IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 9, 2020, accepted September 18, 2020, date of publication September 23, 2020,
date of current version October 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3026080

Efficient Lung Nodule Classification Using Transferable Texture Convolutional Neural Network

IMDAD ALI[1,2], MUHAMMAD MUZAMMIL[2], (Member, IEEE), IHSAN UL HAQ[2], (Member, IEEE), MUHAMMAD AMIR[2], (Member, IEEE), AND SUHEEL ABDULLAH[2], (Member, IEEE)

[1]National Center for Physics, Islamabad 44000, Pakistan
[2]Faculty of Engineering and Technology, International Islamic University, Islamabad, Islamabad 44000, Pakistan

Corresponding author: Muhammad Muzammil (m.muzammil@iiu.edu.pk)

This work was supported in part by the Higher Education Commission under Grant 20-4849/NRPU/R&D/HEC/14/1215.

ABSTRACT Lung nodules are vital indicators for the presence of lung cancer. An early detection enhances the survival rate of the patient by starting treatment at the right time. The detection and classification of malignancy in Computed Tomography (CT) images is a very time-consuming and difficult task for radiologists which lead the researchers to develop algorithms for Computer-Aided Diagnosis (CAD) systems to mitigate this burden. The performance of CAD systems is continuously improving by using various deep learning techniques for screening of lung cancer. In this paper, we proposed transferable texture Convolutional Neural Networks (CNN) to improve the classification performance of pulmonary nodules in CT scans. An Energy Layer (EL) is incorporated in our scheme, which extracts texture features from the convolutional layer. The inclusion of EL reduces the number of learnable parameters of the network, which further reduces the memory requirements and computational complexity. The proposed model has only three convolutional layers and one EL, instead of pooling layer. Overall proposed CNN architecture comprises of nine layers for automatic feature extraction and classification of pulmonary nodule candidates as malignant or benign. Furthermore, the pre-trained model of proposed CNN is also used to handle the smaller dataset classification problem by using transfer learning. This work has been evaluated on publicly available LIDC-IDRI and the LUNGx Challenge database through different evaluation matrices, such as; the accuracy, specificity, error rate and AUC. The proposed model is trained by six-fold cross-validation and achieved an accuracy score of $96.69\% \pm 0.72\%$ with only $3.30\% \pm 0.72\%$ error rate. Whereas, the measured AUC and recall is $99.11\% \pm 0.45\%$ and $97.19\% \pm 0.57\%$, respectively. Moreover, we also tested our proposed technique on the MNIST dataset and achieved state-of-the-art results in terms of accuracy and error rate.

INDEX TERMS Computed tomography, cancer detection, computer aided diagnosis, image classification, machine learning, transfer learning, lung nodule, CNN, LIDC-IDRI, LUNGx challenge.

I. INTRODUCTION

Lung cancer is becoming one of the predominant risks to human health all over the world [1]. During 2018 about 9.6 million people expired in the world due to cancer, whereas, more than 1.7 million people expired due to lung cancer which is approximately 18% of total cancerous deaths. Meanwhile, more than 2 million lung cancer cases were reported in year the 2018 [2]. A report published by the American Cancer Society reveals that the highest death rate of lung cancer is 26% and five years' survival rate is only 18% [2]. The low survival rate is due to the detection of lung cancer in advanced stages as the symptoms are not prominent in early diagnostics. Therefore, early detection of lung cancer is essential to improve the survival rate up to 90%. Primarily, lung cancer is investigated by examination of radiography scans such as X-ray, MRI or Computed Tomography (CT) scan [3]. The key effectiveness of radiography screening depends upon the radiologists who identify the suspicious lesions in the form of lung nodules. This challenging

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

task becomes considerably crucial, specifically for the small lung nodules. From the literature, it is revealed that approximately 68% of lung nodules are correctly detected by single radiologist which can be further improved up to 82% by taking the opinion of second radiologist [4]. Therefore, Computer-Aided Diagnostic (CAD) systems are designed to minimize this load on the radiologists. The most recent CAD systems provide assistance in the screening process by classifying the lung nodule as malignant or benign [2]. The most critical part of nodule classification is the salient feature extraction of malignant and benign nodules because of less distinct properties [5]. Extracted features include size (diameter or volume), shape, texture, and morphology, as well as the volume growth rate of the nodule, with the passage of time.

In recent years, texture features have attained great attention in image classification [6] and also for lesion classification of medical diagnostics [7], [8]. Meanwhile, the deep learning techniques, especially Convolutional Neural Networks (CNNs) have been used recently with some promising results for lung nodule classification [9]–[11]. The CNNs utilize convolutional layers to extract features. The complex features are extracted by the last convolutional layer which are utilized in fully connected layers to extract the complete shape information. The features such as edges are extracted by the first convolutional layers. Whereas, the middle pooling and convolutional layers extract features with considerable complexity. For the texture analysis, the entire object and its complex features are not much useful as compared to the recurring patterns of lower complexity, whereas, the dense features of intermediate convolutional layers accurately represent the texture of the object. Therefore, it is feasible for a classic CNN to explore the texture image properties most efficiently, without changing the architecture [12].

We intended to build a CNN which is capable of learning the texture features and then classifying the lung nodules in CT images, as used in [12], for texture image classification. Therefore, we introduced a texture descriptor named as Energy Layer (EL) after the convolutional layer. We enabled the forward and backward propagation to learn the texture features during the training process. Moreover, the Transfer Learning (TL) technique is also used to investigate the issues of small labeled medical image dataset by using our pre-trained model, therefore, the proposed techniques is named as transferable texture CNN. We also showed that our texture CNN achieved better classification performance on lung CT images with less number of learnable parameters and neurons. The proposed architecture is trained and tested by using six-fold cross-validation for binary classification problem of lung nodule malignancy. Furthermore, the proposed model is also tested on MNIST dataset [13], [14] for classification of hand written digits to ensure its effectiveness. The main contributions of this work are summarized as fallows:

- We proposed the texture CNN for lung nodule classification problem on two medical image dataset; LIDC-IDRI and LUNGx Challenge. The classification accuracy was achieved up to 96.69% for LIDC-IDRI dataset.
- We utilized EL in the proposed texture CNN model, which preserves the texture information, reduces the output vector size and also learns the parameters during forward and backward propagation and hence, increases the overall learning capability of the model.
- We also proposed the TL based model which utilized LIDC-IDRI as source task and LUNGx Challenge dataset as target task. The classification accuracy of LUNGx Challenge dataset was 86.14% without TL which was further improved to 90.91% by using the proposed TL base model.
- The classification accuracy of our proposed texture CNN is 96.69% on LIDC-IDRI dataset.

The rest of this paper is organized as follows: in section II the latest research work is discussed and section III describes the data and methods which include a brief database description and framework of our proposed lung nodule classification scheme, including the importance of the CNN model for nodule classification. Experiments and results are presented in section IV, whereas, section V presents the conclusion of this work.

II. RELATED WORK

A. TRADITIONAL METHODS FOR NODULE CLASSIFICATION

A quick brief review of the literature presented in this subsection shows that various techniques have been proposed for malignancy detection and classification of lung nodules by using traditional methods. For example; Narayanan *et al.* [15] proposed a traditional approch for lung nadule detection and classification by using LUNA16 dataset at various slice thicknesses. They also analyzed the performance of feature selection methods for SVM using same dataset [16]. Sheway *et al.* [17] combined geometric and histogram features for nodule classification with linear SVM, logistic regression, k-Nearest Neighbor (kNN), random forest, and AdaBoost classifiers. Firmino *et al.* [18] extracted the features by using a Histogram of Oriented Gradients (HOG) and watershed techniques and then used a rule-based classifier and SVM for nodule classification. They used the LIDC-IDRI database [19] and achieved accuracy and recall score of 97% and 94.4%, respectively. Farag *et al.* [20] utilized the feature fusion concept for lung nodule classification. They extracted three features by using signed distance transform shape-based feature descriptor, multi-resolution Local Binary Pattern (LBP) and Gabor filter. After that, kNN and SVMs are used for nodule malignancy classification. Shaffie, *et al.* [21] utilized higher-order Markov Gibbs Random Field (MGRF) technique and 3D HOG filter to extract features. Afterwards, they classified the nodules with a stacked auto-encoder after fusion of the extracted features. They used the LIDC-IDRI database and achieved accuracy and recall score of 93.12% and 92.47%, respectively.

