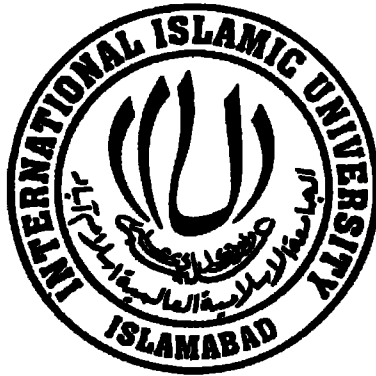


Enhanced Tumour Classification Using Novel Feature Extraction Method



By

Imdad Ali

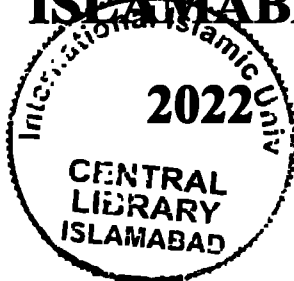
Reg. No. 101-FET/PHDEE/F15

**A dissertation submitted to I.I.U. in partial fulfillment of the
requirements for the degree of**

DOCTOR OF PHILOSOPHY

**Department of Electrical and Computer Engineering
Faculty of Engineering and Technology**

**INTERNATIONAL ISLAMIC UNIVERSITY
ISLAMABAD**



Accession No. 621.381

PhD
621.381

TME

Electronic engineering
Electronics technology

Enhanced Tumour Classification Using Novel Feature Extraction Method



**Imdad Ali
101-FET/PHDEE/F15**

**Submitted in partial fulfillment of the requirements for the PhD degree in Electronic
Engineering at the Department of Electrical and Computer Engineering
Faculty of Engineering and Technology
International Islamic University,
Islamabad**

Supervisor

Prof. Dr. Muhammad Amir

June, 2022

Copyright © 2022 by Imdad Ali

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without permission of the author.

DEDICATED TO

My Teachers,

Parents,

Brothers,

Sisters,

Friends,

Wife,

and Daughter

CERTIFICATE OF APPROVAL

Title of Thesis: Enhanced Tumour Classification Using Novel Feature Extraction Method

Name of Student: IMDAD ALI

Registration No: 101-FET/PhDEE/F15

Accepted by the Department of Electrical and Computer Engineering, Faculty of Engineering and Technology, International Islamic University (IIU), Islamabad, in partial fulfillment of the requirements for the Doctor of Philosophy degree in Electronic Engineering.

Viva voce committee:

Dr. Muhammad Amir (Supervisor)

Professor DECE, FET IIU Islamabad

Dr. Ihsan ul Haq (Co-Supervisor)

Associate Professor DECE, FET IIU Islamabad

Prof. Dr. Aqdas Naveed Malik (Internal)

Professor DECE, FET, IIU Islamabad.

Prof. Dr. Tanveer Ahmed Cheema (External-I)

Professor/Dean DEE, ISRA University Islamabad

Dr. Hafiz Muhammad Faisal Zafar (External-II)

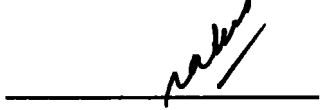
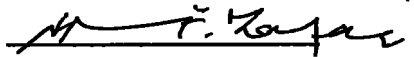
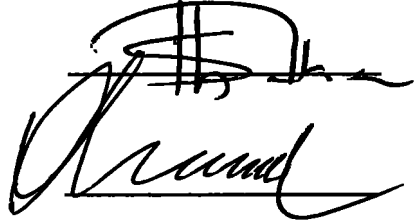
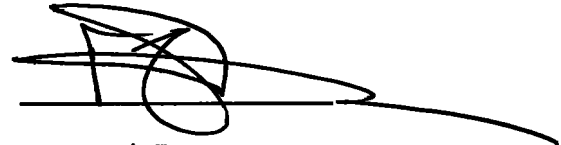
Chief Scientist, PAEC, Islamabad.

Prof. Dr. Suheel Abdullah Malik (Chairman, DEE)

Professor DEE, FET, IIU Islamabad.

Prof. Dr. Nadeem Ahmad Sheikh (Dean, FET)

Professor DME, FET, IIU Islamabad.



Assante

Abstract

Lung cancer has become a major threat to human health worldwide. Early detection of lung cancer can increase the survival rate of the patient by starting the treatment at the right time. Malignancy detection and classification in computed tomography (CT) images is a very tedious and time-consuming task for radiologists, inspiring researchers to develop enhanced techniques for computer-aided diagnosis (CAD). It is observed that the available detection and classification techniques have many false positives and false negatives because the detection algorithms have high sensitivity for lesions. Therefore, some non-lesion structures (e.g., blood vessels) are labelled as lesion wrongly in initial identification step. The objective of this dissertation is to achieve a better diagnosis of lung cancer using enhanced tumour/nodule classification techniques, such as; transfer learning, deep learning, ensemble learning, and medical image fusion with novel feature extraction techniques.

This thesis presents various novel lung nodules malignancy classification techniques for a CAD system. A transferable texture CNN is proposed for efficient lung nodule classification in which an energy layer (EL) is introduced to extract the texture feature map from the convolutional layer. Incorporating EL reduces the proposed network learnable parameters, which further reduces the computational complexity and memory requirements. The pre-trained model of proposed texture CNN is also utilized to tackle the smaller medical image dataset classification issues using transfer learning methodology.

The decision level fusion based on deep feature selection is also proposed to enhance the performance of the CAD system for lungs nodule classification in CT images. First, the performance of AdaBoostM2 and SVM classifiers is evaluated using deep features from

eight state-of-the-art transferable DCNN architectures, which are; VGG-16, VGG-19, GoogLeNet, variants of the residual network (ResNet-18, ResNet-50, and ResNet-101), Inception-V3, and InceptionResNet-V2. After that, the optimal deep features are selected from the DCNNs by identifying the optimal layers, improving classification efficiency. Based on various performance evaluation parameters, it is evident that the proposed lung nodule classification methodologies outperform the state-of-the-art techniques on different datasets, such as; LIDC-IDRI and LUNGx challenge.

LIST OF PUBLICATIONS AND SUBMISSIONS

- [1]. **I. Ali**, M. Muzammil, I. U. Haq, A. A. Khaliq, and S. Abdullah, "Deep Feature Selection and Decision Level Fusion for Lungs Nodule Classification," *IEEE Access*, vol. 9, pp. 18962-18973, 2021.
- [2]. **I. Ali**, M. Muzammil, I. U. Haq, A. A. Khaliq, and S. Abdullah, "Efficient lung nodule classification using transferable texture convolutional neural network," *Ieee Access*, vol. 8, pp. 175859-175870, 2020.
- [3]. M. Muzammil, **I. Ali**, I. U. Haq, A. A. Khaliq, and S. Abdullah, "Pulmonary Nodule Classification Using Feature and Ensemble Learning-Based Fusion Techniques," *IEEE Access*, vol. 9, pp. 113415-113427, 2021.
- [4]. M. Muzammil, **I. Ali**, U. Javed, M. Amir, and I. Ulhaq, "Spatial Stimuli Gradient Based Multifocus Image Fusion Using Multiple Sized Kernels," *Tehnički vjesnik*, vol. 28, pp. 113-122, 2021.
- [5]. **I. Ali**, M. Muzammil, A. Basit, and I. Haq, "Modified Adaptive Predict Hexagon Based Search Motion Estimation Algorithm," *The Nucleus*, vol. 55, pp. 164-170, 2018.
- [6]. M. Muzammil, Z. A. Khan, M. O. Ullah, and **I. Ali**, "Performance analysis of block matching motion estimation algorithms for HD videos with different search parameters," in 2016 International Conference on Intelligent Systems Engineering (ICISE), 2016, pp. 306-311.
- [7]. M. Muzammil, G. Raja, and **I. Ali**, "Field programmable gate array (FPGA) architecture of diamond search motion estimation algorithm for real-time video applications," *NED University Journal of Research*, vol. 12, p. 93, 2015.
- [8]. M. Muzammil, **I. Ali**, M. Sharif, and K. Khalil, "An efficient FPGA architecture for hardware realization of hexagonal based motion estimation algorithm," in 2015 IEEE International Conference on Consumer Electronics-Taiwan, 2015, pp. 422-423.
- [9]. **I. Ali**, G. Raja, M. Muzammil, and A. K. Khan, "Adaptive modified hexagon based search motion estimation algorithm," in 2014 IEEE Fourth International Conference on Consumer Electronics Berlin (ICCE-Berlin), 2014, pp. 147-148.
- [10]. **I. Ali**, M. Muzammil, and G. Raja, "Performance analysis of motion estimation algorithms based on motion activity in video sequences," *Pakistan Journal of Science*, vol. 64, 2012.
- [11]. S. Costantini, Y. Ban, J. Cai, Q. Li, S. Liu, S. Qian, **I. Ali**, et al., "Radiation background with the CMS RPCs at the LHC," *Journal of Instrumentation*, vol. 10, p. C05031, 2015.
- [12]. L. Benussi, S. Bianco, L. Passamonti, D. Piccolo, D. Pierluigi, G. Raffone, **I. Ali**, et al., "Performance of the gas gain monitoring system of the CMS RPC muon detector," *Journal of Instrumentation*, vol. 10, p. C01003, 2015.
- [13]. M. Shopova, A. Aleksandrov, V. Genchev, P. Iaydjiev, M. Rodozov, G. Sultanov, **I. Ali**, et al., "Resistive plate chambers for the LS1 muon upgrade in CMS experiment at LHC," *Bulgarian Chemical Communications*, vol. 47, pp. 186-192, 2015.
- [14]. G. Pugliese, Y. Ban, J. Cai, Q. Li, S. Liu, S. Qian, **I. Ali**, et al., "CMS RPC muon detector performance with 2010-2012 LHC data," *Journal of Instrumentation*, vol. 9, p. C12016, 2014.
- [15]. A. Cimmino, Y. Ban, J. Cai, Q. Li, S. Liu, S. Qian, **I. Ali**, et al., "CMS RPC commissioning of the existing detector during the long shutdown," *Journal of Instrumentation*, vol. 9, p. C10043, 2014.

- [16]. M. S. Kim, Y. Ban, J. Cai, Q. Li, S. Liu, S. Qian, I. Ali, et al., "Web-based monitoring tools for Resistive Plate Chambers in the CMS experiment at CERN," *Journal of Instrumentation*, vol. 9, p. C10031, 2014.
- [17]. J. Goh, M. Kim, Y. Ban, J. Cai, Q. Li, S. Liu, I. Ali, et al., "CMS RPC tracker muon reconstruction," *Journal of Instrumentation*, vol. 9, p. C10027, 2014.
- [18]. S. Colafranceschi, R. Chudasama, L. Pant, A. Mohanty, R. Sehgal, S. Sehgal, I. Ali, et al., "Resistive plate chambers for 2013-2014 muon upgrade in CMS at LHC," *Journal of Instrumentation*, vol. 9, p. C10033, 2014.

The research work presented in this dissertation is based on publications 1 and 2.

ACKNOWLEDGEMENTS

In the name of Allah (SubhanahuWaTa'ala), who is the most kind and the generous. I am grateful to ALLAH SWT for granting me strength, stamina, and direction throughout this work and for thoughts which originated in my mind to complete this research work. Indeed, there is nothing that can happen without HIS will. Peace and blessings of Allah be upon His Last Prophet Muhammad (Sallulah-o-Alaihihe-Wassalam) and all his Sahaba (Razi-Allah-o-Anhum), who sacrificed their lives and wealth for spreading Islam on this globe.

I am grateful to my supervisor, Dr. Muhammad Amir, whose encouragement, concepts, and efforts create an opportunity to complete my advanced education. He appeared as a role model in learning, research, and other parts of life.

I would wish to express gratitude to my co-supervisor, Dr. Ihsan Ul Haq for his enormous guidance, remarks, recommendations, and continuous support throughout this work.

I would also love to pay special thanks to my friend Engr. Muhammad Muzammil and Engr. Shahid Khan for their many extensive, appealing, and rewarding debates, remarks, and recommendations in this work. In every uncertain problem, they came up with solutions and encouraged me to tackle all the challenges. It was difficult to complete my thesis without their assistance. I extend my sincere acknowledgements and gratitude to my associates Dr. Suheel Abdullah, Dr. Umer Javed, Dr. Jawad Ali Shah, and Engr Haris-bin-Nazar for their never-ending assistance and their valuable debates during my advanced studies.

I am surely thankful to my parents (Late), sisters, and brothers, especially Saeed Ahmad and Hafiz Zahoor Ahmad for their affection and sentiments throughout my life. I am also very thankful to my devoted wife for her patience, unending encouragement, and prayers during every stage of my Ph.D. studies. Finally, I am delighted to express my gratitude to my sweet daughter Abiha Fatima, whose sinless and innocent gestures were the source of inspiration for me.

(Imdad Ali)

TABLE OF CONTENTS

Abstract	v
LIST OF PUBLICATIONS AND SUBMISSIONS	vii
ACKNOWLEDGEMENTS	ix
TABLE OF CONTENTS	x
LIST OF FIGRES	xii
LIST OF TABLES	xiii
LIST OF ABBREVIATIONS	xiv
Chapter 1. Introduction	1
1.1 Introduction	1
1.2 Research Objectives	4
1.3 Main Contribution	5
1.4 Thesis organization	7
Chapter 2. Literature Review	9
2.1 Traditional Methods for Lung Nodule Classification	10
2.2 ANN and CNN Based Nodule Classification Techniques	13
2.3 Transfer Learning Based Nodule Classification Techniques	15
2.4 Fusion Based Techniques for Nodule Classification	16
2.5 State-of-the-art DCNNs	17
2.5.1 VGG-Net	17
2.5.2 GoogLeNet	19
2.5.3 Inception-V3	21
2.5.4 ResNet	21
2.5.5 InceptionResNet-V2	22
2.6 Performance Evaluation Metrics	23
2.7 Summary	25
Chapter 3. Transferable Texture CNN for Efficient Lung Nodule Classification	26
3.1 Introduction	26
3.2 Datasets and Methodology	28
3.2.1 Datasets	28
3.2.1.1 LIDC-IDRI Database	28
3.2.1.2 LUNGX Challenge Database	29

3.2.1.3	MNIST dataset	29
3.2.1.4	Image Augmentation.....	30
3.2.1.5	Patch Generation and Data Enhancement.....	30
3.2.2	Architecture of Transferable Texture CNN.....	32
3.2.2.1	Convolutional Layers and Energy Layer	33
3.2.2.2	Batch Normalization and Activation Function	35
3.2.2.3	Pooling Layer.....	36
3.2.2.4	Dropout Regularization.....	37
3.2.2.5	Softmax Classifier and Loss Function	37
3.2.2.6	Back-propagation Algorithm	38
3.2.3	Deep Feature Transfer Technique for Malignancy Classification.....	39
3.3	Training Process.....	41
3.4	Results and Discussion.....	42
3.4.1	The Exploration of Texture CNN Structure.....	42
3.4.2	Performance Evaluation with LIDC-IDRI Database	44
3.4.3	Architecture Complexity Comparison with State-of-the-Art Techniques ..	46
3.4.4	Evaluation of Pre-Trained Mode on LUNGx Challenge Database	47
3.4.5	Performance Validation of Texture CNN on MNIST Dataset.....	48
3.5	Summary	49
Chapter 4. Deep Feature Selection and Decision Level Fusion for Lungs Nodule Classification.....		50
4.1	Introduction	50
4.2	Material and Methods.....	52
4.2.1	Dataset and Pre-Processing	52
4.2.2	Methodology	54
4.2.3	Decision Level Fusion Technique	54
4.4	Experiments and Results	57
4.5	Summary	66
Chapter 5. Conclusions and Future Work.....		68
5.1	Conclusion.....	68
5.2	Future Work	70
REFERENCES		72

LIST OF FIGURES

Fig. 2.1. Traditional Model for Lung Nodule Classification System	10
Fig. 2.2. Architecture of VGG-16 and VGG-19	18
Fig. 2.3. Structure of inception block	20
Fig. 2.4. Basic architecture of GoogLeNet	20
Fig. 2.5. Basic architecture of Inception-V3	21
Fig. 2.6. Architecture of InceptionResNet-V2.....	23
Fig. 3.1. Patch Extraction Process	31
Fig. 3.2. Architecture of Transferable Texture CNN.....	32
Fig. 3.3. Transfer learning methodology using pre-trained texture CNN.....	40
Fig. 3.4. Data distribution detail for training and testing.....	41
Fig. 4.1. Image augmentation	53
Fig. 4.2. Basic scheme of nodule classification by using pre-trained CNN model	54
Fig. 4.3. Decision fusion technique	55
Fig. 4.4. Flow chart of decision fusion	56
Fig. 4.5. Distribution of data for training and testing	57
Fig. 4.6. ROC comparison of SVM and AdaBoostM2 when features from last FC layer	58
Fig. 4.7. AUC comparison of SVM and AdaBoostM2 when features from Last FC layer	59
Fig. 4.8. ROC comparison of SVM and AdaBoostM2 when features from optimal layers	62
Fig. 4.9. AUC comparison of SVM and AdaBoostM2 using features from optimal layers	62
Fig. 4.10. AUC comparison of SVM using features from optimal and last FC layers	63
Fig. 4.11. ROC comparison of decision fusion technique with ResNet-101, GoogLeNet and Inception-V3	65

LIST OF TABLES

Table 2.1. Layer-wise architecture details of VGG-16 and VGG-19	18
Table 2.2. Architectures of ResNet-18, ResNet-50 and ResNet-101	22
Table 2.3. Confusion matrix for two classes.....	23
Table 3.1. Degree of malignancy in LIDC-IDRI database	28
Table 3.2. Layer-wise architecture details of proposed texture CNN.....	33
Table 3.3. Effect of variation in dropout rates with dropout layers of texture CNN	42
Table 3.4. Comparison of classic and texture CNN.....	43
Table 3.5. Classification Score for each class of LIDC-IDRI and LUNGx Challenge	44
Table 3.6. Performance comparison of proposed texture CNN with state-of-the-art traditional methods.....	44
Table 3.7. Performance comparison of proposed texture CNN on the LIDC-IDRI database with state-of-the-art deep learning-based models	45
Table 3.8. The architecture complexity comparisons with state-of-the-art methods.....	46
Table 3.9. Performance comparison of proposed texture CNN with state-of-the-art traditional methods on the LUNGx Challenge database.....	47
Table 3.10. Comparisons with state-of-the-art methods on the MNIST dataset	48
Table 4.1. Nodule classification with SVM when feature the last FC layer.....	59
Table 4.2. Classification results using AdaBoostM2 when features from last FC layer ..	60
Table 4.3. Optimal layers giving the best performance	61
Table 4.4. Nodule classification results by using Optimal Layers with AdaBoostM2.....	63
Table 4.5. Nodule classification results by using optimal layer with SVM	64
Table 4.6. Nodule classification of each class with decision fusion.....	65
Table 4.7. Comparison of proposed decision level fusion with stat-of-the-art techniques	66

LIST OF ABBREVIATIONS

CNN	Convolutional Neural Network
ANN	Artificial Neural Network
DCNN	Deep Convolutional Neural Network
CAD	Computer-Aided Diagnosis
SVM	Support Vector Machine
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
EL	Energy Layer
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
LIDC-IDRI	Lung Image Database Consortium and Image Database Resource Initiative
ResNet	Residual Network
VGG	Visual Geometry Group
TL	Transfer Learning
ROI	Region of Interest
DICOM	Digital Imaging and Communication in Medicine
Hounsfield	Units (HU)
kNN	k-Nearest Neighbor
MC-CNN	Multi-Crop CNN
ELM	Extreme Learning Machine
ReLU	Rectified Linear Unit
FC	Fully Connected
Conv	Convolutional
ROC	Receiver Operating Characteristic Curve
AUC	Area Under ROC Curve
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
HOG	Histogram of Oriented Gradients
MNIST	Modified National Institute of Standards and Technology
PV	Prediction Vector
FV	Feature Vector

Chapter 1.

Introduction

1.1 Introduction

Lung cancer is one of the major threats to human health worldwide and is a leading cause of cancer-related deaths. It was reported in 2018 by the world health organization that approximately 9.6 million people died due to cancer all over the world during the last five years, out of which more than 1.7 million people died due to lung cancer. The death rate is 18% of the total cancer-related deaths. In the meantime, more than two million lung cancer cases were reported during the year 2018. Moreover, as per the report published by the American cancer society, the highest death rate and survival rate is 26% and 18%, respectively [1-3]. The only reason for the low survival rate is diagnosis of cancer in advance stages because the symptoms at early stages are not prominent. Therefore, early diagnostics of lung cancer is of utmost importance to increase the survival rate. Lung cancer is investigated by screening radiograph images, i.e., X-ray, CT or magnetic resonance image (MRI). Usually, manual screening is used to investigate the CT scans slice by slice, which is time-consuming and tedious for radiologists. It requires very high concentration and skill. Furthermore, less experienced radiologists have highly variable detection rates, leading to an increase the false positive detection, especially in subtle cases, when interpretation highly depends on previous experience. Therefore, a conceivable solution to minimize this load on the radiologists is using CAD systems as a second opinion. Some of the studies in the past have shown an improvement in radiologists performance through the use of CAD systems [4]. The CAD systems utilize pattern recognition and machine learning techniques to identify cancerous tissues quicker and

more accurately. These techniques are logistic regression [5], SVM [6], curvelet transform and multi-layer perceptron [7], discrete AdaBoost, and random forest with a heterogeneous feature set composed of geometric [8, 9] and CNNs[9, 10]. To date, different CAD systems such as; neural network-based [11, 12] and traditional models [4, 13] are proposed to detect and classify tumours in medical images.

