KSÜ Mühendislik Bilimleri Dergisi, 28(1), 2025
Araştırma Makalesi

KSU J Eng Sci, 28(1), 2025
Research Article

Kahramanmaras Sutcu Imam University
Journal of Engineering Sciences

Geliş Tarihi : 02.10.2024
Kabul Tarihi :08.11.2024

Received Date : 02.10.2024
Accepted Date : 08.11.2024

AUTOMATIC DETECTION OF MENISCUS TEARS FROM KNEE MRI IMAGES USING DEEP LEARNING: YOLO V8, V9, AND V10 SERIES

DİZ MRI GÖRÜNTÜLERİNDEN MENİSKÜS YIRTIKLARININ DERİN ÖĞRENME İLE OTOMATİK TESPİTİ: YOLO V8, V9 VE V10 SERİLERİ

Mehmet Ali ŞİMŞEK[1*] (ORCID: 0000-0002-6127-2195)
Ahmet SERTBAŞ[2] (ORCID: 0000-0001-8166-1211)

[1] Department of Computer Technologies, Vocational School of Technical Sciences, Tekirdag Namik Kemal University, Tekirdag, TURKEY.
[2] Department of Computer Engineering, Faculty of Engineering, Istanbul University-Cerrahpaşa, Istanbul, TURKEY.

*Sorumlu Yazar / Corresponding Author: Mehmet Ali ŞİMŞEK, masimsek@nku.edu.tr

ABSTRACT

Meniscal tears are a disease that occurs in the knee joint and negatively affects people's mobility. In this study, the performance of the state-of-the-art (SOTA) YOLO (You Only Look Once) models, in particular YOLOv8l, YOLOv8x, YOLOv9c, YOLOv9e, YOLOv10l, and YOLOv10x, for the detection of meniscal tears was investigated. The algorithms were trained and tested with data from magnetic resonance imaging (MRI). In our study, the YOLOv9e model showed the highest performance and achieved the best results in the training phase with a mAP50 of 0.91807, a precision of 0.87684, a recall of 0.93871 and an F1 score of 0.90672. This study makes a unique contribution to the field with its advanced algorithms and comprehensive performance analysis. The findings show that deep learning algorithms are suitable for clinical use in the automatic detection and localization of meniscal tears. In this way, the possibility of early diagnosis increases, and patients can be directed to the right treatment, preventing joint problems that may occur in the future. In future studies, it is aimed to increase the generalization capabilities of the models with larger data sets and different anatomical structures.

Keywords: Osteoarthritis, meniscus tear, magnetic resonance imaging, deep learning, YOLO series,

ÖZET

Menüsküs yırtıkları diz ekleminde meydana gelen ve insanların hareket kabiliyetlerini olumsuz etkileyen bir hastalıktır. Bu çalışmada, menisküs yırtıklarının tespiti amacıyla YOLOv8l, YOLOv8x, YOLOv9c, YOLOv9e, YOLOv10l ve YOLOv10x gibi son teknoloji YOLO (You Only Look Once) modellerinin performansı incelenmiştir. Algoritmalar, manyetik rezonans görüntüleme (MRG) görüntülerinden elde edilen veriler üzerinde eğitilmiş ve test edilmiştir. Çalışmamızda kullanılan YOLOv9e modeli, eğitim sürecinde elde edilen en iyi sonuçlarda 0,91807 mAP50, 0.87684 Precision, 0.93871 Recall ve 0.90672 F1 Score değerleriyle en yüksek başarıyı göstermiştir. Bu çalışma, kullanılan ileri seviye algoritmalar ve kapsamlı performans analizi ile alanda özgün bir katkı sağlamaktadır. Elde edilen bulgular, derin öğrenme algoritmalarının menisküs yırtıklarının otomatik tespiti ve lokalizasyonunda klinik kullanıma uygun olduğunu göstermektedir. Bu sayede erken teşhis olasılığı artmakta ve hastaların doğru tedaviye yönlendirilmesi sağlanarak ilerleyen dönemde oluşabilecek eklem sorunlarının önüne geçilebilmektedir. İlerleyen çalışmalarda daha geniş veri setleri ve farklı anatomik yapılarla yapılacak araştırmalarla modellerin genelleme yeteneklerinin artırılması hedeflenmektedir.

Anahtar Kelimeler: Osteoartrit, menüsküs yırtığı, manyetik rezonans görüntüleme, derin öğrenme, YOLO serileri.

ToCite: ŞİMŞEK, M. A., & SERTBAŞ, A. (2025). AUTOMATIC DETECTION OF MENISCUS TEARS FROM KNEE MRI IMAGES USING DEEP LEARNING: YOLO V8, V9, AND V10 SERIES. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, 28(1), 292-308.

INTRODUCTION

Osteoarthritis (OA) is a degenerative joint disease caused by damage to the articular cartilage and underlying bone, particularly in weight-bearing joints. OA of the knee is the most common type of recurrent arthritis affecting older people. This disease is characterized by the wear and deterioration of the cartilage in the knee joint over time, resulting in significant disability and limitations in the activities of older people(Almajalid et al., 2019; Bilge et al., 2018; Gaj et al., 2020a).

The meniscus is an important component of the human body. Each knee has two fiber-reinforced menisci, one on the inside and one on the outside. The menisci are crescent-shaped, convex, and triangular on the outside, tapering from the outside inwards and covering two thirds of the surface of the tibial plateau. The menisci are flexible and consist of dense and tightly woven collagen fibers to withstand compressive forces. The shock-absorbing properties of the menisci nourish the articular cartilage and stabilize the knee by protecting it from high pressure and distributing weight evenly (Bryceland et al. 2017; Makris et al., 2011; Ölmez et al., 2020). Degeneration or tearing of the meniscal tissue is an important factor in the development of OA. Since the meniscus contributes to the stability and shock absorption of the knee joint, damage to this tissue can lead to wear and degeneration of the articular cartilage, which can trigger or accelerate the OA process. Meniscus tears are common knee injuries and can occur in people of all ages. Meniscus tears can occur as a result of sporting activities, aging, or trauma. Figure 1 shows examples of healthy and tear meniscus. These images visually illustrate the textural differences that our model attempts to recognize and concretize the structural differences between a torn meniscus and a healthy meniscus.

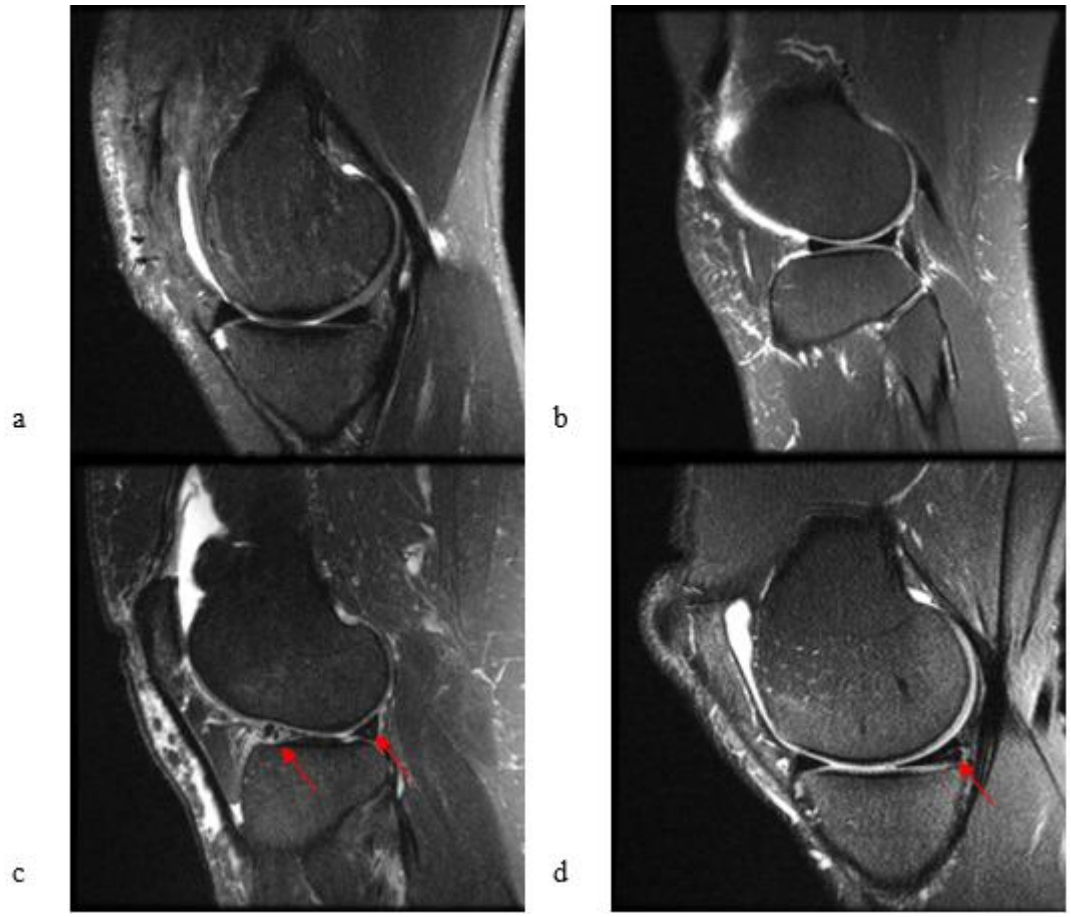

Figure 1. Examples of Healthy and Tear Meniscus. (a) and (b) Show Images of a Healthy Meniscus, While (c) and (d) Show Images of a Tear Meniscus (The Red Arrows Indicate the Tear).

