



Research article

## Detection of the rotator cuff tears using a novel convolutional neural network from magnetic resonance image (MRI)



Mohammad Amin Esfandiari <sup>a</sup>, Mohammad Fallah Tafti <sup>a,\*</sup>,  
Nader Jafarnia Dabanloo <sup>b</sup>, Fereshteh Yousefirizi <sup>c</sup>

<sup>a</sup> Department of Biomedical Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

<sup>b</sup> Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

<sup>c</sup> Department of Integrative Oncology, BC Cancer Research Institute, Canada

---

### ARTICLE INFO

**Keywords:**

Convolutional neural network  
Deep learning  
Rotator cuff tear  
MRI

---

### ABSTRACT

The rotator cuff tear is a common situation for basketballers, handballers, or other athletes that strongly use their shoulders. This injury can be diagnosed precisely from a magnetic resonance (MR) image. In this paper, a novel deep learning-based framework is proposed to diagnose rotator cuff tear from MRI images of patients suspected of the rotator cuff tear. First, we collected 150 shoulders MRI images from two classes of rotator cuff tear patients and healthy ones with the same numbers. These images were observed by an orthopedic specialist and then tagged and used as input in the various configurations of the Convolutional Neural Network (CNN). At this stage, five different configurations of convolutional networks have been examined. Then, in the next step, the selected network with the highest accuracy is used to extract the deep features and classify the two classes of rotator cuff tear and healthy. Also, MRI images are feed to two quick pre-trained CNNs (MobileNetv2 and SqueezeNet) to compare with the proposed CNN. Finally, the evaluation is performed using the 5-fold cross-validation method. Also, a specific Graphical User Interface (GUI) was designed in the MATLAB environment for simplicity, which allows for testing by detecting the image class. The proposed CNN achieved higher accuracy than the two mentioned pre-trained CNNs. The average accuracy, precision, sensitivity, and specificity achieved by the best selected CNN configuration are equal to 92.67%, 91.13%, 91.75%, and 92.22%, respectively. The deep learning algorithm could accurately rule out significant rotator cuff tear based on shoulder MRI.

---

### 1. Introduction

Rotator cuff muscles consist of four scapulae muscles whose tendons are attached to the ball of the humerus. These muscles and tendons enable individuals to lift or rotate their arms and also help firmly fix the ball of the humerus in the shoulder socket. Tendons are sometimes partially injured, leading to shoulder tendonitis, and sometimes completely injured, resulting in tendon detachment (meaning that the tendon has torn away from the bone at the point of attachment to the shoulder) [1,2]. When one or more tendons rupture, the other tendon cannot be attached completely to the humeral head. While most shoulder tendon ruptures occur in the supraspinatus tendon, the other parts of the rotator cuff may also rupture. In many cases, shoulder tendon ruptures start with abrasion.

---

\* Corresponding author.

E-mail address: [m\\_fallah@azad.ac.ir](mailto:m_fallah@azad.ac.ir) (M. Fallah Tafti).

The tendon completely ruptures as the injury progresses. This may also happen while lifting a heavy object. The supraspinatus is among the most prevalent shoulder tendon injuries. Shoulder tendonitis often occurs in older athletes who resume their exercises and tournaments at their sports, including handball, football, and ski [3]. This is particularly rampant among the young wrestlers and weight lifters or the youth who take part in throwing or rocket events. A tentative diagnosis of shoulder injuries is conducted by a clinical examination and taking a history from the patient, and subsequently, by imaging. Two techniques, namely sonography [4–8] or Magnetic Resonance Imaging (MRI) [1,2,9], are generally used to detect the type and severity of the rotator cuff tear. Although this case is normally diagnosed by arthroscopy, the two imaging techniques above are utilized as common modalities as they are non-invasive. Doctors are especially interested in determining the shape of rupture – which plays a critical role in selecting the surgical procedure, the status of the tendons involved, the dimensions and expansion of the rupture, and the evidence of muscular atrophy of MRI report. A problem with MRI-based diagnosis that practically lets the doctor develop a two-dimensional (2D) perception of the location of the injury is that it is impossible to accurately identify the shape and severity of the injury. Since this type of injury has not been commonly treated effectively, an accurate diagnosis can help reduce the number of possible errors.

Recent developments in hardware like the graphical processing units (GPUs) and mathematical methods caused the introduction of deep learning approaches. Deep learning is a division of machine learning techniques that uses several layers to extract deep features. Li et al. [10] proposed a deep learning technique, federated deep learning to detect false data injection in smart grids. For image-based diagnosis in the past, different phases needed to have been taken: pre-processing, feature extraction, feature selection, and classification. However, nowadays, feature extraction, feature selection, and classification are carried out in an integrated and automatic manner by deep learning and a variety of convolutional neural networks (CNNs) [11–16]. CNNs are one of the strongest deep learning methods that extract deep features from images through convolutional layers. Then, these features are learned through several layers to classify multiple classes with eligible loss value via the connected layers. These networks are not sensitive to noise (in some studies, gaussian noise is added to image to generate more images [17–19] and translation. Also, they have been benefited from nonlinear units to process patterns of information in images. Therefore, these automated architecture causes high performance in classification, for this reason, these are used wide spready in almost medical diagnostic studies. For example, CNNs have been used to detect lung cancer [20–22], diagnose Alzheimer's disease [23–28] and segment tissue [29] from MRI.

Previous studies chiefly sought to detect ruptures and shoulder injuries using traditional image processing methods, some of which will be briefly discussed herein. Muto et al. [30] developed a three-dimensional (3D) MRI system to detect the shape and size of the rotator cuff tear from MRI images. In this study, 28 preoperative MRI images were given to two residents and two surgeons (with at least 100 surgery cases). These MRI images were captured using the Medical Image Processing, Analysis, and Visualization (MIPAV) software. Afterward, the doctors were provided with both the 2D and 3D models. This study aimed to examine the effects of whether the images had been 2D or 3D on the results reported. Regarding the rupture shape in the 2D images, the residents and doctors had respectively 64% and 78% accurate reports, whereas the figures increased to 85% and 92% in the 3D model. Besides, concerning the severity and thickness of injury, the number of accurate reports by the residents increased from 71% to 89% for the 2D and 3D images, respectively. Nevertheless, the figure did not show a significant difference for the experienced surgeons. In another study, Gyftopoulos et al. [2] attempted to quantitatively compare the 2D and 3D MRI images to detect the severity and type of the rotator cuff injury. They examined the type and severity of the injury in 34 patients using the 2D MRI images. The 2D images were then converted into 3D images. The results indicated the increased accuracy of the reports from 70% to above 90% upon converting the MRI images into 3D. In another research, Ward et al. [31] intended to detect the shape of rupture from 2D MRI images. They tried to interpolate the MRI images of 73 patients who had undergone arthroscopy using the Amira software to create 3D surfaces. As the 3D images formed, three factors, namely the ratios of the eigenvalues of the covariance matrix of surface points, mean and standard deviation of distances to the centroid, and 3D moment invariants, were selected as features and underwent principal component analysis (PCA). They then reached a detection accuracy of above 80% by employing the support vector machine (SVM) classifier. In another study, Fu et al. [32] used the two-stage pattern recognition method to diagnose the meniscus tear. They initially attempted to detect the meniscal position, and then they extracted 180 features, including the information on the tissue and the spectral features of the image. Then, the number of the features was reduced from 180 to 64 using the sequential floating forward selection (SFFS), and they were classified by the SVM classifier. The results indicated an accuracy of almost 91% and a time of fewer than 6 s for decision-making. In another study, Kim et al. [33] employed the active contour model (using level sets) to classify the MRI images. The accuracy of the study was reported to be 99.5%. In another research, Kang et al. [34] sought to inspect the features of the MRI images of the patients who had undergone the rotator cuff tear surgery. As some patients might have ended up suffering from a shoulder injury after a certain amount of time, their rotator cuff images were investigated to determine whether the images could help us differentiate those individuals who may have experienced the rotator cuff tear once again. Here, 50 patients were inspected, of which 20 individuals contracted rotator cuff tear a second time in less than a year. The radiologist determined the Region Of Interest (ROI) as well as the two slices demonstrating full thickness to be investigated. Among the features of this study were mean, Standard Deviation (SD), skewness, contrast, entropy, and variance. Following a statistical test in the SPSS environment, the values for the features mentioned above seemed to have been significantly higher for those patients who had been injured a second time. These features could demonstrate that the mathematical features obtained from the images could be highly detected, allowing for a second prediction of the probability of reinjury. In another study, Rutten et al. [35] compared the diagnosis of partial or complete tears by MRI and sonography. In this study, which consisted of four steps, 5216 individuals were inspected, of which the MRI images of 275 more suspicious cases were captured. The radiologist then reported the images. The ultrasound seemed to have outperformed MRI in partial tears (89% vs. 67%), whereas the MRI yielded better results than sonography in complete tears (100% vs. 95%). In general, the MR exhibited the satisfactory performance in dealing with intra-articular lesions.