The basic scheme of a classic CAD system includes; image pre-processing, segmentation, feature extraction, and training of a suitable classifier. Feature extraction is the key step of any CAD system. The extracted features may be handcrafted or deep features. The handcrafted features include texture, shape, size (volume or diameter), the nodule volume growth rate with time, and morphology. Recently, the texture features achieved considerable attention in image classification [14] and for lesion classification in medical images [15, 16]. In the meantime, deep learning techniques such as CNNs have been utilized with promising results for pulmonary nodule classification [2, 11, 12, 17]. The convolutional layers of CNNs are used to extract image features. A typical CNN can explore the image texture features most efficiently, without altering their architecture, because the whole object and its complex features are not very useful in texture analysis compared to the repeated patterns of lower complexity. Therefore, we intended to create a CNN capable of learning texture features and then perform lung nodule classification in CT images as used by Andrearczyk and Paul for texture classification [18]. We proposed a transferable texture CNN for lung nodule classification by incorporating an energy layer (EL) as a texture descriptor. The texture features are learned during the training process by enabling forward and backward propagation.

Although the DCNNs show enhancement in natural image classification in the ImageNet large scale visual recognition challenge (ILSVRC) dataset, their performance relies on extensive labeled data for supervised training. The deficiency of labeled medical images decreases the adaption of DCNNs. To tackle this issue, the transfer learning technique is adopted [19]. In this research, we also utilized the pre-trained model of our proposed texture CNN through transfer learning technique for malignancy classification lung nodule in a small medical image dataset.

Fusion techniques, such as feature fusion and decision fusion, are also proposed to enhance CAD systems. These fusion techniques utilize the geometrical size, densities, texture, shape, and appearance features to improve the classification performance of CAD systems [20-22]. Furthermore, researchers are also utilizing the handcrafted features along with the deep feature for medical diagnostics. Wang et al. proposed a fusion technique to fuse the deep and handcrafted features such as; geometric, intensity, contrast to classify the lung nodules in chest X-ray images [23]. Zhang et al. proposed a DCNN to enhance the performance of the proposed model by incorporating a feature fusion technique [24]. In this thesis, we used decision fusion based on the probability score of different classifiers to enhance the classification of the CAD system. We select the optimal deep features from state-of-the-art DCNNs, and then lung nodule classification is done by training AdaBoostM2 and SVM classifiers.

1.2 Research Objectives

Following are the main objectives of our research:

- To improve the existing CAD system for detection and classification of lungs tumour using the latest pattern recognition and machine learning techniques.
- Development, testing, and evaluation of proposed solution for CAD system which will support radiologists to avoid misdiagnosis because of the fatigue, eyestrain, or lack of experience.
- In particular, the goal is to design a classifier for CAD system which can correctly classify the lungs cancerous tissues from the CT image dataset. Moreover, propose a solution that will improve the overall classification accuracy, sensitivity, specificity.

The goal will be achieved using some advanced techniques like:

- Advanced pre-processing methods for noise removal, image enhancement, reduction of artifacts, and improving image quality in terms of brightness, contrast, and exposure enhancement.
- To analyze the role of texture, intensity and gradient features for improving the classification performance.
- To analyze the validity of transfer learning for classification purposes, features extracted from the pre-trained deep network.
- To apply the fusion techniques for the enhancement of existing CAD systems.
- To investigate the effect of advanced classification methods on improving classification accuracy.
- To improve the classification accuracy by training deep CNN from scratch or fine-tuning pre-trained CNN models.

1.3 Main Contribution

Lung cancer has become a major threat to human health worldwide. Early detection of lung cancer can increase the survival rate of the patient by starting the treatment at the right time. The objective of this research is to improve the diagnosis of lung cancer using enhanced tumour or nodule classification techniques. This dissertation describes the development, testing, and evaluation of various lung nodule malignancy classification techniques that will support radiologists in examining radiography scans, such as; CT, X-ray, or MRI. The main contribution of this work is improvement in the malignancy classification of lung nodules using deep learning, transfer learning, ensemble learning, and medical image fusion, which are discussed below:

We have proposed a transferable texture CNN for efficient lung nodule classification in CT images. The overall proposed architecture consists of only nine layers for automatic feature extraction and malignancy classification of a lung nodule as benign or malignant. An EL is incorporated after the last convolution layer of the proposed texture network. The EL extracts the texture features from the convolutional layer. Incorporating EL reduces the learnable parameters, which further reduces the computational complexity and memory requirements without degrading the classification performance. The texture features are learned through forward and backward propagation during the training process. The proposed texture network is tested successfully on the lung image database consortium and image database resource initiative (LIDC-IDRI) dataset for malignancy classification. Moreover, the effectiveness of the proposed model is also tested by classifying the handwritten digits. Furthermore, the smaller dataset classification problems are also

investigated using the transfer learning (TL) technique on the pre-trained model of proposed texture CNN.

Secondly, we proposed a decision level fusion scheme to enhance the classification performance of the CAD system for lungs nodule classification in CT images. The performance of AdaBoostM2 and SVM algorithms is evaluated using deep features from eight state-of-the-art transferable DCNN architectures which are: VGG-16, VGG-19, Inception-V3, GoogLeNet, ResNet-18, ResNet-50, ResNet-101 and InceptionResNet-V2. After that, the optimal deep features are selected from the DCNNs by identifying the optimal layers to improve the classification efficiency.

The significant contributions of this thesis are as follows:

- We proposed the texture CNN for the lung nodule classification problem and evaluated it for two medical image datasets; LIDC-IDRI and LUNGx challenge. The classification accuracy was achieved up to 96.69% for the LIDC-IDRI dataset.
- The EL is incorporated in the proposed texture CNN, which preserves the texture information, reduces the output vector size, and 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 utilizes LIDC-IDRI as the source task and the LUNGx challenge dataset as the target task. The classification accuracy for the LUNGx challenge dataset was 86.14% without TL, which was further improved to 90.91% using the proposed TL-based model.
- Optimum deep features selection from state-of-the-art DCNN to improve lung nodule classification.

- Classifier selection based on the performance of SVM and AdaBoostM2 on the LUNGx challenge dataset.
- The lung nodule classification was performed on deep features from state-of-the-art DCNN models such as: VGG-16, VGG-19, ResNet-18, ResNet-50, ResNet-101, GoogLeNet, InceptionResNet-V2 and Inception-V3. The decision level fusion technique is proposed using top-performing DCNNs such as GoogLeNet, ResNet-101, and Inception-V3.
- Comprehensive performance evaluation of SVM and AdaBoostM2 classifiers based on deep features on LUNGx challenge and LUNA dataset.

1.4 Thesis Organization

The arrangement of the work presented in this thesis is as follows:

Chapter 1 describes the outline of this thesis, containing an introduction of lung cancer, its detection and classification issues, and the research gaps. After that, the research objectives, main contributions, and hypothesis are defined clearly.

Chapter 2 elaborates on the literature review regarding lung nodule detection and classification. Furthermore, the architectures of different state-of-the-art models are presented, which are contemporary research related to our work and utilized in proposed techniques. The evaluation matrices used for the evaluation of the proposed method are also discussed in this chapter.

Chapter 3 describes the efficient lung nodule classification using transferable texture CNN and transfer learning for small lung datasets. Furthermore, the dataset is also described in this chapter. The proposed transferable texture CNN and transfer learning-based techniques were tested on LIDC-IDRI and LUNGx databases.

Chapter 4 includes the decision fusion technique for lung nodule classification. The optimal deep feature selection and performance evaluation of eight state-of-the-art DCNN is also discussed. The LUNGx challenge dataset is used for the performance evaluation.

Chapter 5 describes the conclusion of this dissertation by highlighting the outcomes of the research work, which justifies our significant contribution. Moreover, it also suggests the guidelines for future research work in this field.

Chapter 2.

Literature Review

Cancer is a broad term for a class of diseases characterized by abnormal cells growth, which also invades the healthy cells in the body and forms a solid tumour. Normally, human cells grow and divide to form new cells as the body needs them. The cells grow old or become damaged, die, and the new cells take their place. Whereas, when cancer develops, this orderly process breaks down. As the cells become more and more abnormal, old or damaged cells survive when they should die, and the new cells form when they are not needed. The growth of extra cells are masses of tissue that form solid tumours. A tumour can be benign or malignant. A benign tumour is non-cancerous, usually localized and does not spread to other parts of the body. Most benign tumours respond well to treatment. However, if left untreated, some benign tumours can grow large and lead to serious diseases because of their size. Cancerous tumours are malignant and can spread into or invade nearby tissues [25]. The size of the tumour varies with time from one millimeter to several centimeters (>8cm) [26]. If the tumour size is less than or equal to 30 mm in diameter, it is normally called a nodule. If the nodule is formed in the lungs, it is called the pulmonary nodule [27]. The lung nodules with dimensions greater than 30 mm are known as lung masses with a maximum probability of benign cancerous [2, 28].

Lung cancer is a leading cause of cancer-related deaths with a very low survival rate. The CAD systems are designed to increase the nodule detection rate, reduce the false positive rate and minimize workload on radiologists by assisting in a second opinion in the screening process of medical images [29]. The first CAD system was developed in the late 1980s to detect lung nodule, which was not appealing due to the unavailability of adequate

computational resources for the implementation of advanced image processing techniques [30]. The performance of computer-based image analysis and decision support systems got a high boost after the enhancement in computational resources. However, detecting and classifying small-sized lesion growth at an early stage is a big challenge for the researchers. These small lesions cannot be detected by radiograph images, and by the time these are detected, it is too late for the patient.

Moreover, the available techniques for detection and classification have many false positives and false negatives because the detection algorithms have high sensitivity for lesions. Hence, some non-lesion structures (for example, blood vessels) are labeled as nodules inevitably. Due to these challenges, CAD-based lung nodule detection and classification has become an active research field. Several efforts have been made to investigate these challenges. Various lung nodule classification techniques are proposed in the literature to improve existing CAD systems. These techniques can be divided into five major groups explained below.

2.1 Traditional Methods for Lung Nodule Classification

The basic scheme of the traditional model for lung nodule classification is shown in Fig. 2.1 [31]. The traditional models consist of the following basic steps: pre-processing, nodule detection using segmentation, feature extraction, feature selection, and classification.

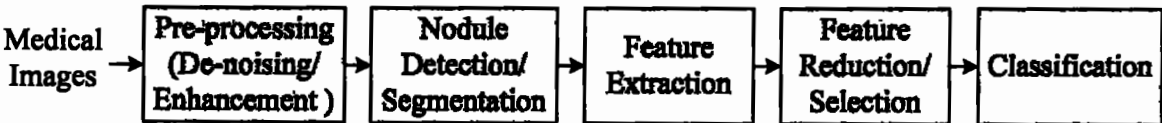


Fig. 2.1. Traditional Model for Lung Nodule Classification System

Researchers are improving the performance of CAD systems by enhancing these steps. The pre-processing includes de-noising and image enhancement techniques. Medical images are often deteriorated by noise due to various sources of interference during image acquisition. The improvement in visual quality and appearance of the images is done by applying image enhancement algorithms. The inappropriate image enhancement application may also increase noise which may suppress minor details and edge sharpness. Therefore, the researchers are doing efforts to use an appropriate pre-processing technique. Schilham et al. used local normalization filtering to remove noise and a global equalization for contrast enhancement in their proposed CAD system [32]. Emre Dandil introduced the image enhancement of medical images to prevent false positive results. In this work, unnecessary noise was removed first by applying the median filter. The unnecessary grains and contrast differences were removed by applying the histogram equalization technique. Furthermore, a Laplacian filter was used to sharpen the contours [33]. In this work, we used adaptive histogram equalization techniques as a pre-processing step to enhance the lung nodules images.

The detection of a nodule is a very important step before their segmentation. As lung nodules have helical and circular structures, circular object detection algorithms like circular Hough transform are being used for lung nodule detection [33]. Moreover, some other techniques like region of interest (ROI) extraction using thresholding can be used to detect the lung nodule. For example, Masood et al. adopted the thresholding technique to detect lung nodules and extracted the ROI around the nodule [34]. The segmentation is the next step after nodule detection. Researchers are utilizing different techniques for the segmentation of nodules. Masood et al. adopted the ROI-based segmentation [34], whereas,

in [33], self-organizing maps [35] were proposed for the lung nodule segmentation. Han et al. presented the vector quantization techniques for accurate segmentation of lung nodules [36]. In this work, we extracted the nodule patches by utilizing the ROI of the required masses.

The salient feature extraction and selection is critical for any classification system. The extracted features consist of shape, size (diameter or volume), morphology, texture, and the volume growth rate of the nodule. Researchers utilize these features to improve the classification task. For example, the shape feature analysis is used by El-Baz et al. for lung nodule diagnostic. The images were segmented with the active contour method, and the texture features of lung nodules were extracted using rubber band straightening transform [37]. The texture features have attained great attention in image classification [14] and lesion classification for medical images [15, 16, 38, 39]. Descriptors of gray-level co-occurrence matrix, discrete wavelet transform, [40], local binary pattern (LBP) [41], higher-order spectra, and histogram of oriented gradients (HOG) [42] are widely used for texture feature representation in medical image processing [43]. In [33], one hundred twenty-three salient features were extracted, and principal component analysis [44] is applied for feature reduction. These features include mean, standard deviation, variance, entropy, skewness, kurtosis, histogram, sharpness, convexity, circularity, texture, and statistical features [45], [46], [47], [48]. In addition, Narayanan et al. [49], [50] also proposed an optimized feature selection-based clustering technique for lung nodules using intensity, gradient, and geometric features.

Selecting a suitable classifier is another important step of a CAD system. The most commonly used classifiers are; linear logistic regression, random forest, k-nearest neighbor

(kNN), extreme learning machine (ELM), LDA, AdaBoost, and SVM [51-53]. In this thesis, we suggested a suitable classifier among the LDA, AdaBoostM2, and SVM based on performance evaluation results.

2.2 ANN and CNN Based Nodule Classification Techniques

ANN is a computational system designed with inspiration from the biological network of the animal brain. It consists of connected nodes, which are the model of neurons in the brain. Based on their learning capability, various lung nodule detection and classification techniques are proposed, such as; YSP Chiou et al. proposed an ANN-based system for lung nodule detection and classification in late 1993 [54]. Penedo et al. proposed two-phase ANN-based lung nodule detection and classification techniques. During the first phase, one ANN performs the detection of a nodule in terms of feature space, whereas, in the next phase, the other ANN performs a classification task [55]. Ashwin et al. proposed an ANN-based technique for efficient and reliable lung nodule diagnosis. The ANN was trained and tested after pre-processing of CT and radio graph images [56]. Many researchers utilized the ANN as a classifier for the classification of lung nodules [57-59].

An ANN can receive only a 1D feature map at its input, whereas the CNNs can take a 2D image or 3D voxel input for the object classification. The CNNs are the deep learning algorithms that take the input image and learn the weights and biases for various objects to detect and classify the objects using ANN at the last layer. The main idea of CNN for pattern recognition is named Neocognitron, which was introduced by Fukushima in 1980. The Neocognitron included two basic layers, which are the convolutional and down-sampling layers. The convolutional layer includes filters, whereas the down-sampling layer computes average activation and helps to classify the object [60] correctly. After that, Yann

LeCun et al. used back-propagation in CNN for learning the coefficients of convolutional filters directly from the input image [61]. The first traditional gradient-based multi-layered CNN, named LeNet-5, was proposed by Yann LeCun et al. in 1998 for handwritten digits classification [62]. The CNNs got a great breakthrough after applying GPUs by Alex Krizhevsky et al. for object recognition in the ImageNet dataset. Their proposed AlexNet was implemented on two GTX 580 GPUs which won the ILSVRC 2012 with a top-5 error rate of 15.3% [63, 64]. After that, various DCNN architectures, such as; VGG-Net [65], GoogLeNet [66], ResNet [67] and InceptionNet [68] achieved promising results on various image classification benchmarks like MS COCO [69], ILSVRC 2012–2017 [64, 70], and CIFAR-10 [71].

The continuous achievements of DCNNs in image classification inspired the researchers to utilize for medical diagnostic [72-75]. Researchers proposed DCNNs based novel techniques for malignancy detection in lungs, such as; R Majidpourkhoei proposed a LeNet based architecture that automatically learns the image features [76]. Shen et al. proposed a multi-crop CNN (MC-CNN) for nodule classification in which salient information of nodules was acquired by cropping specific regions of convolutional feature maps and then max-pooled at multiple stages [11]. Zhang et al. proposed an automated 3D DCNN for detecting and classifying lung nodules [77]. Zhu et al. proposed a fully automated 3D DCNN for lung nodule detection and classification [78]. Recently, Wang et al. proposed a novel DCNN architecture to classify pulmonary nodules in the LUNGx challenge dataset. The classification accuracy, and AUC score were 90.38% and 94.48%, respectively [79]. Cao et al. proposed a dual-stage DCNN for lung nodule detection. The

first stage is for lung nodule detection, whereas the second stage is used for false positive reduction [80].

Similarly, Wei Li et al. proposed a DCNN for the lung nodule classification. They utilized two parallel convolutional layers to extract the feature maps, merged at the first convolutional layer [67]. These DCNN are computationally complex architectures and take more training time due to a large number of learnable parameters. In this thesis, we proposed a transferable texture CNN with a lesser number of parameters that utilized the texture energy for efficient lung nodule classification.

2.3 Transfer Learning-Based Nodule Classification Techniques

The performance of DCNN essentially relies on extensive labeled data for supervised training. The deficiency of labeled medical image dataset slows down the adaption of CNN during the training and testing process. Meanwhile, the manual annotation and labeling of every image to construct a massive medical training database is painful and prohibitive. Second, the training of deep CNN requires large memory and computational resources. The lack of such resources increases the training time of DCNN. Third, the training of DCNN is often complicated due to convergence and overfitting issues which require continuous modification in the learning parameters or architecture of the network to assure that all the layers are learning with approximately equal speed. To tackle these issues, researchers are utilizing different learning techniques, such as; transfer learning and fine-tuning [81-83]. The objective of TL is to transfer knowledge from the source to target domains [84]. As M. Oquab et al. performed training on the source task (ImageNet database), then transferred the pre-trained parameters of CNN to the target task (PASCAL VOC dataset) for object classification [85]. In this thesis, we employed the same strategy

using our pre-trained texture CNN model to classify the pulmonary nodule in a small LUNGx challenge database.

2.4 Fusion Based Techniques for Nodule Classification

The fusion techniques, such as; feature fusion [86], multi-model image fusion [87, 88], and decision fusion, are used to improve the image quality and provide promising results, which motivate the researchers to utilize these techniques in medical diagnostics. Literature shows that various fusion techniques are proposed for medical images to improve the classification accuracy of pulmonary nodules. For example, multi-modal image fusion is employed in medical image processing, in which the medical images from two different modalities, such as; PET – CT, MRI – CT, and MRI – PET, are fused to improve the medical diagnosis [89-92]. The feature fusion techniques are also used to enhance the performance of medical diagnostics, such as; Khan et al. proposed a lung cancer classification technique which performed the serial canonical correlated-based fusion of texture, point, and geometric features to classify pulmonary cancer in the Kaggle challenge database [93]. Wang et al. proposed a deep feature fusion scheme to classify the pulmonary nodules in chest radiograph images by fusing the deep features from pre-trained AlexNet and handcraft features, like; geometric features, contrast, intensity, along with the first order and second-order filter features [23]. Sridar et al. proposed a decision-based fusion technique for fetal ultrasound images plane classification and achieved an accuracy score of 92.00% [94]. Xie et al. proposed a decision level fusion technique for automatic pulmonary nodule classification on the LIDC-IDRI dataset. They trained AdaBoosted back-propagation ANN on shape, texture, and deep features separately and then performed the decision fusion on three outcomes [95]. In this thesis, we used the decision fusion

technique for lung nodule classification in CT images. We showed that the proposed fusion techniques perform well as compared to state-of-the-art techniques.

2.5 State-of-the-art DCNNs

The CNNs got a great breakthrough after applying GPUs in object recognition and classification. Various DCNN architectures (AlexNet, VGG-Net, GoogLeNet, Residual Networks (ResNets) [67] , Inception [96], Xception [97] and Dense Networks [98, 99]) are proposed to improve the classification performance. These architectures have shown promising results in segmentation, recognition, and classification, which inspired us to utilize the state-of-the-art DCNNs. In our research work, we utilized AlexNet, VGG16, VGG19, GoogLeNet, ResNet18, ResNet50, ResNet101, Inception-V3, and InceptionResNet-V2 which are briefly discussed below.

