

Received August 13, 2020, accepted October 17, 2020, date of current version November 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3037745

Multi-Layered Basis Pursuit Algorithms for Classification of MR Images of Knee ACL Tear

ABDUL WAHID¹, JAWAD ALI SHAH², (Senior Member, IEEE), ADNAN UMAR KHAN¹, MUKHTAR ULLAH³, AND MOHD ZAKI AYOB², (Senior Member, IEEE)

¹Department of Electrical Engineering, International Islamic University, Islamabad 44000, Pakistan

²Electronics Section, UniKL British Malaysian Institute, Selangor 53100, Malaysia

³Electrical Engineering Department, National University of Computer and Emerging Sciences, Islamabad 44000, Pakistan

Corresponding author: Jawad Ali Shah (jawad@unikl.edu.my)

This work was supported by the UniKL BMI Kuala Lumpur under Grant FRGS/1/2019/TK04/UNIKL/02/6.

ABSTRACT Deep learning architectures have been extensively used in recent years for the classification of biomedical images to assist clinicians for diagnosis and treatment management of patients with different health conditions. These architectures have demonstrated expert level diagnosis, and in some cases, surpassed human experts in diagnosing health conditions. The automation tools based on deep learning frameworks have the potential to transform all stages of medical imaging pipeline from image acquisition to interpretation and analysis. One of the most common areas where these techniques are applied is knee MR image classification for different types of Anterior Cruciate Ligament (ACL) tears. If properly and timely managed, the diagnosis and treatment of ACL tear can avoid further degradation of patients' knee joints and can also help slow the process of subsequent knee arthritis. In this work, we have implemented a novel classification framework based on multilayered basis pursuit algorithms inspired from recent research work in the area of the theoretical foundation of deep learning with the help of celebrated sparse coding theory. We implement an optimal multi-layered Convolutional Sparse Coding (ML-CSC) framework for classification of a labelled dataset of knee MR images with the coronal view and compare the results with traditional convolutional neural network (CNN) based classifiers. Empirical results demonstrate the effectiveness of the ML-CSC framework and show that the framework can successfully learn distinct features on a small dataset and achieve a good efficiency of more than 92% without employing regularization techniques and extensive training on large datasets. In addition to 95% average accuracy on the presence and absence of ACL tears, the framework also performs well on the imbalanced and challenging classification of partial ACL tear with 85% accuracy.

INDEX TERMS Basis pursuit, iterative shrinkage algorithms, knee MR image classification, multi-layer convolutional sparse coding.

I. INTRODUCTION

One of the most common sports injuries in young adults is anterior cruciate ligament (ACL) tear. A study which spanned over 21 years, discovered an incidence of 68.8 per 100,000 person-years in general population [1]. The diagnosis requires surgical intervention such as reconstruction or enhanced primary repair to avoid further damage and degeneration of injury into osteoarthritis and subsequent chronic instability [2]–[4]. The frequent occurrence of ACL tear in sports community and general public requires accurate

The associate editor coordinating the review of this manuscript and approving it for publication was Chintan Amrit¹.

diagnosis of complete and partial ACL tears. This is also important for therapeutic decision making and avoidance of further damage. In addition to examinations by experienced sports medicine specialists for exams like pivot shift tests, magnetic resonance (MR) imaging is routinely used to complement and confirm clinical diagnosis and assess the status of associated injuries. MR imaging plays a crucial part in diagnosing, treatment planning, treatment delivery and follow-ups. Consensus is building among researchers for stronger need of using automated tools in order to reduce costs, increase efficiency and provide higher diagnostic and prognostic accuracy for clinical decision making. Although for an experienced musculoskeletal (MSK)-trained radiologist,

MR imaging is specific and accurate in diagnosing ACL tears [5], the diagnosis becomes challenging for non-MSK radiologists and clinicians without access to sub-specialty radiology. People who do not have access to specialists for diagnosing ACL tear injuries remain at risk of further deterioration of injuries without timely and proper diagnosis. Deep learning (DL) has emerged as powerful tool for image processing tasks in recent years complimented with the development of graphical processing hardware. The subsequent DL algorithms developed so far have been successfully used in tasks like object detection [6], MR image reconstruction [7], [8] and classification of biomedical images [9], [10]. The primary advantage of learning representations through deep neural networks is their ability to learn semantically meaningful patterns and features in underlying data without explicit human intervention. These models once trained successfully on training datasets, can be effectively used for solution of range of problems like image recognition and image classification on (unseen) test data. With tremendous success of DL architectures i.e. the Convolutional Neural Networks (CNNs) until recent years, their working has largely remained heuristic and a deeper understanding required to model their working and improve performance. Sparse coding theory [11] developed over the last decade has been applied to range of problems in image processing [12], [13]. The theory is based on constructing models that represent signals as linear combinations of few columns, called atoms from a given redundant matrix termed as dictionary. This theory applied successfully in array of image processing tasks over the last decade has been recently extended to explain theoretical foundation of DL. The convolutional sparse coding (CSC) and its multilayered version ML-CSC have been introduced to explain the theoretical foundation of DL and its association with the sparse coding theory. Specifically, the CNNs are interpreted as approximations of multi-layer basis pursuit problem [14], [15].

The wider availability of datasets and the learning ability of DL architectures demonstrate their capacity to become part of biomedical imaging workflow and help revolutionize the healthcare industry especially for communities with limited access to specialized facilities. The deep learning (DL) frameworks currently applied have limited applicability due to,

- The frameworks are mostly applied on general imaging datasets for classification which require large number of labeled images. In case of biomedical images, availability of large datasets labeled from specialist radiologists is a challenging issue.
- Due to limited availability of datasets, the problem of imbalanced classification becomes even more challenging which is crucial for decision making for post diagnosis treatment of patients. Furthermore, recent works mostly address the binary classification problem of presence or absence of ACL tears, wherein the classification task of partial tear presents additional challenge for classification algorithms.

- The DL architectures are mainly trained with heuristic techniques which require theoretical analysis in order to improve feature learning for accurate classification, especially in case of biomedical images where the error margin should be as minimum as possible.
- Lastly medical imaging plays a crucial role in diagnosis, treatment planning, treatment delivery and follow-ups. To increase efficiency, there is an urgent need to integrate DL based automation tools in the form of machine learning, into all stages of the medical imaging pipeline ranging from image acquisition and reconstruction to analysis and interpretation.

In this work we address the above-mentioned gaps with the implementation of a DL classification framework optimally designed and tuned for gray scale MR images obtained at Hospital Kuala Lumpur (HKL) and labeled by expert radiologist for normal, complete and partial ACL tears. Specifically, this work makes following contributions.

- The framework has been optimized successfully on labelled knee MR dataset and an average test set accuracy of 92% has been achieved without adding regularization techniques.
- The unrolling ML-CSC framework with multi-layered iterative thresholding algorithm, its fast version FISTA along with multilayered Basis Pursuit (ML-BP) is demonstrated to perform better than the CNN based classification framework taken as baseline, without increasing network depth.
- The solution to challenging partial ACL tear classification problem, where the classifiers generally do not give good accuracies, is optimized with data augmentation techniques and accuracy of more than 85% on this specific class is achieved for multilayer iterative thresholding algorithm (ML-ISTA) framework, outperforming traditional CNN with same number of parameters.
- Data augmentation technique especially suited for training a DL framework on biomedical images is applied, which shows optimal performance as compared to other transforms used in image classification algorithms.
- Lastly, a classification framework based on a recurrent architecture with the same depth as the generative models described above is trained and analyzed for comparison. The accuracies of all models are compared with CNN, demonstrating the viability and effectiveness of the MR image classification framework.

