Yantaishan Hospital

Research Article

Keywords: knee osteoarthritis, digital model, machine learning, multi-modal

Posted Date: August 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10509154/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Diagnostic model based on CT and plantar pressure features for moderate-to-severe knee osteoarthritis

JiaCheng Wang^{1,2}, Lin Xu², ShengJie Dong¹, ZhengFeng Liu^{1,2}, JiangKun Qu¹,
HaiYang Tang¹, YongQi Wang¹, JunJie Jiang^{1*}

¹Department of Orthopedics, Yantaishan Hospital, Yantai, China

²The Second Clinical Medical School of Shandong Medical and Pharmaceutical
University, Yantai, China

***Corresponding author:**

JunJie Jiang

Department of Orthopedics, Yantaishan Hospital, No. 10087 Science and Technology
Avenue, Laishan District, Yantai 264003, Shandong Province, China

Tel: +86 19861122616

Email: JunjieJiang235@yeah.net

ORCID: JunjieJiang235@yeah.net

Abstract

Background: Knee osteoarthritis (KOA) is one of the most prevalent chronic joint conditions. Accurate stratification of KOA severity is essential for guiding patient-tailored surgical planning. However, there is a notable gap in the assessment and diagnosis of patients with moderate-to-severe KOA (Kellgren–Lawrence [K–L] grades 3 and 4) who are awaiting surgical intervention.

Methods: We constructed a multi-modal classification model integrating computed tomography (CT) imaging, plantar pressure profiling and routine clinical indicators and implemented an interpretable pipeline for automated extraction of the minimum joint space width from CT images and data-driven filtering of plantar pressure features. We further weighted and fused the processed radiomic and biomechanical features with conventional clinical features to train a support vector machine classifier.

Results: The multi-modal model achieved higher accuracy and balanced accuracy than the traditional single-modality imaging model and showed superior classification performance for distinguishing patients with K–L 3 and K–L 4 KOA.

Conclusion: Findings demonstrate that both multi-modal data integration and weighted feature fusion significantly improve KOA classification efficacy.

Clinical trial number: Not applicable.

Keywords

knee osteoarthritis; digital model; machine learning; multi-modal

Background

Knee osteoarthritis (KOA) is a degenerative joint disease affecting the knee articular cartilage, synovium and joint capsule. KOA can lead to joint deformity and disability and is one of the most common chronic diseases worldwide [1–5]. By 2050, ~130 million people worldwide may be affected by it, which poses a substantial challenge

45 to modern medicine [6,7].

46 Surgical treatment is the final-line solution for end-stage KOA. The conventional

47 decision-making process is as follows: doctors use Kellgren–Lawrence (K–L) grading

48 [8,9] to assess the severity of KOA based on weight-bearing knee radiographs and

49 formulate corresponding surgical plans based on clinical guidelines, such as total knee

50 arthroplasty, high tibial osteotomy and unicompartmental knee arthroplasty [10].

51 However, K–L grading, which relies on clinician experience for radiographic

52 interpretation, inherently involves inter-observer subjectivity [11–13]. Insufficient

53 clinical experience may lead to inaccuracies in subsequent surgical planning, resulting

54 in either overexpansion or restriction of surgical indications [14]. Therefore, it is

55 extremely important to use objective grading tools to conduct preliminary severity

56 stratification and screening for patients with K–L 3 or K–L 4 seeking surgical

57 treatment so as to formulate accurate subsequent surgical plans.

58 Regarding the clinical diagnosis and grading of KOA, constructing a machine

59 learning (ML)–based classification model for K–L grading may be a promising

60 solution [15–17]. Most traditional KOA classification models are unimodal,

61 constructed using only a single category of features (e.g. imaging, spatial gait or

62 demographic features). Among these, models that extract imaging features from

63 radiological data achieve better performance than other unimodal models [18–20].

64 Multi-modal models that integrate additional feature modalities demonstrate superior

65 performance over unimodal approaches [21]. However, most studies have focused on

66 the classification of early-stage KOA (K–L 1 or K–L 2), emphasising early detection

and prevention of the disease, with insufficient attention given to patients with moderate-to-severe KOA awaiting surgery (K–L 3 or K–L 4). For these patients, there still exist clear differences in surgical procedures and intraoperative planning between K–L 3 and K–L 4 groups. Accurate stratification of these two grades, therefore, plays an important role in assisting orthopaedic surgeons in delivering optimal patient-tailored surgical decision-making.

Joint space width (JSW) is a well-recognised indicator that indirectly reflects the degree of knee joint cartilage wear. The minimum joint space width (Min-JSW) directly quantifies the core feature of K–L grading (i.e. the severity of joint space stenosis) and is strongly correlated with the K–L grade [22]. Numerous studies have demonstrated that the KOA classification model constructed using the Min-JSW measured from weight-bearing radiographs achieves good diagnostic performance [23]. However, the acquisition of conventional weight-bearing knee radiographs heavily relies on the X-ray beam angle and the patient's knee flexion angle [24,25]. Furthermore, the JSW measurement method specified by the Osteoarthritis Research Society International still introduces subjectivity into Min-JSW assessment, compromising the consistency of subsequent K–L grading [26]. In contrast, knee computed tomography (CT) enables 3D visualisation of the tibiofemoral joint space. Min-JSW measurements from conventional non-weight-bearing CT do not differ significantly from those acquired with weight-bearing CT, but non-weight-bearing CT provides consistently more accurate measurements than traditional weight-bearing radiographs [27].

The relationship between plantar pressure and KOA is relatively complex. KOA-induced knee deformity leads to deviation of lower limb alignment, which, in turn, alters plantar pressure distribution. The relationship can be clearly explained using both anatomical and mechanical mechanisms. In patients with varus-type KOA, the foot centre of pressure shifts outward to varying degrees [28]. The studies by Anzai et al. [29] and Munoz-Organero et al. [30] demonstrated that it is feasible to use plantar pressure parameters to distinguish patients with KOA from individuals without KOA, and KOA progression can be reflected by corresponding changes in plantar pressure. Therefore, we hypothesise that incorporating plantar pressure parameters into the classification model can significantly improve its predictive performance. Multi-dimensional demographic information has been validated as a clinical modal feature closely associated with KOA progression [9,31,32]. In this study, we incorporated relevant demographic features to improve the model's discriminatory performance. We constructed a multi-modal classification model using CT imaging data, plantar pressure measurements and corresponding demographic data as distinct feature modalities to perform K–L grading for patients with moderate-to-severe KOA in order to assist clinicians in subsequent surgical decision-making.

