

Received April 27, 2020, accepted May 25, 2020, date of publication June 1, 2020, date of current version June 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2998808

Deep Convolution Neural Network for Big Data Medical Image Classification

REHAN ASHRAF¹, MUHAMMAD ASIF HABIB¹, MUHAMMAD AKRAM²,
MUHAMMAD AHSAN LATIF³, MUHAMMAD SHERAZ ARSHAD MALIK⁴,
MUHAMMAD AWAIS⁵, SAADAT HANIF DAR⁶, TOQEER MAHMOOD¹,
MUHAMMAD YASIR⁷, AND ZAHOR ABBAS¹

¹Department of Computer Science, National Textile University, Faisalabad 37610, Pakistan

²Department of Software Engineering, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta 87300, Pakistan

³Department of Computer Science, University of Agriculture, Faisalabad 38000, Pakistan

⁴Department of Information Technology, Government College University Faisalabad, Faisalabad 38000, Pakistan

⁵Department of Software Engineering, Government College University Faisalabad, Faisalabad 38000, Pakistan

⁶Department of Software Engineering, Mirpur University of Science and Technology (MUST), Mirpur 10250, Pakistan

⁷Department of Computer Science, University of Engineering and Technology Lahore, Faisalabad 38090, Pakistan

Corresponding author: Muhammad Asif Habib (drasif@ntu.edu.pk)

ABSTRACT Deep learning is one of the most unexpected machine learning techniques which is being used in many applications like image classification, image analysis, clinical archives and object recognition. With an extensive utilization of digital images as information in the hospitals, the archives of medical images are growing exponentially. Digital images play a vigorous role in predicting the patient disease intensity and there are vast applications of medical images in diagnosis and investigation. Due to recent developments in imaging technology, classifying medical images in an automatic way is an open research problem for researchers of computer vision. For classifying the medical images according to their relevant classes a most suitable classifier is most important. Image classification is beneficial to predict the appropriate class or category of unknown images. The less discriminating ability and domain-specific categorization are the main drawbacks of low-level features. A semantic gap that exists between features of low-level as machine understanding and features of human understanding as high-level perception. In this research, a novel image representation method is proposed where the algorithm is trained for classifying medical images by deep learning technique. A pre-trained deep convolution neural network method with the fine-tuned approach is applied to the last three layers of deep neural network. The results of the experiment exhibit that our method is best suited to classify various medical images for various body organs. In this manner, data can sum up to other medical classification applications which supports radiologist's efforts for improving diagnosis.

INDEX TERMS Medical image classification, pre-trained DCNN, convolution neural network, big data, image analysis, image enhancement, biomedical image processing, deep learning.

I. INTRODUCTION

Due to increase in digital devices and advancement of camera technology, there is an exponential increase in the production of medical images. Nowadays modern hospital uses a digital image to predict patient disease intensity. With the rapid development of digital images, image classification has become a significant challenge with a large number of medical images. Therefore, classification methods are required to assign these medical images the most relevant class according

to their similarity. In a medical image classification domain, there are images of different body organs such as computed tomography images (CT-Scans), X-Rays (electromagnetic waves), positron emission tomography (PET) scan is type of imaging test that helps reveal how your organs and tissue of the body are functioning. Magnetic resonance imaging (MRI) is another kind of scan that can produce and gave the complete information and clearly defined the pictures of the parts of the body including brain. Due to large volume of medical images, it is nearly impossible for a doctor or physician to classify images with manual effort. Other than the categories discussed above, there are different images of body parts such

The associate editor coordinating the review of this manuscript and approving it for publication was Liangxiu Han¹.

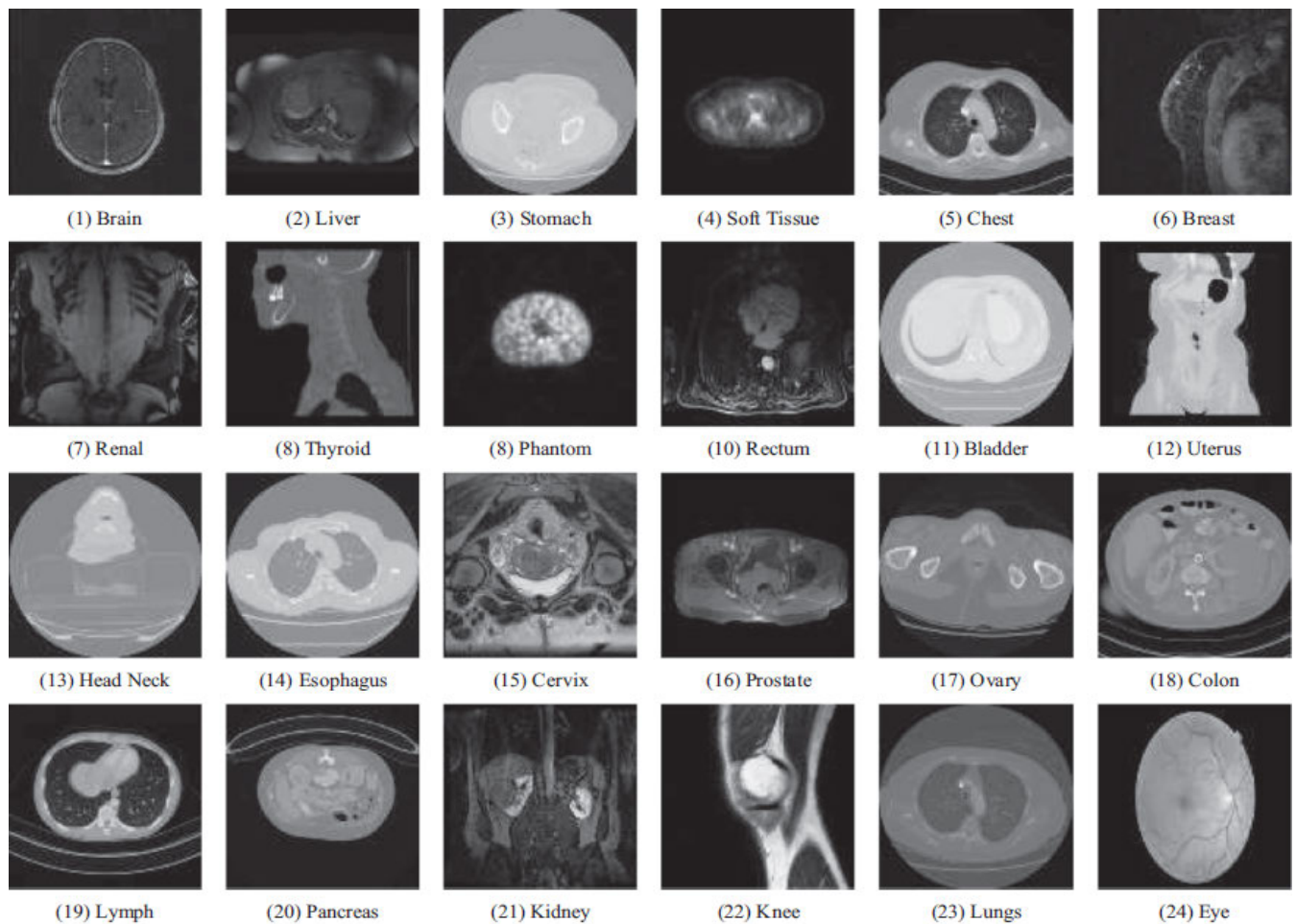


FIGURE 1. Sample images of medical image dataset.

as chest, knee, bladder, eye, breast, colon etc. are examples of this scenario is shown in Figure 1.

In recent years, handcrafted feature representation are widely used in medical domain and many techniques have been reported by using the fusion of low-level features. Low-level features were obtained and evaluated along with single features, texture features, or based on colour features and multiple feature to combine the powerful representation of images. Therefore, classification performance is only dependent on the representation of features which is based on the images. These features are important for the classification but handcrafted features to maintain consciousness of the prior knowledge. This is arduous task to create the hand-craft features and needs labelling the data for training and tuning. A data driven method for image representation and classification exploration, would be more vigorous to the diversity of image modalities and diseases although it has less susceptible to human domain-specific subjectivity.