II. PRIOR AND RELATED WORK

With improvements in hardware for processing of large number of images, it became feasible to train large neural networks for classification tasks on datasets of different sizes. The seminal work of [16] significantly improved state of the art on classification of general images using graphical processing units. The work achieved an error rate of more than 15% on ImageNet dataset. This error rate has been improved significantly since then on general datasets

available for research purpose. In addition to general datasets, P.Rajpurkar *et al.* in [17] implemented a 121 layer deep neural network for radiologist-level pneumonia detection on chest X-rays. The algorithm was trained on large ChestX-ray14 dataset [18]. P. Rajpurkar *et al.* in [19] developed a 33 layer CNN for detecting a wide range of heart arrhythmias from single-lead ECG records. F.Liu *et al.* in [20] used classification performance of DL networks as compared to clinical reports for binary classification (tear or nor tear presence), and concluded that there is no significant difference between the two. The study by P.D. Chang *et al.* in [21], demonstrated the feasibility of a high-performing CNN tool to detect complete ACL injury with over 96% test accuracy for binary classification problem. The study, which excluded cases with partial tear and mucoid pathologies demonstrated the feasibility of high performing CNN tool, with customized CNN architecture and dynamic patched based sampling with five-sliced 3-D input. The results of study in [22] suggested the usefulness of preoperative MRI-detected lateral meniscal extrusion (LME) for estimating lateral meniscus posterior root tear (LMPRT) in injured knees with ACL tear. Although, there is a significant improvement in application to classifications and inverse problems in context of DL architectures, the theoretical foundations of DL largely remain heuristic. One such very useful heuristic technique which is widely applied in DL architectures as regularization to avoid over fitting the model on test data, is dropout. This technique randomly discards activations to improve classification accuracy on tests sets and avoid over-fitting by the learning model. These regularization techniques have been improved recently with proposal of stochastic techniques to further reduce over-fitting by DL networks [23].

The recent research addressing the limitations of DL architectures has focused on the theoretical explanation of the working of deep learning frameworks. In [14], [15], authors elaborated the significance of theoretical understanding of deep learning and proved connection between widely used CNN architectures and celebrated sparse coding theory. The sparse coding theory which has been successfully used in inverse problem in imaging and classification tasks, was shown in [14] to be tightly connected to CNNs. The work established connection between CNN and sparse coding theory and further gave insights to the multilayered version of sparse coding. A further work by [24] pointed out the suboptimal performance of model presented in [14]. The work in [24] analyzed the proposed multilayered basis pursuit in context of combination of synthesis and analysis. Further extending the work on multilayered basis pursuit and its application to explain CNNs and performance on applied problems of classification, J. Sulam *et al.* in [25] introduced a multilayered basis pursuit framework where in an l_1 norm penalty was proposed on intermediate representations of multilayered framework. Reference [25] showed that iterative thresholding algorithms can be used for multilayer basis pursuit and demonstrated the framework effectiveness on classification tasks of general datasets of MNIST, SVHN

and CIFAR-10 with improved performance of thresholding algorithms as compared to the CNNs.

In this work we have implemented an optimal framework for multi-layered basis pursuit algorithms and demonstrate through experiments its applicability to classification of biomedical images. The novel architecture, which is optimized for classification of biomedical images, trained on original dataset of knee MR images, achieves a good average test accuracy of more than 92% and class wise test accuracy of 95%, outperforming traditional CNN without adding regularization parameters and computational complexity. The rest of the paper is organized as follows. Section III gives a brief introduction to clinical background of ACL and its tears, section IV gives brief overview of CNNs, the multilayered sparse coding model and image classification in context of CNN and ML-CSC. Section V gives overviews of iterative thresholding algorithms for single layer basis pursuit and its extended version in multilayer settings along multilayered basis pursuit. Section VI gives the ML-CSC model, its implementations for classification of biomedical image dataset and experimental results of image classification of knee ACL tear.

III. ANTERIOR CRUCIATE LIGAMENT TEAR - BACKGROUND

An anterior cruciate ligament (ACL) is one of the key ligaments that help stabilize the knee joint. These ligaments connect the thighbone (femur) to the shinbone (tibia) (Figure 1). Injuries of ACL are most often a result of low-velocity, noncontact, deceleration injuries and contact injuries with a rotational component. A complete tear is characterized by rupture of the ligaments and partial tear by stretching of the ligaments becoming loose and damaged. The MR images with a partial tear, normal knee, and complete ACL tears are given from dataset used in this work (Figure 2)

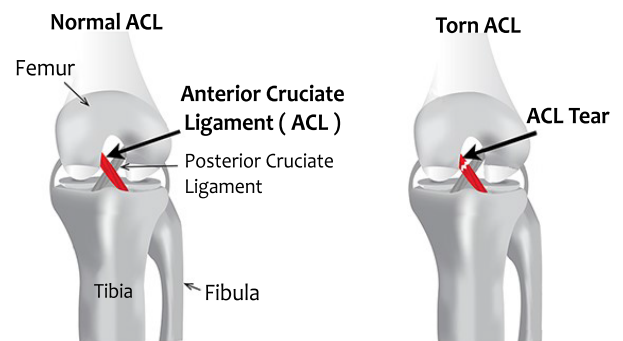


FIGURE 1. Knee anatomy with normal and torn ACL tears (Coronal view).

The diagnosis process involves an emphasis on history and physical examination of affected patients. The Lachman, pivot shift, and anterior drawer tests are three types of physical examinations performed on ACL tear patients for assessment of the injury. Out of these three tests, the anterior drawer test has the highest sensitivity of 94% [26]. MRI examination coupled with physical examination helps clinicians in

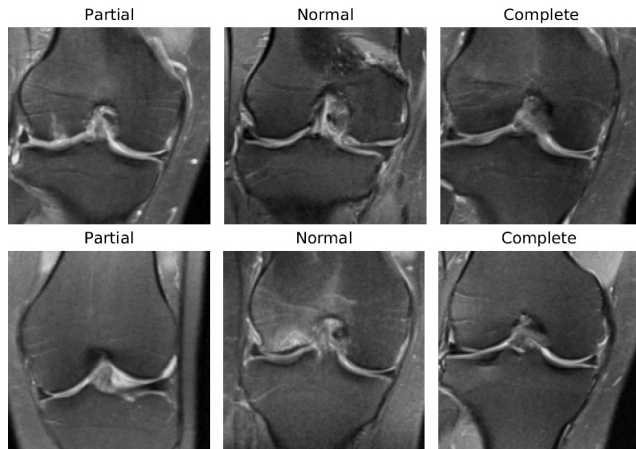


FIGURE 2. Labeled Knee MR images from Knee MR dataset.

identifying ACL tear types in addition to identifying bone bruising, which is present in most of the patients with an ACL tear. Once the ACL tear is diagnosed, the treatment plan is devised by clinicians for rehabilitation or surgical intervention according to patient condition and medical profile. Studies have been reported that, in some cases, an average radiologist must interpret one image every 3-4 seconds in an 8-hour workday to meet the workload demand. Under these conditions, the errors are inevitable for radiology tasks where visual perception and decision making are involved [27]. An integrated AI system in imaging pipeline, which enables the trained radiologist to receive pre-screened images would enable better decision making especially in heavy workloads in addition to helping in the diagnosis of ACL tear of knee injury patients in regions where trained MSK - radiologist are difficult to access. Besides, the ability of machines to scan large amounts of data enables them to generalize the classification algorithms for better decision making.

IV. REPRESENTATION LEARNING AND CLASSIFICATION

The ability of machines learning algorithms that can learn and improve based on experience of the complexity of the problem at hand and their adaptation to that specific problem gives tremendous opportunities in array of applications. The CNNs are fundamental part of representation learning and are briefly explained below.