Methods

Ethical approval

This study was approved by the Ethical Review Committee of Yantaishan Hospital (Ethics Review Approval Letter of Yantaishan Hospital, No. 2026116, Yantai City,

China). The entire research was conducted in strict accordance with the principles set forth in the World Medical Association Declaration of Helsinki. Written informed consent was obtained from all participants, and all experiments and data analyses were conducted in relevant clinical laboratories of the participating hospital.

Participants

The study database included knee CT and X-ray images, corresponding plantar pressure reports and relevant clinical information collected from patients hospitalised with KOA between December 2025 and June 2026. We excluded patients according to the following criteria: (i) missing data for prespecified variables measured on both lower limbs, (ii) age < 50 years at the time of enrolment, (iii) confirmed diagnosis of rheumatic arthritis or post-traumatic arthritis, (iv) any comorbid conditions affecting gait parameters (e.g. osteoarthritis or arthritis of the hip or ankle), (v) prior surgical intervention on the knee of the affected limb and (vi) mild KOA (K–L grade ≤ 2) without surgical indications.

Data collection

Knee CT and X-ray images of the affected limb of all patients were provided by the radiology system of Yantaishan Hospital. Image acquisition was performed uniformly by experienced radiologists following standard operating procedures. The CT dataset included coronal, sagittal and axial images of the affected knee joint acquired with the patient in the supine position. All plain radiographs included weight-bearing anteroposterior and lateral views of the knee joint. Two experienced clinicians independently conducted K–L grading for the knee joint using X-ray imaging under

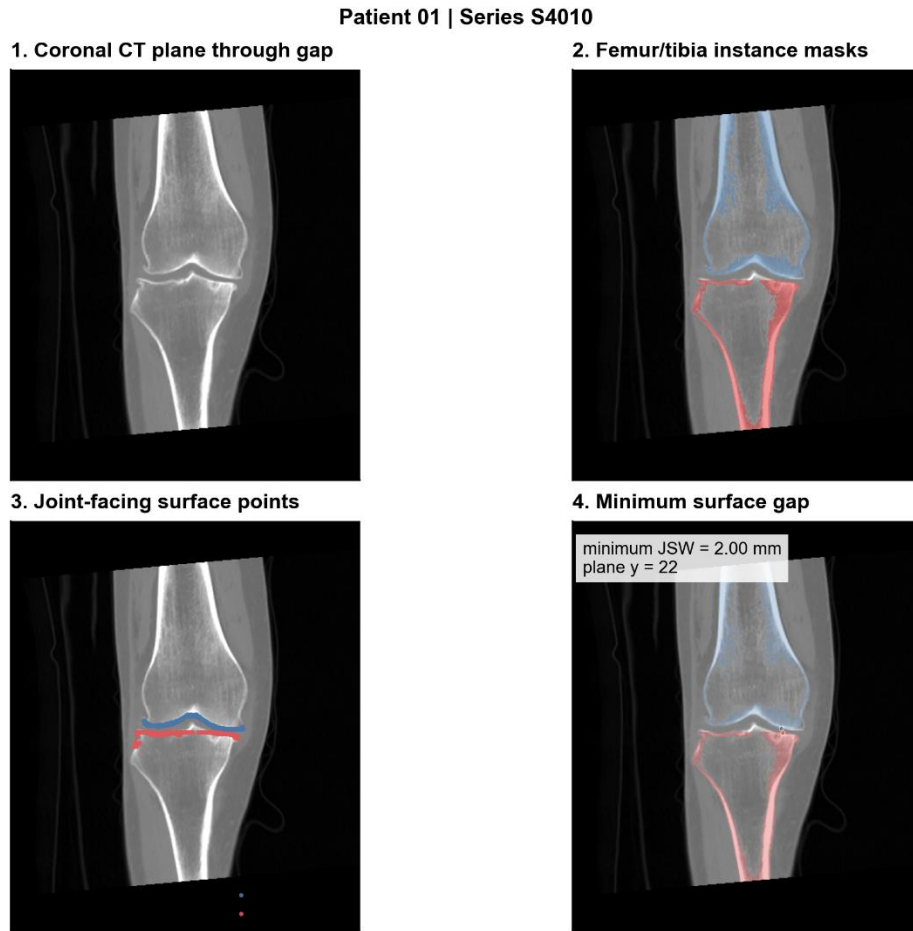
double-blinded conditions. In case of disagreement between the two assessments, a third senior orthopaedic surgeon performed an independent evaluation, and the result was used as the final consensus. Plantar pressure data were collected using the WATMAT plantar pressure gait analysis system (Wenzhou Kangzhilai Technology Co., Ltd., China) and its matching plantar pressure sensors. Plantar pressure outcome measures included the forefoot loading, hindfoot loading, medial loading and lateral loading ratios of the affected limb. All equipment was calibrated by the operator following the manufacturer's instructions to ensure consistent test parameters across all measurements. Prior to data collection, patients were briefed on study precautions, and they completed a warm-up familiarisation phase to adapt to the plantar pressure sensor. Each participant completed three round-trip walks over the pressure plate at a self-selected comfortable pace, maintaining a forward-facing posture, to capture valid measurements from both feet. Finally, the plantar pressure report was printed. Additional patient demographic and clinical data were extracted from the hospital's medical record system to guarantee data accuracy. All feature parameters analysed in this study were extracted from the patients' affected limbs. For patients with bilateral KOA, the side with more severe symptoms was included for analysis.

