

MULTI-MODAL MEDICAL IMAGE ANALYSIS USING DEEP LEARNING-BASED TECHNIQUES

SAKSHI AHUJA



**ELECTRICAL ENGINEERING DEPARTMENT
INDIAN INSTITUTE OF TECHNOLOGY DELHI
JULY 2023**

©Indian Institute of Technology Delhi (IITD), New Delhi, 2024

MULTI-MODAL MEDICAL IMAGE ANALYSIS USING DEEP LEARNING-BASED TECHNIQUES

by

SAKSHI AHUJA

Electrical Engineering Department

Submitted

in fulfillment of the requirements of the degree of Doctor of Philosophy

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI

JULY 2023

Dedicated to my husband
Utkarsh Gautam

Certificate

This is to certify that the thesis entitled "**Multi-modal medical image analysis using deep learning-based techniques**" being submitted by **Mrs. Sakshi Ahuja** to the **Electrical Engineering department, Indian Institute of Technology Delhi**, for the award of the degree of **Doctor of Philosophy** is the record of the bonafide research work carried out by her under our supervision. In our opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted either in part or in full to any other university or institute for the award of any degree or diploma.

(Prof. Bijaya Ketan Panigrahi)

Department of Electrical Engineering

Indian Institute of Technology Delhi

Hauz Khas, New Delhi 110016, India

(Prof. Tapan Kumar Gandhi)

Department of Electrical Engineering

Indian Institute of Technology Delhi

Hauz Khas, New Delhi 110016, India

Date:

Place:

Acknowledgments

First and foremost, I would like to express my profound gratitude to my supervisors **Prof. Bijaya Ketan Panigrahi** and **Prof. Tapan Kumar Gandhi** for their kind encouragement and precious guidance throughout the Ph.D. program. I am indebted to my supervisors for tirelessly playing the role of family elders and providing me with much-needed support during these years. I am truly blessed to have such experienced and supporting supervisors. Their valuable support during COVID times and arranging the necessary resources made me work at home with ease during that tenure as well.

I would like to thank my SRC committee members: Prof. Sumantra Dutta Roy, Prof. Breejesh Lall, and Prof. Anup Singh for their insight and valuable feedback during my research. I am grateful to Dr. Nilanjan Dey, who has provided me with constant motivation to excel in the medical field of research. Sincere thanks to all my labmates and friends especially Shilpi Gupta, Anushikha Singh, and Shefali Gupta for their moral support.

I would like to express my deep sense of gratitude to my grandparents (Late Mrs. Prakash Chhabra, and Late Smt. Mohan Devi Ahuja) who passed away during this Ph.D. journey. I am grateful to my grandfather Dr. Manmohan Chhabra. My grandparents always pushed my limits and supported me in my studies throughout the years.

I would like to thank my husband Utkarsh Gautam, who always believed in me and provided unconditional support. He arranged all the necessary resources for work from home scenario. I would like to mention one special name, my dearest daughter Utrakshi

who is my greatest strength. I would like to thank my parents Om Parkash Ahuja and Hiramani Ahuja, brother Shashwat Ahuja, and sister-in-law Shikha, for their unconditional love, prayer, support, and encouragement at every step of my life. I would like to thank my in-laws for their continuous prayers and kind words.

I am blessed beyond words to God for providing me with this precious opportunity and mercy.

Sakshi Ahuja

Abstract

In the medical field, timely and accurate diagnosis of disease using multi-modal radiographic images is a very challenging task. The major issues associated with the medical field are a) limited availability of the annotated dataset, b) class imbalance, c) data overfitting issues, and d) under-segmentation and over-segmentation of the lesions. To address these issues, transfer learning-based deep learning networks are implemented by the researchers. The transfer learning models have the following advantages a) optimum pre-processing pipeline, b) rapid learning, c) time computation could be adjusted by reducing the layers, and d) requirement of limited training dataset.

Initially, brain MRI scans are examined to diagnose the brain tumor using transfer learning. To address the data overfitting issue associated with the medical field, data augmentation techniques are used on multi-modality radiographic images. In the proposed thesis, wavelet decomposition up to N level along with statistical operation (rotation, translation, shear, etc.) is performed. The performance of the numerous pre-trained deep learning models (VGGs, ResNets, MobileNet, InceptionResNetv2, SqueezeNet, etc.) is compared with augmentation and without augmentation. Further, the impact of various cross-validation techniques, post-processing operations, and optimizers (sgdm, RmsProp, adam) on multi-modality radiographic images is evaluated.

