ORIGINAL RESEARCH

Intelligent detection of knee injury in MRI exam

Sanjana Gupta[1] · Pranav M. Pawar[1] · P. S. Tamizharasan[1]

Abstract Knee injuries are one of the most common injuries that occur, especially among athletes and older people. They are broadly classified into three main kinds of injuries—meniscal tear, anterior cruciate ligament (ACL) tear and abnormality. The best and most preferred method for determining knee injuries' severity is magnetic resonance imaging (MRI). Despite this, knee MRI interpretation is time-consuming and subject to diagnostic error and variability, resulting in many unnecessary surgeries and false-positive predictions. As a result, developing an automated system to interpret knee MRI could help clinicians prioritize patients at a higher risk and make better, more accurate diagnoses. This can be achieved with the help of deep learning methods, which should be capable of automatically learning the layers of features and must be capable of modelling the dynamic relationships between medical images and their interpretations. This paper aims to solve and go through the problem of Knee injuries detection in medical diagnosis and solve the problem by processing MRI by building a multi-model convolutional neural network (CNN) consisting of four pre-trained models—VGG16, VGG19, ResNet152V2, InceptionV3, DenseNet201 to help classify knee injuries from MRI scans into ACL tears, meniscal tears or abnormalities in the knee. The proposed model shows the average highest accuracy, 78.33% using ResNet152V2 as compared with state-of-the-art work for three-class classification of Knee injuries using three different planes of MRI scan.

Keywords Medical imaging · Knee MRI scan · Deep learning · CNN · Transfer learning

1 Introduction

Knee injuries can broadly be classified into three types of namely ACL tear, meniscal tear and abnormal knee. These injuries are so common that around 113,000 ACL tears [1] per year occur in the United States alone, out of which almost 75,000 of those lead to reconstructive surgeries. All three injuries are prevalent among athletes and obese people. These injuries are different from each other. The anatomy of an ACL tear is shown below in Fig. 1a. ACL tears occur due to the stretching, partial tearing, or complete tear of the anterior cruciate ligament. A complete tear is the most common type of damage. Figure 1 below shows how an ACL tear looks like. This injury can only be diagnosed through MRI imaging.

The anatomy of a meniscus tear is shown below in Fig. 1b. This happens when you forcefully try and twist or rotate your knee whilst applying your entire body weight on it, leading to the meniscus tearing. If not appropriately healed, it could lead to osteoarthritis. Figure 1b also demonstrates how a meniscus tear looks like and where the injury occurs. This injury can also be diagnosed through MRI examination.

Another type of injury is an abnormal knee. Abnormal results usually occur due to some pain or injury caused to the knee's ligaments due to damage of the knee cap, bone cancer, inflammation, infection in bone, arthritis or

✉ Pranav M. Pawar
  pranav@dubai.bits-pilani.ac.in

  Sanjana Gupta
  f20180121@dubai.bits-pilani.ac.in

  P. S. Tamizharasan
  tamizharasan@dubai.bits-pilani.ac.in

[1] Department of Computer Science, Birla Institute of Technology and Science Pilani, Dubai Campus, Dubai, UAE
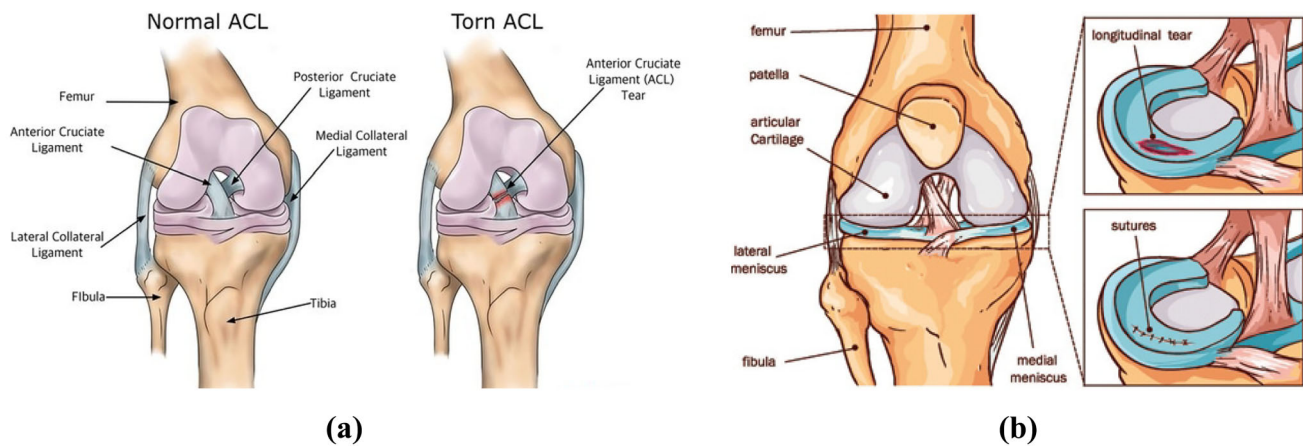
Fig. 1 Anatomy of **a** Anterior Cruciate Ligament tear [2] and Meniscus tear [3]