Feature extraction and data processing

Following the final experimental protocol, we included only patients with K–L 3 and K–L 4 KOA, resulting in a final cohort of 138 participants: 51 (37%) K–L 3 and 87 (63%) K–L 4 cases. The full study workflow consisted of three analytical layers, and all algorithms were implemented in Python 3.14.6 with supporting packages:

CT automated measurement layer: We used SimpleITK to read CT data from Digital Imaging and Communications in Medicine (DICOM) sequences, normalised the spatial orientation of all images and reconstructed the data into a 3D voxel structure. A CT value threshold of 250 HU was set for initial bone tissue screening via standard thresholding. Irrelevant soft tissue was removed, and final bone tissue segmentation was completed by combining connected component analysis. An algorithm was constructed to extract 1D bone/non-bone sequences along the z -direction for each (y, x) longitudinal column. The algorithm identified candidate ‘bone–interval–bone’ structures and recorded the interval centre position ‘centre_z’ and interval thickness ‘gap_vox’. Histogram statistics were calculated for all recorded centre_z values. The highest-frequency region was determined as the most probable tibiofemoral joint zone and recorded as ‘band_centre_z’. Only candidate intervals located near the pre-identified joint zone were retained. This step reduced interference from the patella, osteophytes and distant bone structures away from the articular surface, completing the final screening of the target joint zone. The primary bone structure *superior* to the pre-screened joint zone was defined as the femoral candidate, whereas the primary bone structure *inferior* to the joint zone was defined as the tibial candidate. Next, the inferior surface of the femur (facing the tibia) and the superior surface of the tibia (facing the femur) were extracted before the algorithm calculated the Min-JSW between the two articular surfaces. During distance calculation, all surface points to the vicinity of the pre-screened joint zone were restricted to exclude non-articular surfaces from participating in the measurement. Figure 1 presents the coronal CT, femur/tibia segmentation masks, articular surface point clouds and final minimum gap line from a representative case, illustrating the

177 interpretability of the proposed method.



178

179 **Figure 1** Example of the automatic measurement process of the CT joint space. Blue
180 represents the femur, red represents the tibia and the thin orange line indicates the
181 measurement position of the minimum surface gap. CT: computed tomography.

182

183 *Multi-modal feature layer:* Multi-modal feature construction was performed at this layer.

184 We adopted three distinct modalities of features: CT structural imaging, plantar pressure and
185 general clinical indicator features. For CT image structural features, Min-JSW was extracted
186 as the core quantitative indicator. The plantar pressure modality incorporated loading ratio
187 distribution parameters for each sub-region of the plantar surface. The clinical modality

included general demographic characteristics, such as gender and age. For a complete list of all features, refer to Table 1.

Table 1 Selection and construction of multi-modal features

Modality	Features	Methods
CT structural modality	CT automatic gap measurement	Only Min-JSW was used as the core structural indicator.
Plantar pressure modality	Forefoot loading ratio, hindfoot loading ratio, medial loading ratio and lateral loading ratio	Only high-weight/high-correlation variables were screened within CV folds to avoid data leakage.
Clinical indicator modality	Gender, age and BMI	Gender: male = 1, female = 0. The BMI was calculated as weight/height squared.

BMI, body mass index; CT, computed tomography; CV, cross-validation.

Modelling and fusion layer: We constructed the binary classification model using K–L 3/K–L 4 stratification as the prediction target. Model comparison and optimisation were implemented via stratified cross-validation (CV), intra-fold plantar pressure feature selection, linear support vector machine (SVM) classification and a grid search over modal weighting parameters. We adopted a five-fold stratified CV strategy, which maintained a consistent proportion of K–L 3/K–L 4 samples across each training and test fold, effectively mitigating evaluation fluctuations caused by class imbalance [33,34]. All steps, including standardisation, plantar pressure feature

selection and modal weight optimisation, were completed within the training fold, and
 then fitted transformations were applied to the held-out test fold to prevent test set
 information leakage. Plantar pressure variables were not directly and fixedly input to
 all models. Instead, they were screened independently within each training fold based
 on the absolute correlation strength between each variable and the classification target
 in the current training split, with only the top-ranked features retained (see Figure 2
 for details). The forefoot and hindfoot loading ratios demonstrated significantly
 stronger correlation among all the included plantar pressure variables. Due to the
 limited sample size, we adopted a linear SVM as the classifier, which can produce
 linear models with lower overfitting risk and better interpretability. We used the
 squared hinge loss with L2 regularisation and optimised the objective function via
scipy.optimize.minimize. To evaluate the contribution of different modal combinations,
 we established the following five groups of models: A, ‘CT’; B, ‘All Features’; C,
 ‘Weighted’; D, ‘CT_Pressure’; and E, ‘CT_Clinical’ (see Table 2 for details). Group
 C adopted standardised weighted feature fusion: First, we trained a linear SVM on
 weighted standardised features. Second, we performed a small-scale grid search to
 select optimal modal weights via inner stratified CV within each outer training fold.
 Third, we applied the optimised weights to the held-out outer test fold to generate
 predictions completely free of test set information leakage (see Figure 3 for details of
 the weight selection procedure). Compared with the general clinical indicator feature
 modality, the CT structural imaging and plantar pressure modalities occupied
 substantially higher weight proportions in the trained multi-modal SVM classifier,

exhibiting far greater predictive contribution to the final KOA severity stratification outputs. It is critical to apply modality weights *after* feature standardisation: if weighting is performed before standardisation, the subsequent re-scaling will eliminate the effect of the pre-assigned weights. Therefore, we intentionally ordered the weighting step after standardisation to ensure that the relative importance of different modalities could be correctly reflected in the final model.

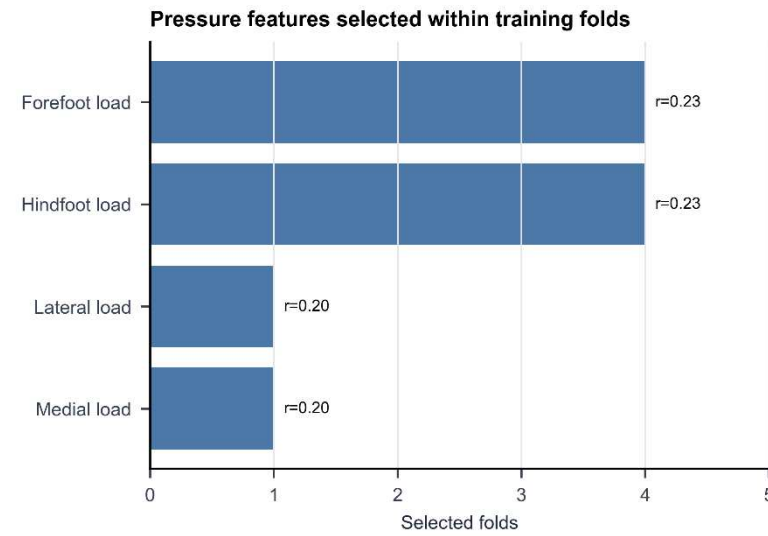


Figure 2 Frequency of plantar pressure feature selection across training folds.

