npj | artificial intelligence

Article

https://doi.org/10.1038/s44387-025-00058-y

A novel deep learning based automatic ultrasonic posterior cruciate ligament clinical assessment tool

Check for updates

Jyun-Ping Kao[1], Jiajing Zhang[2], Wei-Ning Lee[2,3], Chung-Ping Chen[1] & Lawrence Chun-Man Lau[4] ✉

Posterior cruciate ligament (PCL) injuries often present subtle clinical signs and can be overlooked during standard diagnostic procedures, leading to missed or delayed treatment. Although magnetic resonance imaging (MRI) is highly accurate for acute injuries, it is expensive, time-consuming, has low sensitivity for chronic and graft tears, and is limited for patients with metallic implants. Using ultrasound imaging can clearly visualize the intra-articular PCL and holds great promise as an accessible, cost-effective, and real-time alternative with proven high accuracy in both acute and chronic injuries. However, its interpretation is operator-dependent and lacks streamlined quantitative assessment tools. We propose a novel deep learning-based framework for instantaneous, operator-independent quantification of PCL morphology directly from ultrasound images. We employ a Real-Time Detection Transformer that bypasses the need for complex segmentation models, instead deriving clinically meaningful metrics—PCL location, width, and angle through an object detection model and novel pre- and post-processing steps.

The posterior cruciate ligament (PCL) is the strongest and largest ligaments in the knee and thus is extremely important in providing stability for the knee[1]. It acts as a primary stabilizer, ensuring the knee remains steady during movement[2] while preventing posterior tibial translation and serving a secondary role in resisting internal rotation with knee flexion more than 90°[3]. The PCL originates from the anteromedial aspect of the intercondylar notch and inserts posteriorly onto the midline of tibial plateau below the joint line. The PCL composes of an anterolateral bundle (ALB) and a posteromedial bundle (PMB) and together the PCL measures approximately 32–8 mm in length with a cross-sectional area of about 31.2 mm^2 at its mid-substance level, 1.5 times that of the anterior cruciate ligament (ACL)[4].

PCL injury is commonly missed on physical examination (PE) and its true prevalence is unknown[5]. It has been reported that PCL injury occurs in up to 38% of all acute knee injuries and the incidence is even higher up to 56.5% with trauma and together with other ligament injuries in 95% cases[6]. Commonly, patients often do not feel distinct tissue rupture; instead, they experience instability accompanied by mild to moderate effusion[7]. PCL tears can be easily missed during PE[8], and a systematic review has highlighted that the accuracies of many PEs for PCL remain largely unknown, leading to unknown true prevalence of PCL injury and potentials for missed diagnosis[9]. Furthermore, it has been reported that more than 15% of surgically treated PCL patients had previously undergone unnecessary reconstruction of ACL due to misdiagnosis, emphasizing the need for a comprehensive evaluation[10].

Currently, PE with the posterior drawer test is the primary clinical method employed by clinicians to evaluate PCL integrity, measuring the backward movement of the tibia relative to the femur[11,12]. However, early detection of PCL tears can be challenging, and untreated PCL injuries may increase the long-term risk of osteoarthritis (OA). Even with initial detection by PE, imaging modalities are required for confirmation and visualization of the tear[11,12]. If left untreated, PCL injuries can alter knee kinematics, leading to persistent anterior subluxation of the femoral condyle and increasing the risk of OA over time consequently[13].

Diagnostic approaches for PCL injury include posterior stress radiography and magnetic resonance imaging (MRI)[2,7]. MRI offers high accuracy, ranging from 96 to 100% for acute PCL injury diagnosis[7], and can assess associated injuries[7,14]. The normal PCL appears dark on both T1- and T2-weighted MRI scans and is curvilinear in shape. While MRI is widely regarded as the gold standard, it can present with normal PCL appearances even with complete disruption[2]. Although MRI sensitivity for acute PCL tears is acceptable, its sensitivity for chronic (≥6 weeks) primary PCL tears and PCL graft tears can be as low as 62.5% and 18.1%, respectively[15]. Moreover, the signal of acute PCL injury returns to normal within 6 months[16]. The widespread use of MRI for diagnosis of PCL injury is

[1]Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan, China. [2]Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong, China. [3]School of Biomedical Engineering, The University of Hong Kong, Hong Kong, China. [4]Department of Orthopaedics and Traumatology, School of Clinical Medicine, The University of Hong Kong, Hong Kong, China. ✉e-mail: laucml@hku.hk

limited as it is expensive, time-consuming, and contraindicated in patients with certain implants or conditions such as pacemakers, magnetic metal implants, and claustrophobia[2,17]. Besides, knees with metallic prostheses e.g., unicompartmental knee arthroplasty or implants like distal femur plating will create significant metallic artefacts, limiting the visualization of the PCL and its subsequent assessment[18].

Ultrasound (US) is gaining traction as a diagnostic tool for PCL injuries due to its high availability, point-of-care utility, real-time imaging capability, cheap, and lack of ionizing radiation[2,17]. It makes use of the acoustic window over the popliteal fossa, which is quite unique as the PCL, despite being intra-articular, is relatively close to the skin in this region and this area is never covered up by other non-acoustic structures or orthopedic implants, which would otherwise hinder the knee from flexion. These properties guarantee US is possible and reproducible in different clinical situations. A recent review[19] reported that the sensitivity and specificity of point-of-care US ranged from 83.3 to 100% and 86.7 to 100% with 100% accuracy, suggesting substantial potential for cost savings if we use ultrasound to replace MRI for diagnosis of PCL injury.

Recent research has established diagnostic standards for PCL injuries using ultrasound. A PCL width ≥6.5 mm yields 90.6% sensitivity and 86.7% specificity to distinguish acute PCL injured knee from non-injured knees[17]. Furthermore, this 6.5 mm cutoff also demonstrates 83.3% sensitivity and 77.8% specificity for chronic PCL injuries, while a PCL angle (tangent line drawn from the tibial insertion, also known as the angle between the anterior boundary and posterior boundary of PCL) cutoff of 20° shows 88.9% sensitivity and 94.4% specificity for chronic PCL injuries as well[2]. Although these criteria offer direct diagnostic guidance, manual evaluation by physicians remains necessary. Ultrasound measurements must be captured through screenshots or videos, limiting comprehensive assessment. Besides, unlike MRI, ultrasound is operator-dependent, and its interpretation hinges also on the examiner's skill and experience.

Recently, various studies have employed deep learning to assist ultrasound diagnosis, annotation, and segmentation. For example, Sharifi et al.[20] demonstrated automatic detection and segmentation of thyroid nodules, and Zegarra et al.[21] showed effectiveness for identifying normal and abnormal fetal anatomy and measuring fetal biometry. In musculoskeletal ultrasound, segmentation models have been used to delineate osseous and soft-tissue structures, including the radius and tibia[22], cardiac muscle[23], and ligaments[24]. However, when the clinical goal is to compute numeric measurements (e.g., widths or angles), segmentation requires pixel-accurate masks to be reliable[25]. In MSK US, irregular, low-contrast borders, speckle, and acoustic shadowing render boundary delineation subjective and labor-intensive to label and to learn[26]. Furthermore, dense per-pixel decoders substantially increase model parameters and FLOPs, which constrains real-time deployment at the bedside[27]. Although Xue et al.[28] and Shi et al.[29] have recently introduced deep-learning assisted systems that reach near-expert performance for categorical grading of ACL and anterior talofibular ligament (ATFL) injuries on MRI and ultrasound, respectively, these frameworks are designed for status classification and do not return precise, continuous ligament morphometrics (e.g., length, thickness, or laxity). For PCL, prior computer-vision work has largely relied on handcrafted image features rather than true morphometrics. For instance, Zarychta[30] proposed fuzzy C-means guided ROI localization for segmentation and skeletonization of the ligament on MRI and the pathology was then inferred from a shape heuristic to distinguish injured from healthy PCLs cohort. However, this pipeline requires manual seeding, assumes near-orthogonal femorotibial alignment, and degrades in cases with fractures or off-center anatomy and does not yield absolute, continuous ligament measurements.

To overcome these limitations, we propose a detection-based ultrasound measurement framework that quantifies PCL morphology—location, width, and angle—directly from images without segmentation. The key novelty is to replace masks with a compact set of detection targets tailored to the PCL: the overall PCL region, its anterior and posterior margins, and a width proxy. We introduce a simple, reproducible labeling protocol that uses sparse box-level annotations, together with closed-form geometric transformations that convert detections into continuous measurements (millimeter width and angular orientation) using the image calibration metadata. By eliminating heavy per-pixel decoders and optimizing a single-stage detector of a segmentation model with an object detection model, our approach reduces annotation burden and model complexity, and achieves low-latency inference suitable for continuous, real-time scanning. This assessment tool enables objective and operator-independent documentation of the PCL's morphology, including extrema such as maximum width or angle, for subsequent clinical assessment. In this study, we develop and validate a real-time deep-learning framework for automatic PCL detection and quantitative measurement (width and angle) from ultrasound images, and evaluate its accuracy against clinical criteria.