2.5.1 VGG-Net

The VGG-Net architecture was proposed by the visual geometry group (VGG) team for ILSVRC 2014 and won this challenge. It is designed by increasing the depth of the available CNN model up to sixteen for VGG-16 and nineteen for VGG-19, as shown in Fig. 2.2

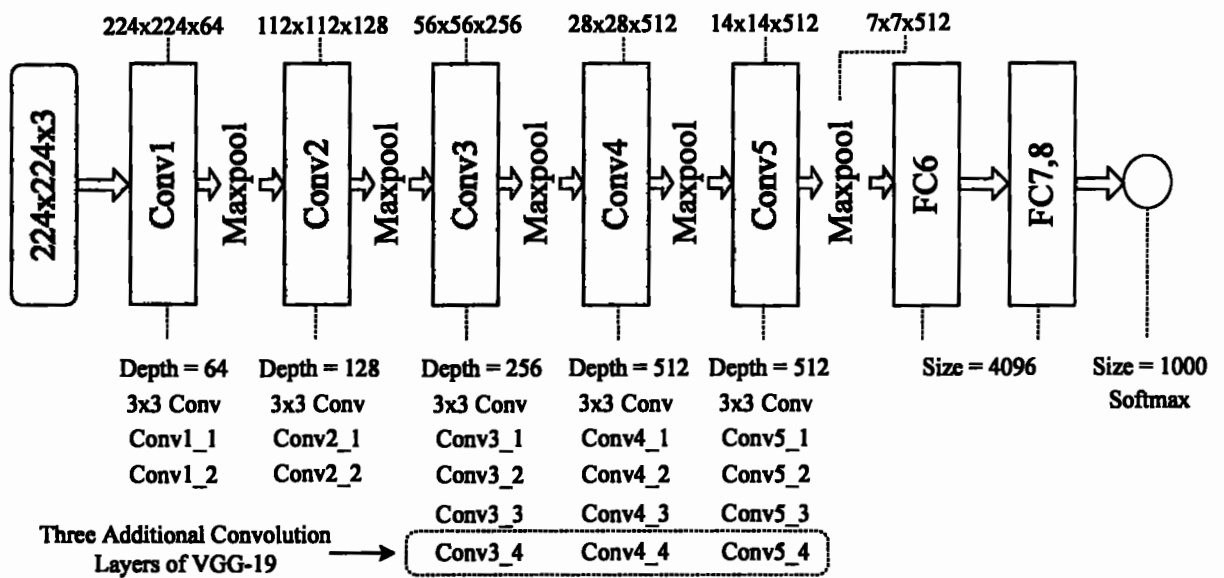


Fig. 2.2. Architecture of VGG-16 and VGG-19

VGG-Net investigates the performance of CNN by increasing its depth. The architectural comparison of VGG-16 and VGG-19 is described in Table 2.1.

Table 2.1. Layer-wise architecture details of VGG-16 and VGG-19

Layer Name	Input Size	Output Size	VGG-16	VGG-19
Conv1	224 × 224	112 × 112	3 × 3, 64 Two Layers, Maxpool	3 × 3, 64 Two Layers, Maxpool
Conv2	112 × 112	56 × 56	3 × 3, 128 Two Layers, Maxpool	3 × 3, 128 Two Layers, Maxpool
Conv3	56 × 56	28 × 28	3 × 3, 256 Three Layers, Maxpool	3 × 3, 256 Four Layers, Maxpool
Conv4	28 × 28	14 × 14	3 × 3, 512 Three Layers, Maxpool	3 × 3, 512 Four Layers, Maxpool
Conv5	14 × 14	7 × 7	3 × 3, 512 Three Layers, Maxpool	3 × 3, 512 Four Layers, Maxpool
FC6	7 × 7 × 512	4096	First Fully Connected Layer	First Fully Connected Layer
FC7	4096	4096	2 nd Fully Connected Layer	2 nd Fully Connected Layer
FC8	4096	4096	Last Fully Connected Layer	Last Fully Connected Layer
Softmax	1000	1000	Classification layer (Softmax)	Classification layer (Softmax)

The architecture of VGG-19 has 144 million parameters, whereas VGG-16 has 138 million parameters. The VGG-16 has 13 convolutional layers, 5 max-pooling layers (2 × 2), and two fully connected layers. The output is a linear layer with a softmax activation

function. All the convolution layers have a ReLU activation function, and dropout regularization is used in fully connected layers. The architecture of VGG-19 is like the VGG-16, except for three additional convolution layers, which are incorporated in convolution blocks named Conv3, Conv4, and Conv5, as shown in Fig. 2.2 and Table 2.1. In our research work, we extracted the deep features from the last fully connected layers of both networks. The output from *FC6* – *FC8* layers of both variants can be computed using the following equations.

$$Y_6 = \sigma(W_6Y_5 + B_6) \quad (2.1)$$

$$Y_7 = \sigma(W_7Y_6 + B_7) \quad (2.2)$$

$$Y_8 = \sigma(W_8Y_7 + B_8) \quad (2.3)$$

where Y_k denotes the output of the k^{th} layer, W_k and B_k are the weights of the k^{th} layer. The output feature map size of *FC6* and *FC7* is 4,096. The deep features from such layers can be used for the classification task. In this work, the deep features are extracted from *FC6* during forward propagation of the input through the pre-trained VGG-Net.

2.5.2 GoogLeNet

The state-of-the-art CNN architecture named GoogLeNet was proposed for the ILSVRC 2014 challenge. It achieved a top-5 classification error of 5.5% and placed at the first position of this challenge. The GoogLeNet is also known as inception because it introduces a new module named inception block. This block concatenates kernels of various sizes into one kernel. The structure of the inception block is shown in Fig. 2.3.

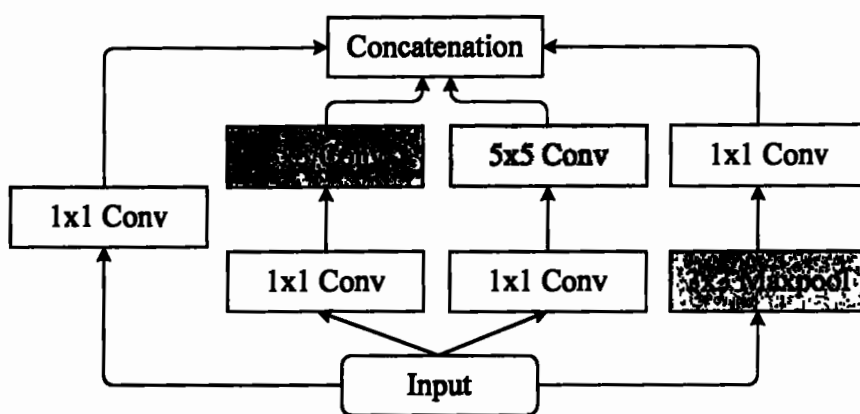


Fig. 2.3. Structure of inception block

It consists of four concurrent branches. The first three branches use convolutional layers with kernel sizes of 1×1 , 3×3 , and 5×5 . The complexity of this model is reduced by convolving the two mid branches with a window size of 1×1 to the input channels. The fourth branch is a 1×1 , convolutional layer, which drives the 3×3 max-pooling layer. The relevant padding is used by all four branches to keep the similar height and width of the input and output. The output of each branch is concatenated to develop the final output of the inception block. The basic architecture of GoogLeNet is shown in Fig. 2.4, which has approximately 6.8 million parameters.

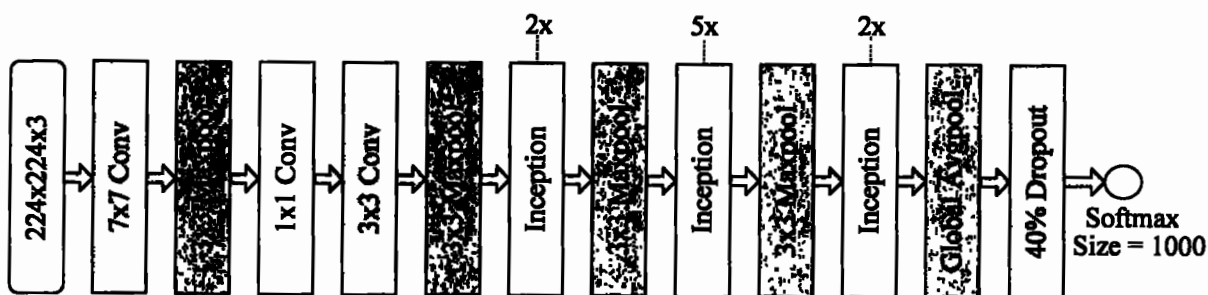


Fig. 2.4. Basic architecture of GoogLeNet

The entire architecture has nine inception blocks (each block has six convolutional layers), two convolutional layers of 7×7 and 3×3 , one convolutional layer of size 1×1 (for dimension reduction), four max-pooling layers, two normalization layers, one average

pooling, and one FC layer. All convolutional layers utilize the ReLU activation function, and dropout regularization is applied in the FC layer. The Softmax activation function is used in the output layer.

2.5.3 Inception-V3

Inception-V3 is the third version of the famous GoogLeNet architecture, which was trained with one million training images of one thousand classes of the ImageNet dataset. In Inception-V3, an inception module concatenates multiple sized convolutional kernels in one kernel. This design reduces the number of learnable parameters, which reduces the complexity of the network. The block diagram of inception-V3 is shown in Fig. 2.5.

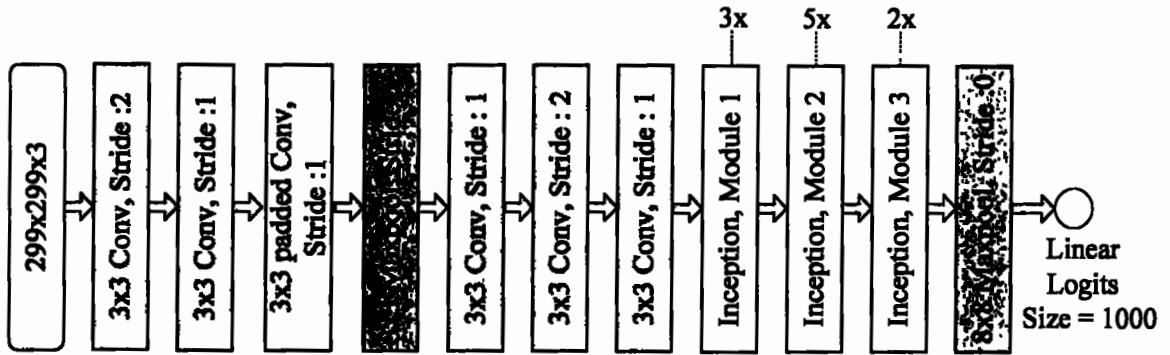


Fig. 2.5. Basic architecture of Inception-V3

2.5.4 ResNet

The basic idea of ResNet is that every layer of the architecture learns from residual functions with reference to its input layer. In this way, the architecture is easily optimized and gains significant accuracy. The ResNet architecture was proposed with different variants in the ILSVRC-2015 competition and placed first. In this work, we utilized pre-trained models of three variants which are ResNet-18, ResNet-50, and ResNet-101. The basic architectures of all the variants are given in Table 2.2.

Table 2.2. Architectures of ResNet-18, ResNet-50 and ResNet-101

Layer Name	Output Size	ResNet-18	ResNet-50	ResNet-101
Conv1	112×112	$7 \times 7, 64, \text{Stride } 2,$ $3 \times 3 \text{ Maxpool, Stride } 2$	$7 \times 7, 64, \text{Stride } 2,$ $3 \times 3 \text{ Maxpool, Stride } 2$	$7 \times 7, 64, \text{Stride } 2,$ $3 \times 3 \text{ Maxpool, Stride } 2$
Conv2_x	56×56	$\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$
Conv3_x	28×28	$\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 4$
Conv4_x	14×14	$\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1,024 \end{bmatrix} \times 23$
Conv5_x	7×7	$\begin{bmatrix} 3 \times 3, & 512 \\ 3 \times 3, & 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2,048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2,048 \end{bmatrix} \times 3$
Pool	$1 \times 1 \times 512$	Average Pool	Average Pool	Average Pool
FC	$512 \times 1,000$	Fully Connected Layer	Fully Connected Layer	Fully Connected Layer
Softmax	1,000	Classification layer	Classification layer	Classification layer

The input size is 224×224 , whereas the first convolution layer ($7 \times 7, 64$ Stride 2) and the last three layers (pooling, FC, and softmax) are fixed for all three discussed architectures. The depth of the network is varied by increasing inner convolution layers.

2.5.5 InceptionResNet-V2

Inception-ResNet-V2 is the variant of Inception-V3 and it also integrates some ideas from ResNet. In InceptionResNet-V2 with the batch normalization is utilized only on the top of the traditional layers. The residual modules are engaged in such a manner that the number of inception blocks and the depth of the network are increased. The basic architecture of InceptionResNet-V2 is shown in Fig. 2.6.

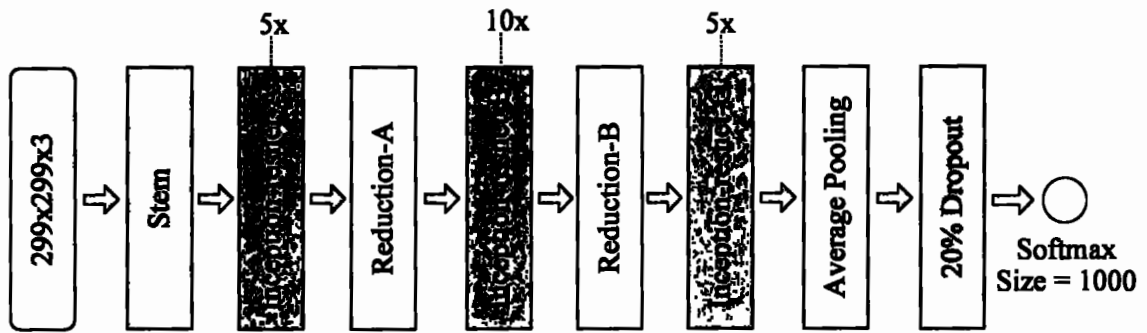


Fig. 2.6. Architecture of InceptionResNet-V2.

The basic architecture consists of one stem block (which has six convolution blocks and one max pool layers), three different types of inception blocks, which are Inception-resnet-A, Inception-resnet-B, and Inception-resnet-C (Inception-resnet-A has five inception modules, each has seven convolution blocks, Inception-resnet-B has ten inception modules, each has five convolution block, Inception-resnet-C has five inception modules, each has four convolution block), two reduction blocks with different convolutional layers, one average pool, and one FC layer. The Softmax function is used in the output layer.

2.6 Performance Evaluation Metrics

In this thesis, we considered Class 0 as benign and class 1 as malignant for binary classification problems. The performance of the binary classification model is described by a confusion matrix shown in Table 2.3.

Table 2.3. Confusion matrix for two classes

Actual Class	Predicted Class	
	Positive (Malignant)	Negative (Benign)
Class 1: Actual Positive (Malignant)	TP	FN
Class 0: Actual Negative (Benign)	FP	TN

For the above confusion matrix, true positive (TP) states that the assessment of experts and classifier prediction is positive, whereas true negative (TN) states that the assessment of

experts and classifier prediction is negative. Similarly, false positive (FP) states that the assessment of experts is negative and classifier predicted as positive, whereas, false negative (FN) states that the assessment of experts is positive and classifier prediction is negative. The quantitative performance of the proposed method is determined by different evaluation matrices, like; Accuracy, Gmean, Precision, Recall, Specificity, Error Rate, receiver operating characteristic curve (ROC), and area under ROC (AUC).

Accuracy: Accuracy is the statistical measure to evaluate the classifier model. It elaborates how well a classifier predicts. It is calculated using the following expression:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (2.4)$$

Precision: It is the positive predictive value and is mathematically defined as:

$$Precision = \frac{(TP)}{(TP + FP)} \quad (2.5)$$

Specificity: Specificity is the measure of the true negative rate. It is the percentage of actual negative cases which are accurately predicted by the classifier and can be calculated using the following equation:

$$Specificity = \frac{(TN)}{(TN + FP)} \quad (2.6)$$

Sensitivity: Sensitivity or recall/true positive rate is the measure of the total percentage of actual positive cases correctly identified by the algorithm. Therefore, it measures the benign or malignant nodules correctly identified by the algorithm. It is calculated using the following equation:

$$Sensitivity = \frac{(TP)}{(TP + FN)} \quad (2.7)$$

Error rate: The error rate of a classifier is defined by the following expression:

$$\text{Error rate} = \frac{(FP + FN)}{(TP + TN + FP + FN)} \quad (2.8)$$

2.7 Summary

In this chapter, we have discussed the importance of the CAD system for lung cancer detection and literature review of different lung nodules classification techniques used in medical diagnostics. Furthermore, different state-of-the-art DCNNs utilized in this thesis are also discussed along with their architectures. Finally, different performance evaluation metrics are discussed which are used to evaluate the quantitative performance of proposed techniques.

In the next Chapter, we comprehensively have discussed the proposed transferable texture CNN and its performance evaluation.

Chapter 3. Transferable Texture CNN for Efficient Lung Nodule Classification

The efficient lung nodule classification using transferable texture CNN and utilization of transfer learning for small lung datasets are discussed in this chapter.

3.1 Introduction

In recent years, texture features have attained great attention in image classification and the lesion classification of medical diagnostics. Han et al. utilized three texture features, such as Haralick, local binary patterns, and Gabor features for lung nodule classification in CT images with the AUC score of 92.70% [15]. Firmino et al. proposed lung cancer detection and classification using HOG and watershed techniques. They utilized SVM and rule-based classifiers for false positive reduction [13]. Tizita and Alhayat utilized the geometric and histogram features of the lung nodule images with linear and non-linear classifiers for malignancy detection [51]. RW de Sousa Costa et al. used texture descriptor in terms of taxonomic diversity index and mean phylogenetic distance, which characterizes the basic structure of the lung nodule [100]. Similarly, Guohui Wei et al. presented lung nodule detection using texture features [101]. Emre Dandil computed the combined shape, intensity, energy, and texture features of the lung nodule. The principal component analysis is used for feature reduction before final classification through a probabilistic neural network [33]. Ahmed Shaffie et al. used HOG and higher-order Markov Gibbs, a random field model, to describe the texture of lung nodules, and the classification is done using a stacked auto-encoder [102]. Recently, deep learning

techniques, especially CNNs, have been used with promising results for lung nodule classification [11, 17, 103].

The CNNs utilize convolutional layers to extract the features. The complex features are extracted by the last convolutional layer, which is utilized in fully connected layers to extract the complete shape information. The gradient features are extracted by the first convolutional layers, whereas the inner pooling and convolutional layers extract features with considerable complexity. For the texture analysis, the entire object and its complex features are not much useful 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 efficiently explore the texture properties without changing the architecture [18].

We intended to build a CNN capable of learning the texture features and then classify the lung nodules in CT images, as used in [18], for texture image classification. Therefore, we introduced a texture descriptor named EL right after the convolutional layer. We enabled the forward and backward propagation to learn the texture features during the training process. Moreover, the TL technique is also used to investigate the issues of the small labeled medical image dataset using our pre-trained model. Therefore, the proposed technique is named transferable texture CNN. We also showed that our texture CNN achieved better classification performance on lung CT images with fewer learnable parameters and neurons. The proposed architecture is trained and tested using six-fold cross-validation for binary classification of lung nodule malignancy. Furthermore, the proposed model is also tested on a modified national institute of standards and technology (MNIST) dataset [62, 104]

3.2 Datasets and Methodology

In this section, we discussed the datasets, patch generation, and image augmentation technique, which is used to increase the size of small biomedical images dataset to meet the training requirements of the proposed CNN model. Afterward, the architecture of our proposed transferable texture CNN and TL methodology are discussed.

3.2.1 Datasets

The performance evaluation of both techniques is done using publicly available LIDCIDRI, LUNGx challenge, and MNIST databases.

3.2.1.1 LIDC-IDRI Database

The LIDC-IDRI is a publicly available database that contains 244,527 thoracic CT scan images of 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 four experienced thoracic radiologists marked. They graded each feature from 1 to 5 annotations [105]. We utilized available XML files and an annotation list [106, 107] 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, given in Table 3.1.

Table 3.1. Degree of malignancy in LIDC-IDRI database

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

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).

3.2.1.2 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 calibration set has ten scans (five females, five males). Five of ten calibration scans contain one confirmed benign nodule, and the other five contain one pathology-confirmed malignant nodule. Whereas the test set has 60 scans with a total of 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

The LUNGx challenge database consists of a single transaxial series with full thoracic coverage for each case. All scans have been obtained on Philips Brilliance scanners with a “D” (over-enhancing), and each scan has a 1 mm slice thickness. The LUNGx challenge has 22,489 CT images in digital imaging and communication in medicine (DICOM) format. Each image file has a Unique Identifier (UID) 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).