Table 2 Models with five different combinations of modal features

Group	Input feature	Purpose
A (CT)	CT-derived Min-JSW	To evaluate the basic discriminative performance of a single structural CT metric

B (All Features)	CT + plantar pressure + clinical variable	To evaluate the performance of conventional full-feature concatenation
C (Weighted)	CT + plantar pressure + clinical variable + modal weight assignment	To evaluate whether weighted multi-modal fusion outperforms simple concatenation
D (CT_Pressure)	CT + plantar pressure	To evaluate whether mechanical load information improves CT-based discrimination
E (CT_Clinical)	CT + clinical variable	To evaluate the incremental value of age, gender and BMI

CT: computed tomography; BMI: body mass index; Min-JSW, minimum joint space width.

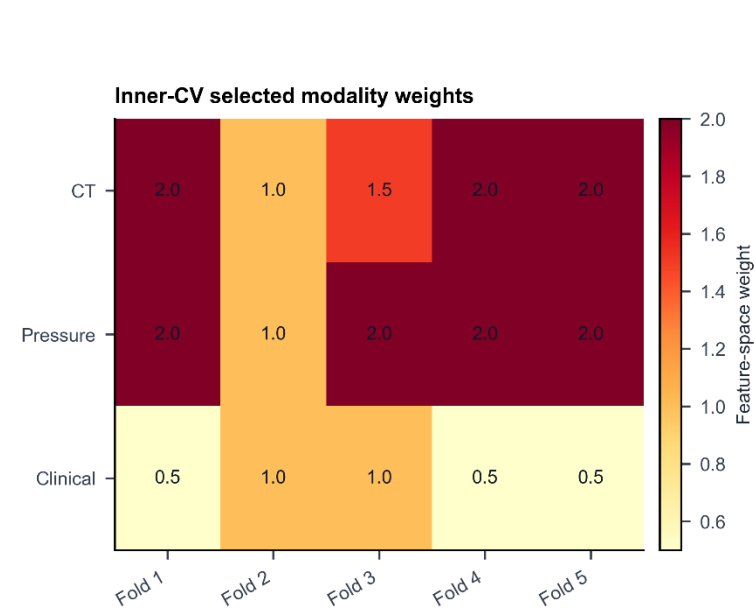


Figure 3 Modal weights selected via inner CV. CT, computed tomography; CV, cross-validation.

Results

Classification performance of the five models

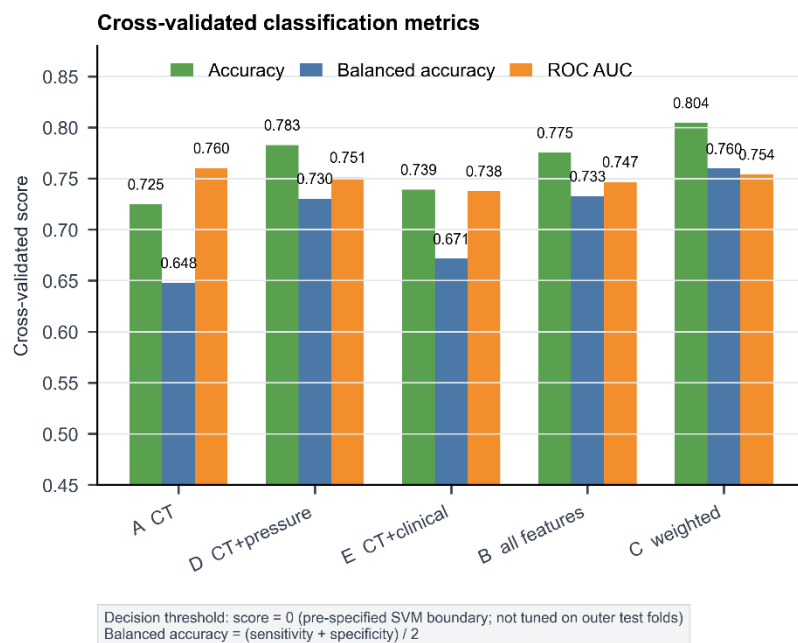
Table 3 and Figures 4–5 present the classification performance of the five models with different feature combinations tested under the same experimental setting. Group C was the multi-modal feature fusion group, which included CT structural imaging, plantar pressure and general clinical indicator modalities and adopted weighted processing at the feature level. The final accuracy and balanced accuracy of group C reached 0.804 and 0.760, respectively, the highest among all groups, and the area under the receiver operating characteristic curve (AUROC) was 0.754, indicating the best overall classification performance. Group A was the traditional imaging feature group, which had the highest AUROC (0.760), but its accuracy and balanced accuracy were relatively low. Specificity increased from 0.35 in group A to 0.59 in group C, indicating that group C effectively reduced the misclassification of K–L 3 as K–L 4. Group B was the direct concatenation group of the three features (accuracy = 0.775, balanced accuracy = 0.733 and AUROC = 0.747), also demonstrating decent classification capability. Groups D and E added one additional type of feature indicator based on the single imaging feature: to investigate the improvement effect of plantar pressure and clinical indicators on the model. Group D, which incorporated plantar pressure features, performed significantly better (accuracy = 0.783, balanced accuracy = 0.730 and AUROC = 0.751) than group E (accuracy = 0.739, balanced accuracy = 0.671 and AUROC = 0.738).