Modern imaging techniques such as MRI, which are widely used in clinical practice, make it possible to examine joint structures. However, the manual measurement of these structures is time-consuming (Almajalid et al., 2019). The difficulty in diagnosing meniscus tears can lead to incorrect treatment or late diagnosis of the condition, which can have a negative impact on the treatment process and even cause permanent damage to the knee joint if left untreated.

Recent advances in deep learning (DL) algorithms have led to significant improvements in the diagnosis of meniscal tears, which present a clinical challenge. These tears are usually detected by MRI. However, they can be difficult to diagnose due to structural defects in the image. DL algorithms, in particular convolutional neural networks (CNN), have shown remarkable performance in accurately detecting these injuries (Botnari et al., 2024; Ozeki et al., 2021).

The success of DL methods, especially in analyzing image data, is quite high. Neural networks and in particular CNN architecture make it possible to detect lesions and anomalies in medical images. Some studies have successfully detected meniscus tears using machine learning (ML) and DL methods (Couteaux et al., 2019; Gaj et al., 2020b; Harman et al., 2023; Hung et al., 2023; Ma et al., 2023; Roblot et al., 2019; Saygili & Albayrak, 2017a, 2020; Saygılı & Albayrak, 2017b; Ying et al., 2024). Further integration of these DL-based technologies into clinical workflows will streamline diagnostic processes and contribute significantly to the reliable detection of meniscal pathologies, improving patient outcomes (Ma et al., 2023).

One of the latest DL methods is YOLO, an object detection algorithm proposed by Redmon et al. in 2016. This algorithm divides an input image into $S \; x \; S$ grid cells and estimates B bounding boxes and the corresponding class probabilities for each cell. YOLO aims at fast and accurate object detection by converting the object detection problem into a regression problem(Redmon et al., 2016; Su et al., 2023). The latest version of YOLO-based algorithms is the YOLOv11 algorithm. In this study, we used the current YOLO algorithms and investigated the performance of these algorithms on meniscal tears.

The use of YOLO series and YOLO-based hybrid architectures in MRI imaging of knees has shown promising results. For example, one study investigated the effectiveness of advanced deep learning models such as YOLOv8 and EfficientNetV2 on sagittal and coronal MRI images of meniscus tears. In this study, the YOLOv8 version was used to localise the meniscus and then the EfficientNetV2 model was used to detect meniscus tears (Güngör et al., 2024). In another study using EfficientNet and YOLO series, the Scaled-YOLOv4 model detects the location of the meniscus and the EfficientNet-B7 model classifies meniscal tears (Chou et al., 2023).

In addition, the SE-YOLOv5 model based on squeeze-and-excitation (SE) inception attention was developed in a study on the automatic detection of cystic lesions in the knee using deep learning methods based on MRI images (Xiongfeng et al., 2022). In another study, a recurrent neural network (RNN) model was proposed for the classification and localization of knee ligament injuries on MRI images. This model was shown to provide better classification and localization compared to YOLOv3 (Zhu et al., 2022).

In a study on the automatic diagnosis of discoid lateral meniscus, it was reported that the YOLOv3 algorithm can diagnose this condition (Li et al., 2021a). In another study, a lesion detection network was trained using the YOLOv5S architecture for the detection of meniscal tears (Li et al., 2021b). In another study, a lesion detection network was trained using the YOLOv5S architecture for the detection of meniscus tears (Zhao et al., 2021). In another study, Darknet-53 proposed a new model that can be integrated into the YOLOv4 algorithm to analyze coronal and sagittal MRI images (Hung et al., 2023).

Although there are already several studies analysing MRI images of the knee using YOLO series, there is still potential for development in this area. Few studies are comparing the performance of the different versions of the YOLO algorithm in detecting meniscal tears. In particular, a comprehensive evaluation of the accuracy, precision, and overall efficiency of the latest versions in detecting meniscal pathologies has not yet been performed. In this study, we focus on the diagnosis of meniscal tears with the current YOLO series (YOLOv8, YOLOv9, YOLOv10) using the MRNet dataset.

The clinical contributions of our work focus on the accurate and rapid diagnosis of meniscal tears. The developed DL-based methods will relieve radiologists and help them to diagnose faster. In particular, early diagnosis of a meniscus tear will ensure that the appropriate treatment plan is implemented in a timely manner and prevent meniscus tears from progressing and causing more serious joint problems. This can improve the patient's general condition, reduce the need for surgery, and speed up the rehabilitation process.

MATERIALS AND METHODS

Selection of the Dataset and Labeling of the Data

The MRNet dataset consists of 1,370 knee MRI examinations performed at Stanford University Medical Centre. The dataset contains 1,104 (80.6%) abnormal exams with 319 (23.3%) ACL tears and 508 (37.1%) meniscal tears; the labels were manually extracted from the clinical reports (Bien et al., 2018). For this study, 250 examinations were randomly selected from the MRNet dataset and 203 examination images were used based on the exclusion criteria. Of these examinations, 97 (47.78%) were examinations without meniscal tears and 106 (52.22%) were examinations with meniscal tears. Figure 2 shows the dataset selection process and exclusion criteria.

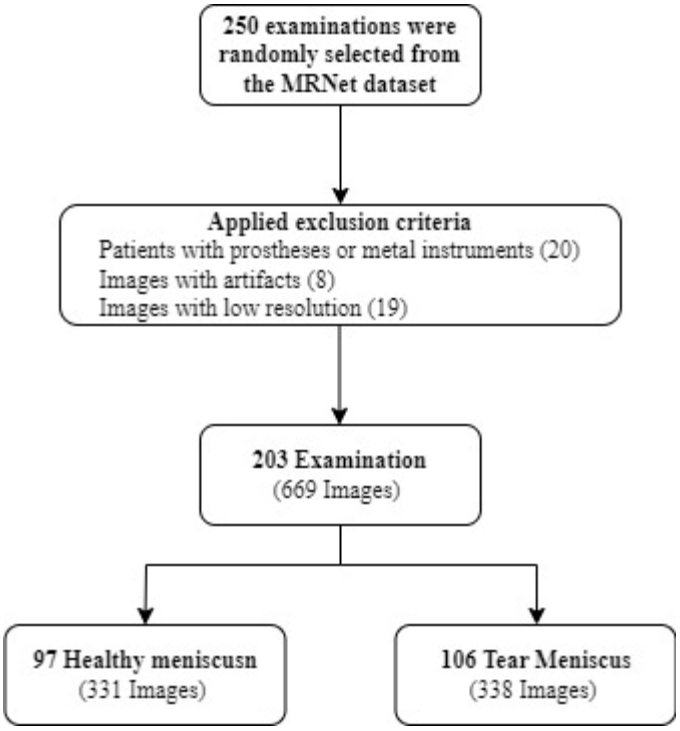

Figure 2. Process of Dataset Creation.

Since the sagittal section provides the most information about meniscal tears (Ma et al., 2023) only sagittal section images were used in this study. Images showing both the anterior and posterior horn in the sagittal plane were selected. From the images of 203 examinations, a total of 669 images were extracted in PNG (Portable Network Graphic) file format. Of these images, 338 (50.52%) were images with meniscal tears and 331 (49.48%) were healthy images without meniscal tears. The dataset was split to 70:25:5 for train, validation, and testing. Table 1 shows the distribution of the dataset.

Table 1. Number of Images in the Train, Validation, and Test Dataset.

	Healthy	Tear	Total
Train	237	232	469
Validation	84	83	167
Test	17	16	33

The meniscus area in images with a meniscus tear is labeled “tear” and the meniscus area in images without a meniscus tear is labeled “no tear”. The anterior horn and posterior horn were labeled separately in images without a meniscal tear, while only the tear area was labeled in images with a meniscal tear. The number of labels in the images of the train, validation, and test sets can be found in Table 2. The online tool RoboFlow was used for labeling the images, sizing, and partitioning the dataset.

Table 2. Number of Tags in the Train, Validation, and Test Dataset.

	Train	Valid	Test
no_tear	417	140	32
tear	333	117	24

Preprocessing of Images

Since the obtained MRI images have different resolutions, they were preprocessed and converted to the input sizes (640x640 pixels) of the YOLO algorithm. This process was performed with RoboFlow.

Experimental Setup and Hyperparameters

All experiments in this study were conducted using the Google Colab platform, where computationally intensive operations are performed by Google servers. The required libraries were provided by this platform. The analysis of the dataset was performed on an NVIDIA A100-SXM4-40GB GPU with the Python-3.10.12 programming language using torch-2.4.1+cu121 with the latest YOLOv8, YOLOv9 and YOLOv10 libraries from Ultralytics. The hyperparameters used in these experiments were kept constant throughout the study and are listed in Table 3.

Table 3. Experimental Hyperparameters.