A. CONVOLUTIONAL NEURAL NETWORKS

The forward pass is the fundamental part of the CNNs, where an input signal X , is convolved with set of learned filters of chosen size giving output as feature maps or kernels. In matrix vector form, this can be written as $W_1^T X$, where W_1 , is a convolutional matrix (transposed) with learned filters as columns with all their shifts. After convolution, a bias term b_1 is added to resulting vector and a nonlinear operation (here Rectified Linear Unit-ReLU) is applied. For a two layers forward pass of CNN, the operation is given by,

$$f(X, [W_i]_{i=1}^2, [b_i]_{i=1}^2) = W_2 = \text{ReLU}(W_2^T \text{ReLU}(W_1^T X + b_1) + b_2), \quad (1)$$

The output of first stage/layer is then treated as input to another stage with convolutional matrix W_2^T and bias term b_2 . The operation is extended up to desired number of layers and feature maps are then used for classification or inverse problems. For the problem of classification, the output of last layer is fed to train a classifier which tries to predict the label $h(X)$ associated with given image X . For given dataset of images $(X_j)_j$, the task of CNN including filters $[W_i]_{i=1}^K$ and biases $[b_i]_{i=1}^K$, parameter of the classifiers U can be written as,

$$\min_{[W_i]_{i=1}^K, [b_i]_{i=1}^K, U} \sum_j l(f(h(X_j), U, f(X_j), [W_i]_{i=1}^2, [b_i]_{i=1}^2)). \quad (2)$$

The task of the optimization algorithm is to minimize the mean of the loss function l .

B. CONVOLUTIONAL SPARSE CODING MODEL-THE MULTILAYERED BASIS PURSUIT

Sparse coding theory works on premise of first learning filters (weights/dictionaries) from given data and then finding their sparse representation from those dictionaries for representation of given images. Once the underlying structure is successfully modeled, the problems of reconstruction on images from noisy measurements, retrieving/reconstructing a signal in compressive sensing domain and classification of test sets on already training dictionaries and sparse maps can be done successfully with the help of different algorithms developed over the years and applied successes fully in different domains. In sparse coding theory, formally, a given signal y admits a sparse representation in terms of a dictionary D , if $y = Dx$, and x is sparse. Given dictionary D , the celebrated basis pursuit problem with l_1 norm penalty is formulated as,

$$\min_x \|x\|_1 \quad \text{s.t.} \quad \|y - Dx\|_2^2. \quad (3)$$

This modeling theory was extended [14] to multilayer settings, providing connection between sparse coding theory and state of the art DL architectures. The traditional sparse coding model assumes the dictionaries without any structure. Whereas in CSC, which is a special form of sparse coding [11], a special structure on learned dictionaries are imposed with filters banded together and concatenated in circulant form. In the multi-layered version of CSC, which is an extended version of CSC, the sparse feature maps thus obtained from one layer are then treated as input to the second layer, and dictionary learning and sparse coding steps are repeated for subsequent layers. The CSC model represents a signal of interest as multiplication of dictionaries D and sparse vectors x . The deep learning problem in context of sparse coding theory, which is shown as theoretical explanation of CNNs [14], can be formulated as follows. For a global signal X , convolutional dictionaries D and sparse vectors x , and k number of layers and cadianity s , the deep pursuit problem is defined as [24]:

$$\text{Find } [x_i]_{i=1}^k \quad \text{s.t.} \quad \|y - D_1 x_1\|_2^2 < \epsilon, \\ x_{i-1} = D_i x_i \quad \forall \quad 2 \leq i \leq k, \\ \|x_i\|_0 \leq s_i \quad \forall \quad 1 \leq i \leq k. \quad (4)$$

A convex relaxation proposed in [25] for deep pursuit problem result in multilayered basis pursuit. For a two-layer model, the problem can be formulated as:

$$\min_x \frac{1}{2} \|y - D_1 D_2 x\|_2^2 + \lambda_1 \|D_2 x\|_1 + \lambda_2 \|x\|_1, \quad (5)$$

In case $\lambda_1, \lambda_2 = 0$ and $\lambda_1 > 0$, the above formulation is equivalent to traditional basis pursuit with global dictionary. With $\lambda_1, \lambda_2 > 0$, analysis priors are imposed on set of sought after representations x with regularized solutions as a result.

C. ML-CSC FOR CLASSIFICATION

Given sparse vectors Γ^* and dictionaries D , the classifications problem can be formulated in deep sparse coding context as:

$$\min_{(D,U)} \sum_j l(f(h(X_j), U, \Gamma^*(X_j), D), \quad (6)$$

where sparse representations Γ are fed to the classifier after dictionary learning, multilayer basis pursuit and training of classifier.

V. ALGORITHMS

DL architectures and algorithms traditionally deal with high dimensional settings where second order methods result in prohibitive computational complexity and slow convergence rates. The proximal gradient descent which uses first order approximations for updating its optimization steps is therefore a suitable choice for multilayer basis pursuit due to its dependence on sparse prior terms instead of the convex term [28]. This algorithm only needs to calculate the sub gradients of convex term, and proximal mapping associated with update depends on sparse prior. The convergence analysis is done in terms of number of iterations of algorithm.

A. LAYERED BASIS PURSUIT

The layered basis pursuit given in [14], addresses sequence of pursuit of the form:

$$\hat{x}_i \leftarrow \arg \min_{x_i} \|\hat{x}_{i-1} - D x_i\|_2^2 + \lambda_i \|x_i\|_1, \quad (7)$$

where $x_0 = y$ and $i = 1$ to k .

These algorithms [29], [30], which present heuristic approximation do not minimize Equation (5) and each layer is required to explain next layer only so cannot be used to generate signal according to multilayer sparse model. Algorithm 1 for layered basis pursuit which seeks sparse

Algorithm 1 ML-BP

Input signal y , dictionaries D_i
 Init Set $\hat{x}_0 = y$,
 1: **for** $i = 1 : k$ **do**
 2: $\hat{x}_i = \begin{cases} \mathcal{H}(D_i^T \hat{x}_{i-1}) \\ P_1(D_i, \hat{x}_{i-1}, \lambda_i) \end{cases}$
 3: return $[\hat{x}_i]_{i=1}^k$ ▷ Set of representations
 4: **end for**

maps x , subject to constraints given in of Equation(4) for P_1 term and thresholding operator \mathcal{H} at each layer of the neural network.

B. ITERATIVE THRESHOLDING ALGORITHMS (ISTA)

Iterative Shrinkage Thresholding Algorithm (ISTA) is a first order method for optimizing functions comprising composite terms originally proposed in [31]. A faster version of this algorithm FISTA, proposed in [32], introduced a momentum term, resulting in improved convergence rates. These algorithms require matrix vector multiplications, therefore are appealing due to low complexity. The ISTA provides convergence in function value in the order of $\mathcal{O}(1/k)$ and its fast version FISTA provides better convergence rate in the order of $\mathcal{O}(1/k^2)$. The proximal gradient method ISTA works by iterating the updates given by the proximal operator. As $g(\cdot)$ in equation (5) is sum of l_1 composite terms so, application of ISTA algorithm is not feasible. Another feasible alternative the generalized LASSO [33] can also be computationally expensive due to requirement of inversions of linear operators during optimization. The iterative algorithm employing re-weighted l_2 norm approaches proposed in [34], for compressive sensing also require iterative matrix inversions and thus is computationally expensive.

Algorithm 2 ISTA

Init $x^0 \in f(x)$
 1: **for** any $k = 0, 1, 2$. **do**
 2: $x^{k+1} = \text{prox}_{\frac{1}{L}} g(x - \frac{1}{L_k} \nabla f(x))$
 3: **end for**

C. MULTI-LAYER ISTA AND FISTA

For a composite model comprising a smooth and convex term $f(x)$ and convex and not necessarily smooth term $g(x)$, the objective function is given by,

$$F(x) = f(x) + g(x),$$

The gradient mapping is the operator given by:

$$G_L^{f,g}(x) = L(x - \text{prox}_{\frac{1}{L}} g(x - \frac{1}{L_k} \nabla f(x))), \quad (8)$$

where L is Lipschitz constant.