The innovative aspect of the present paper is that it proposes a computer-aided diagnostic (CAD) system for the MRI images of the

shoulder. It has been designed based on an optimized convolutional neural network, which automatically diagnoses an individual's shoulder injury (rotator cuff tear) from subject's MRI images. The design of such a deep learning-based model ensures a fast, automated diagnosis and also reduces medical errors. The MRI images collected from a medical center were utilized to validate the designed CAD system. Also, two light weight pre-trained CNNs were used to compare results with the proposed CNN.

## 2. Material and methods

### 2.1. MRI data

The MR data included 75 images from patients suspected to have the rotator cuff tear and 75 images from healthy subjects. These images were recorded from Babak imaging center, Tehran, Iran, and the study was approved by the Iran University of Medical Sciences (IUMS45S438). MRI images were observed by two orthopedic specialist and then tagged as the rotator cuff tear and normal. Images were recorded using the 1.5 T magnet of MAGNETOM Avanto<sup>fit</sup> MRI apparatus. Images were recorded in DICOM format (.dcm) with the size of  $512 \times 512$  according to the resolution of the imaging device and the output provided. The informed consent was obtained from the patient(s) for the publication of their images. [Fig. 1](#) shows an example of images from the two mentioned classes.

### 2.2. Processing methods

The processing method block diagram is shown in [Fig. 2](#). As can be seen, in the first stage, shoulder MRI images were collected from two categories with the label of injury and healthy. These images were observed by an orthopedic specialist and then tagged and used as input in various configuration of CNN. At this stage, several optimal convolutional networks have been selected during five design experiments and the top network. Then, in the next step, the selected network is used to extract the characteristics and classify the two classes of shoulder injury and healthy. Finally, the evaluation is performed using the 5-fold cross validation method.

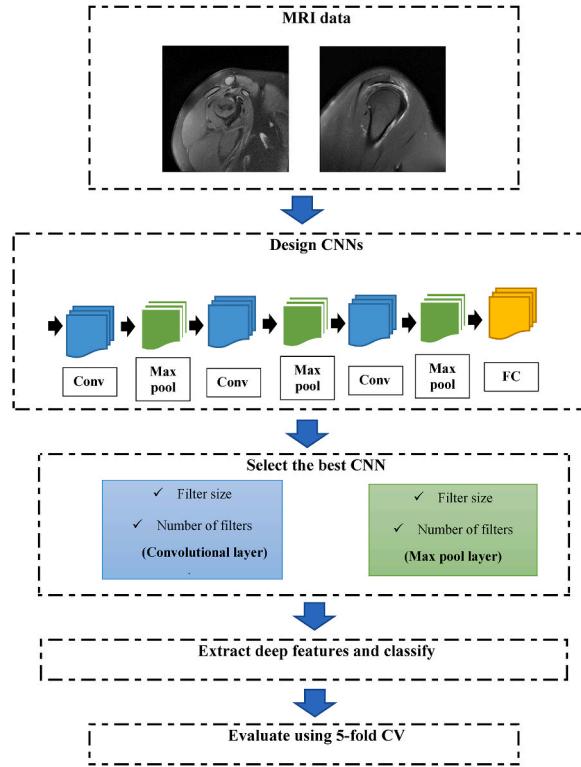
### 2.3. Convolutional neural networks

CNNs are efficient dense networks that hierarchically extract deep features through several convolutional layers and learns deeply pattern of classes. Based on deep learning surveys on medical images [11–16], CNNs are the most powerful deep learning methods and have bellow benefits:

- They are not sensitive to noise on images. As we know, medical images affect by medical imaging system noise.



**Fig. 1.** An example of the MRI images of the two classes of (up) rotator cuff tear and (down) normal.



**Fig. 2.** Block diagram of the CAD system for the diagnosis of shoulder injury from the MRI image.

- They have the ability to extract deep features to solve image classification robustly.

A convolutional neural network is generally a hierarchical neural network whose convolutional layers go alternately with pooling layers, after which a certain number of fully connected layers exist. The main purpose of a convolutional layer is to detect certain features such as edges, lines, and other image elements. For this purpose, several filters are convoluted on the image in several strides to create a feature map. The higher the number of filters in the convolutional layer, the greater the number of features detected. This layer possesses some hyperparameters, including stride, padding, and filters often selected based on empirical knowledge. The pooling layer comes in second, usually located after a convolutional layer, whose aim is to reduce the size of the feature map. To this aim, two operators are utilized: maximum and mean. After the last pooling layer, a fully connected layer is embedded, which converts the 2D feature map into a one-dimensional (1D) feature vector [11].

Each convolutional neural network involves two training phases: feedforward and backpropagation. In the first phase, the input image is applied to the network, i.e., the dot product of the input and neuronal parameters, and finally, implementing convolution operation in each layer. This is followed by calculating the network output. The output is then used to compute the network error to set network parameters, or otherwise speaking, train the network. For this purpose, the network output is compared with the actual output using the cross-entropy error function, and then the error is calculated. This is followed by the backpropagation phase based on the calculated error. In the present paper, network weights are updated by means of the adaptive moment estimation (ADAM) algorithm [36]. This algorithm updates parameters by utilizing a new momentum term and takes the moving and element-wise (element-by-element) averages of the gradients and squares of the parameters [37].

$$m_l = \beta_1 m_{l-1} + (1 - \beta_1) \nabla E(\theta_l) \quad (1)$$

$$v_l = \beta_2 m_{l-1} + (1 - \beta_2) [\nabla E(\theta_l)]^2 \quad (2)$$

where  $l$ ,  $\theta$ , and  $E(\theta)$  are the number of iterations, the parameter vector, and the loss function, respectively. Besides,  $m$  and  $v$  denote the first and second moments, respectively, and  $\beta_1$  and  $\beta_2$  are damping ratios. In moment estimation, moving averages are adaptively used to update network parameters according to Eq (3) [37]:

$$\theta_{l+1} = \theta_l - \frac{\alpha m_l}{\sqrt{v_l} + \epsilon} \quad (3)$$

Once the parameters are updated, the feedforward phase begins. Network training will stop after an appropriate number of these phases are taken.