In the work of [1], author used boundary-weighted adaptive neural network for classification purpose. The suggested method helped to reduce the complexity among prostate and other structures. This method showed more accuracy for segmentation of prostate. However, the evaluation of the

method is performed on small datasets which makes the application limited. Qikui *et al.* used particle swarm optimization with the help of shape features are suggested for segmenting the outer and inner boundaries of the bladder wall [2]. To segregate the weak and strong boundaries the authors utilized gradient information. A segmentation method is proposed to extract intra and inter-slice data of the prostate using the bidirectional convolutional recurrent layers [3].

The main objective of medical image classification research is to accurately classify the medical images of different body organs according to the relevant classes, therefore, training-testing methods are required to solve the problems. For the purpose of classification, representation of images play an important role. Feature representation is a crucial step in medical domain. The learning of our method is done by using a set of training images and features, while the output is reported based on values obtained from the test dataset. To represent the data, low level visual features such as color, shape, spatial layout, and texture are commonly used to decode the images as feature vectors [4], [5]. Most of the low-level features are domain-specific that results in a semantic gap between the visual high level features and their representation. In the last two decades, significant research has

been conducted to reduce the semantic gap which is a major requirement to develop a reliable system closer to the human visual system. Machine learning techniques are applied to the low-level image visuals to perform image classification.

In this regard, the research of computer vision and image classification is shifted to deep learning techniques that are considered as more reliable than traditional machine learning approaches to enhance the performance of image classification and retrieval problems [6]. The working theme of the deep neural network is inspired by the human brain. The deep neural network architecture consists of multiple layers of neurons. The information is passed through the complex networks that work like human brain where neurons are the processing units. Different deep learning techniques are applied in various applications for image and video classification [7], [8] such as visual tracking [9], speech recognition [10] and natural language processing [11]. Inspired from the efficient and reliable performance of deep learning framework, Deep Convolutional Neural Network (DCNN) is deployed for image feature extraction in the medical domain as the applied area of this research. The main theme of this paper is to classify 2D images of the medical image dataset. The dataset used in this research consists of different classes that are assigned labels with respect to the image class. The pre-processing is applied as the first step to resize the images to 224×224 pixels and to remove image noise different operations performed. After pre-processing, all the images are of same size. In this regard, the images are alienated into different classes which is depend on different body part or organ as shown in figure 1. Therefore, a task is to classify each image in to relevant class. DCNN is applied to formulate the features and classifier from the given training data in an end to end learning mechanism. Thereafter, a DCNN framework is applied to classify the images into the relevant class.

The main contributions of this research are mentioned as follows:

- I. Developed a dataset of different body organs using multiple free online databases of medical images.
- II. A novel deep-CNN algorithm is developed by using training-testing method for medical image classification problem.
- III. Accurately classify the image taking as input to a classifier and label with the relevant class name.
- IV. The features which is learned from training images are further used to present a highly efficient medical image classification system based on huge collection of medical dataset.

This remaining paper is organized as follows: Section II presents the literature review regarding deep learning, section III is about the proposed method, section IV describes the experimental results and section V is about the conclusion.

II. LITERATURE REVIEW

In the past decades, two major approaches are commonly used for image classification. In the first approach,

features are extracted, and traditional machine learning approaches are applied for image classification. In the second approach, features are automatically extracted by using a deep CNN through of hidden layers [12], [13]. The traditional approaches are domain-specific and the performance degrades if there is a change in the domain or increase in the number of classes [14]. Low-level features-based on the method such as color, texture, shape is spatial layout are the common examples of these approaches. In the past, significant research is performed by using the low-level visual features and traditional machine learning approaches [14]. The author applied Bag of Visual Words (BOVW) method by using Local Binary Pattern (LBP) and Scale Invariant Feature Transform (SIFT) to classify the medical images. The feature fusion of LBP and SIFT are compared with traditional feature extraction and machine learning approaches [15].

The recent research trends for medical image classification are now shifted to DCNN [16]. Author introduced a method based on hyper dimensional computing using different binary vectors to represent objects [17]. According to the work in [18], the high-dimensional feature vector based on a binary vector with more computations in feature vector computation is beneficial for image classification-based representation.

In the work of [19], author proposed to extract the hand-crafted features which is based on complicated texture and its intensity distribution problem. Author described to resolve the segmentation problem using local smooth regularization technique which is not depended to the process of morphologic methods. In the work of [20], author proposed a clean and effective feature fusion adversarial learning network to mine useful features and relieve over-fitting problems. Firstly, we train a fully convolution autoencoder network with unsupervised learning to mine useful feature maps from our liver lesion data. Secondly, these feature maps will be transferred to our adversarial SENet network for liver lesion classification.

Supervised learning and unsupervised learning are the common trends that are applied to image classification-based problems. In the supervised learning approaches, the data is represented as $D = x, y$ in which x and y are the variables of input features. In training phase, the optimized values are computed so that the best value from the test data can be predicted by using the function $L(Y, \hat{y})$, where, \hat{y} represents the output. In case of unsupervised learning approaches, the unlabeled data is used for training and testing for the classification. Traditional neural networks use transfer functions that are known as tangent hyperbolic function. Multi-Layered Perceptions (MLP) are the example of neural networks with multiple layers. The transfer function can be expressed as:

$$g(x; \theta) = \sigma(Z_P \sigma(Z_P - 1 \dots \sigma(Z_0 x + b_0) + b_P - 1) + b_P) \quad (1)$$

In the above equation, Z_P is the combination of rows Z_x and x is constant, n shows the numbers of bytes while P is the network layer. A Neural Network (NN) with many

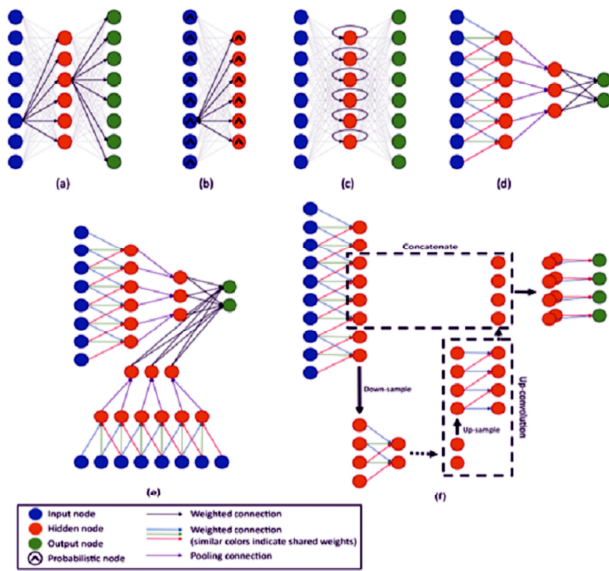


FIGURE 2. 1D NN architecture: a) Auto encoder, b) Boltzmann machine, c) Recurrent NN, (d) CNN, (e) Multistream CNN, f) DCNN.

hidden layers is considered as Deep Neural Network (DNN). In DCNN, activation functions determine the output of a deep learning method that can be expressed as:

$$P(y|x; \theta) = \text{softmax}(x; \theta) = \frac{e^{(Z_i^l)^T x + b_i^l}}{\sum_{m=1}^m e^{(Z_k^p)^T x + b_i^p}} \quad (2)$$

Here, Z_p shows the value of the weights of the output layer that is linked with the layers of MLP. The Maximum likelihood with stochastic gradient is applied to compute the optimal value of θ . The small patches from the dataset are used by stochastic gradient descent for gradient value. The second parameter limits the tag negative log likelihood by using maximum likelihood and is expressed as:

$$\underset{\theta}{\text{argmin}} - \sum_{k=1}^M \log(p(y_k | x_k; \theta)) \quad (3)$$

In [21], one of the main problems while using a DCNN is the high computational cost and requirement of large-scale image dataset for learning. A DNN with optimal parameters setting with enough samples of training data can outperform the existing hand-crafted features. Figure 2 (a-e) represents different configurations of DNN with their respective names/abbreviations.