The ISTA update step for Equation(8) is given by,

$$x^{k+1} = x^k + \frac{1}{L} G_L^{f,g}(x), \quad (9)$$

The optimization problem of multilayer basis pursuit is given by,

$$\min_x F(x) = f(D_2 x) + g_1(D_2 x) + g_2(x),$$

The sparse representations for second layer are given as,

$$F(x_2) = f(D_2 x_2) + g_1(D_2 x_2) + g_2(x_2),$$

The update for the gradient mapping method is given by:

$$x_2^{k+1} = \text{prox}_{t g_2}(x_2^k - t. G_{1/c}^{f(\cdot), g_1(D_2 \cdot)} x_2^k), \quad (10)$$

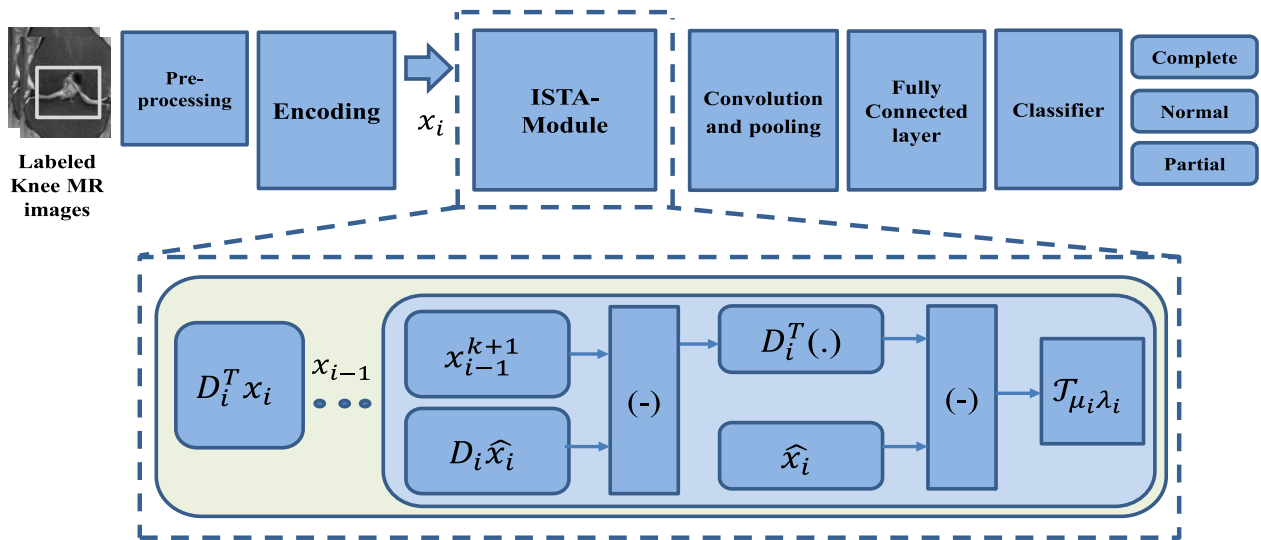


FIGURE 3. Multilayered basis pursuit based framework for MR image classification.

Here c and t are constants with specific bounds for convergence of the subject algorithms. As $g_1(D_2)$ is composite term in Equation (10), in order to avoid calculating its proximal mapping, the term $D_2 x_2$ is approximated with x_1 that is $x_1 = D_2 x_2$. The approximation results in calculation of proximal mapping of x_1 in Equation (10). The update for Equation (10) becomes:

$$x_2^{k+1} = \text{prox}_{t g_2}(x_2^k - t \cdot G_{1/c}^{f(\cdot), g_1(D_2 \cdot)} x_2^k), \quad (11)$$

Consequently, the proximal mapping of composite term after approximation becomes soft thresholding of x_1 which is equal to $x_2 = \mathcal{T}_{t \lambda_1} x_1$. The update step for ML-ISTA after above approximation becomes:

$$x_2^{k+1} \leftarrow \mathcal{T}_{t \lambda_2} [x_2^k - \frac{t}{\mu} D_2^T (x_1^k - \mu D_1^T (D_2 x_1^k - y))]. \quad (12)$$

Algorithms for ISTA and FISTA in multilayer settings are described in Algorithm 3 and Algorithm 4 respectively.

Algorithm 3 ML-ISTA

Input signal y , dictionaries D_i and λ_i ;
 Init Set $x_0^k = y, \forall k$ and $x_L^1 = 0$

- 1: **for** $k = 1:K$ **do**
- 2: $\hat{x}_i \leftarrow D_{(i,L)} x_L^k \quad \forall_i [0, L - 1]$
- 3: **for** $i = 1:L$ **do**
- 4: $x_i^{k+1} \leftarrow \mathcal{T}_{\mu_i \lambda_i}(\hat{x}_i - \mu_i D_i^T (D_i \hat{x}_i - x_{i-1}^{k+1}))$
 ▷ Set of representations
- 5: **end for**
- 6: **end for**

Algorithm 4 ML-FISTA

Input signal y , dictionaries D_i and λ_i ;
 Init: Set $x_0^k = y, \forall k$ and $z = 0$;

- 1: **for** $k = 1:K$ **do**
- 2: $\hat{x}_i \leftarrow D_{(i,L)} z \quad \forall_i [0, L - 1]$
- 3: **for** $i = 1:L$ **do**
- 4: $x_i^{k+1} \leftarrow \mathcal{T}_{\mu_i \lambda_i}(\hat{x}_i - \mu_i D_i^T (D_i \hat{x}_i - x_{i-1}^{k+1}))$
- 5: $t_{k+1} \leftarrow \frac{1 + \sqrt{1 + 4t_k^2}}{2}$
- 6: $z \leftarrow x_L^{k+1} + \frac{t_k - 1}{t_{k+1}} (x_L^{k+1} - x_L^k)$
 ▷ Set of representations
- 7: **end for**
- 8: **end for**

The FISTA algorithms incorporates the momentum term which improves the convergence rate. The framework for classification with iterative thresholding algorithms is given in Figure 3 and pseudocode is presented in Appendix-A. The ISTA module described in Figure 3 computes representations according to Algorithm 3. First the encoded feature maps are backward computed for the three layers framework and the iterations and unfoldings progress according to ML-ISTA algorithm. The (-) sign given in Figure 3 depicts subtraction of resulting representations after convolution and transposed convolution operation are carried out with dictionaries D . The number of unfoldings inside the ISTA module enables the shallow network to increase depth without having any impact on number of parameters.

D. UNFOLDED ITERATIVE ALGORITHMS AS NEURAL NETWORKS

Unfolded iterative algorithms are successfully used in recent research works [35]–[38], for solving sparse recovery problems. To speed up the computational cost associated with

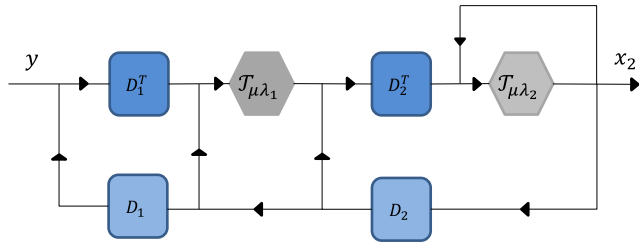


FIGURE 4. A two layer ISTA model as illustrated in [25].

approximation algorithms, the work in [36] showed a combination of optimization and neural networks to produce deterministic functions to successfully approximate parsimonious/sparse models resulting in significant reduction in computational time for applications requiring real-time performance such as image modeling, robust face modeling, audio sources separation and robust speaker recognition. The work in [37], proposed a partial weight coupling structure to learned iterative thresholding algorithms (LISTA) and support selection for improved convergence rate with experimental demonstrations. A two layer ISTA network is given in Figure 4. The classification framework implements a multi-layer ISTA and FISTA framework with two unfoldings.

