IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 20, 2022, accepted March 2, 2022, date of publication March 10, 2022, date of current version March 28, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3158682

Deep Learning-Based Disk Herniation Computer Aided Diagnosis System From MRI Axial Scans

MOHAMMAD ALSMIRAT[1,2], (Member, IEEE), NUSAIBA AL-MNAYYIS[2], MAHMOUD AL-AYYOUB[2], AND ASMA'A AL-MNAYYIS[3]
[1]Department of Computer Science, University of Sharjah, Sharjah, United Arab Emirates
[2]Department of Computer Science, Jordan University of Science and Technology, Irbid 22110, Jordan
[3]Department of Clinical Sciences, Yarmouk University, Irbid 21163, Jordan

Corresponding author: Mohammad Alsmirat (malsmirat@sharjah.ac.ae)

ABSTRACT Computer-aided diagnosis (CAD) systems have been the focus of many researchers in both computer and medical fields. In this paper, we build two convolutional neural network (CNN) based CAD systems for diagnosing lumbar disk herniation from Magnetic Resonance Imaging (MRI) axial scans. The first one is a disk herniation detection CAD system which is a binary CAD system that determines whether the case image contains disk herniation or not. The second system is a disk herniation type classification CAD system which can determine the type of the disk herniation in the image if one exists. To train and test the proposed systems, an image set is built and annotated with the help of a radiologist. In order to get rid of the ''noisy'' parts of the input images and reduce their complexity, we experiment with different ROI extraction methods. The image set is also preprocessed and enlarged using augmentation techniques to make it suitable to be used with CNN. There are many novel aspects of this work. First, the problems of disk herniation detection and recognition from axial scans are not well-studied in the literature. Second, we use deep learning techniques which produces ground-breaking results in many image processing tasks, but are yet to reach their full potential with medical image processing in general. Finally, we explore the use of many innovative techniques such as data augmentation and transfer learning, which greatly improves the accuracy of our models. The results of our systems are impressive with a 95.65% accuracy for the detection problem and a 91.38% accuracy for the recognition problem.

INDEX TERMS Axial scans, augmentation, deep learning, disk herniation, transfer learning.

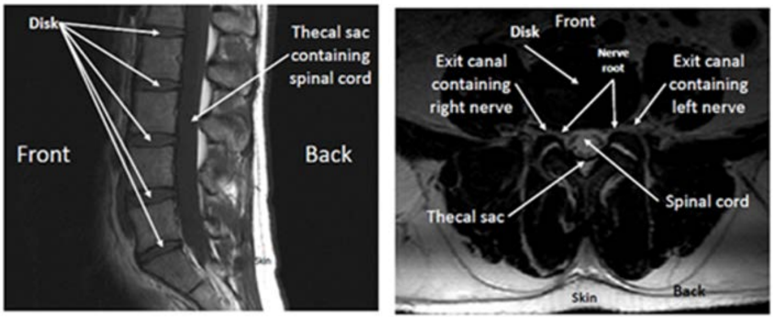
I. INTRODUCTION

Lower Back Pain (LBP) is negatively affecting the lives of people around the world due to the chronic pain it causes. LBP is considered the second most common neurological ailment as 80% of the people experience it in their lives in the United States, according to the American Academy of Orthopedic Surgeons (AAOS) [1]. One of the major causes of LBP is intervertebral disk herniation [2].

This paper focuses on a type of intervertebral disk herniation known by many names such as lumbar disk herniation, sciatica and radiculopathy. Similar to compressing the padding out of a donut, a herniated disk is a disk with its center gel-like material is compressed through fissures in its tough outer wall. In some conditions, pain is caused if this material places enough pressure on spinal nerves which give rise to nerve inflammation and swelling [3]. The reason for building a computer-aided diagnosis (CAD) system is not to replace, but to aid, physicians (in our case, radiologists) in the diagnosis process. Providing the radiologists with quick and accurate aids, such as highlighting the region of interest (ROI) in a medical image, can improve their efficiency. Furthermore, including CAD systems into the diagnosis process can improve many managerial aspects such as generating clinical reports. Finally, CAD systems provide a great platform for teaching, self-learning, and research purposes [4].

The CAD system we are considering in this work can be considered as an image classification system. Image classification is one of the core problems in computer vision and it represents the task of assigning an input image one label from a fixed set of available categories (Classes). In our work, two problems are considered: the first problem is to be able to classify the axial MRI images to abnormal (the disk herniation exists) or normal (the disk herniation not

The associate editor coordinating the review of this manuscript and approving it for publication was Vicente Alarcon-Aquino.



(a) Sagittal View (b) Axial View

FIGURE 1. Spine MRI different views.

exist). The second problem is to recognize the type of the disk herniation. For the second problem, we consider seven different types of abnormalities.

Convolutional neural networks (CNN) are deep (The term ''deep'' usually refers to the use of large number of hidden layers in the neural network) artificial neural networks (NN) that are widely used to solve different image related problems such as classification, object recognition, and many more.

In this paper, We build a CNN for each of the two problems under consideration: disk herniation detection and recognition. We experiment with various issues affecting the accuracy of our models such as ROI extraction and different values for model hyper-parameters.

Due to the limited availability of participating patients and due to the high expenses needed to collect patient images, we could only collect a small image set of usable MRI images within the limited available funds and participating personnel. To overcome the problem of the small size of our dataset and the sensitivity of deep learning techniques in general to the size of the dataset, we utilize a technique called data augmentation. Data augmentation is a technique used to increase the size of the dataset using a set of image transformation techniques before feeding them to train the system. It is very well known in the literature [5] that augmentation always increase the accuracy of CNN because it increases the size of the training data and it is currently one of the well known deep learning hyper parameters that can be tuned to enhance the deep learning accuracy [6].