A. CONVOLUTIONAL NEURAL NETWORK (CNN)

In CNN, the weight values of the networks are set in a way that all operations on the images are performed by the network. Due to this reason, this framework does not require additional detectors if there is a spatial change of an object within the image. The number of parameters of the learning network is also reduced in this case. Figure 2 represents a one-dimensional CNN with all layers of network by using different kernels. The weight W and biases B can be

mathematically expressed as:

$$W = \{w_1, w_2, w_3, \dots, w_k\}$$

$$B = \{b_1, b_2, b_3, \dots, b_k\}$$

The feature vector is generated by using these weights and biases. The process is repeated for each layer of the network in same manner. CNN architecture is different from MLP as the integration of pooling layers in CNN with pixel values of neighborhood use an invariant function such as mean and max. Through this scheme, there is an increase in translation invariance that increases the receptive field of convolutional layers. The last layer on CNN is fully connected and that points to the weight value. In MLP a soft-max function is used for classification problem that is activated in the last layer of the network.

1) MIC USING DEEP CNN

Alex Net and LeNet [22] are considered as popular architectures for the classification of deep CNN. Both are considered as lightweight CNN with two, and five CNN layers with different kernels used for input and output purposes. In Alex Net, hyperbolic functions are not used for activation and linear units are used in CNN layers. Later on, deeper networks with more numbers of the hidden layer are introduced for image classification and analysis [23]. In these architectures, smaller size kernel functions are used with the respective parameters instead of using larger kernels. Due to this reason, these networks require a smaller memory for storage and interfacing. Some of these networks are used in devices with limited memory and computation power like hand-held devices. According to [23]–[25], VGG19 is an example of deep network with 19 layers with small size kernels for all layers. An efficient deep network is the one that requires less training time and contains reduced number of parameters. Christian Szegedy *et al.* proposed a deep network method with 22 layers that is named as GoogleNet and is also known as inception [26].

Wei Shen *et al.* proposed a method that depend on different sizes of convolutional layers with smaller sized kernels and a smaller number of parameters are used to classify multiple classes. In the work of [27], ResNet architecture won the image-Net challenge and different sizes of ResNet blocks. The learning of each layer is dependent on residual blocks using the identity function rule. According to literature [28]–[30], Alex Net and VGG are used in various applications for medical image classification and shown state-of-the-art classification results. The selection of an appropriate network for a specific application of medical image classification is still an open research area. In deep network architectures, different layers of the network are joined at the terminating layers. Multiple tasks are divided among layers by using dual path-way architecture [31]. The context of any digital image provides an important clue about category-wise classification. In deep networks, numbers of patches are increased due to the available information about

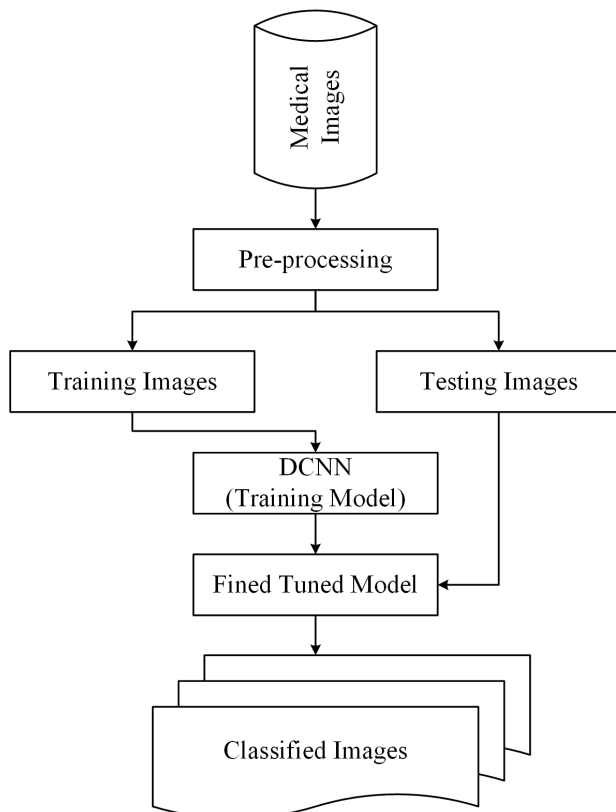


FIGURE 3. The Process of medical image classification based on GoogleNet method.

image context and this operation required more memory. The recent research of medical image classification is focused to enhance the classification accuracy of networks by using the architectures with more number of hidden layers [31]–[33]. Some medical image processing techniques require 3D data and computational complexity increases when deep network is dealing with 3D images. 3D image is processed in the network in the form of slices which carry information travelled within the layers of deep network [34].

B. APPLICATIONS OF DEEP-LEARNING WITH MEDICAL IMAGES

Deep learning is widely used in medical domain and popular architectures for classification of deep CNN.

1) TEXT IMAGE/ CLASSIFICATION

The first main contribution to the field of medical image analysis is the shape of the image. Depending on the application, single or multiple images are taken as the input and output is the diagnosis. In this technique, smaller size datasets are used during iterations, while in computer vision, a large amount of data is used with millions of samples. Deep convolution neural networks are also trained by using transfer learning mode, and for this case, a pre-trained deep-network is used on a new set of images to perform image classification-based tasks. There are two main types of learning which are: 1) use of the pre-trained deep convolution network for feature extraction,

2) Accurate fine-tuning of a trained deep network for image classification-based problem. The key advantage of transfer learning is the low training cost of the network, as the network is pre-trained and requires new samples according to the domain requirement. According to the literature [35], [36], feature extraction with fine-tuning can enhance the classification accuracy of a deep network. The authors reported the results for multi-class knee image dataset [35]. While in [36], results are reported for cytopathology image while dealing with a medical image classification-based problem. According to the work of [30], the pre-trained GoogleNet v3 with fine-tuning can outperform another deep network for the medical image classification case. Stack autoencoder (SAE) and Restricted Boltzmann Machines Simplified (RBMs) are examples of unsupervised classification techniques that are initially used for medical image datasets.

Medical image classification researchers are reported for neuro imaging and procedures based on DBNs and SEAs are used with Magnetic Resonance Imaging (MRI) for the patients who were suffering from Alzheimer disease [37]–[40]. The recent trends for medical image classification are shifted to the use of CNNs [41]. CNNs have shown good results for many cases including MRI and Computed Tomography (CT). 3D image slices have used the input to cure Alzheimer patients. This disease damages the human mind. MRI images are used with a CNNs framework for image classification-based problem [42], [43]. To enhance the performance of image classification, new layers are added in the deep network architecture so that the network can extract features using edge to edge, edge to node and node to graph layers. From all the above-stated discussion it can be concluded that CNNs can be explored further in medical domain for image classification [44], [45].

2) OBJECT AND LESION CLASSIFICATION

For the purpose of object classification, we can divide a medical image into two or more classes such as brain image classification for CT-scan images. Local and global contextual information is extracted from the images to identify the position of lesion that is beneficial for image classification. This approach is unexplored with deep learning algorithms. Many researchers have studied multi-scale fashion using multi-stream algorithms. Shen *et al.* [46] proposed a deep CNN by using fusion patches from start to last layers of the feature vector. Kawahara and Hamarneh [47] increased the number of layers for CNN and address the problem when there is a difference in sizes of image/difference of image resolution. In [48], a hybrid approach is applied by using a combination of CNNs and RNNs to divide multi-level nuclear images. The combination is beneficial as contextual information from different sizes of images can be extracted through this fusion.