To evaluate the robustness of the proposed methodology, a novel COVID-19 infection is diagnosed using chest radiographic images. The medical field suffers from

class imbalance issues, i.e., unequal samples are available in each class. The class imbalance issue is majorly encountered with the severity of COVID-19 infection from chest CT scans. To resolve this issue, Siamese-based networks (P-shot, M-ways approach) with the pre-trained network as the base encoder is implemented. The proposed Siamese-based networks are evaluated on the Indian patients as well.

The proposed classification approach uses deep learning-based techniques for the classification of a variety of diseases, i.e., brain tumor grades (LGG, HGG), and brain tumor types (Glioma, Meningioma, and Pituitary). To segment out the tumor from brain MRI scans, color map-based superpixel technique, and DeepLabv3-based technique is used. To extract the COVID-19 infection patches from the lungs, the novel COVSeg-Net framework is used. Further, the UNet model performance is compared with the COVSeg-Net model for lesion segmentation.

The pre-trained network used in the diagnosis of the disease is considered a black box for clinicians. Clinicians are interested in the analysis of the reason behind the prediction made by deep learning models. Thus, the XAI approach is used to evaluate the role of numerous features in disease localization. The state-of-the-art XAI techniques that are used in the proposed work are LIME, Grad-CAM, and GradCAM++.

In summary, this study evaluates the impact of deep learning-based networks on disease classification, localization, and segmentation from brain MRI and chest CT scans. By analyzing the decision-making process, the XAI tool helps clinicians comprehend the reasoning behind the deep learning model predictions.

सार

चिकित्सा क्षेत्र में मल्टी-मॉडल रेडियोग्राफिक छवियों का उपयोग करके रोग का समय पर और सटीक निदान करना एक बहुत ही चुनौतीपूर्ण कार्य है। चिकित्सा क्षेत्र से जुड़े प्रमुख मुद्दे हैं ए) एनोटेटेड डेटासेट की सीमित उपलब्धता, बी) वर्ग असंतुलन, सी) डेटा ओवरफिटिंग मुद्दे, और डी) घावों का कम-विभाजन और अधिक-विभाजन। इन मुद्दों को संबोधित करने के लिए, शोधकर्ताओं द्वारा ट्रांसफर लर्निंग-आधारित डीप लर्निंग नेटवर्क लागू किए जाते हैं। ट्रांसफर लर्निंग मॉडल के निम्नलिखित फायदे हैं ए) इष्टतम प्री-प्रोसेसिंग पाइपलाइन, बी) तेजी से सीखना, सी) परतों को कम करके समय गणना को समायोजित किया जा सकता है, और डी) सीमित प्रशिक्षण डेटासेट की आवश्यकता।

प्रारंभ में, ट्रांसफर लर्निंग का उपयोग करके ब्रेन ट्यूमर का निदान करने के लिए ब्रेन एमआरआई स्कैन की जांच की जाती है। चिकित्सा क्षेत्र से जुड़े डेटा ओवरफिटिंग मुद्दे को संबोधित करने के लिए, मल्टी-मोडेलिटी रेडियोग्राफिक छवियों पर डेटा संवर्धन तकनीकों का उपयोग किया जाता है। प्रस्तावित थीसिस में, सांख्यिकीय संचालन (रोटेशन, अनुवाद, कतरनी, आदि) के साथ एन स्तर तक तरंगिका अपघटन किया जाता है। कई पूर्व-प्रशिक्षित गहन शिक्षण मॉडल (VGGs, ResNets, MobileNet, InceptionResNetv2, SqueezeNet, आदि) के प्रदर्शन की तुलना संवर्धन और बिना संवर्धन के की जाती है। इसके अलावा, मल्टी-मॉडेलिटी रेडियोग्राफिक छवियों पर विभिन्न क्रॉस-वैलिडेशन तकनीकों, पोस्ट-प्रोसेसिंग ऑपरेशंस और ऑप्टिमाइज़र (एसजीडीएम, आरएमएसप्रॉप, एडम) के प्रभाव का मूल्यांकन किया जाता है।

