Convolutional neural network based diagnosis of bone pathologies of proximal humerus

Aysun Sezer, Hasan Basri Sezer

Version of Record: https://www.sciencedirect.com/science/article/pii/S0925231219304837
Manuscript_a8080460b28bc96e4633165af058fd56

Convolutional Neural Network Based Diagnosis of Bone Pathologies of Proximal Humerus

Aysun Sezer[a,*], Hasan Basri Sezer[b]

[a] *Unité d'Informatique et d'Ingénierie des Systèmes, ENSTA-ParisTech, Université de Paris-Saclay, France*
[b] *University of Medical Sciences, Şişli Hamidiye Etfal Training and Research Hospital, Orthopaedics and Traumatology Clinic, Istanbul, Turkey*

Abstract

MRI is the leading method of evaluation in traumatic shoulder pathologies ranging from soft tissue or bone edema to rupture of tendons or ligaments and subtle fractures of bone. MRI has the power of evaluating both of the soft tissues and bone at the same time. PD weighted MRI sequences are very powerful in demonstrating bone trauma in terms of edema and cortical disruption in the shoulder pathologies. The alterations in the intensity of the metaphyseal bone may be used to predict the presence of the bone trauma or pathologies.

Although it is superior to many other imaging modalities to uncover the effects of trauma on the each anatomical structure of shoulder, PD weighted MRI has innate obstacles to image processing like low signal to noise ratio besides close anatomical relations and unclear bony borders. We proposed an innovative representation of a learning framework for diagnosis of bone lesions from PD MRI sequences that integrates deep learning techniques to automatically learn discriminative features while avoiding the design of specific hand crafted image based feature descriptors. In this study automatically segmented PD weighted shoulder images were evaluated by the proposed convolutional neural network (CNN) to extract features and classify humeral head in three groups as normal, edematous and Hill -Sachs lesion with a success rate of %98.43. Compared to

☆.
*Corresponding author e-mail: Aysun.Sezer@obs-vlfr.fr

the state of art methods, our proposed CNN based diagnosis system is very promising to assist radiologists and orthopedists in decision making.
Keywords: Convolutional neural network, bone pathologies, PD weighted MRI, image classification,

1. Introduction

Traumatic shoulder instability is a very important topic of the orthopaedic practice. It is a common reason for surgery as well in young athletes because of the tendency of the shoulder joint to grow instability due to anatomical changes after each episode [1]. Therefore understanding the pathological conditions leading to instability is very important to prevent further dislocations. The ideal imaging technique for evaluation of shoulder instability must reveal the location and magnitude of all affected anatomical parts to uncover the manner and extent of the trauma.

During a dislocation the magnitude of trauma is often high enough to traumatize bone besides soft tissues surrounding the joint. The recognition of all of the traumatized tissues is essential for understanding the whole picture of affected anatomical elements after a dislocation which is the key for treatment. To localize all of the traumatic changes after a dislocation often required an imaging modality which is capable of visualizing shoulder joint from more than one aspect [2].

The bone is the basic support of the locomotion and provides attachment to surrounding ligaments and muscles. It is a solid tissue which is resistant to trauma due to its hard nature and its relatively deep location in the human body. The response of bone to trauma is peculiar to itself which begins with minor changes like bone edema but eventually ends up with plastic deformation. The soft tissues surrounding the bone plays a barrier role against trauma. Therefore in the presence of bone trauma, soft tissue edema is also anticipated [3]. Moreover if the integrity of vasculature of bone is disrupted, the bleeding leads to hematoma formation in the vicinity of bone rather than inside of

its boundaries because the bone has no space to capacitate bleeding inside its contours [4].

X-ray is the first line of the radiological examination to evaluate bone pathologies. The plain X-ray is very useful in providing the global information of bone in many conditions but sometimes it may not be sufficient due to overlapping bony contours. Computerized tomography (CT) is another method utilizing X-rays in order to produce slices of bone and generating 360^o information. Nonetheless, CT is not as powerful in evaluating soft tissues as it is for bone pathologies. Moreover these imaging modalities are accused of the radiation exposure which has many unwanted effects on the human body. The ultrasound is a valuable and a noninvasive method for evaluation of soft tissues but is ineffective in the examination of bone [5].

The MRI is one of the mostly used imaging modalities of the present time for the diagnosis of musculoskeletal disease. MRI is a safe and a very powerful tool which can be used to detect the early phases of diseases due to its high sensitivity in detection of edema which may result from inflammation or trauma [6]. In disease situations like traumatized or instabile shoulder both of the bone and soft tissue pathologies are critical in the diagnostic evaluation. Therefore it is not hard to expect that the MRI would be the choice of visualization method in shoulder region.

MRI has different sequences all of which have special distinctive capacity for different tissues or pathologies. The proton density (PD) weighted MRI sequence is very useful in evaluating extremities. PD weighted MR images are able to localize bone and soft tissue pathologies at the same time. It provides better anatomical details than T1 sequences. However discrimination of soft tissue and bone is not easy in some parts of the body like shoulder [7].

The philosophy of MRI is based on the assumption that under strong magnetic fields of the MR machine, protons of the different tissues behave differently which can be detected and calculated. There is a high water content in the diseased portions of the body. In the proton density weighted MR images the tissues having high water content are labeled as high intensity values [8].