For medical image classification-based problems, 3D data must be supplied at the input of CNN so that more information can be derived to obtain a high efficiency as output. Different custom architectures are proposed in the literature

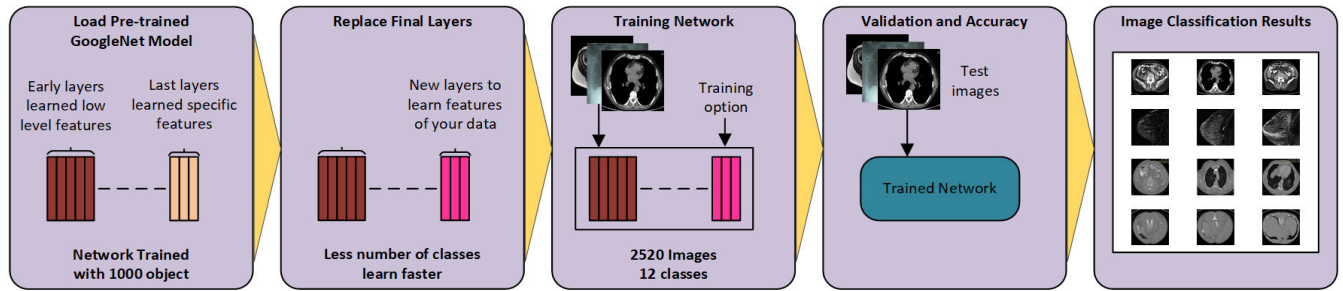


FIGURE 4. GoogleNet fine tuning transfer learning.

to extract the 3D data with CNNs [49]. In [49], multi-level CNNs architecture is applied to obtain the interest points in chest CT-scan images. The researcher used nine patches in convolutional layers and joint the entire fully connected layer just before the output layer. Research for medical image classification domain is also reported while using RBMs [50], [51], SAEs [51] and CSAE [52]. The main difference between these techniques is unsupervised pre-trained networks auto-encoders. In [53], a hybrid approach is proposed that is based on the use of both supervised and unsupervised feature-based training without the use of handcrafted features. The experimental results presented in further validates that the MIL framework outperformed handcrafted features. According to the literature, the medical image annotation is still a challenging research task for the community of computer vision researchers. From the above discussion, it can be concluded that the use of a pre-trained CNN is a recommended framework for classification.

III. PROPOSED METHOD

A. DATASET

In the proposed method, we have selected a publicly available dataset that consists of images with different classes of human body parts such as chest, breast etc. There is a total number of 12 classes and it is important to mention that 11 classes are collected from a public dataset of cancer image archives, while 12 classes are collected from Messidor that is collected from an open-access website of knee images. Each class contains 300 images and for the whole dataset, there are 3600 images with 12 different classes. We used a training-testing framework; therefore, we selected a random number of 70 percent and 30 percent as train-test ratio in our case. Therefore, 2520 images are used for training and 1080 are used for testing. There is no common image used for both the training and testing process. In the first step, all images are converted from DICOM (Digital Imaging and Communication in Medicine) to JPG format. Due to the increase in variation, there can be a more complex feature matrix for the neural network to classify more accurately. To handle this issue the intensity normalization was applied before feeding the data to the convolutional neural network.

B. PRE-PROCESSING

Nowadays, deep learning is widely used in many applications, which are based on artificial neural networks and

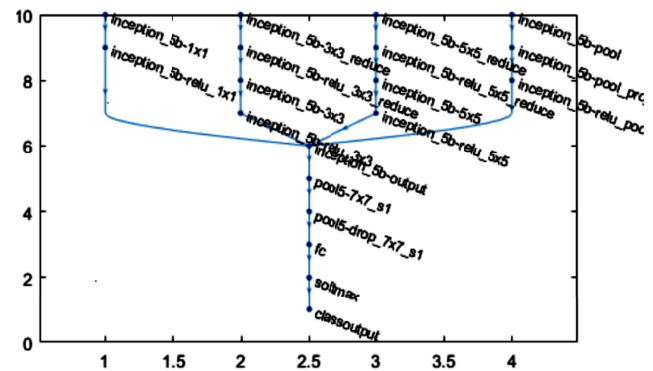


FIGURE 5. Layers of GoogleNet method after transfer learning.

treated like as human brain. The feed forward neural networks, comprising of many hidden layers are good examples of models with deep architecture. All the images are resized to achieve the same distance scale that is physical space represented by each pixel in an image. After applying the normalization, each image was converted to 224×224 dimension, an input image is 224×224 which is fed to first convolution layers. The first convolutional layer is applied using the kernel of 4×4 with the same padding, stride of 1 and 8 filters by using the activation function of ReLu. All the images are resized to 224×224 pixels so that they can be used as an input for GoogleNet method as shown in Figure 3.

C. REUSE PRE-TRAINED NETWORK

Characteristically, DCNN image classification techniques have depended two phases, one is feature extraction and second is classification module. To extract the feature by providing the training images and learned after that classifier of softmax layer is used from the training image data in an end-to-end learning framework. The model used for training consisted of multiple layers, out of which five are convolutional layers and three are fully connected layers, as depicted in Figure 4. The algorithm of deep learning learns from the images as low-level, mid-level and abstract features directly which is against for handcrafted features. A training set of images are used as an input for GoogleNet pre-trained method to extract the deep features. In our case, we used 144 layers including convolution layers, fully connected layers and applied set of images with different modalities. We relocated features of the GoogleNet method by applying both

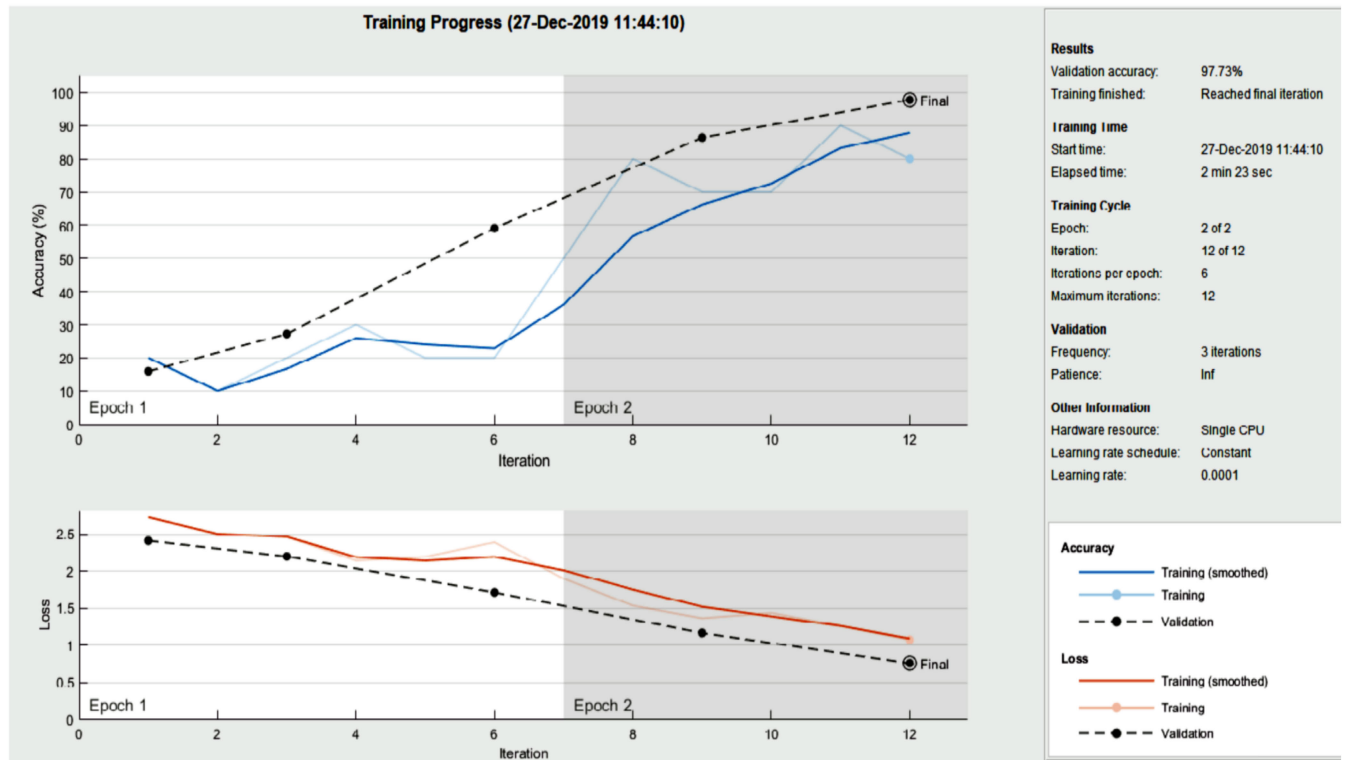


FIGURE 6. Implementation of proposed method using MATLAB.