प्रस्तावित पद्धति की मजबूती का मूल्यांकन करने के लिए, छाती की रेडियोग्राफिक छवियों का उपयोग करके एक नए COVID-19 संक्रमण का निदान किया जाता है। चिकित्सा क्षेत्र वर्ग असंतुलन के मुद्दों से ग्रस्त है, अर्थात्, प्रत्येक वर्ग में असमान नमूने उपलब्ध हैं। छाती सीटी स्कैन से सीओवीआईडी -19 संक्रमण की गंभीरता के साथ वर्ग असंतुलन का मुद्दा प्रमुख रूप से सामने आया है। इस समस्या को हल करने के लिए, आधार एनकोडर के रूप में पूर्व-

प्रशिक्षित नेटवर्क के साथ सियामी-आधारित नेटवर्क (पी-शॉट, एम-वेज़ दृष्टिकोण) लागू किया गया है। प्रस्तावित सियामीज़-आधारित नेटवर्क का मूल्यांकन भारतीय रोगियों पर भी किया जाता है।

प्रस्तावित वर्गीकरण दृष्टिकोण विभिन्न प्रकार की बीमारियों, यानी ब्रेन ट्यूमर ग्रेड (एलजीजी, एचजीजी), ब्रेन ट्यूमर प्रकार (ग्लियोमा, मेनिंगियोमा और पिट्यूटरी) के वर्गीकरण के लिए गहन शिक्षण-आधारित तकनीकों का उपयोग करता है। मस्तिष्क एमआरआई स्कैन से ट्यूमर को अलग करने के लिए, कलर मैप-आधारित सुपरपिक्सेल तकनीक और डीपलैबवी3-आधारित तकनीक का उपयोग किया जाता है। फेफड़ों से COVID-19 संक्रमण पैच निकालने के लिए, नवीन COVSeg-Net फ्रेमवर्क का उपयोग किया जाता है। इसके अलावा, यूनेट मॉडल के प्रदर्शन की तुलना घाव विभाजन के लिए COVSeg-Net मॉडल से की जाती है।

रोग के निदान में प्रयुक्त पूर्व-प्रशिक्षित नेटवर्क को चिकित्सकों के लिए ब्लैक बॉक्स माना जाता है। चिकित्सक गहन शिक्षण मॉडल द्वारा की गई भविष्यवाणी के पीछे के कारण के विश्लेषण में रुचि रखते हैं। इस प्रकार, XAI दृष्टिकोण का उपयोग रोग स्थानीयकरण में कई विशेषताओं की भूमिका का मूल्यांकन करने के लिए किया जाता है। प्रस्तावित कार्य में उपयोग की जाने वाली अत्याधुनिक XAI तकनीकें LIME, Grad-CAM और GradCAM++ हैं।

संक्षेप में, यह अध्ययन मस्तिष्क एमआरआई और छाती सीटी स्कैन से रोग वर्गीकरण, स्थानीयकरण और विभाजन पर गहन शिक्षण-आधारित नेटवर्क के प्रभाव का मूल्यांकन करता है। निर्णय लेने की प्रक्रिया का विश्लेषण करके, XAI उपकरण चिकित्सकों को गहन शिक्षण मॉडल भविष्यवाणियों के पीछे के तर्क को समझने में मदद करता है।

साक्षी आहूजा

Table of Contents

Certificate	i
Acknowledgments	ii
Abstract	iv
List of Figures	xiii
List of Tables	xvi
Abbreviations	xviii
1 Introduction	1
1.1 Pre-requisite	1
1.2 Motivation	2
1.3 Disease diagnosis	4
1.3.1 Brain tumor diagnosis using MRI scans.....	4
1.3.2 COVID-19 diagnosis.....	6
1.4 Literature review	6
1.4.1 Brain tumor diagnosis via deep learning.....	7
1.4.2 COVID-19 diagnosis.....	11

1.5	Scope of work.....	12
1.6	Thesis outline	13
2	Transfer learning	15
2.1	Introduction	15
2.2	Proposed methodology	16
2.3	Experimentation	20
2.3.1	Dataset details.....	21
2.3.2	Performance evaluation parameters.....	21
2.3.3	Limitations.....	24
2.4	Conclusions.....	25
3	Impact of data augmentation	27
3.1	Introduction	27
3.2	Proposed methodology.....	27
3.2.1	Tumor classification.....	28
3.2.2	Tumor localization.....	33
3.2.3	Tumor segmentation.....	34
3.3	Experimentation.....	35
3.3.1	Performance evaluation parameters.....	35
3.3.2	Tumor localization using feature maps.....	37
3.3.3	Tumor segmentation using superpixel.....	37
3.4	Ablation study.....	39
3.4.1	Performance of pre-trained networks.....	39