Table 3 Comparison of the performance of SVM classification models constructed from different groups of modal vectors

Group	<i>N</i>	<i>T</i>	<i>T</i>	<i>F</i>	<i>F</i>	Accu	Balanced	Sensi	Speci	Preci	Re	<i>F</i> ₁ -	AU
		<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	racy	accuracy	tivity	ficity	sion	call	scor	RO
												e	C
A (CT)	1	8	1	3	5	0.72	0.648	0.943	0.353	0.71	0.9	0.81	0.76
	3	2	8	3		5				3	43	2	0
	8												
D	1	8	2	2	6	0.78	0.730	0.931	0.529	0.77	0.9	0.84	0.75
(CT_Pres	3	1	7	4		3				1	31	4	1
sure)	8												
E	1	8	2	3	6	0.73	0.671	0.931	0.412	0.73	0.9	0.81	0.73
(CT_Clin	3	1	1	0		9				0	31	8	8
ical)	8												
B (All	1	7	2	2	9	0.77	0.733	0.897	0.569	0.78	0.8	0.83	0.74
Features)	3	8	9	2		5				0	97	4	7
	8												
C	1	8	3	2	6	0.80	0.760	0.931	0.588	0.79	0.9	0.85	0.75
(Weighte	3	1	0	1		4				4	31	7	4
d)	8												

AUROC, area under the receiver operating characteristic curve; CT, computed tomography; FN, false negative; FP, false positive SVM, support vector machine; TN, true negative; TP, true positive.

265



266

267 **Figure 4** Results of the K–L 3/K–L 4 binary classification model. Accuracy, balanced
268 accuracy and AUROC. AUROC, area under the receiver operating characteristic
269 curve; CT, computed tomography; CV, cross-validation; ROC, receiver operating
270 characteristic curve.

271

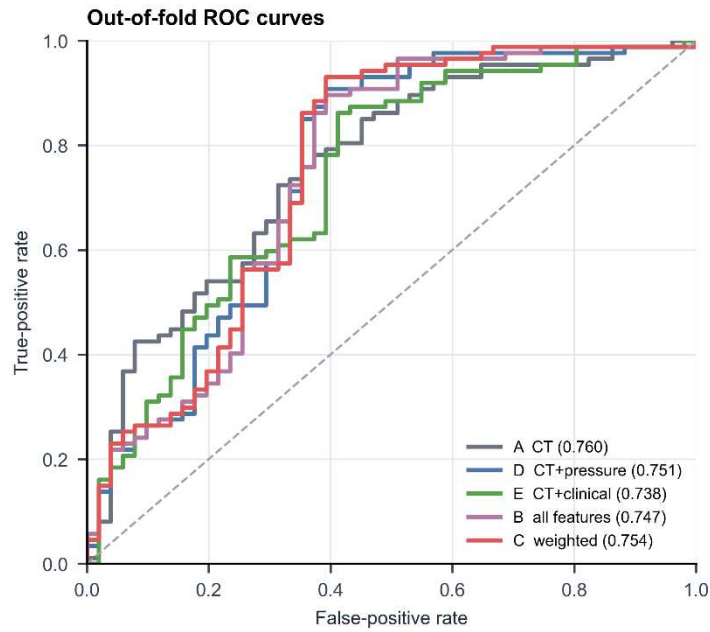


Figure 5 Outer-fold ROC curves. CT, computed tomography; ROC, receiver operating characteristic.

Figure 6 shows the comparison of classification performance among multiple groups. The AUROC marginally differed between groups C and A, indicating that the ranking ability of weighted feature fusion in group C was not significantly compromised. Instead, it achieved superior thresholded classification gain.

Weighted multimodal model improves thresholded classification more than ROC ranking

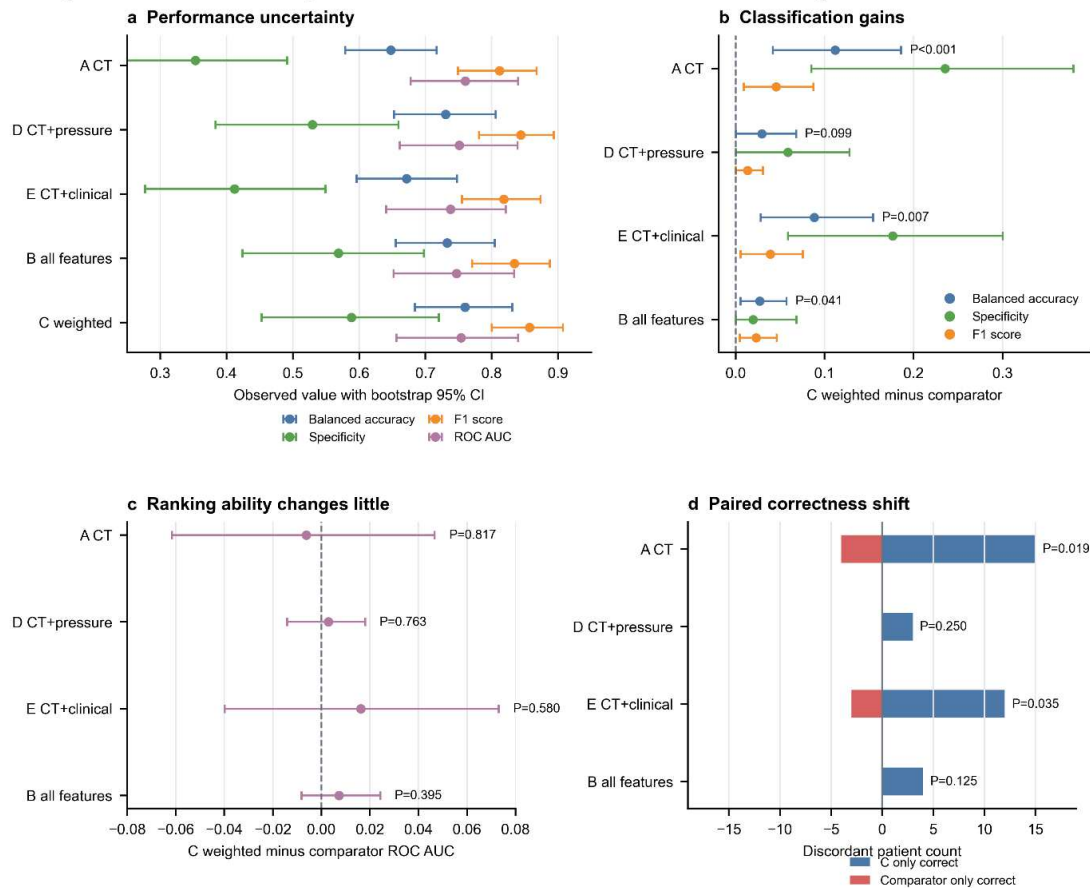


Figure 6 Weighted multi-modal model improved thresholded classification more than ROC ranking. **(a)** Uncertainty range of performance across all models, **(b)** difference in classification performance between group C and other control models, **(c)** difference in AUROC between group C and control models, **(d)** number of patients correctly predicted by the group C and control models. AUROC, area under the receiver operating characteristic curve; CI, confidence interval; CT, computed tomography; ROC, receiver operating characteristic curve.