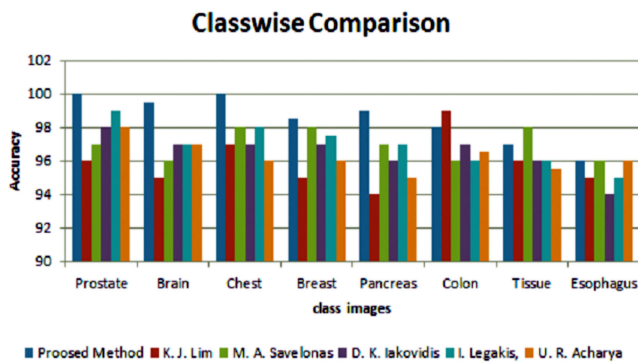


FIGURE 7. Classwise comparison of the proposed method and state-of-the-art methods.

training and validation from our medical images dataset. The aim of soft-function is to perform re-learning according to 12 classes of the dataset. The effective rate to start training for method layers is from 0.01 to 0.0001 using stochastic gradient descent. GoogleNet is already trained for 1000 different semantic classes by using features from a large-scale dataset. Transfer learning is a technique of deep networks through which we can re-train a network according to the new levels through fine-tuning of parameters. In fine tuning process, features are extracted from the given set of images. Max-pooling layer is applied to reduce the size of the input using the kernel of 4×4 , same padding, the stride of 1 and 8 filters. The output of this pooling layers gives 112×112 dimensional output. The centers of neighborhood neurons in the kernel map is a distance which is called stride. The output

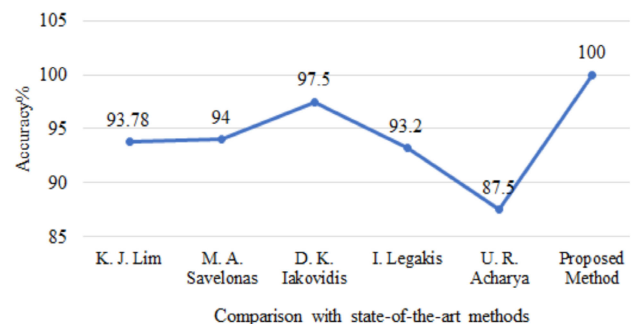


FIGURE 8. Classification results of the proposed method with state-of-the-art methods.

of first convolutional layer is fed to a nonlinearity after that the spatial max pooling layer for summarizing neighboring neurons. Rectified linear unit (ReLU) is used for nonlinearity to the output of all fully connected layers. The training and validation results from the proposed DCNN model is shown in Figure 11.

For a pre-trained GoogleNet method, images are used as input to replace the last three layers that are loss3-classifier, prob, and output, with the aim to re-join these layers with rest of the network. In our case, we added a fully connected layer, a soft-max layer, and a classification output layer to the pre-trained GoogleNet method. Soft max function to a N-dimensional vector, which scales the values of the vector in the range of (0, 1) and its summation gives a value of 1 to represent the each class. The last fully connected layer is of the same size that is a number of classes on our dataset that is 12.

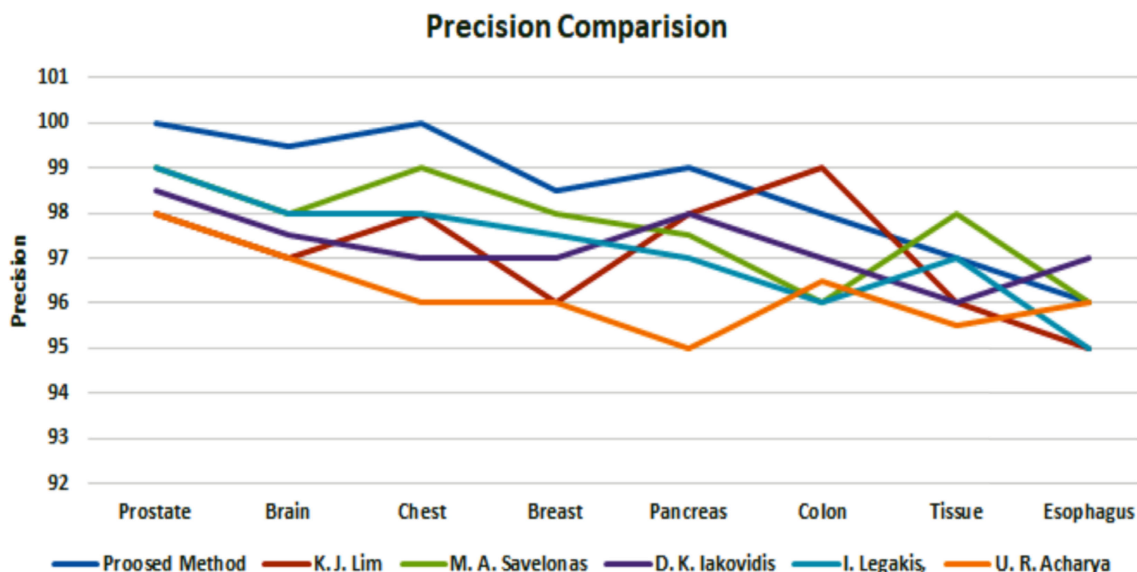


FIGURE 9. Precision analysis of the proposed method with state-of-the-art methods.

To enhance the learning process of GoogleNet, we increased the learning rate factors of the fully connected layer by connecting the transferred layer with a remaining network as shown in Figure 5. For learning process, stochastic gradient descent method is used with a low learning rate of 0.0001 with maximum of 30 epochs. In stochastic gradient descent method used an objective function which is known as negative log likelihood and it is very efficient in learning discriminative linear classifiers under a convex loss function like SVM or Logistic Regression. This function is widely used for classification problems. Once the DCNN model is successfully optimized and trained for classifying the medical images.

IV. EXPERIMENTAL RESULTS

In this manuscript, a popular and widely-used deep learning technique has been used for developing and training the proposed deep convolution neural network for classifying the medical images. In this regard, we performed image pre-processing as the first step and later performed fine-tuning to enhance the classification accuracy with reduced training time. The artifacts in the images are removed through pre-processing. Table 1 presents a short summary with a number of medical images that are used for training and validation. The training time computed in our experiments is a bit higher, as we are using the CPU-based framework as shown in Figure 6. We used two epochs, for each epoch the network performs 252 iterations. Therefore, the method takes 504 iterations to complete the training process. The validation process reflects the efficiency and accuracy of our proposed method.

In the training phase of transfer learning mode, we used the parameters that can avoid over-fitting of the network as classification accuracy of deep network decrease due to over-fitting [54]. The techniques such as re-scaling and rotation can be used to retrain a deep network. In our case,

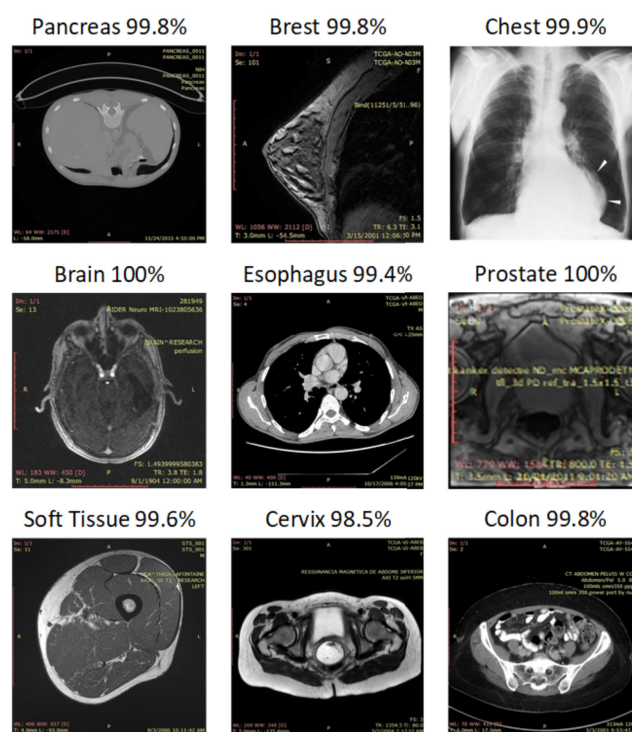


FIGURE 10. The classification results using DCNN.

we used cropping feature only in order to avoid over-fitting. In training phase, first three layers has been used regularization layer to avoid the problem of over-fitting. The experimental results show that the proposed method in this paper outperformed state-of-the-art methods in term of classification accuracy as shown in Figure 8. Table 3 presents a detailed comparison of the proposed method with state-of-the-art methods. Figure 7 shows the comparison accuracy of the proposed method with state-of-the-art methods. Figure 9 shows the performance of proposed method in term of precision with

TABLE 1. Training and testing data.