3.4.2	Impact of augmentation techniques.....	40
3.4.3	Performance on the BraTS dataset.....	40
3.4.4	CAD tool.....	41
3.4.5	Limitations.....	43
3.5	Conclusions.....	44
4	COVID-19 infection diagnosis	45
4.1	Introduction.....	45
4.2	Proposed methodology	46
4.2.1	Dataset details.....	47
4.2.2	Pre-processing.....	47
4.2.3	Data augmentation.....	48
4.2.4	Transfer learning.....	50
4.2.5	Grad-CAM localization.....	52
4.3	Experimentation.....	52
4.3.1	Limitations	55
4.4	Conclusions.....	55
5	McS-Net for COVID-19 infection classification	56
5.1	Introduction.....	56
5.2	Proposed methodology	57
5.2.1	Dataset	57
5.2.2	Pre-processing.....	58
5.2.3	Pre-trained deep learning network.....	59

5.2.4	Nearest neighbor classification.....	61
5.3	Results and discussion.....	62
5.3.1	Performance evaluation parameters.....	62
5.3.2	Grad-CAM based visualization.....	64
5.3.3	Indian patients' diagnosis.....	65
5.4	Ablation study and limitations.....	66
5.5	Conclusion.....	67
6	Segmentation of abnormality	69
6.1	Introduction	69
6.2	DeepLabv3+ for tumor segmentation	69
6.2.1	Dataset.....	71
6.2.2	Pre-processing.....	71
6.2.3	Data augmentation	71
6.2.4	DeepLabv3+ segmentation.....	72
6.3	COVID-19 infection segmentation.....	76
6.4	Limitations	77
6.5	Conclusions.....	78
7	Role of XAI in medical image analysis	79
7.1	Introduction	79
7.2	XAI for COVID-19 diagnosis	80
7.2.1	Stage-1: Lungs segmentation.....	81
7.2.2	Stage-2: Transfer learning.....	84

7.2.3	Stage-3: Infection localization	85
7.2.4	Stage-4: Post-processing.....	86
7.3	Ablation study.....	88
7.3.1	Limitations.....	89
7.4	Conclusions	90
8	Conclusion and future scope	91
8.1	Conclusion.....	91
8.2	Future scope	93
	Bibliography	94
	Work published based on this thesis	105
	Technical biography of the author	107

List of Figures

2.1	Schematic hierarchy of the deep learning-based models.....	16
2.2	Block diagram of the two-phase brain tumor diagnosis.....	17
2.3	Convergence graph of transfer learning based VGG-19 pre-trained network	22
2.4	First layer activation of VGG-19 on the considered BraTS 2019 brain MRI dataset.....	23
2.5	CAM on the considered BraTS 2019 brain MRI database.....	24
2.6	Pictorial representation of the colormap-based superpixel segmentation technique.....	25
3.1	Self-explanatory representation of the three-phase brain tumor diagnosis model on the figshare dataset.....	28
3.2	Brief details of three types of tumors in multi-dimensional views from Dataset B.....	29
3.3	ROC (FP vs. TP) of the DarkNets with and without augmentation.....	36
3.4	Convergence graph for best performing pre-trained DarkNet-53 model.....	36
3.5	Tumor localization using deeper layer ‘res23’ of DarkNet-53 pre-trained model.....	38
3.6	Multi-class brain tumor segmentation from the multi-orientation views using colormap-based superpixel technique.....	38
3.7	Self-explanatory block diagram of the CAD tool for brain tumor diagnosis.....	42
3.8	Deep learning-based CAD tool for brain tumor diagnosis.....	43
4.1	Self-explanatory block diagram for the 3-phase COVID-19 diagnosis.....	46
4.2	Chest CT scans of a) COVID-19 and b) non-COVID patients.....	47
4.3	Pictorial representation of the data augmentation using stationary wavelet decomposition for $t=3$	49