B. DEEP CNN METHODS FOR NODULE CLASSIFICATION

Many researchers employed deep CNN for lung nodule classification inspired by the recent achievements of deep CNN on various image classification benchmarks, like ImageNet and MS COCO database [22]. Shen *et al.* [10] proposed a Multi-Crop CNN (MC-CNN) to acquire nodule salient information by cropping specific regions of convolutional feature maps. MC-CNN architecture achieved 87.14% nodule classification accuracy for the LIDC-IDRI database. Nasrullah *et al.* [4] proposed a Customized Mixed link Network (CMixNet), deep CNN based model for lung nodule detection and classification. The network was trained and tested on the LIDC-IDRI database and achieved 94% recall and 91% specificity. Al-Shabi *et al.* [23] proposed Gated-Dilated (GD) Networks for lung nodule malignancy classification. The GD network model utilized various dilated convolutions for scale variation instead of max-poolings. The classification performance is evaluated on the LIDC-IDRI database achieved the accuracy and AUC score of 92.57% and 95.14%, respectively. A 2D DCNN architecture with 15 layers is proposed by Tran *et al.* [24] for automatic feature extraction and nodule classification on the LIDC-IDRI dataset. They utilized focal loss function in the training process and improved the classification accuracy up to 97.2%. Xie *et al.*, proposed a state-of-the-art feature fusion and fused texture, shape and deep model-learned information (Fuse-TSD) techniques for nodule classification [25]. They performed classification task with the AUC score of 96.65% on the LIDC-IDRI dataset.

C. SMALL DATABASE ISSUES IN LUNG NODULE CLASSIFICATION

CNNs gradually construct higher-level features from the group of pixels normally found in medical images. However, they are able to extract these features more effectively when more and more training examples are available. In addition to the privacy issues with medical images, the lack of labeled medical database affects the adaption of CNN. Therefore, several efforts have been made to overcome the problem of small medical images data by using transfer learning techniques [26], [27]. Zhao et. al. used transfer learning and fine-tuning scheme for lung nodule classification task. They fine-tunned all the layers of their four CNN architectures, which were pre-trained over the natural images [28]. The multi-view knowledge-based collaborative (MV-KBC) deep model is proposed by Xie *et al.* [29] for malignancy classification. They achieved an accuracy score of 91.60% on the LIDC-IDRI database. Xie *et al.* further extended the same idea with the name of multi-view knowledge-based semi-supervised adversarial classification (MK-SSAC) [30] and improved the accuracy score up to 92.53%. Dey *et al.* [31] proposed four two-pathway Convolutional Neural Networks (CNN) and achieved state-of-the-art accuracy on the LIDC-IDRI database. Furthermore, they use their pre-trained model to handle small dataset problems by using transfer learning and achieved a good accuracy score on their own lung database.

III. DATA AND METHOD

We proposed the texture CNN and also implemented a deep feature TL technique for lunge nodule malignancy classification in CT images. The performance evaluation of both the techniques is done by using publicly available LIDC-IDRI and LUNGx Challenge database. In this section, we discussed the image augmentation technique first, which is used to increase the size of small biomedical images dataset to meet the training requirements of the proposed CNN model and then database and patch generation technique is described. Afterwards, the architecture of our proposed CNN and TL methodology are discussed.

A. LIDC-IDRI DATABASE

The LIDC-IDRI is a publicly available database that contains 244,527 thoracic CT scan images of the 1,010 cases. The x and y-axis coordinates and the boundary information of each nodule are available in associated extensible markup language (XML) annotation files. The XML files also contain semantic diagnostic features which were marked by four experienced thoracic radiologists. They graded each feature from 1 to 5 annotations. [32]. We utilized available XML files and an annotation list provided in [33] to decide which annotation is assigned to the related nodule. Radiologists classified the degree of malignancy for each pulmonary nodule from 1 to 5 categories, which are given in Table 1. In this work, the first three categories (1-3) are recognized as benign (Class 0), whereas, the latter two categories (4,5) are identified as malignant (Class 1).

TABLE 1. Degree of malignancy for each pulmonary nodules.

Occurrence of Cancer	Degree of malignancy
Highly unlikely for cancer	1
Moderately unlikely for cancer	2
Indeterminate likelihood	3
Moderately suspicious for cancer	4
Highly suspicious for cancer	5

B. LUNGx CHALLENGE DATABASE

This database was introduced for nodule classification instead of nodule detection. Therefore, it was mainly focused on the automatic classification of lung modules as malignant or benign in CT images. The LUNGx challenge has a set of calibration and testing scans with online available CSV files containing nodule locations. The set of calibration has ten scans (five females, five males). Out of ten calibration scans, five contain one confirmed benign nodule, and the other five contain one pathology-confirmed malignant nodules. Whereas, the test set has 60 scans which have total 73 nodules. Out of these 60, 13 scans have two nodules. The total of 60 test scans, 23 males and 37 females contained 37 benign and 36 malignant nodules [34], [35].

The LUNGx challenge database consists of a single trans-axial series with full thoracic coverage for each individual case. All scans have been obtained on Philips Brilliance scanners with a "D" (over enhancing) and each scan has 1mm slice thickness. The LUNGx challenge has 22,489 CT images which are in Digital Imaging and Communication in Medicine (DICOM) format. Each image file has a Unique Identifier (UID) that is assigned according to the DICOM standard. To achieve a proper anatomy-based sequencing of the images, the slice number is acquired from the DICOM tag (0020,0013).

C. IMAGE AUGMENTATION

The huge amount of sample data can effectively improve the deep CNN training and testing accuracy by reducing the loss function, and ultimately improving the robustness of networks. Image augmentation is a very good technique to boost the performance of a deep network with very small training data. Image augmentation artificially creates training images by using different image processing operations, such as; translation, resize, random rotation, flips and shear, etc. In this paper, the size of the dataset $\mathcal{D} = \{X_i : 1 \leq i \leq N\}$ is increased by using translation, random rotation, and flip image processing operations to create artificial training images for our proposed Deep CNN. Where N is the total number of images.

D. PATCH GENERATION AND DATA ENHANCEMENT

The LIDC-IDRI and LUNGx challenge database comprises of a heterogeneous set of scans that are acquired by using various reconstruction and acquisition parameters. All the slices are in the DICOM format, having a size of 512 × 512 at a pixel depth of 16 bits. To normalize the pixels, all CT images are first converted to Hounsfield (HU) scales by using the available information of the series header (0028, 1052) and (0028, 1053) in the DICOM and then transformed to a range of (0, 1) from (-1000, 500 HU). After this, the image patches are created in two phases. In the first phase, Region of Interest (ROI) around the nodule is extracted by acquiring the central coordinates (x, y, z) and slice number of malignant and benign nodules from the associated XML file. Then we acquired the voxel coordinates by taking some pixels around the central coordinates with respect to slice thickness. The nodule size is between 3mm to 30mm and slice thickness varies from 0.6mm to 5mm for the LIDC-IDRI database. In the second phase, we extracted all the patches by using voxel coordinates which were extracted in the first phase. We used the same central coordinates (x, y) for each slice during the extraction of every patch. The patch extraction process is illustrated in Figure 1.

In this way, a total of 19,388 patches of size 64 × 64 were extracted from 1,010 cases of LIDC-IDRI database for benign and malignant nodules and named as class 0 and class 1, respectively. Similarly, for the LUNGx Challenge database, we acquired 480 patches for class 1 and 663 patches for class 0.

