

Hierarchical classification of anterior cruciate ligament using deep learning for athletes healthcare

Xuejiao Yan¹, Lei Xiao^{2*} 

¹Sichuan Railway College, Chengdu 611732, China

²Chengdu Technological University, Chengdu 611730, China

Article Info



Article Type:
Original Article

Article History:
Received: 1 Mar. 2025
Revised: 8 May 2025
Accepted: 26 Jun. 2025
ePublished: 4 Nov. 2025

Keywords:
ACL injury classification
Deep learning
Knee MRI analysis
Hierarchical CNN
Athlete healthcare
Sports medicine

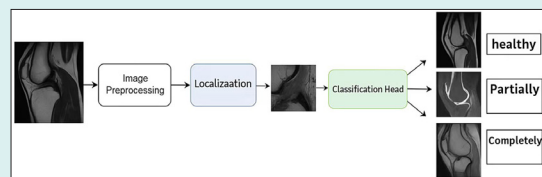
Abstract

Introduction: Accurate and automated assessment of anterior cruciate ligament (ACL) injuries in MR images is essential for athlete healthcare and rapid diagnosis of knee injuries. However, challenges such as the small size of the ligament, variations in MR image quality, and complex anatomical structures complicate the classification process.

Methods: In this study, we propose a hierarchical deep learning model for the detection and classification of ACL injuries. The model consists of two main phases: ACL segmentation and injury classification. In the first phase, we employ an encoder-decoder architecture with attention mechanisms to accurately identify the ACL region in MR images, while suppressing background noise. Skip connections are used to preserve spatial details and improve segmentation accuracy. In the second phase, the segmented ACL region is input into a hierarchical convolutional neural network (CNN) for classification. Dense blocks are incorporated to maximize feature reuse, while max-pooling and global average pooling (GAP) layers help to reduce overfitting and improve feature extraction.

Results: The proposed method was evaluated on a knee MRI dataset and compared with other state-of-the-art approaches. Our model demonstrated high accuracy in both segmentation and classification tasks, owing to the integration of attention mechanisms and hierarchical feature extraction.

Conclusion: This approach offers a robust solution for the automated assessment of ACL injuries, providing clinicians and sports medicine specialists with a reliable tool for more efficient and accurate diagnosis.



Introduction

Sports injuries are a significant challenge in the field of athlete healthcare, with anterior cruciate ligament tears being one of the most common and severe injuries. This type of injury frequently occurs in sports that involve sudden changes in speed and direction, such as football, basketball, and skiing. The timely and accurate diagnosis of anterior cruciate ligament (ACL) injuries is crucial for effective treatment and reducing long-term complications. Traditional diagnostic methods, such as clinical examinations and MR imaging, have limitations, including high costs, dependency on specialist expertise, and limited accessibility. In recent years, the emergence of artificial intelligence (AI) and deep learning technologies has introduced innovative solutions for diagnosing and classifying sports injuries. Deep learning models can analyze medical images and biomechanical data to

identify and categorize ACL injuries with high accuracy. However, developing an automated hierarchical classification system for ACL injuries presents several challenges, such as data imbalance, variations in injury characteristics, and the complexity of hidden patterns in medical imaging data. To address these challenges, researchers have explored various approaches, including convolutional neural networks, transfer learning models, and the integration of multimodal data (such as MR images and clinical information). In this study, we propose a hierarchical deep learning-based classification model for ACL injury detection, aiming to improve diagnostic accuracy and reduce reliance on traditional methods. This approach has the potential to serve as an assistive tool for medical professionals and sports specialists, ultimately enhancing athlete healthcare and injury management.

The diagnosis and classification of anterior cruciate



*Corresponding author: Lei Xiao, Email: xlei1@cdtu.edu.cn



© 2025 The Author(s). This work is published by BioImpacts as an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>). Non-commercial uses of the work are permitted, provided the original work is properly cited.

ligament injuries have been extensively explored using deep learning techniques, particularly in conjunction with magnetic resonance imaging. Recent advancements in AI have led to the development of fully automated systems for ACL tear detection, segmentation, and severity staging. This section reviews significant contributions in this field, focusing on segmentation techniques, classification models, severity assessment, and multimodal learning approaches.

Numerous studies have employed deep learning-based methodologies to detect and classify ACL injuries from MRI scans. Bien et al¹ introduced a fully automated system utilizing convolutional neural networks for ACL tear diagnosis, achieving accuracy levels comparable to those of experienced radiologists. Yao et al² proposed an efficient deep learning approach that required minimal preprocessing, demonstrating superior performance over conventional methods. Thomas et al³ explored an improved deep convolutional neural network (CNN) model for distinguishing ACL tears from osteoarthritis, highlighting the importance of robust feature extraction. Further advancements in this domain include Liu et al,⁴ who developed an end-to-end deep learning model for ACL segmentation and severity staging, enabling precise injury localization. The U-Net architecture, initially proposed by Ronneberger et al,⁵ has been widely adopted for medical image segmentation, including ACL tear detection. Chen et al⁶ refined the U-Net model by integrating attention mechanisms, significantly enhancing segmentation accuracy.

Several other studies have introduced novel approaches to classification. Gong et al⁷ combined CNNs with recurrent neural networks (RNNs) to assess ACL injury severity from MRI scans. Wang et al⁸ investigated transformer-based architectures for severity classification, yielding promising results in complex MRI datasets. Zhang et al⁹ proposed a multi-scale feature fusion approach to enhance severity staging performance. Gupta et al¹⁰ incorporated clinical data with MR images using a hybrid CNN-LSTM model, achieving higher diagnostic accuracy. Additional techniques have been developed to optimize ACL injury detection and classification. He et al¹¹ leveraged transfer learning from large-scale medical imaging datasets to improve classification robustness. Huang et al¹² introduced a self-supervised learning approach to mitigate data scarcity issues in ACL injury detection. Kim et al¹³ compared CNN-based models with radiologist assessments, reporting a reduction in diagnostic errors. Patel et al¹⁴ emphasized the role of AI-assisted diagnosis in alleviating the workload of medical professionals.

Automated segmentation is a critical component of ACL injury assessment. Singh et al¹⁵ employed ensemble learning techniques to reduce model variance and enhance classification robustness. Xie et al¹⁶ applied semi-supervised learning to improve classification

performance when labeled data availability was limited. Park et al¹⁷ integrated generative adversarial networks (GANs) to refine ACL segmentation, effectively reducing artifacts in low-quality MRI scans. Additionally, Zhao et al¹⁸ explored explainable AI techniques to generate interpretable severity assessments for clinical applications. Li et al¹⁹ introduced meta-learning strategies to adapt severity classification models to varying MRI protocols and patient demographics. Sun et al²⁰ developed a reinforcement learning-based segmentation framework, dynamically optimizing model parameters during training. Recent studies have investigated multimodal data fusion and transfer learning to enhance ACL injury diagnosis. Ahmed et al²¹ proposed a multi-view CNN model that integrates MRI sequences from different planes to improve classification accuracy. Lee et al²² introduced an interpretable deep learning model for ACL tear diagnosis, providing clinicians with enhanced transparency in decision-making. Feng et al²³ employed contrastive learning to improve feature representation in ACL classification, outperforming traditional handcrafted feature extraction techniques. Several studies have also explored domain adaptation methods. Zhang et al²⁴ utilized capsule networks for ACL tear classification, demonstrating robust feature learning. Han et al²⁵ implemented federated learning for ACL tear detection across multiple institutions, enabling privacy-preserving model training while enhancing generalization performance. Xu et al²⁶ investigated multi-scale CNN architectures, achieving improved performance on heterogeneous MRI datasets.

Deep learning-based methods have consistently outperformed traditional machine learning and manual assessment techniques. Li et al²⁷ explored knowledge distillation techniques to optimize deep learning models for ACL tear classification. Wu et al²⁸ investigated data augmentation strategies, including synthetic MRI generation, to improve segmentation accuracy in limited datasets. Shi et al²⁹ applied graph neural networks for ACL injury prediction, leveraging spatial relationships within MRI features. Further improvements in model efficiency and clinical deployment have been explored. Gao et al³⁰ integrated radiology reports with MRI data to enhance classification accuracy. Lin et al³¹ examined cross-domain transfer learning, utilizing knowledge from related musculoskeletal imaging tasks. Cheng et al³² introduced a self-attention mechanism for effectively fusing multi-sequence MRI data in ACL diagnosis. Li et al³³ proposed a deep ensemble learning strategy to enhance ACL injury detection robustness. Sun et al³⁴ developed an edge AI solution for real-time ACL tear detection on portable medical devices, facilitating on-site athlete assessment. Zhou et al³⁵ applied Bayesian deep learning techniques to estimate uncertainty in ACL injury classification, providing clinicians with confidence measures for AI-assisted

diagnoses. An end-to-end deep learning model, DCLU-Net, was proposed for the simultaneous segmentation and classification of ACL injuries. By incorporating radiomic features, the model achieved classification accuracies of 90% for intact ACLs, 82% for partial tears, and 92% for complete ruptures. Additionally, the use of supervised learning techniques reduced the reliance on extensive manual annotations.³⁶ In another recent study, a fully automated deep learning framework was introduced, comprising two models: ACL-DNet for segmentation and ACL-SNet for classification. The classification model achieved a sensitivity and specificity of 97%, along with an area under the receiver operating characteristic curve (AUC) of 0.99, outperforming experienced orthopedic specialists in diagnostic accuracy.³⁷ Furthermore, a modified 3D ResNet architecture was employed for the detection and classification of ACL injuries. This model achieved a peak accuracy of 97.15% using custom data splits and demonstrated substantial improvements over conventional three-class classifiers, thereby enhancing diagnostic precision in sports medicine applications.³⁸ In summary, deep learning has significantly advanced ACL injury detection, segmentation, and severity staging. While CNN-based models remain predominant, emerging techniques such as transformers, contrastive learning, and multimodal fusion are enhancing diagnostic accuracy. Future research should focus on improving model generalizability, interpretability, and real-time clinical deployment to optimize athlete healthcare and injury prevention strategies. Given the challenges outlined in this paper, a method based on Hierarchical Classification of Anterior Cruciate Ligament Injuries Using Deep Learning for Athlete Healthcare is proposed. This approach aims to address the difficulties associated with small ligament size, varying MR image quality, and complex anatomical features, offering a more accurate and efficient way to detect and classify ACL injuries. The key innovations of this method include the use of an encoder-decoder model with an attention mechanism for precise segmentation of the ACL region, which allows the model to focus on relevant areas and filter out background noise. Additionally, the hierarchical deep learning

architecture enhances the classification performance by leveraging dense blocks, maximizing feature reuse, and improving gradient flow. These innovations contribute to superior results in both segmentation and classification, outperforming several existing models

Following is the structure of the article. After the introduction, the paper provides a detailed explanation of the proposed method. Following that, the dataset used for the experiments is introduced, and the method is thoroughly evaluated from various aspects. Finally, the conclusion is presented.