4.4	Sample training data after combined statistical augmentation on CT scan images.....	49
4.5	Self-explanatory block diagram for the proposed ResNet-18.....	51
4.6	Skip connection scenario in the ResNet18 network.....	51
4.7	Convergence graph for accuracy vs iterations and loss vs. iterations for the finest performing ResNet18 model.....	53
4.8	ROC curve for the pre-trained ResNet18 model.....	53
4.9	Performance evaluation using Glyph plot.....	54
5.1	The pictorial representation of the proposed McS-Net for the severity of COVID-19 infection classification.....	57
5.2	Dataset description of the multi-class classification of severity of COVID-19 infection from chest CT scan.....	58
5.3	Brief description of the ResNet-18 architecture used as the base encoder network.....	60
5.4	Schematic representation of COVID-19 infection severity level classification using McS-Net.....	61
5.5	Visualization of the COVID-19 infection in multiple classes using Grad-CAM.....	65
5.6	COVID-19 infection visualization using grad-CAM on the Indian dataset.....	65
6.1	Detailed hierarchy of the proposed methodology for brain tumor segmentation using semantic segmentation.....	70
6.2	Data augmentation via geometrical operation implemented on the pre-processed FLAIR MRI slice.	72
6.3	The layer-wise architecture of DeepLabv3+ with variable loss-based brain tumor segmentation.....	73
6.4	Pictorial representation of the proposed methodology (a) FLAIR MRI with gold standard mask in blue color, (b) predicted label using MobileNetv2+DL, (c) predicted label using MobileNetv2+TL, (d) predicted label using MobileNetv2+CoL (e) predicted label using ResNet18+DL (F) predicted label using ResNet18+TL, (g) predicted label using ResNet18+CoL.....	74

6.5	Comparative analysis of the proposed deeplabv3+-based segmentation approach with different base encoders	75
6.6	Flow diagram for steps-wise implementation of the lung patches segmentation from the CT scan image.....	76
7.1	Flow chart of the 4-stage COVID-19 diagnosis using chest CT scans.....	81
7.2	Sample CT scans after preprocessing a) original CT scan b) after histogram equalization, c) cropping operation.....	82
7.3	The detailed architecture of the UNet model	83
7.4	Sample chest CT scans a) original chest CT scans, and b) segmented lungs using chest CT scans.....	84
7.5	Channel activation maps obtained using VGG-16 pre-trained model using XAI Grad-CAM, Grad-CAM++, and LIME.....	87
7.6	Channel activation maps obtained using ResNet50 pre-trained model using XAI Grad-CAM, Grad-CAM++, and LIME.....	87
7.7	GradCAM++ implementation on the Figshare dataset using pre-trained MobileNet.....	89

List of Tables

2.1	Layers and activations details of pre-trained VGG-19 model.....	19
3.1	Hyper-parameters settings for pre-trained DarkNets model.....	32
3.2	Comparative analysis of the DarkNet model on the testing set (N-Normal, M-Meningioma, G-Glioma, P-Pituitary).....	37
3.3	Performance comparison of different pre-trained networks on the T1W-CE MRI dataset.....	39
3.4	Performance comparison of optimization functions on hybrid dataset.....	39
3.5	Evaluation of the data augmentation technique on the proposed methodology for brain tumor classification on the hybrid dataset.....	40
3.6	Performance evaluation parameters obtained for multi-modality MRI using the proposed methodology for binary classification of tumor grades.....	41
4.1	Comparative analysis of the pre-trained networks implemented on the considered dataset of chest CT scans.....	54
5.1	Global performance evaluation parameters for 15-shot 5-ways simases based nets with 5 random trials conducted with the ResNet18 base encoder.....	63
5.2	Performance evaluation parameters for the variation in severity of the COVID-19 infection on the MosMed dataset (base encoder ResNet18).....	64
5.3	Performance comparison of the 5-ways 2-shot McS-Net with different pre-trained	

encoder networks.....	66
5.4 Comparative analysis of the 5-ways M-shot McS-Net with ResNet18-based encoder	67
6.1 Performance evaluation parameters for the COV-Seg Net without augmentation..	77
6.2 Performance evaluation parameters obtained by COVSeg-NET with augmentation techniques.....	77
7.1 Comparative analysis of the segmentation model used in the proposed methodology	83
7.2 Performance evaluation parameters of the transfer learning-based model.....	84
7.3 Comparative analysis of the proposed methodology on state-of-the-art XAI techniques.....	88
7.4 Quantitative evaluation of the performance of numerous deep learning-based techniques with and without augmentation.....	89

Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
CT	Computed Tomography
DWT	Discrete Wavelet Transform
DL	Deep Learning
ET	Enhancing Tumor
fMRI	Functional Magnetic Resonance Imaging
IoU	Intersection of Union
McS-Net	Multi-class Siamese Network
MRI	Magnetic Resonance Imaging
PCR	Polymerase Chain Reaction
TC	Tumor Core
WT	Whole Tumor