Class	Images in each Class	Training	Validation
Brain	300	210	90
Breast	300	210	90
Chest	300	210	90
Bladder	300	210	90
Colon	300	210	90
Cervix	300	210	90
Eye	300	210	90
Esophagus	300	210	90
Lymph	300	210	90
Pancreas	300	210	90
Prostate	300	210	90
Soft Tissue	300	210	90
Total Images	3600	2520	1080

TABLE 2. Proposed method class wise accuracy results.

Image Class	Accuracy
Prostate	100
Brain	100
Chest	99.9
Breast	98.8
Pancreas	99.8
Colon	97.8
Soft Tissue	98.6
Esophagus	99.4

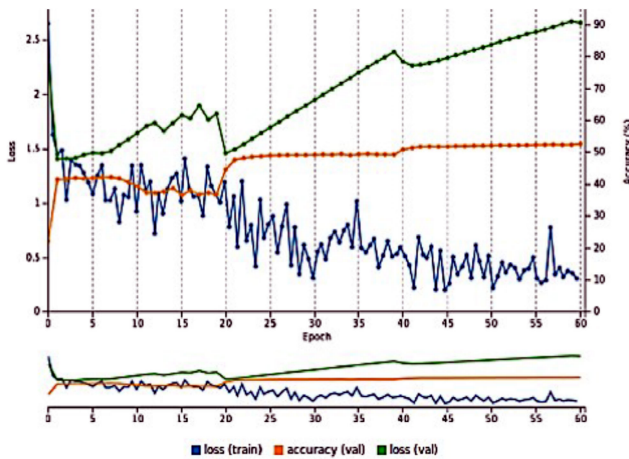


FIGURE 11. Error with each epoch in Training and Validation for classifying medical image modalities.

other state-of-the-art methods. Figure 10 shows class-wise classification accuracy results that are obtained while using 70-30 training-testing ratio. Figures 12 and 13 show that proposed method is achieved highest accuracy with other methods. Table 2 shows the class-wise classification accuracy values; we achieved 100% accuracy for the case prostate and brain, 99.9% accuracy for the brain, 99.8% accuracy for chest, breast and pancreas, 99.6 for soft-tissues while 99.4 for the class esophagus. All of these results show that proposed method based on deep convolution neural network is more reliable and efficient than the state-of-the-art methods in term of classification accuracy.

MATLAB is used to implement the proposed method with following hardware and software specification: Windows 7, Intel core i5, 1.60 GHz-2.30GHz with 4GB RAM. The training process took 682 minutes and 10 seconds. Table 3 shows

TABLE 3. Accuracy comparison of the proposed method with state-of-the-art methods.

Methods	Feature	Technique	Testing Samples	Accuracy
[55]	Shape features	ANN	190 thyroid lesions	93.78%
[56]	histogram	ANN	190 thyroid lesions	93.78%
[57]	Hybrid features	SVM	250 thyroid ultra-sound images	97.50%
[58]	Texture features	SVM	142 Longitudinal images	93.20%
[59]	Higher-order features	Linear Discriminant analysis	98 module image sequences	87.50%
Proposed method	Global features	Softmax GoogleNet	1080 different medical images	97.73%

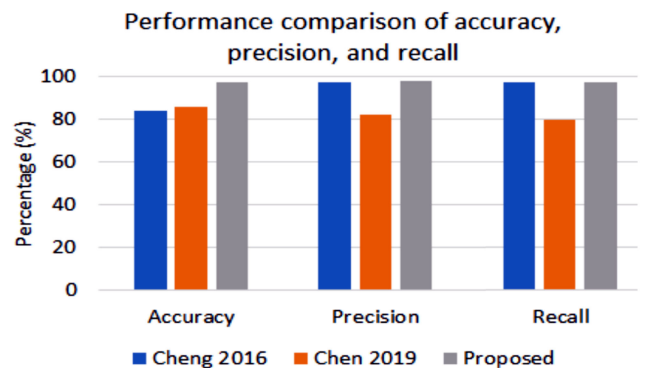


FIGURE 12. Performance comparison of accuracy, precision, and recall.

the results of the Brain class, prostate class achieved 100% accuracy it means that DCNN classified the images according to their class correctly, while remaining classes like chest, colon, esophagus and soft tissue got 98.9%, 97.8%, 98.4 and 96% and 99.6% accuracy. The overall performance of the proposed method accuracy is 98%. To evaluate the performance of proposed method for classification task in terms of precision, recall and accuracy.

A. SYSTEM AND RUNNING TIME

The simulations have been performed on Dell Inspiron 5520 Laptop having Intel Core i3 CPU with clock speed of 2.40 GHz with a RAM of 4.00 GB. The execution time of Proposed method is 682 min 10 sec. The proposed method has been evaluated in terms of classification results.

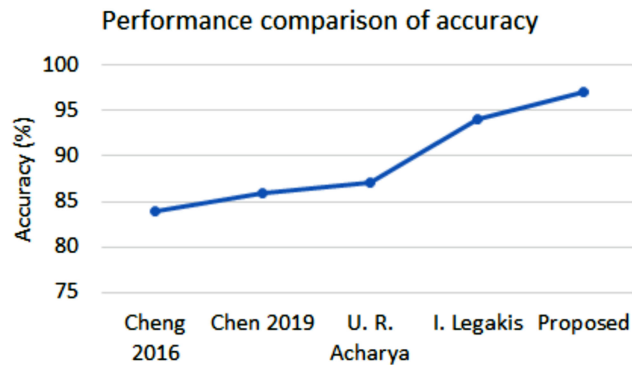


FIGURE 13. Performance comparison of accuracy with other methods.

V. CONCLUSION AND FUTURE WORK

For classification purpose, deep learning based framework for medical image classification by training the images is proposed. In this regard, diagnosis is one of the main requirements of the existing era and investigated or examine to specific diseases. The use of computer-aided tools and reliable image analysis are the main aspects that can increase the efficiency of doctors and physicians. It is a requirement of the current era to develop such image processing methods that can help doctors in various fields of medical science. Such methods are beneficial to save human life and it is evident that diseases can be predicted before they can affect the human body. Since the last few decades, the computer vision research community is trying to reduce this gap by developing automated systems which can process medical images using machines to make decisions. We have proposed a novel deep convolution network-based approach that is assist of doctors and physicians in making reasonable decisions. The results obtained from the proposed method outperformed state-of-the-art methods that is reported for the same dataset. In future, we aim to explore large-scale image datasets for medical image classification and detection problems.