Proposed Method

Accurate localization of the anterior cruciate ligament plays a crucial role in the automated detection and classification of ACL injuries in MR images. However, detecting small objects in medical imaging presents several challenges due to the complexity of anatomical structures and variations in image contrast. Traditional deep learning models often struggle to accurately identify small objects like the ACL, as critical features may be lost during the down-sampling process. To overcome this challenge, we propose an encoder-decoder-based model that leverages attention mechanisms to enhance feature extraction. This model effectively transfers spatial and contextual information through skip connections, improving segmentation accuracy while preserving fine structural details. Once the ACL is precisely localized, the next step is to classify its condition into three main categories: healthy, partial tear, and complete tear. This classification is of significant clinical importance, as different injury levels require distinct treatment strategies. To achieve this, we employ a deep convolutional neural network with dense blocks, which enables the extraction of discriminative features from MRI scans. The proposed model integrates a combination of convolutional layers, max-pooling, dense feature extraction, and global average pooling. This combination enhances classification accuracy and improves the model's ability to distinguish between different ACL conditions. Fig. 1 illustrates the overall stages of the proposed method. Next, each step will be explained in more detail.

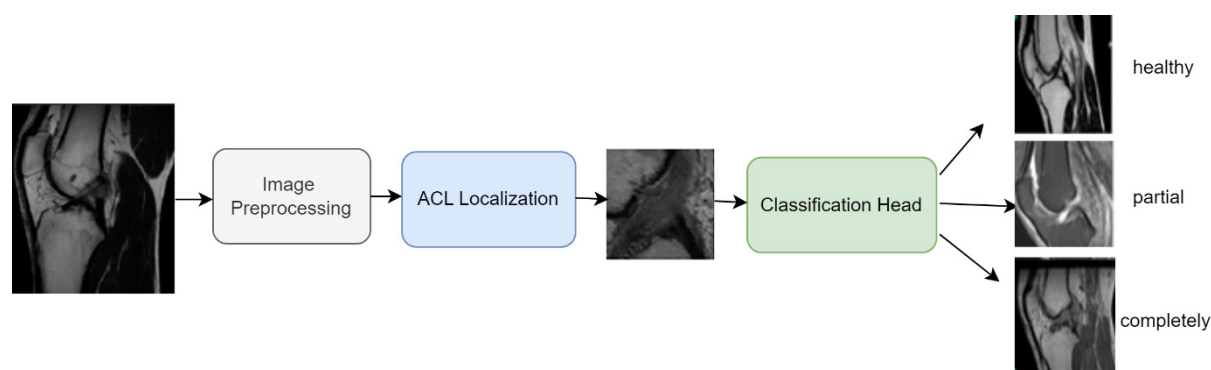


Fig. 1. Main steps of the Hierarchical ACL tear model.

Image preprocessing

In the proposed method, the data preprocessing phase involves resizing all MR images to a fixed dimension of 256×256 pixels to ensure uniformity in input dimensions. This step enhances the performance of the neural network by allowing it to extract relevant features without being affected by variations in image size. In addition to resizing, other preprocessing techniques such as pixel intensity normalization are applied to homogenize brightness and contrast levels, and noise reduction is performed to eliminate irrelevant information.

ACL localization

After preprocessing, the next phase is the localization of the anterior cruciate ligament. This step is crucial as it serves as the foundation for subsequent injury classification.

The proposed method utilizes an encoder-decoder architecture enhanced with attention mechanisms to effectively detect and segment the ACL region. The encoder extracts hierarchical features from the input MR scan, while the attention mechanism ensures a focused feature representation by highlighting relevant regions. Finally, the decoder reconstructs the spatial details to generate an accurate segmentation mask. This phase is illustrated in Fig. 2.

Encoder module

The encoder module is responsible for extracting hierarchical feature representations from the input MR image. It consists of multiple convolutional layers followed by down-sampling operations, which progressively reduce the spatial dimensions while increasing the depth of the feature maps. This hierarchical extraction enables the

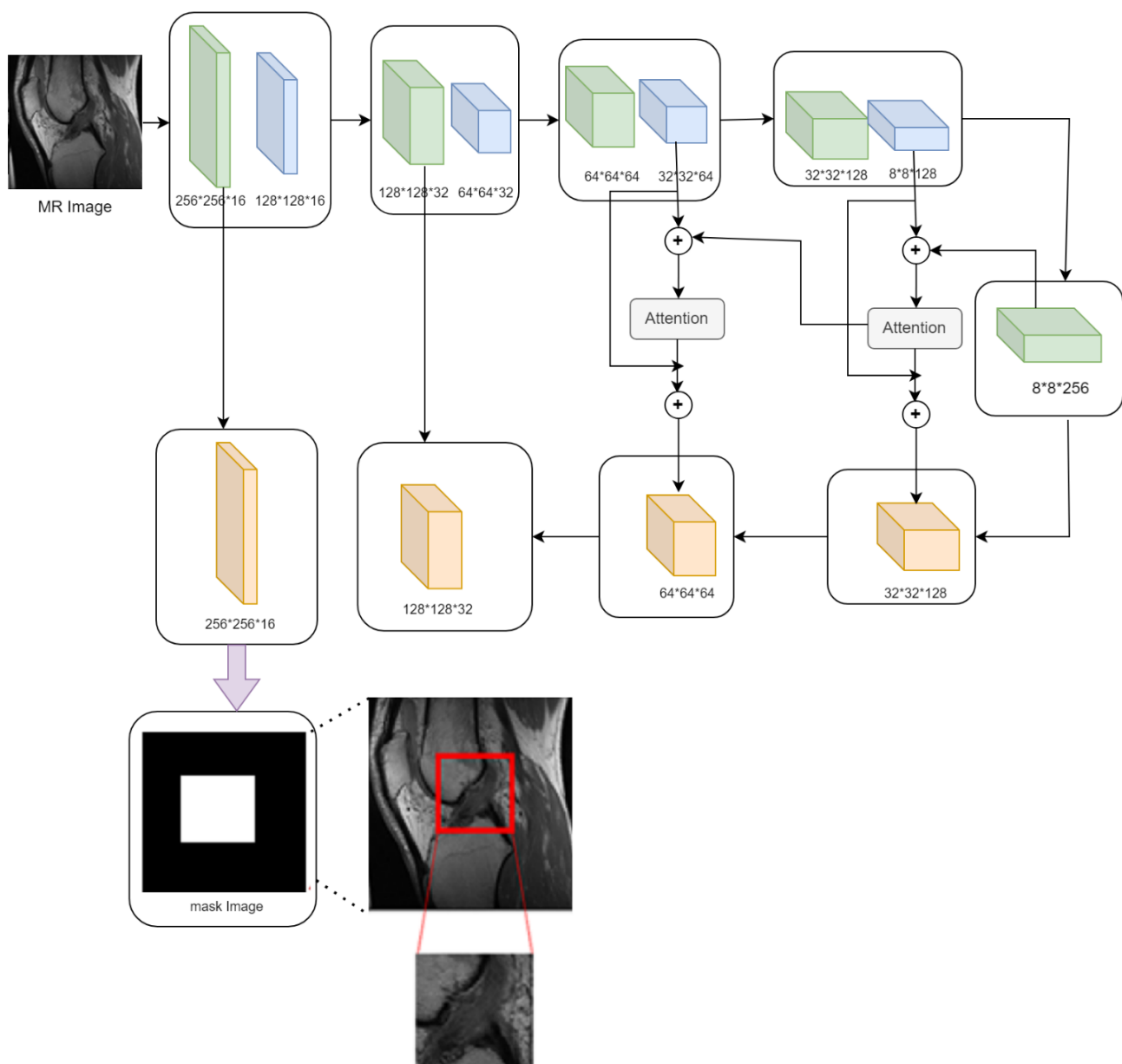


Fig. 2. Anterior cruciate ligament (ACL) localization using an attention-based encoder-decoder architecture.

model to capture both low-level and high-level anatomical structures. Given an input MR image I of size $H \times W$, the first convolutional layer applies a set of k convolutional filters with kernel size (f, f) to extract initial features:

$$F_1 = \sigma(W_1 * I + b_1) \quad (1)$$

Where $*$ represents the convolution operation, W_1 and b_1 are the weights and bias of the first convolutional layer, σ is the activation function (ReLU), F_1 is the resulting feature map. As the image progresses through deeper layers, high-level contextual information is extracted:

$$F_l = \sigma(W_l * F_{l-1} + b_l) \quad (2)$$

Where l represents the layer index in the encoder. To reduce spatial dimensions, down-sampling is performed using max-pooling or stride convolution:

$$F'_l = \text{MaxPool}(F_l, p) \quad (3)$$

Where p is the pooling window size. These operations compress the image representation while preserving essential features, enabling the model to learn patterns critical for ACL localization.