## 2.4. The architecture of the proposed CNN

In the present paper, five experiments were designed to reach the best CNN architecture. The configuration was also designed based on trial and error. That is, the arrangement and number of convolutional layers and their hyperparameters and the number of padding and fully connected layers were examined during the five experiments to reach the highest accuracy. To ensure an identical experimental setup, the input MRI images into the networks were randomly selected. To put it another way, the training and validation sets were selected at random. The first set was used to train neural networks with multiple configurations and the second set to validate them. Similar or mirror padding was employed in such networks. In all experiments, ten training epochs were assumed. In the first experiment, a convolutional neural network with two convolutional layers with 8 and 32  $3 \times 3$  filters was designed. Given the low number of convolutional layers, small-size filter or window dimensions should usually be used to help further examine the details. In the second experiment, three convolutional layers with 8, 32, and 64  $3 \times 3$  filters were assumed. Here, the purpose was to explore the effect of the increased number of filters on shoulder injury diagnosis. In the third experiment, as better results were yielded, the fourth convolutional layer with a larger number of windows (i.e., 128) was considered. In the fourth experiment, as the accuracy decreased, a convolutional layer with a fewer number of filters (i.e., 4) was examined. Finally, in the fifth experiment, another convolutional layer with a higher number of filters (i.e., 16) was explored (as the accuracy decreased in the last two experiments, the subsequent experiment was not carried out). In all structures, a max pool layer was embedded after each convolutional layer. [Table 1](#) summarizes the details of the convolutional layers in the five experiments.

## 2.5. Pre-trained CNNs

Pre-trained CNNs have specific architecture that are trained on the huge ImageNet database and provide weights for new problems. In this study, two quick and small pre-trained CNN, SqueezeNet [38] and MobileNetv2 [39] were applied to MRI images to compare with the proposed CNN in diagnosing the rotator cuff tear problem. These CNNs preserve the accuracy while have lower parameters than very deep CNNs like the DenseNet [40]. These CNNs were used previously to solve medical diagnose from MRI or CT images [41–46].

## 2.6. Statistical measures

In this paper, the evaluation has been done in two stages, the evaluation in the first stage has been done during the selection of the better convolutional network and in the second stage after the selection of the better network. In the first stage, only the criterion of accuracy has been calculated, but in the second stage, in addition to accuracy, three important criteria of precision, sensitivity and specificity have been calculated. In the second stage, in order to evaluate the performance of shoulder injury diagnosis (rotator cuff tear), MRI images were divided into two categories of training and testing with a ratio of 80 to 20%, and this division was repeated until all the images were included in the test set. So, we had 5 folds containing 30 images and each time one-fold was used for testing and 4 folds for training. Then, each time the amount of accuracy, precision, sensitivity and specificity of the classification of 2 classes of rotator cuff tear and healthy is calculated according to Eq (4) to Eq (7) [47].

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (4)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (7)$$

If it is assumed that rotator cuff tear is the positive class, and normal is the negative class. Therefore, true positive (TP) is the number of rotator cuff tear that are correctly predicted, and the false positive (FP) is the number of rotator cuff tear that are predicted wrongly. True negative (TN) is the number of correctly predicted normal, and false-negative (FN) is the number of incorrectly predicted normal.

**Table 1**  
Details of the design of the convolutional neural network in the five experiments.

Experiment number	Number of the convolutional layer	Number of Filters
1	2	8, 32
2	3	8, 32, 64
3	4	8, 32, 64, 128
4	5	8, 32, 64, 128, 4
5	6	8, 32, 64, 128, 4, 16

### 3. Results and discussion

The results of the convolutional neural network design in the five experiments, as discussed in Section 2.4, were initially demonstrated to help determine the superior network in Section 3.1. The selected convolutional neural network was then employed to explore shoulder injury (rotator cuff tear), the results of which on 150 MRI images are presented in Section 3.2. To carry out the experiments, a laptop for home use equipped with an Intel Core™ i5 and 8 GB RAM with Windows 10 OS was used. Moreover, the proposed algorithm was implemented in MATLAB 2020 (version b).

#### 3.1. The design of the convolutional neural network

A convolutional neural network was designed based on trial and error according to maximum accuracy in diagnosing shoulder injury from MRI images during the five experiments listed above. On the other hand, the number of convolutional layers and their filters was evaluated by trial and error. The size of the convolutional layer filters was empirically assumed to be  $3 \times 3$  to reduce the computational cost. The size of the filters in the max pool layer was also empirically considered to be  $3 \times 3$ , with the stride and padding assumed to be 2 and 0, respectively. The ADAM training algorithm was employed to update the network parameters. The hyperparameters were selected with try and error. The optimum learning rate had been found such that it did not be very low or very high. If the learning rate is very low, it will slowly reach the best loss value or maybe overfitting can occur, and if it is very high the training will have diverged. The number of epochs, mini-batches, the learning rate, and the damping factor were selected to be 10, 32, 0.0001, and 0.99, respectively.

Fig. 3 presents the results of the first experiment on a given network with two convolutional layers (8 and 32  $3 \times 3$  filters, respectively). The filter size was selected as inspired by the convolutional layers of pre-trained networks such as ResNet. As indicated, the accuracy diagram for the training set has converged to an accuracy of approximately 90% after ten epochs. Finally, as shown, the neural network with the specified architecture for the validation set, which was not used for training, has reached an accuracy of 83.33%.

The results of the second experiment under similar conditions for a network with 3 convolutional layers with 8, 32, and 64  $3 \times 3$  filters are shown in Fig. 4. As shown, the addition of a third layer to this number of  $3 \times 3$  filters has proved effective, and the accuracy for the validation set has increased from 83.33% to 86.67%.

The process continued in accordance with the second experiment, and a convolutional layer with 128  $3 \times 3$  filters was added.

Accordingly, the third experiment was conducted using four convolutional layers with 8, 32, 64, and 128 [ $3 \times 3$ ] filters. The results of the experiment are presented in Fig. 5. These results reveal that the addition of a layer with 128  $3 \times 3$  filters led to a decrease in the accuracy value (76.67%).