Therefore the detection of the white colored items means that there is increased density of protons in that tissue which is the sign of the disease process in the PD weighted sequences of MRI.

Bone edema is present in many painful conditions of shoulder like trauma, avascular necrosis etc. [9, 10]. In the traumatized shoulder the magnitude of trauma is somewhat directly proportional to the pathological changes in the humeral head. For example after a shoulder dislocation the trauma to the humeral head has a sequential effect on the bone. The humeral head may stay unaffected, may be edematous or with greater forces there may be cortical disruptions of bone which may lead to instability of the shoulder joint or fracture of bone. Therefore the detection of presence of edema in the humeral bone and its distribution may be used for prediction of direction and magnitude of trauma and concomitant pathologies of the other parts of shoulder joint. Not surprisingly the PD weighted MRI sequence is the imaging modality of choice by many clinicians and radiologists in order to investigate traumatic instability of shoulder in terms of the bone edema or cortical deformations of the humeral head bone like Hill-Sachs lesions [11].

The Hill-Sachs lesion (HSL) is the cortical defect of the posterosuperior portion of humeral bone which is mostly encountered after anterior dislocation of shoulder joint. The presence of bone edema in the posterosuperior portion of humeral head in PD sequences after anterior dislocation of shoulder joint may lead to clinical suspicion of the presence of HSL [9, 2]. Traumatic deformation of the humeral head also means that there is an alteration of shape of the humeral bone coinciding with bone edema. Nevertheless the location of the bone edema which is nearby the cortical deformation may jeopardize the visual detection of the cortical disruption of bone in the PD weighted MRI sequences even by the visual inspection which is critical in both the therapeutic and legal means.

Alterations of the shape of the humeral head may have additive effect on instability of the shoulder. Clinical importance of the detection of HSL in the presence of anterior shoulder instability is the bipolar instability pattern which may require additive surgery of the humeral head [2]. The computer

aided diagnosis (CAD) of the cortical disruption in the metaphyseal bone in the PD sequences of MRI may help clinicians and radiologists to diagnose these clinically significant defects of humeral head more precisely. A previous study to discriminate edematous bone from normal bone in the axial PD weighted MR images with a CAD system was done by the authors [12]. There is no other CAD system addressing the bony pathologies of shoulder using PD weighted MR images.

In our previous study, we reduced Rician noise which is present in the MRI and extended the use of speckle reducing anisotropic diffusion (SRAD) method in PD-weighted MR images by estimating speckle scale function from the region of interest. We applied region based active contour method to segment the normal and edematous humeral heads and measured performance of Hermite-based texture features in classification of edematous and normal humeral bones. We concluded that Hermite based textures features are more robust than curvelet, contourlet and gray level co-occurrence matrix-based texture feature descriptors to classify edematous and normal images [12, 34].

In this study we proposed a fully automatic CAD system based on CNN for diagnosis of bone pathologies of proximal humerus as normal, edematous and humeral head with Hill-Sachs lesion. We extended our dataset by adding MR images including Hill-Sachs lesion in the dataset. The whole dataset was segmented with the same manner. As would be predicted the presence of edema might interfere with the diagnosis of subtle fractures like a small Hill-Sachs lesion. Therefore the success rate of the classification of Hermite based features decreased to 94.7% in the extended dataset compared to the 98.23% classification rate of the proposed system on the previous dataset which included only the edematous and normal humeral heads.

This study has contributions by several aspects:

- To the best of our knowledge there is no study in the literature addressing computer based diagnosis of pathologies of bony structures of shoulder utilising CNN.

- We proposed a new computer aided diagnosis system for detection of bone pathologies of shoulder which is unique in both utilising PD weighted MR images and introducing deep learning to the diagnosis of bone pathologies from MRI and studying classification of shoulder pathologies from automatically segmented image patches.

- This study includes comparative results of state of art hand-crafted methods with CNN.

- A new dataset composed of axial PD weighted shoulder MR images of 219 patients were included in the study.

The remainder of paper is organized as follows: Section III describes the dataset and proposed method in detail. Experimental result and discussion are given in Section IV and V.

2. Related Work

The recognition and classification of diseases from the different imaging modalities have been proposed for many fields of medicine in the last decades. With the development of computer vision technologies, many morphological features can easily be extracted from images and used as the basis for classification by machine learning algorithms. These cover first order statistical moments, Histograms of Oriented Gradients (HOG) [13], bag of features [14, 15], Local Binary Pattern (LBP) [16], Gray Level Co-occurrence Matrices (GLCM) [17] etc. called as hand crafted methods. However many CAD systems in literature are based on hand-crafted, shape and texture features from the spatial domain and-or frequency domain [18, 19]. The hand-crafted shape or texture features are dependent on the training level of the evaluator. Multiscale geometric analysis algorithms such as curvelet, contourlet, shearlet transforms are successfully used to extract shape and texture feature from image with promising results. However this handcrafted texture and shape feature extraction

approach is somewhat heuristic of which the success of the system depends basically on the insight and experience of the researchers. Therefore the sensitivity and specificity rates of many of these proposed systems are not in the desired level to be utilized in the clinical field.