Hyperparameter	Value
workers	2
batch	12
device	0
epochs	100
lr0	0.01
lrf	0.01
momentum	0.95
weight_decay	0.0001
warmup_epochs	10
warmup_momentum	0.5
warmup_bias_lr	0.1
optimizer	SGD

The hyperparameters used in this study were carefully selected to efficiently train the model and optimize its overall performance. The batch size was set to 12 to compensate for memory usage during training, while the value of epochs was set to 100 to ensure that the model processes the data appropriately. The learning rate (lr0) was set to 0.01 and the momentum to 0.95 to achieve rapid convergence of the model. The weight decay was set to 0.0001 to increase the generalization ability of the model. At the beginning of the training process, a low momentum (0.5) and learning rate (0.1) were used for 10 epochs in the warm-up phase so that the model learns more stably. Finally, the SGD optimization algorithm was used to update the model parameters. The balanced use of these hyperparameters contributed to the high performance of the model in detecting meniscus tears.

YOLO Series

YOLO is an object recognition algorithm proposed by Redmon et al. in 2015. YOLO aims at fast and accurate object recognition by converting the object recognition problem into a regression problem (Redmon et al., 2015; Su et al., 2023). The latest version of YOLO-based algorithms is YOLOv11. Over the years, numerous versions of YOLO have been released with innovative techniques to improve performance. Figure 3 shows the timeline of YOLO versions.

The three current versions of the YOLO algorithms and the models with the highest number of parameters (YOLOv8l, YOLOv8x, YOLOv9c, YOLOv9e, YOLOv10l, and YOLOv10x) were used in this study. The reason for choosing these models is that they have the potential to learn more complex structures and provide more accurate results due to their high number of parameters. Especially when detecting subtle structural changes such as meniscus tears, the deeper learning capacity of larger models can improve accuracy and allow better generalization to complex images.

YOLOv8(Jocher et al., 2023; Sohan et al., 2024; Terven et al., 2023), 2023 is an object detection model released by Ultralytics in 2023 that offers significant improvements over previous versions. This model offers flexibility through multiple scaled versions for different application domains, including YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large) and YOLOv8x (extra-large). YOLOv8 introduces the C2f module, which significantly improves its architecture and extends the CSP (Cross-Stage Partial) layer. This module enables the combination of high-level features with contextual information, improving the accuracy of the model. YOLOv8 also

improves the overall accuracy of the model by moving to a decoupled header structure that independently handles object presence, classification, and regression tasks(Alif & Hussain, 2024; Sapkota et al., 2024; Wang & Liao, 2024).

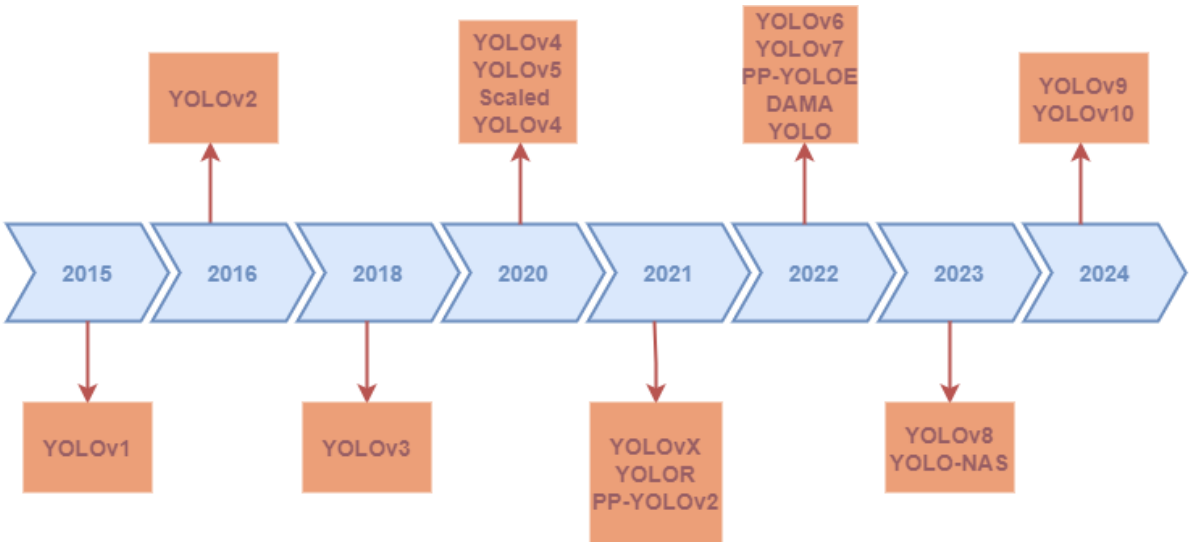

Figure 3. Timeline of the YOLO Versions: YOLOv1 Through YOLOv10.

YOLOv9 is the next generation object recognition model, launched in 2024, and builds on the benefits of previous versions. It aims to preserve data integrity in deep learning processes by introducing Programmable Gradient Information (PGI) and the Generalized Efficient Layer Aggregation Network (GELAN) architecture, an innovative approach to reduce information loss. YOLOv9 is designed to prevent data corruption - a common problem with deep neural networks - while ensuring a better flow of information and improving object recognition capabilities in complex and dynamic environments (Alif & Hussain, 2024; A. Wang et al., 2024; Wang & Liao, 2024).

YOLOv10(A. Wang et al., 2024; Wang & Liao, 2024) is the latest object recognition model released in 2024. It offers significant innovations and combines all the advantages of the previous versions. It significantly reduces recognition time by eliminating the need for traditional NMS (Non-Maximum Suppression). YOLOv10 manages to increase detection accuracy without compromising speed by optimizing the training protocol with a double label assignment strategy. The architecture of the model includes innovative components such as lightweight classifier heads and discrete subsampling with spatial channels that minimize information loss (Alif & Hussain, 2024; Sapkota et al., 2024; A. Wang et al., 2024; Wang & Liao, 2024).

YOLOv8, YOLOv9, and YOLOv10 can be seen as successive advances in the field of object recognition in the context of deep learning. Each model improves the functions of the previous versions and offers greater speed, accuracy, and efficiency. While YOLOv8 provides flexibility with architectural improvements and multiscale versions, YOLOv9 improves performance with innovative approaches to reduce information loss. YOLOv10 combines all the advantages of the previous versions and offers innovative features that reduce recognition time and increase accuracy. These three models play an important role in the development of object recognition technology and pave the way for future applications.

Metrics of Success

The metrics Recall (R), Precision (P), Average Precision (AP), Mean Average Precision (mAP) and F1 score were used to evaluate the performance of the model. These metrics are the most used metrics for evaluating object recognition models and their expressions are given in Equations 1-5.

$$R = \frac{TP}{TP+FN} \tag{1}$$

$$P = \frac{TP}{TP+FP} \tag{2}$$

$$AP = \int_0^1 P(r)dr \tag{3}$$

$$mAP = \frac{\sum_{i=1}^{N} AP_i}{N} \quad (4)$$

$$F1\ score = 2\frac{P\ x\ R}{P + R} \quad (5)$$

True Positive (TP) refers to the number of instances in the dataset that the model correctly predicts as positive, while False Negative (FN) describes cases where the model incorrectly classifies an instance that should be positive as negative. False Positive (FP) is when the model incorrectly predicts a sample as positive when it should be negative. Recall indicates how many instances the model correctly predicts as positive and evaluates the ratio between TP and FN. Precision indicates how many of the instances that the model predicts as positive are correct and measures the trade-off between TP and FP. Ideally, a model should have high precision and recall values, but it can often be difficult to achieve a balance between these two metrics. The mean average precision (mAP) represents the average of the average precision (AP) values calculated for each class. AP is a metric that is calculated considering the precision and recall of the model at a given threshold and is usually presented as an average of the performance obtained for different classes.

F1 score is a metric that evaluates the trade-off between precision and recall. This metric is calculated as the harmonic mean of precision and recall and measures the model's performance on false-positive and false-negative predictions in a balanced way. The F1 score is an important indicator, especially for unbalanced datasets, as it more accurately reflects the overall performance of the model.

RESULTS

This study evaluates the performance of the YOLOv8, YOLOv9 and YOLOv10 series for the detection of meniscal tears based on the results from the training, validation, and testing phases. The performance measures such as mAP, precision, recall, and F1 score of each of the models in the training process were examined and qualitative and quantitative analyses were performed based on the classification results obtained in the test set.

A total of 6 experiments were carried out with the two versions of the YOLOv8, YOLOv9, and YOLOv10 series with the most parameters. Table 4 shows the comparison of the number of parameters, GFLOPs, and processing time for these trainings. The YOLOv8x model has the highest value with 68.1 million parameters, while the YOLOv9c and YOLO10l models have the lowest values with around 25 million parameters. In terms of FLOPs, the YOLOv8x model has the highest value with 257.4 GFLOPs. This indicates that the model is more computationally intensive, while the YOLOv9c model requires the least computing power with 102.3 GFLOPs. In terms of processing time, the Yolov8l model requires the shortest time (13 min 8 sec), while the YOLOv9e model requires the longest time (27 min 40 sec). In particular, the YOLOv9c model stands out as a computationally efficient model with its small number of parameters and FLOPs, while models such as YOLOv8x and YOLOv9e, as complex structures that require more calculations, have longer processing times.

Table 4 Comparison of the Number of Parameters, FLOP and Processing Time of the Train.