To enhance the accuracy of our models, we use transfer learning which recently become popular approach in deep learning. Transfer learning is to reuse pre-trained models in neural networks that are used to accomplish different task [7]. The use of this technique in the proposed models significantly improves the results.

The results of our experimentation revealed many interesting observations. Comparing the different ROI extraction techniques shows that, although the fixed cropping technique is the most basic one, it gives the best results. It increases the accuracy of the detection system to 95.65%. On the other hand, the accuracy of the detection system when using one of the other two ROI extraction techniques (the heuristic based technique and the SIFT based technique) is 89.13%. For the type classification system, the accuracy reaches 91.38% with the use of data augmentation and transfer learning.

The rest of this paper is organized as follows. Section II discusses background knowledge necessary to understand this work. Section III presents a group of related works and defines the new value that is added through this paper. Section IV describes the dataset in detail. Section V explains the methodology and how we implement the proposed work. Section VI lists and analyzes the experiment results for the proposed work. Finally, Section VII summarizes the achieved work and lists the possible work that can be done in the future.

II. BACKGROUND INFORMATION

A. LUMBER DISK ANATOMY

There are two MRI views that are usually used to diagnose lumbar disk herniation; sagittal and axial [8]. Figure 1 shows these two views. Sagittal view is usually used to determine if the lumbar disks of different vertebrae are normal or not and can be used to locate the herniated disks (if any) [8]. The axial view scans usually provide more information about disk problems [8]. Each axial view scan contains only one disk. By examining these scans, physicians can determine if the disk is herniated or not in addition to determining the type of the disk herniation.

In this work, we are interested in diagnosing the types of disk herniation. In general, herniated (Abnormal) disk is divided into two types. [8]:

1) Diffuse: Diffuse disk bulge that can compress thecal sac and cause narrowing of both exit canals. In this type, changes occur in the basic disk structure and the disk bulges from all sides relative to the thecal sac region, which can cause narrowing or cutting (in the difficult cases) of both nerve exit canals and compressing on both nerve roots (the problem in all regions). Sometimes the disk may bulge without significant compression on the thecal sac or narrowing in the exit canals and which indicates Mild Diffuse form that could be considered within normal cases.
2) Focal: disk protrusion in a specific focal area. This form includes different types such as:

- Left: disk protrusion in the left side of spinal canal. In this type, there could be narrowing of the left exit canal with possible pressure on the spinal nerve roots located there.
- Right: disk protrusion in the right side of spinal canal. In this type, there could be narrowing of the right exit canal with possible pressure on the spinal nerve roots located there.
- Central: disk protrusion on the central region of the spinal canal, where the spinal cord or nerve roots are located. In some cases, diffuse and focal types can coexist in the same case.

B. ROI EXTRACTION TECHNIQUES

The Region of Interest (ROI) is a subset of an image identified for a particular purpose. ROI extraction sometimes is the hardest stage in a CAD system and it greatly determines its accuracy. Different ROI extraction techniques are used in this paper. The first used technique is the fixed cropping technique which, simply, crops a square that represent 40% of the the image around the middle area. After that, the cropped image is resized to 100×100 in order to be able to feed it into the CNN. This is done over 2 steps because images in the image set may have different dimensions. The second ROI extraction technique, used and tested in this work, is the extraction technique that was proposed by Alawneh *et al.* [9]. The third technique tested in the work is the ROI extraction technique proposed by Al-Ayyoub *et al.* [10] which is a SIFT based extraction technique.

III. LITERATURE REVIEW

Most of the work that can be found in literature deals with sagittal view of the disk area for diagnosis purposes [1], [11]–[24], segmentation and visualization of the disk [20], [25]–[30] and few researches used neural network in their work [31]–[33] for different purposes such as lower back pain diagnosis in general or segmentation. Table 1 summarizes some of the details of these works.

According to our knowledge, only [9], [34] try to perform diagnosis on MRI axial images and only [35] try to use CNN in the diagnosis of lumbar disk herniation from MRI axial scans.

Salehi *et al.* [35] used Alexnet based CNN to classify the cases MRI axial scan into one of 4 classes: Normal, Bulge, Protrusion and Extrusion. They achieved an accuracy of 87.75%. The works in [9] are based on only image processing techniques and they try to classify the cases into 4 classes and they have achieved around 75% accuracy. In this work, we classify cases into more classes and we have achieved better accuracy.

Al Kafri *et al.* [34] in 2017 used the Centroid Distance Function to detect the lumbar disk herniation using the axial view MRI images.This work is only an initial work as the authors assumed a perfect segmentation of the disk area which is from our experience is a very hard thing to achieve.

IV. THE DATASET

In this paper, a dataset is created consisting of axial MRI scans of human spines images acquired from King Abdullah University Hospital (KAUH), Irbid, Jordan. The original scans are stored as DICOMDIR files. DICOM viewer is used to extract the images in JPEG format after the radiologist specified the best slice to be extracted and its diagnosis. Our dataset consists of 164 512×512 images distributed among the different classes under consideration as shown in Table 2.

V. THE PROPOSED DEEP LEARNING BASED DISK HERNIATION DIAGNOSING SYSTEMS

In this paper, we have built two CNN based diagnosing systems. The first one can diagnose the cases as normal or abnormal, and we call it Disk Herniation System, while the second one can classify the type of the abnormal cases as well, and we call it Disk Herniation Type Classification System.