Since its introduction deep learning has become one of the trend topics in the field of computer perception. Because deep CNNs are capable of overcoming the limitation of former medical CAD systems by extracting discriminative shape and texture features according to their varying depth and breadth architecture. Deep CNNs are used in the form of adjustable architectures to solve different image feature extraction and classification problems. Shallow networks have more capacity for learning low level features such as edge information while deep networks are more suitable for learning high level features like shape information. By constructing multi-layered networks, more complex features and irregular structures of the medical images can be learned by the CNNs. As a consequence deep CNNs have an applaudable capacity of representing signals in object recognition, image segmentation and classification fields which makes them promising tools for medical diagnosis systems even with small amounts of training samples. Thus it is not surprising that CNNs drawn the attention of the researchers who work in many fields.

There are valuable studies in this new research area with different medical imaging modalities in the literature. Zhang et al. proposed a diagnosis system based on two layers deep learning architecture for classifying breast tumors with shear-wave elastography [20]. Accuracy of their system was 93.4%. Hao et al. localized the fetal abdonominal standart plane from ultrasound videos with a success rate of 90% by using a deep CNN [21]. Arevalo et al. classified breast cancer lesions from manually segmented mammography films with a deep CNN model with a success rate of 82% [22]. Ma et al. classified thyroid nodules from ultrasound image patches with fusion of two pre-trained CNNs with an average success rate of 83.02% [23]. Marios et. al. classified lung CT images with five convolutional layers and a Leaky activation layer with a performance rate of 85% [24].

There are a few studies in the literature employing deep CNNs on musculoskeletal images of different modalities in the literature. Roth et al. used deep CNNs for detection of sclerotic spinal metastasis from spinal CT images and reported 83% success rate from a study group of 59 patients [25]. Geng et al. studied bone tumors with deep CNNs in the segmentation of scintigraphy images and reported 88% success rate according to dice metric [26] . Different CNN architectures were applied for vertebrae localization and segmentation with the CT [27, 28] and MRI [29, 30]. Age assessment studies were reported from hand MRI [31] and X-ray [32] with deep learning. Prasoon et al studied knee cartilage segmentation using multi-stream CNNs in knee MRI [33].

3. Material and Methods

3.1. Image dataset and pre-processing operation

We included 219 shoulder MR images of randomly selected patients who admitted with painful shoulder due to traumatic and acute conditions. 38 of the patients had the complaint of anterior shoulder instability. 100 patients were labeled as having bone edema by the orthopedist. All of the included images were 1.5 Tesla PD weighted axial MRI images. The slice thickness was 4 millimeters and the size of the images were 256x256 pixels.

A healthy humeral head, being composed of cancellous bone, has an almost homogenous architecture presented as a dark round bony structure (Fig 1a). An edematous humeral head contains intensity changes in the form of whitish patchy irregularities not specific to any area in the humeral head (Fig 1b). A Hill-Sachs lesion is a more localized version of humeral head abnormality which is located in the posterosuperior aspect of the humeral head and causing not only increased intensity due to bone edema but also disruption of cortical bone in the form of a depression fracture (Fig 1c, Fig 2).

2D MR images were used in this study instead of 3D MR images because 2D MR images demonstrate the inner structure of the bone where the bone edema exist. Moreover 2D MRI slices are also more suitable to the scope of

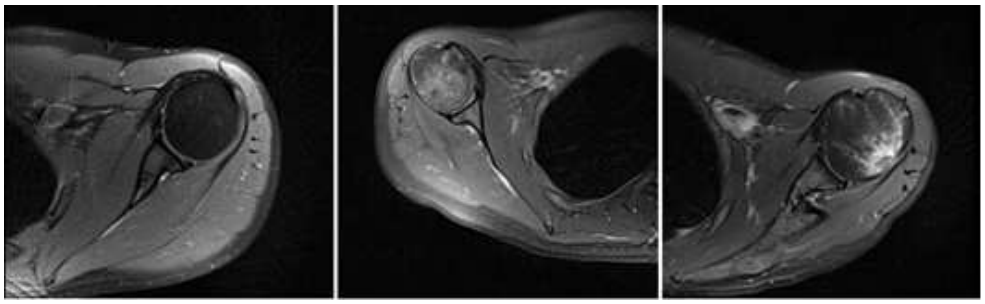

Figure 1: Axial PD MR images of humeral heads representing a) normal b) edematous c) Hill-Sachs lesion

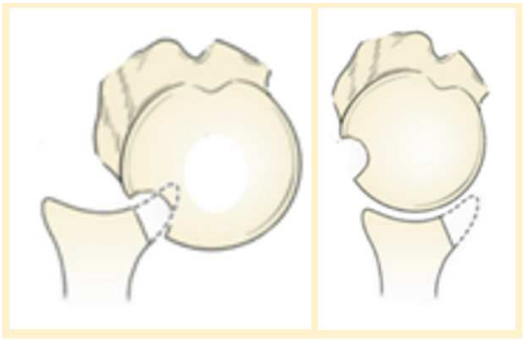

Figure 2: Demonstration of a) anterior shoulder dislocation and b)Hill-Sachs lesion

this study due to their capability of representing both of the cortical disruption and intensity changes of the cancellous bone of the humeral head simultaneously. The same as the many clinicians preference, we also selected PD weighted images due to their superiority to the other imaging modalities and other sequences of MRI in presenting bone edema and the anatomic details. Besides the advantages of PD weighted images in the shoulder region, there are many difficulties of segmentation of humeral head in this sequence, as well.