degeneration of the knee with age. Figure 2 below shows an MRI of an abnormal knee. MRI is used for the diagnosis of an abnormal knee.

Therefore, to diagnose all of these injuries, an MRI is used, a radiology imaging technique that creates a picture of the anatomy and physical processes of the body. MRI scans produce a set of various image slices of the organ stacked up together in a voluminous manner so that it looks like an animation of the organ, and the organs can be viewed from different levels of depth on several axes of the organ. 3D representation of MRI images exhibits intrinsic features that helps the deep neural networks to learn effectively [5].

There are three planes in a knee MRI scan, namely, sagittal, axial, and coronal, and to diagnose an injury correctly, the doctor needs to examine the knee MRI scans from all three axes to achieve a good global view. Figure 3 below shows the three planes of the MRI scan of a knee. But even MRI scans have a few significant drawbacks—it's significantly time-intensive and considering the number of knee injuries that occur. There is a need for a quicker system to help diagnose them, in subject to diagnostic error and variability. Not all the images produced of each plane

Fig. 2 MRI scan of abnormal knee [4]

have the same number of slices, leading to difficulty caused when plotting the slices of MRI and the global view that a radiologist needs to make a proper diagnosis cannot be achieved, especially when there are a lot of slices in an MRI scan. It is necessary to go over each one of them to make diagnosis.

Deep learning techniques and approaches are well equipped for modelling the intricate connections between medical images and their interpretations and give quick preliminary results following MRI exams because these deep learning approaches can intelligently learn features layer by layer, thereby improving the quality of MRI diagnosis in the absence of radiologists and specialists. Clinical experts can improve the quality and consistency of MRI interpretation by providing predictions from a deep learning model.