Attention mechanism

One of the key challenges in ACL localization is the small size of the ligament relative to the entire knee structure. To address this, attention mechanisms are integrated into the model to selectively focus on the most relevant regions while suppressing irrelevant background information. These attention layers refine high-level features by weighting the most informative areas, ensuring that the ACL remains the focal point of the model's learning process. Furthermore, the attended features are combined with the encoder outputs through residual connections, preserving spatial information and improving feature propagation. By dynamically adjusting feature weights based on learned importance, the model prioritizes the ACL region, leading to more precise localization. Given an input feature map F , attention weights are computed using a scaled dot-product attention mechanism:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

Where:

- $Q = W_Q F$, $K = W_K F$, $V = W_V F$ are the query, key, and value matrices,
- d_k is the scaling factor (dimensionality of the key),
- The *softmax* function ensures that attention weights sum to 1.

The attended feature map is then computed as:

$$F_{\text{attended}} = \alpha F + (1 - \alpha) \cdot \text{Attention}(Q, K, V) \quad (5)$$

Where α is a learnable parameter that balances original and attended features. To preserve spatial information, attended features are combined with encoder outputs via residual connections:

$$F_{\text{skip}} = F_{\text{encoder}} + F_{\text{attended}} \quad (6)$$

Decoder module

The decoder module reconstructs spatial details of the ACL region from the encoded feature representations. Using up-sampling layers and convolutional operations, the decoder progressively restores the spatial resolution of the input image. Feature refinement is achieved by integrating high-resolution spatial information from the skip connections, ensuring that fine anatomical details are preserved. Skip connections play a vital role in mitigating information loss, as they enable direct information transfer between corresponding encoder and decoder layers. This allows the model to retain structural details necessary for accurate segmentation. The final layer of the decoder produces a binary segmentation mask, highlighting the predicted ACL region. This mask undergoes further refinement to ensure precise localization, which is essential for subsequent injury classification and severity assessment. Given an encoded feature map F_{enc} , the up-sampling operation expands the spatial resolution:

$$F_{\text{upsampled}} = \text{UpSample}(F_{\text{enc}}, s) \quad (7)$$

Where s is the scaling factor. This operation is followed by convolutional refinement:

$$F_{\text{dec}} = \sigma(W_{\text{dec}} * F_{\text{upsampled}} + b_{\text{dec}}) \quad (8)$$

Skip connections play a crucial role in merging high-resolution spatial features from the encoder with up-sampled decoder features:

$$F_{\text{final}} = F_{\text{dec}} + F_{\text{skip}} \quad (9)$$

Finally, the segmentation output is obtained through a sigmoid activation function:

$$M_{\text{ACL}} = \sigma(W_o * F_{\text{final}} + b_o) \quad (10)$$

Where M_{ACL} represents the binary mask highlighting the ACL region.

Classification head

After accurately localizing the anterior cruciate ligament through segmentation, the next critical step is classifying the injury type into three main categories: healthy, partial tear, and complete tear. A healthy ACL appears intact with no signs of damage, whereas a partial tear involves damage to some ligament fibers while maintaining partial

structural integrity. In contrast, a complete tear results in a fully ruptured ligament, causing a discontinuity in its structure. To achieve robust and precise classification, a deep Convolutional Neural Network is employed, integrating dense connectivity, feature down-sampling, and global average pooling. This architecture ensures effective feature extraction from the segmented ACL region, enabling fine-grained differentiation between injury types and facilitating accurate clinical diagnosis. The classification phase is shown in Fig. 3.

Feature extraction using convolutional layers

The classification process begins by extracting features from the localized ACL region obtained in the segmentation phase. This extracted ACL patch is fed into a deep CNN consisting of multiple Conv2D layers, each followed by batch normalization and a ReLU activation function. These layers progressively learn both low-level and high-level features that distinguish between healthy and injured ligaments. Given an input ACL patch I_{ACL} of size $H \times W$, the first convolutional layer applies k filters of size (f, f) to extract basic features:

$$F_1 = \sigma(BN(W_1 * I_{ACL} + b_1)) \quad (11)$$

Where $*$ denotes the convolution operation, W_l and b_l are the weight matrix and bias, BN represents batch normalization, which stabilizes training and speeds up convergence and σ is the ReLU activation function:

$$\sigma(x) = \max(0, x) \quad (12)$$

As the image propagates through deeper convolutional layers, the feature extraction process continues:

$$F_l = \sigma(BN(W_l * F_{l-1} + b_l)) \quad (13)$$

Where F_l represents the feature map extracted at the l -th layer. To reduce computational complexity and improve feature robustness, max-pooling layers are employed at multiple stages in the network. Max-pooling helps retain the most prominent features while reducing spatial dimensions. The pooling operation is defined as:

$$F'_l = \text{MaxPool}(F_l, p) \quad (14)$$

Where p is the pooling window size (2×2). This operation helps, reduce overfitting by forcing the network to focus on important regions and provide translational invariance, making classification robust to small variations in the ACL's position. To improve feature extraction efficiency, dense blocks are integrated into the CNN architecture. In a dense block, each layer receives input from all preceding layers, ensuring maximum feature reuse and gradient flow:

$$F_d^l = \sigma(BN(W_d^l * [F_c^1, F_c^2, \dots, F_c^{l-1}] + b_d^l)) \quad (15)$$

Where, F_d^l is the feature map at layer l within the dense block. The concatenation operation $[\cdot]$ ensures that each layer receives all previous feature maps as input. The integration of dense blocks provides several advantages.

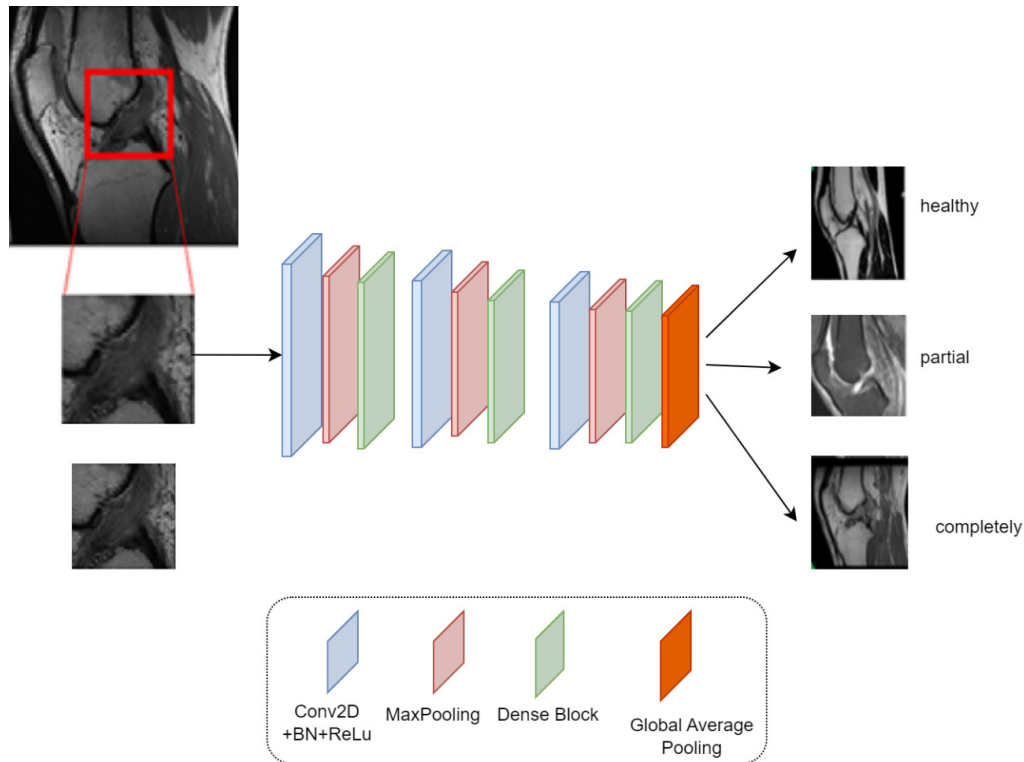


Fig. 3. ACL Injury classification using a CNN with dense blocks.

First, it prevents information loss, ensuring that fine details related to ACL injuries are preserved throughout the network. Second, it enhances gradient flow, leading to improved training stability and faster convergence. Finally, dense connectivity reduces redundant feature extraction, making the network more parameter-efficient while maintaining high performance in distinguishing between healthy, partial tear, and complete tear cases. After extracting dense features, the network applies Global Average Pooling to convert feature maps into a compact representation:

$$F_{GAP} = \frac{1}{N} \sum_{i=1}^N F_d^i \quad (16)$$

Where N is the total number of spatial locations in the feature map. Unlike fully connected layers, GAP minimizes overfitting while preserving spatial relationships. The final feature vector is passed through a *softmax* classifier, which assigns probabilities to each of the three ACL injury categories:

$$P(y_i) = \frac{e^{W_y^i F_{GAP}}}{\sum_j e^{W_y^j F_{GAP}}} \quad (17)$$

Where $P(y_i)$ represents the probability of ACL being in class i (healthy, partial tear, or complete tear). The predicted class is determined as:

$$\hat{y} = \arg_i \max P(y_i) \quad (18)$$

Where $\hat{y}$ is the final classification result.