3.2. Segmentation of Humerus Bone and Determination of ROI

The segmentation of the normal humeral head is a challenging task due to anatomical complexity of the shoulder. Humeral bone serves as an attachment of complex ligamentous and tendinous structures of the shoulder which has nearly

the same intensity as the cortical bone in the PD weighted images. From the point of view of segmentation there is innate poor signal to noise ratio (SNR) and the blurred transition zones in the PD weighted images. Moreover in a traumatized shoulder the altered anatomical shape and texture of the humeral bone due to cortical discrepancies and resulting edema may further complicate a successful segmentation operation.

In the presence of such difficulties we proposed a strategy to facilitate the segmentation process by decreasing noise before the segmentation process and also decreasing the size of target area by determining the area of region of interest. For this purpose we applied circular Hough transform (CHT) to automatically localize the humeral head. Afterwards we applied active contours without edges described by Chan-Vese (CV) in the area determined by the CHT. The CV method uses the statistical information to evolve the curve with the level set function to segment images. It is a well-known and successful method to segment inhomogeneous images.

The segmentation success of this approach is prone to the intensity inhomogeneities and the determination of the initial contour. To overcome both of these problems the CHT was to be applied successfully. The CHT was first used to exclude unnecessary fields from the image in which large amounts of intensity inhomogenous tissues inhabitate. Secondly the location of the initial contour was automatically determined with the center point function of the CHT which is indeed the center of the humeral head. By this way we successfully segmented the humeral head despite the presence of local intensity inhomogeneity problem.

The segmentation result is demonstrated in the Fig 3. for three groups of humeral head. The image patches were constructed according to segmented humeral head regions (for further detail we refer to read [34]).

3.3. Convolutional Neural Networks

CNNs are multilayered mechanisms that work in synergy. The main structures that constitute a CNN are convolution layer, activation function and pooling layer which have a cascaded mechanism of functioning that serves the output

Figure 3: Segmentation results of humeral heads labelled as a) normal, b) edematous and c) Hill-Sachs lesion

of one layer to the following layer as an input. According to the requirements of the specific pattern recognition problem there may be a number of intermediate layers to extract effective low level features which are used to build high level features. CNNs have a learning capability by analyzing the hierarchy of these effective features.

Not so different from the mechanism of human visual conception every neuron of a layer is connected to its own specific region of image which commonly overlap to other neuronal output. CNNs can better represent the texture and geometrical structure of an image by analyzing the overlapping information of local areas defined by each neuron of the layer.

CNNs have hierarchical organization structure which can be multiplied in order to increase the depth and breadth of the network to represent images that have a more complex structure. There is no strict rule of estimating the optimum depth and breadth configuration of the CNNs to achieve a desired goal proven mathematically. Therefore it is an expert dependent path which is principally a trial and error minimization approach for prediction of the optimum architecture of the CNN. The training of CNNs, like other artificial neural networks is through minimizing a loss function which produces information to feedback the CNN.

The convolutional layer is the key block of CNN which has a particular importance in extracting features from the raw image. It uses learnable filters

each of which represents a small portion of the image called as a local receptive field. Every kernel is slidable on the image, the magnitude of which is defined by the stride parameter. According to the width and height of the input volume the convolving operation produces a feature map. The specific type of features produced by different filters as an input are learned by the network which will be used to produce the full output volume through the stacking of all of the activation maps in the next step.

Activation function uses the feature maps produced by the convolutional layer as input information. In case of non-linear problems an activation function which adds non-linearity to the system is required. Recently the rectified linear unit (RELU) is the mostly used function in this step. RELU replaces negative values of the feature maps with zero which considerably increase the speed of the convergence of the CNN [35].

The pooling layer is another important parameter of the CNN which is actually a form of down-sampling. It provides to reduce feature maps with using several non-linear functions of which the max and average pooling are the mostly used. The pooling layer prevents overfitting by decreasing the amount of parameters and computation. The output of the convolution and pooling layers are high level features extracted from image. The fully connected layer of CNN utilizes these discriminative features as input and provides classification of the test data according to training dataset.

3.4. Proposed Convolutional Neural Network

The proposed CNN consisted of five convolution and three max pooling layers. Each image patch of humeral head was convolved with the kernel of size 3x3 in the first convolution layer to obtain feature maps. 192 feature maps of size 64x64 were generated by the first convolution layer. Output of the first convolution layer was given to the second convolution layer as an input to extract useful low level features. At this stage 64 feature maps of size 64x64 were obtained. Pooling operation was generated after suppression of non-negative values in the feature maps by rectified linear unit. Max-pooling was applied to

Table 1: Detail of proposed CNN components (Conv: Convolution; MP: Max poling layer; FC: Fully connected layer)

Details of proposed CNN1 architecture				Details of proposed CNN2 architecture			
Layers	Input size	CNN1	Stride	Layers	Input size	CNN2	Stride
Conv1	64x64	3x3x64	1	Conv1	64x64	3x3x192	1
MP1	32x32	2x2	2	Conv2	64x64	5x5x64	1
Conv2	32x32	5x5x128	1	MP1	32x32	2x2	2
MP2	16x16	2x2	2	Conv3	32x32	7x7x128	1
Conv3	16x16	7x7x256	1	Conv4	32x32	5x5x64	1
MP3	8x8	2x2	2	MP2	16x16	2x2	2
Conv4	8x8	9x9x512	1	Conv5	16x16	9x9x256	1
MP4	4x4	2x2	2	Conv6	16x16	5x5x64	
FC1	128	-		MP3	8x8	2x2	2
				FC1	128	-	-

feature maps with the stride (S) size of S=2 to make input representation in smaller size.