Hence to tackle these issues which doctors and radiologists face while diagnosing a knee MRI scan, the contribution of paper is to propose deep learning model using transfer learning. The proposed model applies five pretrained CNNs, namely VGG16, VGG19, ResNet152V2, InceptionV3, and DenseNet201 on the MRNet Dataset to classify the knee injury into the three types of knee injuries which were discussed earlier. This would help reduce the time of diagnosis per patient allowing a quick switch from one patient to another. Still, it would also give lesser false positive predictions with better accuracy rates so the patient could avoid invasive knee surgeries. All in all, this model will work efficiently at a much faster pace. Also, the proposed model in the paper is doing a three-class classification on all the three planes of a knee MRI scan. The past work done in this area does not cover every plane and measuring accuracy by considering all considered models. They perform classification on only one plane out of the three planes of an MRI scan. The average accuracies and F1 score of each model are: VGG16—75.279%, VGG19—
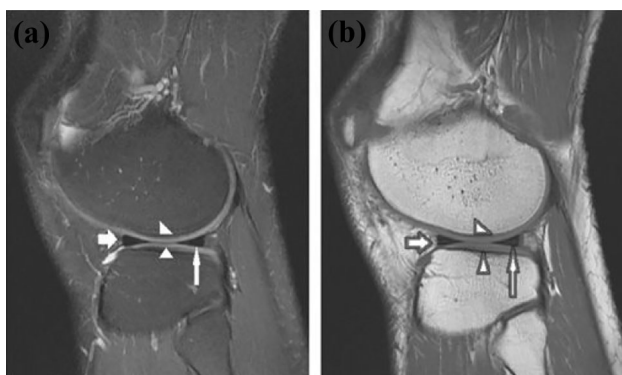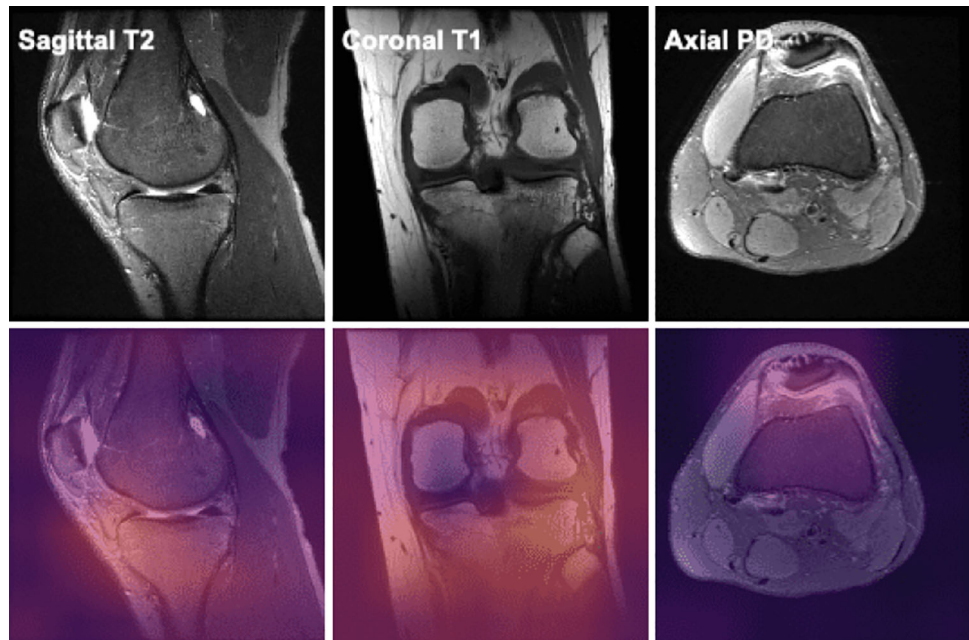
Fig. 3 The three planes of the MRI scan of a knee [6]

77.5%, ResNet152V2—78.33%, DenseNet201—71.64%, InceptionV3—71.39%. The performance of proposed model is also compared with state-of-the-art deep learning model in [7]. The model proposed in paper shows better accuracy as compared with state-of-the-art [7].

The remainder of this paper is organized as follows: Section 2 discusses the related work with a comparative review of related work; Section 3 describes the research methodology with dataset details, dataset pre-processing, and system architecture. Section 4 provide the results and discussion of developed model; Section 5 concludes the work with future work.

2 Related works

Deep learning has not been used a lot for disease detection for MRI. It is usually very challenging because it requires analysis of complex abnormalities on multiple sections of different image datasets [7–9]. In [8], the author Liu F detected the cartilage lesions within the knee joints using a deep learning approach. Using segmentation and CNNs for classification, a fully automated deep learning-based cartilage lesion detection system was developed. The sensitivity and specificity of the cartilage lesion detection system at the optimal threshold according to the Youden index were 84.1% and 85.2%, respectively, for evaluation 1 and evaluation 2 were 80.5% and 87.9%, respectively. In Pedoia et al. [7], proposes deep learning models for detecting a stage the severity of meniscus and patellofemoral cartilage lesions. The authors showed that detecting meniscal and patellar cartilage lesions using a fully