	Parameters/M	**FLOPs/G**	**Time/min**
YOLOv8l	43,6	164.8	13 min 08 sec
YOLOv8x	68,1	257.4	16 min 56 sec
YOLOv9c	25,3	102.3	15 min 42 sec
YOLOv9e	57,4	189.1	27 min 40 sec
YOLOv10l	25,7	126.3	16 min 56 sec
YOLOv10x	31,6	169.8	18 min 44 sec

Table 5 shows the values of the performance indicators Last and Best achieved by the models during the training process. In comparison between the models, the YOLOv9e model achieved the highest performance metrics overall. In both the last training results and the best results, this model had the highest values in terms of mAP50 (Last: 0.90308, Best: 0.91807) and recall (Last: 0.90522, Best: 0.93871), showing significant success in recognizing complex objects. The high mAP50 metric indicates that the intersection between the recognized area and the labeled area is high. On the other hand, the YOLOv8l and YOLOv8x models stand out as lower cost models that offer a balanced performance in terms of precision and F1 score values. The YOLOv10x model, on the other hand, generally has lower accuracy values. Considering the high number of parameters and FLOPs, it can be said that the YOLOv9e model is preferable for complex and challenging object detection problems in terms of accuracy and efficiency.

The YOLOv9e model showed the highest overall performance among the models tested and achieved the best values. In particular, it achieved a peak mAP50 of 0.91807, a precision of 0.87684, a recall of 0.93871, and an F1 score of 0.90672 during training. These metrics show that the YOLOv9e model is particularly effective at recognizing complex objects and has high accuracy in matching detected regions to labeled regions in the dataset. This performance makes it particularly suitable for demanding applications, such as the automatic detection and localization of meniscus tears.

Table 5. Comparison of the Last and Best Performance Metrics Achieved by the Models During the Training Process.

	LAST				BEST			
	mAP50	**Precision**	**Recall**	**F1 Score**	**mAP50**	**Precision**	**Recall**	**F1 Score**
YOLOv8l	0.88859	**0.83977**	0.86968	0.85446	0.89789	0.87019	0.91230	0.89075
YOLOv8x	0.89370	0.83390	0.87592	0.85439	0.89896	0.85991	0.92659	0.89201
YOLOv9c	0.88149	0.83096	0.85610	0.84334	0.90505	0.85839	0.93654	0.89576
YOLOv9e	**0.90308**	0.82948	**0.90522**	**0.86570**	**0.91807**	**0.87684**	**0.93871**	**0.90672**
YOLOv10l	0.89527	0.83659	0.85668	0.84652	0.89527	0.86764	0.89942	0.88324
YOLOv10x	0.86525	0.83454	0.85458	0.84444	0.88694	0.86887	0.90943	0.88869

During training, performance analysis was performed using graphs showing the mAP50, precision, recall, and F1 score values of the models during each epoch. Figure 4 shows the mAP50 metric values of the models during training. In the beginning, all models increase their performance at a certain rate, but this increase slows down over time. After the 40th epoch, the learning rate slows down for all models. After 80 epochs, the models show that the increase in performance slows down and eventually reaches a constant level. This saturation point indicates that the model will no longer achieve a significant improvement in performance with further training. Although the YOLOv10l model achieves higher mAP50 values with fewer epochs, the YOLOv9e model achieves the best mAP50 value at the end of training.

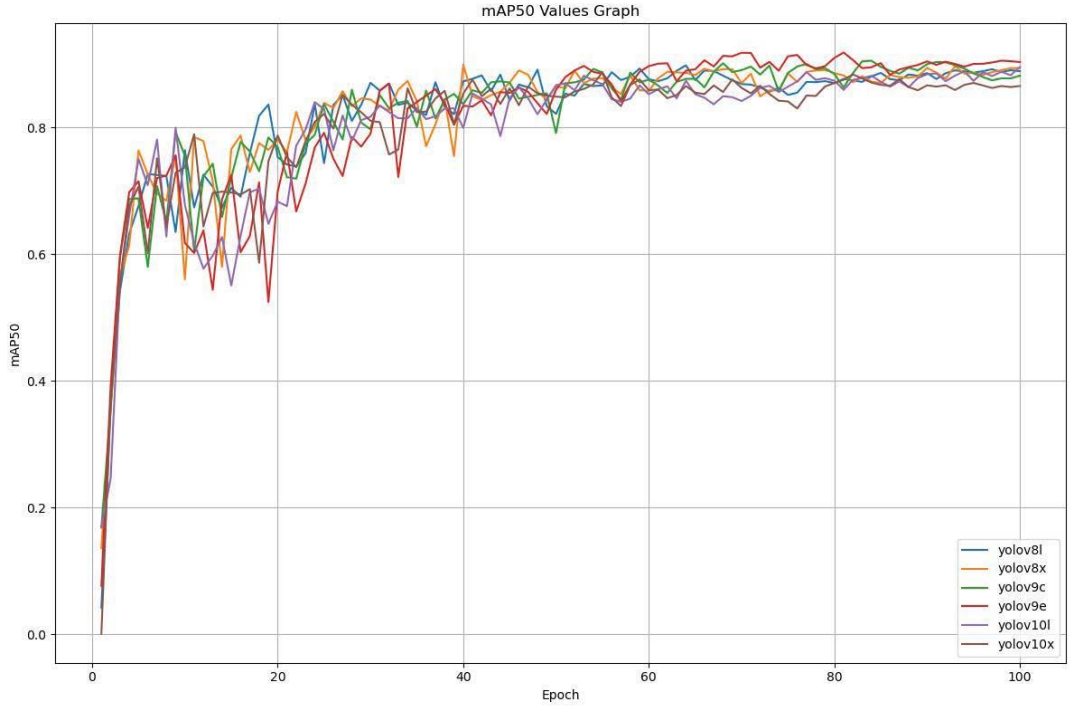

Figure 4. Variation of the mAP50 Metric Over 100 Epochs.

Figure 5 shows the precision metric values of the models during training. It can be seen that the precision values of all models increase at the beginning of the training process, but after 80 epochs the increase slows down or even stops. The YOLOv9c and YOLOv9e models showed good overall performance among their peers and achieved higher accuracy values without overfitting. Although YOLOv8l has the highest value at the end of training, YOLOv9e has the highest value throughout training. Eliminating the drop of the YOLOv9e model at the end of training indicates that this model will be the optimal choice for solving the meniscus tear recognition problem.

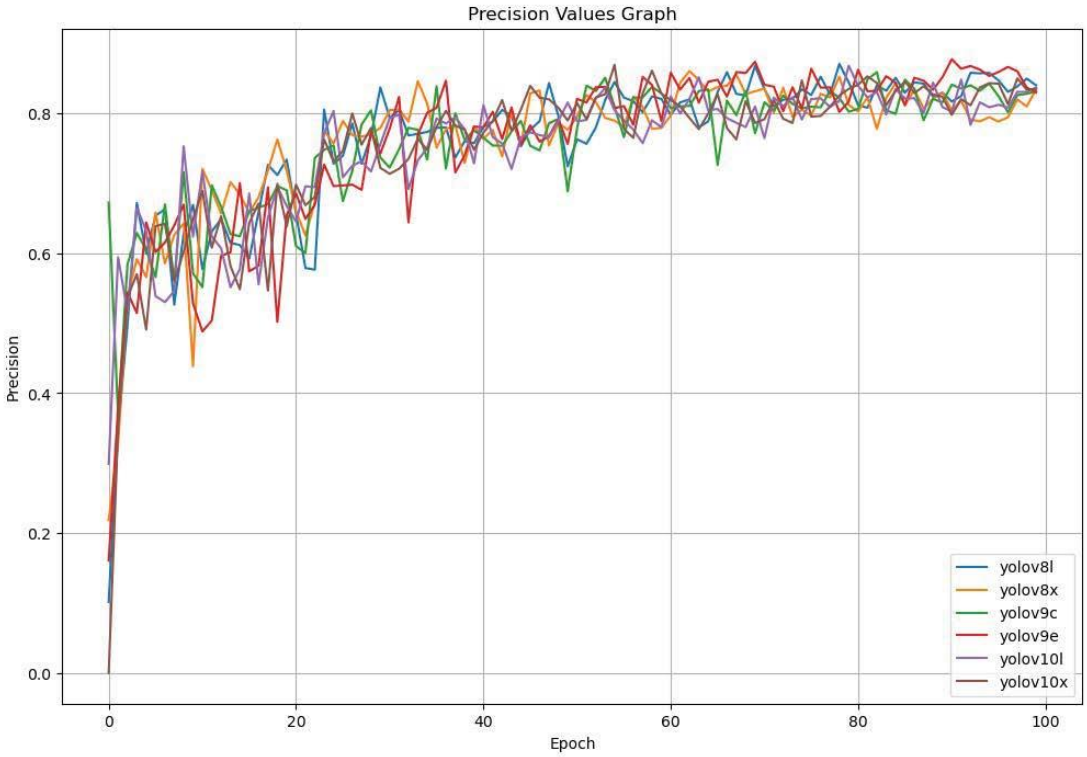

Figure 5. Variation of the Precision Metric Over 100 Epochs.

Figure 6 shows the recall metric values of the models during training. Although the models YOLOv8l and YOLOv8x achieve higher recall values in a shorter time than the other models, YOLOv9e provides the best recall value. Although all models show high performance in the first epochs, it can be observed that the performance decreases after the trial.