Discussion

Principal findings

Most of the previous studies on the relationship between JSW/plantar pressure and KOA grading were conducted from the perspective of single-modal features and did not cover the research direction of multi-modal combinations [23,35]. Our classification model not only combines JSW and plantar pressure with the clinical feature modality but also realises the weighted fusion of different features to further improve classification performance. The final results of our study show that the multi-modal groups, especially the fusion group C, achieved better accuracy and balanced accuracy than the traditional imaging-only group A, further corroborating previous findings that to a certain extent, introducing appropriate modal features effectively enhances the model's classification performance [36–39]. The highest accuracy and balanced accuracy achieved by group C demonstrate that weighted multi-modal fusion can better balance classification decisions between K–L 3 and K–L 4 with a fixed classification threshold. Especially, the higher specificity in group C compared with group A indicated that group C effectively reduced misclassification of K–L 3 as K–L 4. Since the dataset in our study contained more K–L 4 samples, simply maximising the recognition rate of K–L 4 tended to result in the over-classification of K–L 3 as K–L 4. By incorporating plantar pressure and clinical background information, group C achieved improved recognition performance for K–L 3 cases. The lower AUROC of group C compared with group A can likely be attributed to the fact that the AUROC quantifies a model's ability to rank sample-wise risk rather than classification accuracy under a fixed decision threshold. Generally, CT-derived Min-JSW inherently has a direct structural association with K–L grading [23]. Therefore,

when used as a single modality, its grading signal is distinct. Incorporating additional modalities can not only improve classification performance under a fixed threshold but also may also introduce extra noise, meaning the overall ranking ability does not necessarily improve further. Therefore, the observed result (group C achieved the best classification performance, whereas group A yielded the highest AUROC) is logically consistent. Figure 6 shows the comparison of classification performance among multiple groups. The marginal difference in AUROC between groups C and A showed that the ranking ability of weighted feature fusion in group C was not significantly compromised but, instead, achieved superior thresholded classification gain. This finding further corroborates the feasibility of the curated modality combination and designated fusion strategy in the KOA classification framework.

Group B was the multi-modal feature concatenation group, with accuracy, balanced accuracy and AUROC all lower than those of group C, indicating that compared with simple feature concatenation, weighted feature fusion substantially improves the model's classification capability [21] by allowing the model to automatically adjust different modalities' contribution, enhancing the fusion strategy's data adaptability and further unleashing the overall model performance. Group D achieved significantly higher accuracy and balanced accuracy than those of group A, indicating that plantar pressure features complement CT JSW parameters and can reflect abnormal load distribution and functional status [35,40]. Compared with group A, the results of group E demonstrated that adding clinical indicators improves performance compared to using a single imaging modality, but the performance gain is limited.

Comparing groups B and D further indicated that simply adding raw clinical indicators does not necessarily improve classification performance significantly and may even introduce noise. This result can likely be attributed to three reasons: (i) Both K–L 3 and K–L 4 represent advanced stages of KOA, where structural joint space information is more directly correlated with plantar pressure variations [23,35]; (ii) clinical variables are characterised by lower dimensionality and coarser information granularity; and (iii) the limited sample size led to a weak linear association between general clinical information variables and outcome grading. Therefore, a multi-modal classification model’s performance depends on both the number of input features and pre-weighted fusion of multi-source features [21].

Strengths and limitations

In relation to cost and radiation exposure, the majority of studies have extracted Min-JSW from plain radiographs of outpatient cohorts [41]. For surgical candidates who undergo more comprehensive pre-operative examinations, however, CT is a superior alternative as it provides far more precise anatomical information. In this study, we innovatively integrated CT-derived Min-JSW extraction into the predictive modelling framework and implemented a fully automated extraction pipeline, constructing a complete closed-loop workflow spanning from raw CT DICOM ingestion, standardised 3D measurement and structured result write-back to subsequent modelling analysis, eliminating manual intervention in the entire process. For CT-derived Min-JSW extraction, we adopted a rule-driven 3D instance segmentation pipeline combined with surface measurement. No manual segmentation annotations

are required, making the process well-suited for most scenarios lacking manually labelled masks. Moreover, all processing steps (feature standardisation, plantar pressure feature selection and weight optimisation) are completed within the training folds, effectively preventing data leakage. Furthermore, our proposed method offers superior interpretability: CT features are explicit millimetre-level JSW measurements, a linear SVM classifier is adopted and both fusion weights and plantar pressure feature selection results are fully traceable.

This study also has several limitations. The K–L 3 sample size was relatively small, leading to lower model specificity; future studies should include additional K–L 3 samples. Currently, plantar pressure feature selection is performed via correlation magnitude ranking within each training fold; in future work, stability selection, L1 regularisation and embedded feature selection can be explored as alternative strategies. Constrained by practical clinical settings, all samples in this study were collected from a single institution, with no external validation performed. Therefore, the generalisability and transferability of our findings cannot be confirmed; future work should expand the study cohort to include patient data from multiple regions across different geographical locations.

Conclusions

K–L grading of KOA relies on clinician experience for radiographic interpretation and can be subjective, leading to inaccuracies in subsequent surgical planning. Knee CT enables 3D visualisation of the tibiofemoral joint space. We constructed a KOA

classification model based on CT structural imaging, plantar pressure and general clinical indicator modalities to address the inaccuracy in grading patients with K–L 3 and K–L 4 KOA and assist clinicians in formulating corresponding surgical decisions. The unified classification framework demonstrated a novel structural-mechanical-clinical multi-modal combination approach. This work will assist clinicians in developing more rational and personalised surgical strategies and reduce the incidence of inappropriate KOA treatment.