Given limitations in current PE and MRI imaging for PCL injury but its potential high prevalence[5], a quick, inexpensive, objective, bedside imaging tool would be highly beneficial for the clinical assessment of PCL injury by clinicians as a first-line assessment tool and US potentially can fulfill this role if its measurement can be objectively produced. To address the issues mentioned above, we propose a novel deep learning-based measurement system to quantify the PCL's morphology—including the location, width, and angle—directly from ultrasound images simultaneously. The measurement system, implemented with an object detection model specifically designed for continuous ultrasound measurement, is computationally efficient and can operate in a real-time manner such that continuous documentation of the PCL's location, width, and angle is made possible for subsequent assessment. This allows precise documentation of the PCL morphology at every instant, for instance, the maximum width and angle, potentially minimizing operator dependency in detection and interpretation. Our approach ultimately aims to facilitate clinical and research use of point-of-care ultrasound for musculoskeletal assessment by maximizing the objectivity of ultrasound measurement and reporting. In this study, we aim to develop and validate a real-time deep learning framework for automatic PCL detection and quantitative measurement (width and angle) from ultrasound images, and to evaluate its accuracy against clinical criteria.

Results

PCL detection

Our system achieves near-perfect detection of the PCL locations across all validation folds, as summarized in Table 1. This remarkable performance can be attributed to the standardized data acquisition protocol, wherein the ultrasound probe was systematically placed over the PCL region under carefully controlled clinical conditions. By ensuring that every captured frame prominently features the PCL along with the adjacent femur and tibia, the dataset effectively minimizes instances of covered up or absent PCL.

The high precision and recall values indicate that our model rarely produces false positives or overlooks true PCL instances. Correspondingly, the consistently high mAP50 and mAP50:95 scores highlight the model's ability to accurately delineate the ligament boundaries despite potential variability in ligament morphology or subtle image artifacts. These results underscore the reliability and robustness of our real-time detection framework for identifying the PCL location.

Table 1 | Performance metrics for PCL detection across five validation folds, including Precision, Recall, mAP50 and mAP50:95

Fold	Images	Precision	Recall	mAP50	mAP50-95
Fold-1	569	0.999	1	0.995	0.627
Fold-2	578	0.999	1	0.995	0.570
Fold-3	562	0.998	1	0.995	0.620
Fold-4	598	0.999	1	0.995	0.531
Fold-5	569	0.999	1	0.995	0.498

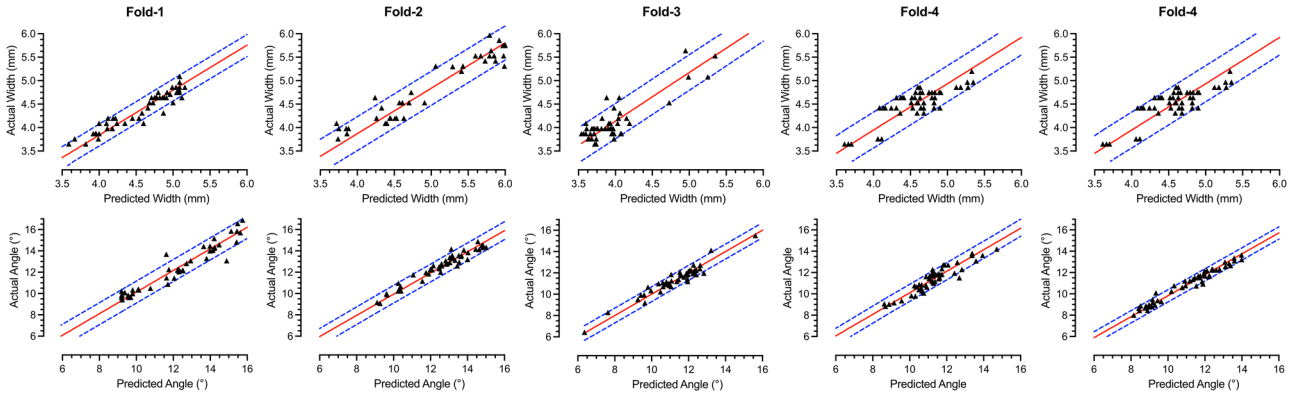

Fig. 1 | The regression plots of PCL width (the first row) and angle (the second row) measurement in the 5-fold cross-validation. Scatters denote the actual values associated with each of the predicted values, where the linear regression results are shown as red solid lines and 90% of the prediction bands are bounded by blue dashed lines.

Table 2 | Error metrics for PCL width measurement across 5-fold, showing mean error (%), standard deviation (STD), and 95% confidence intervals (CIs)

Fold	Mean Error (%)	STD (%)	95% CI Upper Bound (%)	95% CI Lower Bound (%)
Fold-1	4.01	3.34	4.94	3.08
Fold-2	2.91	4.66	4.22	1.61
Fold-3	-3.31	4.96	-1.92	-4.70
Fold-4	1.17	5.04	2.58	-0.25
Fold-5	-0.81	5.29	0.67	-2.29

The STD and CIs are derived from the distribution of the mean error and all metrics are in %.

Table 3 | Error metrics for PCL angle measurement across 5-fold, showing mean error (%), standard deviation (STD), and 95% confidence intervals (CIs)

Fold	Mean Error (%)	STD (%)	95% CI Upper Bound (%)	95% CI Lower Bound (%)
Fold-1	-1.51	4.67	-0.21	-2.82
Fold-2	0.33	3.72	1.37	-0.72
Fold-3	-0.40	3.53	0.59	-1.39
Fold-4	-0.95	3.99	0.17	-2.07
Fold-5	1.50	3.08	2.36	0.64

The STD and CIs are derived from the distribution of the mean error and all metrics are in %.

PCL width measurement

Regression analysis and error analysis are performed to evaluate our performance in estimating the PCL width. The first row in Fig. 1 shows the regression plots of PCL width measurement. The slopes of 5-fold cross-validations are 0.960 (95% CI: 0.951 ~ 0.968), 0.968 (95% CI: 0.956 ~ 0.979), 1.035 (95% CI: 1.019 ~ 1.050), 0.987 (95% CI: 0.973 ~ 1.001), 1.012 (95% CI: 0.997 ~ 1.027), respectively. Slopes close to 1 suggest substantial agreement between the model-predicted PCL width and the actual ones.

As for error analysis, mean error, STD, and 95% CI are derived from the distribution of prediction errors (i.e., the difference between the predicted width and the ground truth). As detailed in Table 2, the mean errors across the five folds ranged from -3.31 to 4.01%, indicating a slight underestimation trend. The maximum mean error of 4.01% across the entire cross-validation suggests that the model is clinically reliable. Despite variability in PCL morphology among participants and the inherent challenges posed by non-linear and indistinct PCL boundaries, the model consistently maintained prediction errors within ± 5%.

The STD values, spanning from 3.34 to 5.29, further demonstrate that large or outlier-level errors are relatively uncommon, with most predictions clustering close to the ground truth. This consistency highlights the model's robustness in accommodating diverse anatomical variations.

The narrow interval of 95% CIs across five folds ensures that the predicted widths are not only accurate on average but also reliable. From a clinical perspective, within the 95% CIs in our model can predict the width values within ± 5% (approximately ± 0.325 mm of the threshold of PCL injury) and enable identifying subtle morphological changes that might signal early pathology or guide treatment decisions.

PCL angle measurement

Regression analysis and error analysis are also performed to evaluate our performance in estimating the PCL angle. The second row in Fig. 1 shows the regression plots of PCL angle measurement. The slopes of 5-fold cross-validations are 1.013 (95% CI: 0.999 ~ 1.027), 0.994 (95% CI: 0.983 ~ 1.005), 1.000 (95% CI: 0.990 ~ 1.010), 1.010 (95% CI: 0.998 ~ 1.021), 0.982 (95% CI: 0.974 ~ 0.991), respectively. Slopes close to 1 suggest substantial agreement between the model-predicted PCL angle and the actual ones.

Similarly to the evaluation of PCL width, the error analysis of PCL angle was based on the distribution of angular prediction errors. As noted in Table 3, all performance metrics are derived from the distribution of prediction errors (i.e., predicted angle minus ground-truth angle).