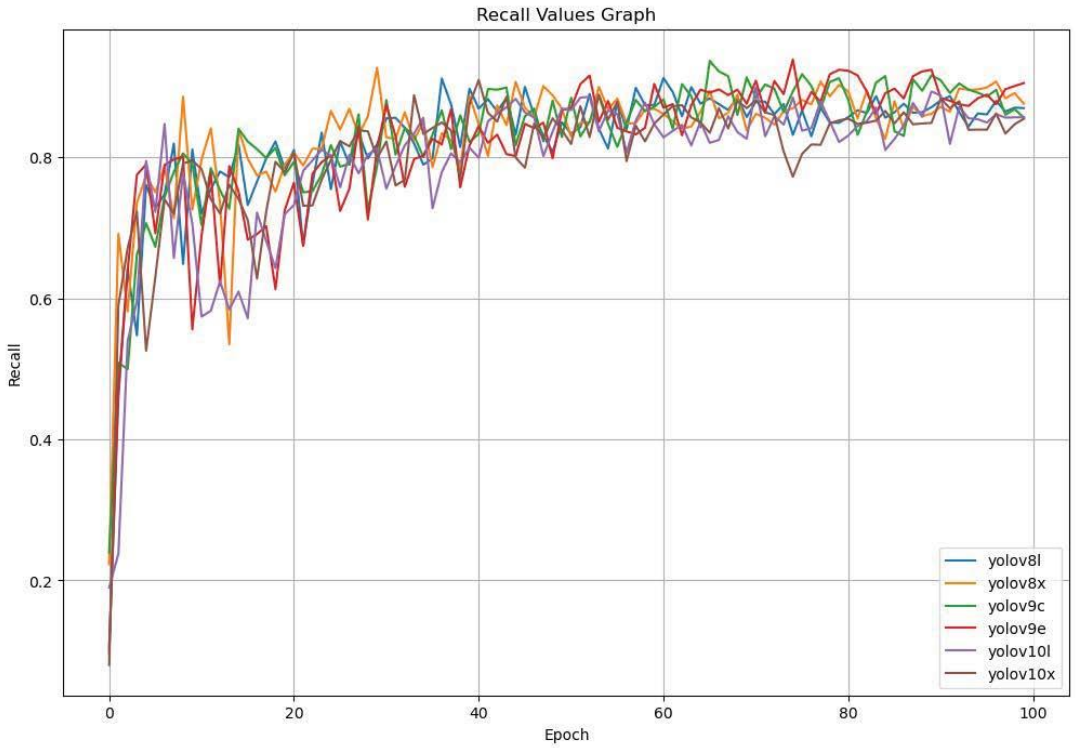

Figure 6. Variation of the Recall Metric Over 100 Epochs.

Figure 7 shows the F1 score of each model during training. This metric is used to balance precision and recall and is suitable for models with an unbalanced class distribution. It may improve rapidly at a certain point during training. This indicates that the learning process of the model is efficient. It shows that the learning process is high in the first

epochs of training. The fact that the F1 score of the YOLOV9e model increases faster compared to other models may indicate that this model has a high learning capacity and is trained more effectively.

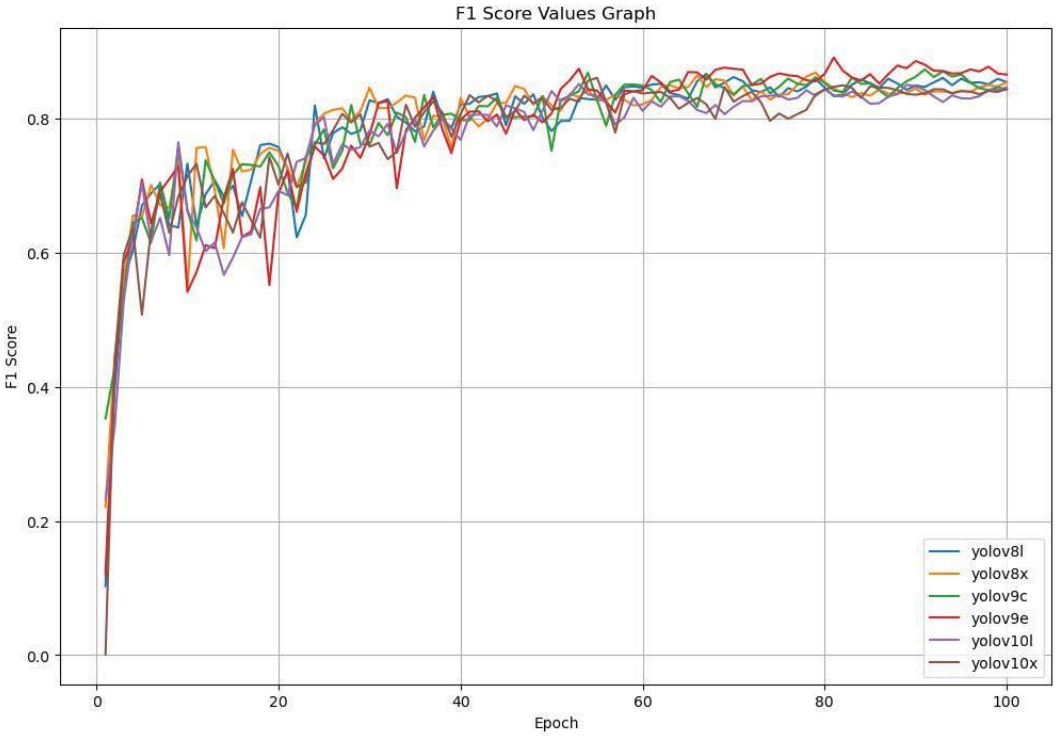

Figure 7. Variation of the F1 Score Metric Over 100 Epochs.

The YOLOv8, YOLOv9 and YOLOv10 series provide reliable and very accurate results for the detection of meniscal tears. In particular, the high values achieved for critical metrics such as mAP50 and recall speak for the use of these models for real-time applications in the medical field.

Table 6 shows the precision, recall and AP50 values of the models used in the train set for the classes no_tear and tear. The YOLOv9c and YOLOv9e models are characterized by their strong performance, especially in the recognition of complex objects. The YOLOv9c model achieved the best results in the "no_tear" class with an AP50 value of 0.909. The same model also showed a balanced performance in the precision and recognition values and achieved successful recognition in both classes. The YOLOv9e model also has a high recall value of 0.950 in the "no_tear" class and a high performance in the "tear" class with a precision of 0.888 and an AP50 of 0.895. The YOLOv10l model achieved the highest AP50 value (0.906), especially in the recognition of healthy images and proved to be an ideal model for this class.

Table 6. Performance Comparison of the Models Used in the Train Set for the Classes.

Model	Class	Precision	Recall	AP50
YOLOv8l	**no_tear**	0.819	0.893	0.887
	tear	0.828	0.855	0.898
YOLOv8x	**no_tear**	0.803	0.903	0.888
	tear	0.804	0.841	0.884
YOLOv9c	**no_tear**	0.810	0.943	**0.909**
	tear	**0.906**	0.786	0.899
YOLOv9e	**no_tear**	0.805	**0.95**	0.894
	tear	0.888	**0.88**	0.895
YOLOv10l	**no_tear**	**0.841**	0.893	0.886
	tear	0.833	0.821	**0.906**
YOLOv10x	**no_tear**	0.839	0.900	0.876
	tear	0.847	0.786	0.865

Table 7 shows the precision, recall and AP50 values of the models used in the validation set for the classes no tear, and tear. The YOLOv9e and YOLOv9c models perform well on the data without meniscus tears, while the

performance of these models is lower on the data with meniscus tears. If we evaluate Table 6 and Table 7 together, we see that a good balance is achieved between the train and validation sets of the YOLOv9e and YOLOv9c models.

Table 7. Performance Comparison of the Models Used in the Validation Set for the Classes.

Model	Class	Precision	Recall	AP50
YOLOv8l	**no_tear**	0.820	0.893	0.887
	tear	0.837	0.863	**0.906**
YOLOv8x	**no_tear**	0.797	0.896	0.88
	tear	0.803	0.838	0.885
YOLOv9c	**no_tear**	0.807	0.943	**0.909**
	tear	**0.906**	0.786	0.899
YOLOv9e	**no_tear**	0.805	**0.95**	0.894
	tear	0.888	**0.879**	0.895
YOLOv10l	**no_tear**	**0.840**	0.893	0.885
	tear	0.833	0.821	0.906
YOLOv10x	**no_tear**	0.839	0.892	0.875
	tear	0.847	0.786	0.865