Given the large number of filters (i.e., 128), fewer filters were added. The fourth experiment was carried out by adding 4 filters of the same size, the results of which are depicted in Fig. 6. As reflected, the addition of a layer with 4 filters has proved effectiveness, leading to an improvement in the accuracy of the validation set to 90%. The experiment was repeated to ensure that we have reached

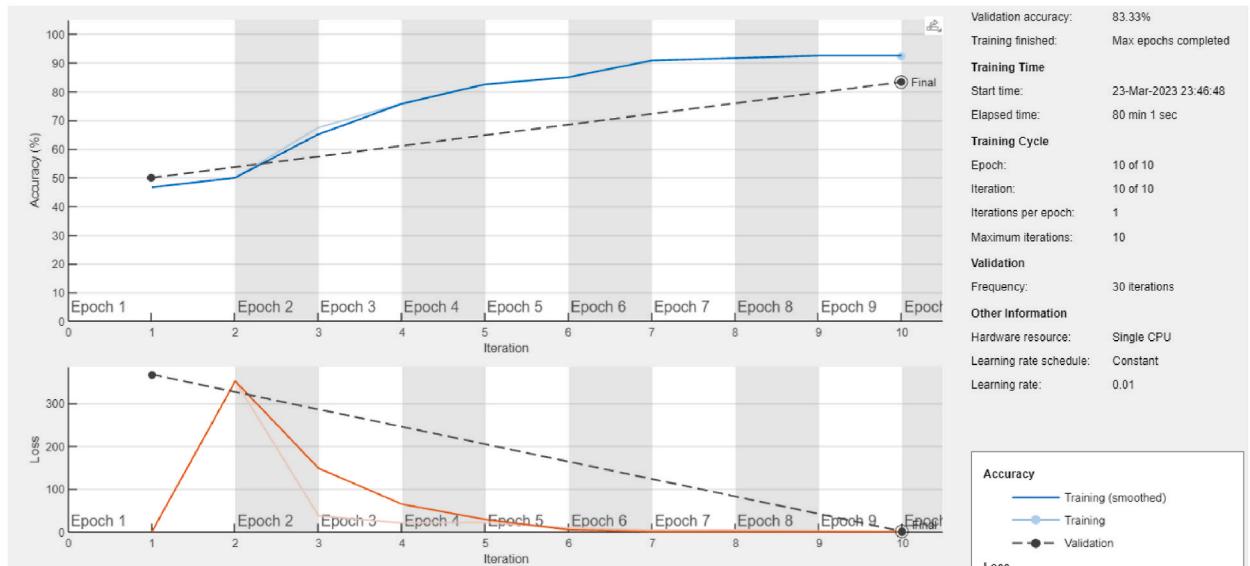


Fig. 3. The results of the first experiment on a given network with two convolutional layers (8 and 32  $3 \times 3$  filters, respectively). Up panel: the accuracy curve in 10 epochs; Down panel: the loss function curve in those epochs. Here, the blue (up panel) and red (down panel) lines denote accuracy and loss values in training, respectively. Besides, the dotted lines in both panels demonstrate the results of the validation set. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

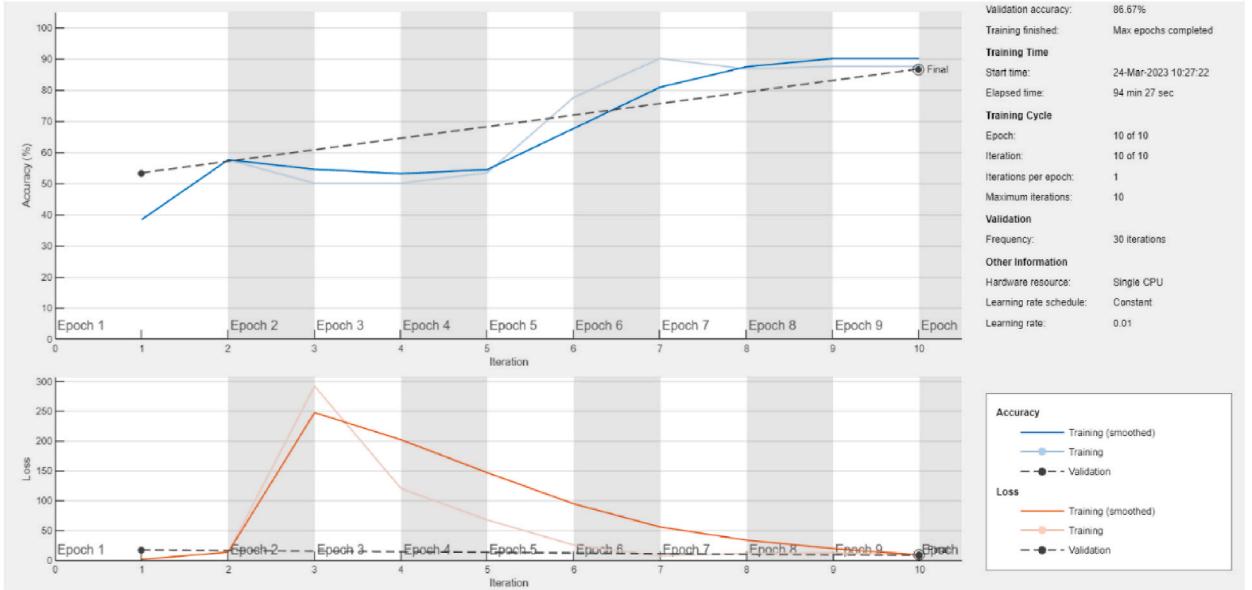


Fig. 4. The results of the second experiment for a network with 3 convolutional layers with 8, 32, and 64  $3 \times 3$  filters, respectively.

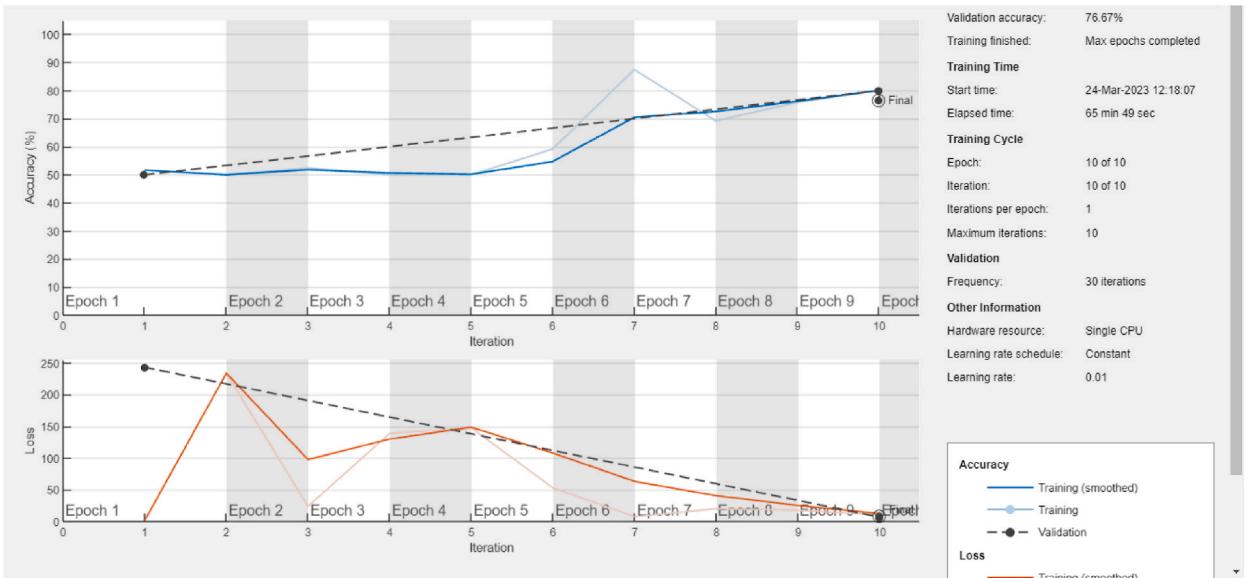


Fig. 5. The results of the third experiment for a network with 4 convolutional layers with 8, 32, 64, and 128  $3 \times 3$  filters, respectively.

the best configuration. In the fifth experiment, 16 filters of equal size were added. With this layer, the number of convolutional layers amounted to 6, and the experiment was repeated under the same conditions. The result of this experiment is displayed in Fig. 7, indicating that increasing the number of layers to 5 and placing 16  $3 \times 3$  filters following the fourth experiment resulted in an increase in the network accuracy, which reached 90% for the validation set.

