

الجمهورية الجزائرية الديمقراطية الشعبية
People's Democratic Republic of Algeria
وزارة التعليم العالي والبحث العلمي
Ministry of Higher Education and Scientific Research

MOHAMED KHIDER UNIVERSITY – Biskra
FACULTY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF ELECTRICAL
ENGINEERING
Ref:.....



جامعة محمد خيضر
ببسكرة كلية العلوم والتكنولوجيا
.....
المرجع:.....

Thesis submitted for obtaining the
degree of doctorate in electrical
engineering.

Option: Electronic.

Deep Segmentation for early diagnosis in Medical Imaging

Presented by

Meryem KETFI.

Thesis defended publicly on 06/02/2025

The jury members:

Pr. AbdErrazak DEBILOU
Pr. Mebarka BELAHCENE
Dr. Youcef BRIK
Pr. AbdElMalik OUAMANE

Professor
Professor
MCA
Professor

President
Supervisor
Examinator
Examinator

University of Biskra
University of Biskra
University of Msila
University of Biskra

ACKNOWLEDGEMENTS

First and foremost, I express my gratitude to Allah for guiding me throughout my research journey and facilitating the completion of my thesis.

My sincere appreciation extends to Professor **Mebarka BELAHCENE**, my esteemed supervisor, whose unwavering support, insightful suggestions, and motivation have been invaluable throughout my research and thesis development.

I am thankful to Professor **AbdErrazak DEBILOU** for presiding over my Thesis Committee.

I extend my heartfelt thanks to all members of the jury. Special acknowledgment is due to Professor **Youcef BRIK** and **AbdeIMalik OUAMANE** for graciously agreeing to assess my thesis, offering valuable advice, insightful comments, and posing thoughtful questions.

In a broader sense, I wish to express my gratitude to everyone who has provided encouragement and support, both technically and personally, during these intense years of dedicated work.

DEDICATION

I dedicate this modest work to:

My beloved parents.

My husband: This thesis is a culmination of your love, encouragement, and unwavering belief in me. Your support has been my greatest strength, propelling me through challenges and inspiring me to achieve my goals. With heartfelt gratitude, I dedicate this work to you.

Sisters and brothers.

My children.

All who supported me.

Publications & Scientific Communications-----

A. International Publications

[1] M. Ketfi, M. Belahcene and S. Bourennane, Transfer Learning Fusion and Stacked Auto-Encoders for Viral Lung Disease Classification, New Generation Computing. (2024). <https://doi.org/10.1007/s00354-024-00247-4>

[2] Meryem, KETFI, Mebarka BELAHCENE, and Salah BOURENNANE. "OCAE and OUNET: Standard automatic optimization for medical image segmentation." Multimedia Tools and Applications (2024): 1-34. <https://link.springer.com/article/10.1007/s11042-024-20287-8>

B. International Communications

[1] M. Ketfi, M. Belahcene and S. Bourennane, Efficient Diagnosis COVID-19 using Gabor and Transfer Learning, 2021 1st International Conference On Cyber Management And Engineering (CyMaEn) (pp. 1-6). IEEE, Rouen, France.

doi: [10.1109/CyMaEn50288.2021.9497304](https://doi.org/10.1109/CyMaEn50288.2021.9497304)

[2] M. Ketfi, M. Belahcene and S. Bourennane, "Efficient Transfer Learning for COVID-19 Diagnosis using X-Ray and Computed Tomography Images," 2022 International Conference of Advanced Technology in Electronic and Electrical Engineering (ICATEEE) 2022, pp. 1-5, Msila, Algeria.

doi: [10.1109/ICATEEE57445.2022.10093757](https://doi.org/10.1109/ICATEEE57445.2022.10093757).

[3] Ketfi M, Belahcene M, Bourennane S (2020) Image Enhancement and Hybrid Deep Features for COVID19 Diagnosis (CEE' 2023), Algiers, Algeria, 09-10 May 2023. <http://www.emp.mdn.dz/events/cee/program.html>

[4] Ketfi, Meryem, Mebarka Belahcene, and Salah Bourennane. "Automated Detection and Segmentation of Breast Cancer in Ultrasound Images with Deep Learning." 2024 8th International Conference on Image and Signal Processing and their Applications (ISPA). IEEE, 2024, Biskra, Algeria. <https://ieeexplore.ieee.org/abstract/document/10536843>

Abstract

Every year millions around the world could be saved if they had access to faster and more accurate diagnosis of their disease. Advances in Artificial Intelligence (AI) are set to revolutionize the healthcare industry. For the year 2018, the Data Science Bowl brought together many Data Scientists from around the world to tackle one of the biggest challenges for biologists and doctors: designing image segmentation algorithms to automate the detection of nuclei in biological cells. It is no coincidence that segmentation has become a hot topic in medical image processing. Thus, researchers and doctors can understand the underlying biological processes and speed up medical diagnosis.

A difficulty related to the data lies in the heterogeneity of the images. Indeed, the images are very varied, have different magnifications, different colorizations, and contain different cell types. Several strategies will therefore be required to overcome these imbalances and the size of the data set. A key characteristic of convolutional neural networks (CNNs) and Deep Learning (DL), in general, is the assumption of spatial invariance in image features. We are interested in the same patterns to be recognized in the different parts of the image. Technically, this amounts to having all the neurons of the same layer share similar weights, which considerably reduces the number of network parameters. It should be noted that this hypothesis could nevertheless limit the exploitation of very specific structures in an image, such as the geometry of an organ (face in biometrics).