In the second stage, we applied again two cascading convolution layers once again but with different parameters to detect different high level features as edges, circular shape structure of humeral head. 128 features maps of size 32x32 were generated with parameters F=7x7, S=1 and padding (P) size of P=1 in the first, and 256 features maps of size 16x16 F=9x9, S=1, P=0 in the second convolution layer. In the last stage one convolution layer was used to obtain 64 feature maps of size 16x16 with F=5x5, S=1, P=0. Finally, a fully connected layer with 128 hidden units with max-out activation function was stacked to label each image patches as normal, edematous and Hill-Sachs based on training dataset (Table 1).

We tested different CNNs architectures to recognize humeral head bone. We realized that applying convolution layer one after another instead of directly

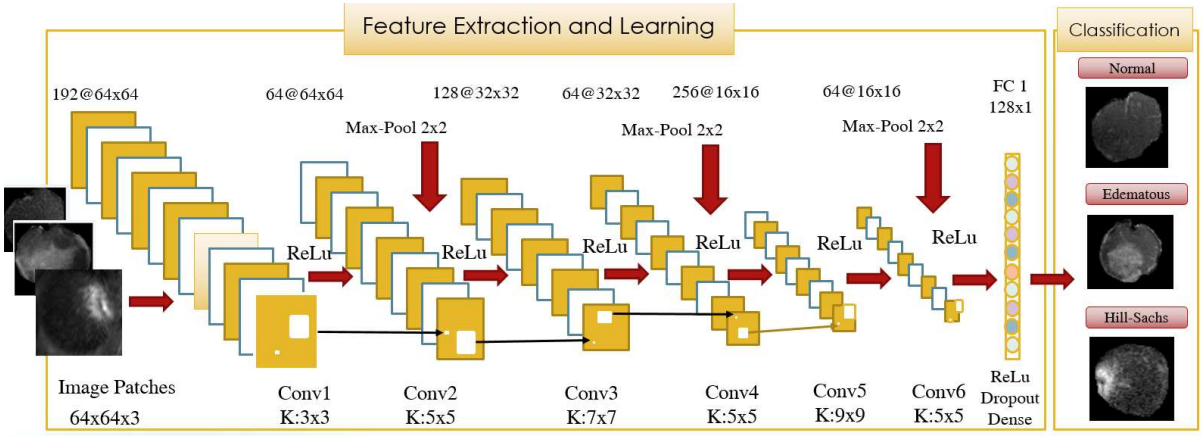

Figure 4: Demonstration of the proposed CNN architecture for classification of humeral head (Conv: Convolution; MP: Max pooling layer; FC: Fully connected layer)

reducing the dimension by application of pooling layer provides to extract more discriminative features. Therefore CNN2 architecture was more useful to distinguish HSL with edema than the proposed CNN1 structure (Fig 4).

4. Experimental Results

Classification of PD weighted shoulder MR images is a difficult mission due to both anatomical complexity of the shoulder region and image quality of the selected MR sequence. Discrimination of soft tissues covering the humeral head which are represented by a very close intensity level to the bone is already a problematic task. In addition the low signal to noise ratio of the PD weighted sequences complicates the segmentation and classification of humeral head. In order to decrease the workload and increase the precisions of the classification we automatically segmented humeral head using region based active contour model and created image patches which include humeral head and minimal amount of surrounding tissues. Each image patch has a size of 64x64 pixels. Each image patch was validated and also labeled by the expert.

The success of diagnostic system depends on the selected approaches of feature extraction as well as the segmentation success of the humeral head. Our

proposed diagnosis system of bone pathologies of proximal humerus is based on CNN. CNNs are capable of learning discriminative features according to their architecture. The size and the number of filters used in the convolution layer of CNN has an important role of capturing the distinctive features during the training process. It is possible to construct many different CNN architectures each of which has a different behavior and success rates for a given condition. However it is still a trial and error approach to define the most fitting CNN architecture for a specific problem.

Table 2: The confusion matrix of overall accuracy for our CNN1 architecture with 10 iterations.

Humerus Type	Classification accuracy of humeral head by our proposed CNN1 (%)			Confusion matrix for our proposed CNN1 (%)		
	Normal	Edematous	HSL	Normal	Edematous	HSL
Normal	**100**	0	0	**79**	0	0
Edematous	2.53	97.80	0	2	89	0
HSL	0	1.09	97.95	0	1	48

Table 3: The confusion matrix of overall accuracy for our CNN2 architecture with 10 iterations.

Humerus Type	Classification accuracy of humeral head by our proposed CNN2 (%)			Confusion matrix for our proposed CNN2 (%)		
	Normal	Edematous	HSL	Normal	Edematous	HSL
Normal	**100**	0	0	**79**	0	0
Edematous	1.26	98.90	0	1	90	0
HSL	0	0	**100**	0	0	**49**

The focus of this study is to discriminate between normal, edematous humeral heads and humeral heads with Hill-Sachs lesion. All of the experiments were carried out on a server having a NVIDIA Geforce GTX Titan X (6 GB on board memory). The CNN algorithms were performed with the TensorFlow.