Fig. 8 shows the selected structure of CNN with six convolutional layers, four max pool layer and one fully connected layer. Other layers such as batch normalization and ReLu are showed. This structure obtained the highest accuracy of 96.67%.

### 3.2. Rotator cuff tear diagnosis from MRI images

As can be seen in Fig. 2, we aimed to design a system for diagnosing the severity of shoulder injury from MRI images. For this purpose, 74 images of the shoulder injury and 75 healthy images were validated during a given process. Of these, 80% were used to train the selected convolutional neural network and 20% for validation. Accordingly, 15 images belonging to each class were examined

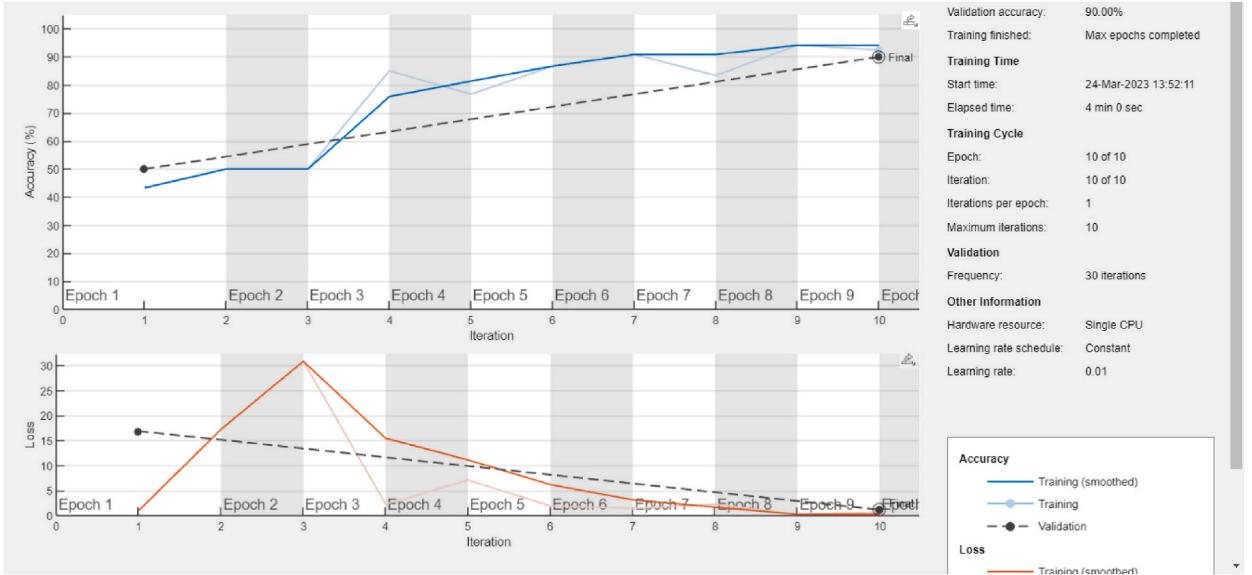


Fig. 6. The results of the fourth experiment for a network with 5 convolutional layers with 8, 32, 64, 128, and 4  $3 \times 3$  filters, respectively.

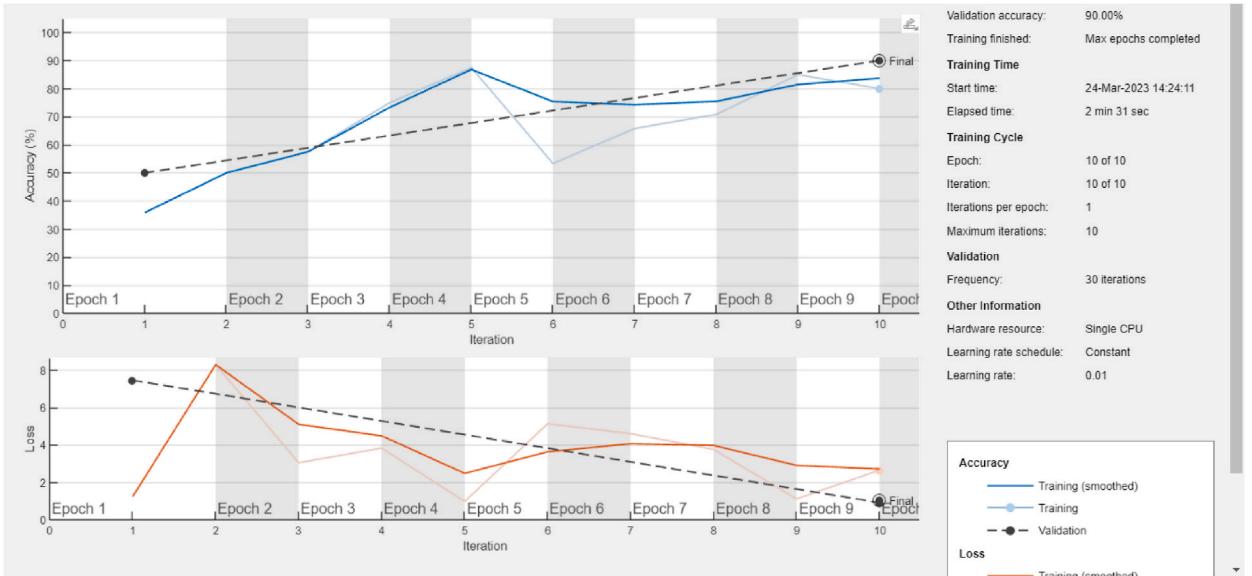


Fig. 7. The results of the fifth experiment for a network with 6 convolutional layers with 8, 32, 64, 128, 4, and 16  $3 \times 3$  filters, respectively.

during five stages for validation. Fig. 9 illustrates the process of training the selected convolutional neural network (Fig. 8) on the second epoch of MRI images with the highest accuracy on the test set (96.67%).