Across the five folds, the mean error values range from -1.51 to 1.50%. This indicates that for the majority of cases, the model's angle predictions are within less than a single degree of the true angle on average. The minor deviations caused by the model are likely clinically acceptable, given that small angular differences are often within the range of intra- and inter-operator variability associated with manual ultrasound assessments[2].

The STDs of the angle errors, which lie between 3.08 and 4.67, reflect the variability in the model's predictions. Although these STDs indicate that individual angle predictions can sometimes differ from the ground truth, the distribution of errors remains relatively tight and does not suggest large or erratic outliers. This modest dispersion implies that, while perfect angular fidelity is challenging, particularly in ultrasound imaging where probe orientation, soft tissue acoustics, and subtle anatomical shifts can influence the measured angle, the model maintains a commendable level of consistency.

The 95% CIs further contextualize the reliability of the mean error estimates. These narrow CIs highlight that our sample-based mean errors are stable estimates, and the true mean error is unlikely to deviate substantially from the reported values. This level of statistical certainty is important in clinical applications, reinforcing that the model's angular predictions are not only numerically close to the ground truth but also remain reliably.

In essence, the PCL angle detection performance, as quantified by mean errors, STDs, and 95% CIs, reflects a system that can reliably capture subtle orientation details of the ligament. Although improvements remain possible, the current level of performance is already sufficient to enhance

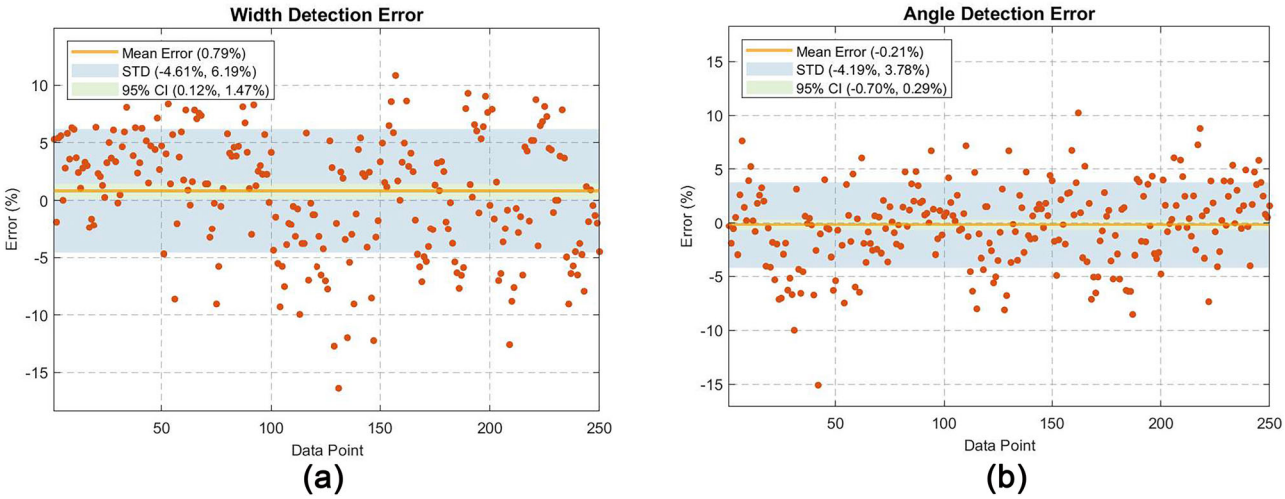

Fig. 2 | Error distributions for PCL width and angle measurements aggregated across all five folds (totaling 250 data points). a Width. **b** Angle. Each dot represents a single prediction error (%) relative to the ground truth. The horizontal yellow line indicates the mean error. The narrow green-shaded region is the 95% Confidence Intervals (CIs) of the mean, representing the precision of this average error. The wider blue-shaded region denotes the Standard Deviation (STD) of the individual errors, representing their overall spread.

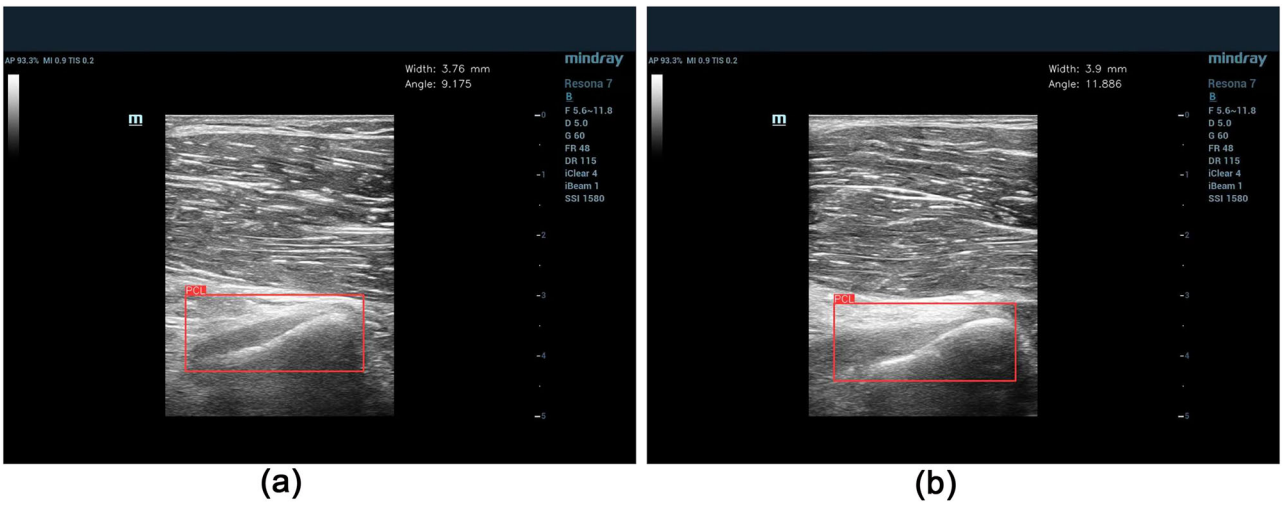

Fig. 3 | Examples of Real-time inference results. The red bounding box in **a** and **b** indicates the location of PCL and PCL's width (in millimeters) and angle (in degrees) after conversion displayed on the upper right of the US image.

ultrasound-based examinations by providing a reproducible, operator-independent angle measurement that complements the model's high-accuracy location and width estimations.

Combined performance across all folds

While the preceding analyses have focused on fold-by-fold evaluation metrics, it is instructive to consider the aggregated performance across all five folds to gain a more holistic picture of the model's overall reliability. Figure 2 presents scatter plots illustrating the distribution of prediction errors for PCL width and angle measurements over a consolidated dataset of 250 data points (50 from each fold, totaling all five folds). Each data point on the x-axis corresponds to an individual prediction, while the y-axis indicates the relative error (%) between the predicted and ground-truth measurements.

For width estimation, the combined mean error was 0.79%, with a STD of approximately 5.40% (spanning -4.61 to 6.19%). Notably, the 95% CI margin was confined between 0.12 and 1.47%, underscoring the statistical stability of the model's width predictions. This suggests that even when pooling all folds together, the system maintains a bias-free and closely clustered distribution of width prediction errors, allowing clinicians to have high confidence in the model's measurements.

Similarly, for angle estimation, the aggregated mean error was -0.21%, with an STD of around 3.98% (ranging from -4.19 to 3.78%). The 95% CI margin for angle predictions remained tight at -0.70 to 0.29%. Such stability is especially valuable in clinical scenarios where small angular differences can influence diagnostic decisions or treatment planning.

In combination, these aggregated results reinforce the model's robustness and generalizability. By examining the entire dataset as a single cohort, we confirm that the system's predictive performance is not unduly influenced by particular folds or subgroups of patients. Instead, the model consistently delivers accurate and reliable estimates of PCL width and angle.

Real-time evaluation

Figure 3 showcases representative examples of the proposed system's real-time inference capabilities. In these instances, the model processes incoming ultrasound frames and instantaneously displays bounding boxes and annotations for the PCL location, PCL width, and PCL posterior and anterior bounds. To prevent the bounding boxes for PCL width

and boundaries from overlapping the PCL and influencing diagnosis, we hide these boxes and quantify their width and angle through post-processing. The measurements are then displayed in the top left corner of the current ultrasound frame, and the annotated measurements are continuously updated.

Unlike measuring methods restricted to offline or post-hoc analyses, our system effectively adapts to slight probe movements, varying imaging angles, and subtle fluctuations in tissue echogenicity. This dynamic evaluation underscores the model's capacity to maintain stable performance under realistic clinical conditions, maximizing the advantages of ultrasound imaging.