Table 4: Classification accuracy of several state of art methods

Methods For Diagnosis of Bone Pathologies	Overall accuracy rate of the system (%)
PHOG evaluated on original data set	85 ± 0.30
Gray level co-occurence matrix (GLCM)	91.6 ± 0.26
Hermite based GLCM	94.7 ± 0.20
PHOG + GLCM (composite kernel)	94 ± 0.17
AlexNet evaluated on original data set	85.48 ± 0.05
GoogLeNet evaluated on original data set	82.25 ± 0.06
Proposed CNN2	98.43 ± 0.02

To evaluate the reliability and stability of our proposed diagnosis system which was based on CNN, the data set was split into two parts as 70% for training and 30% for testing. The images employed in the training set were excluded from the test data. Random sampling was done ten times in order to produce different training and testing sets from the data-set in order to decrease the selection-bias.

Table 2 and Table 3 demonstrate the confusion matrix of the proposed CNN1 and CNN2 method for classifying the normal, edematous humeral heads and humeral heads having Hill-Sachs lesion. The accuracy rate of CNN1 was 95.31% while the accuracy rate of CNN2 was 98.43 % for ten iterations. Although the defective and edematous humeral heads were supposed to be more difficult to differentiate visually by experts, our proposed CNN2 had 100% accuracy rate for diagnosis of Hill-Sachs lesion. Only one edematous image was confused with the normal humeral head image group.

We evaluated traditional feature extraction methods as pyramid HOG, GLCM and Hermite based GLCM in order to compare the performance of our proposed CNN method. Each image patch of shoulder MRI was divided into cells and histogram of gradient directions were calculated for each cells. Pyramid of HOG is a useful shape feature extraction method but its success rate was 85% for our

dataset.

Although the main differences of HSL group from normal or edematous groups was the shape of the humeral head the texture information also has a positive effect on the recognition. Normal humeral head images has a more homogeneous distribution of intensity values than HSL and edematous humeral head images. Thus both of the shape and texture features might be accepted as essential components of the proposed shoulder bone recognition system.

We calculated conditional probability density function of gray level pair with different offset and orientation parameters to better define second order statistical texture features. We concatenated extracted GLCM features with Pyramid HOG feature by composite kernel. By this way the overall accuracy rate for classification of humeral head pathologies increased up to 94 %. In order to obtain more robust texture features, we calculated second order GLCM features as entropy, autocorrelation, energy, maximum probability... etc by using Hermite coefficients which were proven to be more powerful than curvelet and contourlet transforms [12].

Instead of exploiting traditional texture and shape features we classified our dataset with pre-trained CNNs of GoogLeNet and AlexNet which are two successful CNNs used very commonly for pattern recognition problems. Nonetheless optimization of the hyperparameters and determining the optimal learning rate of AlexNet and GoogLenet for different layers are very challenging tasks due to the limited knowledge about the relevance of the large number of architectural and training hyperparameters. We followed the approach of Sharif et al. in which they assumed to utilise very small learning rates compared to default. After meticulous parameter tuning, we obtained the most successful model for our problem, which was very difficult to further improve. The accuracy rates of GoogLeNet and AlexNet were 82.25% and 85.48% respectively as demonstrated in Table 4. The GoogLeNet and AlexNet were also popular for their success in classification of images however their success rates were not in desired levels in differentiation of edematous humeral heads from humeral heads with Hill-Sachs lesion. Another possible reason might be the fact that GoogLenet and AlexNet

were pre-trained on natural image datasets which are completely different from the MR images of the humeral head. GoogLeNet and AlexNet were originally designed for the fixed image dimension of 256x256x3 pixels whereas our dataset was composed of MR images of 64x64x3 pixels.

Utilizing pre-trained CNNs may be a solution for problem of limited number of samples. Unfortunately there is no existing pre-trained CNN model in the literature on medical images or images with similar structure. Our proposed CNN architecture was performing well despite the fact that it is not deep as GoogLeNet and AlexNet.

5. Discussion

When the extent of trauma reaches to the bone not only the surrounding tissues but also the trabecular bone inside the cortical borders present abnormalities. MRI is a very powerful tool for imaging of shoulder pathologies because of its high capacity of representing pathologies both inside and outside the bone. The PD weighted MR images has the ability to demonstrate the extent of trauma by means of edema of the tissues. The reaction of many tissues to trauma is edema which is represented as increased intensities in PD weighted MRI sequence. Therefore as the edema reaches to the bone the increased intensity in other words "noise in the image" scatters through the borders of bone. To differentiate the tissue characteristics in such a noisy condition is not easy to overcome in an area of complex anatomical architecture.

A successful automatic classification of the pathological humeral head requires a successful identification of the target anatomical region. We automatically segmented humeral head and produced image patches from axial PD weighted MR images of shoulder. Each image patch then was given to the proposed CNN to classify the humeral heads according to the extracted features.

CNN is a kind of self-learning mechanism which uses filters to represent signals each of which serves as an input to another. The role of a filter is to sort the given signal according to image characteristics and to produce quantitative

values. The CNN learns these values regardless of the location of the object and recognize these patterns in the image. We hypothesized to construct a CNN ultrastructure containing a cascaded convolutional layers to extract low level features from the MR images of the humeral head. Using multiple convolutional layers in each stage increases the effectiveness of extraction of low level texture features which will serve to construct the high level features. This approach enabled our CNN to produce high quality representative features (Table 1).