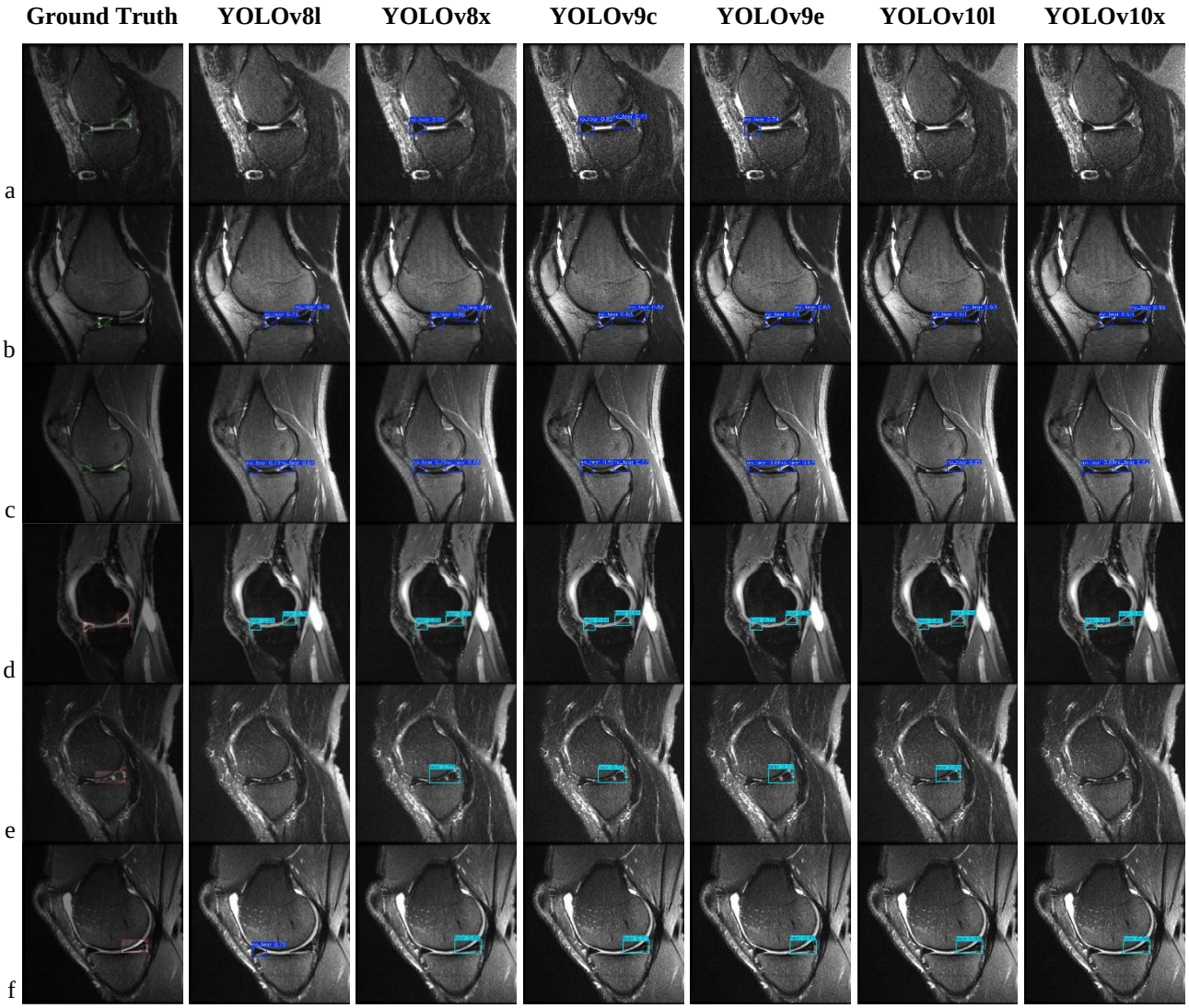

Figure 8. Comparison of the Prediction Results of the Models Used with Ground Truth.

Examples of the prediction results of the architectures YOLOv8, YOLOv9, and YOLOv10 are shown in Figure 8. From the MRI images in the sagittal plane of the test dataset, 3 healthy images and 3 images with meniscal tears were selected. Although the meniscus in the image in Figure 8a is healthy in both the anterior and posterior horns, this region could not be correctly predicted in the YOLOv8l, YOLOv10l, and YOLOv10x models. In the architectures YOLOv8x and YOLOv9e, only the anterior horn could be predicted correctly. In the image in Figure 8c, both horns

of the meniscus are healthy, but only the posterior horn was incorrectly predicted in the YOLOv10l model. In the image in Figure 8e, there is a vertical meniscal tear in the posterior horn. This tear was not correctly predicted by the YOLOv8l and YOLOv10x models. Finally, the image in Figure 8f shows a horizontal meniscus tear in the posterior horn that was not predicted by the YOLOv8x model but correctly identified the anterior horn as healthy. The same tear was not correctly predicted in the YOLOv10x model. Looking at the accuracy of the predictions for the test dataset shown in Table 8 by class, it can be seen that the YOLO models provide results with high accuracy in most cases and the number of correct predictions is high. However, in this section, the cases with incorrect predictions are highlighted.

The test dataset consists of 33 images. Table 8 shows the number of correct and incorrect detections of the YOLO algorithms according to the classes tear and no tear in the test set. In general, all models achieved a high number of correct detections in both classes. YOLOv9c performed the best with 0 false detections in the classes tear and no tear, while YOLOv8x performed the worst with 5 false detections. All models appear to be successful in detecting meniscus tears.

Table 8. Number of Correct and Incorrect Recognitions of YOLO Models According to the Classes "Tear" and "No_Tear" in the Test Set.

	YOLOv8x			**YOLOv8l**	
	Correct	**Incorrect**		**Correct**	**Incorrect**
tear	24	0	**tear**	22	2
no_tear	31	1	**no tear**	29	3
	YOLOv9c			**YOLOv9e**	
	Correct	**Incorrect**		**Correct**	**Incorrect**
tear	24	0	**tear**	24	0
no_tear	32	0	**no tear**	31	1
	YOLOv10l			**YOLOv10x**	
	Correct	**Incorrect**		**Correct**	**Incorrect**
tear	24	0	**tear**	23	1
no_tear	29	3	**no tear**	29	3

DISCUSSION

In this study, the most parameterized (YOLOv8x, YOLOv8l, YOLOv9c, YOLOv9c, YOLOv9e, YOLOv10l, YOLOv10x) models of the current YOLO algorithms YOLOv8, YOLOv9, and YOLOv10 are examined for their success in detecting meniscal tears.

Examinations were randomly selected from the examination images in the MRNet dataset and a suitable dataset for the study was created through the necessary pre-processing. As can be seen in Figure 2, the number of healthy images and images with meniscus tears are close to each other. Looking at Table 1 and Table 2, it can be said that the created dataset is balanced. Each class is evenly distributed in the train, validation, and test sets. This is to reduce the risk of overfitting the models, achieve better generalization ability, and obtain consistent metric values. This helps to obtain more reliable and accurate predictions when fitting the models to real data.

During the training process, the models were evaluated using the metrics mAP50, precision, recall, and F1 score. The scores of the individual classes in both the train and validation process were calculated using the metrics precision, recall AP50. The comparison of the last and best performance metrics achieved by the models in the training process is shown in Table 5. The YOLOv9e model is one of the models with the most parameters, as can be seen in Table 4. Due to its large number of parameters and compact structure, YOLOv9e seems to be the most successful model. The YOLOv9e model seems to give the best results for both the last metric values and the best metric values. YOLOv9c has relatively few parameters, but the model is free of unnecessary parameters and its compact structure ensures that the model performs more optimally. Considering the performance of the metrics, the YOLOv9c model seems to be very balanced, and considering the number of parameters, FLOPs, and performance results, it can be said that it can be preferred in real-time applications. Compared to the other models, it has the least training time, and the training time is shorter than the other models (except YOLOv8l), indicating that it generalizes quickly. YOLOv9 shows better results in object recognition thanks to its more complex structures and attention mechanisms in the backbone and head networks.

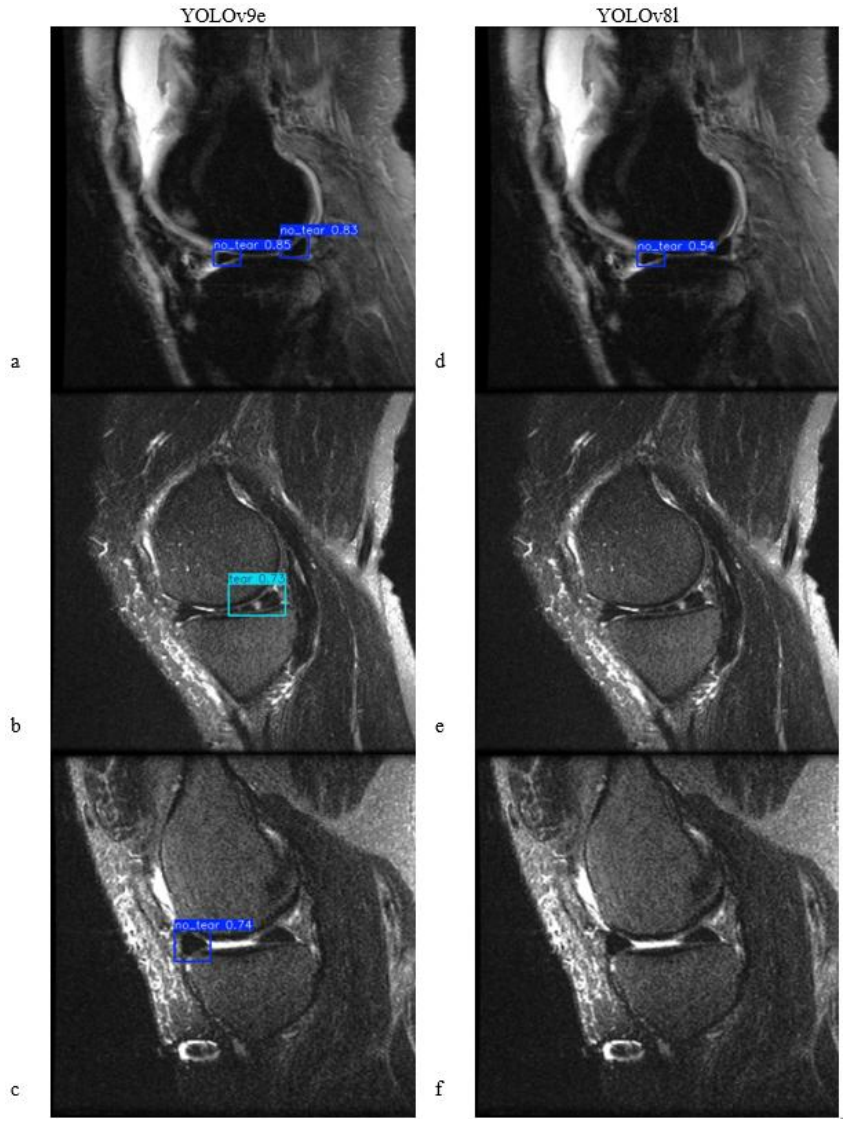

Figure 9. Comparison of Prediction Results Between the YOLOv9e and YOLOv8l Models on MRI Images. Images (a), (b), and (c) Show the Predictions of YOLOv9e, While Images (d), (e), and (f) Show the Predictions of YOLOv8l.