Discussion

In this study, we develop a deep learning-based PCL assessment tool that automatically produces critical information for the diagnosis of PCL injury according to current ultrasonic diagnostic criteria. This study introduces a novel labeling method and dedicated post-processing steps that tackle the inherent limitations of classical detection models, which typically output only bounding boxes and confidence scores. In contrast, we designed specialized labeling methods to capture both the anterior and posterior boundaries of the PCL, allowing us to derive its width and angle directly from bounding box geometry. Rather than pursuing a conventional segmentation pathway, which often suffers from high computational cost and limited real-time feasibility, our approach harnesses bounding box diagonal lines to delineate the two ligament boundaries (i.e., the anterior and posterior margins), and a fixed bounding box dimension to calculate the ligament width at 2 cm distal to the tibial insertion. These tailored annotations are then transformed via our post-processing algorithm anchored by (1) and (2) to yield clinically interpretable measurements of PCL width (in millimeters) and angle (in degrees). This explicit linking of bounding box geometry to actual PCL morphology marks a significant advancement over traditional object detection models, which typically provide only box coordinates and require additional, often complicated processing steps to translate those coordinates into clinically useful parameters. By leveraging these carefully designed annotations, our framework circumvents the need for pixel-level segmentation that can introduce contour irregularities and high computational overhead, which is impractically cumbersome in real-time ultrasound settings. Classic segmentation methods usually demand fine-grained, manual delineation of ligament outlines that must be integrated into complex neural networks for boundary refinement. In contrast, our strategy pins down the most clinically relevant aspects—boundary positions, ligament width, and ligament angle—directly through bounding box features that are considerably easier for a detection transformer to learn. Consequently, the final post-processing step utilizes the diagonal lines of our bounding box to calculate the PCL's angle and employs the bounding box dimensions to rescale the ligament width into millimeters. This integrated approach enhances both inference and interpretability, providing a direct, operator-independent method for measuring changes in the PCL's morphology over time or across different ultrasound frames. Through this tailored algorithm, we are able to generate the essential measurement for the diagnosis of PCL tears in a computationally efficient manner such that an ultrasound machine equipped with an average, low-cost processing unit can perform the deep learning-based assessment.

Across all five validation folds, both the width and angle predictions consistently exhibit low standard deviations and narrow CIs, indicating that most errors are small (large erratic deviations are uncommon), although this does not entirely preclude the presence of occasional outlier cases. Given that commercial ultrasound machines can capture over 30 frames per second, and our model processes each frame independently, these low STDs and narrow CIs ensure reliable performance for real-time clinical assessments. Moreover, to aid diagnosis, the assessment tool only needs to detect an abnormal finding in a single frame to prompt further clinical evaluation or additional diagnostics. This reliance on any abnormal instance, rather than continuous aggregated measurements, maintains high clinical relevance with acceptable margins of error.

Since we are capturing the location, width, and angle of PCL in real-time, we can detect and record automatically the maximal width and angle of PCL during the whole course of the US assessment. Since the diagnostic criteria of PCL injury are based on the maximal width and/or angle of PCL above the cut-off of ≥6.5 mm and/or angle of 20° respectively[2], we can review all the detected values after performing the US assessment to see whether any detected values are above the diagnostic cut-off and determine whether it fulfills the diagnosis of PCL injury. This instantaneous measurement value capturing will significantly detail the US assessment unprecedentedly. Notably, all healthy participants in our dataset had PCL width <6.5 mm and angle <20° as measured by our system, consistent with their non-injured status. Without this automatic assessment tool, it would rely on the skills of the clinicians to detect the position of the probe showing the maximum width/angle of PCL, hold the probe still, capture the US frame, and analyze it subsequently. Even the US machine can videotape the US imaging, if the frame showing the maximum width/angle of PCL lasts only for a split second it may not be able to be traced back in subsequent video analysis. Therefore, our automatic assessment tool will largely minimize the dependence on the US skill of the clinicians, potentially making it more operator-independent. From a clinical perspective, the ability to measure the PCL angle with such modest error and consistent reliability offers a new dimension of quantitative assessment. While a mean absolute error of around one degree may not drastically alter a clinical decision—especially compared to large-scale structural measurements like ligament thickness—it can help refine diagnostic criteria, follow-up assessments, and even surgical planning. The stable and predictable nature of these errors also suggests that future refinements to the training data or model architecture, or the integration of additional anatomical landmarks, could further reduce this residual variability. In practical clinical use, we anticipate that clinicians will position the ultrasound probe above the patient's PCL with reasonable accuracy, ensuring that the ligament is consistently visible. However, real-world scenarios may involve significantly poor probe placement, with the probe capturing predominantly tibial or femoral regions or not close to the PCL at all. In such situation, our automatic assessment tool would not be able to salvage such operator error. One of our limitations is that our training dataset did not include such off-target imaging to indicate these errors, which would further enhance the model's adaptability and robustness in dynamic, real-time examinations and inform operators of potential errors. Another limitation is the small sample size we used, however, this sample size is comparable to number of subjects used to derive the ultrasound diagnostic criteria of PCL injury[2,17]. Besides, during model training, we use each instant frame such that the total number of images trained was up to approximately 1310 images, and this method was proven to be successful in our previous publication[31]. Our tool is also limited in that it only directly produces the measurement of PCL's width and angle, without giving a definite diagnosis of any PCL injury. We deliberately make this design because current ultrasonic diagnosis criteria of PCL injury are mainly published by Asian authors[2,17] and whether these criteria apply in Caucasians/black populations with potentially larger knees and PCLs remain unknown.

Therefore, only producing the measurement of PCL's location, width and angle may enhance the applicability of the tool to different populations in the future. As a matter of fact, we think the current limited research and clinical popularity in using ultrasound for diagnosing PCL injury may stem from the difficulty of achieving reproducible PCL's location, width, and angle and many clinicians are not confident in using US for assessing PCL injury. Our automatic PCL assessment tool bridges this clinical technical gap and potentially fosters the development in this clinical and research area. For instance, we can use this tool to produce a more objective measurement of PCL among different ethnicities to see the applicability of the current PCL diagnostic criteria to different populations. Furthermore, given the sensitivity of MRI for chronic (≥6 weeks) primary PCL tears and PCL graft tears can be as low as 62.5% and 18.1%, respectively[15], our automatic ultrasonic PCL assessment tool may potentially perform better than MRI and warrant further investigation. Of note, US can be combined with

dynamic testing, e.g. adding posteriorly directed force to exaggerate the PCL laxity, which would further improve the sensitivity of US compared to static MRI. Our automatic ultrasonic PCL assessment tool lays the foundation to develop further more sophisticated diagnostic criteria and methods for PCL injury by instantaneously capturing the PCL's measurement, which may potentially give rise to a new dynamic assessment criteria for PCL injury, reproducible with this novel, nearly operator-independent assessment tool.

In this study, we present a novel deep learning-based framework that enables real-time, quantitative evaluation of PCL morphology from ultrasound images. By leveraging our labeling method and dedicated post-processing steps, this approach effectively eliminates the need for traditional segmentation models and can be seamlessly adapted to various medical imaging modalities. In doing so, it facilitates reliable, point-of-care assessments of ligament anatomy and function.

Our framework demonstrated near-perfect performance in identifying PCL regions. The mean width errors ranged from -3.31 to 4.01%, with STD between 3.34 and 5.29%. For ligament angle measurements, mean errors ranged from -1.51 to 1.50%, with STD values between 3.08% and 4.67% across a 5-fold cross-validation. These results underscore the model's ability to consistently produce clinically acceptable measurements, even in the face of variable PCL morphologies and suboptimal image clarity.

By integrating established PCL injury criteria—such as ligament width and angle thresholds—our approach can readily function as an injury detection tool. The system provides immediate, operator-independent measurements, thus enhancing the diagnostic utility of ultrasound imaging. This advancement could support ultrasound as a practical, cost-effective alternative or adjunct to MRI in PCL injury evaluation, streamlining clinical decision-making and patient management.

In summary, incorporating real-time, quantitative analysis into standard ultrasound imaging holds profound clinical significance. This method can enhance diagnostic accuracy, improve patient care efficiency, and reduce reliance on more complex and expensive imaging modalities. Future work will refine calibration methodologies for width and angle measurements and broaden the scope of analysis to additional musculoskeletal structures. Ultimately, this approach may become an indispensable tool in clinical practice, aiding timely PCL injury detection, informing targeted treatment strategies, and guiding long-term patient outcomes. Moreover, this technology holds the potential to significantly improve healthcare delivery in rural or low-income regions, where expert sonographers are scarce and patients often must travel long distances for specialized evaluations.