3.2.1.3 MNIST dataset

The MNIST dataset is a handwritten digits dataset constructed by the National Institute of standards and technology (NIST); hence, known as modified NIST or MNIST. It consists of 10,000 labeled tests and 60,000 labeled training images. The vectors size for

each black and white image is 784 [62, 108]. Each consist of 28×28 pixels. Therefore, we first convert it into a fixed-size image before training and testing the proposed methodology.

3.2.1.4 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 using different image processing operations, such as; translation, resize, random rotation, flips, and shear, etc. In this work, the size of the dataset $D = \{X_i: 1 \leq i \leq N\}$ is increased 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.

3.2.1.5 Patch Generation and Data Enhancement

The LIDC-IDRI and LUNGx challenge database comprises a heterogeneous set of scans acquired using various reconstruction and acquisition parameters. In the CT images, the air is available in the lungs with a mean intensity of -1000 Hounsfield units (HU) approximately, and most of the tissue have intensity in the range of -910 HU to -500 HU. Moreover, the blood, bone, and chest wall are much denser (above -500 HU) [109]. Therefore, to normalize the pixels, all CT images are first converted to HU scales 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).

In both databases, all the slices are available in the DICOM format, having a size of 512×512 at a pixel depth of 16 bits. The image patches are created in two phases after HU transformation. In the first phase, ROI around the nodule is extracted by acquiring the central coordinates (x, y, z) and slicing the 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 using voxel coordinates 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 Fig. 3.1.

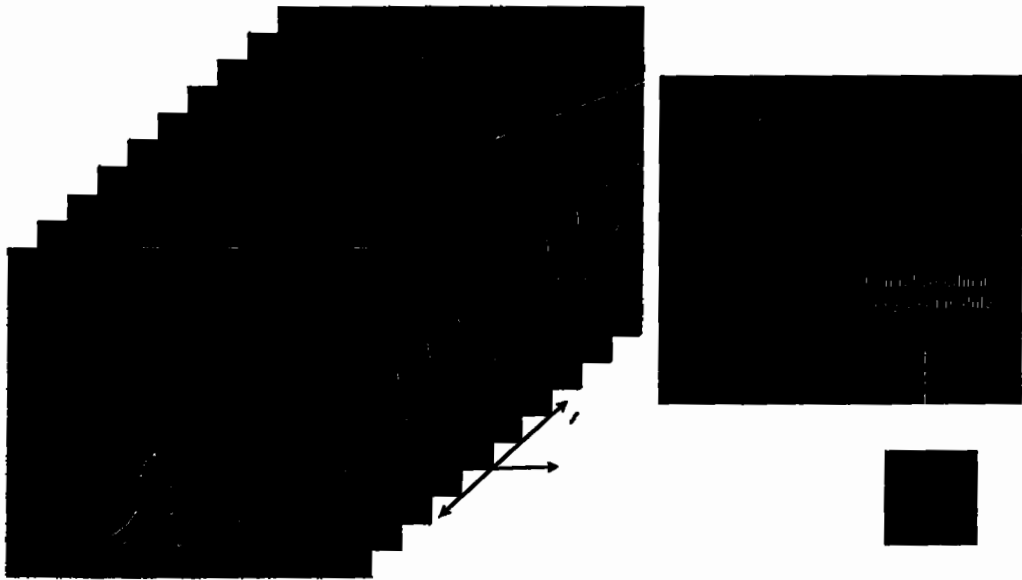


Fig. 3.1. Patch Extraction Process

In this way, a total of 19,388 patches of size 64×64 were extracted from 1,010 cases of the 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.

3.2.2 Architecture of Transferable Texture CNN

Keeping in view the following three essential features of the image, deep CNN has been developed. First, some discriminational patterns have a very small size than the actual image, but if their size equals the size of the convolution filter mask, then the said patterns can be found by the convolution filter. Second, some shapes 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 critical for the max-pooling layer and do not alter the shape of the input image. These pixels are utilized in biomedical image classification. Fig. 3.2 shows the overall architecture of the proposed texture CNN for lung nodule classification.

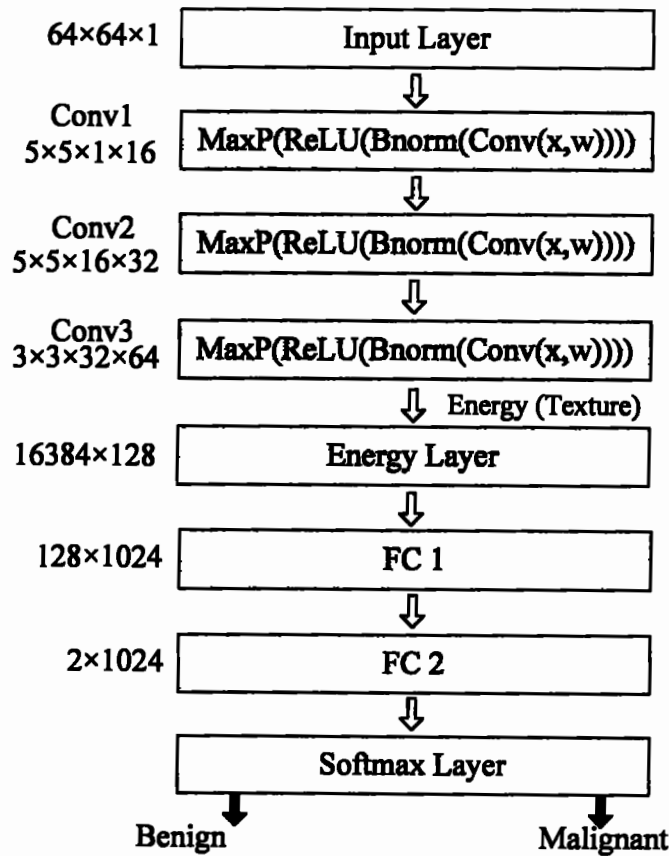


Fig. 3.2. Architecture of Transferable Texture CNN

The proposed CNN has two convolutional layers, followed by the normalization and the pooling layers. Whereas the third convolutional layer drives the EL. Finally, the Softmax is utilized with the fully connected layer to classify lung nodules. Moreover, the layer-wise dimensional details are given in Table 3.2, including network layer details like kernel, stride, and padding for each layer.

Table 3.2. Layer-wise architecture details of proposed texture CNN

Layer	Input size	Kernel	Stride	Padding	Output size	Learnable parameters
Conv1	$64 \times 64 \times 1$	5×5	$[1 \ 1]$	$[1 \ 1 \ 1 \ 1]$	$62 \times 62 \times 16$	Bias: $1 \times 1 \times 16$, Weights: $5 \times 5 \times 1 \times 16$
Pool1	$62 \times 62 \times 16$	2×2	$[2 \ 2]$	$[1 \ 1 \ 1 \ 1]$	$31 \times 31 \times 16$	Weights:0
Conv2	$31 \times 31 \times 16$	5×5	$[1 \ 1]$	$[1 \ 1 \ 1 \ 1]$	$30 \times 30 \times 32$	Bias: $1 \times 1 \times 32$, Weights: $5 \times 5 \times 16 \times 32$
Pool2	$30 \times 30 \times 32$	2×2	$[2 \ 2]$	$[1 \ 1 \ 1 \ 1]$	$16 \times 16 \times 32$	Weights:0
Conv3	$16 \times 16 \times 32$	3×3	$[1 \ 1]$	$[1 \ 1 \ 1 \ 1]$	$16 \times 16 \times 64$	Bias: $1 \times 1 \times 64$, Weights: $3 \times 3 \times 32 \times 64$
EL	$16 \times 16 \times 64$	NA	NA	NA	128×1	Bias: 128×1 , Weights: $128 \times 163,84$
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 input and output dimensions of each layer are also mentioned in Table 3.2. We used the following mathematical relation to compute the output size of any convolutional layer:

$$OutputSize = \frac{S_i - S_o + 2P}{\zeta + 1} \quad (3.1)$$

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

3.2.2.1 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 an intermediate layer to extract the texture features. It has 64 output channels and a 3×3 kernel size. The number of learnable parameters for convolutional layers is only 31,744, computed using the following mathematical expression.

$$\theta_C = (S_k \times \zeta + 1) \times N_C \quad (3.2)$$

$$\theta_C = S_k \times \zeta \times N_C + N_C \quad (3.3)$$

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

Each convolutional layer computes the output of neurons 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 an output in a volume of $32 \times 32 \times 16$ with 16 kernels. Let χ be an input feature map and be the weights, then the output of the neurons at first convolutional layer is given by equation 3.4.

$$Y^k = f(\chi^k * \omega^k + b^k) \quad (3.4)$$

where Y^k is the output feature map of the convolutional layer for k th 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 to classify 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 an 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 equation 3.5.

$$E(\chi, \theta) = \sigma(\sum_{i=1}^n \omega_i^T \chi_i + b) \quad (3.5)$$

where $E(\chi, \theta)$ is the output of EL, n is the number of 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 the previous layer and also learns during forward and backward propagation. Therefore, the EL enhances the overall learning capability of the network in addition to the reduction of vector size for the 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 using equation 3.6.

$$\theta_{EL} = \delta^n \times \delta^{n-1} \quad (3.6)$$

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

3.2.2.2 Batch Normalization and Activation Function

The batch normalization is used between the convolutional and ReLU layers to speed up the training process and minimize the sensitivity of network initialization [110, 111]. 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 using the following equations.

$$\mu_B = \frac{1}{m} \sum_i^m \chi_i \quad (3.7)$$

$$\sigma_B = \frac{1}{m} \sum_i^m (\chi_i - \mu_B)^2 \quad (3.8)$$

where μ_B and σ_B are the mean and variance of mini-batch, whereas, m is the mini-batch size of i input feature element. The value of m is selected as 64. After computing μ_B and σ_B , the batch normalization is computed using equation 3.9.

$$Y_i = \frac{(\chi_i - \mu)}{\sqrt{\sigma^2 + \epsilon}} \gamma + b \quad (3.9)$$

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 [112]. The ReLU layer is used as a piecewise function, such as $\max(0; x)$ thresholding at zero. Equation 3.10 is used as an activation function, whereas, expression 3.11 represents the output of the ReLU layer.

$$Y_{i,j,k} = \max\{0, \chi_{i,j,k}\} \quad (3.10)$$

$$Y_{ReLU} = \text{ReLU} \left(B_{norm}(\text{Conv}(\chi, \omega)) \right) \quad (3.11)$$

In equation 3.9, $Y_{i,j,k}$ 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.

3.2.2.3 Pooling Layer

A pooling layer reduces the feature map size and ultimately reduces the computations and weights, leading to overfitting 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 3.12.

$$Y_{POOL} = \text{MaxP} \left(\text{ReLU} \left(B_{norm}(\text{Conv}(\chi, \omega)) \right) \right) \quad (3.12)$$

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

$$Y^k = \max(0, \sum_{k=1}^p \chi^{k-1} \omega^k) \quad (3.13)$$

where Y^k is the output feature map for k^{th} channel and χ is the input feature map. Whereas ω is the kernel for the maxpooling layer, and p represents the pooling size. Two maxpooling 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 equation 3.13.

3.2.2.4 Dropout Regularization

We used dropout regularization to prevent the overfitting of training data, as it eliminates the random subset of parameters iteratively during the weight update process. As the fully connected layer has the maximum number of parameters over the entire network, it goes under the influence of overfitting 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 rates.

3.2.2.5 Softmax Classifier and Loss Function

The softmax is used as a classifier that 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 given in equation 3.14, also computes the compatibility

of the available set of parameters, analogous to the ground truth labels of the training dataset.

$$\Gamma_L = \psi_{y_i} + \text{Log} \sum_j \exp(\psi_j) \quad (3.14)$$

where Γ_L is the total loss, and ψ_j is the j^{th} 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 using the following softmax function:

$$Y_i = \frac{\exp(\psi_{y_i})}{\sum_j \exp(\psi_j)} \quad (3.15)$$

3.2.2.6 Back-propagation Algorithm

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

$$L = -\frac{1}{|\chi|} \sum_{i=1}^{|\chi|} \ln(P(y^i, |\chi|^i)) \quad (3.16)$$

where L is the cost function which is calculated iteratively. The network parameters (θ) are updated with stochastic gradient descent with momentum technique given in equation 3.17.

$$\theta(t+1) = \theta(t) - \left(\lambda \frac{\partial L}{\partial \theta} - \alpha \theta(t) + \beta \lambda \theta(t) \right) \quad (3.17)$$

where α represents the momentum rate, whereas, λ denotes the learning rate, which accelerates the learning procedure and leads to coping 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 gradient descent to tune the network parameters and train a CNN.

3.2.3 Deep Feature Transfer Technique for Malignancy Classification

The performance of different machine learning techniques essentially relies on extensive labeled data for supervised training. Whereas, deficiency of the labeled medical database for training and testing reduces the adaption of CNN. Simultaneously, manually annotating and labeling every data item to construct an immense training database from miscellaneous domains is painful and prohibitive, particularly for the medical image databases that also have their 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 labeled data of various domains. Therefore, the idea of transferring features is utilized to study a discriminative and robust model in the presence of variable tests and training distributions known as TL [84]. 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 [85]. The same strategy is employed in this work for lung nodule malignancy classification using our pre-trained CNN model. The platform is introduced between deep learning and TL for lung nodule classification. Fig. 3.3 shows the proposed TL methodology.

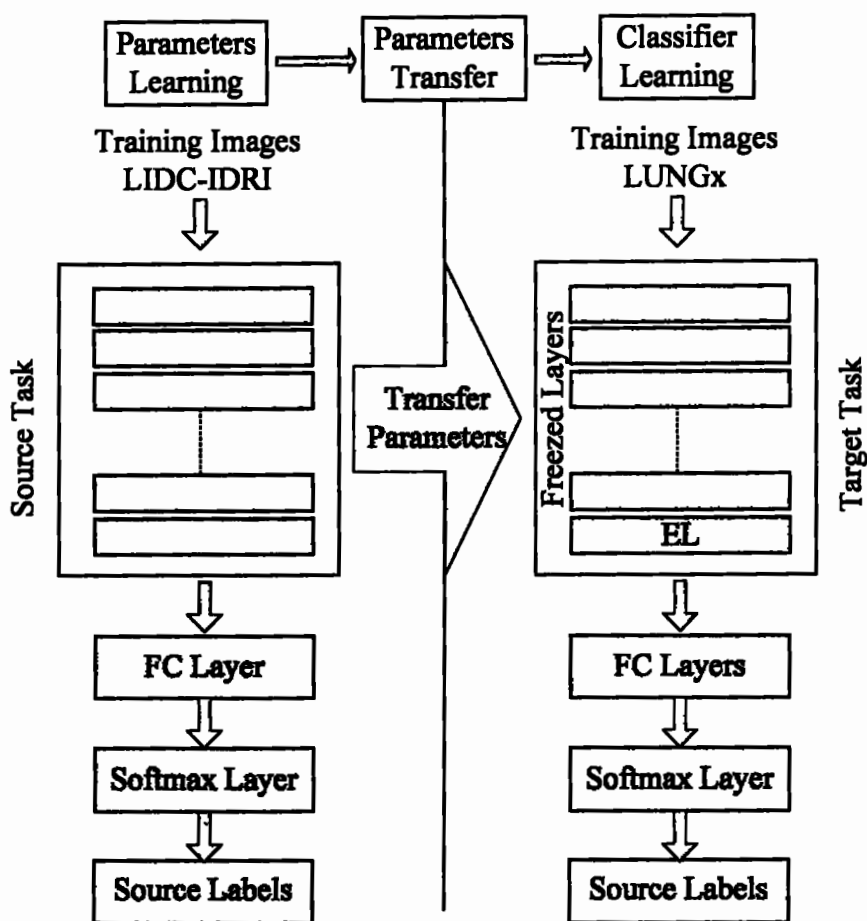


Fig. 3.3. Transfer learning methodology using pre-trained texture CNN

It is a more accurate and stable TL-based classifier model, which learns the significant features of the biomedical image without considering the rich labeled biomedical image dataset. Initially, the network is trained using GPU on the source task (top row of Fig. 3) with many data samples such as an 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 Fig. 3). Here, the source task is the 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 the last two fully connected layers for LUNGx challenge images.

3.3 Training Process

The proposed CNN model is trained and tested on a publicly available LIDC-IDRI [106, 107] database using a six-fold cross-validation strategy. A total of 925,632 image patches of the LIDC-IDRI database is divided into six subsets. Then the six-fold cross-validation is carried out by taking five subsets of data as training and the remaining one as testing to compute the performance of our proposed texture CNN. Furthermore, to avoid the overfitting of the model and monitor the training process, 20% of each k-fold training data is used to validate 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 Fig. 3.4.

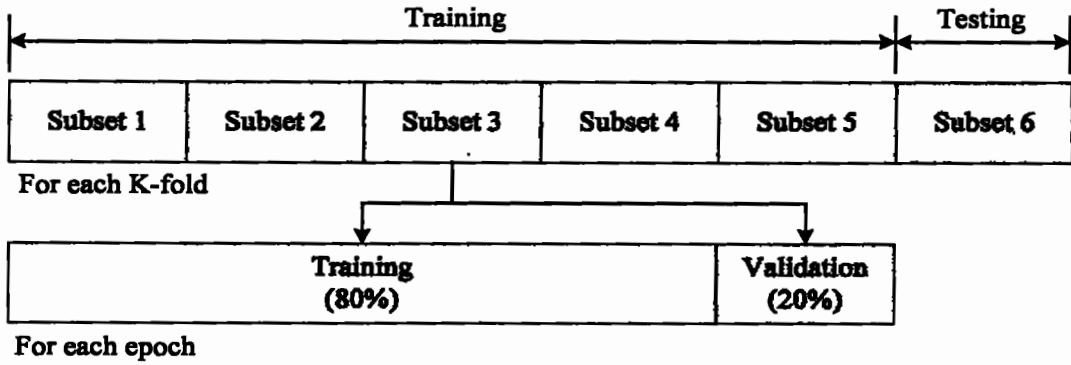


Fig. 3.4. Data distribution detail for training and testing

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 a 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 at 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 and 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 chapter 2.

3.4 Results and Discussion

The implementation of the proposed texture CNN is done with a server having an 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 performance evaluation on the LIDC-IDRI and LUNGx challenge database was performed. After that, we also validated the effectiveness of our model on the MNIST dataset.

3.4.1 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 by changing the dropout layers. Table 3.3 shows the comparison of the results.

Table 3.3. Effect of variation in dropout rates with dropout layers of texture CNN

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%

We sustained all the neurons for the next coming layer when we used no dropout layer. In such a case, the classification accuracy is lower due to overfitting. Furthermore, classification accuracy remained low at the dropout rate of 0.5% and 0.6% due to the withholding of extra neurons. The maximum accuracy is achieved at the dropout value of 0.2%, and 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 3.4.

Table 3.4. Comparison of classic and texture CNN

Models	Dataset	Accuracy(%)	Recall(%)	Error Rate(%)
Texture CNN	LIDC-IDRI	96.69 $\pm$ 0.12	96.05 $\pm$ 0.37	03.30 $\pm$ 0.06
Classic CNN	LIDC-IDRI	92.08 $\pm$ 0.23	95.12 $\pm$ 0.25	07.98 $\pm$ 0.10
Texture CNN	LUNGx Challenge	86.14 $\pm$ 0.21	88.76 $\pm$ 0.22	13.85 $\pm$ 0.19
Classic CNN	LUNGx Challenge	85.71 $\pm$ 0.23	87.77 $\pm$ 0.21	14.85 $\pm$ 0.21

The results show that the proposed texture CNN performed well compared to the classic configuration CNN for both 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 3.4, it can be observed that Texture CNN has a significant improvement in classification accuracy and other metrics, as compared to the classic CNN configuration for both databases. Moreover, we also measured the classification accuracy for each class. These results are given in Table 3.5.

Table 3.5. Classification Score for each class of LIDC-IDRI and LUNGx challenge

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 3.5, 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.

3.4.2 Performance Evaluation with LIDC-IDRI Database

The proposed model achieved comparable classification results on LIDC-IDRI Database. Table 3.6 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 3.6. Performance comparison of proposed texture CNN with state-of-the-art traditional methods

Models	Accuracy (%)	Recall (%)	Specificity (%)	AUC (%)
Narayanan et al. [50]	-	87.86	-	-
Farag et al. [39]	-	-	-	99.00
Han et al. [15]	-	89.35	86.02	94.05
Dhara et al. [31]	-	89.73	86.36	95.05
Shaffie et al. [102]	93.12	92.47	93.60	97.53
Costa et al. [100]	91.81	93.42	91.21	94.00
Sheway et al. [8]	84.00	82.00	93.00	94.00
Wei et al. [101]	87.65	89.30	86.00	94.20
Proposed Model	96.69	96.05	97.37	99.11

The given results show that the achieved accuracy, recall, specificity, and AUC score are 96.69%, 97.16%, 97.19%, and 99.11%, respectively. 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 3.7 shows the performance comparison of the proposed model with various existing state-of-the-art deep learning-based models, like deep fully CNN (DFCNet) [34], Fuse-TSD algorithm [95], MV-KBC learning model [113], MK-SSAC model [114], and GD network [115], etc.