automated deep-learning pipeline is possible. The work shows accuracies of 80.74%, 78.02%, and 75.00% for normal, small, and complex large lesions, respectively. In Fang et al. [9], created three separate CNNs. They made three classification models: one that selected only those image sections containing ACL from the entire MRI dataset. Second, that isolated the region of the intercondylar notch that included the ACL on the chosen image sections to taper the range of information, and a third that created a classification model that finally evaluated the presence or absence of ACL tear on the selected image sections. The DenseNet provided the best diagnostic performance, but their model had loads of limitations as training three different models is a burden. But the biggest drawback was that this could only detect ACL tears that had full-thickness ligament injuries and not the ones with partial tears or intra-substance sprains of the ACL. The reason for it is that detecting these is much more challenging than full-thickness tears because the changes shown by these in the contour and signal intensity of the injured ligament are much subtler [10, 11]. The specificity and sensitivity achieved by this model is 96%, and AUC is 0.98. In [12], a unique CNN architecture called ELNet was implemented, which was lightweight, unlike the one discussed above. It remained lightweight because of the novel integration of multi-slice normalization and BlurPool operations. This model was robust regardless of a highly unbalanced distribution making it very helpful when the number of cases is large. It also helps locate tears on the most significant slice, but the only drawback is that it does not incorporate all the three planes of the MRI—sagittal, axial and coronal. In [13] this paper, a customized 14 layers

ResNet-14 architecture of CNN was used with the help of class balancing and data augmentation. The result was calculated using sensitivity, specificity, accuracy, precision, and F1 score of the CNN, resulting in the following AUCs: 0.980 for healthy ACL, 0.970 for partially torn ACL and 0.999 for fully torn ACL. This benefited our assertion made in the beginning that the deep learning approach can be used to detect and evaluate ACL injuries automatically. In [14], used features extraction, histogram-oriented gradient descriptor and gist descriptor techniques were on the dataset. The area under the receiver operating characteristic curve (AUC) achieved for this model was 0.894 for injury detection 0.943 for total rupture. This was achieved by combining SVM and random forest (RF). AlexNet architecture of CNN was avant-garde in works [15] related to deep learning's use in the ACL tear detection using MRI. With the help of transfer learning ImageNet [16], it extracted the features of the MRNet dataset. In Jaskaran and Sandeep [17], have used machine learning techniques such as decision trees, SVMs, k-nearest and Markov's process to detect ACL tears in sports injuries. Still, they concluded that AI techniques were giving better results as their method gave an accuracy of only 54%. In Chang et al. [18], developed a convolutional Siamese network to predict unilateral knee pain using MRI scans and achieved a great AUC value of 0.8. In Lim et al. [19], proposed using deep neural networks with scaled PCA to detect Osteoarthritis using statistical data and achieved an AUC of 78%. In Wahid et al. [20], implemented a multi-layered convolutional sparse coding to classify the MRI scan as an ACL tear but only for the coronal plane. Although it achieved a good accuracy of 85%, it was not of great use in diagnosis as it examined only one of the three planes and only one kind of injury. Effective object detection model such as You Only Look at Once (YOLO) with CNNs can be used to localize the object (area where the features reveal disease) [21].

3 Research methodology

The methodology that was applied to solve the problem of knee injury detection using MRI is shown in Fig. 4. This discussion of the section is divided into three main parts—dataset considered, data pre-processing and system architecture.

3.1 Dataset

The dataset used is the MRNet dataset [22], consisting of 1370 knee MRI exams performed at Stanford University Medical Centre. The dataset contains 1,104 (80.6%) abnormal exams, with 319 (23.3%) ACL tears and 508

Fig. 4 Overview of used methodology

(37.1%) meniscal tears. The samples have been split into a training set (1130 exams, 1088 patients) and a test set (120 exams, 113 patients) per plane, which means axial, coronal, and sagittal. The resolution of each image is 1500 × 2000 pixels.

3.2 Data pre-processing

First, the dataset is loaded, and it is divided into two parts: the test set, which has 1130 slices of the MRI scan, and the other part being the test set, which has 120 slices of the MRI scans per plane. Then through data augmentation, a function was defined that ensures that the MRI scan input has dimensions s × 256 × 256 × 3 where s is the number of slices in the MRI scan and 3 represents the number of color channels per slice and then call the function for both test and train data. This is done so because the CNN architecture typical to all the networks is MRNet, and the input of MRNet has the dimensions s × 256 × 256 × 3.

3.3 System architecture

Figure 5 shows the system architecture of the model. Around 1130 images per plane were fed as input for training the model and test the model on 120 images. The data was pre-processed with the help of data augmentation to tackle an unequal number of slices in an MRI scan. The five pre-trained networks—VGG16, VGG19, InceptionV3, ResNet152V2 and DenseNet201-are used to perform transfer learning.