A. DISK HERNIATION DETECTION

For the disk herniation detection problem, we first create a convolution neural network (CNN) in MATLAB using the following steps:

1) Loading the dataset and storing it as an image datastore object. The images are used as gray images in this part of the work.
2) Dividing the dataset into training and validation datasets. This step is implemented in MATLAB using 'splitEachLabel' function which splits the datastore into two new datastores. The data is split randomly to 75% as training dataset and the remaining as the test dataset.
3) Since our dataset is not large enough, the real-time MATLAB image augmenter is used. It automatically applies different random transformations on the training images to create many new versions of each image with the same label. This way, the size of the dataset is significantly increased. We use the 'augmentedImageDatastore' function to generate batches of augmented image data with 'imageDataAugmenter' object to configure image data augmentation and specify its parameters. Since the used dataset in this paper is very sensitive to image transformations, We use the augmentation parameters provided by MATLAB with care. For all augmentation parameters, we visually inspect the impact of these parameters to decide on the suitable ranges as detailed in Table 3.
4) Design the CNN architecture. In this part of this paper, the architecture of the CNN is designed as a modified version of the AlexNet architecture as shown in Figure 2. It consists of 26 layers of seven different types distributed as five Convolutional layers, three Max Pooling layers, five Batch Normalization layers, seven ReLU layers, two Dropout layers with 0.5 probability, three fully Connected layers and a single SoftMax

TABLE 1. Summary of previous works on disk herniation.

Work Reference	Work Title	Year	Methodology	Number of Images used	Results
[15]	Lumbar disc localization and labeling with a probabilistic model on both pixel and object features	2008	Propose a two-level probabilistic model for disc localization and labeling in MRI saggital scans.	20 normal cases	96%
[16]	Abnormality detection in lumbar discs from clinical MR images with a probabilistic model	2009	Propose a probabilistic model for detection of abnormal discs from clinical T2weighted MR scans (saggital MRI scans).	80 clinical cases	91%
[17]	Computer-aided diagnosis of lumbar disc pathology from clinical lower spine MRI	2010	Propose a probabilistic model for detection of abnormal discs from clinical T2weighted MR scans (saggital MRI scans).	80 clinical cases	91%
[13]	Automatic diagnosis of lumbar disc herniation with shape and appearance features from MRI	2010	Propose a method for automatic diagnosis of lumbar disc herniation from MRI saggital scans that utilizes a Gibbs distribution for classification of discs using appearance and shape features.	33 clinical MRI cases	91%
[14]	Toward a clinical lumbar CAD: herniation diagnosis	2011	A Bayesian-based classifier with a Gibbs distribution was designed and implemented for diagnosing lumbar disc herniation from sagittal MRI scans. Each disc is segmented with a gradient vector flow active contour model (GVF-snake) to extract shape features that feed the classifier.	65 clinical cases	92.5%
[18]	Computer-aided diagnosis for lumbar mri using heterogeneous classifiers	2011	Propose a lumbar herniation diagnosis system for clinical MRI saggital scans which is based on three steps: 1) automatically label the five lumbar intervertebral discs in a sagittal MRI slice using a probabilistic model and then extract an ROI for each disc using an Active Shape Model. 2) generate relevant intensity and texture features from each disc ROI. 3) Construct five different classifiers (SVM, PCA+LDA, PCA+Naive Bayes, PCA+QDA, PCA+SVM) and combine them in a majority voting scheme.	35 clinical cases	94.85%
[19]	Lumbar spinal stenosis CAD from clinical MRM and MRI based on inter- and intra-context features with a two-level classifier	2011	Propose a method to diagnose lumbar spinal stenosis (LSS), a narrowing of the spinal canal, from magnetic resonance myelography (MRM) images.	55 cases	91.3%
[20]	Computer aided diagnosis system for lumbar spine	2011	Combines the works in [14], [16]–[19]	not provided	not provided
[25]	Disc herniation diagnosis in MRI using a CAD framework and a two-level classifier	2012	A computer-aided diagnosis framework for lumbar spine with a two-level classification scheme for disc herniation diagnosis was developed using heterogeneous classifiers: a perceptron classifier, a least mean square classifier, a support vector machine classifier, and a k-Means classifier. Each classifier makes a diagnosis based on a feature set generated from regions of interest that contain vertebrae, a disc, and the spinal cord in saggital scan. Then, an ensemble classifier makes a final decision using score values of each classifier.	70	99%
[1]	A new approach to automatic disc localization in clinical lumbar MRI: Combining machine learning with heuristics	2012	Propose a localization method which outputs a tight bounding box for each disc in an MRI saggital scan. HOG (Histogram of Oriented Gradients) features along with SVM (Support Vector Machine) as classifier are used and combined with some heuristics.	53 cases	99%
[21]	Lumbar spine disc herniation diagnosis with a joint shape model	2014	A classifier is designed to automatically detect disc Herniation using shape potentials that are extracted by jointly applying an active shape model (ASM) and a gradient vector flow snake model (GVF-snake). The ASM roughly segments the disc by the detection of a certain point distribution around the disc. Then, this point distribution is used to initialize a GVF-snake model to delineate the posterior disc segment. Then the set of shape potentials is extracted for our Gibbs-based classifier.	65 cases	93.9%

Layer, in addition to the input layer and the classification layer.

5) Training the network. Before training the network using 'trainNetwork' function, the training options must be specified. The first option we specify is choosing the solver, which is the Stochastic Gradient Descent (SGD) optimization algorithm. In each step, the solvers update the parameters by uses a subset of the data, called a mini-batch. The mini-batch size is specified using the 'MiniBatchSize' option. Each parameter update is called an iteration. A full pass through the entire dataset is called an epoch and the option 'MaxEpochs' specifies the maximum number of epochs to train for. In addition to that, there is an

TABLE 2. The classes in the disk herniation recognition part.