Loss function and optimization

Training a deep learning model for ACL injury classification requires an effective loss function and optimization strategy to ensure accurate convergence and prevent overfitting.

Categorical cross-entropy loss function

Since the classification task involves three mutually exclusive classes (Healthy, Partial Tear, Complete Tear), the categorical cross-entropy loss function is used to measure the divergence between the predicted probability distribution and the ground-truth labels. This loss function is defined as:

$$L = - \sum_{i=1}^C y_i \log P(y_i) \quad (19)$$

Where:

- C is the number of classes (Healthy, Partial Tear, Complete Tear).
- y_i represents the ground-truth label for class i , which is one-hot encoded (i.e., if the true label is class j , then

$y_i = 1$ and all other $y_i = 0$).

- $P(y_i)$ is the predicted probability that the model assigns to class i , obtained from the *softmax* activation function in the final layer:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^C e^{z_j}} \quad (20)$$

Where z_i is the network's output (logit) for class i , and the denominator ensures that the sum of probabilities across all classes is 1. The cross-entropy loss penalizes incorrect predictions more when the model assigns high confidence to the wrong class, which helps in better training.

Optimization with Adam optimizer

To efficiently minimize the loss function and improve model convergence, the Adam (adaptive moment estimation) optimizer is used. Adam dynamically adjusts the learning rate for each parameter during training by computing adaptive moment estimates. The parameter update rule for Adam is given as:

$$\theta_{t+1} = \theta_t - \eta \frac{m_t}{\sqrt{v_t} + \epsilon} \quad (21)$$

Where θ_t represents the weight parameters at time step t , which are updated in each training iteration. η is the learning rate, controlling the step size of weight updates. m_t is the first moment estimate (moving average of gradients), calculated as:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (22)$$

Where g_t is the gradient of the loss with respect to the weights at time step t , and β_1 is a decay factor. ($\beta_1 = 0.9$). v_t is the second moment estimate (moving average of squared gradients), given by:

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (23)$$

Where β_2 is another decay factor ($\beta_2 = 0.999$). ϵ is a small constant added to avoid division by zero ($\epsilon = 10^{-8}$). Adam combines the benefits of momentum-based updates (using m_t) and RMSProp (adaptive scaling with v_t), leading to faster convergence, better stability, and improved generalization in deep networks. To prevent overfitting and improve generalization, L2 regularization (also known as weight decay) is applied, modifying the loss function as:

$$L_{reg} = L + \lambda \sum_i \|\theta_i\|^2 \quad (24)$$

Where λ is a regularization coefficient that controls

the penalty on large weights. This encourages the model to keep weight values small, improving generalization on unseen MRI scans. Additionally, dropout layers are introduced during training, randomly setting some neuron activations to zero, which forces the network to learn more robust and distributed representations.

Results

Dataset

The dataset used in this study consists of 917 sagittal plane DICOM MRI scans of the knee, obtained from the Clinical Hospital Center of Rijeka archiving and communication system. The images are 12-bit grayscale and are accompanied by their corresponding ACL diagnosis labels. All scans were acquired using a Siemens Avanto 1.5T MRI scanner (manufactured in Muenchen, Germany) between 2007 and 2010. The imaging protocol included proton density-weighted fat suppression, ensuring high contrast for ligament visualization. The dataset is categorized into three classes: healthy (0 label), partial tear (1 label), and complete tear (2 label). The dataset comprises 690 healthy samples, 172 partial tear cases and 55 complete ruptures, stored in pickle format for efficient processing. Additionally, metadata, including diagnostic details, is provided in CSV format for reference.³⁹ Given the class imbalance in the dataset, where the number of healthy samples significantly exceeds the partial tear and complete tear cases, data augmentation techniques are employed to balance the dataset. Augmentation strategies such as rotation, flipping, contrast adjustment, and elastic deformations are applied to increase the number of samples in the underrepresented classes. After augmentation, the total number of training samples per class was balanced to approximately 720 images for each category (Healthy, Partial Tear, Complete Tear). This ensures a more balanced distribution, preventing bias toward the majority class and enhancing the model's ability to generalize across different ACL injury severities. Fig. 4 illustrates an example of the augmented data.

Evaluation metrics

To assess the performance of the proposed ACL localization and classification model, standard evaluation metrics are employed. Given the class distribution in the dataset, the following metrics are used to ensure a comprehensive evaluation:

Accuracy: Measures the overall percentage of correctly classified samples. It is defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (25)$$

Where TP and TN represent true positives and true negatives, while FP and FN denote false positives and false negatives, respectively.

Precision: Evaluates the proportion of correctly predicted positive instances among all positive predictions. It is calculated as:

$$Precision = \frac{TP}{TP + FP} \quad (26)$$

Precision is crucial in medical diagnosis, as a high precision value ensures fewer false positives, reducing unnecessary medical interventions.

Sensitivity: Measures the ability of the model to identify actual positive cases (injured ACLs). It is given by:

$$Sensitivity = \frac{TP}{TP + FN} \quad (27)$$

High recall ensures that the model does not miss a significant number of ACL injuries.

Specificity: Measures the proportion of actual negative cases (healthy ACLs) correctly identified by the model:

$$Specificity = \frac{TN}{TN + FP} \quad (28)$$

This metric is crucial in medical imaging, ensuring that healthy individuals are not misclassified as injured.

F1-Score: Provides a balanced measure of precision and recall:

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (29)$$

This metric is especially useful in datasets with class imbalance, such as the current dataset where healthy cases significantly outnumber injured ones.

AUC-ROC: Measures the model's ability to distinguish between classes. A higher AUC value indicates better discrimination between healthy, partial tear, and complete tear cases.

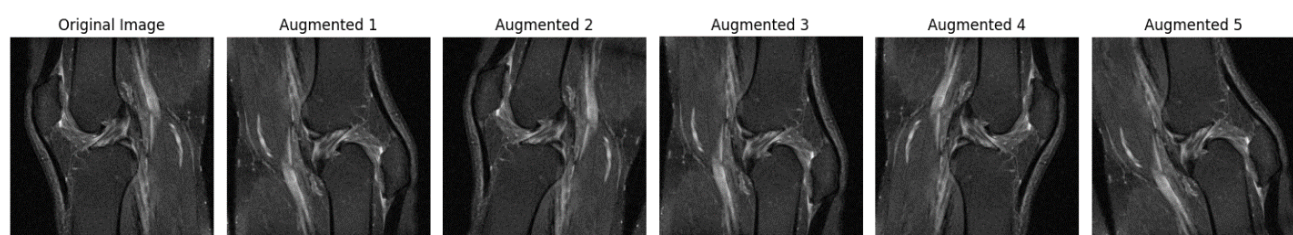


Fig. 4. Example of augmented MRI scans for class balancing.

Experimental results

In this section, the performance of the proposed method is evaluated based on standard evaluation metrics for machine learning models. The metrics include overall accuracy, precision, recall, specificity, and F1-score for each of the classes: Healthy, Partial, and Complete. This analysis helps assess the model's accuracy in classifying different classes and identifies its strengths and weaknesses. Fig. 5a shows the Confusion Matrix, which illustrates the distribution of the model's predictions compared to the true labels. For evaluation purposes, a subset of 300 samples (comprising 140 healthy, 80 partial tears, and 80 complete tears) was randomly selected from the dataset. This number was chosen to maintain

a relatively balanced class distribution in the evaluation phase and to reduce bias, allowing for a more accurate assessment of the model's performance across all classes. This matrix allows us to analyze the accuracy of the model's predictions and the types and frequencies of its errors. As can be seen, the proposed model has correctly classified the majority of samples, with minimal misclassification in each class. A detailed examination of this matrix helps identify in which classes the model makes the most errors and whether specific types of data pose challenges for the model. Fig. 5b presents the bar plot of evaluation metrics, which enables direct comparison of evaluation metrics across the different classes. It is evident that the precision and recall values are high for all three classes, indicating