List of abbreviations

KOA	Knee osteoarthritis
K–L	Kellgren–Lawrence
CT	Computed tomography
ML	Machine learning
JSW	Joint space width
Min-JSW	Minimum joint space width
DICOM	Digital imaging and communications in medicine
BMI	Body mass index
CV	Cross-validation
SVM	Support vector machine
AUROC	Area under the receiver operating characteristic curve
FN	False negative
FP	False positive
TN	True negative
TP	True positive

403 ROC Receiver operating characteristic curve

404

405 **Declarations**

406

407 **Ethics approval and consent to participate**

408 This study was approved by the Ethical Review Committee of Yantaishan Hospital
409 (Ethics Review Approval Letter of Yantaishan Hospital, No. 2026116, Yantai City,
410 China). The entire research was conducted in strict accordance with the principles set
411 forth in the World Medical Association Declaration of Helsinki. Written informed
412 consent was obtained from all participants, and all experiments and data analyses
413 were conducted in relevant clinical laboratories of the participating hospital.

414

415 **Consent for publication**

416 Consent for publication was obtained from all participants.

417

418 **Availability of data and materials**

419 All datasets involved in this study were obtained from the hospital, including patients'
420 CT images, X-ray radiographs, plantar pressure reports and other basic clinical
421 information, which are available from the corresponding author upon reasonable
422 request.

423

424 **Competing interests**

The authors declare that they have no competing interests

Funding

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

Authors' contributions

J.W.: Conceptualization, Validation, Formal analysis, Writing - original draft. L.X.: Methodology, Validation, Formal analysis, Writing - review & editing, Supervision. S.D.: Validation, Formal analysis, Writing - review & editing, Supervision. Z.L.: Validation, Data curation. J.Q.: Validation, Data curation. H.T.: Validation, Data curation. Y.W.: Validation, Data curation. J.J.: Conceptualization, Methodology, Validation, Writing - review & editing, Supervision.

Acknowledgements

The authors would like to thank the Department of Joint Surgery of Yantaishan Hospital for its support in data acquisition for this study.

References

1. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. *Arthritis Rheum.* 1998;41:1343-55.

446 2. Kaufman KR, Hughes C, Morrey BF, Morrey M, An KN. Gait characteristics of
447 patients with knee osteoarthritis. *J Biomech.* 2001;34:907-15.

448 3. Glyn-Jones S, Palmer AJ, Agricola R, Price AJ, Vincent TL, Weinans H, et al.
449 Osteoarthritis. *Lancet.* 2015;386:376-87.

450 4. Sharma L. Osteoarthritis of the knee. *N Engl J Med.* 2021;384:51-9.

451 5. Felson DT, Naimark A, Anderson J, Kazis L, Castelli W, Meenan RF. The
452 prevalence of knee osteoarthritis in the elderly. the Framingham osteoarthritis study.
453 *Arthritis Rheum.* 1987;30:914-8.

454 6. Choi JC, Jeong MY, Kim YJ, Kim KG. Artificial intelligence model assists knee
455 osteoarthritis diagnosis via determination of K-L grade. *Diagnostics.* 2025;15:1220.

456 7. Wu Y, Xiang Z, Tan Y, Hu J, Chen D, Zhao J, et al. Study on multimodal spatially-
457 constrained contrastive learning for knee osteoarthritis severity grading. *J Biomed*
458 *Inform.* 2025;172:104962.

459 8. Yoo TK, Kim DW, Choi SB, Oh E, Park JS. Simple scoring system and artificial
460 neural network for knee osteoarthritis risk prediction: a cross-sectional study. *PLoS*
461 *One.* 2016;11:e0148724.

462 9. Lazzarini N, Runhaar J, Bay-Jensen AC, Thudium CS, Bierma-Zeinstra SMA,
463 Henrotin Y, et al. A machine learning approach for the identification of new

464 biomarkers for knee osteoarthritis development in overweight and obese women.
 465 Osteoarthritis Cartilage. 2017;25:2014-21.

466 10. Burgio C, Bosco F, Cobisi C, Zepeda K, Giustra F, Capella M, et al. Surgical
 467 treatment of varus unicompartmental knee osteoarthritis: indications, trends, and
 468 outcomes—a narrative review. Arch Orthop Trauma Surg. 2026;146:227.

469 11. Gossec L, Jordan JM, Mazzuca SA, Lam M-A, Suarez-Almazor ME, Renner JB,
 470 et al. Comparative evaluation of three semi-quantitative radiographic grading
 471 techniques for knee osteoarthritis in terms of validity and reproducibility in 1759 X-
 472 rays: report of the OARSI-OMERACT task force. Osteoarthritis Cartilage.
 473 2008;16:742-8.

474 12. Sheehy L, Culham E, McLean L, Niu J, Lynch J, Segal NA, et al. Validity and
 475 sensitivity to change of three scales for the radiographic assessment of knee
 476 osteoarthritis using images from the Multicenter Osteoarthritis Study (MOST).
 477 Osteoarthritis Cartilage. 2015;23:1491-8.

478 13. Culvenor AG, Engen CN, Øiestad BE, Engebretsen L, Risberg MA. Defining the
 479 presence of radiographic knee osteoarthritis: a comparison between the Kellgren and
 480 Lawrence system and OARSI atlas criteria. Knee Surg Sports Traumatol Arthrosc.
 481 2015;23:3532-9.

- 482 14. Riddle DL, Perera RA, Jiranek WA, Dumenci L. Using surgical appropriateness
483 criteria to examine outcomes of total knee arthroplasty in a United States sample.
484 *Arthritis Care Res.* 2015;67:349-57.
- 485 15. Kwok CK, Guehring H, Aydemir A, Hannon MJ, Eckstein F, Hochberg MC.
486 Predicting knee replacement in participants eligible for disease-modifying
487 osteoarthritis drug treatment with structural endpoints. *Osteoarthritis Cartilage.*
488 2020;28:782-91.
- 489 16. Heisinger S, Hitzl W, Hobusch GM, Windhager R, Cotozana S. Predicting total
490 knee replacement from symptomology and radiographic structural change using
491 artificial neural networks-data from the osteoarthritis initiative (OAI). *J Clin Med.*
492 2020;9:1298.
- 493 17. Urish KL, Keffalas MG, Durkin JR, Miller DJ, Chu CR, Mosher TJ. T2 texture
494 index of cartilage can predict early symptomatic OA progression: data from the
495 osteoarthritis initiative. *Osteoarthritis Cartilage.* 2013;21:1550-7.
- 496 18. Joseph GB, McCulloch CE, Nevitt MC, Neumann J, Gersing AS, Kretschmar M,
497 et al. Tool for osteoarthritis risk prediction (TOARP) over 8 years using baseline
498 clinical data, X-ray, and MRI: data from the osteoarthritis initiative. *J Magn Reson*
499 *Imaging.* 2018;47:1517-26.