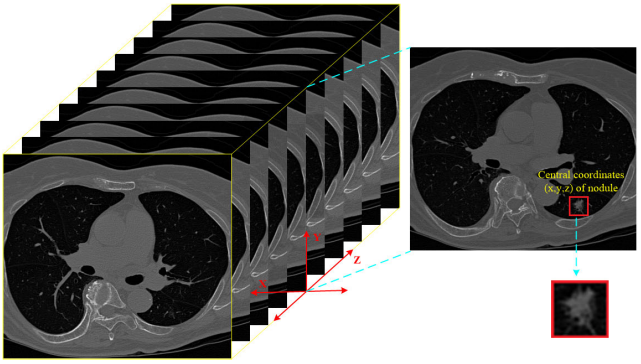

FIGURE 1. Patch extraction process.

E. PROPOSED TRANSFERABLE TEXTURE CNN ARCHITECTURE

Keeping in view the following three essential features of the image, deep CNN has been developed. First, some discriminational patterns having a very small size than the actual image, but if their size equals to the size of the convolution filter mask, then the said patterns can be found by the convolution filter. Second, some shape or patterns are available in different areas of the image, such patterns can also be identified by the convolution of the complete input image. Third, the sub-sampling pixels are very critical for the max-pooling layer and they do not alter the shape of the input image. These pixels are utilized in biomedical image classification. Figure 2 shows the overall architecture of the proposed texture CNN for lung nodule classification.

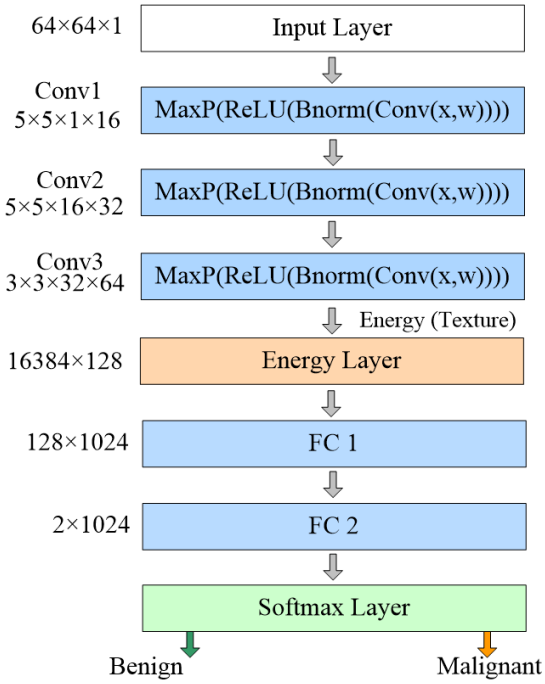

FIGURE 2. Proposed transferable texture cnn architecture.

The proposed CNN has two convolutional layers, each followed by the normalization and the pooling layers. Whereas,

TABLE 2. Layer-wise architecture details of proposed texture CNN.

Layer	Input size	Kernal	Stride	Padding	Output size	Learnable parameters
Conv1	64×64×1	5×5	[1 1]	[1 1 1 1]	62×62×16	Bias: 1×1×16,Weights: 5×5×1×16
Pool1	62×62×16	2×2	[2 2]	[1 1 1 1]	32×32×16	Weights:0
Conv2	31×31×16	5×5	[1 1]	[1 1 1 1]	30×30×32	Bias: 1×1×32, Weights: 5×5×16×32
Pool2	30×30×32	2×2	[2 2]	[1 1 1 1]	16×16×32	Weights:0
Conv3	16×16×32	3×3	[1 1]	[1 1 1 1]	16×16×64	Bias: 1×1×64, Weights: 3×3×32×64
EL	16×16×64	NA	NA	NA	128×1	Bias: 128×1,Weights: 128×16384
Dropout	128×1	NA	NA	NA	128×1	NA
FC1	128×1	NA	NA	NA	1024×1	Bias: 1024×1, Weights: 1024×128
Dropout	1024×1	NA	NA	NA	1024×1	NA
FC2	1024×1	NA	NA	NA	2×1	Bias: 2×1,Weights: 2×1024

the third convolutional layer drives the EL. Finally, the softmax is utilized with the fully connected layer for the classification of lung nodules. Moreover, the layer-wise dimensional details are given in Table 2 which include network layer details like kernel, stride, and padding for each layer.

The input and output dimensions of each layer are also mentioned in Table 2. We used the following mathematical relation to compute the output size of any convolutional layer:

$$OutputSize = \frac{S_i - S_o + 2P}{\zeta + 1} \tag{1}$$

where P is padding, S_i is input size, S_o is the filter size and ζ is the value of stride.

F. CONVOLUTIONAL LAYERS AND ENERGY LAYER

Only the three convolutional layers are used in the proposed model. The kernel size for the first two layers is 5×5, whereas, the output channels are 16 and 32, respectively. The third convolutional layer is considered as intermediate layer to extract the texture features. It has 64 output channels and 3×3 kernel size. The number of learnable parameters for convolutional layers are only 31,744 which are computed by using the following mathematical expression.

$$\theta_C = (S_k \times \zeta + 1) \times N_c \tag{2a}$$
$$\theta_C = N_k \times \zeta \times N_c + N_c \tag{2b}$$

where θ_C is learnable parameters of CNN layer, S_k is kernel size, N_c is number of channels and ζ is stride.

Each convolutional layer computes the output of neurons that are connected to the input and computation is a dot product among their weights and a small area of input where it is connected. The first convolutional layer produces output in volume of $32 \times 32 \times 16$ with 16 kernels. Let χ be an input feature map and ω be the weights, then the output of the neurons at first convolutional layer is given by equation 3.

$$Y^k = f(\sum_k \chi^k * \omega^k + b^k) \tag{3}$$

where Y^k is the output feature map of the convolutional layer for *kth* input and b is the bias term, whereas, $*$ represents the 2D convolution operation. The CNN usually combines the dense orderless features by sharing the weight of the convolutional layer. These features are combined within the CNN for the classification of lung nodule images. Therefore, an energy descriptor is desired at the output of the last convolutional layer, which can learn the texture features during forward and backward propagation. Keeping in view the requirement of energy descriptor, an energy layer is incorporated after the third convolutional layer, which works as the dense orderless texture descriptors. The connection between the EL and the last convolutional layer is given by the equation 4.

$$E(\chi, \theta) = \sigma(\sum_{i=1}^{n} \omega_i^T \chi_i + b) \tag{4}$$

where $E(\chi, \theta)$ is the output of EL, n is the number input connections and ω is the weight vector of EL which is randomly initialized during the start of training. The interconnections between the EL and the FC layers are much smaller as compared to the interconnections of the last classic convolutional layer, which leads to the reduction of the learnable parameters. Furthermore, EL preserves the energy/texture information of previous layer and also learns during forward and backward propagation. Therefore, the EL enhances overall learning capability of network in addition to reduction of vector size for next fully connected layer. This also reduces the complexity of the proposed network without compromising the accuracy. We compared the learnable parameters of the proposed CNN with EL and without EL structure. The learnable parameters of the EL are computed by using equation 5.

$$\theta_{EL} = \delta^n \times \delta^{n-1} \tag{5}$$

where θ_{EL} is learnable parameters of EL, δ^n is the neurons of current fully connected layer and δ^{n-1} is neurons of the previous fully connected layer. Then we computed the learnable parameters of proposed CNN with and without EL, that are, 2, 263, 170 and 16, 812, 034, respectively. By incorporating the EL, the learnable parameters were reduced by 86% as compared to the classic CNN configuration.

G. BATCH NORMALIZATION AND ACTIVATION FUNCTION

The BNL is used between the convolutional and ReLU layers to speed up the training process and minimize the sensitivity of network initialization [36], [37]. The purpose of BNL is to eliminate the internal covariate shift. It is done by taking batch-wise mean and standard deviation normalization.