Methods

Data acquisition

Participants for this study were recruited in The University of Hong Kong, comprising 14 healthy volunteers with normal knee joints with no history of fractures or injuries around the knees. The age range of participants was 21–29 years, with a mean age of 23.4 years and a male-to-female ratio of 5:2. Ultrasound imaging was performed using a Mindray Resona 7 system (Mindray Medical, Shenzhen, China) equipped with an L11-3U linear transducer with bandwidth 3.0–10.0 MHz by Orthopedic Surgeon specialized in knee surgery (LCM Lau). Each of the 14 participants underwent seven scanning sessions, yielding a total of 108 video recordings, with each video having an average duration of approximately 6 s (180 frames). This comprehensive dataset was utilized for subsequent analysis and evaluation. Data was collected in a prone position and the knee was fully extended. The probe was positioned longitudinally in the popliteal region, extending from the intercondylar area of the femur to the proximal tibia following previously described ultrasound assessment methods[2,17]. For each video, an Orthopedic Surgeon specialized in knee surgery (LCM Lau) selected a single optimal frame by visual inspection using a priori criteria, in which the PCL is clearly delineated by its anterior and posterior borders for measurement in clinical. All annotations (location, width, and angle) were defined on this operator-selected optimal frame.

Data labeling

To train a supervised computer vision model (such as RT-DETR[32]), we prepared images paired with annotations for training. The annotations of PCL's location, width, and angle were performed by an orthopedic knee specialist (LCM Lau), who also ensured the annotations' consistency. These annotations were stored using a standardized format that includes the bounding box's center coordinates as well as its width and height, represented as $\widehat{\mathbf{Y}}$. The label for the object detection task was parameterized by:

$$\widehat{\mathbf{Y}}_i = \{(\widehat{c}_i, \widehat{x}_i, \widehat{y}_i, \widehat{w}_i, \widehat{h}_i) | i = 1, \ldots, N\},$$

where $\widehat{c}_i$ represents the class label, $(\widehat{x}_i, \widehat{y}_i)$ denotes the center coordinates, and $\widehat{w}_i$ and $\widehat{h}_i$ were the width and height of the bounding box, respectively. In our model, $N = 4$, corresponded to four objects of interest, including the PCL location, PCL posterior bound (i.e., posterior margin of PCL), PCL anterior bound (i.e. anterior margin of PCL), and PCL width.

For PCL location labeling, as shown in Fig. 4, the posterior and anterior boundaries correspond to the PCL's posterior and anterior margins, respectively. The left boundary was the PCL's femoral origin point, while the right boundary corresponds to its tibial insertion point. To label the width, as shown in Fig. 5a, we marked the bounding box's right boundary as the actual width measured between the PCL's posterior and anterior boundaries, taken at 2 cm from the PCL's tibial insertion (i.e., the right boundary of the labeled bounding box), according to M Kimura's and LY Wang's method[2,17].

For angle labeling, as shown in Fig. 5b, we fitted one bounding box's diagonal line to the PCL's anterior boundary and another bounding box's diagonal line to the PCL's posterior boundary. During inference, we can estimate the predicted width by scaling the bounding box dimensions according to each DICOM file's calibration factor, defined by the ultrasound machine's internal parameters. Besides, given the predicted bounding box of posterior bound and anterior bound PCL defined by the coordinates of upper right $(x_{upper-right}, y_{upper-right})$ and lower left $(x_{lower-left}, y_{lower-left})$ points, the predicted angle can be determined by calculating the angle formed by the slope of two diagonal lines as shown in equation (1).

$$m = \frac{y_{upper-right} - y_{lower-left}}{x_{upper-right} - x_{lower-left}} \tag{1}$$

After achieving the slope, we can derive the PCL angle between two diagonal lines from the posterior boundary and anterior boundary of PCL by equation (2).

$$\theta = \arctan\left|\frac{m_{uppper-bound} - m_{lower-bound}}{1 + m_{lower-bound} \cdot m_{uppper-bound}}\right| \tag{2}$$

All ground-truth bounding boxes were labeled using the open-source tool YOLO MARK[33], and measurements were made on the raw DICOM images with the open-source MicroDicom Viewer[34], which can measure PCL width and angle directly. To prioritize consistency and accuracy, a single expert performed all annotations under a fixed protocol as shown in Figs. 4, 5a, b. All measurements were taken at specific, reproducible locations using consistent anatomic landmarks to ensure comparability and reliability across subjects and time points, supporting measurement reliability and validity[35].

Model architecture

Our system adopts the Real-Time Detection Transformer (RT-DETR)[32,36] as illustrated in Fig. 6. RT-DETR is an end-to-end object detection framework that directly predicts a set of object instances without requiring Non-Maximum Suppression (NMS). By formulating detection as a bipartite set-matching problem, RT-DETR circumvents the architectural overhead that often plagues conventional detectors. This end-to-end paradigm is well-aligned with recent advances in transformer-based architectures that have

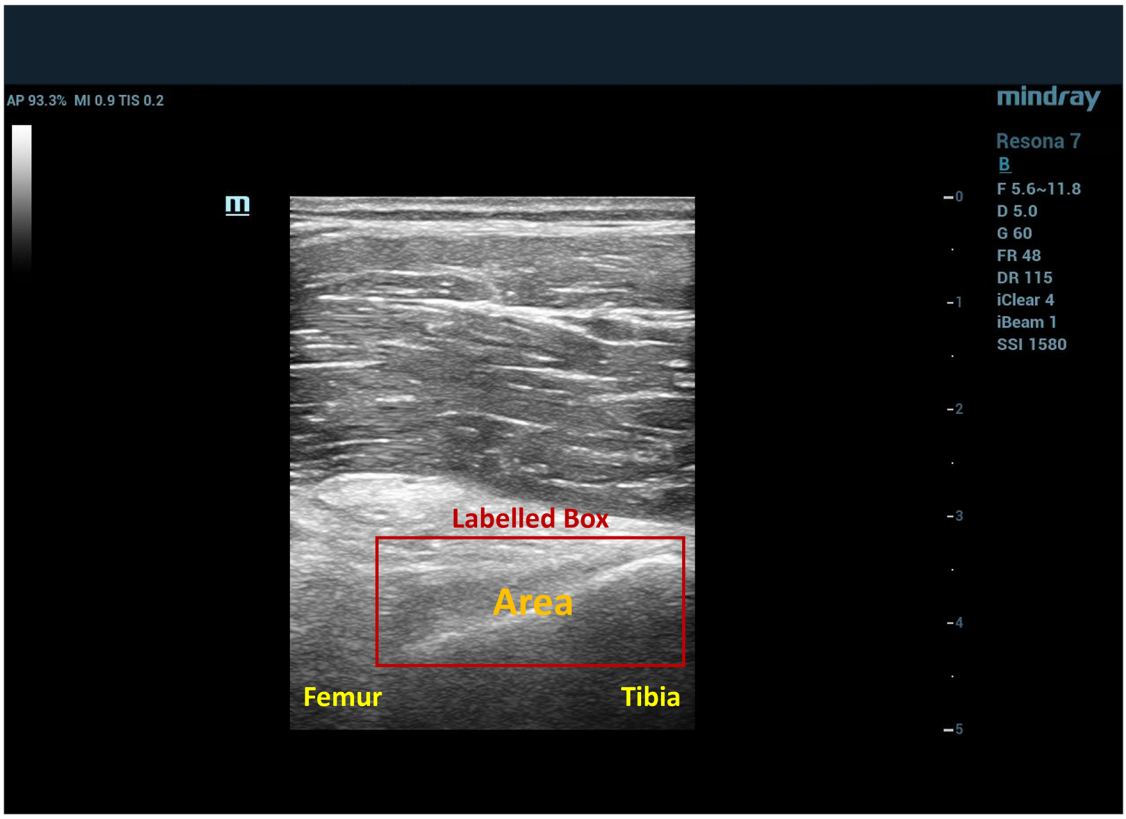

Fig. 4 | Illustration of PCL locationh labeling.