VI. EXPERIMENTS AND RESULTS ON KNEE MR DATASET

We use dataset of 623 MR images comprising 205 (complete tear), 205 (normal), and 213 (partial tear) images with coronal view. Data collected in the study include adult patients, aged between 18 to 40 years (Male and Female), with Proton density (PD)-weighted images and fat saturation. The images were labeled by certified MSK-radiologist at Hospital Kuala Lumpur (HKL) and have been used in [39], for classification of MR images with CNN. An 80-20 split is applied for train and test. This work does not employ regularization techniques of dropout and batch normalization in order to provide clear experimental setup and demonstrate its effectiveness on a framework with application to biomedical image classification. All algorithms use three convolutional layers, with filter size of 5 in each layer, and number of feature maps of 16,32 and 32 size for layer one, two and three respectively. These parameters have been empirically experimented for optimal performance on this dataset. Similarly, all algorithms use learning rate of 0.001 and batch size of 3. The optimizer parameters of weight decay (an l_2 weight regularization) and learning rate scheduler values have also been kept same for all algorithms. All models have been trained with stochastic gradient descent. Table-1 gives precision, recall, average accuracies and F-1 scores for baseline CNN, and All-Free learning framework and proposed ML-ISTA, ML-FISTA and ML-BP with network unfoldings. The classification metrics of framework with highest average accuracy has been highlighted to emphasize the effectiveness and better accuracy of proposed frameworks as compared to baseline. Class-wise accuracies, average accuracies and test losses of CNN and proposed algorithms are given in Table-2, with emphasis on

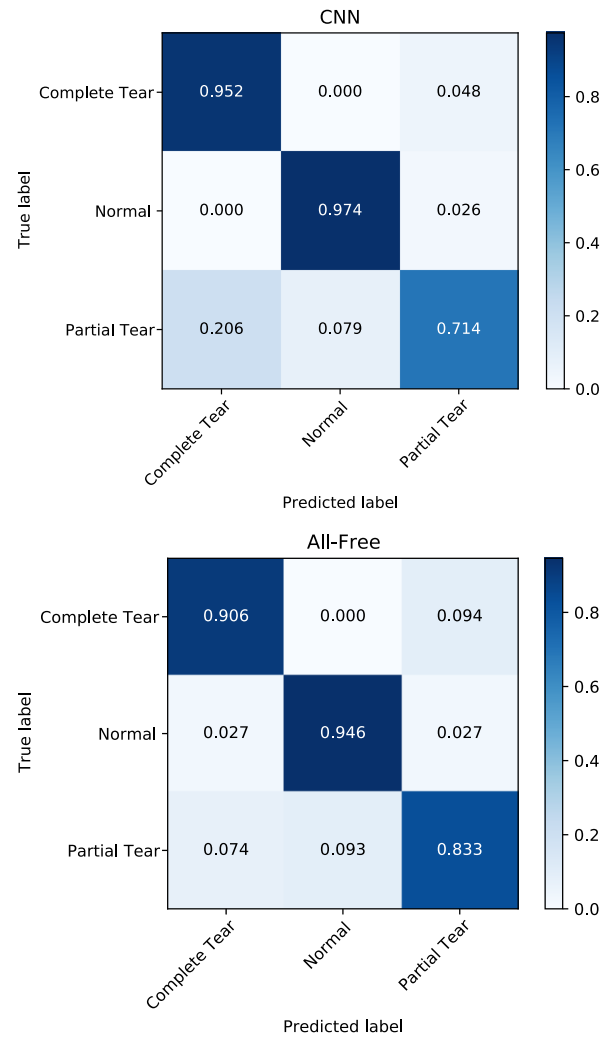


FIGURE 5. Confusion matrices of CNN and All-Free framework.

framework with better classification accuracies on complete ACL tear and partial ACL tear.

A. MULTI-LAYERED ITERATIVE THRESHOLDING ALGORITHMS WITH UNFOLDINGS

The results of classifier based on features extracted by multi-layered iterative thresholding algorithms are given in Figure 6 and Figure 7. The classifier performance is given for two unfoldings (1 and 2) and further increase in unfolding value results in divergence of algorithms. The training accuracy of implemented frameworks is given in Figure 9 for unfoldings 1 and 2. Train losses and validation losses for number of unfoldings are depicted in Figure 10 and Figure 11 respectively. The empirical results show improvement in learning performance of proposed classification framework, sharp decrease in loss curves and better classification accuracy as unfoldings are increased. The ML-ISTA framework with two unfoldings outperforms CNN and ML-FISTA with reasonable margins as shown in Table-1.

TABLE 1. Precision, Recall, Accuracy and F-1 scores of iterative thresholding algorithms with unfoldings.

| Algorithms | Precision | Recall | Accuracy | F1-Score |
|----------------------------|--------------|--------------|--------------|--------------|
| CNN | 87.79 | 90.30 | 88.19 | 88.34 |
| All Free | 88.77 | 89.50 | 89.89 | 89.58 |
| ML-FISTA-unfolding= 1 | 82.29 | 82.18 | 82.00 | 82.14 |
| ML-FISTA-unfolding = 2 | 90.45 | 90.55 | 89.43 | 90.45 |
| ML-ISTA-unfolding = 1 | 87.50 | 88.10 | 88.00 | 87.30 |
| ML-ISTA-unfolding = 2 | 91.00 | 91.29 | 91.96 | 91.03 |
| ML-BP-unfolding = 1 | 92.42 | 92.68 | 92.20 | 92.42 |
| ML-BP-unfolding = 2 | 90.01 | 90.01 | 90.90 | 89.58 |

TABLE 2. Class wise accuracies and test loss.

| Algorithms | Partial ACL tear | Complete ACL tear | Average accuracy | Test Loss |
|-------------------|------------------|-------------------|------------------|---------------|
| CNN | 71 | 95 | 88 | 0.1291 |
| All Free | 71 | 97 | 89 | 0.0814 |
| ML-ISTA | 85 | 93 | 92 | 0.0785 |
| ML-FISTA | 73 | 98 | 89 | 0.1845 |
| Layered BP | 82 | 100 | 92 | 0.0814 |

B. LAYERED BASIS PURSUIT

The layered basis pursuit algorithm is incorporated with same architecture and hyper parameters of baseline CNN and results are added for comparison. The algorithm as proposed in [14], is implemented with each iterative shrinkage iteration unrolling at each individual layer. Whereas in case of ML-ISTA and ML-FISTA, unrolling of iterations is done for entire multi-layer basis pursuit problem. The experimental framework uses two iterations and results are given for comparison with CNN in Figure 8. The ML-BP framework with single unfolding has comparable results with ML-ISTA framework and outperforms baseline with reasonable margins in terms of average accuracy.

C. AN ADAPTIVE LEARNING FRAMEWORK

In addition to three generative models described above, an all free learning framework consisting three layers with same number of feature maps as CNN is also implemented. In this framework, the dictionaries and corresponding representations are adaptively learned for subject dataset. The all free model is trained on same number of layers and relevant parameters for a recurrent architecture. Framework for ML-ISTA, ML-FISTA and layered BP have same number of parameters as of CNN, whereas all free framework has $\mathcal{O}(LK)$ parameters. L is number of layers (here 3 layers are used) and K is the number of unfoldings. The results for classification accuracies for baseline CNN and all-free recurrent architecture are given in Figure 5, Table-1 and Table-2. The number of iterations of the framework slightly improves the accuracies on given classes as observed in ISTA and FISTA architectures.

D. DATA AUGMENTATION

MR data used in the architecture is first center cropped for image size of 320×320 and then normalized with mean

and standard deviation of the dataset. Of many transforms available in Pytorch, the center crop proves very effective in this architecture and learning curve for training dataset follows steady pattern of decrease with increase in number of epochs.