For batch normalization computation, mean and variance are calculated by using the following equations.

$$\mu_B = \frac{1}{m}\sum_{i}^{m}\chi_i \tag{6}$$

$$\sigma_B = \frac{1}{m}\times\sum_{i}^{m}(\chi_i - \mu_B)^2 \tag{7}$$

where μ_B and σ_B are the mean and variance of mini-batch, whereas, m is the mini-batch size of χ_i input feature element. The value of m is selected as 64. After computing μ_B and σ_B, the batch normalization is computed by using equation 8.

$$Y_i = \frac{(\chi_i - \mu)}{\sqrt{\sigma^2 + \epsilon}}\gamma + b \tag{8}$$

where γ and b are initial values of learnable parameters for each output.

The rectified linear unit (ReLU) is used as an activation function at the output of the convolutional layer to avoid the vanishing gradient problem and boost-up the learning speed [38]. The ReLU layer is used as piecewise function, such as $max(0, x)$ thresholding at zero. Equation 9 is used as an activation function, whereas, expression 10 represents the output of the ReLU layer.

$$Y_{i,j,k} = max\{0, \chi_{i,j,k}\} \tag{9}$$

$$Y_{ReLU} = ReLU(Bnorm(Conv(x, w))) \tag{10}$$

In equation 9, $Y_{i,j,k}$ is the output feature element and $\chi_{i,j,k}$ is the input feature element. The i and j are index values of pixels for k^{th} channel image.

H. POOLING LAYER

A pooling layer reduces the feature map size and ultimately reduces the computations and weights which leads to control over-fitting the network. In this work, every feature map from consecutive convolutional layers is directly pooled by computing the maximum of its ReLU output as given in expression 11.

$$Y_{POOL} = MaxP(ReLU(Bnorm(Conv(x, w)))) \tag{11}$$

The max-pooling is done by using following mathematical expression:

$$Y^k = max(0, \sum^{p}\chi^{k-1}\omega^k) \tag{12}$$

where Y^k is the output feature map for kth channel and χ is the input feature map. Whereas, ω is the kernel for the max-pooling layer and p represents the pooling size. Two max-pooling layers are available in our architecture and the kernel size of each layer is 2×2. The max-pooling layer operates individually on each depth slice of the input feature map and resizes it in the spatial domain by utilizing the equation 12.

I. DROPOUT REGULARIZATION

We used dropout regularization to prevent the over-fitting of training data, as it eliminates the random subset of parameters in iterative manner during the weight update process. As the fully connected layer has the maximum number of parameters over the entire network, therefore, it goes under the influence of over-fitting on training data. Therefore, the dropout regularization layer is added after the fully connected layer. In this work, we also explored our technique with different dropout regularization rate.

J. SOFTMAX CLASSIFIER AND LOSS FUNCTION

The softmax is used as a classifier which utilizes the log loss as a loss function. The probability value of softmax varies between 0 and 1, which is the confidence score for binary classes. The loss function which is given in equation14, also computes the compatibility of the available set of parameters, analogous to the ground truth labels of the training dataset.

$$\Gamma_L = \psi_{y_i} + log\sum_{j} exp(\psi_j) \tag{13}$$

where Γ_L is the total loss, and ψ_j is the jth element of the vector from class scores ψ. Moreover, the regularization term also confirms that the weights are well distributed. The objective of the classifier is to narrow down the difference between the probabilities of the actual label and predicted label, which are computed by using the following softmax function:

$$Y_i = \frac{exp^{\psi_{y_i}}}{\sum_j exp(\psi_j)} \tag{14}$$

K. BACK-PROPAGATION ALGORITHM

The proposed texture CNN was trained by using back-propagation algorithm. Let, $\theta = (\omega_i, b_i)$ be the network parameters which are updated by using following decreasing cost function between the ground truth and the training results:

$$L = -\frac{1}{|\chi|}\sum_{i=1}^{|\chi|} ln\left(P\left(y^i|\chi^i|\right)\right) \tag{15}$$

where L is the cost function which is calculated in an iterative manner. The network parameters (θ) are updated with stochastic gradient descent with momentum technique which is given in equation 16.

$$\theta(t+1) = \theta(t) - (\lambda\frac{\partial L}{\partial\theta} - \alpha\Delta\theta(t) + \beta\lambda\theta(t)) \tag{16}$$

where α represents the momentum rate, whereas, λ denotes the learning rate, which accelerates the learning procedure and leads to cope with the global minimum of the given loss function. The β represents the weight decay rate, which minimizes the decaying weight parameters nearly zero during each iteration, which causes to improve the learning efficiency of the entire network parameters. The back-propagation becomes even more effective when using the gradient descent to tune the network parameters and train a CNN.

L. DEEP FEATURE TRANSFER TECHNIQUE FOR MALIGNANCY CLASSIFICATION

The performance of different machine learning techniques essentially relies on extensive labelled data for supervised training. Whereas, deficiency of the labelled medical database for training and testing reduces the adaption of CNN. Simultaneously, the manually annotating and labelling of every data item to construct an immense training database from miscellaneous domains is really painful and prohibitive, particularly for the medical image databases which also have their own distinct privacy issues. Hence, there is a powerful inspiration to construct a classifier via deep feature transfer for the biomedical image classification problem, by taking advantage of rich labelled data of various domains. Therefore, the idea of transferring features is utilized to study a discriminative and robust model in the presence of variable test and training distributions known as TL [39]. The objective of TL is to transfer deep features from the source to target domains for the classification task. M. Oquab *et al.* performed training on the source task (ImageNet database), then transferred the pre-trained parameters of CNN to the target task for object classification [40]. The same strategy is employed in this work for lung nodule malignancy classification by using our pre-trained CNN model. The platform is introduced between deep learning and TL for lung nodule classification. Figure 3 shows the proposed TL methodology.

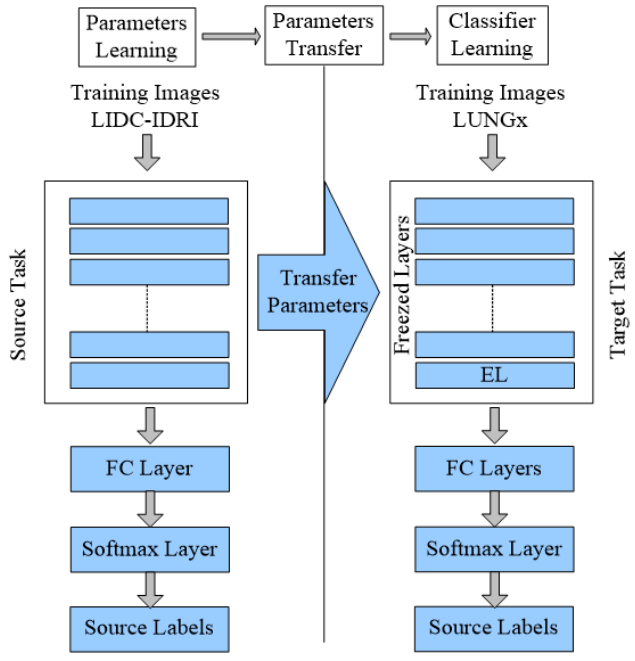

FIGURE 3. Transfer learning methodology using pre-trained texture CNN.

It is more accurate and stable TL based classifier model, which learns the significant features of the biomedical image, without considering the rich labelled biomedical image dataset. Initially, the network is trained by using GPU on the source task (top row of Figure 3) with the large number of data samples such as augmented dataset. Then, the pre-trained parameters of the internal convolutional layers and the first fully connected layer are transferred to the target task (bottom row of Figure 3). Here, the source task is LIDC-IDRI database, whereas, the target task is the LUNGx challenge database. The features are extracted from EL and then weights and biases are fine tuned by retraining last two fully connected layers for LUNGx challenge images.

M. TRAINING PROCESS AND EVALUATION METRICS

The proposed CNN model is trained and tested on a publicly available LIDC-IDRI [19] database using six-fold cross-validation strategy. The total of 925,632 image patches of LIDC-IDRI database are divided into six subsets. Then the six-fold cross validation is carried out by taking five subsets of data as training and remaining one as testing to compute the performance of our proposed texture CNN. Furthermore, to avoid the over fitting of model and monitoring the training process, 20% of each k-fold training data is utilized for validation of the proposed model. The validation is done at the end of the training epoch. The data distribution details of each training fold are illustrated in Figure 4.