Table 3.7. Performance comparison of proposed texture CNN on the LIDC-IDRI database with state-of-the-art deep learning-based models

Models	Accuracy (%) (Mean SD)	Recall (%) (Mean SD)	Specificity (%) (Mean SD)	AUC (%) (Mean SD)
CMixNet [30]	88.79	93.97	89.83	88.79
LdcNet-FL [116]	97.20	96.00	97.30	98.20
LdcNet-CE [116]	95.60	90.20	96.00	95.60
J. Lyu et al. [2]	92.19	92.10	91.50	97.05
S. Ghosal et al. [117]	95.30	95.00	-	97.00
Local-Global [118]	88.46	88.66	-	95.62
MC-CNN [11]	87.14	77.00	93.00	93.00
Fuse TSD [95]	89.53±0.09	84.19±0.09	89.85±0.25	96.65±0.01
Feature fusion [95]	89.05±0.03	84.33±0.02	91.12±0.19	96.45±0.02
MV-KBC [113]	91.60±0.15	86.52±0.25	94.00±0.03	95.70±0.24
Y. Xie et al. [119]	93.40±0.01	91.43±0.02	94.09±0.02	97.78±0.0001
GD network [115]	92.57±2.47	92.21±4.84	-	95.14±0.78
MK-SSAC [114]	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

The results presented in Table 3.7 show that the proposed model performs better than all other deep learning 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 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.

3.4.3 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) [116], and the customized mixed link network (CMixNet) [30] as given in Table 3.8.

Table 3.8. The architecture complexity comparisons with state-of-the-art methods

Models	Neurons	Parameters	Accuracy
CMixNet [30]	14,725,632	>14.7	88.79%
LdcNet-CE [116]	1,008,898	3,292,763	95.60%
LdcNet-FL [116]	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 3.8, it can be observed that the proposed model has a lesser number of neurons and learnable parameters, which leads to a reduction in complexity. It is noteworthy to mention here that the reduction of neurons and learnable parameters are due to the incorporation of EL. Therefore, the EL reduced the complexity of the network without degrading the classification accuracy.

3.4.4 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 the proposed CNN are compared with MV-KBC [113], fine-tuned MK-SSAC [114], CADx using SVM with tree parzen estimator (TPE), gradient tree boosting (XGBoost) with TPE [120], a lung nodule classification scheme proposed by Mizuho Nishio et al. [121]. Table 3.9 compares our proposed texture CNN on the LUNGx challenge database.

Table 3.9. Performance comparison of proposed texture CNN with state-of-the-art traditional methods on the LUNGx challenge database

Models	Accuracy (%)	Recall (%)	Specificity (%)	AUC (%)
SVM (TPE) [120]	82.00	-	-	85.00
XGBoost (TPE) [120]	86.84	-	-	89.60
Nishio et al. [121]	-	86.70	-	83.70
MK-SSAC [114],	77.26	87.22	67.57	78.83
MV-KBC [113],	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 the 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 except 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 compared to our trained model, which proves the effectiveness of TL methodology.

3.4.5 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 to validate the performance of our proposed transferable texture CNN. We successfully trained and tested our proposed model and compared the results with state-of-the-art techniques. These results are given in Table 3.10.

Table 3.10. Comparisons with state-of-the-art methods on the MNIST dataset

Models	Accuracy	Error rate
Tabik et al. [122]	-	0.10%
Skouson et al. [123]	99.20%	-
Simonovsky et al. [124]	99.37%	-
Klokov et al. [125]	99.10%	-
Grover et al. [126]	99.54%	0.51%
Qi et al. [127]	99.50%	0.51%
Proposed Model	99.89%	0.12%

The results show that the proposed texture CNN also performed well 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. Furthermore, it is also mentioned that Grover et al. achieved recall and specificity of 97.73% and 99.74%, respectively. However, the achieved recall and specificity scores for both metrics by our proposed model are 99.94% and 99.93%, respectively, which reflect the effectiveness of the proposed texture CNN.

3.5 Summary

In this chapter, utilization of energy in transferable texture CNN and transfer learning is discussed. We 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 using various evaluation metrics. The results were compared with the 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 the error rate of 96.69%, 96.05%, 97.37%, 99.11%, and 3.30%, respectively, on the LIDC-IDRI database. The learned features of EL were analyzed, and it was shown that the EL extracted texture from the convolutional layer. The EL also reduced the number of learnable parameters of the network, which minimized the memory requirements and complexity of CNN.

Furthermore, we explored our pre-trained model to handle the smaller dataset classification problem 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 a 0.12% error rate.

The next chapter presents the deep feature selection and decision level fusion technique for lungs nodule classification.

Chapter 4. Deep Feature Selection and Decision Level

Fusion for Lungs Nodule Classification

The decision level fusion technique is proposed to improve the performance of the CAD system for lung nodule classification. This chapter presents the deep feature selection and the decision level fusion. The deep features are extracted by identifying the optimal layers, which improve the performance of classifiers. Then, the performance of SVM and AdaBoostM2 is also evaluated on the basis of deep features extracted from the state-of-the-art transferable architectures (such as; VGG- 16, VGG-19, GoogLeNet, Inception-V3, ResNet-18, ResNet-50, ResNet-101 and InceptionResNet-V2). After that, the performance of the SVM and AdaBoostM2 classifier is analyzed as a function of deep features. The best performing classifier is utilized in our proposed decision level fusion technique.

4.1 Introduction

The classification performance of a typical CAD system depends on lung nodule feature extraction, based on intensity values, shapes, densities, texture, generic and deep features. Therefore, researchers utilized such features, for example, [20], [105], and [149] used size features, whereas, [21], [22], [105], [150-154] used intensity features. The research work presented in [22], [31], [119], [39], [95], [105], [150], [153], [155], [156] utilized shape features, whereas, [157, 158] used density features in their proposed CAD systems. The research work presented in [20], [22], [31], [119], [150], [152], [153] Feature-based lung nodule classification [155], [156], [159], [160] utilized texture features. Similarly, some CAD systems utilized generic features, including LBP [39], [154] scale-invariant feature transform (SIFT) [161-163], Gabor [39], [154], [160], [161], [164] HOG

[22], [156], [161], [8] and speeded-up robust features [165]. Furthermore, Gupta et al. proposed a traditional technique for automatic lung disease detection in CT images using four image features: Gabor, Zernike, Tamura and Haralick features [166]. Meanwhile, DCNN has attracted researchers for the last few years to utilize the deep features in medical image processing due to the continuous improvement in image recognition and classification, as compared to the handcrafted features [81], [95], [131] [167]. Researchers are utilizing deep features from pre-trained DCNN for medical diagnostic, such as; Rajaraman et al. proposed a deep feature-based technique for malaria parasite detection in blood smear images [168]. Chen et al. utilized HOG and deep features for the study of lung nodules [156]. Raj et al. [169] proposed optimal feature selection based deep learning algorithm for medical image classification. They used opposition-based crow search algorithm for optimal feature selection from gray-level and texture features. Finally, the deep network was used for medical image classification. Xie et al. also proposed a transferable multi-model ensemble algorithm based on deep features from ResNet-50 to classify the lung nodules [119]. Similarly, Abraham et al. utilized deep features from AlexNet and VGG-16 for the same application [132]. Whereas, in our work we performed optimum deep feature selection to identify the optimal layers in eight state-of-the-art DCNNs. The deep feature selection is accomplished by carrying-out extensive experiments.

Furthermore, different fusion techniques, such as; feature fusion [23], multi-model image fusion [89], [90], and decision fusion, were proposed for the improvement of CAD systems. Zhang et al. utilized a DCNN model along with the feature fusion to improve the performance of the classification task [24]. Xie et al. proposed a decision level fusion

technique for automatic pulmonary nodule classification. They trained the AdaBoosted back-propagation NN using texture features, and then decision fusion was performed on the outcomes of three classifiers. They also performed the performance analysis of feature fusion and decision fusion and found that the decision fusion performed better [95]. Therefore, we utilized the decision fusion technique in our proposed methodology. Furthermore, the techniques presented in [23], [24], [119], [95], [169] require more training time and memory requirement due to extra fine-tuning or training of DCNN from scratch. Whereas, in our case, we extract the deep features from DCNN during forward propagation using the TL technique.

In this chapter, we presented a decision level fusion technique for the lung nodule classification. First, we evaluated the performance of SVM and AdaBoostM2 algorithms based on the optimal deep features from VGG- 16, VGG-19, GoogLeNet, Inception-V3, ResNet-18, ResNet-50, ResNet-101 and InceptionResNet-V2. Then, we utilized three types of deep features (from GoogLeNet, ResNet-101, and Inception-V3) based on best performance results. For these features, we trained an SVM classifier, and the decision level fusion is performed by estimating the probability scores.

4.2 Material and Methods

4.2.1 Dataset and Pre-Processing

The LUNGx challenge dataset is used to validate the proposed methodology. It has a set of calibration and testing scans. The calibration set has 10 scans (five females, five males). Five of the 10 calibration scans contain one confirmed benign nodule, and the other five contain one pathology-confirmed malignant nodule. The test set has 60 scans which

have a total of 73 nodules. Whereas in 13 scans, each case has two nodules. The total 60 test scans of 23 males and 37 females contain 37 benign nodules and 36 malignant nodules (including 9 non-small cell carcinomas, 15 adenocarcinomas, 1 squamous cell carcinoma, 7 mini cell carcinomas, 2 nodules dubious for malignancy, and 2 carcinoid tumours) [170], [171].

The nodule information is given in the online available annotated CSV file of each scan. The CSV file has an instance number, which is the slice number of each scan, and coordinates of the origin of every nodule. We search the given instance number in the DICOM tag to access the appropriate nodule contained slice. The coordinate information (x, y, z) of ROI around the nodule was used to crop the nodule patch. The patches of size 64×64 were extracted from a voxel. This patch size was selected because all candidates of a nodule would be fully accommodated in this area, as it is noted that the size of the biggest nodule was 30mm in our selected dataset. We extracted 1,144 patches, including 664 patches for class 0 (benign) and 480 patches for class 1 (malignant). Furthermore, the image augmentation is achieved by the flip, rotation, scaling, and translation operations, as shown in Fig. 4.1.

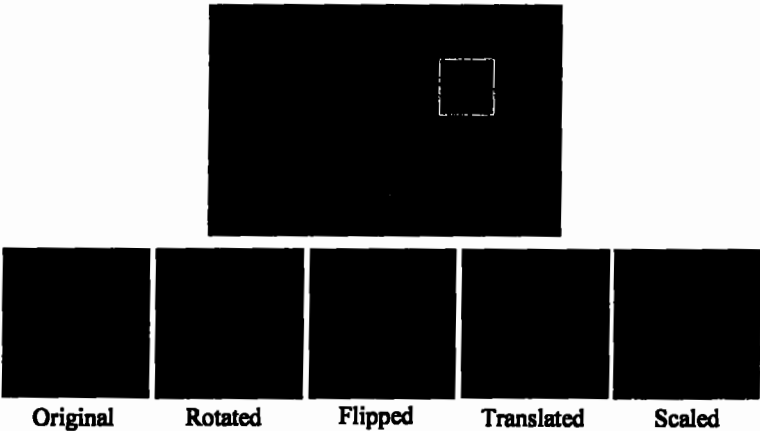


Fig. 4.1. Image augmentation

The flipped patches are extracted by randomly flipping the normal ROI patches horizontally or vertically. The rotated patches are extracted after random rotation of complete slice by 0 to 359°. The scaling is done by increasing the patch size to 10% of the normal ROI patch and then rescale to 64 × 64 patch size. The translated patches are extracted by randomly shifting ROI about 10 to 20 pixels around available coordinate information.

4.2.2 Methodology

The basic strategy of our work is similar to any classic medical image classification technique as Fig. 4.2

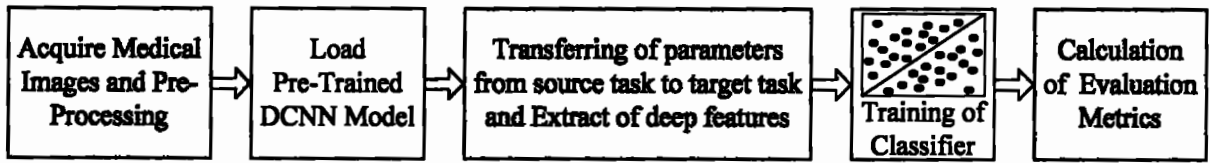


Fig. 4.2. Basic scheme of nodule classification using pre-trained CNN model

After acquiring medical images from the LUNGx database, pre-processing techniques are applied for contrast enhancement. The features are extracted from each DCNN during forward propagation. The SVM and AdaBoostM2 classifiers are trained simultaneously, and the final prediction is done based on the probability scores of each class. We also analyzed the effect of deep features on classification accuracy for each classifier.

4.2.3 Decision Level Fusion Technique

The decision level fusion technique fuses the classification results acquired by different features or data, independently. In this work, we proposed a novel technique that fuses three deep feature maps at the decision level to classify the lung nodule as malignant or benign. The proposed scheme is illustrated in Fig. 4.3.

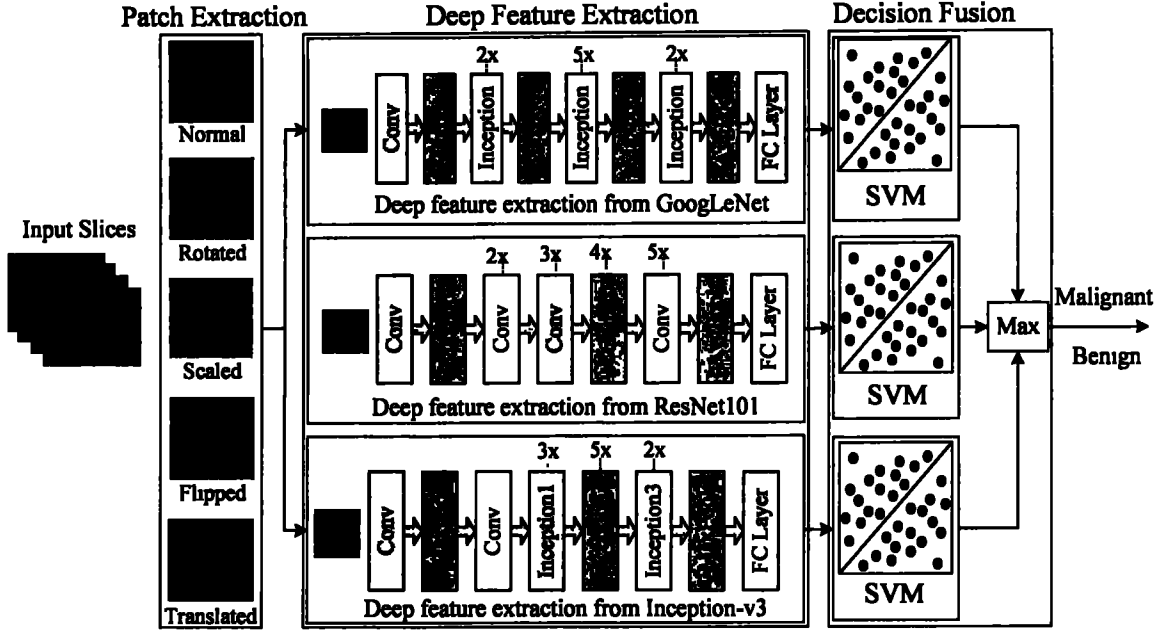


Fig. 4.3. Decision fusion technique

The decision fusion is based on the maximum likelihood of three classification results from the deep features of ResNet-101, GoogLeNet, and Inception-V3. Let us have the following training dataset:

$$\{(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \dots \dots \dots (X_N, Y_N)\} \quad (4.1)$$

where $X_j \in R$ is a deep feature vector of j^{th} image patch, whereas, Y_j denotes the associated class label and N is the total number of training image patches. We extracted three deep feature maps from ResNet-101, GoogLeNet, and Inception-V3 for each lung nodule image patch. If a lung nodule ϕ , consists of i slices, then we extracted i deep feature maps for one nodule. We obtained the following three prediction vectors after SVM training with deep feature maps.

$$\psi_G = P_{jC}(\phi_i), C \in \{malignant; benign\} \quad (4.2)$$

$$\psi_R = P_{jC}(\phi_i), C \in \{malignant; benign\} \quad (4.3)$$

$$\psi_I = P_{jC}(\phi_i), C \in \{malignant; benign\} \quad (4.4)$$

where ψ_G , ψ_R , and ψ_I are the prediction vectors obtained after training SVM on deep features from GoogLeNet, ResNet-101, and Inception-V3, respectively. Each element of $P_{jC}(\phi_i)$ denotes the likelihood of nodule ϕ_i class as malignant or benign predicted from the j^{th} deep feature map of i^{th} slice. The prediction vector is updated based on the maximum likelihood of each feature map, as given in equation 4.5.

$$\psi_F = \text{argmax} (\psi_G, \psi_R, \psi_I) \quad (4.5)$$

where ψ_F is the final prediction vector. The labels were assigned according to the prediction score of each class in the final prediction vector, as given in equation 4.6.

$$\bar{Y}_G = \begin{cases} \text{Malignant}, & \text{if Score} > 0.5 \\ \text{Benign}, & \text{otherwise} \end{cases} \quad (4.6)$$

The flowchart of the proposed methodology is illustrated in Fig. 4.4.

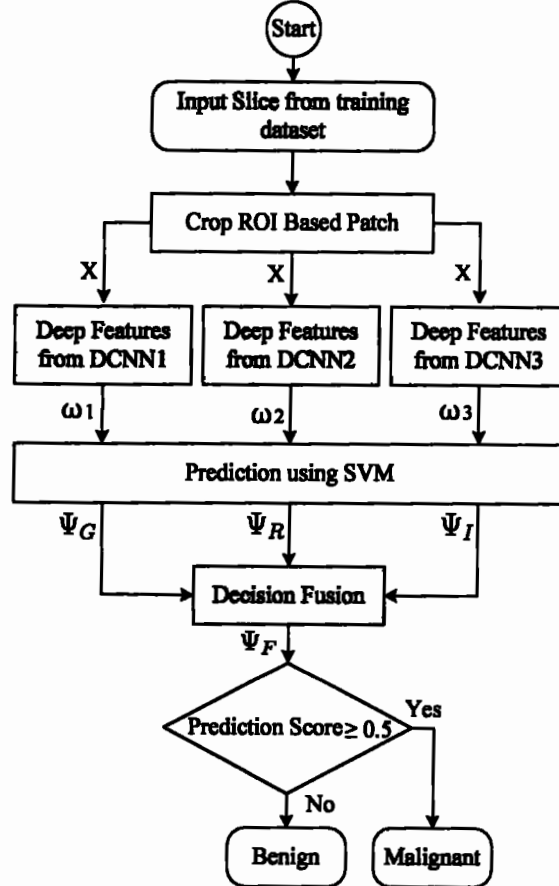


Fig. 4.4. Flow chart of decision fusion

We show that our proposed technique effectively classifies the malignant and benign nodules when tested on the LUNGx challenge dataset.

4.4 Experiments and Results

The lung nodule classification is done using decision level fusion after selecting a deep feature map from state-of-the-art DCNNs. The whole experiment was carried out using Matlab R2019b. All experiments were performed with Intel(R) Core(TM) i7-8550 processor, 8GB. RAM, and one NVIDIA GeForce MX150 GPU with compute capability of 6.1.

4.4.1 Nodule Classification with SVM and AdaBoostm2 using Deep Features

First, the lung nodule classification is done by SVM and AdaBoostM2 classifier after extracting deep features from DCNNs. For the training of the classifier, we used four-fold cross-validation. We divided the dataset into four subsets, three subsets used for training and one for the test, as shown in Fig. 4.5.