Class	Frequency
Left	10
Right	5
Central	7
Diffuse	35
Diffuse+Left	4
Diffuse+Right	4
Diffuse+Central	6
Normal	93

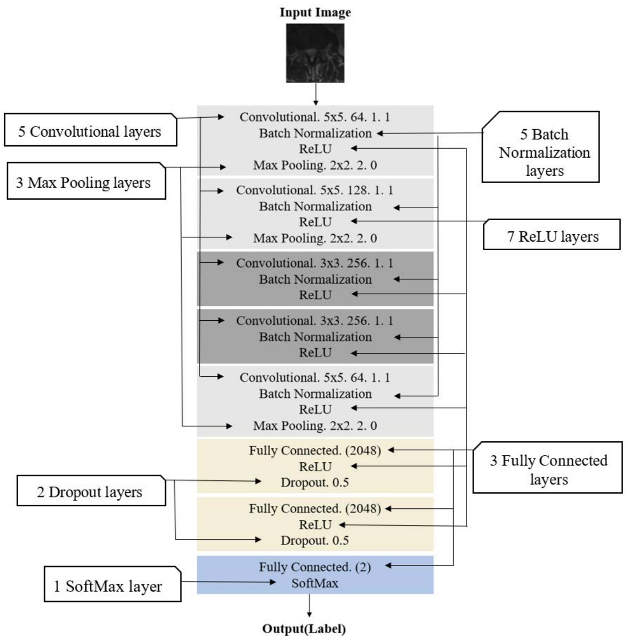


FIGURE 2. The architecture of disk herniation detection CNN.

option to specify the initial learning rate (the amount the weights are updated with), and 'Verbose' which is an indicator to display training progress information. The training of the network stops if the training reaches the maximum number of epochs or if the validation loss error value becomes greater or equal to the current value in five iterations (five is the default; you can change it using the 'ValidationPatience' option).

1) ANALYSIS OF DATA AUGMENTATION TECHNIQUES USING T-TEST

To study the impact of data augmentation techniques on our network, the network is trained and tested with and without data augmentation. After that, a paired t-test is applied to determine whether there is a statistically significant evidence that data augmentation improves the performance of the detection CNN. The CNN is trained 30 times with augmented data and 30 times without augmented data. The test accuracy is recorded for each training session. Hence, we have 30 pairs of accuracies to be used in the t-test.

The t-test stands on two hypotheses. The first one is the Null Hypothesis, which assumes that there is no difference between the means of the two groups of accuracies we have. It is represented by Equation 1. The second hypothesis is the Alternative Hypothesis that assumes there is a difference between the means of the two groups of accuracies we have [36]. It is represented by the Equation 2.

$$H0: \quad \mu_1 = \mu_2 \tag{1}$$

$$H1: \quad \mu_1 \neq \mu_2 \tag{2}$$

where μ_1 stands for the mean for the first events and μ_2 stands for the mean of the second event. To compute the t-test, the following are needed.

1) The degree of freedom (df), which represents the number of independent pieces of information that can go into the estimate of a parameter. It is equal to $n - 1$, where n is the sample size, which is 30 in our 30 case.
2) The significance level (α), which refers to the likelihood that the random sample you choose (for example, experiment accuracies) is not representative of the population. The lower the significance level, the more confident you can be in replicating your results. The most commonly used significance levels in the literature are 0.05 and 0.01 levels. In our work, α is chosen to be 0.05.
3) P-value, which is the smallest level of significant that would lead to rejection of the null hypothesis $H0$ with the given data, where if it is lower than α the $H0$ is rejected. Equation 3 illustrates how to compute this value.

$$P = 2P(t_{n-1} > |t_0|) \tag{3}$$

4) The Confidence Interval (CI) of a mean is a region within which a score may be said to fall with a certain amount of "confidence." The CI uses sample size and standard deviation to generate a lower and upper limits such that you can be 95% sure will include any sample you take from a set of data. This interval is calculated by the following equation.

$$\bar{x} = t_{\alpha/2,n-1}S/\sqrt{n} \leq \mu \leq \bar{x} + t_{\alpha/2,n-1}S/\sqrt{n} \tag{4}$$

t value, which is needed to find the probability of the significant value that must be compared with α to be able to judge the hypothesis. It is computed by the following equation.

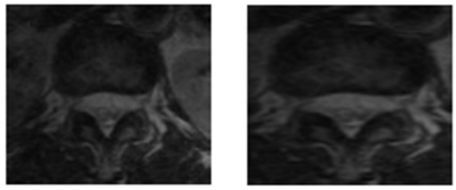
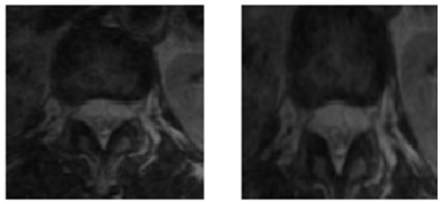
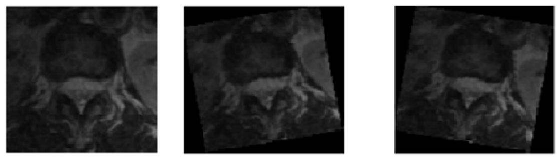
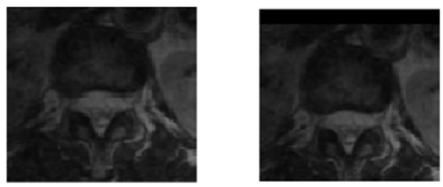
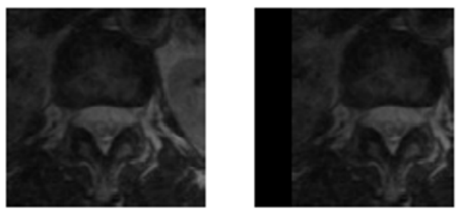
$$t = \frac{x_d - H0_{value}}{s_d/n} \tag{5}$$

where x_d is the mean of the difference of the two sample of the data, s_d is the standard deviation of the difference of the two sample of the data. Finally, Microsoft Excel was used to perform this task.