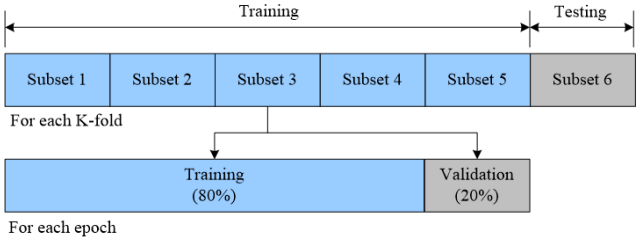

FIGURE 4. Data distribution detail for training and testing.

The training process is repeated six-time and each time the weights from the network are reinitialized randomly and then the model is trained end-to-end for 300 epochs using back-propagation algorithm. The learning rate (λ) of the model is set to 0.001 which decreases after every 2500 iterations. The decreasing factor of λ is 1×10^{-1}. The value of momentum rate (α) and weight decay rate (β) is 9×10^{-1} and 2×10^{-4}, respectively. Furthermore, the value of mini-batch size is kept 64 during back-propagation. It is to be noted that the training process becomes smooth after passing the sixty epochs. The improvement in accuracy becomes negligible which leads to the end of the successful training process. The same training procedure is also adopted to evaluate the performance of the proposed model on the LUNGx challenge database as well as on the MNIST database. The quantitative performance of the proposed method is determined by computing accuracy, recall, precision, specificity, and error rate. The details of these evaluation metrics are given in Table 3.

IV. RESULTS AND DISCUSSIONS

The implementation of the proposed texture CNN is done with a server having Intel(R) Core(TM) i7-8700 processor, 16GB RAM, and one NVIDIA TITAN Xp GPU with 12 GB RAM and compute capability of 6.1. In this work, we explored the texture CNN architecture, then the

TABLE 3. Details of evaluation metrics.

Evaluation Metrics	Mathematical Expression	Comment
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	It shows how well a classifier predicts
Recall	$\frac{TP}{TP+FN}$	The percentage of correctly identified actual positive cases
Specificity	$\frac{TN}{TN+FP}$	It is true negative rate
Error Rate	$\frac{FP+FN}{TP+FP+FN+TN}$	

performance evaluation on the LIDC-IDRI and LUNGx challenge database is performed. After that, we also validated the effectiveness of our model on the MNIST dataset.

A. THE EXPLORATION OF TEXTURE CNN STRUCTURE

For the proposed texture CNN model, first, we evaluated the performance with the different dropout layer configurations to find the appropriate value of the dropout rate for each layer. To evaluate the performance of the network, we compared it with different dropout rates. Table 4 shows the comparison of the results.

TABLE 4. Comparison of different dropout layers and dropout rates of texture CNN, the best result is mentioned in bold.

Dropout Layers	Dropout Rate	Accuracy	Error Rate
0	0	95.36%	4.63%
1	0.20%	96.54%	3.46%
2	**0.20%**	**96.69%**	**3.30%**
2	0.50%	95.03%	4.79%
2	0.60%	94.63%	5.37%

Whenever we used no dropout layer, we sustained all the neurons for the next coming layer. In such case, the classification accuracy is lower due to over-fitting. Furthermore, classification accuracy remained low at dropout rate of 0.5% and 0.6% due to withholding of extra neurons. The maximum accuracy is achieved at dropout value of 0.2%, the results are shown in bold. In this case, we kept 80% of the neurons for the next layer.

We also evaluated the performance of our model with and without EL (i.e. texture CNN and classic CNN configuration) to study the effect of EL on nodule classification performance. The comparison of both configurations is given in Table 5.

TABLE 5. Comparison of different dropout layers and dropout rates of texture CNN, the best result is mentioned in bold.

Models	Dataset	Accuracy(%)	Recall(%)	Error Rate(%)
Texture CNN	LIDC	96.69±0.12	96.05±0.37	3.30±0.06
Classic CNN	LIDC	92.08±0.23	95.12±0.25	7.98±0.10
Texture CNN	LUNGx	86.14±0.21	88.76±0.22	13.85±0.19
Classic CNN	LUNGx	85.71±0.23	87.77±0.21	14.85±0.21

The results show that the proposed texture CNN performed well as compared to the classic configuration CNN for both the databases. As the EL is identical to the average pooling and worked as dense orderless texture descriptors, it learned texture features during forward and backward propagation, which improved the classification performance.

From Table 5, it can be observed that the Texture CNN has significant improvement in classification accuracy and other metrics, as compared to the classic CNN configuration for both the databases. Moreover, we also measured the classification accuracy for each class. These results are given in Table 6.

TABLE 6. Classification accuracy score for each class for LIDC-IDRI and LUNGx Challenge dataset.

Models	Dataset	Accuracy(%)
Malignant	LIDC-IDRI	97.03%
Benign	LIDC-IDRI	96.00%
Malignant	LUNGx	85.86%
Benign	LUNGx	86.48%

From the results shown in Table 6, it can be observed that the proposed texture CNN classified the malignant nodules more accurately for both the databases, as compared to benign nodules. The classification accuracy score of the malignant nodules is 97.03% for the LIDC-IDRI dataset, whereas, it is 86.48% for the LUNGx challenge dataset.

B. PERFORMANCE EVALUATION WITH LIDC-IDRI DATABASE

The proposed model achieved comparable classification results on LIDC-IDRI Database. Table 7 shows the performance comparison of the proposed texture CNN with state-of-the-art traditional lung nodule classifications methods in terms of classification accuracy, recall, specificity, and area under the curve (AUC) scores.

TABLE 7. Performance comparison of proposed texture CNN with state-of-the-art traditional methods.

Models	Accuracy(%)	Recall(%)	Specificity(%)	AUC(%)
Narayanan et al. [41]	×	87.86	×	×
Farag et al. [20]	×	×	×	99.00
Han et al. [7]	×	89.35	86.02	94.05
Dhara et al. [42]	×	89.73	86.36	95.05
Shaffie et al. [21]	93.12	92.47	93.60	97.53
Costa et al. [43]	91.81	93.42	91.21	94.00
Sheway et al. [17]	84.00	82.00	93.00	94.00
Wei et al. [44]	87.65	89.30	86.00	94.20
Proposed Model	**96.69**	**96.05**	**97.37**	**99.11**

From the given results, it can be observed that the achieved accuracy, recall, specificity and AUC score are 96.69%, 97.16%, 97.19% and 99.11%, respectively. All of these metrics are better than the rest of the traditional lung nodule classifications methods under consideration.

After comparing with the traditional approaches, we also compared our texture CNN with deep learning-based models to prove the effectiveness of the model. Table 8 shows the performance comparison of the proposed model with various existing state-of-the-art deep learning-based models, like Fuse-TSD algorithm [25], deep fully convolutional neural network (DFCNet) [45], MV-KBC learning model [29], MK-SSAC model [30] and GD network [23] etc.

The results presented in Table 8 show that the proposed model performs better as compared to all other deep learning

TABLE 8. Performance comparison of proposed CNN on the LIDC-IDRI database with state-of-the-art methods.