Pooling operation is a very important part of CNN which reduces workload and prevents overfitting by decreasing the size of feature maps. However an aggressive pooling operation may result with loss of important information produced by convolutional layers which may be useful in representing the image details. Therefore we avoided an assertive pooling operation which enabled us to handle translation invariance problem while protecting important features.

Sensitivity and specificity rates of a CAD system are very important in order to evaluate its validity which is the key measure for the usability in the clinical field. Sensitivity (true positive rate) is the measure of correctly labelling the true positives. Whereas specificity (true negative rate) is the ability of a test to correctly classify a person as healthy when there is no illness. Our proposed CAD system classified humeral heads with a sensitivity rate of 100% (95% CI 97.38% to 100%) and a specificity rate of 98.75% (95% CI 93.23% to 99.97%) within a 95% confidence level. The proposed system correctly labelled all of the patients with Hill-Sachs lesion but only one patient having edema was labelled as normal. Therefore our system may be regarded as precise in determining Hill-Sachs lesions. Although the sensitivity rate of the system is at desired levels the specificity level is to be improved in order to overcome misclassification problem which may lead to misdiagnosis. For this reason many medical diagnosis systems reported in the literature which were proposed to handle diverse pathologies in different anatomical regions had no chance for use in real life.

Although the specificity rate of the system is at desired levels the sensitivity level is to be improved in order to overcome misclassification problem which may lead to overdiagnosis. For this reason many medical diagnosis systems reported

in the literature proposed to handle diverse pathologies in different anatomical regions had no chance for use in real life.

The difference between an edematous humeral head from a normal humeral head is the texture information. On the other hand discrimination of Hill-Sachs lesion from the edematous humeral heads requires successful extraction of shape features besides texture features because the traumatized humeral head with a Hill-Sachs deformity mostly contains cortical deformation in addition to bone edema. The high success rate of the proposed system in diagnosing a Hill-Sachs lesion might be due to its excellent capacity to differentiate the shape features. According to our knowledge this is the first study in the literature handling classification of MR images of shoulder pathologies by using CNNs

According to our knowledge this is the first study in the literature handling classification of MR images of shoulder pathologies by using CNNs. Although MRI is the gold standard in the diagnosis of bone edema, automatic classification of bone pathologies especially from PD weighted MR images was not studied possibly due to the challenges in the segmentation and also the heterogeneity of the magnitude and distribution of edema in the bone.

6. Conclusions

Humeral head defects is a challenging topic in the orthopedic practice. There are ongoing studies on this subject in the orthopaedic literature which address both of the humeral and glenoid defects in the shoulder also called bipolar lesions. The treatment of the bipolar lesions is still under research. This study is concentrated on the humeral pathologies and detection of humeral bone defects. Further research may be on measuring the bone defects and the combined effect of bipolar bone lesions on the MR images. We proposed a CNN architecture to classify bone pathologies of the humeral head. To increase the success of the classification by the CNN, increasing either the number of filters or the number of convolution layers may be addressed. The small number of data present in our dataset was one of the limitations of this study. The repetition of convolution

layers in each step with different size and number of filters besides avoiding an aggressive reduction in the pooling layer in such a small dataset are the key points of this CNN architecture.

References

[1] R. J. Robins, J. H. Daruwalla, S. C. Gamradt, E. C. McCarty, J. L. Dragoo, R. E. Hancock, J. A. Guy, G. A. Cotsonis, J. W. Xerogeanes, A. C. Group, et al., Return to play after shoulder instability surgery in national collegiate athletic association division i intercollegiate football athletes, The American journal of sports medicine 45 (10) (2017) 2329–2335.

[2] K. P. Lim, I. S. Lee, I.-B. Kim, Intra-articular lesions and clinical outcomes in traumatic anterior shoulder dislocation associated with greater tuberosity fracture of the humerus, 20 (4) (2017) 195–200.

[3] T. Ruedi, R. Buckley, C. Moran, AO principles of fracture management, books and DVD, Thieme, 2007.

[4] H. Schell, G. Duda, A. Peters, S. Tsitsilonis, K. Johnson, K. Schmidt-Bleek, The haematoma and its role in bone healing, Journal of experimental orthopaedics 4 (1) (2017) 5.

[5] M. Mujoomdar, E. Russell, F. Dionne, K. Moulton, C. Murray, S. McGill, K. Lambe, Optimizing health system use of medical isotopes and other imaging modalities.

[6] D. T. Felson, S. McLaughlin, J. Goggins, M. P. LaValley, M. E. Gale, S. Totterman, W. Li, C. Hill, D. Gale, Bone marrow edema and its relation to progression of knee osteoarthritis, Annals of internal medicine 139 (5_Part_1) (2003) 330–336.

[7] D. Nascimento, G. Suchard, M. Hatem, A. de Abreu, The role of magnetic resonance imaging in the evaluation of bone tumours and tumour-like lesions, Insights into imaging 5 (4) (2014) 419–440.

[8] P. S. Tofts, Pd: proton density of tissue water, Quantitative MRI of the Brain: Measuring Changes Caused by Disease (2003) 83–109.

[9] M. Jana, S. Gamanagatti, Magnetic resonance imaging in glenohumeral instability, World journal of radiology 3 (9) (2011) 224.