For each pre-trained model, the training layer is being froze to preserve the knowledge that these networks contain, and then the output is taken. A dense layer is added with one neuron within the sigmoid activation function to get the probability. After which the output of each model is taken and added a global average pooling layer, a dropout of 0.6 is added, and with the dense layer, the probability is obtained. After doing this, it's checked how well the model predicts abnormal cases, predicting ACL cases and predicting meniscus cases at each plane by making individual predictions for each. Then all the three predictions are combined, which leads to the combining of all the probabilities. If the combined probability comes out to be greater than 1 then it is appended to 1; otherwise, it is appended to 0, where 1 stand for it being positive for the injury and 0 indicates that it is injury-free. Then the predicted probabilities are converted into a list, i.e., a 1D array, and compare with the dataset's column containing the actual probability values of the dataset, which then converted into a 1D array to calculate each pre-trained network to see which performed the best. After calculating each ensemble's accuracy and F1 scores, the accuracy graphs and training loss graphs were plotted, thereby concluding our classification ensemble.

4 Results and discussion

The model's mainly aim to perform multiclass classification on the MRI scans provided by the MRNet. The implemented model examines all three planes of an MRI scan. Results for each plane were given for diagnosis of the three kinds of injuries—ACL tear, meniscus tear, or abnormal injury. The results obtained for each model in each plane for all three kinds of injuries is shown in Table 1. The discussion on Table 1 results is as follows,

VGG19 gave the highest accuracy of 87.5% compared to all the other pre-trained networks for abnormalities in the knee. VGG19 is popular for providing higher accuracies on large scale image recognition settings and training the deeper networks by pretraining on shallower versions. ResNet152V2 gave the highest accuracy of 83.33% compared to all the other pre-trained networks for detecting ACL tears in the knee, and it also shows the highest average accuracy of 78.33%. ResNet152V2 gives good results as it uses skip connection, making it possible to train on way deeper networks. It makes it easier to copy activations from layer to layer, preserving information through every layer, and these skip connections facilitate both features constructed in shallow and deep networks. DenseNet201 gave the highest accuracy of 70% compared to all the other pre-trained networks for detecting meniscus tears. DenseNet201 gives good results because it is well suited for smaller datasets as instead of adding the activations produced by one layer to later layers, they are concatenated together.

Table 2 shows the comparative average accuracy of all implemented model and comparison of the obtained results with state-of-the art results in [7]. The average accuracy of the RestNet152V2 model considered in work is highest
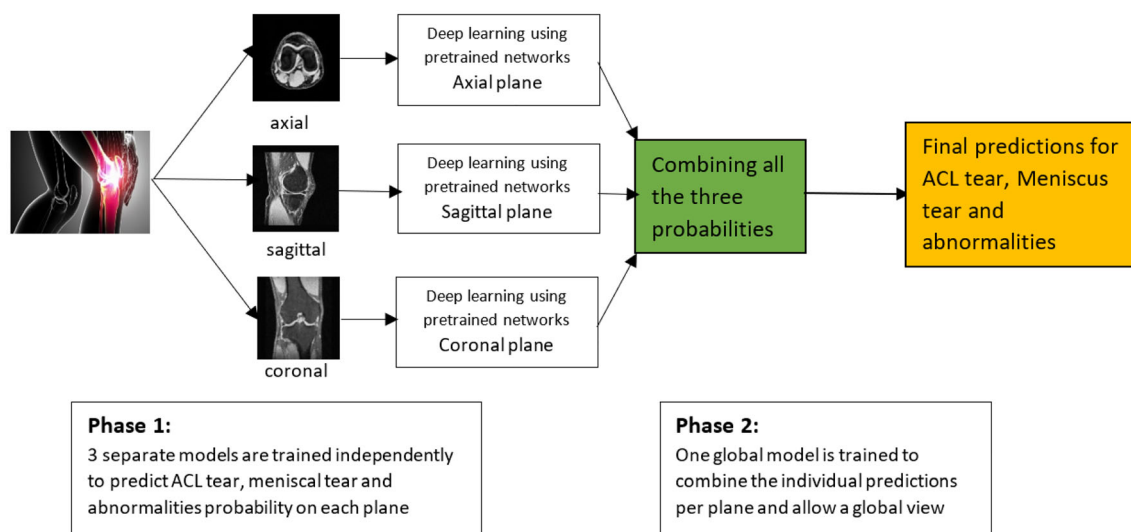