Models	Accuracy(%) [Mean± SD]	Recall(%) [Mean± SD]	Specificity(%) [Mean±SD]	AUC(%) [Mean±SD]
CMixNet [4]	88.79%	93.97%	89.83%	88.79%
LdcNet-FL [24]	**97.20%**	96.00%	**97.30%**	98.20%
LdcNet-CE [24]	95.60%	90.20%	96.00%	95.6%
J. Lyu, et al. [2]	92.19%	92.10%	91.50%	97.05%
S. ghosal, et al. [46]	95.30%	95.00%	×	97.00%
Local–Global [22]	88.46%	88.66%	×	95.62%
MC-CNN [10]	87.14%	77.00%	93.00%	93.00%
Fuse TSD [25]	89.53±0.09	84.19±0.09	89.85±0.25	96.65±0.01
Feature fusion [25]	89.05±0.03	84.33±0.02	91.12±0.19	96.45±0.02
MV-KBC [29]	91.60±0.15	86.52±0.25	94.00±0.03	95.70±0.24
Y. Xie, et al. [47]	93.40±0.01	91.43±0.02	94.09±0.02	97.78±0.0001
GD network [23]	92.57±2.47	92.21±4.84	×	95.14±0.78
MK-SSAC [30]	92.53±0.05	84.94±0.17	96.28±0.08	95.81±0.19
Proposed Model	**96.63**±0.12	**96.05**±0.37	**97.37**±0.27	**99.11**±0.15

techniques except LdcNet-FL which has a bit higher accuracy and specificity score. LdcNet-FL computed the mentioned score with approximately 3.3 million learnable parameters and one million neurons. It is noteworthy to mention here, that our proposed texture CNN computed the marginal lower accuracy with considerably lesser learnable parameters and neurons. The comparison of learnable parameters and neurons of the proposed texture CNN with other deep learning techniques is given in the subsequent subsection.

C. EVALUATION OF PRE-TRAINED MODE ON LUNGx CHALLENGE DATABASE

First we, trained our proposed texture CNN from scratch on the LUNGx challenge database and achieved comparable results. After that we used our pre-trained model of the LIDC-IDRI database to investigate the small dataset training issue of CNN by implementing the TL methodology. The classification results of proposed CNN are compared with Xie *et al.* [30], Fine Tuned MK-SSAC, MV-KBC, and Mizuho Nishio, *et al.* [48], who proposed two classification models, which are: CADx using SVM with tree parzen estimator (TPE) and gradient tree boosting (XGBoost) with TPE. Table 9 shows the comparison of our proposed texture CNN on the LUNGx Challenge database.

TABLE 9. Performance comparison of proposed texture CNN with state-of-the-art traditional methods.

Models	Accuracy(%)	Recall(%)	Specificity(%)	AUC(%)
SVM (TPE) [48]	82.00	×	×	85.00
XGBoost (TPE) [48]	86.84	×	×	89.60
Nishio et al. [49]	×	86.70	×	83.70
Xie et al. [30]	77.26	87.22	67.57	78.83
MV-KBC [29]	75.62	87.22%	64.32	76.85
Proposed without TL	86.14%	88.76%	93.11%	92.63%
Proposed with TL	**90.91%**	**91.39%**	**90.46%**	**94.14%**

The achieved classification score of our proposed texture CNN without TL (trained from scratch) for accuracy, recall, specificity, and AUC score on LUNGx database are 86.14%, 88.76%, 93.11% and 92.63%, respectively, which show that the proposed CNN performed better than all the other considered techniques expect XGBoost (TPE) in terms of accuracy score only. Furthermore, the results show that the implementation of TL methodology with a pre-trained model significantly improved the accuracy as compared to our trained model, which proves the effectiveness of TL methodology.

D. ARCHITECTURE COMPLEXITY COMPARISON WITH STATE-OF-THE-ART TECHNIQUES

The architecture complexity is based on activation functions like neurons and learnable parameters. We computed the total number of neurons and learnable parameters of the proposed model and compared them with the recent proposed state-of-the-art techniques like LdcNet with cross-entropy loss (LdcNet-CE), LdcNet with Focal Loss (LdcNet-FL) [24] and customized mixed link network (CMixNet) [4] as given in Table 10.

TABLE 10. Comparisons with state-of-the-art methods in terms of neurons and learnable parameter.

Models	Neurons	Parameters	Accuracy
CMixNet [4]	14,725,632	>14.7	88.79%
LdcNet-CE [24]	1,008,898	3,292,763	95.60%
LdcNet-FL [24]	1,008,898	3,292,763	97.20%
Classic Configuration	2,263,170	16,812,034	94.005%
Proposed Model	**107,842**	**2,263,170**	**96.69%**

From Table 10, it can be observed that the proposed model has a lesser number of neurons and learnable parameters which leads to a reduction in the complexity. It is noteworthy to mention here that the reduction of neurons and learnable parameters is due to the incorporation of EL. Therefore, the EL reduced the complexity of the network without degrading the classification accuracy.

E. PERFORMANCE VALIDATION OF TEXTURE CNN ON MNIST DATASET

The proposed texture CNN was successfully trained and tested for lung nodule malignancy classification. In addition, we also validated our proposed model on the MNIST dataset which is a handwritten digits dataset. It consists of 10,000 labeled test images and 60,000 labeled training images. We trained and tested our proposed model successfully and compared the results with state-of-the-art techniques. These results are given in Table 11.

TABLE 11. Comparisons with state-of-the-art methods on the MNIST dataset.

Models	Accuracy	Error rate
Tabik,et al. [50]	×	0.10%
Skouson et al. [51]	99.20%	×
Simonovsky et al. [52]	99.37%	×
Klokov et al. [53]	99.10%	×
Grover et al. [54]	99.54%	0.51%
Qi et al. [55]	99.50%	0.51%
Proposed Model	**99.89%**	**0.12%**

The results show that the proposed texture CNN also performed well as compared to the other techniques. It can also be observed that the proposed texture CNN computed the marginal lower error rate of 0.02% than Tabik *et al.* [50]. Furthermore, it is also mentioned that Grover *et al.* [54] achieved recall and specificity of 97.73% and 99.74%, respectively. However, the achieved recall and speficicity scores for both metrics by our proposed model are 99.94% and 99.93%, respectively, which reflect the effectiveness of the proposed texture CNN.

V. CONCLUSION

In this paper, we have proposed a transferable texture CNN architecture for lung nodule malignancy classification tasks. We introduced the EL, which removes the overall shape information and explores the texture features. Experimental results show the effectiveness of the proposed technique for benign and malignant nodules classification, without nodule segmentation or any complex pre-processing. After successful training, we evaluated the performance of the proposed network by using various evaluation metrics. The results were compared with state-of-the-art lung nodule classification methods. The results show that our proposed texture CNN architecture performed well for approximately all the evaluation metrics. The training was done successfully by six-fold cross-validation and achieved an accuracy, recall, specificity, AUC, and error rate of 96.69%, 96.05%, 97.37%, 99.11%, and 3.30%, respectively, on LIDC-IDRI database. The learned features of EL were analyzed and it was noted that the EL extracted texture from the convolutional layer. The EL also reduced the number of learnable parameters of the network, which leads to minimize the memory requirements and complexity of CNN. Furthermore, we explored our pre-trained model to handle the smaller dataset classification problem by using TL. We also show that our pre-trained model achieved better results than the compared techniques on a small LUNGx Challenge database. Moreover, we also validated the effectiveness of our proposed texture CNN on the MNIST dataset, as our model achieved 99.89% accuracy with only 0.12% error rate.

REFERENCES

[1] M. G. Kris, L. E. Gaspar, J. E. Chaft, E. B. Kennedy, C. G. Azzoli, P. M. Ellis, S. H. Lin, H. I. Pass, R. Seth, and F. A. Shepherd, "Adjuvant systemic therapy and adjuvant radiation therapy for stage I to IIIA completely resected non-small-cell lung cancers: American society of clinical oncology/cancer care Ontario clinical practice guideline update," *J. Clin. Oncol.*, vol. 35, no. 25, pp. 2960–2974, 2017.

[2] J. Lyu, X. Bi, and S. H. Ling, "Multi-level cross residual network for lung nodule classification," *Sensors*, vol. 20, no. 10, p. 2837, May 2020.

[3] G. A. P. Singh and P. Gupta, "Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans," *Neural Comput. Appl.*, vol. 31, no. 10, pp. 6863–6877, 2018.

[4] N. Nasrullah, J. Sang, M. S. Alam, M. Mateen, B. Cai, and H. Hu, "Automated lung nodule detection and classification using deep learning combined with multiple strategies," *Sensors*, vol. 19, no. 17, p. 3722, Aug. 2019.