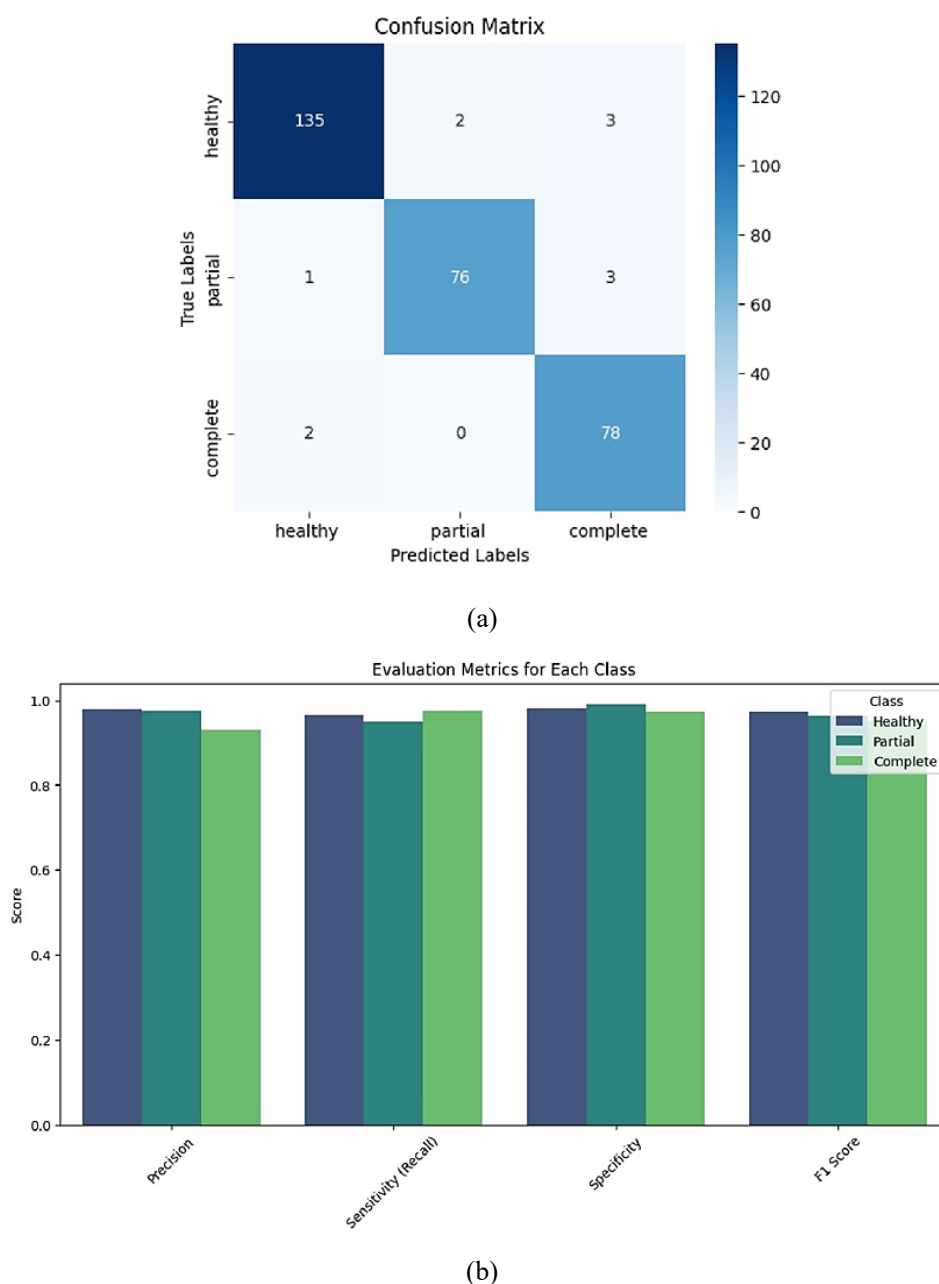


Fig. 5. (a) Confusion matrix for evaluating model performance in classifying the three classes: healthy, partial and complete (b) bar chart of evaluation metrics for each class.

the model's ability to reduce type I and type II errors. The high F1-score also suggests a balanced performance between precision and recall, confirming the model's consistent ability.

Given that the proposed method involves a considerable number of hyperparameters that significantly affect the model's performance, it is essential to clearly specify their values. Table 1 presents the hyperparameter settings used in our approach. These values were determined based on extensive experiments and empirical tuning to achieve optimal performance. All reported results and analyses in this study are based on these hyperparameter configurations.

To evaluate the effectiveness of the proposed hierarchical architecture, its performance was compared to a flat CNN model without the segmentation stage. As shown in Table 2, the proposed method outperformed the flat CNN in all evaluation metrics, including accuracy, sensitivity, precision, specificity, and the AUC.

Given the imbalanced nature of the dataset in this study, the use of two metrics, receiver operating characteristic (ROC) curve and precision-recall (PR) curve, was crucial for a more accurate evaluation of the model's performance. In classification problems with imbalanced data, especially when the number of samples for one or more classes is significantly smaller than the others, traditional metrics like accuracy may lead to misleading performance evaluations. For instance, a high accuracy rate may result from simply predicting the majority class, without adequately addressing the minority class. Therefore, metrics such as AUC for the ROC curve and average precision (AP) for the PR curve were employed to assess the model's ability to distinguish between classes,

particularly for the minority classes. In this study, the ROC curves were computed for each of the classes (Healthy, Partial, Complete), and the Area Under the Curve was evaluated separately for each. The results obtained from the AUC show the model's ability to distinguish between the different classes. Specifically, the AUC for the Healthy class was 0.98, for Partial it was 0.94, and for Complete it was 0.93. These results indicate that the model performed exceptionally well in distinguishing the Healthy class with the highest AUC, while maintaining acceptable performance for the Partial and Complete classes. Notably, while the AUC for the minority classes (Partial and Complete) is lower, it still reflects good classification performance, considering the inherent class imbalance. On the other hand, the PR curves were also plotted for each class, and Average Precision was computed for each. The obtained values of AP were 0.96 for the Healthy class, 0.89 for the Partial class, and 0.85 for the Complete class. These values demonstrate that the model, particularly in identifying positive samples of the Healthy class, exhibited strong performance. However, the performance for the Partial and Complete classes, which have fewer samples, showed a slight decrease. This is especially important in imbalanced data settings, as the AP metric can reveal the model's shortcomings in identifying the minority classes. Specifically, while the model achieves high precision in detecting healthy ACLs, its precision slightly decreases for the smaller classes, especially in recognizing partial and complete ACL tears. Ultimately, the results from these two analyses emphasize the importance of using ROC and PR metrics for evaluating classification models with imbalanced datasets. These metrics, in particular, help assess the model's ability to accurately identify minority classes. The results of these analyses are shown in Fig. 6a and 6b, where the ROC and Precision-Recall curves for each class are fully displayed. These plots provide additional insights into the model's performance, offering a clearer understanding of its ability to recognize minority classes in imbalanced data scenarios.

Training prosses

To evaluate the training process and the accuracy of the proposed model, two key graphs were analyzed: the loss curve and the accuracy curve. In Fig. 7a, the cross-entropy loss reduction over 50 training epochs is depicted. This graph includes two curves representing the training loss and test loss. As shown, the loss value steadily decreases in both datasets, indicating a gradual and optimal learning process of the proposed model. Additionally, towards the final epochs, the gap between the two curves narrows,

Table 1. Hyperparameter settings for the proposed ACL classification model

Hyperparameter	Value
Image size	256 × 256
Loss	Categorical cross-entropy
Number of dense blocks	3
Dropout rate	0.3
Optimizer	Adam
Learning rate	0.001
ϵ (Adam parameter)	1e-8
β_2 (Adam parameter)	0.999
β_1 (Adam parameter)	0.9
L2 Regularization (λ)	0.0001
Number of epochs	50
Batch size	32

Table 2. Comparison between flat CNN and the proposed hierarchical method based on evaluation metrics

Method	Accuracy (%)	Sensitivity (%)	Precision (%)	Specificity (%)	AUC
Flat CNN (without segmentation)	88.1	85.2	86.5	89.3	0.89
Proposed Hierarchical Method	96.3	96.3	96.0	97.1	0.97

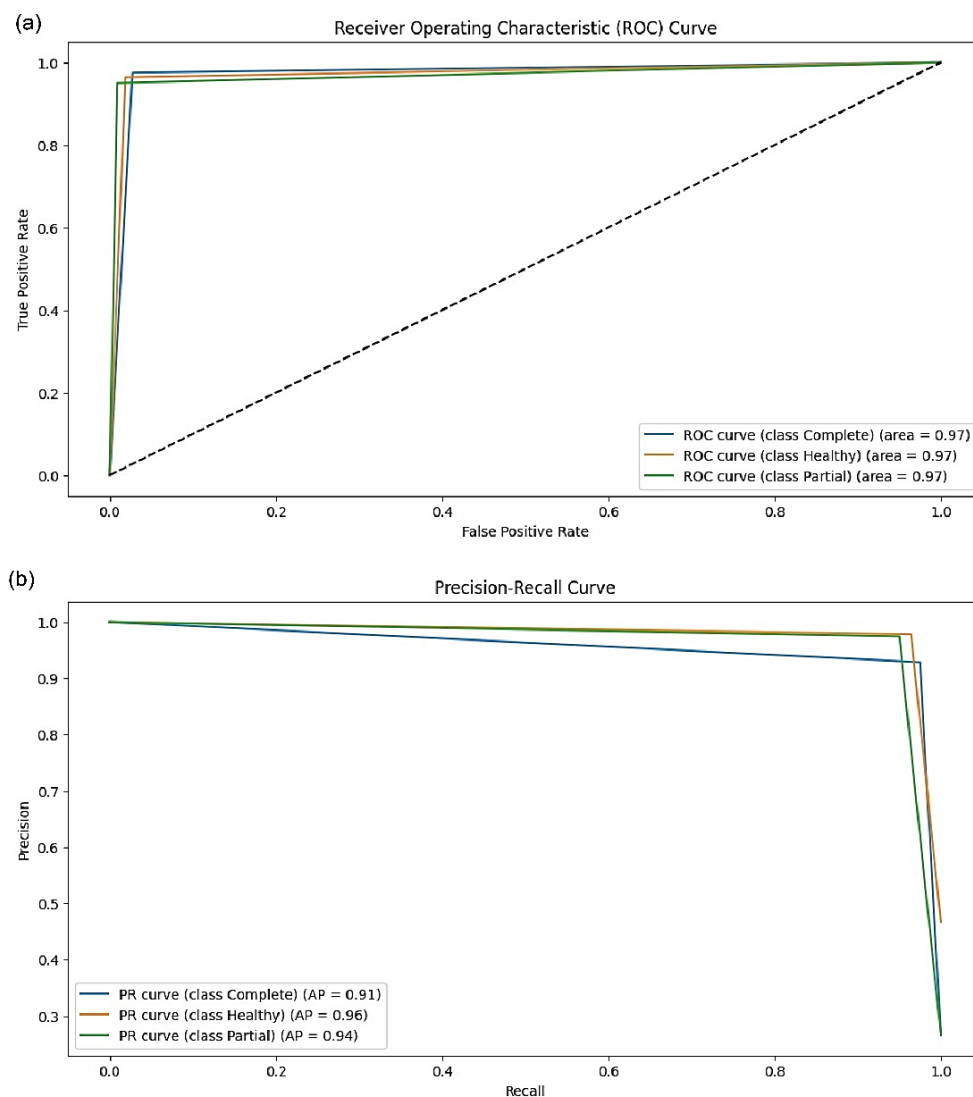


Fig. 6. (a) ROC curves for each class (Healthy, Partial, Complete) with AUC values. (b) Precision-recall curves for each class (Healthy, Partial, Complete) with average precision (AP) values.

suggesting a reduced risk of overfitting and demonstrating that the model generalizes well to unseen data. In Fig. 7b, the accuracy curve illustrates the performance of the model throughout the training process. The graph shows that the accuracy steadily increases for both the training and test datasets, ultimately reaching over 85%. This upward trend confirms that the model has successfully learned to differentiate between various ACL injury categories. Moreover, the convergence of the two curves in the later training stages indicates that the model does not suffer from significant mismatches between training and test performance, further reinforcing its reliability on real-world data. Overall, the continuous loss reduction and increasing accuracy confirm the model's effectiveness in learning complex ACL injury features. The narrowing gap between training and test curves indicates that the model generalizes well without overfitting. Additionally, achieving over 95.7 accuracy on the test set highlights the model's high capability in automatic ACL injury

classification from MR images. These results suggest that the proposed approach can serve as a precise and efficient diagnostic tool to assist clinicians in evaluating ACL injuries.