B. DISK HERNIATION TYPE CLASSIFICATION

Because the CNN used for disk herniation detection gives excellent results for detection purposes, it was the first CNN that is tested, for the disk herniation type classification, after modifying the classification layer to be able to classify the available different disk herniation classes. Unfortunately, the

TABLE 3. Summery of used MATLAB augmentation parameters.

Parameter	Impact	Range used	Example
'RandXScale'	Specifies the range of the X dimension random scale applied to the input image.	$[1, 1.5]$	 (a) Original image (b) 1.5 X scaling image
'RandYScale'	Specifies the range of the Y dimension random scale applied to the input image.	$[1, 1.5]$	 (a) Original image (b) 1.5 Y scaling image
'RandRotation'	Specifies the range, in degrees, of random rotations that are applied to the input image.	$[-10, 10]$	 (a) Original image (b) Rotate 10 degree angle (c) Rotate -10 degree angle
'RandXTranslation'	Specifies the range, in pixels, of random translation in the X dimension that is applied to the input image.	$[-10, 10]$	 (a) Original image (b) 10 pixels translation in X dimension
'RandYTranslation'	Specifies the range, in pixels, of random translation in the Y dimension that is applied to the input image.	$[-10, 10]$	 (a) Original image (b) 10 pixels translation in Y dimension

detection CNN failed to reach any reasonable accuracy level. For that reason, Many modifications are made in order to improve the results. These modifications includes changing with many parameters turnings, preprocessing of the image set, pre-augmentation on the image set that is inspired by [5], and followed the suggestions in Brownlee's blog [37]. The following are the details of the full proposed model.

To train the proposed model, we use modified dataset after performing pre-augmentation. This process is performed using rotating images around their center with a specifide angle. Each image in the "minority" classes, which are: Right, Left, Central, Diffuse+Left, Diffuse+Right and Diffuse+Central, is rotated 10 degrees to create a new image and -10 degrees to create another. Then, the resulting images are added to the data. So, the counts of the images in these classes become 15 images in the Right class, 30 in the Left class, 21 in the Central class, 12 in the Diffuse+Left class, 12 in the Diffuse+Right class and 18 in the Diffuse+Cenral class. Consequently, we have 236 in total.

The pretrained model used in the proposed model is the pretrained AlexNet model that is pretrained using the ImageNet dataset. It consists of 23 layers in addition to the image input layer and the output classification layer. Illustration of the used AlexNet is shown in Figure 3. AlexNet is trained on

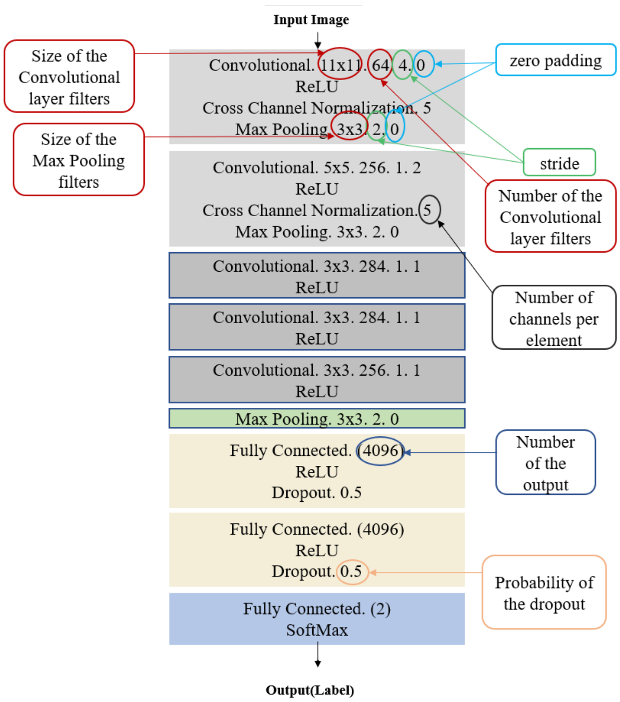


FIGURE 3. The architecture of AlexNet.

our data with an initial learning rate of 0.001, max epochs of 40, and mini-batch size of 16.

VI. ANALYSIS AND RESULTS

MATLAB is used to build the CNN and realize the proposed architecture. Python language is used to prepare the dataset.

A. DISK HERNIATION DETECTION RESULT

This section presents the results of our disk herniation detection experiments, which are divided into two parts. The first one is dedicated to show the impact of the ROI extraction techniques on the results and the second one is dedicated to show the impact of the data augmentation technique.

1) RESULTS OF USING DIFFERENT ROI EXTRACTION TECHNIQUES

After we train the detection CNN on the images extracted by fixed cropping technique, we get a test accuracy of 95.65 % and a loss error of 0.18. While training the network on the images using the heuristic based ROI extraction technique, we get 89.13% test accuracy and 0.42 loss error. Finally, the results of using the SIFT based ROI extraction technique is a test accuracy of 89.13% test accuracy with 0.37 loss error. Figure 4 shows these results.

After careful investigation in our image set, it is found that the reason behind the differences in the results of using fixed cropping and of using the other two techniques is that the fixed cropping technique produce an ROI that includes all the important regions, while the other ROI extraction techniques produce an ROI that includes only parts of the important regions for some images.