[5] A. Snoeckx, P. Reyntiens, D. Desbuquoit, M. J. Spinhoven, P. E. Van Schil, J. P. van Meerbeeck, and P. M. Parizel, "Evaluation of the solitary pulmonary nodule: Size matters, but do not ignore the power of morphology," *Insights Into Imag.*, vol. 9, no. 1, pp. 73–86, Feb. 2018.

[6] S. Jia, B. Deng, J. Zhu, X. Jia, and Q. Li, "Local binary pattern-based hyperspectral image classification with superpixel guidance," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 749–759, Feb. 2018.

[7] F. Han, H. Wang, G. Zhang, H. Han, B. Song, L. Li, W. Moore, H. Lu, H. Zhao, and Z. Liang, "Texture feature analysis for computer-aided diagnosis on pulmonary nodules," *J. Digit. Imag.*, vol. 28, no. 1, pp. 99–115, Feb. 2015.

[8] N. Banerjee, N. Dimitrova, V. Varadan, S. Kamalakaran, A. Janevski, and S. Maity, "Integrated phenotyping employing image texture features," U.S. Patent 9 552 649, Jan. 24, 2017.

[9] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," in *Proc. Int. Conf. Inf. Process. Med. Imag.* Cham, Switzerland: Springer, 2015, pp. 588–599.

[10] W. Shen, M. Zhou, F. Yang, D. Yu, D. Dong, C. Yang, Y. Zang, and J. Tian, "Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification," *Pattern Recognit.*, vol. 61, pp. 663–673, Jan. 2017.

[11] F. Ciompi, B. de Hoop, S. J. van Riel, K. Chung, E. T. Scholten, M. Oudkerk, P. A. de Jong, M. Prokop, and B. V. Ginneken, "Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box," *Med. Image Anal.*, vol. 26, no. 1, pp. 195–202, Dec. 2015.

[12] V. Andrearczyk and P. F. Whelan, "Using filter banks in convolutional neural networks for texture classification," *Pattern Recognit. Lett.*, vol. 84, pp. 63–69, Dec. 2016.

[13] L. Yann, C. Corinna, and J. B. Christopher. (1998). *Mnist Dataset of Handwritten Digits*. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[14] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[15] B. N. Narayanan, R. C. Hardie, and T. M. Kebede, "Performance analysis of a computer-aided detection system for lung nodules in ct at different slice thicknesses," *J. Med. Imag.*, vol. 5, no. 1, p. 014504, 2018.

[16] B. N. Narayanan, R. C. Hardie, and T. M. Kebede, "Performance analysis of feature selection techniques for support vector machine and its application for lung nodule detection," in *Proc. IEEE Nat. Aerosp. Electron. Conf. (NAECON)*, Jul. 2018, pp. 262–266.

[17] T. Nesibu Shewaye and A. Ali Mekonnen, "Benign-malignant lung nodule classification with geometric and appearance histogram features," 2016, *arXiv:1605.08350*. [Online]. Available: http://arxiv.org/abs/1605.08350

[18] M. Firmino, G. Angelo, H. Morais, M. R. Dantas, and R. Valentim, "Computer-aided detection (CADe) and diagnosis (CADx) system for lung cancer with likelihood of malignancy," *Biomed. Eng. OnLine*, vol. 15, no. 1, p. 2, Dec. 2016.

[19] S. G. Armato, III, *et al.*, "The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans," *Med. Phys.*, vol. 38, no. 2, pp. 915–931, Jan. 2011.

[20] A. A. Farag, A. Ali, S. Elshazly, and A. A. Farag, "Feature fusion for lung nodule classification," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 10, pp. 1809–1818, Oct. 2017.

[21] A. Shaffie, A. Soliman, H. A. Khalifeh, M. Ghazal, F. Taher, A. Elmaghraby, R. Keynton, and A. El-Baz, "Radiomic-based framework for early diagnosis of lung cancer," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1293–1297.

[22] M. Al-Shabi, B. L. Lan, W. Y. Chan, K.-H. Ng, and M. Tan, "Lung nodule classification using deep local–global networks," *Int. J. Comput. Assist. Radiol. surgery*, vol. 14, no. 10, pp. 1815–1819, 2019.

[23] M. Al-Shabi, H. Kuan Lee, and M. Tan, "Gated-dilated networks for lung nodule classification in CT scans," 2019, *arXiv:1901.00120*. [Online]. Available: http://arxiv.org/abs/1901.00120