Fig. 5 System architecture

1820

Int. j. inf. tecnol. (June 2022) 14(4):1815–1821

Table 1 Accuracy and F1 score for each model with respect to each injury

Models	Abnormalities		ACL tear		Meniscus tear	
	Accuracy (%)	F1 score (%)	Accuracy (%)	F1 score (%)	Accuracy (%)	F1 score (%)
VGG16	85.0	91.0	76.67	69.5	64.167	51.68
VGG19	**87.50**	**92.46**	80.0	78.18	65.0	56.25
ResNet152V2	85.83	91.45	**83.33**	**80.39**	65.83	51.764
DenseNet201	80.83	87.83	64.1	72.25	**70.0**	**80.21**
IneptionV3	82.5	89.65	71.67	59.52	60.0	58.62

Bold value shows the highest accuracy and F1 score obtained for each class for different models.

Table 2 Comparative average accuracy

Model	Average accuracy
VGG16	75.279
VGG19	77.5
DenseNet201	71.64
IneptionV3	71.39
ResNet152V2	**78.33**
[7]	**77.92**

Bold value shows the highest accuracy and F1 score obtained for each class for different models.

compared with the state-of-the-art model in [7]. The highest accuracy achieved by the VGG19 model in the proposed work is 87.5% (for abnormalities), whereas the highest accuracy achieved by [7] is 80.74% (for abnormalities).

5 Conclusions

The main goal of this paper was to build a deep learning model that helps in the detection of the three kinds of a knee injury, diagnosing all the three planes of a knee MRI scan. Using the proposed ensemble, for the detection desired results were achieved using data augmentation, transfer learning, and fine-tuning on pre-trained networks. Five pre-trained networks were used. All the five pre-trained networks gave good enough results for all three kinds of knee injuries while examining each axis of the knee MRI scan. VGG19 performed best for detecting abnormalities in the knee. The accuracies and F1-scores achieved by VGG19 for abnormalities were higher than

those achieved by any other pre-trained networks. For abnormal injuries, achieved an accuracy of 87.5% and achieved an F1-score of 92.46%. ResNet152V2 performed best for detecting ACL tears in the knee. The accuracy and F1-score achieved by ResNet152V2 for ACL tears were higher than those achieved by any other network. For ACL tears, achieved an accuracy of 83.33% and F1-score of 80.39%. DenseNet201 performed the best for meniscus tears and gave results better than all the other pre-trained networks used. The accuracies and F1-score for meniscus tears achieved by DenseNet201 are 70% and 80.21%, respectively. The novelty of this research is that it not only focuses on detecting the type of knee injury and classifying it into the three types of injuries but it also diagnosis each plane of the MRI scan and tells which pre-trained network out of the five pre-trained networks used performs the best making this model very beneficial to radiologists for making a more accurate diagnosis and saving the radiologists a lot of diagnosis time. At the same time, it tackles the issue of an unequal number of slices of an MRI scan, making a diagnosis for radiologists easier and ultimately helps to prioritize high-risk patients while saving many from unnecessary invasive knee surgeries. The proposed technique is also compared with state-of-the-art work and it shows better accuracy than them. The future scope of this study can be extended to experiment with other pre-trained networks and developed an efficient model for knee injury detection.

Declarations

Conflict of interest The author(s) declare that they have no conflicts of interests.