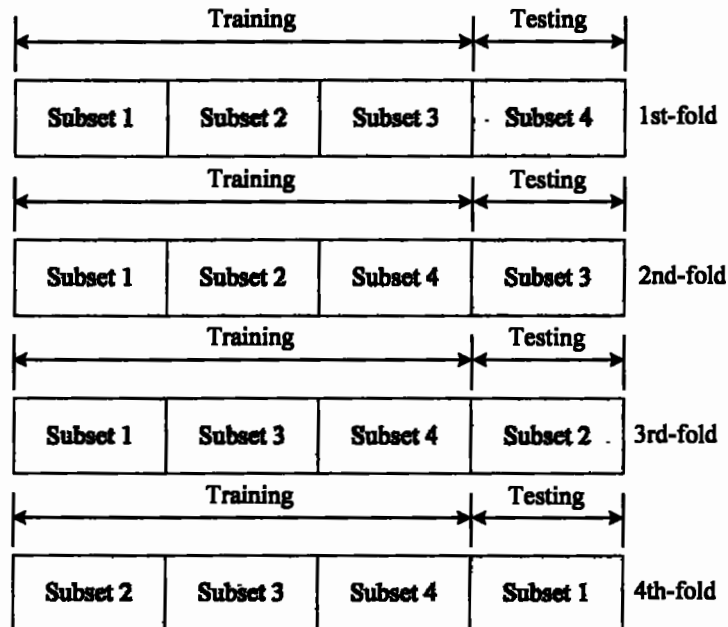


Fig. 4.5. Distribution of data for training and testing

First, we performed ROC analysis by computing ROC curves of all DCNN models with SVM and AdaBoostM2. The results are obtained by extracting features from the last FC layer of each DCNN model. For example, FC2 of VGG-Net, fc1000 of ResNet, loss3-classifier of GoogLeNet, and predictions layer of both inception DCNNs. The ROC comparison of all DCNNs using SVM and AdaBoostM2 is shown in Fig. 4.6.

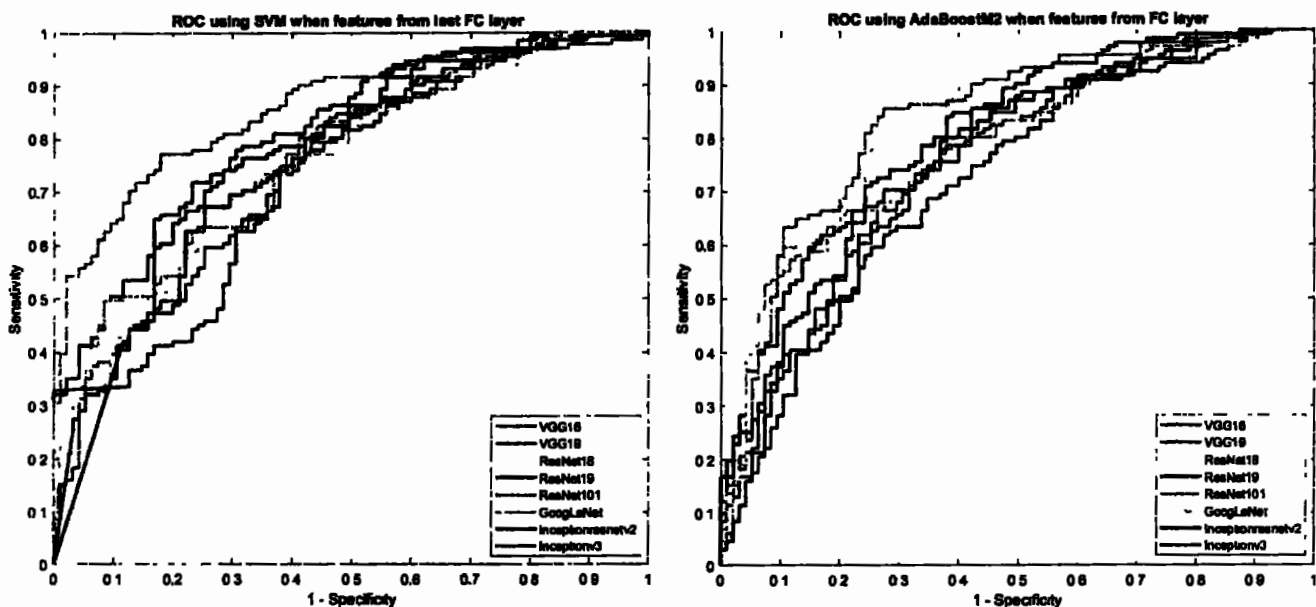


Fig. 4.6. ROC comparison of SVM and AdaBoostM2 when features from the last FC layer

We also computed the AUC value for each DCNN model with SVM and AdaBoostM2 classifiers to analyze the effect of each classifier on lung nodule classification. This analysis revealed which classifier is more effective on extracted features for each DCNN model. The AUC plots of SVM and AdaBoostM2 for all DCNN models are shown in Fig. 4.7.

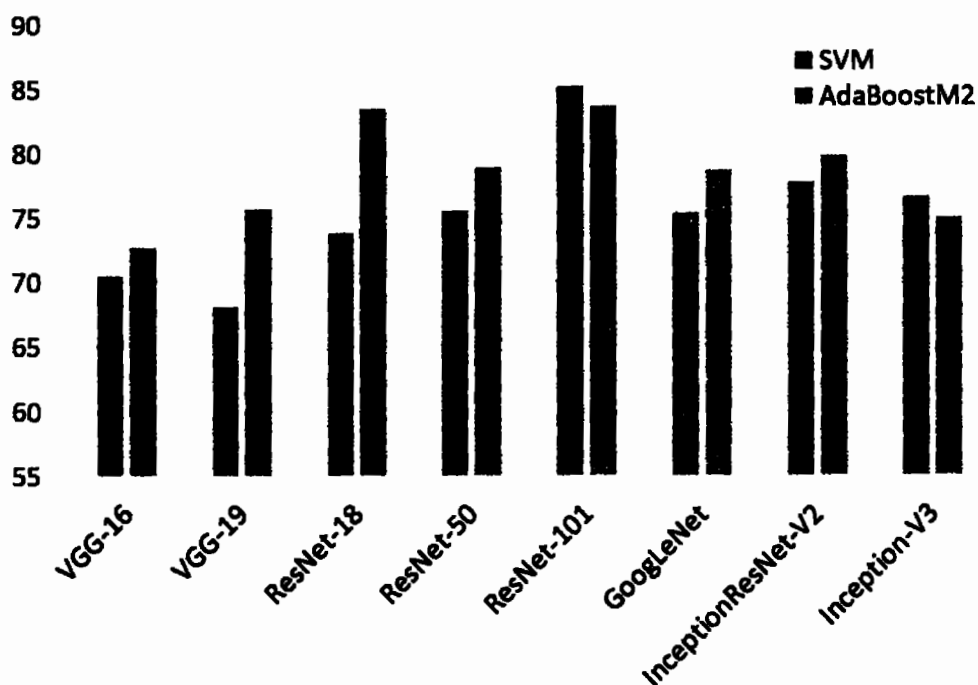


Fig. 4.7. AUC comparison of SVM and AdaBoostM2 when features from Last FC layer

It can be observed that the AdaBoostM2 performed well on the features extracted from the last fully connected layers of all DCNNs, except ResNet-101. The features extracted from ResNet-101 provide the maximum AUC score using the SVM classifier. We also evaluated the performance of lung nodule classification with SVM and AdaBoostM2. Table 4.1 shows the lung nodule classification results obtained by the SVM classifier.

Table 4.1. Nodule classification with SVM when feature the last FC layer

Models	Accuracy	Gmean	Precision	Recall	Error Rate
VGG-16	65.27%	63.22%	70.31%	90.99%	34.73%
VGG-19	64.05%	61.30%	67.45%	73.28%	35.95%
ResNet-18	67.70%	63.60%	69.81%	78.82%	32.2%
ResNet-50	66.92%	67.22%	77.47%	61.07%	33.08%
ResNet-101	76.88%	77.34%	86.48%	71.56%	23.12%
GoogLeNet	67.37%	67.71%	77.07%	62.92%	32.63%
Incep.ResNet-V2	72.01%	66.09%	71.00%	87.79%	27.99%
Inception-V3	71.13%	71.23%	78.86%	69.08%	28.87%

The results show that the features from ResNet-101 perform better than the rest of the models in terms of accuracy using the SVM classifier. The maximum accuracy with ResNet-101 is 76.88%, whereas InceptionResNet-V2 and Inception-V3 have 72.01% and 71.13%, respectively. Similarly, the ResNet-101 also performs better than other models in terms of Gmean, precision, and Error Rate. It can also be observed that the SVM classifier had a minimum classification score when deep features were extracted from the last layer of VGG-19. After this detailed analysis, we repeated the same experiment for the AdaBoostM2 classifier. Table 4.2 shows the lung nodule classification results with the AdaBoostM2 classifier.

Table 4.2. Classification results using AdaBoostM2 when features from last FC layer

Models	Accuracy	Gmean	Precision	Recall	Error Rate
VGG-16	66.81%	65.38%	70.90%	72.52%	33.19%
VGG-19	70.35%	69.09%	73.88%	75.57%	29.65%
ResNet-18	75.22%	72.82%	75.86%	83.97%	24.78%
ResNet-50	69.91%	68.01%	72.66%	77.10%	30.09%
ResNet-101	78.32%	77.94%	82.03%	80.15%	21.68%
GoogLeNet	71.24%	69.54%	73.91%	77.86%	28.76%
Incep.ResNet-V2	73.01%	70.89%	74.65%	80.92%	26.99%
Inception-V3	72.57%	70.81%	74.82%	79.39%	27.43%

It can be observed that the deep features extracted from ResNet-101 produced better results with the AdaBoostM2 classifier. The mean accuracy score of four-fold is 78.32%, and the error rate is 21.68%. Similarly, the ResNet-18 is second and InceptionResNet-V2 at third position in terms of accuracy and error rate score. Furthermore, VGG-16 is in the last position with the AdaBoostM2 classifier and achieved an accuracy score of 66.81%.

The results in Table 4.1 and Table 4.2 show that the AdaBoostM2 performed well compared to the SVM classifier when deep features were extracted from ResNet-101.

4.4.2 Optimum Deep Feature Selection

We further studied the AdaBoostM2 and SVM classifiers by optimizing the deep parameters. First, we extracted features from FC6 of VGG-16 and trained the AdaBoostM2 classifier. In this case, we observed that the accuracy of VGG-16 was improved to 69.91%, with an error rate of 30.08%. After this experiment, the same test was also performed using the SVM classifier. The acquired accuracy score was increased from 65.27% to 72.56%. Therefore, we decided to select the optimum deep features for all DCNN models by identifying the optimal layer for feature extraction, which can provide the best performance using AdaBoostM2 and SVM. For this purpose, we evaluated the performance of each DCNN by extracting features from different layers and identifying the optimal layer. The optimal layers for deep feature extraction, which provides the best performance results, are given in Table 4.3.

Table 4.3. Optimal layers give the best performance

Model	Layer Number	Optimal Layer
VGG-16	32	'pool5'
VGG-19	38	'pool5'
GoogLeNet	139	'inception_5b-output'
Inception-V3	312	'mixed10'
ResNet-18	67	'res5b_relu'
ResNet-50	173	'activation_49_relu'
ResNet-101	343	'res5c_relu'
Incep.ResNet-V2	820	'conv_7b_ac'

Fig. 4.8 showed the ROC comparison of SVM and AdaBoostM2 when features were extracted from optimal layers. We can observe that the ROC plot with optimal layers (in Fig. 4.8) is improved from the previous plot, as shown in Fig. 4.6.

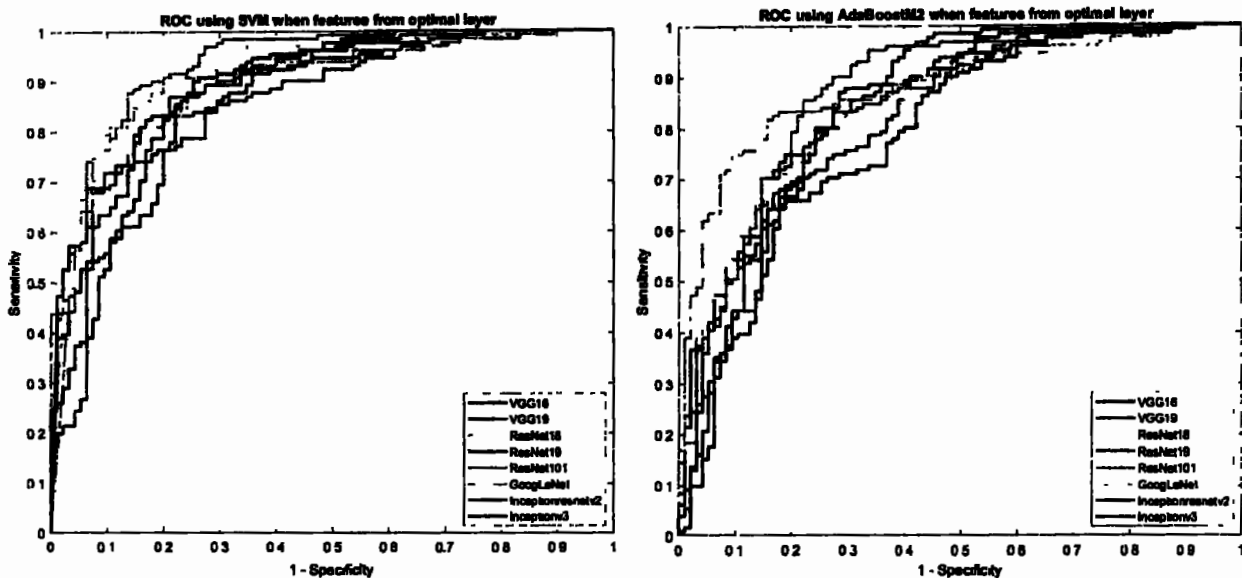


Fig. 4.8. ROC comparison of SVM and AdaBoostM2 when features from optimal layers
Moreover, we also analyzed the SVM and AdaBoostM2 by comparing their AUC values.

The AUC comparison of SVM and AdaBoostM2 is shown in Fig. 4.9.

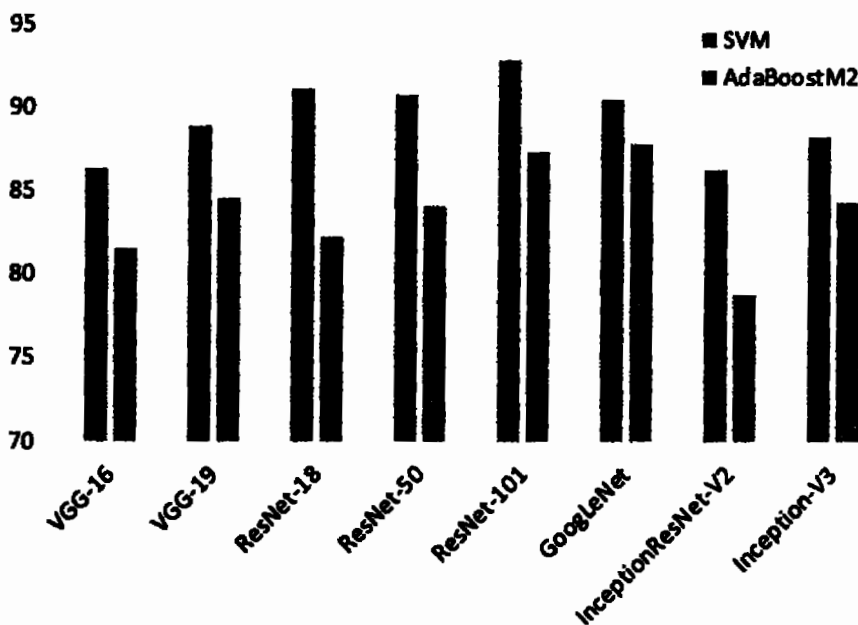


Fig. 4.9. AUC comparison of SVM and AdaBoostM2 using features from optimal layers
It is noted that the AUC results for both SVM and AdaBoostM2 using features from optimal layers are improved. Furthermore, SVM performed well as compared to the AdaBoostM2. The maximum achieved AUC score of SVM using features from ResNet-

101 is 92.77%. Furthermore, we also analyzed the improvement in AUC using features from the optimal layer using the SVM classifier. Fig. 4.10 shows the AUC comparison results of the SVM classifier for features from optimal and last FC layers.

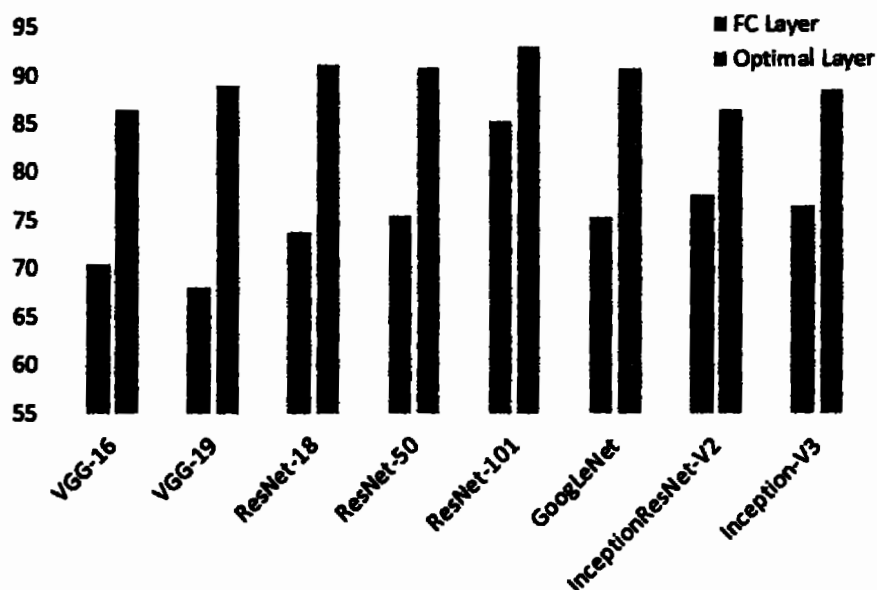


Fig. 4.10. AUC comparison of SVM using features from optimal and last FC layers

The remaining evaluation metrics are also evaluated for the classifiers with features from optimal layers. The classification results using the AdaBoostM2 classifier with features from optimal layers are given in Table 4.4.

Table 4.4. Nodule classification results using Optimal Layers with AdaBoostM2

Models	Accuracy (%) Mean±Std	Gmean (%) Mean±Std	Precision (%) Mean±Std	Recall (%) Mean±Std	Error Rate (%) Mean±Std
VGG-16	73.01±0.32	71.63±0.34	75.70±0.25	78.62±0.29	27.01±0.32
VGG-19	77.00±0.24	75.20±0.32	78.00±0.27	84.00±0.28	23.00±0.24
ResNet-18	76.11±0.26	73.20±0.19	75.80±0.20	86.30±0.19	23.90±0.26
ResNet-50	77.88±0.18	77.00±0.17	80.50±0.19	81.70±0.21	22.10±0.18
ResNet-101	79.20±0.17	79.30±0.19	84.40±0.17	78.60±0.12	20.08±0.17
GoogLeNet	82.30±0.20	82.30±0.19	86.40±0.22	82.40±0.18	17.70±0.20
InceptionResNet-V2	70.79±0.15	69.43±0.19	74.10±0.20	76.30±0.21	29.20±0.15
Inception-V3	78.32±0.21	76.71±0.24	79.28±0.23	84.73±0.22	21.68±0.21

It is noted that the acquired results were significantly improved as compared to the previous ones, as mentioned in Table 4.2. For example, for the case of VGG-16, the accuracy increased from 66.81% to 73.01%, and the error rate decreased from 33.19% to 27.01%. Similarly, the other evaluation metrics like Gmean, precision, recall, and F1 score were also improved. Furthermore, the GoogLeNet accuracy score was improved from 71.24% to 82.30%. After significantly improving classification results using AdaBoostM2 with features extracted from optimal layers, we utilized the same features with SVM. Table 4.5 shows the lung nodule classification results using the features extracted from optimal layers with the SVM classifier.

Table 4.5. Nodule classification results using optimal layer with SVM

Models	Accuracy (%) Mean±Std	Gmean (%) Mean±Std	Precision (%) Mean±Std	Recall (%) Mean±Std	Error Rate (%) Mean±Std
VGG-16	77.43±0.96	77.38±0.74	82.39±0.38	77.67±2.19	22.56±0.96
VGG-19	80.90±0.66	81.20±0.44	87.10±1.16	78.60±2.57	19.10±0.66
ResNet-18	82.20±0.22	81.50±0.71	81.20±1.17	85.30±2.19	17.80±0.22
ResNet-50	81.42±1.08	81.30±1.07	85.50±2.00	81.90±3.10	18.60±1.08
ResNet-101	86.28±0.82	85.90±0.46	88.20±1.34	88.20±0.98	13.70±0.82
GoogLeNet	83.40±0.44	83.80±0.51	89.50±0.76	80.90±0.01	16.60±0.44
Incep.ResNet-V2	80.31±0.85	80.10±0.72	84.20±0.24	81.30±1.46	19.70±0.85
Inception-V3	82.63±1.42	82.30±1.07	85.50±0.63	84.40±3.33	17.40±1.42

The results show that all the evaluation metrics were improved using features from optimal layers. In this case, the deep features from ResNet-101 provided better results than the other DCNN models. The maximum accuracy was 86.28%, and the error rate was 13.70%. GoogLeNet at second position with an accuracy of 83.40% and an error rate of 16.60%. VGG-16 was at last position with an accuracy of 77.43% and the error rate of 22.56%, but it can be noted that in the previous case, the accuracy was 65.27% with an error rate of 34.73%.