REFERENCES

- [1] Q. Zhu, B. Du, and P. Yan, "Boundary-weighted domain adaptive neural network for prostate MR image segmentation," *IEEE Trans. Med. Imag.*, vol. 39, no. 3, pp. 753–763, Mar. 2020.
- [2] Q. Zhu, B. Du, P. Yan, H. Lu, and L. Zhang, "Shape prior constrained PSO model for bladder wall MRI segmentation," *Neurocomputing*, vol. 294, pp. 19–28, Jun. 2018.
- [3] Q. Zhu, B. Du, B. Turkbey, P. Choyke, and P. Yan, "Exploiting interslice correlation for MRI prostate image segmentation, from recursive neural networks aspect," *Complexity*, vol. 2018, pp. 1–10, Feb. 2018.
- [4] K. Kranthi Kumar and T. V. Gopal, "A novel approach to self order feature reweighting in CBIR to reduce semantic gap using relevance feedback," in *Proc. Int. Conf. Circuits, Power Comput. Technol. (ICCPCT)*, Mar. 2014, pp. 1437–1442.
- [5] R. Ashraf, K. B. Bajwa, and T. Mahmood, "Content-based image retrieval by exploring banded regions through support vector machines," *J. Inf. Sci. Eng.*, vol. 32, no. 2, pp. 245–269, 2016.
- [6] J. Wan, D. Wang, S. C. H. Hoi, P. Wu, J. Zhu, Y. Zhang, and J. Li, "Deep learning for content-based image retrieval: A comprehensive study," in *Proc. ACM Int. Conf. Multimedia (MM)*, 2014, pp. 157–166.
- [7] F. Shaukat, G. Raja, R. Ashraf, S. Khalid, M. Ahmad, and A. Ali, "Artificial neural network based classification of lung nodules in ct images using intensity, shape and texture features," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 10, pp. 4135–4149, 2019.
- [8] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1725–1732.
- [9] G. Wu, W. Lu, G. Gao, C. Zhao, and J. Liu, "Regional deep learning model for visual tracking," *Neurocomputing*, vol. 175, pp. 310–323, Jan. 2016.
- [10] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.
- [11] S. Zhou, Q. Chen, and X. Wang, "Active deep learning method for semi-supervised sentiment classification," *Neurocomputing*, vol. 120, pp. 536–546, Nov. 2013.
- [12] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587.
- [13] H.-Y. Shi, "Pancreatic carcinosarcoma: First literature report on computed tomography imaging," *World J. Gastroenterol.*, vol. 21, no. 4, p. 1357, 2015.
- [14] M. R. Zare, W. C. Seng, and A. Mueen, "Automatic classification of medical X-ray images using a bag of visual words," *IET Comput. Vis.*, vol. 7, no. 2, pp. 105–114, Apr. 2013.
- [15] S. Khan, S.-P. Yong, and J. D. Deng, "Ensemble classification with modified SIFT descriptor for medical image modality," in *Proc. Int. Conf. Image Vis. Comput. New Zealand (IVCNZ)*, Nov. 2015, pp. 1–6.
- [16] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany: Springer, 2013, pp. 411–418.
- [17] P. Kanerva, "Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors," *Cognit. Comput.*, vol. 1, no. 2, pp. 139–159, Jun. 2009.
- [18] O. Yilmaz, "Machine learning using cellular automata based feature expansion and reservoir computing," *J. Cellular Automata*, vol. 10, pp. 435–472, Sep. 2015.
- [19] D. Cheng, G. Meng, G. Cheng, and C. Pan, "SeNet: Structured edge network for Sea-Land segmentation," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 2, pp. 247–251, Feb. 2017.
- [20] P. Chen, Y. Song, D. Yuan, and Z. Liu, "Feature fusion adversarial learning network for liver lesion classification," in *Proc. ACM Multimedia Asia ZZZ*, Dec. 2019, pp. 1–7.
- [21] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2007, pp. 153–160.
- [22] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.
- [23] D. E. Arroyo, "Visualizando neuronas en redes neuronales convolucionales," Public Univ. Navarre, Pamplona, Spain, Tech. Rep. 2454/33694, 2019.
- [24] S. Al-Hadhrani, S. Altuwaijri, N. Alkharashi, and R. Ouni, "Deep classification technique for density counting," in *Proc. 2nd Int. Conf. Comput. Appl. Inf. Secur. (ICCAIS)*, May 2019, pp. 1–6.
- [25] R. Ashraf, M. Ahmed, U. Ahmad, M. A. Habib, S. Jabbar, and K. Naseer, "MDCBIR-MF: Multimedia data for content-based image retrieval by using multiple features," *Multimedia Tools Appl.*, vol. 79, pp. 1–27, Jun. 2018.
- [26] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
- [27] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [28] R. Ashraf, M. Ahmed, S. Jabbar, S. Khalid, A. Ahmad, S. Din, and G. Jeon, "Content based image retrieval by using color descriptor and discrete wavelet transform," *J. Med. Syst.*, vol. 42, no. 3, p. 44, Mar. 2018.
- [29] Y. Liu, K. Gadepalli, M. Norouzi, G. E. Dahl, T. Kohlberger, A. Boyko, S. Venugopalan, A. Timofeev, P. Q. Nelson, G. S. Corrado, J. D. Hipp, L. Peng, and M. C. Stumpe, "Detecting cancer metastases on gigapixel pathology images," 2017, *arXiv:1703.02442*. [Online]. Available: <http://arxiv.org/abs/1703.02442>

- [30] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, p. 115, 2017.
- [31] K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Med. Image Anal.*, vol. 36, pp. 61–78, Feb. 2017.
- [32] R. Ashraf, K. Bashir, A. Irtaza, and M. Mahmood, "Content based image retrieval using embedded neural networks with banditized regions," *Entropy*, vol. 17, no. 6, pp. 3552–3580, May 2015.
- [33] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1252–1261, May 2016.
- [34] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Berlin, Germany: Springer, 2013, pp. 246–253.
- [35] J. Antony, K. McGuinness, N. E. O'Connor, and K. Moran, "Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 1195–1200.
- [36] E. Kim, M. Corte-Real, and Z. Baloch, "A deep semantic mobile application for thyroid cytopathology," *Med. Imag., Inform., Next Gener. Innov., Int. Soc. Opt. Photon.*, vol. 9789, Apr. 2016, Art. no. 97890A.
- [37] T. Brosch and R. Tam, "Manifold learning of brain MRIs by deep learning," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Berlin, Germany: Springer, 2013, pp. 633–640.
- [38] S. M. Plis, D. R. Hjelm, R. Salakhutdinov, E. A. Allen, H. J. Bockholt, J. D. Long, H. J. Johnson, J. S. Paulsen, J. A. Turner, and V. D. Calhoun, "Deep learning for neuroimaging: A validation study," *Frontiers Neurosci.*, vol. 8, p. 229, Aug. 2014.
- [39] H.-C. Shin, M. R. Orton, D. J. Collins, S. J. Doran, and M. O. Leach, "Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1930–1943, Aug. 2013.
- [40] H.-I. Suk, S.-W. Lee, and D. Shen, "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis," *NeuroImage*, vol. 101, pp. 569–582, Nov. 2014.
- [41] A. Menegola, M. Fornaciali, R. Pires, S. Avila, and E. Valle, "Towards automated melanoma screening: Exploring transfer learning schemes," 2016, *arXiv:1609.01228*. [Online]. Available: <http://arxiv.org/abs/1609.01228>
- [42] A. J. Hoffman and R. R. Singleton, "On moore graphs with diameters 2 and 3," in *Selected Papers Of Alan J Hoffman: With Commentary*. Singapore: World Scientific, 2003, pp. 377–384.
- [43] A. Payan and G. Montana, "Predicting Alzheimer's disease: A neuroimaging study with 3D convolutional neural networks," 2015, *arXiv:1502.02506*. [Online]. Available: <http://arxiv.org/abs/1502.02506>
- [44] J. Kawahara, C. J. Brown, S. P. Miller, B. G. Booth, V. Chau, R. E. Grunau, J. G. Zwicker, and G. Hamarneh, "BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, Feb. 2017.
- [45] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3156–3164.
- [46] 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.
- [47] J. Kawahara and G. Hamarneh, "Multi-resolution-tract cnn with hybrid pretrained and skin-lesion trained layers," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, Cham, Switzerland: Springer, 2016, pp. 164–171.
- [48] X. Gao, S. Lin, and T. Y. Wong, "Automatic feature learning to grade nuclear cataracts based on deep learning," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 11, pp. 2693–2701, Nov. 2015.
- [49] D. Nie, H. Zhang, E. Adeli, L. Liu, and D. Shen, "3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Cham, Switzerland: Springer, 2016, pp. 212–220.
- [50] G. van Tulder and M. de Bruijne, "Combining generative and discriminative representation learning for lung CT analysis with convolutional restricted Boltzmann machines," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1262–1272, May 2016.
- [51] Q. Zhang, Y. Xiao, W. Dai, J. Suo, C. Wang, J. Shi, and H. Zheng, "Deep learning based classification of breast tumors with shear-wave elastography," *Ultrasonics*, vol. 72, pp. 150–157, Dec. 2016.
- [52] M. Kallenberg, K. Petersen, M. Nielsen, A. Y. Ng, P. Diao, C. Igel, C. M. Vachon, K. Holland, R. R. Winkel, N. Karssemeijer, and M. Lillholm, "Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1322–1331, May 2016.
- [53] Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, and E. I.-C. Chang, "Deep learning of feature representation with multiple instance learning for medical image analysis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 1626–1630.
- [54] D. Soekhoe, P. Van Der Putten, and A. Plaat, "On the impact of data set size in transfer learning using deep neural networks," in *Proc. Int. Symp. Intell. Data Anal.*, Cham, Switzerland: Springer, 2016, pp. 50–60.
- [55] K. J. Lim, C. S. Choi, D. Y. Yoon, S. K. Chang, K. K. Kim, H. Han, S. S. Kim, J. Lee, and Y. H. Jeon, "Computer-aided diagnosis for the differentiation of malignant from benign thyroid nodules on ultrasonography," *Acad. Radiol.*, vol. 15, no. 7, pp. 853–858, 2008.
- [56] M. A. Savelonas, D. E. Maroulis, D. K. Iakovidis, and N. Dimitropoulos, "Computer-aided malignancy risk assessment of nodules in thyroid US images utilizing boundary descriptors," in *Proc. Panhellenic Conf. Inform.*, Aug. 2008, pp. 157–160.
- [57] D. K. Iakovidis, E. G. Keramidas, and D. Maroulis, "Fusion of fuzzy statistical distributions for classification of thyroid ultrasound patterns," *Artif. Intell. Med.*, vol. 50, no. 1, pp. 33–41, Sep. 2010.
- [58] I. Legakis, M. A. Savelonas, D. Maroulis, and D. K. Iakovidis, "Computer-based nodule malignancy risk assessment in thyroid ultrasound images," *Int. J. Comput. Appl.*, vol. 33, no. 1, pp. 29–35, 2011.
- [59] U. R. Acharya, S. V. Sree, G. Swapna, S. Gupta, F. Molinari, R. Garberoglio, A. Witkowska, and J. S. Suri, "Effect of complex wavelet transform filter on thyroid tumor classification in three-dimensional ultrasound," *Proc. Inst. Mech. Eng. H, J. Eng. Med.*, vol. 227, no. 3, pp. 284–292, Mar. 2013.