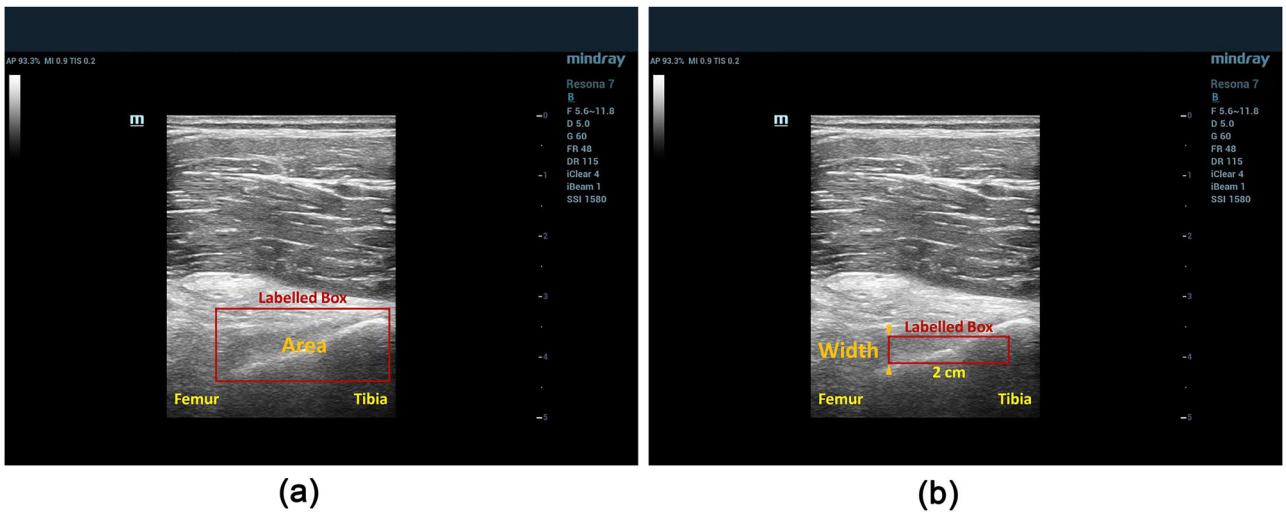

Fig. 5 | Illustration of PCL width and angle labeling. a Width labeling - the PCL's width is measured 2 cm above the tibial insertion by the bounding box's right boundary; **b** Angle labeling - lines are fitted along the PCL's anterior and posterior boundaries to measure the angle between them.

demonstrated remarkable robustness, accuracy, and computational efficiency.

The RT-DETR pipeline extracts multi-scale deep feature representations from the final three stages of a convolutional backbone. These multi-level features {*S*3, *S*4, *S*5} are subsequently processed by an Efficient Hybrid Encoder, which comprises two key modules: Attention-based Intra-scale Feature Interaction (AIFI) and CNN-based Cross-scale Feature Fusion (CCFF). AIFI is a module that employs a single-scale transformer mechanism to refine high-level representations within *S*5. By attending over itself, *S*5 captures global context and inter-relations among semantic concepts corresponding to PCL of interest. CCFF integrates {*S*3, *S*4} with the enhanced *S*5 via hierarchical convolutional fusion. This integration process aggregates low-level geometric cues and high-level semantic information, resulting in a unified set of image features encapsulating rich anatomical details.

Once the integrated feature representation is derived from the encoder, RT-DETR's uncertainty-minimal Query Selection mechanism will select a fixed subset of high-quality encoder features to serve as object queries for the decoder. By explicitly optimizing a feature uncertainty measure during training, this approach ensures that the queries fed into the decoder are both category-aware and localization-aware, thus furnishing the decoder with superior initialization points.

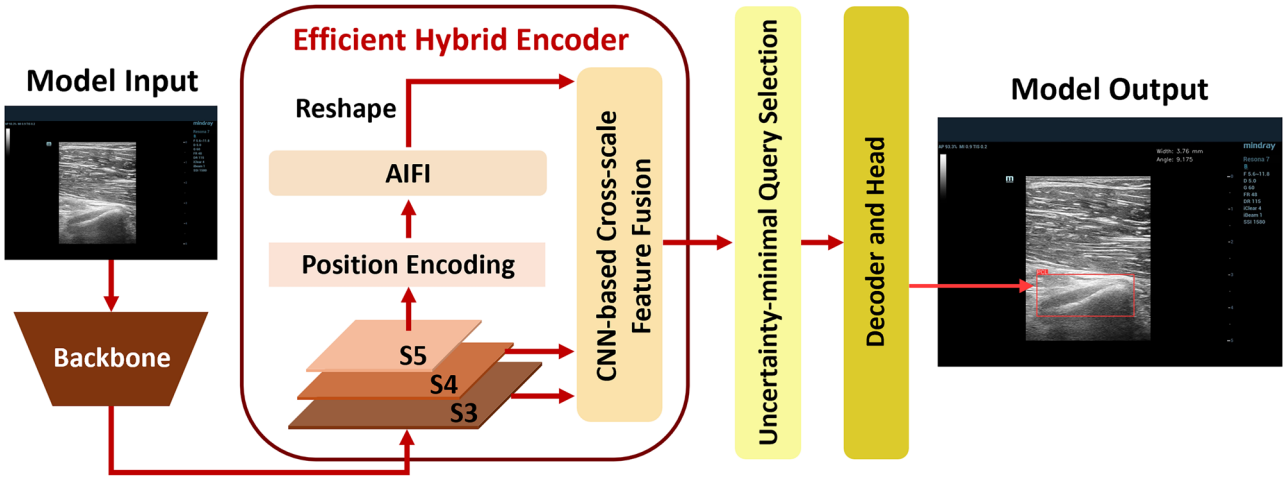

Fig. 6 | Architecture of RT-DETR model[32].

The transformer-based decoder then refines these object queries through iterative cross-attention and feed-forward layers. Each decoder layer progressively enhances the query embeddings to predict accurate bounding boxes and categories. The final output is directly matched to ground-truth annotations via a global, one-to-one bipartite matching. This matching fosters a direct correspondence between ground-truth $(\widehat{c}_i, \widehat{x}_i, \widehat{y}_i, \widehat{w}_i, \widehat{h}_i)$ and predicted instances $(c_i, x_i, y_i, w_i, h_i, s_i)$, eliminating the need for NMS and its associated computational overhead and hyperparameter tuning.

In our framework, the input to the model is an ultrasound frame, denoted as $\mathbf{I} \in \mathbb{R}^{H \times W \times C}$, where **H, W**, and **C** represent the height, width, and number of channels, respectively. In our experiments, we set **H** = 640, **W** = 640, and **C** = 3. The predicted output **Y** is parameterized by:

$$\mathbf{Y}_i = \{(c_i, x_i, y_i, w_i, h_i, s_i) | j = 1, \ldots, N\},$$

where c_i is the predicted class label, (x_i, y_i) are the predicted center coordinates, w_i and h_i are the predicted width and height of the bounding box, and s_i is the confidence score for the detection.

In essence, integrating AIFI and CCFF in the encoder, uncertainty-minimal query selection, and an end-to-end transformer decoder allow RT-DETR to excel in complex ultrasound imaging environments. The model effectively leverages both fine-grained spatial cues and higher-order semantic information to produce reliable, real-time predictions of PCL morphology. This architectural design, grounded in recent top-tier research advancements, paves the way for precise, operator-independent ligament quantification in clinical settings.

Training details

To train the RT-DETR model, we employ a composite loss function that addresses both classification and bounding box regression in a unified framework. Specifically, we use a categorical cross-entropy loss for classification and a combined L1 (Mean Absolute Error) plus Generalized Intersection over Union (GIoU)[37] loss for bounding box regression. This design ensures that the model can accurately identify the target ligament classes (e.g., PCL boundaries) while simultaneously refining spatial localization.

The classification loss, denoted as $\mathcal{L}_{cls}$, is defined by:

$$\mathcal{L}_{cls} = -\sum_k \widehat{p}_k \log(p_k) \tag{3}$$

where p_k is the predicted probability of class k, and $\widehat{p}_k$ is the corresponding ground-truth label. This term penalizes misclassifications by comparing the predicted class distribution to the true distribution, thereby guiding the network toward accurate class predictions. For bounding-box regression, let $\mathbf{b} = (x_i, y_i, w_i, h_i)$ and $\widehat{\mathbf{b}} = (\widehat{x}_i, \widehat{y}_i, \widehat{w}_i, \widehat{h}_i)$ denote the predicted and ground-truth bounding boxes, respectively. We first apply an L1 loss:

$$\| \mathbf{b} - \widehat{\mathbf{b}} \|_1 = \sum_{i \in \{x,y,w,h\}} | b_i - \widehat{b}_i | \tag{4}$$

This simple yet effective penalty rapidly drives the predicted coordinates toward the ground-truth values, establishing a coarse but robust alignment.