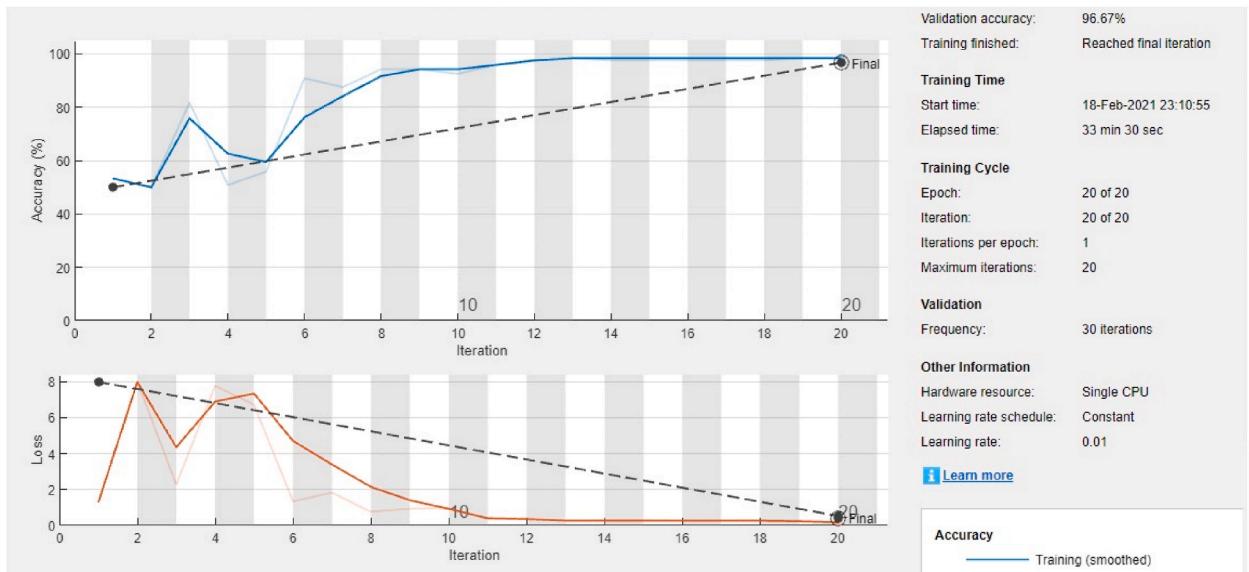
Table 2 lists the results of the rotator cuff tear diagnosis from MRI images using the selected CNN structure. As observed, the average accuracy, precision, sensitivity and specificity on 5-folds achieved using the selected CNN structure were 92.67%, 91.13%, 91.75% and 92.22%, respectively. Also, you can observe that, the two pre-trained CNNs achieved lower accuracies than the proposed method (91.25% for MobileNetv2 and 89.76% for SqueezeNet).

Finally, a specific Graphical User Interface (GUI) was designed in the MATLAB environment for simplicity, which allows for testing by detecting the image class. Moreover, the technique of random image selection of the training and validation sets were used for training. The designed GUI is illustrated in Fig. 10.

The diagnosis of shoulder injury from MRI images in the domain of Artificial Intelligence (AI) techniques is among medical engineering problems. Medical image processing for diagnostic purposes in various problems has long been the subject of much research on AI and medical engineering [12,13]. This is a classification problem in which the system uses input images to automatically identify



**Fig. 8.** The architecture of the superior convolutional neural network.



**Fig. 9.** The process of training for shoulder injury diagnosis with the selected convolutional neural network on the second epoch of shoulder MRI images.

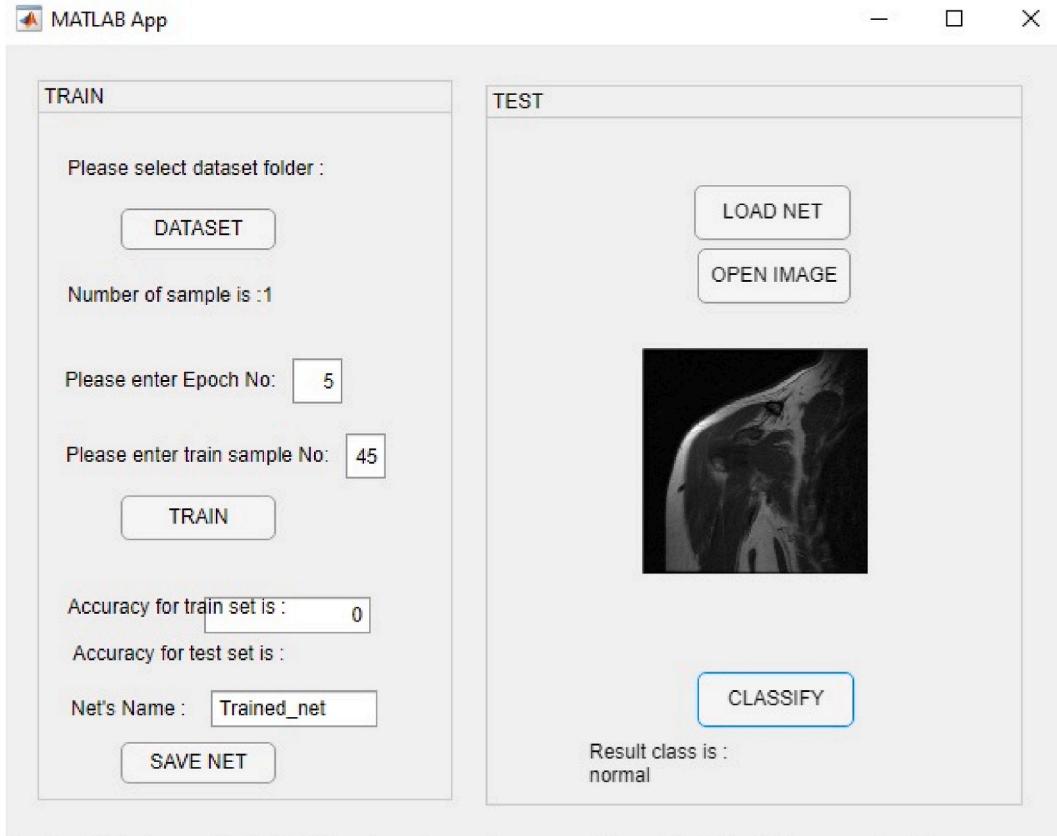
if there is a problem and sometimes determine the type of the problem. This paper discussed how shoulder injury is diagnosed. This problem includes two classes: healthy and shoulder injury. The purpose was to design a CAD system to automatically diagnose the injury from the input MRI images.

One of the limitations of this study was the number of MRI images, we had not access to higher MRI images. This is a common issue in biomedical area. Another is the processing unit, we had CPU instead of GPU, therefore, it takes more time to run CNNs on MR

**Table 2**

The results of the rotator cuff tear diagnosis from MRI images using the proposed CNN and pre-trained CNNs.

Method	Fold	Accuracy	Precision	Sensitivity	Specificity
The proposed CNN	1	83.33%	82.56%	82.34%	83.21%
	2	96.67%	96.23%	95.57%	96.24%
	3	96.67%	96.12%	95.45%	96.32%
	4	93.33%	92.58%	92.53%	92.76%
	5	93.33%	93.17%	92.84%	92.56%
	Avg ± std	92.67% ± 1.93	91.13% ± 1.92	91.75% ± 1.64	92.22% ± 5.31
MobileNetv2	Avg ± std	91.25% ± 3.25	90.45% ± 2.57	91.78% ± 3.10	91.27% ± 2.65
SqueezeNet	Avg ± std	89.76% ± 3.54	90.45% ± 3.34	89.12% ± 3.02	89.65% ± 2.60



**Fig. 10.** The design of a GUI for shoulder injury diagnosis from MRI images.

images. Also, high complexity of designing an effective CNN and computational cost to diagnose the rotator cuff tear from MRI images. Advantages of this study are

- design an automatic GUI to diagnose the rotator cuff tear, this case can be generalized to other medical decision problems.
- design a shallow and quick CNN comparable to deep and quick pre-trained CNNs

#### 4. Conclusion

This study attempted to employ the collected set of images to automatically diagnose shoulder injury (rotator cuff tear) from MRI images using a GUI based on the powerful and robust category of deep learning methods i.e., convolutional neural networks. The proposed CNN could diagnose shoulder injury from MRI images with acceptable accuracy (i.e., 92.67%). This result was approximately higher than the two light-weight pre-trained CNNs (SqueezeNet and MobileNetv2) that was used to compare the performance. The proposed approach and MRI images together with the experiments conducted could effectively pave the way for future research in this domain and other similar domains. The processing method that was proposed in this manuscript has great value to be used on large scale for clinical modules in MRI apparatus, especially for evaluate athletes. CNNs have proven their high performance during the

COVID-19 time in intelligence diagnosis systems and now, in this study prove another time.