References

1. Nathan S, McMurray NAB, Samuel Fischer NDS, Timothy EH (2020) Investigation of primary and second anterior cruciate ligament tears using a geographic database. Int J Sports Phys Ther 15(4):593
2. Anatomy of an anterior cruciate ligament tear (2020) https://www.jaipurjoints.com/acl-tear/. Accessed 7 Jul 2021
3. Anatomy of a meniscus tear (2020) https://www.verywellhealth.com/difference-between-meniscus-tear-and-cartilage-tear-2549642. Accessed 7 Jul 2021
4. MRI of an abnormal knee (2017) https://radiologykey.com/knee-mri/. Accessed 7 Jul 2021
5. Yeoh PSQ, Lai KW, Goh SL, Hasikin K, Hum YC, Tee YK, Dhanalakshmi S (2021) Emergence of deep learning in knee osteoarthritis diagnosis. Comput Intell Neurosci 2021:2
6. The three planes of the MRI scan of a knee (2020) https://towardsdatascience.com/deep-learning-and-medical-imaging-part-1-explore-the-mrnet-mri-dataset-of-knee-injuries-f519d063165. Accessed 7 Jul 2021
7. Pedoia V, Norman B, Mehany SN, Bucknor MD, Link TM, Majumdar S (2019) 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 49(2):400–410. https://doi.org/10.1002/jmri.26246
8. Liu F, Zhou Z, Samsonov A et al (2018) Deep learning approach for evaluating knee MR Images: achieving high diagnostic performance for cartilage lesion detection. Radiology 289(1):160–169
9. Fang L, Bochen G, Zhaoye Z, Alexey S, Humberto R, Kevin L, Ruchi S, Andrew K, John K, Ali G, Richard K (2019) Fully automated diagnosis of anterior cruciate ligament tears on knee MR images by using deep learning. Radiol Artif Intell 1:3
10. Vahey TN, Broome DR, Kayes KJ, Shelbourne KD (1991) Acute and chronic tears of the anterior cruciate ligament: differential features at MR imaging. Radiology 181(1):251–253
11. Van Dyck P, Vanhoenacker FM, Gielen JL et al (2011) Three tesla magnetic resonance imaging of the anterior cruciate ligament of the knee: can we differentiate complete from partial tears? Skelet Radiol 40(6):701–707
12. Chen-Han T, Nahum K, Raquel K (2020) Knee Injury detection using MRI with efficiently-layered network (EL Net). Proc Mach Learn Res 121:784–794
13. Javed AM, Mohd RM, Salim N et al (2021) Efficient detection of knee anterior cruciate ligament from magnetic resonance imaging using deep learning approach. Diagnostics 11:105. https://doi.org/10.3390/diagnostics11010105
14. Štajduhar I, Mamula M, Miletić D, Ünal G (2017) Semi-automated detection of anterior cruciate ligament injury from MRI. Comput Methods Programs Biomed 140:151–164. https://doi.org/10.1016/j.cmpb.2016.12.006
15. Bien N, Rajpurkar P, Ball RL, Irvin J, Park A et al (2018) Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. PLoS Med 15(11). https://doi.org/10.1371/journal.pmed.1002699
16. Russakovsky O, Deng J, Su H et al (2015) ImageNet large scale visual recognition challenge. Int J Comput Vis 115:211–252. https://doi.org/10.1007/s11263-015-0816-y
17. Jaskaran K, Sandeep S (2020) Anterior cruciate ligament knee injury detection of sports persons using machine learning techniques. Int J Innov Technol Explor Eng 9:6
18. Chang GH, Felson DT, Qiu S, Guermazi A, Capellini TD, Kolachalama VB (2020) Assessment of knee pain from MR imaging using a convolutional Siamese network. Eur Radiol 30(6):3538–3548. https://doi.org/10.1007/s00330-020-06658-3
19. Lim J, Kim J, Cheon S (2019) A deep neural network-based method for early detection of osteoarthritis using statistical data. Int J Environ Res Public Health 16(7):1281. https://doi.org/10.3390/ijerph16071281
20. Wahid A, Shah JA, Khan AU, Ullah M, Ayob MZ (2020) Multi-layered basis pursuit algorithms for classification of MR images of knee ACL tear. IEEE Access 8:205424–205435. https://doi.org/10.1109/ACCESS.2020.3037745
21. Wang Y, Wang X, Gao T, Du L, Liu W (2021) An automatic knee osteoarthritis diagnosis method based on deep learning: data from the osteoarthritis initiative. J Healthcare Eng 2021:2
22. MRNet Dataset (2019) https://stanfordmlgroup.github.io/competitions/mrnet/. Accessed 7 Jul 2021