In this study, statistical significance of the results from the proposed method was evaluated using 5-fold cross-validation. This approach was applied to examine the stability and accuracy of the model across different datasets and to assess its generalization ability to new, unseen data. Using this technique, the model was trained and evaluated five times on different subsets of the data. The results from this experiment are shown in Fig. 8. The plot presents the average values of various evaluation metrics, including precision, sensitivity, specificity, F1-score, and accuracy, for the three classes: Healthy, Partial, and Complete, along with the variance of each of these metrics. Specifically, for the Healthy class, the precision is 0.978, sensitivity is 0.964, specificity is 0.981, F1-score is 0.971, and accuracy is 0.977. For the Partial class, precision

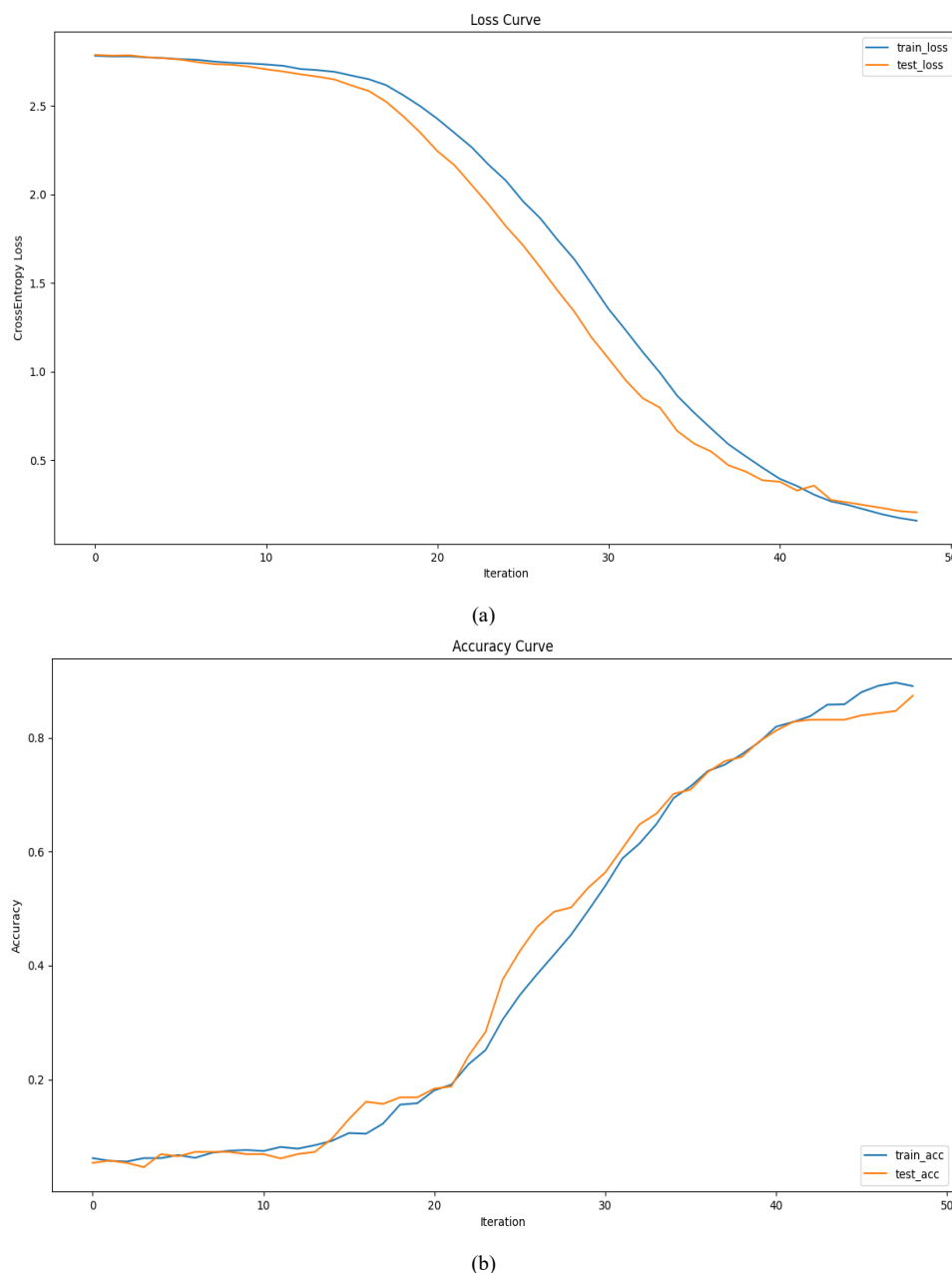


Fig. 7. (a) Error Plot of the Proposed Method for Training and Evaluation Data. (b) Model Accuracy Progression During Training

is 0.974, sensitivity is 0.950, specificity is 0.990, F1-score is 0.962, and accuracy is 0.973. For the Complete class, precision is 0.928, sensitivity is 0.975, specificity is 0.972, F1-score is 0.951, and accuracy is 0.939. The variance of these metrics for all three classes ranges between 0.001 and 0.009, indicating stable results. The use of 5-fold cross-validation means that the model was trained and evaluated five times on different portions of the data, which allows for a more comprehensive and accurate assessment. The low variance in the metric values suggests that the model demonstrates consistent performance and that the results are not due to random fluctuations. In other words, the model exhibits high stability, and these results are both reliable and statistically significant. This analysis indicates that the proposed model demonstrates meaningful and

stable performance across all metrics, reinforcing the model's credibility and its ability to generalize.

Comparison with other methods

The proposed method in this study was implemented and evaluated on a specific dataset, and its results were compared with existing methods. This comparison was based on standard evaluation metrics, including accuracy, recall, precision, and F1 score. The results of this experiment are presented in Table 3. According to the obtained results, the proposed method exhibited strong performance across all evaluation metrics. The overall accuracy of the proposed model was 96.33%, which demonstrates its high capability in correctly classifying samples across all classes. This accuracy outperforms

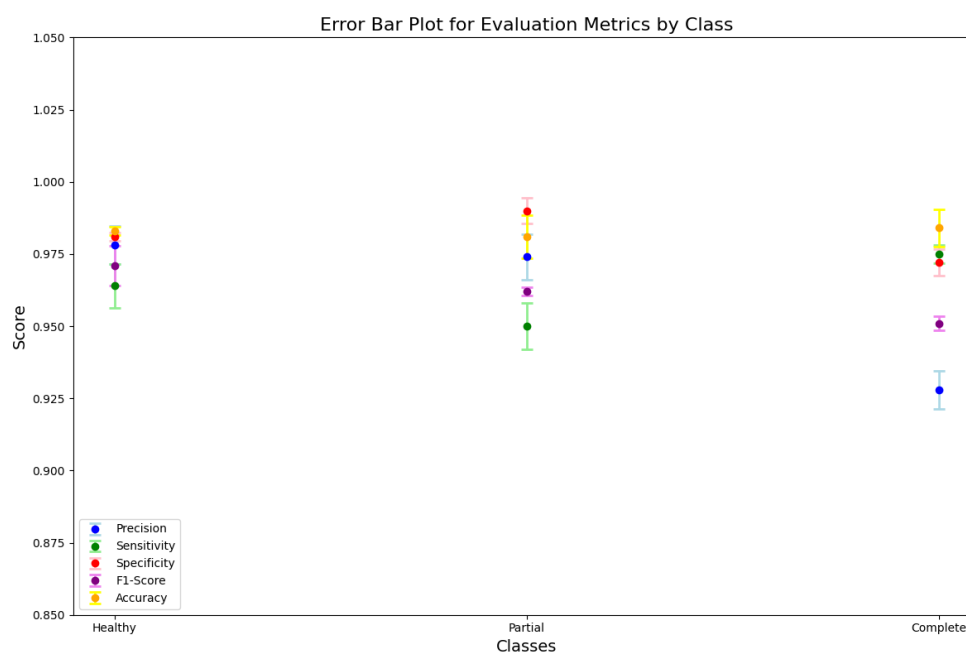


Fig. 8. Cross-validation results with $k=5$, Error bar plot of evaluation metrics including Precision, Sensitivity, Specificity, F1-Score and Accuracy for three different classes along with the variance of each metric.