In the future, we will focus on combining CNNs and long short-term memory method to increase the performance of diagnosis. Combining these two deep learning methods demonstrated high performance in diagnosis medical issues.

## Author contribution statement

Mohammad Amin Esfandiari: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Mohammad Fallah Tafti: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Nader Jafarnia Dabanloo; Fereshteh Yousefirizi: Conceived and designed the experiments; Wrote the paper.

## Data availability statement

Data will be made available on request.

## Conflicts of interest

There are no conflicts of interest.

## Funding statement

The authors did not receive support from any organization for the submitted work.

## Declaration of interest

There are no declarations of interest.

## Acknowledgement

We would like to thank South Tehran Branch, Islamic Azad University due to their support.

## References

- [1] Y. Morag, J.A. Jacobson, B. Miller, M. De Maeseneer, G. Girish, D. Jamadar, MR imaging of rotator cuff injury: what the clinician needs to know, Radiographics 26 (4) (2006) 1045–1065, <https://doi.org/10.1148/rgr.264055087>.
- [2] S. Gyftopoulos, S.L.S. Beltran, K. Gibbs, L. Jazrawi, P. Berman, J. Babb, R. Meislin, Rotator cuff tear shape characterization: a comparison of two-dimensional imaging and three-dimensional magnetic resonance reconstructions, J. Shoulder Elbow Surg. 251 (2016) 22–30, <https://doi.org/10.1016/j.jse.2015.03.028>.
- [3] A.T. Pennock, J. Dwek, E. Levy, P. Stearns, J. Manning, M.M. Dennis, A. Davis-Juarez, T. Bastrom, K.S. Taylor, Shoulder MRI abnormalities in asymptomatic little league baseball players, Orthop. J. Sports Med. 6 (2018).
- [4] B.E. Park, W.S. Jang, S.K. Yoo, Texture analysis of supraspinatus ultrasound image for computer aided diagnostic system, Health. Inform. Res. 22 (4) (2016) 299–304, <https://doi.org/10.4258/hir.2016.22.4.299>.
- [5] R.F. Chang, C.C. Lee, C.M. Lo, Computer-aided diagnosis of different rotator cuff lesions using shoulder musculoskeletal ultrasound, Ultrasound, Med. Biol. 42 (2016) 2315–2322, <https://doi.org/10.1016/j.ultrasmedbio.2016.05.016>.
- [6] R.F. Chang, C.C. Lee, C.M. Lo, Quantitative diagnosis of rotator cuff tears based on sonographic pattern recognition, PLoS One 14 (2019), e0212741, <https://doi.org/10.1371/journal.pone.0212741>.
- [7] A.P. Apostolopoulos, A. Stavros, R.K. Yallafragada, S. Khan, J. Nadjafi, T. Balfousias, S.T. Selvan, The sensitivity of magnetic resonance imaging and ultrasonography in detecting rotator cuff tears, Cureus 11 (2019), <https://doi.org/10.7759/cureus.4581>.
- [8] A. Teng, F. Liu, D. Zhou, T. He, Y. Chevalier, R.M. Klar, Effectiveness of 3-dimensional shoulder ultrasound in the diagnosis of rotator cuff tears: a meta-analysis, Medicine 97 (2018).
- [9] Y.K. Shin, K.N. Ryu, J.S. Park, W. Jin, S.Y. Park, Y.C. Yoon, Predictive factors of retear in patients with repaired rotator cuff tear on shoulder MRI, Am. J. Roentgenol. 210 (2018) 134–141, <https://doi.org/10.2214/AJR.17.17915>.
- [10] Y. Li, X. Wei, Y. Li, Z. Dong, M. Shahidehpour, Detection of false data injection attacks in smart grid: a secure federated deep learning approach, IEEE Trans. Smart Grid 13 (2022) 4862–4872, <https://doi.org/10.1109/TSG.2022.3204796>.
- [11] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, M.S. Lew, Deep learning for visual understanding: a review, Neurocomputing 187 (2016) 27–48, <https://doi.org/10.1016/j.neucom.2015.09.116>.
- [12] C. Cao, F. Liu, H. Tan, D. Song, W. Shu, W. Li, Y. Zhou, X. Bo, Z. Xie, Deep learning and its applications in biomedicine, Genom, Proteom. Bioinform. 16 (2018) 17–32, <https://doi.org/10.1016/j.gpb.2017.07.003>.
- [13] G. Litjens, Thijss Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. Van Der Laak, B. Van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88, <https://doi.org/10.1016/j.media.2017.07.005>.
- [14] M. Bakator, D. Radosav, Deep learning and medical diagnosis: a review of literature, Multimodal Technol. Interact. 2 (2018) 47, <https://doi.org/10.3390/mti2030047>.
- [15] A. İşin, C. Direkoglu, M. Şah, Review of MRI-based brain tumor image segmentation using deep learning methods, Procedia Comput. Sci. 102 (2016) 317–324, <https://doi.org/10.1016/j.procs.2016.09.407>.
- [16] O. Ronneberger, P. Fischer, T. Brox, U-net, Convolutional Networks for Biomedical Image Segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, 2015, pp. 234–241, [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [17] M. Momeny, A.A. Neshat, M.A. Hussain, S. Kia, M. Marhamati, A. Jahanbakhshi, G. Hamarneh, Learning-to-augment strategy using noisy and denoised data: improving generalizability of deep CNN for the detection of COVID-19 in X-ray images, Comput. Biol. Med. 136 (2021), 104704, <https://doi.org/10.1016/j.combiomed.2021.104704>.