[10] P. J. MacMahon, W. E. Palmer, Magnetic resonance imaging in glenohumeral instability, Magnetic Resonance Imaging Clinics 20 (2) (2012) 295–312.

[11] S. Demehri, N. Hafezi-Nejad, E. K. Fishman, Advanced imaging of glenohumeral instability: the role of mri and mdct in providing what clinicians need to know, Emergency radiology 24 (1) (2017) 95–103.

[12] A. Sezer, H. B. Sezer, S. Albayrak, Hermite-based texture feature extraction for classification of humeral head in proton density-weighted mr images, Neural Computing and Applications 28 (10) (2017) 3021–3033.

[13] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, Vol. 1, IEEE, 2005, pp. 886–893.

[14] M. R. Zare, A. Mueen, W. C. Seng, Automatic classification of medical x-ray images using a bag of visual words, IET Computer Vision 7 (2) (2013) 105–114.

[15] B. André, T. Vercauteren, A. M. Buchner, M. B. Wallace, N. Ayache, Learning semantic and visual similarity for endomicroscopy video retrieval, IEEE Transactions on Medical Imaging 31 (6) (2012) 1276–1288.

[16] Z. Guo, L. Zhang, D. Zhang, A completed modeling of local binary pattern operator for texture classification, IEEE Transactions on Image Processing 19 (6) (2010) 1657–1663.

[17] S. Mishra, B. Majhi, P. K. Sa, L. Sharma, Gray level co-occurrence matrix and random forest based acute lymphoblastic leukemia detection, Biomedical Signal Processing and Control 33 (2017) 272–280.

[18] P. Amorim, T. Moraes, D. Fazanaro, J. Silva, H. Pedrini, Electroencephalogram signal classification based on shearlet and contourlet transforms, Expert Systems with Applications 67 (2017) 140–147.

[19] U. R. Acharya, U. Raghavendra, H. Fujita, Y. Hagiwara, J. E. Koh, T. J. Hong, V. K. Sudarshan, A. Vijayananthan, C. H. Yeong, A. Gudigar, et al., Automated characterization of fatty liver disease and cirrhosis using curvelet transform and entropy features extracted from ultrasound images, Computers in biology and medicine 79 (2016) 250–258.

[20] Q. Zhang, Y. Xiao, W. Dai, J. Suo, C. Wang, J. Shi, H. Zheng, Deep learning based classification of breast tumors with shear-wave elastography, Ultrasonics 72 (2016) 150–157.

[21] H. Chen, D. Ni, J. Qin, S. Li, X. Yang, T. Wang, P. A. Heng, Standard plane localization in fetal ultrasound via domain transferred deep neural networks, IEEE journal of biomedical and health informatics 19 (5) (2015) 1627–1636.

[22] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, M. A. G. Lopez, Representation learning for mammography mass lesion classification with convolutional neural networks, Computer methods and programs in biomedicine 127 (2016) 248–257.

[23] J. Ma, F. Wu, J. Zhu, D. Xu, D. Kong, A pre-trained convolutional neural network based method for thyroid nodule diagnosis, Ultrasonics 73 (2017) 221–230.

[24] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung pattern classification for interstitial lung diseases using a deep convolutional neural network, IEEE transactions on medical imaging 35 (5) (2016) 1207–1216.

[25] H. R. Roth, J. Yao, L. Lu, J. Stieger, J. E. Burns, R. M. Summers, Detection of sclerotic spine metastases via random aggregation of deep convolu-

tional neural network classifications, in: Recent advances in computational methods and clinical applications for spine imaging, Springer, 2015, pp. 3–12.

[26] S. Geng, S. Jia, Y. Qiao, J. Yang, Z. Jia, Combining cnn and mil to assist hotspot segmentation in bone scintigraphy, in: International Conference on Neural Information Processing, Springer, 2015, pp. 445–452.

[27] H. Chen, C. Shen, J. Qin, D. Ni, L. Shi, J. C. Cheng, P.-A. Heng, Automatic localization and identification of vertebrae in spine ct via a joint learning model with deep neural networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 515–522.

[28] W. Shen, F. Yang, W. Mu, C. Yang, X. Yang, J. Tian, Automatic localization of vertebrae based on convolutional neural networks, in: Medical Imaging 2015: Image Processing, Vol. 9413, International Society for Optics and Photonics, 2015, p. 94132E.

[29] D. Forsberg, E. Sjöblom, J. L. Sunshine, Detection and labeling of vertebrae in mr images using deep learning with clinical annotations as training data, Journal of digital imaging 30 (4) (2017) 406–412.

[30] R. Korez, B. Likar, F. Pernuš, T. Vrtovec, Model-based segmentation of vertebral bodies from mr images with 3d cnns, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 433–441.

[31] D. Štern, C. Payer, V. Lepetit, M. Urschler, Automated age estimation from hand mri volumes using deep learning, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 194–202.

[32] C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, R. Leonardi, Deep

learning for automated skeletal bone age assessment in x-ray images, Medical image analysis 36 (2017) 41–51.

[33] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, M. Nielsen, Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network, in: International conference on medical image computing and computer-assisted intervention, Springer, 2013, pp. 246–253.

[34] A. Sezer, H. B. Sezer, S. Albayrak, Segmentation of bone with region based active contour model in pd weighted mr images of shoulder, Computational and mathematical methods in medicine 2015.

[35] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (7553) (2015) 436.