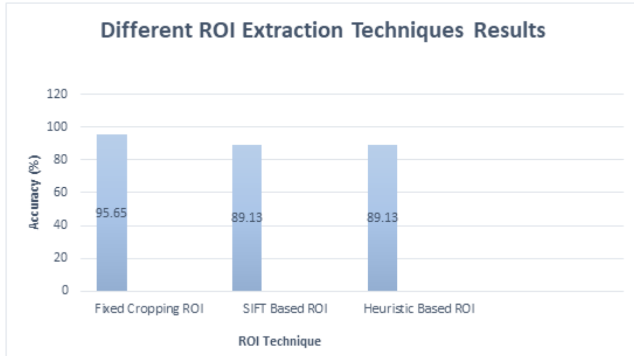


FIGURE 4. Accuracies when using different ROI extraction techniques.

TABLE 4. The results of applying the t-test.

Differences Statistics Paired				
Mean (Xd)	Std Deviation (Sd)	df	t	P
6.52	8.39	29	4.26	0.0002

	Left	Right	Central	Diffuse	Diffuse +Left	Diffuse +Right	Diffuse +Central	Normal
Left	7	0	0	0	0	0	0	0
Right	0	4	0	0	0	0	0	0
Central	0	0	5	0	0	0	0	0
Diffuse	0	0	0	9	0	0	0	0
Diffuse+Left	0	0	0	0	3	0	0	0
Deffuse+Right	0	0	0	0	1	2	0	0
Diffuse+Central	0	0	0	0	0	0	4	0
Normal	1	0	1	2	0	0	0	19

FIGURE 5. The confusion matrix of disk herniation recognition.

2) RESULTS OF THE DATA AUGMENTATION TECHNIQUE

In order to prove that data augmentation is indeed useful, we perform 30 pairs of experiments, where each pair is simply training the disk herniation detection CNN with and without augmented data. Then, we use Microsoft Excel to conduct a t-test and compute all values related to it. The results are shown in Table 4. As we can see in the previous table, the mean and the standard deviation of the difference between the two samples of the data are 8.39 and 5.60427 respectively. The t value is 4.26 and the probability of the t value is 0.0002, which is significantly smaller than 0.05. Therefore, the null hypothesis is rejected, which means that there is a huge difference between the augmented data accuracy and the original data accuracy. There is a strong evidence that the Alternative Hypotheses is acceptable in our case and the accuracies returned when using the augmented data is better than the ones returned when using the original data.

B. DISK HERNIATION TYPE CLASSIFICATION RESULTS

The best accuracy we achieve in this part from training the network architecture described in section V.B is 91.38% and a loss rate of around 0.103. The confusion matrix of the different classes is shown in Figure 5.

As shown in Figure 5, the number of images in the test set is 58 images. 54 images are classified correctly (sum of the images inside the diagonal) while only 4 images are classified incorrectly.

VII. CONCLUSION AND FUTURE WORK

In this paper, CNN models are built for the automatic diagnosis of disk herniation, by detecting it or recognizing its type. The best model for detecting disk herniation has 95.56% accuracy, while the best model for recognizing the type of disk herniation has 91.38% accuracy. Different ROI extraction techniques were tested, the results showed that although the fixed cropping ROI extraction technique is the most direct, obvious and generic approach, it gave the best results (95.65% accuracy) whereas accuracies were 89.13% for each of heuristic based and SIFT based techniques. In addition, the models discussed in this paper benefited from hot and promising techniques such as data augmentation and transfer learning. The results showed a clear advantage of using these techniques.

As the future work, we intend to increase the size of our dataset, especially, for the "minority" classes. Moreover, we intend to use unsupervised deep learning networks, such as DBN [38] and Gated autoencoder [39], to extract good features and use them with different classifiers such as SVM [40] and others. Finally, Generative Adversarial Networks (GANs) [33] seems like a very promising future direction for our work.

REFERENCES

[1] S. Ghosh, M. R. Malgireddy, V. Chaudhary, and G. Dhillon, "A new approach to automatic disc localization in clinical lumbar MRI: Combining machine learning with heuristics," in *Proc. 9th IEEE Int. Symp. Biomed. Imag. (ISBI)*, May 2012, pp. 114–117.

[2] S. Ghosh, R. S. Alomari, V. Chaudhary, and G. Dhillon, "Composite features for automatic diagnosis of intervertebral disc herniation from lumbar MRI," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 5068–5071.

[3] *Herniated Cervical Disc*. Accessed: Nov. 27, 2018. [Online]. Available: http://www.mayfieldclinic.com/PE-HCDisc.htm

[4] K. Doi, "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential," *Comput. Med. Imag. Graph.*, vol. 31, nos. 4–5, pp. 198–211, Jun. 2007.

[5] M. Ebrahim, M. Alsmirat, and M. Al-Ayyoub, "Performance study of augmentation techniques for HEp2 CNN classification," in *Proc. 9th Int. Conf. Inf. Commun. Syst. (ICICS)*, Apr. 2018, pp. 163–168.

[6] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, p. 60, Dec. 2019.

[7] J. Brownlee. (2017). *A Gentle Introduction to Transfer Learning for Deep Learning*. Accessed: Nov. 27, 2018. [Online]. Available: https://machinelearningmastery.com/transfer-learning-for-deep-learning/

[8] M. Weyreuther, C. E. Heyde, M. Westphal, J. Zierski, U. Weber, and B. Herwig, *MRI Atlas: Orthopedics and Neurosurgery, The Spine*, 2007th ed. Springer, 2006.

[9] K. Alawneh, M. Al-dwiekat, M. Alsmirat, and M. Al-Ayyoub, "Computer-aided diagnosis of lumbar disc herniation," in *Proc. 6th Int. Conf. Inf. Commun. Syst. (ICICS)*, Apr. 2015, pp. 286–291.