E. CLASS IMBALANCE

When there is an imbalance in classes, DL frameworks give poor classification accuracy for certain classes. For the knee ACL tear classification problem, the classifiers give poor accuracy for the challenging problem of partial ACL tear classification as compared to the other two classes. To circumvent this issue, the oversampling of the partial tear category is done during the training phase. This technique significantly increases test accuracy for partial ACL tear class, when equally applied for CNN as well as the proposed framework, resulting in improved results of the proposed framework with two unfoldings in comparison to other classification algorithms. The accuracies and effectiveness of biomedical classification algorithms can be further improved by incorporating the learning framework with different datasets comprising different age profiles (young and aged population) and gender profiles for a more specific and accurate diagnosis.

F. DEEPER LEARNING ARCHITECTURE AND THE CHALLENGE OF OVER FITTING

The DL architecture chosen for this problem of classification, the size of filters and number of feature maps were observed to be optimal for this dataset of knee MR images. Deeper architectures for CNN, with higher number of feature maps, resulted in under fitting of the learning model. The effect of performance of classifier is given to demonstrate the effectiveness of ISTA, FISTA and layered basis pursuit with unfoldings, which uses the algorithms iterations to extend the depth of network without increasing number of parameters.

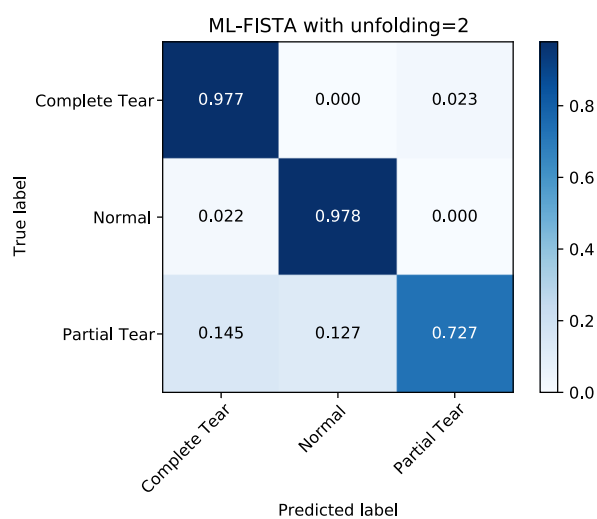
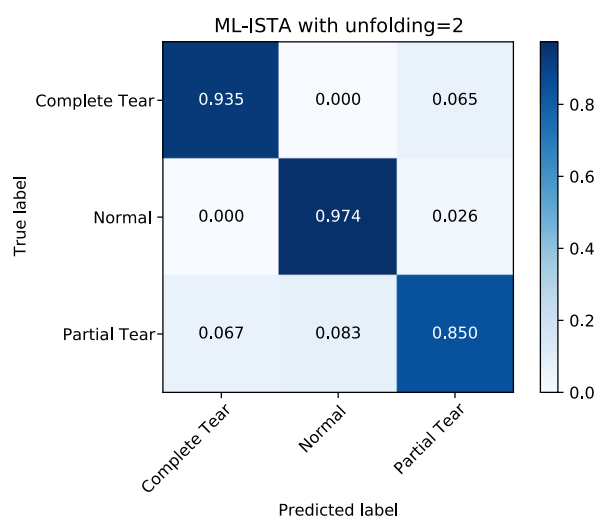
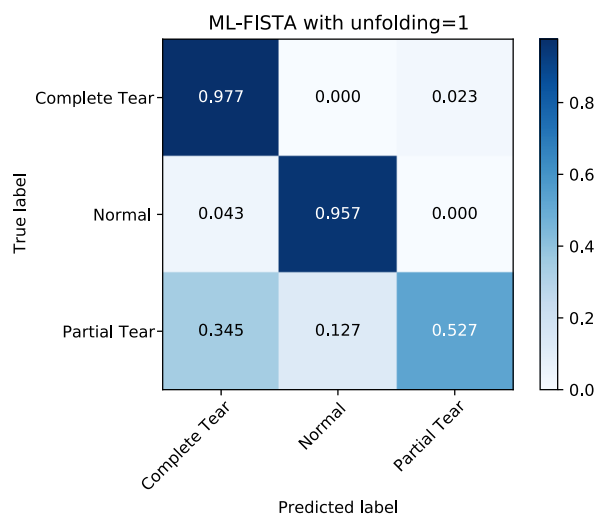
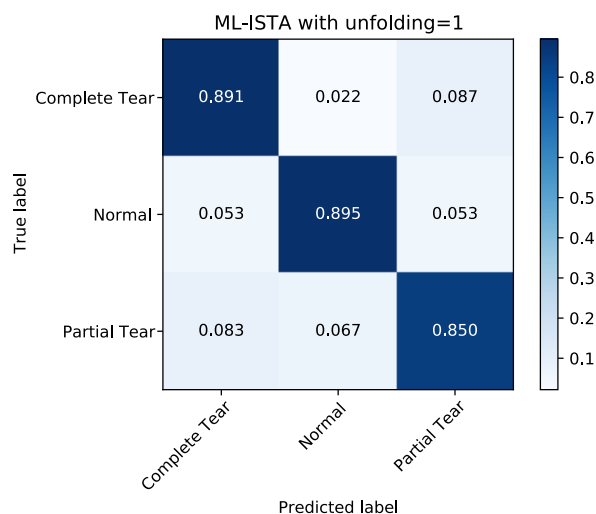


FIGURE 6. Confusion matrices of ML-ISTA with unfoldings.

FIGURE 7. Confusion matrices of ML-FISTA with unfoldings.

With improved results in terms of accuracy of classifiers as shown in confusion matrices given in Figure 6, 7 and 8, the unfolding increases the classification performance of the implemented framework especially on imbalanced class of partial ACL tear. Unlike the CNN, where the increase in depth of neural network results in under fitting of learning model, this framework successfully can be implemented with two unfolding for ISTA (Figure-6), FISTA (Figure 7) and L-BP (Figure 8).

G. DISCUSSION

The challenging problem of identification of partial ACL tear, which is characterized by stretching and weakening of the knee ligaments is diagnosed by clinicians with clinical tests along with MR imaging and arthroscopic examinations. In our work, the MR images with coronal PD were used for training and testing the framework, as the coronal imaging plane is mainly used by radiologists to trace ACL fibers from origin to insertion. In the proposed framework, the partial

tear class is successfully identified with 85% accuracy by ML-ISTA and 82% accuracy by ML-BP. The cases of complete ACL tear class which are characterized by the rupture of the knee ligaments are identified with 98% accuracy by ML-FISTA followed by ML-BP framework which has an accuracy of 97%. Overall, the ML-BP algorithm results in the highest average classification accuracy on all classes as shown in Table 1 and Table 2.

Generally, the presence of notch origin tears makes the diagnosis of a complete ACL tear difficult for radiologists to detect in clinical settings. Another possible reason for misclassification of complete ACL tear class is the mild focal intrasubstance degeneration rather than a complete tear.