Table 3. Comparison of state-of-the-art works with our proposed model

Author, Year	Model	Dataset	Output	Evaluation				
				Accuracy	Sensitivity	Precision	Specificity	AUC
Bien et al, ⁴⁰ 2018	AlexNet	MRNet 1370 exams	ACL tear	0.867	0.759	-	0.968	0.965
			abnormal	0.850	0.879	-	0.714	0.937
			meniscus tear	0.725	0.892	-	0.741	0.847
Chang et al, ⁴¹ 2019	Dynamic patch + ResNet	260 MRI coronal volumes	partial AC, full torn	0.967 -	1.00	0.938	0.933	-
			full thickness	-	0.92	-	0.92	0.95
			ACL tear, Intact ACL	-	0.96	-	0.96	0.98
Liu et al, ¹ 2019	DensNet	sagittal MR 175 (exams)	ACL tear, Intact ACL	-	0.89	-	0.88	0.90
			Intact ACL	-	0.22	-	0.90	-
			Partial tear	-	0.75	-	1.00	-
Namiri et al, ⁷ 2019	2D CNN	NIH MRI 1243 (exams)	Full tear	-	0.82	-	0.94	-
			Intact ACL	-	0.22	-	0.90	-
			Partial tear	-	0.75	-	1.00	-
Zhang et al, ⁴² 2020	3D DenseNet	sagittal MR 408 (exams)	ACL tears	0.957	0.976	0.940	0.944	0.960
			Intact ACL	0.943	0.952	0.952	0.909	0.946
			Intact ACL	0.899	0.912	0.869	0.886	0.859
Tsai et al, ⁴⁴ 2020	EfficientNet	MRNet 1370	abnormal	0.917	0.968	-	0.72	0.941
			ACL tear	0.904	0.923	-	0.891	0.960
Irmakci et al, ⁴³ 2020	GoogleLeNet	MRNet 1370 exams	abnormal	0.833	0.978	-	0.280	0.909
			ACL tear	0.808	0.666	-	0.924	0.890
Mazhar et al, ² 2021	Customized ResNet-14	KneeMRI 917 exams	ACL Intact	0.92	0.89	0.89 0.92	0.93	0.98
			partial tear	0.91	0.87	0.87	0.92	0.97
			ruptured	0.93	0.99	0.96	0.99	0.99
Li et al, ⁴⁵ 2021	Multi-modal feature fusion Deep CNN	MRI group + Arthroscopy group ACL:60 cases	Grade 0 Grade 1 Grade II Grade III	92.1%	96.7%	-	90.6 %	0.963
Proposed method	Encoder-Decoder-Attation	KneeMRI 917 exams	Healthy	0.977	0.964	0.978	0.981	0.97
			partial	0.973	0.950	0.974	0.990	0.97
			Complete	0.939	0.975	0.928	0.972	0.97

several other methods, such as AlexNet (86.7%) and ResNet (90.4%). In terms of recall, the proposed model showed high sensitivity, especially in the "Complete" class, with a value of 97.5%, which is higher than the recall of other models, such as GoogleLeNet (66.6%) and AlexNet (88.9%). This high performance in identifying instances from various classes indicates that the proposed model is better at classifying samples with greater precision. Additionally, the F1 score of the proposed model also demonstrated strong performance, especially in identifying the "Healthy" and "Partial" classes, with F1 scores of 97.12% and 96.20%, respectively. These values are notably higher than those of other models, such as DenseNet (95%) and ResNet (94.3%). Overall, the results indicate that the proposed method outperforms many existing models across all evaluation metrics and demonstrates a high ability to accurately classify and identify different classes.

Our approach leverages an encoder-decoder architecture enhanced with attention mechanisms, which significantly improves the localization and segmentation of the ACL. The attention mechanism allows the model to prioritize important regions, ensuring precise feature extraction, while skip connections help prevent the loss of fine details during down-sampling. These features enable the model to effectively distinguish between different ACL injuries, such as partial and complete tears, with high accuracy. Additionally, the use of a deep convolutional network with dense blocks facilitates robust feature extraction, further improving the model's ability to classify ACL conditions accurately.

However, there are some limitations. One of the main challenges is the computational complexity of the model, which requires significant hardware resources. Moreover, the model's sensitivity to the quality and variability of training data may affect its performance, particularly in clinical settings where MRI images vary in quality or are obtained from different devices. These limitations could impact the model's scalability and training time. Despite these challenges, our method provides substantial improvements in ACL injury detection and classification, offering a valuable tool for clinical decision-making.

Conclusion

In this study, a deep learning-based hierarchical classification model was proposed for the detection and classification of anterior cruciate ligament injuries in MR images. By leveraging advanced techniques such as attention mechanisms and hierarchical architecture in convolutional neural networks, the proposed model demonstrated outstanding performance in both ACL segmentation and injury classification stages. The experimental results showed that the proposed model achieved high accuracy compared to other methods, particularly in identifying different types of ACL

Research Highlights

What is the current knowledge?

- Traditional methods like MRI scans and clinical assessments are commonly used for ACL injury diagnosis.
- Deep learning models have been explored for ACL injury detection, focusing on MRI segmentation and classification.
- CNNs have been utilized for ACL segmentation and severity staging with varying success.
- Attention mechanisms and transfer learning have improved segmentation and classification accuracy in ACL injury diagnosis.
- Multimodal approaches combining MRI and clinical data have enhanced diagnostic performance.

What is new here?

- Proposes a hierarchical deep learning-based model for precise ACL injury detection and classification.
- Integrates an attention mechanism into an encoder-decoder architecture for accurate ACL localization.
- Uses dense blocks and global average pooling in the classification phase for improved feature extraction.
- Focuses on overcoming challenges like small ligament size and image quality variations in ACL imaging.
- Introduces a fully automated, hierarchical classification system for better clinical utility and athlete healthcare.

injuries, and it excelled in precision, recall, F1 score, and sensitivity metrics across various classes. These results clearly indicate that the proposed method can serve as an efficient and effective tool for physicians and sports medicine specialists in the rapid and accurate diagnosis of ACL injuries. Furthermore, the model's ability to classify various injury types, specially complete tears, significantly enhances diagnostic workflows, potentially improving clinical outcomes.

Authors' Contribution

Conceptualization: Lei Xiao.

Data curation: Xuejiao Yan.

Formal analysis: Xuejiao Yan, Lei Xiao.

Investigation: Xuejiao Yan, Lei Xiao.

Writing-original draft: Xuejiao Yan.

Writing-review & editing: Lei Xiao.

Competing Interests

All authors declare that they have no conflict of interest.

Ethical Approval

Not applicable.

Funding

Not applicable.

References

1. Liu F, Guan B, Zhou Z, Samsonov A, Rosas H, Lian K, et al. Fully automated diagnosis of anterior cruciate ligament tears on knee MR images by using deep learning. *Radiol Artif Intell* 2019; 1: 180091. doi: 10.1148/ryai.2019180091.