4.4.3 Results for Decision Level Fusion

The decision level fusion was applied on ResNet-101, GoogLeNet, and Inception-V3, which were at the top level in lung nodule classification performance. The classification results of decision fusion were based on the maximum computed likelihood score among both DCNN models. The classification accuracy for each benign and malignant class of the LUNGx challenge dataset is given in Table 4.6.

Table 4.6. Nodule classification of each class with decision fusion

Models	Accuracy	Error Rate
Benign	90.70%	9.20%
Malignant	90.20%	9.78%

The results show that the proposed decision fusion technique accurately classified each class with an accuracy score of 90.70% for benign and 90.20% for malignant. Moreover, the ROC plot comparison of the decision fusion technique, ResNet-101, GoogLeNet, and Inception-V3 is shown in Fig. 4.11.

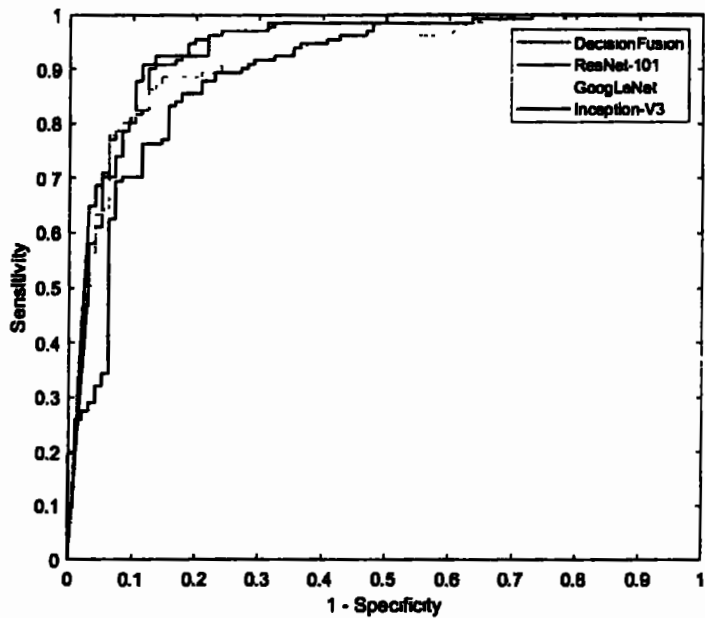


Fig. 4.11. ROC comparison of decision fusion technique with ResNet-101, GoogLeNet, and Inception-V3

It can be observed that the area under ROC of proposed techniques is improved as compared to the ResNet-101, GoogLeNet, and Inception-V3. The comparison of our proposed decision level fusion with the stat-of-the-art techniques is given in Table 4.7.

Table 4.7. Comparison of proposed decision level fusion with stat-of-the-art techniques

Models	Accuracy (%)	Recall (%)	Specificity (%)	AUC (%)
Nishio et al. [121]	—	86.70	74.40	83.70
SVM (TPE) [120]	82.00	—	—	85.00
XGBoost (TPE) [120]	86.84	—	—	89.60
Wang et al. [79]	90.38	88.87	92.40	94.48
MV-KBC [113]	75.62±1.15	87.22±7.24	64.32±7.00	76.85±0.17
MK-SSAC [113]	76.16±0.75	86.11±7.61	66.49±8.46	77.59±0.39
Xie et al. [114]	77.26±0.75	87.22±7.25	67.57±6.34	78.83±0.75
Texture CNN [133]	86.14±1.40	88.76±1.90	93.11±1.45	92.63±0.75
Proposed	90.46±0.25	90.10±0.44	92.56±0.18	94.49±0.11

It can be observed that the proposed technique outperforms in terms of accuracy, recall, and AUC with the achieved score of 90.46±0.25%, 90.10±0.44%, and 94.46±0.11%, respectively. The proposed decision fusion is at second position in terms of specificity with a score of 92.56±0.18%, but the specificity score deviation is very low compared to the Texture CNN. Furthermore, we also compute the classification accuracy score of the proposed technique based on the simple average of the prediction scores, which is 89.10%.

4.5 Summary

In this chapter, we proposed a decision level fusion technique for lung nodule classification to improve the classification performance of the CAD system. First, we

evaluated the performance of SVM and AdaBoostM2 for lungs nodule classification, based on deep features, which are extracted from eight state-of-the-art transferable architectures. We selected the optimum deep features by identifying the optimal layers for feature extraction with a series of experiments by monitoring the variation in the classification performance. We also showed an improvement in AUC compared to the features from optimal and FC layers. The optimum deep features improved the classification performance. For example, the classification accuracy of ResNet-101 was improved from 76.88% to 86.28%, and the accuracy of GoogLeNet increased from 67.37% to 83.40%. We also observed that the SVM outperformed AdaBoostM2 with deep features from optimal layers from all eight DCNNs. Finally, we showed that the proposed decision level fusion technique performed well as compared to the recent state-of-the-art techniques. The achieved classification score was $90.46 \pm 0.25\%$, $90.10 \pm 0.44\%$, $92.56 \pm 0.18\%$, and $94.49 \pm 0.11\%$ in terms of accuracy, recall, specificity, and AUC, respectively.

Chapter 5. Conclusions and Future Work

This chapter describes the conclusion inferred from the proposed deep learning methodologies for lung nodule classification. Moreover, this chapter also includes the guidelines for researchers interested in utilizing the proposed methodologies in medical diagnostics or various engineering fields.

5.1 Conclusion

Lung nodules are vital indicators for the presence of lung cancer. Early detection enhances the survival rate of the patient by starting the treatment at the right time. The detection and classification of malignancy in CT images is a time-consuming and difficult task for the radiologists, leading the researchers to develop CAD systems to mitigate this burden. In this dissertation, we studied various methodologies to improve the performance of the CAD system for lung nodule classification.

The efficient lung nodule classification was performed using the proposed transferable texture CNN. The utilization of EL in the proposed texture CNN model contributes to lung nodule malignancy classification. The EL preserves the texture information, reduces the output vector size, and learns the parameters during forward and backward propagation, increasing the overall learning capability of the model. The EL also reduces the number of learnable parameters of the network, which minimizes the memory requirements and complexity of CNN. Experimental results show the effectiveness of the proposed technique for benign and malignant nodules classification, without nodule segmentation or any complex pre-processing. The proposed texture CNN was tested on two medical image datasets: LIDC-IDRI and LUNGx challenge. The achieved classification accuracy, recall,

specificity, AUC, and an error rate of 96.69%, 96.05%, 97.37%, 99.11%, and 3.30%, respectively, on the LIDC-IDRI database. In addition, the proposed model has fewer learnable parameters than LdcNet-FL, which reduces the memory requirement and computational complexity of the algorithm. Moreover, the effectiveness of our proposed texture CNN is also validated on the MNIST dataset, as our model achieved 99.89% accuracy with only a 0.12% error rate.

We also used a pre-trained model of texture CNN to handle the smaller medical dataset classification problem using TL. This model utilizes LIDC-IDRI as the source task and the LUNGx challenge dataset as the target task. The classification accuracy of the LUNGx challenge dataset was 86.14% without TL, which was further improved to 90.91% using the proposed TL base model. We also show that our pre-trained model achieved better results than the state-of-the-art techniques, such as; fine-tuned MK-SSAC, MV-KBC, CADx using SVM with tree parzen estimator (TPE) and gradient tree boosting (XGBoost) with TPE.

A decision level fusion technique was also proposed to improve the performance of the CAD system for lung nodule classification. First, we evaluated the performance of SVM and AdaBoostM2 algorithms based on the deep features of the state-of-the-art transferable architectures (such as; VGG-16, VGG-19, GoogLeNet, Inception-V3, ResNet-18, ResNet-50, ResNet-101 and InceptionResNet-V2). Then, we analyzed the performance of the SVM and AdaBoostM2 classifier as a function of deep features. We selected the optimum deep features by identifying the optimal layers with a series of experiments by monitoring the variation in the classification performance. We showed the improvement in AUC compared to the features from optimal and FC layers. We also showed that the

optimal deep features improved the performance of the classifiers, for example, the classification accuracy of ResNet-101 was improved from 76.88% to 86.28%.

Similarly, the accuracy of GoogLeNet was increased from 67.37% to 83.40%. Finally, we showed that the proposed decision level fusion technique performed well as compared to the recent state-of-the-art methods. The achieved classification score was $90.46 \pm 0.25\%$, $90.10 \pm 0.44\%$, $92.56 \pm 0.18\%$, and $94.49 \pm 0.11\%$ in terms of accuracy, recall, specificity, and AUC, respectively, on the LUNGx challenge dataset.

5.2 Future Work

The guidelines for future work in the field of medical image diagnostics are as follows:

- The proposed methodologies are only tested on lung CT images, so the proposed work may also be applied to other modalities, such as; X-rays and PET-CT.
- The presented techniques are proposed for lung nodule classification, the same work can be extended for the classification of other cancer types, such as; brain, breast, or prostate cancers.
- The transferable texture CNN is trained and tested on the available LIDC-IDRI database for only lung cancer classification. It can be trained on an artificial medical image dataset generated by generative adversarial networks and tested on the original medical dataset for further enhancement in classification accuracy. The same trained model can also be utilized for the classification of other types of cancers using transfer learning.
- The decision fusion is proposed based on the probability score of the DCNN. The deep feature fusion technique can be adopted by including deep features from the

same deeper models, such as; GoogLeNet, ResNet101, Inception-V3, and InceptionResNet-V2.

- The hybrid feature fusion, such as; deep feature and handcraft feature fusion, can be utilized for medical diagnostics.
- The ensemble learning-based techniques can be adopted with different classifiers such as; ELM, KNN, random forest, and binary trees which take the deep feature from GoogLeNet, ResNet101, Inception-V3, and InceptionResNet-V2

REFERENCES

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a cancer journal for clinicians*, vol. 68, pp. 394-424, 2018.
- [2] J. Lyu, X. Bi, and S. H. Ling, "Multi-level cross residual network for lung nodule classification," *Sensors*, vol. 20, p. 2837, 2020.
- [3] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA: a cancer journal for clinicians*, vol. 69, pp. 7-34, 2019.
- [4] X. Wang, K. Mao, L. Wang, P. Yang, D. Lu, and P. He, "An appraisal of lung nodules automatic classification algorithms for CT images," *Sensors*, vol. 19, p. 194, 2019.
- [5] S. K. Mandal, "Performance Analysis Of Data Mining Algorithms For Breast Cancer Cell Detection Using Naïve Bayes, Logistic Regression and Decision Tree," *International Journal Of Engineering And Computer Science*, vol. 6, pp. 20388-20391, 2017.
- [6] Y. Sui, Y. Wei, and D. Zhao, "Computer-aided lung nodule recognition by SVM classifier based on combination of random undersampling and SMOTE," *Computational and mathematical methods in medicine*, vol. 2015, 2015.
- [7] M. Jadoon, Q. Zhang, I. U. Haq, A. Jadoon, A. Basit, and S. Butt, "Classification of mammograms for breast cancer detection based on curvelet transform and multi-layer perceptron," *Biomedical Research*, vol. 28, 2017.
- [8] T. N. S. a. A. A. Mekonnen, "Benign-Malignant Lung Nodule Classification with Geometric and Appearance Histogram Features," 26 May 2016.
- [9] E. Adetiba and O. O. Olugbara, "Lung cancer prediction using neural network ensemble with histogram of oriented gradient genomic features," *The Scientific World Journal*, vol. 2015, 2015.
- [10] K. Liu and G. Kang, "Multiview convolutional neural networks for lung nodule classification," *International Journal of Imaging Systems and Technology*, vol. 27, pp. 12-22, 2017.
- [11] W. Shen, M. Zhou, F. Yang, D. Yu, D. Dong, C. Yang, *et al.*, "Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification," *Pattern Recognition*, vol. 61, pp. 663-673, 2017.
- [12] W. J. Sori, J. Feng, and S. Liu, "Multi-path convolutional neural network for lung cancer detection," *Multidimensional Systems and Signal Processing*, vol. 30, pp. 1749-1768, 2019.
- [13] 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," *Biomedical engineering online*, vol. 15, pp. 1-17, 2016.

- [14] S. Jia, B. Deng, J. Zhu, X. Jia, and Q. Li, "Local binary pattern-based hyperspectral image classification with superpixel guidance," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, pp. 749-759, 2017.
- [15] F. Han, H. Wang, G. Zhang, H. Han, B. Song, L. Li, *et al.*, "Texture feature analysis for computer-aided diagnosis on pulmonary nodules," *Journal of digital imaging*, vol. 28, pp. 99-115, 2015.
- [16] N. Banerjee, N. Dimitrova, V. Varadan, S. Kamalakaran, A. Janevski, and S. Maity, "Integrated phenotyping employing image texture features," ed: Google Patents, 2017.
- [17] F. Ciompi, B. de Hoop, S. J. van Riel, K. Chung, E. T. Scholten, M. Oudkerk, *et al.*, "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," *Medical image analysis*, vol. 26, pp. 195-202, 2015.
- [18] V. Andrearczyk and P. F. Whelan, "Using filter banks in convolutional neural networks for texture classification," *Pattern Recognition Letters*, vol. 84, pp. 63-69, 2016.
- [19] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, *et al.*, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE transactions on medical imaging*, vol. 35, pp. 1285-1298, 2016.
- [20] J. Wang, X. Liu, D. Dong, J. Song, M. Xu, Y. Zang, *et al.*, "Prediction of malignant and benign of lung tumor using a quantitative radiomic method," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 1272-1275.
- [21] H. Kim, C. M. Park, E. J. Hwang, S. Y. Ahn, and J. M. Goo, "Pulmonary subsolid nodules: value of semi-automatic measurement in diagnostic accuracy, diagnostic reproducibility and nodule classification agreement," *European radiology*, vol. 28, pp. 2124-2133, 2018.
- [22] S. K. Dilger, A. Judisch, J. Uthoff, E. Hammond, J. Newell, and J. C. Sieren, "Improved pulmonary nodule classification utilizing lung parenchyma texture features," in *Medical Imaging 2015: Computer-Aided Diagnosis*, 2015, p. 94142T.
- [23] C. Wang, A. Elazab, J. Wu, and Q. Hu, "Lung nodule classification using deep feature fusion in chest radiography," *Computerized Medical Imaging and Graphics*, vol. 57, pp. 10-18, 2017.
- [24] G. Zhang, D. Zhu, X. Liu, M. Chen, L. Itti, Y. Luo, *et al.*, "Multi-scale pulmonary nodule classification with deep feature fusion via residual network," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-12, 2018.
- [25] P. W. Sperduto, T. J. Yang, K. Beal, H. Pan, P. D. Brown, A. Bangdiwala, *et al.*, "Estimating survival in patients with lung cancer and brain metastases: an update of the graded prognostic assessment for lung cancer using molecular markers (Lung-molGPA)," *JAMA oncology*, vol. 3, pp. 827-831, 2017.

- [26] P. J. Eifel, M. Morris, J. T. Wharton, and M. J. Oswald, "The influence of tumor size and morphology on the outcome of patients with FIGO stage IB squamous cell carcinoma of the uterine cervix," *International Journal of Radiation Oncology* Biology* Physics*, vol. 29, pp. 9-16, 1994.
- [27] A. R. Larici, A. Farchione, P. Franchi, M. Ciliberto, G. Cicchetti, L. Calandriello, *et al.*, "Lung nodules: size still matters," *European Respiratory Review*, vol. 26, 2017.
- [28] D. M. Hansell, A. A. Bankier, H. MacMahon, T. C. McLoud, N. L. Muller, and J. Remy, "Fleischner Society: glossary of terms for thoracic imaging," *Radiology*, vol. 246, pp. 697-722, 2008.
- [29] K. Suzuki, "A review of computer-aided diagnosis in thoracic and colonic imaging," *Quantitative imaging in medicine and surgery*, vol. 2, p. 163, 2012.
- [30] 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, p. 3722, 2019.
- [31] 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," *Journal of digital imaging*, vol. 29, pp. 466-475, 2016.
- [32] A. M. Schilham, B. Van Ginneken, and M. Loog, "A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database," *Medical Image Analysis*, vol. 10, pp. 247-258, 2006.
- [33] E. Dandil, "A computer-aided pipeline for automatic lung cancer classification on computed tomography scans," *Journal of healthcare engineering*, vol. 2018, 2018.
- [34] A. Masood, B. Sheng, P. Li, X. Hou, X. Wei, J. Qin, *et al.*, "Computer-assisted decision support system in pulmonary cancer detection and stage classification on CT images," *Journal of biomedical informatics*, vol. 79, pp. 117-128, 2018.
- [35] T. Kohonen, "Self-organizing maps of massive databases," *International journal of engineering intelligent systems for electrical engineering and communications*, vol. 9, pp. 179-186, 2001.
- [36] H. Han, L. Li, F. Han, B. Song, W. Moore, and Z. Liang, "Fast and adaptive detection of pulmonary nodules in thoracic CT images using a hierarchical vector quantization scheme," *IEEE journal of biomedical and health informatics*, vol. 19, pp. 648-659, 2014.
- [37] B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and M. M. Goodsitt, "Computerized characterization of masses on mammograms: The rubber band straightening transform and texture analysis," *Medical Physics*, vol. 25, pp. 516-526, 1998.
- [38] S. G. Kandemirli, S. Chopra, S. Priya, C. Ward, T. Locke, N. Soni, *et al.*, "Presurgical detection of brain invasion status in meningiomas based on first-order histogram based texture analysis of contrast enhanced imaging," *Clinical Neurology and Neurosurgery*, vol. 198, p. 106205, 2020.

- [39] A. A. Farag, A. Ali, S. Elshazly, and A. A. Farag, "Feature fusion for lung nodule classification," *International journal of computer assisted radiology and surgery*, vol. 12, pp. 1809-1818, 2017.
- [40] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, pp. 610-621, 1973.
- [41] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, pp. 971-987, 2002.
- [42] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, 2005, pp. 886-893.
- [43] K. S. Vidya, E. Ng, U. R. Acharya, S. M. Chou, R. San Tan, and D. N. Ghista, "Computer-aided diagnosis of myocardial infarction using ultrasound images with DWT, GLCM and HOS methods: a comparative study," *Computers in biology and medicine*, vol. 62, pp. 86-93, 2015.
- [44] H. Çamdevýren, N. Demýr, A. Kanik, and S. Keskýn, "Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-a in reservoirs," *Ecological Modelling*, vol. 181, pp. 581-589, 2005.
- [45] U. Akilandeswari, R. Nithya, and B. Santhi, "Review on feature extraction methods in pattern classification," *European Journal of Scientific Research*, vol. 71, pp. 265-272, 2012.
- [46] Y. Mingqiang, K. Kidiyo, and R. Joseph, "A survey of shape feature extraction techniques," *Pattern recognition*, vol. 15, pp. 43-90, 2008.
- [47] D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," *Canadian Journal of remote sensing*, vol. 28, pp. 45-62, 2002.
- [48] G. Van de Wouwer, P. Scheunders, and D. Van Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE transactions on image processing*, vol. 8, pp. 592-598, 1999.
- [49] 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," *Journal of Medical Imaging*, vol. 5, p. 014504, 2018.
- [50] 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 Analysis and Applications*, vol. 22, pp. 559-571, 2019.
- [51] T. N. Shewaye and A. A. Mekonnen, "Benign-malignant lung nodule classification with geometric and appearance histogram features," *arXiv preprint arXiv:1605.08350*, 2016.
- [52] T. W. Way, L. M. Hadjiiski, B. Sahiner, H. P. Chan, P. N. Cascade, E. A. Kazerooni, *et al.*, "Computer-aided diagnosis of pulmonary nodules on CT scans:

- Segmentation and classification using 3D active contours," *Medical physics*, vol. 33, pp. 2323-2337, 2006.
- [53] P. Monkam, S. Qi, M. Xu, H. Li, F. Han, Y. Teng, *et al.*, "Ensemble learning of multiple-view 3D-CNNs model for micro-nodules identification in CT images," *IEEE Access*, vol. 7, pp. 5564-5576, 2018.
 - [54] Y. Chiou, Y. Lure, and P. Ligomenides, "Neural network image analysis and classification in hybrid lung nodule detection (HLND) system," in *Neural Networks for Signal Processing III-Proceedings of the 1993 IEEE-SP Workshop*, 1993, pp. 517-526.
 - [55] M. G. Penedo, M. J. Carreira, A. Mosquera, and D. Cabello, "Computer-aided diagnosis: a neural-network-based approach to lung nodule detection," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 872-880, 1998.
 - [56] S. Ashwin, J. Ramesh, S. A. Kumar, and K. Gunavathi, "Efficient and reliable lung nodule detection using a neural network based computer aided diagnosis system," in *2012 International Conference on Emerging Trends in Electrical Engineering and Energy Management (ICETEEEM)*, 2012, pp. 135-142.
 - [57] S. Gunasundari and S. Baskar, "Application of Artificial Neural Network in identification of lung diseases," in *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)*, 2009, pp. 1441-1444.
 - [58] E. Dandıl, M. Çakiroğlu, Z. Ekşi, M. Özkan, Ö. K. Kurt, and A. Canan, "Artificial neural network-based classification system for lung nodules on computed tomography scans," in *2014 6th International conference of soft computing and pattern recognition (SoCPaR)*, 2014, pp. 382-386.
 - [59] R. Arulmurugan and H. Anandakumar, "Early detection of lung cancer using wavelet feature descriptor and feed forward back propagation neural networks classifier," in *Computational vision and bio inspired computing*, ed: Springer, 2018, pp. 103-110.
 - [60] K. Fukushima and S. Miyake, "Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition," in *Competition and cooperation in neural nets*, ed: Springer, 1982, pp. 267-285.
 - [61] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, *et al.*, "Backpropagation applied to handwritten zip code recognition," *Neural computation*, vol. 1, pp. 541-551, 1989.
 - [62] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, pp. 2278-2324, 1998.
 - [63] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097-1105, 2012.
 - [64] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009, pp. 248-255.