The performance of the models in the two classes tear and no tear was analyzed during training (Table 6) and validation (Table 7). The class predictions generally show high accuracy. However, it was found that some models made relatively many false detections in the no_tear class (e.g. the YOLOv8x model). The models show good results in detecting meniscus tears in both training and validation.

Low false positive (FP) rates in medical image analysis, especially in the tear class, are of great clinical importance. It helps to ensure an accurate diagnosis. Failure to detect a meniscal tear (FN) may result in the patient not receiving treatment or delayed treatment. Incorrect detections in both classes can have a significant impact on medical decisions, so the models need to be evaluated in this respect. Table 8 shows the number of correct and incorrect detections of the YOLO models according to the tear and no_tear classes in the test set. YOLOv9c showed the best performance with 0 errors in the tear and no_tear classes, while YOLOv8x showed the worst performance with 5 false detections. The analysis of Table 8 shows that the models are more successful in detecting meniscal tears than in detecting a healthy meniscus. The low contrast and ambiguity of the meniscus structure in some regions of the images lead to inaccurate predictions. In particular, the models did not capture enough discriminative features when localizing vertical and horizontal tears, which also led to incorrect predictions. In addition, the generalization performance decreased because the models failed to detect some small tears or focused on specific tear types during training. Although the dataset is balanced, the type of tears in the images with meniscal tears varies.

Figure 8 compares example prediction results of the models used with real data. The image in Figure 8a was either not predicted or only the anterior horn was predicted in all models except YOLOv9c. The low resolution of the image

and the presence of artifacts mean that the models cannot make accurate predictions. Figure 9 displays the prediction results on the test dataset for YOLOv9e, which achieved the highest performance among the models tested, and YOLOv8l, which showed one of the lower performances. In these images, YOLOv9e (a, b, c) generally provides higher confidence levels and more accurate predictions compared to YOLOv8l (d, e, f). The YOLOv9e model makes correct predictions in Figures 9a and 9b, while in Figure 9c, it accurately detects the anterior horn but fails to identify the healthy meniscus in the posterior horn. On the other hand, the YOLOv8l model correctly identifies the anterior horn in Figure 9d but fails to detect the posterior horn and makes no predictions in Figures 9e and 9f. Analyzing these examples reveals that image quality, contrast, and the presence of artifacts play a significant role in prediction accuracy. Future improvements in image preprocessing and model fine-tuning could help mitigate these issues, enabling more reliable detection of meniscal tears.

In the literature, there are CNN-based studies on the detection, localization, and characterization of meniscal tears (Couteaux et al., 2019; Hung et al., 2023; Ma et al., 2023). There are also studies using YOLO series, which also form the basis of this study, but they are not at a sufficient level. Some of these studies also propose YOLO-based hybrid architectures (Chou et al., 2023; Güngör et al., 2024; Zhao et al., 2021). Some studies propose architectures developed using YOLO series (Chou et al., 2023; Xiongfeng et al., 2022). In these studies, YOLO versions YOLOv3(Li et al., 2021b; Zhu et al., 2022), YOLOv4 (Chou et al., 2023), YOLOv5(Xiongfeng et al., 2022) and YOLOv8 (Güngör et al., 2024) have been used. Although there have been several studies analysing MRI images of the knee using YOLO series, both the developments in the YOLO series and the inadequacy of the existing literature suggest that there is still potential for development in this area. The application of the current YOLO series in this study will contribute to the literature. A summary and comparison of these studies can be found in Table 9, which highlights the models, architectures, and objectives of existing approaches to meniscal tear detection.

Table 9. Comparison of Studies on Models for the Detection of Meniscus Tears.

Study	Model Used	Architecture	Objective	Description
(Couteaux et al., 2019; Hung et al., 2023; Ma et al., 2023)	CNN-based models	Standard CNN architecture	Meniscus tear detection, localization, and characterization	CNN-based models were utilized for meniscus tear detection, but detailed localization and characterization were limited.
(Chou et al., 2023; Güngör et al., 2024; Zhao et al., 2021)	YOLO + Hybrid	YOLO-based hybrid architectures	Meniscus tear detection and localization	Proposed YOLO-based hybrid architecture, but studies in this area are still developing within the literature.
(Chou et al., 2023; Xiongfeng et al., 2022)	Enhanced YOLO	Custom architectures using YOLO series	Meniscus tear detection	Custom YOLO-based architectures with modifications to various YOLO versions were proposed.
(Li et al., 2021b; Zhu et al., 2022)	YOLOv3	YOLOv3 standard	Meniscus tear detection	YOLOv3 was employed as a foundational model for meniscus detection.
(Chou et al., 2023; Xiongfeng et al., 2022)	YOLOv4	YOLOv4 standard	Meniscus tear detection	YOLOv4 was utilized but showed limited performance compared to newer models in the literature.
(Xiongfeng et al., 2022)	YOLOv5	YOLOv5 standard	Meniscus tear detection	YOLOv5 showed improved accuracy and was more effective for meniscus tear detection.
(Güngör et al., 2024)	YOLOv8	YOLOv8 standard	Meniscus tear detection	YOLOv8 was employed, but its application in the literature remains relatively limited.
Our Study	YOLOv8 YOLOv9 YOLOv10	YOLOv8 standard YOLOv9 standard YOLOv10 standard	Meniscus tear detection and localization	Utilizing the YOLOv9e model for improved accuracy in meniscus tear detection, our study achieved high precision and recall values, indicating the model's reliability for clinical applications.

The current study has some limitations. Firstly, the dataset used consists of 203 examinations and 669 images. Medical imaging often requires large datasets to effectively train deep learning models. In addition, the quality of images can vary significantly due to differences in imaging equipment, settings, and patient conditions. These differences can affect the performance of the model. Secondly, the MRNet dataset also contains low resolution images and the dataset consists of only sagittal slices. Although the models have shown great potential for analyzing meniscal tears, these limitations need to be addressed and the models need to be successfully integrated into clinical applications.

CONCLUSION

In this study, the YOLOv8, YOLOv9, and YOLOv10 algorithms were used to detect meniscus tears, and the results were compared. The results show that all three models can detect meniscal tears with high accuracy. In particular, the YOLOv9e model performed best in both the tear and no tear classes, making it the most successful model for clinical applications. The YOLOv9e model achieved the highest performance in the training phase with a mAP50 of 0.91807, a precision of 0.87684, a recall of 0.93871, and an F1 score of 0.90672.

This work makes several important contributions to the field of automated medical imaging analysis and meniscal tear detection. First, it provides a detailed evaluation of SOTA YOLO models for meniscal tear detection. The findings demonstrate that the YOLOv9e model, with its high accuracy metrics, can be a reliable tool for clinical diagnosis, potentially aiding radiologists by reducing time and improving diagnostic consistency. Additionally, this study introduces a robust methodology for high-quality data preparation and model optimization, which may serve as a reference for future applications in radiologic imaging. By addressing the challenges of meniscal tear localization and classification, this research sets a foundation for extending deep learning techniques to other joint pathologies, thus expanding the role of AI in radiologic diagnostics.

In future studies, expanding the dataset by increasing the number and variety of samples and including images of different anatomical structures could improve the generalizability and robustness of the model. In addition, improving the quality and resolution of the MRI data could allow the model to capture finer structural details, which could increase detection accuracy. Future research could also focus on improving the interpretability of the model so that clinicians can better understand the basis of the predictions. Beyond just localization, research could also be conducted into categorizing different types of meniscal tears according to their severity and characteristics. Finally, the development of hybrid models that combine YOLO with other deep learning architectures could improve the model's ability to handle complex cases and expand its application to other areas of knee and joint pathology.

In summary, the YOLO model series provides successful results in the automatic detection and localization of meniscal tears and is considered a powerful option to support radiologists in the clinical setting.

Contribution and Acknowledgements

We thank radiologist Dr. Hadi SASANİ for his valuable contribution to this study. His expertise and support played an important role in the success of our study.

REFERENCES

Alif, M. A. R., & Hussain, M. (2024). YOLOv1 to YOLOv10: A comprehensive review of YOLO variants and their application in the agricultural domain.