[24] G. S. Tran, T. P. Nghiem, V. T. Nguyen, C. M. Luong, and J.-C. Burie, "Improving accuracy of lung nodule classification using deep learning with focal loss," *J. Healthcare Eng.*, vol. 2019, pp. 1–9, Feb. 2019.
[25] Y. Xie, J. Zhang, Y. Xia, M. Fulham, and Y. Zhang, "Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT," *Inf. Fusion*, vol. 42, pp. 102–110, Jul. 2018.
[26] I. Banerjee, A. Crawley, M. Bhethanabotla, H. E. Daldrup-Link, and D. L. Rubin, "Transfer learning on fused multiparametric MR images for classifying histopathological subtypes of rhabdomyosarcoma," *Comput. Med. Imag. Graph.*, vol. 65, pp. 167–175, Apr. 2018.
[27] P. M. Cheng and H. S. Malhi, "Transfer learning with convolutional neural networks for classification of abdominal ultrasound images," *J. Digit. Imag.*, vol. 30, no. 2, pp. 234–243, Apr. 2017.
[28] X. Zhao, S. Qi, B. Zhang, H. Ma, W. Qian, Y. Yao, and J. Sun, "Deep CNN models for pulmonary nodule classification: Model modification, model integration, and transfer learning," *J. X-Ray Sci. Technol.*, vol. 27, no. 4, pp. 615–629, Sep. 2019.
[29] Y. Xie, Y. Xia, J. Zhang, Y. Song, D. Feng, M. Fulham, and W. Cai, "Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT," *IEEE Trans. Med. Imag.*, vol. 38, no. 4, pp. 991–1004, Apr. 2019.
[30] Y. Xie, J. Zhang, and Y. Xia, "Semi-supervised adversarial model for benign–malignant lung nodule classification on chest CT," *Med. Image Anal.*, vol. 57, pp. 237–248, Oct. 2019.
[31] R. Dey, Z. Lu, and Y. Hong, "Diagnostic classification of lung nodules using 3D neural networks," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 774–778.
[32] S. Shen, S. X. Han, D. R. Aberle, A. A. T. Bui, and W. Hsu, "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification," 2018, *arXiv:1806.00712*. [Online]. Available: http://arxiv.org/abs/1806.00712
[33] A. P. Reeves and A. M. Biancardi, "The lung image database consortium (LIDC) nodule size report," Cornell Univ. Vis. Image Anal. Group, Ithaca, NY, USA, Tech. Rep. 2011-10-27-2, 2011.
[34] S. G. Armato, L. Hadjiiski, G. D. Tourassi, K. Drukker, M. L. Giger, F. Li, G. Redmond, K. Farahani, J. S. Kirby, and L. P. Clarke, "Guest editorial: LUNGx challenge for computerized lung nodule classification: Reflections and lessons learned," *J. Med. Imag.*, vol. 2, no. 2, Jun. 2015, Art. no. 020103.
[35] S. G. Armato, K. Drukker, F. Li, L. Hadjiiski, G. D. Tourassi, R. M. Engelmann, M. L. Giger, G. Redmond, K. Farahani, J. S. Kirby, and L. P. Clarke, "LUNGx challenge for computerized lung nodule classification," *J. Med. Imag.*, vol. 3, no. 4, Dec. 2016, Art. no. 044506.
[36] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: http://arxiv.org/abs/1502.03167
[37] C. Cheng, J. Li, Y. Liu, M. Nie, and W. Wang, "Deep convolutional neural network-based in-process tool condition monitoring in abrasive belt grinding," *Comput. Ind.*, vol. 106, pp. 1–13, Apr. 2019.
[38] S. Pang, Z. Yu, and M. A. Orgun, "A novel end-to-end classifier using domain transferred deep convolutional neural networks for biomedical images," *Comput. Methods Programs Biomed.*, vol. 140, pp. 283–293, Mar. 2017.
[39] M. Long, H. Zhu, J. Wang, and M. I. Jordan, "Deep transfer learning with joint adaptation networks," 2016, *arXiv:1605.06636*. [Online]. Available: http://arxiv.org/abs/1605.06636
[40] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1717–1724.
[41] B. N. Narayanan, R. C. Hardie, T. M. Kebede, and M. J. Sprague, "Optimized feature selection-based clustering approach for computer-aided detection of lung nodules in different modalities," *Pattern Anal. Appl.*, vol. 22, no. 2, pp. 559–571, May 2019.
[42] A. K. Dhara, S. Mukhopadhyay, A. Dutta, M. Garg, and N. Khandelwal, "A combination of shape and texture features for classification of pulmonary nodules in lung CT images," *J. Digit. Imag.*, vol. 29, no. 4, pp. 466–475, Aug. 2016.
[43] R. W. de Sousa Costa, G. L. F. da Silva, A. O. de Carvalho Filho, A. C. Silva, A. C. de Paiva, and M. Gattass, "Classification of malignant and benign lung nodules using taxonomic diversity index and phylogenetic distance," *Med. Biol. Eng. Comput.*, vol. 56, no. 11, pp. 2125–2136, Nov. 2018.
[44] G. Wei, H. Ma, W. Qian, F. Han, H. Jiang, S. Qi, and M. Qiu, "Lung nodule classification using local kernel regression models with out-of-sample extension," *Biomed. Signal Process. Control*, vol. 40, pp. 1–9, Feb. 2018.
[45] A. Masood, B. Sheng, P. Li, X. Hou, X. Wei, J. Qin, and D. Feng, "Computer-assisted decision support system in pulmonary cancer detection and stage classification on CT images," *J. Biomed. Informat.*, vol. 79, pp. 117–128, Mar. 2018.
[46] S. S. Ghosal, I. Sarkar, and I. El Hallaoui, "Lung nodule classification using convolutional autoencoder and clustering augmented learning method (CALM)," in *Proc. HSDM@ WSDM*, 2020, pp. 19–26.
[47] Y. Xie, Y. Xia, J. Zhang, D. D. Feng, M. Fulham, and W. Cai, "Transferable multi-model ensemble for benign-malignant lung nodule classification on chest ct," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2017, pp. 656–664.
[48] M. Nishio, M. Nishizawa, O. Sugiyama, R. Kojima, M. Yakami, T. Kuroda, and K. Togashi, "Computer-aided diagnosis of lung nodule using gradient tree boosting and Bayesian optimization," *PLoS ONE*, vol. 13, no. 4, Apr. 2018, Art. no. e0195875.
[49] M. Nishio and C. Nagashima, "Computer-aided diagnosis for lung cancer: Usefulness of nodule heterogeneity," *Acad. Radiol.*, vol. 24, no. 3, pp. 328–336, 2017.
[50] S. Tabik, R. F. Alvear-Sandoval, M. M. Ruiz, J.-L. Sancho-Gómez, A. R. Figueiras-Vidal, and F. Herrera, "MNIST-NET10: A heterogeneous deep networks fusion based on the degree of certainty to reach 0.1% error rate. Ensembles overview and proposal," *Inf. Fusion*, vol. 62, pp. 73–80, Oct. 2020.
[51] M. B. Skouson, B. J. Borghetti, and R. C. Leishman, "Ursa: A neural network for unordered point clouds using constellations," in *Proc. Sci. Inf. Conf.* Cham, Switzerland: Springer, 2019, pp. 447–457.
[52] M. Simonovsky and N. Komodakis, "Dynamic edge-conditioned filters in convolutional neural networks on graphs," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 3693–3702.
[53] R. Klokov and V. Lempitsky, "Escape from cells: Deep KD-networks for the recognition of 3D point cloud models," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 863–872.
[54] D. Grover and B. Toghi, "Mnist dataset classification utilizing K-NN classifier with modified sliding-window metric," in *Proc. Sci. Inf. Conf.* Cham, Switzerland: Springer, 2019, pp. 583–591.
[55] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5099–5108.

IMDAD ALI received the B.Sc. degree in electrical engineering from the Federal Urdu University of Arts, Science and Technology, Pakistan, in 2009, and the M.Sc. degree in electrical engineering from the University of Engineering and Technology, Taxila, Pakistan, in 2014. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, International Islamic University, Islamabad, Pakistan. He is also a Senior Engineer with the National Centre for Physics, Quaid-e-Azam University, Pakistan. His research interests include computer added biomedical diagnostic using deep learning, compress sensing in digital image processing, and video coding.

MUHAMMAD MUZAMMIL (Member, IEEE) received the B.Sc. degree in electrical engineering from the Federal Urdu University of Arts Science and Technology, Islamabad, Pakistan, and the M.Sc. degree in electrical engineering with specialization in digital techniques from the University of Engineering and Technology Taxial, Pakistan, in 2014. He has been serving as a Lecturer with the Department of Electrical Engineering (DEE), Faculty of Engineering and Technology (FET), International Islamic University (IIU), Islamabad, Pakistan. He is currently a Ph.D. Scholar with the DEE, FET, IIU, Islamabad. His research interests include image processing, image fusion, medical image analysis, and deep learning.

IHSAN UL HAQ (Member, IEEE) received the Ph.D. degree in electronics and information engineering from Beihang University, Beijing, China, in 2009, and the M.S. degree in electronics from Quad-i-Azam University, Islamabad, Pakistan, in 2004. He has been working as an Associate Professor with the Department of Electrical Engineering, Faculty of Engineering and Technology, International Islamic University, Islamabad. He has supervised six Ph.D. and 25 M.S. students. He has published 35 journal and 15 proceeding articles. His research interests include image processing, hyperspectral image processing, medical image processing, signal processing, deep learning, and communication.

MUHAMMAD AMIR (Member, IEEE) received the B.S. degree in electrical engineering from the University of Engineering and Technology, Peshawar, in 1997, the M.S. degree in electrical engineering from the University of Engineering and Technology, Taxila, in 2003, and the Ph.D. degree in electronic engineering from International Islamic University, Islamabad, Pakistan, with a specialization in signal and image processing. From 1997 to 2001, he has served in industry as a PLC Engineer. From 2002 to 2007, he worked with Pakistan Telecommunication Company as an Assistant Divisional Engineer. In 2007, he joined International Islamic University, Islamabad as an Assistant Professor. He worked as the Head of the Department of Electrical Engineering, and the Director of Quality Enhancement Cell of the university. He is currently a Professor of electrical engineering and the Dean of the Faculty of Engineering and Technology. He was the Co-Chair of International Conference on Intelligent Systems Engineering 2017 (ICISE 2017) and the Technical Committee Co-Chair of ICISE 2016. He is the Chairman of the Board of Faculty as well as a member of Board of Studies, Board of Advanced Studies, and Research and Academic Council.

SUHEEL ABDULLAH (Member, IEEE) received the B.E. degree in electrical and electronics from Bangalore University, Bengaluru, India, in 1997, the M.S. degree in electronic engineering from Muhammad Ali Jinnah University (MAJU), Pakistan, and the Ph.D. degree in electronic engineering from International Islamic University, Islamabad (IIUI), Pakistan. Since 2007, he has been with IIUI, where he is currently serving as the Chairman/Associate Professor with the Department of Electrical Engineering (DEE). He has authored more than 20 journal and conference publications. He is also a reviewer of some journals. He has served as a member/chair in some international conferences. His research interests include control systems, numerical investigation of nonlinear problems, application of nature-inspired metaheuristic algorithms for solving nonlinear problems, and signal processing.