To further enhance localization, we incorporate a GIoU term, an extension of the standard Intersection over Union (IoU). IoU and GIoU are defined as:

$$\mathrm{IoU}(\mathbf{b}, \widehat{\mathbf{b}}) = \frac{|\mathbf{b} \cap \widehat{\mathbf{b}}|}{|\mathbf{b} \cup \widehat{\mathbf{b}}|} \tag{5}$$

$$\mathrm{GIoU}(\mathbf{b}, \widehat{\mathbf{b}}) = \mathrm{IoU}(\mathbf{b}, \widehat{\mathbf{b}}) - \frac{|\mathbf{C} \backslash (\mathbf{b} \cup \widehat{\mathbf{b}})|}{|\mathbf{C}|} \tag{6}$$

where b and $\widehat{b}$ represent the predicted and ground-truth boxes, respectively, and C denotes the smallest convex set enclosing both b and $\widehat{b}$. By incorporating the surrounding region, GIoU penalizes partial overlaps more aggressively than standard IoU, thereby encouraging more precise boundary alignment.

The final loss function for RT-DETR is given by the weighted sum:

$$\mathcal{L}_{\mathrm{RT-DETR}} = \lambda_{\mathrm{cls}} \mathcal{L}_{\mathrm{cls}}(\mathbf{p}, \widehat{\mathbf{p}}) + \lambda_{\mathrm{bbox}} \| \mathbf{b} - \widehat{\mathbf{b}} \|_1 + \lambda_{\mathrm{giou}} \left[1 - \mathrm{GIoU}(\mathbf{b}, \widehat{\mathbf{b}}) \right] \tag{7}$$

where the hyperparameters

$$\lambda_{\mathrm{cls}} = 1 \, \lambda_{\mathrm{bbox}} = 5 \, \lambda_{\mathrm{giou}} = 2$$

control the relative contributions of the classification, L1, and GIoU terms, respectively. This holistic approach allows the network to learn both accurate class assignments and precise object localization simultaneously to yield accurate PCL detections and measurements in real-time.

We conducted patient-wise, stratified 5-fold cross-validation to robustly assess generalization while preventing information leakage across folds. In each fold, we used approximately 1,310 images from 10 participants for training and 570 images from 4 participants for validation, preserving an overall 7:3 ratio and maintaining the class/view distribution. We selected

k = 5 as a bias-variance and feasibility trade-off tailored to our cohort: (i) patient-level splitting prevents correlated samples from the same subject appearing in both sets; (ii) with k = 5, the validation cohort per fold is sufficiently large to stabilize estimates, whereas larger k would yield very small validation cohorts (1–2 participants) and higher fold-to-fold variance, and smaller k would increase estimator bias; and (iii) the computational cost remains practical, enabling repeated runs and comprehensive uncertainty reporting across folds.

For parameter optimization, we employ the official default training configuration[32,36], with Stochastic Gradient Descent (SGD) with a learning rate of 0.01, momentum of 0.9, and a weight decay of 5×10^{-4} and kept it fixed for all five folds to ensure a fair comparison and reproducibility. Training is performed over 200 epochs on an NVIDIA GeForce RTX 3090 GPU (24 GB VRAM), completing within approximately two hours per fold.

To enhance model resilience against diverse medical imaging scenarios, we incorporate data augmentation techniques commonly used in medical image analysis. These include mosaic augmentation (randomly merging multiple images into a single composite sample), random horizontal flipping, random scaling, and adjustments to brightness and contrast and all listed augmentations were applied with set probabilities, ensuring each training image experienced a random combination of transformations. Such augmentations effectively simulate tissue's echogenicity and probe placement angles to handle real-world challenges in ultrasound imaging detection.

Evaluation metrices

To evaluate the performance of our system in detecting and quantifying PCL morphology, six evaluation metrics are involved. For PCL location detection, precision, recall, and mean Average Precision (mAP) at both the 0.5 (mAP50) and 0.5 to 0.95 (mAP50:95) Intersection-over-Union (IoU) thresholds are evaluated.

As shown in Eq. (8) ~ (9), precision measures the proportion of detected PCL regions that are correctly identified, while recall evaluates the percentage of true PCL instances successfully detected.

$$\text{Precision}(\mathbf{Y}, \widehat{\mathbf{Y}}) = \frac{TP}{TP + FP}, \tag{8}$$

$$\text{Recall}(\mathbf{Y}, \widehat{\mathbf{Y}}) = \frac{TP}{TP + FN}, \tag{9}$$

where TP (True Positives) is the count of correctly predicted PCL location when $c_i = \widehat{c}_i$, FP (False Positives) is the count of incorrectly predicted PCL locations when $c_i \neq \widehat{c}_i$, and FN (False Negatives) is the count of actual PCL locations that were not predicted, i.e. $\widehat{c}_i$ is PCL but c_i does not match. Eq. (10) shows the average precision AP_t at a specific IoU threshold t. As shown in Eq. (11) ~ (12), the $\text{mAP}_{0.5}$ condense an averaged detection performance at 50% IoU threshold over four objects of interest; $\text{mAP}_{0.5:0.95}$ further evaluates mAP across a range of IoU thresholds from 50% to 95%. Those two metrics capture how well the model localizes the ligament boundaries under varying conditions of spatial overlap.

$$\text{AP}(\mathbf{Y}, \widehat{\mathbf{Y}})_t = \sum_{k=1}^{K} \text{Precision}(\mathbf{Y}, \widehat{\mathbf{Y}})_k \cdot \Delta\text{Recall}(\mathbf{Y}, \widehat{\mathbf{Y}})_k, \tag{10}$$

$$\text{mAP}_{0.5}(\mathbf{Y}, \widehat{\mathbf{Y}}) = \frac{1}{N} \sum_{n=1}^{N} \text{AP}^n_{0.5}, \tag{11}$$

$$\text{mAP}_{0.5:0.95}(\mathbf{Y}, \widehat{\mathbf{Y}}) = \frac{1}{10} \sum_{t=0.5}^{0.95} \text{mAP}_t(\mathbf{Y}, \widehat{\mathbf{Y}}), \tag{12}$$

where K is the number of distinct recall levels, $\Delta\text{Recall}_k = \text{Recall}_k - \text{Recall}_{k-1}$ is the change in recall between consecutive levels, $N = 4$ denotes four objects of interest.

Beyond detection, we assessed the performance of PCL width and angle measurements via regression analysis and error analysis. In our error analysis, we calculated the mean error, which indicates the systematic bias of the model's measurements and the standard deviation (STD), which measures the spread of individual prediction errors. A small STD signifies that large, outlier-level errors are rare. Additionally, to evaluate the precision of our estimate of the model's bias, we derived the 95% confidence intervals (CIs) for the mean error. These statistical bounds reflect the uncertainty in our estimation of the mean, with narrow CIs suggesting a stable and reliable measure of the model's central tendency. In regression analysis, the slopes of the linear fits are used to measure the agreements between the predicted and actual values, where a slope approaching 1 indicates a high consistency. 90% of prediction bands and 95% CIs are also calculated to indicate the stability of the measurement.

Clinically anchored acceptance criteria

Prior to any model evaluation, we pre-specified clinically anchored acceptance targets tied to decision thresholds that are used in ultrasonographic screening for chronic PCL injury[2]. The primary reference reports an optimal cut-off of 6.5 mm for PCL thickness (uninjured: 5.8 ± 1.2 mm; injured: 8.1 ± 1.9 mm) and 20° for the PCL angle. Anchoring to these thresholds, we defined a tolerance as measurement errors that are small relative to the decision boundary and thus unlikely to alter clinical classification except for cases extremely close to the boundary.

Operationally, we evaluated errors as percentages and mapped them to absolute deviations at the clinical cut-offs to facilitate interpretation in clinical units (millimeters and degrees). A priori, we set a conservative tolerance of ± 10% of the decision cut-off (i.e., ± 0.65 mm at 6.5 mm; ± 2° at 20°), and required the fold-wise 95% confidence intervals (CIs) of the mean error (bias) to lie within these margins. This tolerance is substantially smaller than between-group differences reported in the literature (e.g., radiographic posterior laxity 7.5 ± 3.5 mm in chronic PCL vs. 1.2 ± 1.0 mm in controls), providing clinical assurance that observed errors would not obviate the diagnostic signal[2].

Data availability

The datasets generated and analyzed during the current study involve human participants and are therefore not publicly available due to privacy and institutional governance constraints. De-identified ultrasound frames and corresponding annotations sufficient to verify the main findings will be made available from the corresponding author upon reasonable request.

Code availability

The underlying code for this study and the pretrained weights are available from the corresponding author on reasonable request. To preserve patent novelty and comply with institutional policies, unrestricted public release of the full tool and trained model weights is not possible until the institutional intellectual property (IP) filing is complete.