- 500 19. Tiulpin A, Thevenot J, Rahtu E, Lehenkari P, Saarakkala S. Automatic knee
501 osteoarthritis diagnosis from plain radiographs: a deep learning-based approach. *Sci*
502 *Rep.* 2018;8:1727.
- 503 20. Yeoh PSQ, Lai KW, Goh SL, Hasikin K, Hum YC, Tee YK, et al. Emergence of
504 deep learning in knee osteoarthritis diagnosis. *Comput Intell Neurosci.*
505 2021;2021:4931437.
- 506 21. Liu L, Chang J, Zhang P, Ma Q, Zhang H, Sun T, et al. A joint multi-modal
507 learning method for early-stage knee osteoarthritis disease classification. *Heliyon.*
508 2023;9:e15461.
- 509 22. Gale DR, Chaisson CE, Totterman SM, Schwartz RK, Gale ME, Felson D.
510 Meniscal subluxation: association with osteoarthritis and joint space narrowing.
511 *Osteoarthritis Cartilage.* 1999;7:526-32.
- 512 23. Cheung JC, Tam AY, Chan LC, Chan PK, Wen C. Superiority of multiple-joint
513 space width over minimum-joint space width approach in the machine learning for
514 radiographic severity and knee osteoarthritis progression. *Biology.* 2010;10:1107.
- 515 24. Hunter DJ, Zhang YQ, Tu X, LaValley M, Niu JB, Amin S, et al. Change in joint
516 space width: hyaline articular cartilage loss or alteration in meniscus? *Arthritis*
517 *Rheum.* 2006;54:2488-95.

- 518 25. Kothari MD, Rabe KG, Anderson DD, Nevitt MC, Lynch JA, Segal NA, et al.
519 The relationship of three-dimensional joint space width on weight bearing CT with
520 pain and physical function. *J Orthop Res.* 2019; doi:10.1002/jor.24566.
- 521 26. Tiulpin A, Saarakkala S. Automatic grading of individual knee osteoarthritis
522 features in plain radiographs using deep convolutional neural networks. *Diagnostics.*
523 2020;10:932.
- 524 27. Fritz B, Fritz J, Fucentese SF, Pfirrmann CWA, Sutter R. Three-dimensional
525 analysis for quantification of knee joint space width with weight-bearing CT:
526 comparison with non-weight-bearing CT and weight-bearing radiography.
527 *Osteoarthritis Cartilage.* 2022;30:671-80.
- 528 28. Lidtke RH, Muehleman C, Kwasny M, Block JA. Foot center of pressure and
529 medial knee osteoarthritis. *J Am Podiatr Med Assoc.* 2010;100:178-84.
- 530 29. Anzai E, Yamashita K, Matsunaga Y, Sato M, Menz HB, Ohta Y. Foot posture
531 and dynamic foot function in people with and without knee osteoarthritis.
532 *Osteoarthritis Cartilage.* 2017;25:S131-2.
- 533 30. Munoz-Organero M, Littlewood C, Parker J, Powell L, Grindell C, Mawson S.
534 Identification of walking strategies of people with osteoarthritis of the knee using
535 insole pressure sensors. *IEEE Sens J.* 2017;17:3909-20.

- 536 31. Kerkhof HJM, Bierma-Zeinstra SMA, Arden NK, Metrustry S, Castano-
537 Betancourt M, Hart DJ, et al. Prediction model for knee osteoarthritis incidence,
538 including clinical, genetic and biochemical risk factors. *Ann Rheum Dis*.
539 2014;73:2116-21.
- 540 32. Halilaj E, Le Y, Hicks JL, Hastie TJ, Delp SL. Modeling and predicting
541 osteoarthritis progression: data from the osteoarthritis initiative. *Osteoarthritis*
542 *Cartilage*. 2018;26:1643-50.
- 543 33. Padman R, Bai X, Airolidi EM. A new machine learning classifier for high
544 dimensional healthcare data. *Stud Health Technol Inform*. 2007;129:664-8.
- 545 34. Yang X, Qiu H, Wang L, Wang X. Predicting colorectal cancer survival using
546 time-to-event machine learning: retrospective cohort study. *J Med Internet Res*.
547 2023;25:e44417.
- 548 35. Wang A, Li D, Fan N, Yuan S, Wu Q, Fu Z, et al. Piezoresistive-based gait
549 monitoring technique for the recognition of knee osteoarthritis patients. *IEEE Access*.
550 2022;10:123874-84.
- 551 36. Hu W, Meng X, Bai Y, Zhang A, Qu G, Cai B, et al. Interpretable multimodal
552 fusion networks reveal mechanisms of brain cognition. *IEEE Trans Med Imaging*.
553 2021;40:1474-83.

554 37. Zhang D, Wang Y, Zhou L, Yuan H, Shen D. Multimodal classification of
555 Alzheimer's disease and mild cognitive impairment. *NeuroImage*. 2011;55:856-67.

556 38. Hu W, Cai B, Zhang A, Calhoun VD, Wang Y-P. Deep collaborative learning
557 with application to the study of multimodal brain development. *IEEE Trans Biomed*
558 *Eng*. 2019;66:3346-59.

559 39. Liu J, Tan G, Lan W, Wang J. Identification of early mild cognitive impairment
560 using multi-modal data and graph convolutional networks. *BMC Bioinform*.
561 2020;21:123.

562 40. Levinger P, Menz HB, Fotoohabadi MR, Feller JA, Bartlett JR, Bergman NR.
563 Foot posture in people with medial compartment knee osteoarthritis. *J Foot Ankle*
564 *Res*. 2010;3:29.

565 41. Conrozier T, Lequesne M, Favret H, Taccoen A, Mazières B, Dougados M, et al.
566 Measurement of the radiological hip joint space width. An evaluation of various methods of
567 measurement. *Osteoarthritis Cartilage*. 2001;9:281-6.

568

569