[10] M. Al-Ayyoub, N. Al-Mnayyis, M. A. Alsmirat, K. Alawneh, Y. Jararweh, and B. B. Gupta, "SIFT based ROI extraction for lumbar disk herniation CAD system from MRI axial scans," *J. Ambient Intell. Humanized Comput.*, to be published, doi: 10.1007/s12652-018-0750-2.

[11] R. S. Alomari, J. J. Corso, V. Chaudhary, and G. Dhillon, "Automatic diagnosis of lumbar disc herniation with shape and appearance features from MRI," *Proc. SPIE*, vol. 7624, Mar. 2010, Art. no. 76241A.

[12] R. S. Alomari, J. J. Corso, V. Chaudhary, and G. Dhillon, "Toward a clinical lumbar CAD: Herniation diagnosis," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 6, no. 1, pp. 119–126, Jan. 2011.

[13] J. J. Corso, R. S. Alomari, and V. Chaudhary, "Lumbar disc localization and labeling with a probabilistic model on both pixel and object features," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2008, pp. 202–210, doi: 10.1007/978-3-540-85988-8_25.

[14] A. Raja'S, J. J. Corso, V. Chaudhary, and G. Dhillon, "Abnormality detection in lumbar discs from clinical MR images with a probabilistic model," *Injury*, vol. 4, p. 3, Jun. 2009.

[15] R. Alomari, J. Corso, V. Chaudhary, and G. Dhillon, "Computer-aided diagnosis of lumbar disc pathology from clinical lower spine MRI," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 5, pp. 287–293, Sep. 2010.

[16] S. Ghosh, R. S. Alomari, V. Chaudhary, and G. Dhillon, "Computer-aided diagnosis for lumbar MRI using heterogeneous classifiers," in *Proc. IEEE Int. Symp. Biomed. Imag.: Nano Macro*, Mar. 2011, pp. 1179–1182.

[17] J. Koh, R. S. Alomari, V. Chaudhary, and G. Dhillon, "Lumbar spinal stenosis CAD from clinical MRM and MRI based on inter-and intra-context features with a two-level classifier," *Proc. SPIE*, vol. 7963, Mar. 2011, Art. no. 796304.

[18] R. S. Alomari, V. Chaudhary, and G. Dhillon, "Computer aided diagnosis system for lumbar spine," in *Proc. 4th Int. Symp. Appl. Sci. Biomed. Commun. Technol. (ISABEL)*, 2011, p. 145.

[19] R. S. Alomari, J. J. Corso, V. Chaudhary, and G. Dhillon, "Lumbar spine disc herniation diagnosis with a joint shape model," in *Computational Methods and Clinical Applications for Spine Imaging* (Lecture Notes in Computational Vision and Biomechanics), vol. 17, J. Yao, T. Klinder, and S. Li, Eds. Cham, Switzerland: Springer, doi: 10.1007/978-3-319-07269-2_8.

[20] M.-D. Tsai, S.-B. Jou, and M.-S. Hsieh, "A new method for lumbar herniated inter-vertebral disc diagnosis based on image analysis of transverse sections," *Computerized Med. Imag. Graph.*, vol. 26, no. 6, pp. 369–380, Dec. 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0895611102000332

[21] M. Hsieh and M. Tsai, "Three-dimensional surgery simulation system," U.S. Patent 7 121 832, Oct. 17, 2006. [Online]. Available: http://www.google.it/patents/US7121832

[22] S. K. Michopoulou, L. Costaridou, E. Panagiotopoulos, R. Speller, G. Panayiotakis, and A. Todd-Pokropek, "Atlas-based segmentation of degenerated lumbar intervertebral discs from MR images of the spine," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 9, pp. 2225–2231, Sep. 2009.

[23] J. Koh, V. Chaudhary, and G. Dhillon, "Disc herniation diagnosis in MRI using a CAD framework and a two-level classifier," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 7, no. 6, pp. 861–869, Nov. 2012, doi: 10.1007/s11548-012-0674-9.

[24] A. B. Oktay, N. B. Albayrak, and Y. S. Akgul, "Computer aided diagnosis of degenerative intervertebral disc diseases from lumbar MR images," *Computerized Med. Imag. Graph.*, vol. 38, no. 7, pp. 613–619, Oct. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0895611114000573

[25] V. C. Chetan Bhole and S. Kompalli, "Context sensitive labeling of spinal structure in mr images," *Proc. SPIE*, vol. 7260, Feb. 2009, Art. no. 72603P, doi: 10.1117/12.812191.

[26] C. Chevrefils, F. Chériet, G. Grimard, and C.-E. Aubin, "Watershed segmentation of intervertebral disk and spinal canal from MRI images," in *Image Analysis and Recognition*, M. Kamel and A. Campilho, Eds. Berlin, Germany: Springer, 2007, pp. 1017–1027.

[27] S. Ghosh, M. R. Malgireddy, V. Chaudhary, and G. Dhillon, "A supervised approach towards segmentation of clinical MRI for automatic lumbar diagnosis," in *Computational Methods and Clinical Applications for Spine Imaging*, J. Yao, T. Klinder, and S. Li, Eds. Cham, Switzerland: Springer, 2014, pp. 185–195.

[28] M. Muja and D. G. Lowe, "Scalable nearest neighbor algorithms for high dimensional data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 11, pp. 2227–2240, Nov. 2014.

[29] A. Al-Mnayyis, S. A. Alasal, M. Alsmirat, Q. B. Baker, and S. AlZu'bi, "Lumbar disk 3D modeling from limited number of MRI axial slices," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 10, no. 4, p. 4101, Aug. 2020.