- [18] D.M. Montserrat, Q. Lin, J. Allebach, E.J. Delp, Training object detection and recognition CNN models using data augmentation, *Electron. Imag.* 10 (2017) 27–36, <https://doi.org/10.2352/ISSN.2470-1173.2017.10.IMAWM-163>.
- [19] Z. Wang, J. Yang, H. Jiang, X. Fan, CNN training with twenty samples for crack detection via data augmentation, *Sensors* 20 (2020) 4849, <https://doi.org/10.3390/s20174849>.
- [20] W. Alakwaa, M. Nassem, A. Badr, Lung cancer detection and classification with 3D convolutional neural network (3D-CNN), *Int. J. Adv. Comput. Sci. Appl.* 8 (2017), <https://doi.org/10.14569/IJACSA.2017.080853>.
- [21] M. Nishio, O. Sugiyama, M. Yakami, S. Ueno, T. Kubo, T. Kuroda, T. Kaori, Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning, *PLoS One* 13 (2018), e0200721, <https://doi.org/10.1371/journal.pone.0200721>.
- [22] S. Liu, Y. Xie, A. Jirapatnakul, A.P. Reeves, Pulmonary nodule classification in lung cancer screening with three-dimensional convolutional neural networks, *J. Med. Imaging* 4 (2017), 041308, <https://doi.org/10.1117/1.JMI.4.4.041308>.
- [23] B. Khagi, C.G. Lee, G.R. Kwon, Alzheimer's Disease Classification from Brain MRI Based on Transfer Learning from CNN, 2018 11th Biomedical Engineering International Conference, BMEICON, 2018, pp. 1–4, <https://doi.org/10.1109/BMEICON.2018.8609974>.
- [24] W. Lin, T. Tong, Q. Gao, D. Guo, X. Du, Y. Yang, G. Guo, M. Xiao, M. Du, X. Qu, Alzheimer's disease neuroimaging initiative convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment, *Front. Neurosci.* 12 (2018) 777, <https://doi.org/10.3389/fnins.2018.00777>.
- [25] J. Islam, Y. Zhang, Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks, *Brain. Inform.* 5 (2018) 1–14, <https://doi.org/10.1186/s40708-018-0080-3>.
- [26] S. Silvia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. Filippi, Alzheimer's Disease Neuroimaging Initiative, Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks, *NeuroImage Clin.* 21 (2019), 101645, <https://doi.org/10.1016/j.nicl.2018.101645>.
- [27] S. Sarraf, D.D. DeSouza, J. Anderson, G. Tofighi, DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI, *bioRxiv* (2017), 070441, <https://doi.org/10.1101/070441>.
- [28] W. Lin, T. Tong, Q. Gao, D. Guo, X. Du, Y. Yang, G. Guo, Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment, *Front. Neurosci.* 12 (2018) 777, <https://doi.org/10.3389/fnins.2018.00777>.
- [29] F. Liu, Z. Zhou, H. Jang, A. Samsonov, G. Zhao, R. Kijowski, Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging, *Magn. Reson. Med.* 79 (2018) 2379–2391, <https://doi.org/10.1002/mrm.26841>.
- [30] T. Muto, H. Inui, H. Tanaka, Development of three-dimensional rotator cuff tendon magnetic resonance imaging system, 2325967117S00367, *Orthop. J. Sports. Med.* 5 (2017).
- [31] A.D. Ward, G. Hamarneh, M. E. Schweitzer, Anatomical Shape Analysis: Exploring the Relationship between Shape and Pathology.
- [32] J.C. Fu, C.C. Lin, C.N. Wang, Y.K. Ou, Computer-aided diagnosis for knee meniscus tears in magnetic resonance imaging, *J. Ind. Prod. Eng.* 30 (2013) 67–77, <https://doi.org/10.1080/10170669.2012.761285>.
- [33] S. Kim, L. Deukhee, S. Park, K.S. Oh, S.W. Chung, Y. Kim, Automatic segmentation of supraspinatus from MRI by internal shape fitting and autocorrection, *Comput. Meth. Prog. Biomed.* 140 (2017) 165–174, <https://doi.org/10.1016/j.cmpb.2016.12.008>.
- [34] Y. Kang, G.Y. Lee, J.W. Lee, E. Lee, B. Kim, S.J. Kim, J.M. Ahn, H.S. Kang, Texture analysis of torn rotator cuff on preoperative magnetic resonance arthrography as a predictor of postoperative tendon status, *Korean J. Radiol.* 18 (2017) 691–698, <https://doi.org/10.3348/kjr.2017.18.4.691>.
- [35] M.J.C.M. Rutten, G.J. Spaargaren, T. van Loon, M.C. de Waal Malefijt, L.A. Kiemeneij, G.J. Jager, Detection of rotator cuff tears: the value of MRI following ultrasound, *Eur. Radiol.* 20 (2010) 450–457, <https://doi.org/10.1007/s00330-009-1561-9>.
- [36] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, 2014, <https://doi.org/10.48550/arXiv.1412.6980> arXiv preprint arXiv:1412.6980.
- [37] Y. Wang, Z. Xiao, G. Cao, A convolutional neural network method based on Adam optimizer with power-exponential learning rate for bearing fault diagnosis, *J. Vibroeng.* 24 (2022) 666–678, <https://doi.org/10.21595/jve.2022.22271>.
- [38] F.N. Iandola, S. Han, M.W. Moskewicz, K. Ashraf, W.J. Dally, Kurt Keutzer, SqueezeNet, AlexNet-level Accuracy with 50x Fewer Parameters And < 0.5 MB Model Size, 2016, <https://doi.org/10.48550/arXiv.1602.07360> arXiv preprint arXiv:1602.07360.
- [39] M. Sandler, H. Andrew, M. Zhu, A. Zhmoginov, L.C. Chen, MobilenetV2: inverted residuals and linear bottlenecks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 4510–4520, <https://doi.org/10.48550/arXiv.1801.04381>.
- [40] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700–4708, <https://doi.org/10.48550/arXiv.1608.06993>.
- [41] N. Ullah, J.A. Khan, M.S. Khan, W. Khan, I. Hassan, M. Obayya, N. Negam, A.S. Salama, An effective approach to detect and identify brain tumors using transfer learning, *Appl. Sci.* 12 (2022) 5645, <https://doi.org/10.3390/app12115645>.
- [42] S.M. Ashhar, S.S. Mokri, A.A. Abd Rahni, A.B. Huddin, N. Zulkarnain, N.A. Azmi, T. Mahalethchumy, Comparison of deep learning convolutional neural network (CNN) architectures for CT lung cancer classification, *Int. J. Adv. Tech. Eng. Exp.* 8 (2021) 126, <https://doi.org/10.19101/IJATEE.2020.S1762126>.
- [43] Md Manjurul Ahsan, Redwan Nazim, Zahed Siddique, Pedro Huebner, Detection of COVID-19 patients from CT scan and chest X-ray data using modified MobileNetV2 and LIME, *Healthcare* 9 (2021) 1099, <https://doi.org/10.3390/healthcare9091099>.
- [44] S.Y. Lu, S.H. Wang, Y.D. Zhang, A classification method for brain MRI via MobileNet and feedforward network with random weights, *Pattern, Recog. Lett.* 140 (2020) 252–260, <https://doi.org/10.1016/j.patrec.2020.10.017>.
- [45] M.M. Ahsan, K.D. Gupta, M.M. Islam, S. Sen, M. Rahman, M.S. Hossain, Study of Different Deep Learning Approach with Explainable AI for Screening Patients with COVID-19 Symptoms: Using CT Scan and Chest X-Ray Image Dataset, 2020, <https://doi.org/10.3390/make2040027> arXiv preprint arXiv:2007.12525.
- [46] H.H.N. Alrashedy, A.F. Almansour, D.M. Ibrahim, M.A.A. Hammoudeh, BrainGAN: brain MRI image generation and classification framework using GAN architectures and CNN models, *Sensors* 22 (2022) 4297, <https://doi.org/10.3390/s22114297>.
- [47] M. Sokolova, L. Guy, A systematic analysis of performance measures for classification tasks, *Inf. Process. Manag.* 45 (2009) 427–437, <https://doi.org/10.1016/j.ipm.2009.03.002>.