2. Awan MJ, Mohd Rahim MS, Salim N, Mohammed MA, Garcia-Zapirain B, Abdulkareem KH. Efficient Detection of Knee Anterior Cruciate Ligament from Magnetic Resonance Imaging Using Deep Learning Approach. *Diagnostics (Basel)* **2021**; 11: 105. doi: 10.3390/diagnostics11010105.
3. Awan MJ, Mohd Rahim MS, Salim N, Rehman A, Nobanee H, Shabir H. Improved deep convolutional neural network to classify osteoarthritis from anterior cruciate ligament tear using magnetic resonance imaging. *J Pers Med* **2021**; 11: 1163. doi: 10.3390/jpm11111163.
4. Dung NT, Thuan NH, Van Dung T, Van Nho L, Tri NM, Vy VP, et al. End-to-end deep learning model for segmentation and severity staging of anterior cruciate ligament injuries from MRI. *Diagn Interv Imaging* **2023**; 104: 133–41. doi: 10.1016/j.diii.2022.10.010.
5. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing; **2015**. p. 234–41. doi: 10.1007/978-3-319-24574-4_28.
6. Awan MJ, Mohd Rahim MS, Salim N, Rehman A, Garcia-Zapirain B. Automated knee MR images segmentation of anterior cruciate ligament tears. *Sensors (Basel)* **2022**; 22: 1552. doi: 10.3390/s22041552.
7. Namiri NK, Flament I, Astuto B, Shah R, Tibrewala R, Caliva F, et al. Deep learning for hierarchical severity staging of anterior cruciate ligament injuries from MRI. *Radiol Artif Intell* **2020**; 2: e190207. doi: 10.1148/ryai.2020190207.
8. Chan S, Zhang M, Zhi YY, Razmjoo S, El-Sherbeeney AM, Lin L. Improved anterior cruciate ligament tear diagnosis using gated recurrent unit networks and Hybrid Tasmanian Devil Optimization. *Biomed Signal Process Control* **2024**; 95: 106309. doi: 10.1016/j.bspc.2024.106309.
9. Wang D, Guo L, Zhong J, Yu H, Tang Y, Peng L, et al. A novel deep-learning based weighted feature fusion architecture for precise classification of pressure injury. *Front Physiol* **2024**; 15: 1304829. doi: 10.3389/fphys.2024.1304829.
10. Cai W, Zhu J, Zhuang X. Combining Medical Images and Biomechanical Data in Sports Injury Prediction Models. *Res Sq* [Preprint]. September 10, **2024**. Available from: <https://www.researchsquare.com/article/rs-4884511/v1>.
11. Haddadian J, Balamurali M. Transfer learning and data augmentation in the diagnosis of knee MRI. In: Long G, Yu X, Wang S, eds. *AI 2021: Advances in Artificial Intelligence*. Cham: Springer International Publishing; **2022**. p. 452–63. doi: 10.1007/978-3-030-97546-3_37.
12. Manna S, Bhattacharya S, Pal U. Self-Supervised Representation Learning for Detection of ACL Tear Injury in Knee MR Videos. *ArXiv* [Preprint]. July 15, **2020**. Available from: <https://arxiv.org/abs/2007.07761>.
13. Alam MJ, Talukdar MS, Hasan MN, Jennifer SS, Morshed MA, Reza AW, et al. Evaluation of customized ACL-Net model: a comparative analysis of CNN and image enhancement techniques for classifying ACL injuries. In: *2024 IEEE International Conference on Computing, Applications and Systems (COMPAS)*. Cox's Bazar: IEEE; **2024**. p. 1–5. doi: 10.1109/compas60761.2024.10796820.
14. Guan J, Li Z, Sheng S, Lin Q, Wang S, Wang D, et al. An artificial intelligence-driven revolution in orthopedic surgery and sports medicine. *Int J Surg* **2025**; 111: 2162–81. doi: 10.1097/jvs.0000000000002187.
15. Zhang T, Ye Z, Cai J, Chen J, Zheng T, Xu J, et al. Ensemble algorithm for risk prediction of clinical failure after anterior cruciate ligament reconstruction. *Orthop J Sports Med* **2024**; 12: 23259671241261695. doi: 10.1177/23259671241261695.
16. Shan Lee VL, Gan KH, Tan TP, Abdullah R. Semi-supervised learning for sentiment classification using small number of labeled data. *Procedia Comput Sci* **2019**; 161: 577–84. doi: 10.1016/j.procs.2019.11.159.
17. Wang J, Luo J, Hounye AH, Wang Z, Liang J, Cao Y, et al. A decoupled generative adversarial network for anterior cruciate ligament tear localization and quantification. *Neural Comput Appl* **2023**; 35: 19351–64. doi: 10.1007/s00521-023-08776-7.
18. Calderón-Díaz M, Silvestre Aguirre R, Vázquez JP, Yáñez R, Roby M, Querales M, et al. Explainable machine learning techniques to predict muscle injuries in professional soccer players through biomechanical analysis. *Sensors (Basel)* **2023**; 24: 119. doi: 10.3390/s24010119.
19. Sehanobish A, Kannan K, Abraham N, Das A, Odry B. Meta-Learning Pathologies from Radiology Reports Using Variance Aware Prototypical Networks. *ArXiv* [Preprint]. October 22, **2022**. Available from: <https://arxiv.org/abs/2210.13979>.
20. Xue Y, Yang S, Sun W, Tan H, Lin K, Peng L, et al. Approaching expert-level accuracy for differentiating ACL tear types on MRI with deep learning. *Sci Rep* **2024**; 14: 938. doi: 10.1038/s41598-024-51666-8.
21. Li F, Zhai P, Yang C, Feng G, Yang J, Yuan Y. Automated diagnosis of anterior cruciate ligament via a weighted multi-view network. *Front Bioeng Biotechnol* **2023**; 11: 1268543. doi: 10.3389/fbioe.2023.1268543.
22. Jeon Y, Yoshino K, Hagiwara S, Watanabe A, Quek ST, Yoshioka H, et al. Interpretable and lightweight 3-D deep learning model for automated ACL diagnosis. *IEEE J Biomed Health Inform* **2021**; 25: 2388–97. doi: 10.1109/jbhi.2021.3081355.
23. Liu W, Lin W, Zhuang Z, Miao K. Domain-adaptive framework for ACL injury diagnosis utilizing contrastive learning techniques. *Electronics* **2024**; 13: 3211. doi: 10.3390/electronics13163211.
24. Pedoia V, Norman B, Mehany SN, Bucknor MD, Link TM, Majumdar S. 3D convolutional neural networks for detection and severity staging of meniscus and PFJ cartilage morphological degenerative changes in osteoarthritis and anterior cruciate ligament subjects. *J Magn Reson Imaging* **2019**; 49: 400–10. doi: 10.1002/jmri.26246.
25. Park H, Lee J. HQK-FL: hybrid-quantum-key-based secure federated learning for distributed multi-center clinical studies. *Hum Centric Comput Inf Sci* **2023**; 13: 1–17. doi: 10.22967/hcis.2023.13.045.
26. Awan MJ, Mohd Rahim MS, Salim N, Nobanee H, Asif AA, Attiq MO. MGACA-Net: a novel deep learning based multi-scale guided attention and context aggregation for localization of knee anterior cruciate ligament tears region in MRI images. *PeerJ Comput Sci* **2023**; 9: e1483. doi: 10.7717/peerj-cs.1483.
27. Ying M, Wang Y, Yang K, Wang H, Liu X. A deep learning knowledge distillation framework using knee MRI and arthroscopy data for meniscus tear detection. *Front Bioeng Biotechnol* **2023**; 11: 1326706. doi: 10.3389/fbioe.2023.1326706.
28. Wang X, Wu Y, Li J, Li Y, Xu S. Deep learning-assisted automatic diagnosis of anterior cruciate ligament tear in knee magnetic resonance images. *Tomography* **2024**; 10: 1263–76. doi: 10.3390/tomography10080094.
29. Wang J, Luo J, Liang J, Cao Y, Feng J, Tan L, et al. Lightweight attentive graph neural network with conditional random field for diagnosis of anterior cruciate ligament tear. *J Imaging Inform Med* **2024**; 37: 688–705. doi: 10.1007/s10278-023-00944-4.
30. Ajuied A, Wong F, Smith C, Norris M, Fitzgerald R, Harris R. Review on the role of deep learning models in medical image classification for diagnosing anterior cruciate ligament injuries. *Curr Med Imaging* **2025**; 1: 172–88.
31. Karam C, El Zini J, Awad M, Saade C, Naffaa L, El Amine M. A progressive and cross-domain deep transfer learning framework for wrist fracture detection. *J Artif Intell Soft Comput Res* **2021**; 12: 101–20. doi: 10.2478/jaiscr-2022-0007.
32. Zhou X, Yu Y, Feng Y, Ding G, Liu P, Liu L, et al. Attention mechanism based multi-sequence MRI fusion improves prediction of response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Radiat Oncol* **2023**; 18: 175. doi: 10.1186/s13014-023-02352-y.
33. Kim DH, Chai JW, Kang JH, Lee JH, Kim HJ, Seo J, et al. Ensemble deep learning model for predicting anterior cruciate ligament tear

- from lateral knee radiograph. *Skeletal Radiol* **2022**; 51: 2269-79. doi: 10.1007/s00256-022-04081-x.
34. Wu Z, Huang Z, Tang N, Wang K, Bian C, Li D, et al. Research on sports injury rehabilitation detection based on IoT models for digital health care. *Big Data* **2025**; 13: 144-60. doi: 10.1089/big.2023.0134.
 35. Dhakal P, Joshi SR. Uncertainty estimation in detecting knee abnormalities on MRI using Bayesian deep learning. *Proc IOE Grad Conf* **2021**; 10: 1375-82.
 36. Mead K, Cross T, Roger G. MRI deep learning models for assisted diagnosis of knee pathologies: a systematic review. *Eur Radiol* **2025**; 35: 2457-2469. doi: 10.1007/s00330-024-11105-8
 37. Fritz B, Fritz J. Artificial intelligence for MRI diagnosis of joints: a scoping review of the current state-of-the-art of deep learning-based approaches. *Skeletal Radiol* **2022**; 51: 315-29. doi: 10.1007/s00256-021-03830-8.
 38. Voinea ȘV, Gheonea IA, Teică RV, Florescu LM, Roman M, Selișteanu D. Refined detection and classification of knee ligament injury based on ResNet convolutional neural networks. *Life (Basel)* **2024**; 14: 478. doi: 10.3390/life14040478.
 39. Štajduhar I, Mamula M, Miletic D, Ünal G. Semi-automated detection of anterior cruciate ligament injury from MRI. *Comput Methods Programs Biomed* **2017**; 140: 151-64. doi: 10.1016/j.cmpb.2016.12.006.
 40. Bien N, Rajpurkar P, Ball RL, Irvin J, Park A, Jones E, et al. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. *PLoS Med* **2018**; 15: e1002699. doi: 10.1371/journal.pmed.1002699.
 41. Chang PD, Wong TT, Rasiej MJ. Deep learning for detection of complete anterior cruciate ligament tear. *J Digit Imaging* **2019**; 32: 980-6. doi: 10.1007/s10278-019-00193-4.
 42. Zhang L, Li M, Zhou Y, Lu G, Zhou Q. Deep learning approach for anterior cruciate ligament lesion detection: evaluation of diagnostic performance using arthroscopy as the reference standard. *J Magn Reson Imaging* **2020**; 52: 1745-52. doi: 10.1002/jmri.27266.
 43. Irmakci I, Anwar SM, Torigian DA, Bagci U. Deep learning for musculoskeletal image analysis. In: *2019 53rd Asilomar Conference on Signals, Systems, and Computers*. Pacific Grove, CA: IEEE; **2019**. p. 1481-5. doi: 10.1109/ieeeeconf44664.2019.9048671.
 44. Tsai CH, Kiryati N, Konen E, Eshed I, Mayer A. Knee Injury Detection using MRI with Efficiently-Layered Network (ELNet). *ArXiv [Preprint]*. May 6, **2020**. Available from: <https://arxiv.org/abs/2005.02706>.
 45. Li Z, Ren S, Zhou R, Jiang X, You T, Li C, et al. Deep learning-based magnetic resonance imaging image features for diagnosis of anterior cruciate ligament injury. *J Healthc Eng* **2021**; 2021: 4076175. doi: 10.1155/2021/4076175.