- [65] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [66] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1-9.
- [67] W. Li, P. Cao, D. Zhao, and J. Wang, "Pulmonary nodule classification with deep convolutional neural networks on computed tomography images," *Computational and mathematical methods in medicine*, vol. 2016, 2016.
- [68] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818-2826.
- [69] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, *et al.*, "Microsoft coco: Common objects in context," in *European conference on computer vision*, 2014, pp. 740-755.
- [70] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, pp. 211-252, 2015.
- [71] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," 2009.
- [72] X. Wang, Y. Lu, Y. Wang, and W.-B. Chen, "Diabetic retinopathy stage classification using convolutional neural networks," in *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, 2018, pp. 465-471.
- [73] P. Teare, M. Fishman, O. Benzaquen, E. Toledano, and E. Elnekave, "Malignancy detection on mammography using dual deep convolutional neural networks and genetically discovered false color input enhancement," *Journal of digital imaging*, vol. 30, pp. 499-505, 2017.
- [74] J.-Z. Cheng, D. Ni, Y.-H. Chou, J. Qin, C.-M. Tiu, Y.-C. Chang, *et al.*, "Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans," *Scientific reports*, vol. 6, pp. 1-13, 2016.
- [75] W. Sun, B. Zheng, and W. Qian, "Computer aided lung cancer diagnosis with deep learning algorithms," in *Medical imaging 2016: computer-aided diagnosis*, 2016, p. 97850Z.
- [76] R. Majidpourkhoei, M. Alilou, K. Majidzadeh, and A. Babazadehsangar, "A novel deep learning framework for lung nodule detection in 3d CT images," *Multimedia Tools and Applications*, pp. 1-17, 2021.
- [77] G. Zhang, L. Lin, and J. Wang, "Lung Nodule Classification in CT Images Using 3D DenseNet," in *Journal of Physics: Conference Series*, 2021, p. 012155.
- [78] W. Zhu, C. Liu, W. Fan, and X. Xie, "Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018, pp. 673-681.

- [79] Y. Wang, H. Zhang, K. J. Chae, Y. Choi, G. Y. Jin, and S.-B. Ko, "Novel convolutional neural network architecture for improved pulmonary nodule classification on computed tomography," *Multidimensional Systems and Signal Processing*, vol. 31, pp. 1163-1183, 2020.
- [80] H. Cao, H. Liu, E. Song, G. Ma, X. Xu, R. Jin, *et al.*, "A two-stage convolutional neural networks for lung nodule detection," *IEEE journal of biomedical and health informatics*, vol. 24, pp. 2006-2015, 2020.
- [81] H. Greenspan, B. Van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE transactions on medical imaging*, vol. 35, pp. 1153-1159, 2016.
- [82] H. B. Nazar, M. Moetesum, S. Ehsan, I. Siddiqi, K. Khurshid, N. Vincent, *et al.*, "Classification of graphomotor impressions using convolutional neural networks: An application to automated neuro-psychological screening tests," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 2017, pp. 432-437.
- [83] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, *et al.*, "Convolutional neural networks for medical image analysis: Full training or fine tuning?," *IEEE transactions on medical imaging*, vol. 35, pp. 1299-1312, 2016.
- [84] M. Long, H. Zhu, J. Wang, and M. I. Jordan, "Deep transfer learning with joint adaptation networks," in *International conference on machine learning*, 2017, pp. 2208-2217.
- [85] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1717-1724.
- [86] H.-H. Yang, K.-C. Huang, and W.-T. Chen, "Laffnet: A lightweight adaptive feature fusion network for underwater image enhancement," *arXiv preprint arXiv:2105.01299*, 2021.
- [87] I. Ali, U. Javed, M. Amir, and I. Ulhaq, "Spatial Stimuli Gradient Based Multifocus Image Fusion Using Multiple Sized Kernels," *Tehnički vjesnik*, vol. 28, pp. 113-122, 2021.
- [88] S. Bhat and D. Koundal, "Multi-focus image fusion techniques: a survey," *Artificial Intelligence Review*, pp. 1-53, 2021.
- [89] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Information Fusion*, vol. 36, pp. 191-207, 2017.
- [90] H. Cui, X. Wang, J. Zhou, S. Eberl, Y. Yin, D. Feng, *et al.*, "Topology polymorphism graph for lung tumor segmentation in PET-CT images," *Physics in Medicine & Biology*, vol. 60, p. 4893, 2015.
- [91] M. Yin, X. Liu, Y. Liu, and X. Chen, "Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsamped shearlet transform domain," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, pp. 49-64, 2018.

- [92] Y. Na, L. Zhao, Y. Yang, and M. Ren, "Guided filter-based images fusion algorithm for CT and MRI medical images," *IET Image Processing*, vol. 12, pp. 138-148, 2018.
- [93] M. A. Khan, S. Rubab, A. Kashif, M. I. Sharif, N. Muhammad, J. H. Shah, *et al.*, "Lungs cancer classification from CT images: An integrated design of contrast based classical features fusion and selection," *Pattern Recognition Letters*, vol. 129, pp. 77-85, 2020.
- [94] P. Sridar, A. Kumar, A. Quinton, R. Nanan, J. Kim, and R. Krishnakumar, "Decision fusion-based fetal ultrasound image plane classification using convolutional neural networks," *Ultrasound in medicine & biology*, vol. 45, pp. 1259-1273, 2019.
- [95] 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," *Information Fusion*, vol. 42, pp. 102-110, 2018.
- [96] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-first AAAI conference on artificial intelligence*, 2017.
- [97] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251-1258.
- [98] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700-4708.
- [99] G. Huang, Z. Liu, G. Pleiss, L. Van Der Maaten, and K. Weinberger, "Convolutional networks with dense connectivity," *IEEE transactions on pattern analysis and machine intelligence*, 2019.
- [100] 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," *Medical & biological engineering & computing*, vol. 56, pp. 2125-2136, 2018.
- [101] G. Wei, H. Ma, W. Qian, F. Han, H. Jiang, S. Qi, *et al.*, "Lung nodule classification using local kernel regression models with out-of-sample extension," *Biomedical Signal Processing and Control*, vol. 40, pp. 1-9, 2018.
- [102] A. Shaffie, A. Soliman, H. A. Khalifeh, M. Ghazal, F. Taher, A. Elmaghraby, *et al.*, "Radiomic-based framework for early diagnosis of lung cancer," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019, pp. 1293-1297.
- [103] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," in *International conference on information processing in medical imaging*, 2015, pp. 588-599.

- [104] Y. LeCun, "The MNIST database of handwritten digits," <http://yann.lecun.com/exdb/mnist/>, 1998.
- [105] S. Shen, S. X. Han, D. R. Aberle, A. A. Bui, and W. Hsu, "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification," *Expert systems with applications*, vol. 128, pp. 84-95, 2019.
- [106] 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.*, pp. 10-27, 2011.
- [107] A. Reeves and A. Biancardi, "The Lung Image Database Consortium (LIDC) Nodule Size Report, Release: 2011-10-27," URL <http://www.via.cornell.edu/lidc>, 2011.
- [108] L. Deng, "The mnist database of handwritten digit images for machine learning research [best of the web]," *IEEE Signal Processing Magazine*, vol. 29, pp. 141-142, 2012.
- [109] C. Karthikeyan, B. Ramadoss, and S. Baskar, "Segmentation algorithm for CT images using morphological operation and artificial neural network," *international journal of signal processing, image processing and pattern recognition*, vol. 5, pp. 115-122, 2012.
- [110] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International Conference on Machine Learning*, 2015, pp. 448-456.
- [111] 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," *Computers in Industry*, vol. 106, pp. 1-13, 2019.
- [112] S. Pang, Z. Yu, and M. A. Orgun, "A novel end-to-end classifier using domain transferred deep convolutional neural networks for biomedical images," *Computer methods and programs in biomedicine*, vol. 140, pp. 283-293, 2017.
- [113] Y. Xie, Y. Xia, J. Zhang, Y. Song, D. Feng, M. Fulham, *et al.*, "Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT," *IEEE transactions on medical imaging*, vol. 38, pp. 991-1004, 2018.
- [114] Y. Xie, J. Zhang, and Y. Xia, "Semi-supervised adversarial model for benign-malignant lung nodule classification on chest CT," *Medical image analysis*, vol. 57, pp. 237-248, 2019.
- [115] M. Al-Shabi, H. K. Lee, and M. Tan, "Gated-dilated networks for lung nodule classification in CT scans," *IEEE Access*, vol. 7, pp. 178827-178838, 2019.
- [116] 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," *Journal of healthcare engineering*, vol. 2019, 2019.
- [117] S. S. Ghosal, I. Sarkar, and I. El Hallaoui, "Lung Nodule Classification Using Convolutional Autoencoder and Clustering Augmented Learning Method (CALM)," in *HSDM@ WSDM*, 2020, pp. 19-26.

- [118] M. Al-Shabi, B. L. Lan, W. Y. Chan, K.-H. Ng, and M. Tan, "Lung nodule classification using deep local-global networks," *International journal of computer assisted radiology and surgery*, vol. 14, pp. 1815-1819, 2019.
- [119] 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 *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2017, pp. 656-664.
- [120] M. Nishio, M. Nishizawa, O. Sugiyama, R. Kojima, M. Yakami, T. Kuroda, *et al.*, "Computer-aided diagnosis of lung nodule using gradient tree boosting and Bayesian optimization," *PloS one*, vol. 13, p. e0195875, 2018.
- [121] M. Nishio and C. Nagashima, "Computer-aided diagnosis for lung cancer: usefulness of nodule heterogeneity," *Academic radiology*, vol. 24, pp. 328-336, 2017.
- [122] 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," *Information Fusion*, vol. 62, pp. 73-80, 2020.
- [123] M. B. Skouson, B. J. Borghetti, and R. C. Leishman, "Ursa: A neural network for unordered point clouds using constellations," in *Science and Information Conference*, 2019, pp. 447-457.
- [124] M. Simonovsky and N. Komodakis, "Dynamic edge-conditioned filters in convolutional neural networks on graphs," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 3693-3702.
- [125] R. Klokov and V. Lempitsky, "Escape from cells: Deep kd-networks for the recognition of 3d point cloud models," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 863-872.
- [126] D. Grover and B. Toghi, "MNIST dataset classification utilizing k-NN classifier with modified sliding-window metric," in *Science and Information Conference*, 2019, pp. 583-591.
- [127] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," *arXiv preprint arXiv:1706.02413*, 2017.
- [128] Y. Ren, L. Zhang, and P. N. Suganthan, "Ensemble classification and regression-recent developments, applications and future directions," *IEEE Computational intelligence magazine*, vol. 11, pp. 41-53, 2016.
- [129] T. G. Dietterich, "Ensemble methods in machine learning," in *International workshop on multiple classifier systems*, 2000, pp. 1-15.
- [130] N. Altman and M. Krzywinski, "Ensemble methods: bagging and random forests," *Nature Methods*, vol. 14, pp. 933-935, 2017.
- [131] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, "Lung pattern classification for interstitial lung diseases using a deep convolutional

- neural network," *IEEE transactions on medical imaging*, vol. 35, pp. 1207-1216, 2016.
- [132] G. K. Abraham, P. Bhaskaran, and V. Jayanthi, "Lung nodule classification in CT images using convolutional neural network," in *2019 9th International Conference on Advances in Computing and Communication (ICACC)*, 2019, pp. 199-203.
 - [133] I. Ali, M. Muzammil, I. U. Haq, A. A. Khaliq, and S. Abdullah, "Efficient lung nodule classification using transferable texture convolutional neural network," *Ieee Access*, vol. 8, pp. 175859-175870, 2020.
 - [134] D. Ciregan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *2012 IEEE conference on computer vision and pattern recognition*, 2012, pp. 3642-3649.
 - [135] S. P. Kannoja and G. Jaiswal, "Ensemble of hybrid CNN-ELM model for image classification," in *2018 5th international conference on signal processing and integrated networks (SPIN)*, 2018, pp. 538-541.
 - [136] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, p. e1249, 2018.
 - [137] A. A. A. Setio, A. Traverso, T. De Bel, M. S. Berens, C. Van Den Bogaard, P. Cerello, *et al.*, "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge," *Medical image analysis*, vol. 42, pp. 1-13, 2017.
 - [138] S. G. Armato III, G. McLennan, L. Bidaut, M. F. McNitt-Gray, C. R. Meyer, A. P. Reeves, *et al.*, "The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans," *Medical physics*, vol. 38, pp. 915-931, 2011.
 - [139] W. Zuo, F. Zhou, and Y. He, "An Embedded Multi-branch 3D Convolution Neural Network for False Positive Reduction in Lung Nodule Detection," *Journal of digital imaging*, vol. 33, 2020.
 - [140] K. Zuiderveld, "Contrast limited adaptive histogram equalization," *Graphics gems*, pp. 474-485, 1994.
 - [141] N. Khosravan and U. Bagci, "S4ND: Single-shot single-scale lung nodule detection," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018, pp. 794-802.
 - [142] R. Gruetzemacher, A. Gupta, and D. Paradice, "3D deep learning for detecting pulmonary nodules in CT scans," *Journal of the American Medical Informatics Association*, vol. 25, pp. 1301-1310, 2018.
 - [143] F. V. Farahani, A. Ahmadi, and M. F. Zarandi, "Lung nodule diagnosis from CT images based on ensemble learning," in *2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2015, pp. 1-7.
 - [144] H. A. Ahmed and S. A. Mahmood, "A deep learning technique for lung nodule classification based on false positive reduction," *Journal of Zankoy Sulaimani*, pp. 107-116.

- [145] Q. Dou, H. Chen, L. Yu, J. Qin, and P.-A. Heng, "Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection," *IEEE Transactions on Biomedical Engineering*, vol. 64, pp. 1558-1567, 2016.
- [146] P. Zhai, Y. Tao, H. Chen, T. Cai, and J. Li, "Multi-Task Learning for Lung Nodule Classification on Chest CT," *IEEE Access*, vol. 8, pp. 180317-180327, 2020.
- [147] A. Naik, D. R. Edla, and V. Kuppili, "Lung Nodule Classification on Computed Tomography Images Using Fractalnet," *Wireless Personal Communications*, pp. 1-21, 2021.
- [148] A. Mobiny, P. Yuan, P. A. Cicalese, S. K. Moulik, N. Garg, C. C. Wu, *et al.*, "Memory-augmented capsule network for adaptable lung nodule classification," *IEEE Transactions on Medical Imaging*, 2021.
- [149] N. Sun, D. Yang, S. Fang, and H. Xie, "Deep convolutional nets for pulmonary nodule detection and classification," in *International Conference on Knowledge Science, Engineering and Management*, 2018, pp. 197-208.
- [150] W. Peng, R. V. Mayorga, and E. M. Hussein, "An automated confirmatory system for analysis of mammograms," *Computer methods and programs in biomedicine*, vol. 125, pp. 134-144, 2016.
- [151] K. Z. Htwe, K. Yamamori, T. Katayama, and T. M. Kyi, "Automated lung nodule classification by artificial neural network and fuzzy inference system," in *2016 IEEE 5th Global Conference on Consumer Electronics*, 2016, pp. 1-2.
- [152] H. Jiang, H. Ma, W. Qian, G. Wei, X. Zhao, and M. Gao, "A novel pixel value space statistics map of the pulmonary nodule for classification in computerized tomography images," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 556-559.
- [153] C.-H. Chen, C.-K. Chang, C.-Y. Tu, W.-C. Liao, B.-R. Wu, K.-T. Chou, *et al.*, "Radiomic features analysis in computed tomography images of lung nodule classification," *PloS one*, vol. 13, p. e0192002, 2018.
- [154] X.-X. Li, B. Li, L.-F. Tian, and L. Zhang, "Automatic benign and malignant classification of pulmonary nodules in thoracic computed tomography based on RF algorithm," *IET Image Processing*, vol. 12, pp. 1253-1264, 2018.
- [155] J. Song, H. Liu, F. Geng, and C. Zhang, "Weakly-supervised classification of pulmonary nodules based on shape characters," in *2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*, 2016, pp. 228-232.
- [156] S. Chen, J. Qin, X. Ji, B. Lei, T. Wang, D. Ni, *et al.*, "Automatic scoring of multiple semantic attributes with multi-task feature leverage: A study on pulmonary nodules in CT images," *IEEE transactions on medical imaging*, vol. 36, pp. 802-814, 2016.
- [157] V. Le, D. Yang, Y. Zhu, B. Zheng, C. Bai, Q. Nguyen, *et al.*, "Automated classification of pulmonary nodules for lung adenocarcinomas risk evaluation: an

- effective CT analysis by clustering density distribution algorithm," *Journal of Medical Imaging and Health Informatics*, vol. 7, pp. 1753-1758, 2017.
- [158] G. Wei, H. Ma, W. Qian, and S. Qi, "A content-based image retrieval scheme for lung nodule classification," *Current Medical Imaging*, vol. 13, pp. 210-216, 2017.
 - [159] J. Gong, J.-Y. Liu, X.-W. Sun, B. Zheng, and S.-D. Nie, "Computer-aided diagnosis of lung cancer: the effect of training data sets on classification accuracy of lung nodules," *Physics in Medicine & Biology*, vol. 63, p. 035036, 2018.
 - [160] G. Wei, H. Cao, H. Ma, S. Qi, W. Qian, and Z. Ma, "Content-based image retrieval for lung nodule classification using texture features and learned distance metric," *Journal of medical systems*, vol. 42, pp. 1-7, 2018.
 - [161] M. Keming and D. Zhuofu, "Lung nodule image classification based on ensemble machine learning," *Journal of Medical Imaging and Health Informatics*, vol. 6, pp. 1679-1685, 2016.
 - [162] F. Zhang, Y. Song, W. Cai, M.-Z. Lee, Y. Zhou, H. Huang, *et al.*, "Lung nodule classification with multilevel patch-based context analysis," *IEEE Transactions on Biomedical Engineering*, vol. 61, pp. 1155-1166, 2014.
 - [163] J. Yuan, X. Liu, F. Hou, H. Qin, and A. Hao, "Hybrid-feature-guided lung nodule type classification on CT images," *Computers & Graphics*, vol. 70, pp. 288-299, 2018.
 - [164] A. Farag, A. Ali, J. Graham, S. Elhabian, A. Farag, and R. Falk, "Feature-based lung nodule classification," *Advances in Visual Computing*, pp. 79-88, 2010.
 - [165] K. Mao and Z. Deng, "Lung nodule image classification based on local difference pattern and combined classifier," *Computational and mathematical methods in medicine*, vol. 2016, 2016.
 - [166] N. Gupta, D. Gupta, A. Khanna, P. P. Rebouças Filho, and V. H. C. de Albuquerque, "Evolutionary algorithms for automatic lung disease detection," *Measurement*, vol. 140, pp. 590-608, 2019.
 - [167] S. Zhang, F. Sun, N. Wang, C. Zhang, Q. Yu, M. Zhang, *et al.*, "Computer-aided diagnosis (CAD) of pulmonary nodule of thoracic CT image using transfer learning," *Journal of digital imaging*, vol. 32, pp. 995-1007, 2019.
 - [168] S. Rajaraman, S. K. Antani, M. Poostchi, K. Silamut, M. A. Hossain, R. J. Maude, *et al.*, "Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images," *PeerJ*, vol. 6, p. e4568, 2018.
 - [169] R. J. S. Raj, S. J. Shobana, I. V. Pustokhina, D. A. Pustokhin, D. Gupta, and K. Shankar, "Optimal feature selection-based medical image classification using deep learning model in internet of medical things," *IEEE Access*, vol. 8, pp. 58006-58017, 2020.
 - [170] S. G. Armato III, L. Hadjiiski, G. D. Tourassi, K. Drukker, M. L. Giger, F. Li, *et al.*, "SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset. The Cancer Imaging Archive," ed, 2015.

- [171] S. G. Armato III, L. Hadjiiski, G. D. Tourassi, K. Drukker, M. L. Giger, F. Li, *et al.*, "LUNGx Challenge for computerized lung nodule classification: reflections and lessons learned," *Journal of Medical Imaging*, vol. 2, 2015.