Received: 9 January 2025; Accepted: 26 November 2025;
Published online: 07 January 2026

References

1. Pache, S. et al. Posterior cruciate ligament: current concepts review. *Arch. Bone Jt. Surg.* **6**, 8 (2018).
2. Kimura, M. et al. The use of ultrasonography as an effective screening tool for chronic posterior cruciate ligament injuries. *J. Med. Ultrason.* **51**, 109–115 (2024).
3. Wu, F., Colak, C. & Subhas, N. Preoperative and postoperative magnetic resonance imaging of the cruciate ligaments. *Magn. Reson. Imaging Clin.* **30**, 261–275 (2022).
4. Harner, C. D. et al. The human posterior cruciate ligament complex: an interdisciplinary study: ligament morphology and biomechanical evaluation. *Am. J. sports Med.* **23**, 736–745 (1995).
5. Patel, D. V., Allen, A. A., Warren, R. F., Wickiewicz, T. L. & Simonian, P. T. The nonoperative treatment of acute, isolated (partial or complete)

posterior cruciate ligament-deficient knees: an intermediate-term follow-up study. *HSS J.* **3**, 137–146 (2007).
6. Fanelli, G. C. & Edson, C. J. Posterior cruciate ligament injuries in trauma patients: part ii. *Arthrosc. J. Arthroscopic Relat. Surg.* **11**, 526–529 (1995).
7. Colvin, A. C. Meislin, R. J. Posterior cruciate ligament injuries in the athlete. *Bull. NYU Hosp. Joint Dis.* **67**, 291–298 (2009).
8. Verhulst, F. V. & MacDonald, P. Diagnosing pcl injuries: history, physical examination, imaging studies, arthroscopic evaluation. *Sports Med. Arthrosc. Rev.* **28**, 2–7 (2020).
9. Kopkow, C., Freiberg, A., Kirschner, S., Seidler, A. & Schmitt, J. Physical examination tests for the diagnosis of posterior cruciate ligament rupture: a systematic review. *J. Orthop. sports Phys. Ther.* **43**, 804–813 (2013).
10. Strobel, M. J. & Weiler, A. Management of the posterior cruciate ligament-deficient knee. *Tech. Orthop.* **16**, 167–194 (2001).
11. Gupta, A., Bagri, N., Spalkit, S. & Jana, M. Technical and interpretive pitfalls in pediatric musculoskeletal ultrasound. *J. Ultrasound Med.* **42**, 1639–1646 (2023).
12. Rubinstein JR, R. A., Shelbourne, K. D., McCarroll, J. R., VanMeter, C. D. & Rettig, A. C. The accuracy of the clinical examination in the setting of posterior cruciate ligament injuries. *Am. J. sports Med.* **22**, 550–557 (1994).
13. Logan, M., Williams, A., Lavelle, J., Gedroyc, W. & Freeman, M. The effect of posterior cruciate ligament deficiency on knee kinematics. *Am. J. sports Med.* **32**, 1915–1992 (2004).
14. Wind Jr, W. M., Bergfeld, J. A. & Parker, R. D. Evaluation and treatment of posterior cruciate ligament injuries: revisited. *Am. J. sports Med.* **32**, 1765–1775 (2004).
15. DePhillipo, N. N. et al. Posterior tibial translation measurements on magnetic resonance imaging improve diagnostic sensitivity for chronic posterior cruciate ligament injuries and graft tears. *Am. J. Sports Med.* **46**, 341–347 (2018).
16. Jung, Y.-B. et al. Characterization of spontaneous healing of chronic posterior cruciate ligament injury: analysis of instability and magnetic resonance imaging. *J. Magn. Reson. Imaging* **27**, 1336–1340 (2008).
17. Wang, L.-Y. et al. Evaluating posterior cruciate ligament injury by using two-dimensional ultrasonography and sonoelastography. *Knee Surg. Sports Traumatol. Arthrosc.* **25**, 3108–3115 (2017).
18. Koff, M. F., Burge, A. J. & Potter, H. G. Clinical magnetic resonance imaging of arthroplasty at 1.5 t. *J. Orthop. Res.* **38**, 1455–1464 (2020).
19. Stoianov, A. G. et al. Point-of-care ultrasound for the evaluation and management of posterior cruciate ligament injuries: a systematic review. *Diagnostics* **13**, 2352 (2023).
20. Sharifi, Y. et al. Deep learning on ultrasound images of thyroid nodules. *Biocybern. Biomed. Eng.* **41**, 636–655 (2021).
21. Ramirez Zegarra, R. & Ghi, T. Use of artificial intelligence and deep learning in fetal ultrasound imaging. *Ultrasound Obstet. Gynecol.* **62**, 185–194 (2023).
22. Luan, K., Li, Z. & Li, J. An efficient end-to-end cnn for segmentation of bone surfaces from ultrasound. *Computerized Med. Imaging Graph.* **84**, 101766 (2020).
23. Guo, Y., Du, G.-Q., Xue, J.-Y., Xia, R. & Wang, Y. -h A novel myocardium segmentation approach based on neutrosophic active contour model. *Comp. Methods Prog. Biomed.* **142**, 109–116 (2017).
24. Singh, V. et al. Calcaneofibular ligament ultrasound image segmentation based on advanced image processing techniques. In *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 1–4 (IEEE, 2017).
25. Shen, W. et al. A survey on label-efficient deep image segmentation: Bridging the gap between weak supervision and dense prediction. *IEEE Trans. pattern Anal. Mach. Intell.* **45**, 9284–9305 (2023).
26. Patil, P. & Dasgupta, B. Role of diagnostic ultrasound in the assessment of musculoskeletal diseases. *Therapeutic Adv. Musculoskelet. Dis.* **4**, 341–355 (2012).
27. Ramakrishnan, A. H. et al. A systematic survey on segmentation algorithms for musculoskeletal tissues in ultrasound imaging. *Arch. Comput. Methods Eng.* **32**, 1–34 (2024).
28. Xue, Y. et al. Approaching expert-level accuracy for differentiating acl tear types on mri with deep learning. *Sci. Rep.* **14**, 938 (2024).
29. Shi, X. et al. Deep learning-enabled ultrasound for advancing anterior talofibular ligament injuries classification: a multicenter model development and validation study. *Academic Radiology* (2025).
30. Zarychta, P. Features extraction in anterior and posterior cruciate ligaments analysis. *Computerized Med. Imaging Graph.* **46**, 108–120 (2015).
31. Kao, J.-P., Chung, Y.-C., Hung, H.-Y., Chen, C.-P. & Chen, W.-S. Lora-enhanced RT-DETR: first low-rank adaptation based detr for real-time full body anatomical structures identification in musculoskeletal ultrasound. *Computerized Med. Imag. Graph.* **124**, 102583 (2025).
32. Zhao, Y. et al. Detrs beat yolos on real-time object detection. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 16965–16974 (IEEE, 2024).
33. Bochkovskiy, A. Yolo_mark [software]. https://github.com/AlexeyAB/Yolo_mark GitHub repository, commit ea049f3. Accessed 2024-10-25 (2019).
34. Ltd., M. Microdicom dicom viewer [software]. https://www.microdicom.com/downloads.html Version 2024.2 Released 2024-06-06. Accessed 2025-09-23 (2024).
35. Hutchinson, M. R. & Bae, T. S. Reproducibility of anatomic tibial landmarks for anterior cruciate ligament reconstructions. *Am. J. Sports Med.* **29**, 777–780 (2001).
36. Ultralytics. Ultralytics [software]. https://github.com/ultralytics/ultralytics Version 8.2.1; accessed on 25 Oct 2024 (2024).
37. Rezatofighi, H. et al. Generalized intersection over union: A metric and a loss for bounding box regression. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 658–666 (IEEE, 2019).

Acknowledgements

This research received no specific grant from any funding agency in the public, private, or not-for-profit sectors. We extend our heartfelt gratitude to all the volunteer participants from The University of Hong Kong and National Taiwan University who generously contributed their time and effort to undergo ultrasound examinations and provide valuable data for our research.

Author contributions

J.P.K., J.Z., and L.C.M.L. conducted the experiment. L.C.M.L. supervised the entire work. W.N.L. supervised the experiments and provided the equipment for ultrasound examinations. C.P.C. provided the computing resources for model training. J.P.K., J.Z., and L.C.M.L. wrote the main manuscript text. J.Z. prepared Figure 1. J.P.K. prepared Figures 2–6.All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Lawrence Chun-Man Lau.