As MRI based pathology is localized to small regions of interest, the image crop operation applied in our work significantly improves the learning network training accuracy. This insight can be used to further improve the framework with the incorporation of training images comprising of sagittal and axial planes, which are part of standard knee imaging protocol

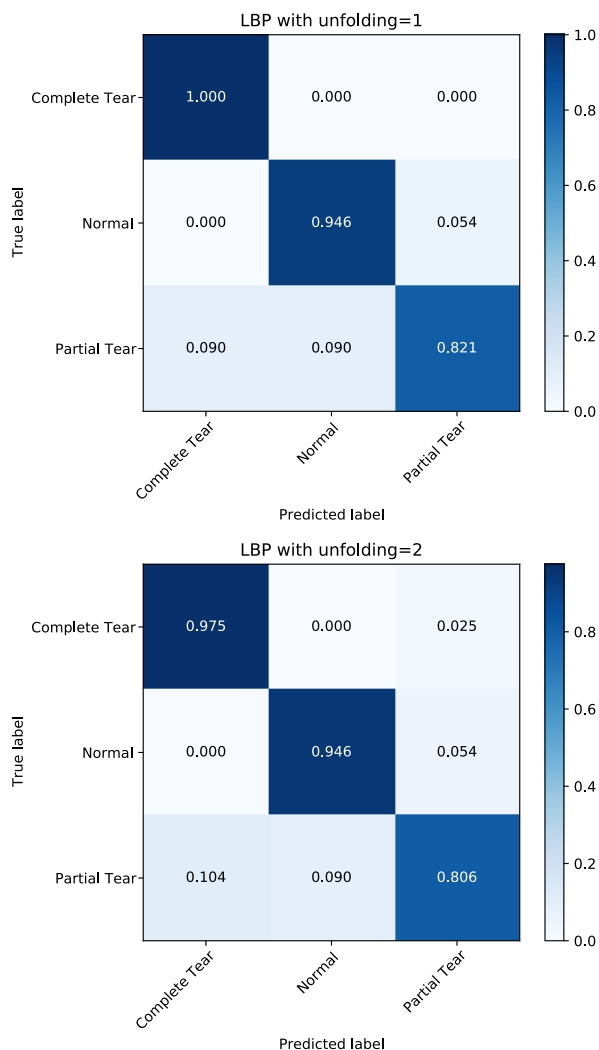


FIGURE 8. Confusion matrices of ML-BP with unfoldings.

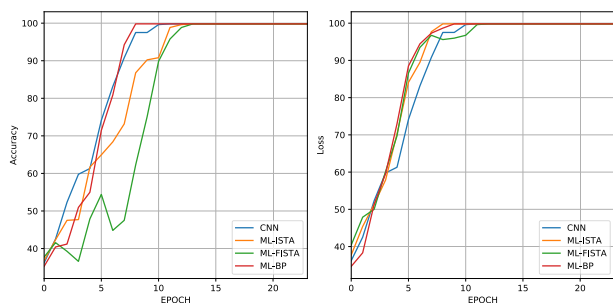


FIGURE 9. Training accuracy for unfolding = 1 (left), and unfolding = 2 (right).

used in clinical applications. The evaluation and interpretation of three-dimensional (3D) data is another unique feature associated with cross-sectional imaging. For musculoskeletal injuries, the combination of 3D contextual information of ligaments in the imaging pipeline is especially useful for

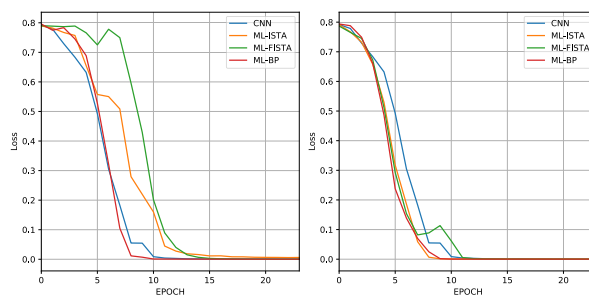


FIGURE 10. Training loss for unfolding = 1 (left), and unfolding = 2 (right).

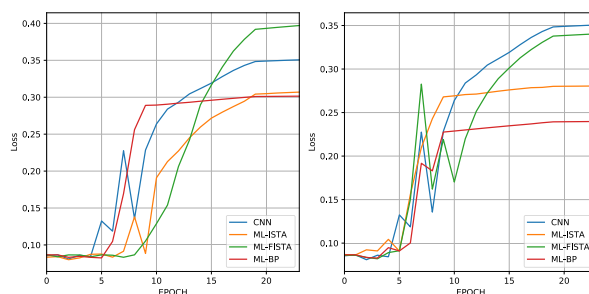


FIGURE 11. Validation loss for unfolding = 1 (left), and unfolding = 2 (right).

diagnosing ACL tears. Furthermore, the performance and generalizability of the framework may be improved with the incorporation of different magnetic field strengths, scanning protocols, and vendors of MRI scanners.

VII. CONCLUSION

We implemented a multi-layered convolutional sparse coding (ML-CSC) framework employing iterative thresholding pursuit algorithms and demonstrated their effectiveness in terms of classification accuracy in comparison to traditional CNN based frameworks. Algorithms of gradient mapping schemes like iterative thresholding algorithm (ISTA), fast iterative thresholding algorithm (FISTA) along with multi-layered basis pursuit were implemented for feature extraction and training of the classifier. The framework was applied to a labeled dataset of knee MR images for classification and accuracies were given for different types of ACL tears. In absence of larger labeled datasets, this work demonstrated the effectiveness of the classification framework’s learning capability with the same number of features as the baseline CNN, and without adding regularization hyperparameters and computational complexity to the neural network architecture. The framework also demonstrated the effectiveness of unfolding on neural networks’ performance, improving classification accuracies on imbalanced classification problem of partial ACL tear. In future work, the research can be extended to design more generalized classifiers using transfer learning, which can adapt to different datasets without requiring training from scratch and use them for improvement in performance of classifiers of biomedical images.

APPENDIX A

Pseudocode for Classification Framework

• Training the Framework with Train Data:
Input: Labeled MR images**Output:** Model containing parameters of neural network for classification

Data augmentation-Centercrop and Normalize dataset

for Unfolding = 1,2 **do**

Find set of representations with CNN,

ML-ISTA, ML-FISTA and ML-BP

end for

• Classify Test set with trained model:
Input: Labeled MR test images, Trained model**Output:** Classification results**for** i = 1:n **do** ▷ n is the size of test set

Calculate Confusion matrix, Test loss, accuracy,

Precision, Recall and F-1 Score

Results

end for

ACKNOWLEDGMENT

The authors would like to thank Electronics Section, Universiti Kuala Lumpur British Malaysia Institute and Radiology department Hospital Kuala Lumpur(HKL) for their support in conducting this research work.