[30] S. A. Alasal, M. Alsmirat, and A. Al-Mnayyis, "Improving radiologists 'and orthopedists' QoE in diagnosing lumbar disk herniation using 3D modeling," *Int. J. Elect. Comput. Eng. (2088–8708)*, vol. 11, no. 5, pp. 4336–4344, 2021.

[31] D. G. Bounds, P. J. Lloyd, B. Mathew, and G. Waddell, "A multilayer perceptron network for the diagnosis of low back pain," in *Proc. Int. Conf. Neural Netw.*, vol. 2, Jul. 1988, pp. 481–489.

[32] A. Jamaludin, T. Kadir, and A. Zisserman, "SpineNet: Automated classification and evidence visualization in spinal MRIs," *Med. Image Anal.*, vol. 41, no. 18, pp. 63–73, Oct. 2017.

[33] Z. Han, B. Wei, A. Mercado, S. Leung, and S. Li, "Spine-GAN: Semantic segmentation of multiple spinal structures," *Med. Image Anal.*, vol. 50, pp. 23–35, Dec. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S136184151830642X

[34] A. S. A. Kafri, S. Sudirman, A. J. Hussain, P. Fergus, D. Al-Jumeily, H. A. Smadi, M. Khalaf, M. Al-Jumaily, W. Al-Rashdan, and M. Bashtawi, "Detecting the disc herniation in segmented lumbar spine MR image using centroid distance function," in *Proc. 10th Int. Conf. Develop. eSystems Eng. (DeSE)*, Jun. 2017, pp. 9–13.

[35] E. Salehi, S. Khanbare, H. Yousefi, H. Sharpasand, and O. Sojoodi Sheyjani, "Deep convolutional neural networks for automated diagnosis of disc herniation on axial MRI," in *Proc. Sci. Meeting Elect.-Electron. Biomed. Eng. Comput. Sci. (EBBT)*, Apr. 2019, pp. 1–6.

[36] (2018). *Student's t-Distribution–From Wolfram Mathworld*. Accessed: Nov. 27, 2018. [Online]. Available: http://mathworld.wolfram.com/Studentst-Distribution.html

[37] J. Brownlee. (2016). *How to Improve Deep Learning Performance*. Accessed: Nov. 27, 2018. [Online]. Available: https://machinelearningmastery.com/improve-deep-learning-performance/

[38] G. E. Hinton, "Deep belief networks," *Scholarpedia*, vol. 4, no. 5, p. 5947, May 2009.

[39] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, New York, NY, USA, 2008, pp. 1096–1103, doi: 10.1145/1390156.1390294.

[40] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Jul. 1998.

MOHAMMAD ALSMIRAT (Member, IEEE) received the bachelor's degree in computer science from the Jordan University of Science and Technology, in 2002, the master's degree in computer science from the New York Institute of Technology, in 2003, and the Ph.D. degree in computer engineering from Wayne State University, in 2013. From October 2003 to December 2006, he was a full-time Lecturer with the Department of Computer Science, Jordan University of Science and Technology. He then joined the Department of Electrical and Computer Engineering (ECE), Wayne State University, as a Graduate Research Assistant, in January 2007. From August 2007 to December 2012, he worked as a Graduate Teaching Assistant with the ECE Department. After that, he worked as a Research and Development Engineer with General Motors for a period of nine months until he moved back to Jordan to join the Department of Computer Science, Jordan University of Science and Technology, as an Assistant Professor. In 2018, he was promoted to an Associate Professor rank. In 2020, he joined the Department of Computer Science, University of Sharjah, as an Associate Professor, where he is currently working there. He published more than 65 research papers in conferences and journals. His research interests include multimedia system and networks, image processing and computer vision, parallelization using GPU, and machine learning.

NUSAIBA AL-MNAYYIS received the Bachelor of Computer Engineering and master's degrees in computer science from the Jordan University of Science and Technology (JUST), in 2013 and 2018, respectively. Since 2019, she has been a Programmer with the Information Systems Department, King Abdullah University teaching hospital (KAUH). She also worked as a part-time Lecturer with the College of Computer and Information Technology, Jordan University of Science and Technology, from 2018 to 2019. She is currently a Computer Engineer with JUST. Her research interests include image processing, deep learning, and computer-aided diagnosis (CAD) systems.

MAHMOUD AL-AYYOUB received the Ph.D. degree in computer science from Stony Brook University, in 2010. He is currently a Professor of computer science with the Jordan University of Science and Technology (JUST). He is also working as the Director of the Center of Excellence for Innovative Projects. His research interests include cloud computing, high-performance computing, machine learning, and AI.

ASMA'A AL-MNAYYIS received the Bachelor of Medicine, Bachelor of Surgery (MBBS) degree from the Jordan University of Science and Technology (JUST), in 2005, and the High Specialization degree in diagnostic radiology from the Jordan University of Science and Technology (JUST) and its teaching hospital (King Abdullah University Hospital KAUH), in 2010. She received her clinical fellowship in musculoskeletal and spinal radiology from the University of Milan and Galeazzi Orthopedic Institute, Italy, in 2017. She has a Full Registration with the Jordanian Medical Council and the Jordanian Society of Radiology and worked as a Consultant Radiologist in different hospitals in Jordan. She is currently an Associate Professor and a Consultant Radiologist. Since 2017, she has been with the Department of Clinical Sciences, College of Medicine, Yarmouk University, as a Member of the Staff. She also worked as a part-time with the College of Medicine, JUST, teaching musculoskeletal radiology for undergraduate students and postgraduate radiology residents. Her research interests include musculoskeletal, radiology, cross sectional imaging, CAD radiology, artificial intelligence, imaging in metabolic diseases, and female imaging. She is also a Registered Member of the European Society of Radiology and the European Society of Musculoskeletal Radiology and fulfilled all the required European Board of Radiology Exams.