Almajalid, R., Shan, J., Zhang, M., Stonis, G., & Zhang, M. (2019). Knee Bone Segmentation on Three-Dimensional MRI. 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), 1725–1730. IEEE. https://doi.org/10.1109/ICMLA.2019.00280

Bien, N., Rajpurkar, P., Ball, R. L., Irvin, J., Park, A., Jones, E., … Lungren, M. P. (2018). Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet. PLOS Medicine, 15(11), e1002699. https://doi.org/10.1371/journal.pmed.1002699

Bilge, A., Ulusoy, R. G., Üstebay, S., & Öztürk, Ö. (2018). Osteoarthritis. Kafkas Journal of Medical Sciences, 8(50), 133–142. https://doi.org/10.5505/kjms.2016.82653

Botnari, A., Kadar, M., & Patrascu, J. M. (2024). A Comprehensive Evaluation of Deep Learning Models on Knee MRIs for the Diagnosis and Classification of Meniscal Tears: A Systematic Review and Meta-Analysis. Diagnostics, 14(11), 1090. https://doi.org/10.3390/diagnostics14111090

Bryceland, J. K., Powell, A. J., & Nunn, T. (2017). Knee Menisci. CARTILAGE, 8(2), 99–104. https://doi.org/10.1177/1947603516654945

Chou, Y.-T., Lin, C.-T., Chang, T.-A., Wu, Y.-L., Yu, C.-E., Ho, T.-Y., … Kuang-Sheng Lee, O. (2023). Development of artificial intelligence-based clinical decision support system for diagnosis of meniscal injury using magnetic resonance images. Biomedical Signal Processing and Control, 82, 104523. https://doi.org/10.1016/j.bspc.2022.104523

Couteaux, V., Si-Mohamed, S., Nempont, O., Lefevre, T., Popoff, A., Pizaine, G., … Boussel, L. (2019). Automatic Knee Meniscus Tear Detection and Orientation Classification with Mask-RCNN. Diagnostic and Interventional Imaging, 100(4), 235–242. https://doi.org/10.1016/j.diii.2019.03.002

Gaj, S., Yang, M., Nakamura, K., & Li, X. (2020a). Automated cartilage and meniscus segmentation of knee MRI with conditional generative adversarial networks. Magnetic Resonance in Medicine, 84(1), 437–449. https://doi.org/10.1002/mrm.28111

Gaj, S., Yang, M., Nakamura, K., & Li, X. (2020b). Automated Cartilage and Meniscus Segmentation of Knee MRI with Conditional Generative Adversarial Networks. Magnetic Resonance in Medicine, 84(1), 437–449. https://doi.org/10.1002/mrm.28111

Güngör, E., Vehbi, H., Cansın, A., & Ertan, M. B. (2024). Achieving high accuracy in meniscus tear detection using advanced deep learning models with a relatively small data set. Knee Surgery, Sports Traumatology, Arthroscopy. https://doi.org/10.1002/ksa.12369

Harman, F., Selver, M. A., Baris, M. M., Canturk, A., & Oksuz, I. (2023). Deep Learning-Based Meniscus Tear Detection From Accelerated MRI. IEEE Access, 11, 144349–144363. https://doi.org/10.1109/ACCESS.2023.3344456

Hung, T. N. K., Vy, V. P. T., Tri, N. M., Hoang, L. N., Tuan, L. Van, Ho, Q. T., … Kang, J. (2023). Automatic Detection of Meniscus Tears Using Backbone Convolutional Neural Networks on Knee MRI. Journal of Magnetic Resonance Imaging, 57(3), 740–749. https://doi.org/10.1002/jmri.28284

Jocher, G., Chaurasia, A., & Qiu, J. (2023). Yolov8 by Ultralytics.

Li, X., Sun, Y., Jiao, J., Wu, H., Yang, C., & Yang, X. (2021a). Automatic Discoid Lateral Meniscus Diagnosis from Radiographs Based on Image Processing Tools and Machine Learning. Journal of Healthcare Engineering, 2021, 1–7. https://doi.org/10.1155/2021/6662664

Li, X., Sun, Y., Jiao, J., Wu, H., Yang, C., & Yang, X. (2021b). Automatic Discoid Lateral Meniscus Diagnosis from Radiographs Based on Image Processing Tools and Machine Learning. Journal of Healthcare Engineering, 2021, 1–7. https://doi.org/10.1155/2021/6662664

Ma, Y., Qin, Y., Liang, C., Li, X., Li, M., Wang, R., … Jiang, Y. (2023). Visual Cascaded-Progressive Convolutional Neural Network (C-PCNN) for Diagnosis of Meniscus Injury. Diagnostics, 13(12), 2049–2062. https://doi.org/10.3390/diagnostics13122049

Makris, E. A., Hadidi, P., & Athanasiou, K. A. (2011). The knee meniscus: Structure–function, pathophysiology, current repair techniques, and prospects for regeneration. Biomaterials, 32(30), 7411–7431. https://doi.org/10.1016/j.biomaterials.2011.06.037

Ölmez, E., Akdoğan, V., Korkmaz, M., & Er, O. (2020). Automatic Segmentation of Meniscus in Multispectral MRI Using Regions with Convolutional Neural Network (R-CNN). Journal of Digital Imaging, 33(4), 916–929. https://doi.org/10.1007/s10278-020-00329-x

Ozeki, N., Seil, R., Krych, A. J., & Koga, H. (2021). Surgical treatment of complex meniscus tear and disease: state of the art. Journal of ISAKOS, 6(1), 35–45. https://doi.org/10.1136/jisakos-2019-000380

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). You Only Look Once: Unified, Real-Time Object Detection.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779–788.

Roblot, V., Giret, Y., Bou Antoun, M., Morillot, C., Chassin, X., Cotten, A., … Fournier, L. (2019). Artificial Intelligence to Diagnose Meniscus Tears on MRI. Diagnostic and Interventional Imaging, 100(4), 243–249. https://doi.org/10.1016/j.diii.2019.02.007

Sapkota, R., Qureshi, R., Calero, M. F., Badjugar, C., Nepal, U., Poulose, A., … Karkee, M. (2024). YOLOv10 to Its Genesis: A Decadal and Comprehensive Review of The You Only Look Once (YOLO) Series.

Saygili, A., & Albayrak, S. (2017a). Meniscus segmentation and tear detection in the knee MR images by fuzzy c-means method. 2017 25th Signal Processing and Communications Applications Conference (SIU), 1–4.

Saygili, A., & Albayrak, S. (2020). Knee Meniscus Segmentation and Tear Detection from MRI: A Review. Current Medical Imaging Formerly Current Medical Imaging Reviews, 16(1), 2–15. https://doi.org/10.2174/1573405614666181017122109

Saygili, A., & Albayrak, S. (2017b). A new computer-based approach for fully automated segmentation of knee meniscus from magnetic resonance images. Biocybernetics and Biomedical Engineering, 37(3), 432–442. https://doi.org/10.1016/j.bbe.2017.04.008

Sohan, M., Sai Ram, T., & Rami Reddy, Ch. V. (2024). A Review on YOLOv8 and Its Advancements. In Data Intelligence and Cognitive Informatics (pp. 529–545). https://doi.org/10.1007/978-981-99-7962-2_39

Su, Y., Cheng, B., Conference, Y. C.-2023 I. 6th I., & 2023, undefined. (n.d.). Detection and Recognition of Traditional Chinese Medicine Slice Based on YOLOv8. Ieeexplore.Ieee.OrgY Su, B Cheng, Y Cai2023 IEEE 6th International Conference on Electronic Information, 2023•ieeexplore.Ieee.Org. Retrieved from https://ieeexplore.ieee.org/abstract/document/10245026/

Terven, J., Córdova-Esparza, D.-M., & Romero-González, J.-A. (2023). A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. Machine Learning and Knowledge Extraction, 5(4), 1680–1716. https://doi.org/10.3390/make5040083

Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOv10: Real-Time End-to-End Object Detection.

Wang, C.-Y., & Liao, H.-Y. M. (2024). YOLOv1 to YOLOv10: The fastest and most accurate real-time object detection systems.

Xiongfeng, T., Yingzhi, L., Xianyue, S., Meng, H., Bo, C., Deming, G., & Yanguo, Q. (2022). Automated detection of knee cystic lesions on magnetic resonance imaging using deep learning. Frontiers in Medicine, 9. https://doi.org/10.3389/fmed.2022.928642

Ying, M., Wang, Y., Yang, K., Wang, H., & Liu, X. (2024). A Deep Learning Knowledge Distillation Framework using Knee MRI and Arthroscopy Data for Meniscus Tear Detection. Frontiers in Bioengineering and Biotechnology, 11. https://doi.org/10.3389/fbioe.2023.1326706

Zhao, R., Zhang, Y., Yaman, B., Lungren, M. P., & Hansen, M. S. (2021). End-to-End AI-based MRI Reconstruction and Lesion Detection Pipeline for Evaluation of Deep Learning Image Reconstruction.

Zhu, K., Chen, Y., Ouyang, X., White, G., & Agam, G. (2022). Fully RNN for knee ligament tear classification and localization in MRI scans. Electronic Imaging, 34(14), 227-1-227–6. https://doi.org/10.2352/EI.2022.34.14.COIMG-227