REFERENCES

- [1] T. L. Sanders, H. Maradit Kremers, A. J. Bryan, D. R. Larson, D. L. Dahm, B. A. Levy, M. J. Stuart, and A. J. Krych, "Incidence of anterior cruciate ligament tears and reconstruction: A 21-year population-based study," *Amer. J. Sports Med.*, vol. 44, no. 6, pp. 1502–1507, 2016.
- [2] G. S. Perrone, B. L. Proffen, A. M. Kiapour, J. T. Sieker, B. C. Fleming, and M. M. Murray, "Bench-to-bedside: Bridge-enhanced anterior cruciate ligament repair," *J. Orthopaedic Res.*, vol. 35, no. 12, pp. 2606–2612, Dec. 2017.
- [3] H. Louboutin, R. Debarge, J. Richou, T. A. S. Selmi, S. T. Donell, P. Neyret, and F. Dubrana, "Osteoarthritis in patients with anterior cruciate ligament rupture: A review of risk factors," *Knee*, vol. 16, no. 4, pp. 239–244, Aug. 2009.
- [4] K. Satku, V. Kumar, and S. Ngoi, "Anterior cruciate ligament injuries. To counsel or to operate?" *J. Bone Joint Surg. Brit. Volume*, vol. 68, no. 3, pp. 458–461, 1986.
- [5] J. H. Mink, T. Levy, and J. V. Cruess, "Tears of the anterior cruciate ligament and menisci of the knee: MR imaging evaluation," *Radiology*, vol. 167, no. 3, pp. 769–774, Jun. 1988.
- [6] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019.
- [7] M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: A review," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 85–95, Nov. 2017.
- [8] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 2, pp. 491–503, Feb. 2018.
- [9] J. B. S. Carvalho, J.-M. Moreira, M. A. T. Figueiredo, and N. Papanikolaou, "Automatic detection and segmentation of lung lesions using deep residual CNNs," in *Proc. IEEE 19th Int. Conf. Bioinf. Bioengineering (BIBE)*, Oct. 2019, pp. 977–983.
- [10] Y. Zhou, X. He, L. Huang, L. Liu, F. Zhu, S. Cui, and L. Shao, "Collaborative learning of semi-supervised segmentation and classification for medical images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 2079–2088.
- [11] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, 1st ed. Cham, Switzerland: Springer, 2010.
- [12] J. Wright, A. Y. Yang, A. Ganesh, S. Shankar Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
- [13] W. Luo, J. Li, J. Yang, W. Xu, and J. Zhang, "Convolutional sparse autoencoders for image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 7, pp. 3289–3294, Jul. 2018.
- [14] V. Pappas, Y. Romano, and M. Elad, "Convolutional neural networks analyzed via convolutional sparse coding," *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 2887–2938, 2017.
- [15] V. Pappas, Y. Romano, J. Sulam, and M. Elad, "Theoretical foundations of deep learning via sparse representations: A multilayer sparse model and its connection to convolutional neural networks," *IEEE Signal Process. Mag.*, vol. 35, no. 4, pp. 72–89, Jul. 2018.
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [17] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng, "CheXNet: Radiologist-level pneumonia detection on chest X-Rays with deep learning," 2017, *arXiv:1711.05225*. [Online]. Available: <http://arxiv.org/abs/1711.05225>
- [18] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2097–2106.
- [19] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," 2017, *arXiv:1707.01836*. [Online]. Available: <http://arxiv.org/abs/1707.01836>
- [20] F. Liu, B. Guan, Z. Zhou, A. Samsonov, H. Rosas, K. Lian, R. Sharma, A. Kanarek, J. Kim, A. Guermazi, and R. Kijowski, "Fully automated diagnosis of anterior cruciate ligament tears on knee MR images by using deep learning," *Radiol. Artif. Intell.*, vol. 1, no. 3, May 2019, Art. no. 180091.
- [21] P. D. Chang, T. T. Wong, and M. J. Rasiej, "Deep learning for detection of complete anterior cruciate ligament tear," *J. Digit. Imag.*, vol. 32, no. 6, pp. 980–986, Dec. 2019.
- [22] Y. Kamatsuki, T. Furumatsu, M. Fujii, Y. Kodama, S. Miyazawa, T. Hino, and T. Ozaki, "Complete tear of the lateral meniscus posterior root is associated with meniscal extrusion in anterior cruciate ligament deficient knees," *J. Orthopaedic Res.*, vol. 36, no. 7, pp. 1894–1900, Jul. 2018.
- [23] N. Khan, J. Shah, and I. Stavness, "Bridgeout: Stochastic bridge regularization for deep neural networks," *IEEE Access*, vol. 6, pp. 42961–42970, 2018.
- [24] A. Aberdam, J. Sulam, and M. Elad, "Multi-layer sparse coding: The holistic way," *SIAM J. Math. Data Sci.*, vol. 1, no. 1, pp. 46–77, Jan. 2019.
- [25] J. Sulam, A. Aberdam, A. Beck, and M. Elad, "On multi-layer basis pursuit, efficient algorithms and convolutional neural networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 8, pp. 1968–1980, Aug. 2020.
- [26] H. Makhmalbaf, A. Moradi, S. Ganji, and F. Omidi-Kashani, "Accuracy of Lachman and anterior drawer tests for anterior cruciate ligament injuries," *Arch. Bone Joint Surg.*, vol. 1, no. 2, p. 94, 2013.
- [27] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. W. L. Aerts, "Artificial intelligence in radiology," *Nature Rev. Cancer*, vol. 18, no. 8, pp. 500–510, 2018.
- [28] A. Beck, *First-Order Methods in Optimization* (MOS-SIAM Series on Optimization), vol. 25. Philadelphia, PA, USA: SIAM, 2017.
- [29] J. Zhang and B. Ghanem, "ISTA-net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1828–1837.
- [30] C. Murdock, M. Chang, and S. Lucey, "Deep component analysis via alternating direction neural networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 820–836.
- [31] I. Daubechies, M. DeFrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, 2004.

[32] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, Jan. 2009.

[33] R. J. Tibshirani and J. Taylor, "The solution path of the generalized lasso," *Ann. Statist.*, vol. 39, no. 3, pp. 1335–1371, Jun. 2011.

[34] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 3869–3872.

[35] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. 27th Int. Conf. Mach. Learn.*, 2010, pp. 399–406.

[36] P. Sprechmann, A. M. Bronstein, and G. Sapiro, "Learning efficient sparse and low rank models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1821–1833, Sep. 2015.

[37] X. Chen, J. Liu, Z. Wang, and W. Yin, "Theoretical linear convergence of unfolded ISTA and its practical weights and thresholds," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 9061–9071.

[38] P. Ablin, T. Moreau, M. Massias, and A. Gramfort, "Learning step sizes for unfolded sparse coding," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 13100–13110.

[39] M. H. Razali, S. M. Sazwan, M. Mahmood, D. Nazri, J. Ali, and M. Z. Ayob, "Anterior cruciate ligament (ACL) coronal view injury diagnosis system using convolutional neural network," in *Proc. 2nd Int. Conf. Electron. Electr. Eng. Technol.*, Sep. 2019, pp. 118–122.



ADNAN UMAR KHAN received the B.S. degree in electrical and electronic engineering from Eastern Mediterranean University, Cyprus, in 1994, the M.S. degree in communication system from the University of Portsmouth, U.K., in 1995, and the Ph.D. degree from De Montfort University, U.K. He is currently working as an Assistant Professor with the Department of Electrical Engineering, International Islamic University, Islamabad, Pakistan.



ABDUL WAHID received the B.S. degree in electronic engineering from NED University Karachi, in 1999, and the M.S. degree in telecom engineering from U.E.T Peshawar, in 2007. He is currently pursuing the Ph.D. degree in electrical engineering with International Islamic University Pakistan. He has been an Assistant Professor with the Balochistan University of Information Technology, Engineering, and Management Sciences, Quetta, since January 2013. His research interests

include sparse signal processing, computer vision, and deep learning for biomedical imaging applications.



MUKHTAR ULLAH received the B.S. degree (Hons.) in electrical engineering from the University of Engineering and Technology Peshawar, Pakistan, in 1999, the M.Sc. degree in advanced control and systems engineering from the University of Manchester, U.K., in 2002, and the Ph.D. degree in systems biology and bioinformatics from the University of Rostock, Germany, in 2009. His current research interests include stochastic modeling, signal processing, machine learning, systems and controls, and interdisciplinary research. He is currently a Professor with FAST National University Islamabad.

with FAST National University Islamabad.



JAWAD ALI SHAH (Senior Member, IEEE) received the B.S. degree (Hons.) in electrical engineering and the M.S. degree (Hons.) in telecommunication engineering from the University of Engineering and Technology Peshawar, Pakistan, in 2001 and 2007, respectively, and the Ph.D. degree from International Islamic University Islamabad, Pakistan, in 2015. He is currently an Assistant Professor with the British Malaysian Institute, UniKL Malaysia. His current research

interests include sparse signal processing, compressed sensing, and machine learning. He won two gold medals and presidential award in his academic career. He was the recipient of a six-month fellowship by Higher Education Commission Pakistan under the International Research Support Initiative Program (IRSIP) to pursue his research work at UC Denver/Anschutz Medical Campus, Aurora, CO, USA.



MOHD ZAKI AYOB (Senior Member, IEEE) received the B.Eng. degree (Hons.) in electronic and computer from Monash University Australia, in 1997, the M.Sc. degree in mechatronics (signal and systems) from IIUM, Malaysia, in 2006, and the Ph.D. degree from the University of York, U.K., in 2014. His research interests include biomedical image processing, machine learning, pattern recognition, and AI-based agricultural solutions. He is currently a Senior Lecturer

with UniKL BMI, Malaysia.

...