REHAN ASHRAF received the M.S. degree from the Center for Advanced Studies in Engineering (CASE), Islamabad, Pakistan, in 2011, and the Ph.D. degree from the University of Engineering and Technology Taxila, Pakistan. He was a Full Time Ph.D. Scholar with the University of Engineering and Technology Taxila. He is currently an Assistant Professor with the Department of Computer Science, Faculty of Science, National Textile University, Faisalabad. He performed his research with the Computer Vision Group, UET Taxila. His research interests include content-based image retrieval (CBIR), image/video processing, machine learning, computer vision, artificial intelligence, and medical image processing. He serves as a Reviewer for *Technical Journal* (UET Taxila), *IEEE Access*, *Multimedia Tools and Applications*, *Eurasip Journal on Image and Video Processing*, *KSII Transaction on Internet and Information Systems*, and so on.



MUHAMMAD ASIF HABIB received the Ph.D. degree from JKU, Linz, Austria. He is currently an Associate Professor with the Department of Computer Science, National Textile University, Faisalabad, Pakistan. His research interests include information network security, authorization, role-based access control, the IoT, cloud and grid computing, association rule mining, recommender systems, wireless sensor networks, block chain, and vehicular networks. He serves as a Technical Reviewer for top tier journals and conferences.



MUHAMMAD AKRAM received the B.S. degree in computer engineering from the Balochistan University of Information Technology, Engineering, and Management Sciences, Quetta, Pakistan, the M.S. degree in software engineering from Hamdard University, Pakistan, and the Ph.D. degree in computer engineering from Sungkyunkwan University, South Korea, in 2018. He is currently an Assistant Professor with the Department of Software Engineering, Balochistan University of Information Technology, Engineering, and Management Sciences. His research interests include wireless sensor networks, machine learning, and context-aware computing.



MUHAMMAD AHSAN LATIF received the Ph.D. degree in computer science from Alpen-Adria University, Klagenfurt, Austria, in 2011. He is currently an Assistant Professor with the Department of Computer Science, University of Agriculture, Faisalabad, Pakistan. He is also Leading a Research Group aiming to map crop health under different nonlinear conditions using unmanned aerial systems and vision sensors. His contributions are research articles, book chapters, a complete book on digital image processing, and various research projects. His main research interests include image processing, computer vision, and artificial intelligence in the agricultural context.



MUHAMMAD SHERAZ ARSHAD MALIK received the Ph.D. degree in information technology. He has more than seven years of research and industrial experience. He is currently an Assistant Professor with the Department of Information Technology, Government College University Faisalabad, Pakistan. His research interests include information visualization, temporal data, data analytics, and the Internet of Things.



MUHAMMAD AWAIS received the master's and M.Phil. degrees in computer science from the University of Agriculture, Faisalabad, Pakistan, in 2001 and 2004, respectively, the M.Phil. degree in applied information science from Albert-Ludwigs University, Freiburg, Germany, in 2008, and the Ph.D. degree in robotics from the University of Bayreuth, Germany, in 2013. In 2005, he joined the Government College University Faisalabad, as a Lecturer and an Assistant Professor, in 2016. His current research interests include machine learning application in software engineering, human-robot interaction, intention estimation, and machine learning.



SAADAT HANIF DAR received the Ph.D. degree in electrical engineering from Riphah International University Islamabad, Pakistan, in 2018. He is currently an Associate Professor and the Chairperson with the Department of Software Engineering, Mirpur University of Science and Technology (MUST), Mirpur. He has authored number of scientific publications in the field of wearable antennas for biocompatible applications. His research interests include wearable antennas, body area networks, biomedical systems, RFIDs, and modeling of electromagnetic.



TOQEER MAHMOOD received the M.S. degree in computer engineering from the Center for Advanced Studies in Engineering (CASE), Islamabad, Pakistan, in 2010, and the Ph.D. degree in computer engineering from the University of Engineering and Technology Taxila, Pakistan, in 2017. He is currently an Assistant Professor with the Department of Computer Science, National Textile University (NTU), Faisalabad. He has authored or coauthored many scientific articles in conferences and journals of international repute. His research interests include image processing, image retrieval, steganography, and numerical techniques with particular attention to multimedia forensics. He serves as a Reviewer for *Technical Journal* (UET Taxila), *Journal of King Saud University Computer and Information Sciences*, *ETRI Journal*, *Journal of Information Security and Applications*, *Australian Journal of Forensic Sciences*, *Forensic Science International*, and so on.



MUHAMMAD YASIR received the M.S. degree in computer science from SZABIST, Karachi, Pakistan. He is currently pursuing the Ph.D. degree in computer science with National Textile University, Faisalabad, Pakistan. He is also an Assistant Professor with the Department of Computer Science, University of Engineering and Technology Lahore, Faisalabad, Pakistan. His research interests include data science, machine learning, agent-based modeling and simulation, network security, cryptography, and the IoT.



ZAHOOOR ABBAS received the M.Sc. degree in computer science from the University of Agriculture, Faisalabad, Pakistan, in 2016, and the M.S. degree in computer science from National Textile University, Faisalabad, in 2019. He is currently a SSE in computer science with the School Education Department, Punjab. His research interests include deep learning and medical imaging.

...