The work developed in this thesis brings several innovative advancements in this context. First, for classification, we used the Transfer Learning (TL) model called VGG16 after extracting parameters using the best Gabor filters to simulate retinal performance. Then, we explored a new approach to image enhancement with the extraction of global and textural features based on DL. Initially, chest X-ray and computed tomography images were preprocessed and enhanced using histogram equalization (HE). Next, global and local features were extracted using hybrid feature descriptors such as MobileNetV2 via Local Binary Pattern (LBP) models and Gabor filters. Concatenation of the best models for optimal feature extraction was employed, and DL methods for deep feature extraction and data reduction were applied for optimal classification. To validate this work, we first tested this approach on the COVID19 database (collected during the epidemic period) as well as on various types of pneumonia.

For segmentation, we used the metaheuristic algorithm Particle Swarm Optimization (PSO) to improve the performance of our segmentation system. Two types of optimizations were studied: autoencoder optimization for image denoising before feeding it into the UNET model, and accuracy optimization of UNET. The results obtained are promising. All our experiments were conducted on different datasets, including COVID19, Viral Pneumonia, Breast Cancer, Skin Cancer, and a synthetic retinal database. The outcomes were satisfactory and promising, with potential for further improvement through enhanced detection and preprocessing techniques.

Keywords: Classification, Segmentation, Deep Learning, Scanner, Medical Imaging.

Résumé

Chaque année des millions dans le monde pourraient être sauvées s'ils avaient accès à un diagnostic plus rapide et précis de leur maladie. Les avancées de l'Intelligence Artificielle sont vouées à bouleverser le monde de la santé. Pour l'année 2018 le Data Science Bowl a réuni de nombreux Data Scientist du monde entier pour s'attaquer à un des plus grands défis pour les biologistes et médecins : concevoir un algorithme de segmentation d'image pour automatiser la détection des noyaux dans les cellules biologiques. Ce n'est pas par hasard que la segmentation est devenue un sujet d'actualité dans le traitement des images médicales. Ainsi, les chercheurs et les médecins peuvent comprendre les processus biologiques sous-jacents et accélérer le diagnostic médical.

Une difficulté liée aux données réside dans l'hétérogénéité des images. En effet, les images sont très variées, ont des grossissements différents, des colorisations différentes, et contiennent différents types de cellules. Plusieurs stratégies seront donc nécessaires pour surmonter ces déséquilibres et la taille de l'ensemble des données. La particularité d'un réseau de neurones convolutifs ou deep learning, en général, vient de l'hypothèse d'invariance spatiale des caractéristiques utilisées dans l'image : on s'intéresse aux mêmes motifs à reconnaître dans les différentes parties de l'image. Techniquement, cela revient à ce que tous les neurones d'une même couche partagent des poids similaires, ce qui réduit considérablement le nombre de paramètres du réseau. Il convient de noter que cette hypothèse peut néanmoins limiter l'exploitation de structures très spécifiques dans une image, comme la géométrie d'un organe (visage en biométrie).

Le travail développé dans cette thèse apporte plusieurs avancées innovantes dans ce contexte. Tout d'abord, pour la classification, nous avons utilisé le modèle de TL nommé VGG16 après avoir extrait les paramètres à l'aide des meilleurs filtres de Gabor visant à simuler les performances de la rétine. Ensuite, nous avons étudié une nouvelle approche d'amélioration d'image avec extraction de caractéristiques globales et texturales basée sur l'apprentissage profond (DL). Dans un premier temps, les images de radiographie thoracique et de tomodensitométrie ont été prétraitées et améliorées à l'aide de l'égalisation de l'histogramme (HE). Ensuite, les caractéristiques globales et locales ont été extraites à l'aide de descripteurs de caractéristiques hybrides tels que MobileNetV2 via les modèles Local Binary Pattern (LBP) et les filtres Gabor. La concaténation des meilleurs modèles pour une extraction optimale des caractéristiques a été utilisée, et des méthodes de Deep Learning (DL) pour l'extraction de caractéristiques profondes et la réduction des données en profondeur ont été appliquées pour une classification optimale. Pour valider ce travail, nous avons testé cette approche d'abord sur la base de données COVID19 (pendant la période de l'épidémie) ainsi que sur différentes pneumonies.

Pour la segmentation, nous avons utilisé l'algorithme métaheuristique Particle Swarm Optimization (PSO) dans le but d'améliorer les performances de notre système de segmentation. Deux types d'optimisations ont été étudiés : l'optimisation de l'autoencodeur pour le débruitage de l'image avant de l'injecter dans le modèle UNET, et l'optimisation de la précision de UNET. Les résultats obtenus sont encourageants. Toutes nos expériences ont été réalisées sur différents ensembles de données (COVID19, Pneumonie virale, Cancer du sein, Cancer de la peau, base de données rétinienne synthétique). Des résultats satisfaisants et prometteurs ont été obtenus et pourraient être encore améliorés grâce à de meilleures détections et prétraitements.

Mot clés : Classification, Segmentation, Deep Learning, Scanner, Imagerie Médicale.

المخلص

في كل عام، يمكن إنقاذ الملايين حول العالم إذا تمكنوا من الوصول إلى تشخيص أسرع وأكثر دقة لمرضهم. من المتوقع أن يؤدي التقدم في مجال الذكاء الاصطناعي إلى إحداث تغيير جذري في عالم الصحة.

في عام 2018، اجتمع العديد من علماء البيانات من جميع أنحاء العالم لمعالجة أحد أكبر التحديات التي يواجهها علماء الأحياء والأطباء: تصميم خوارزمية تجزئة الصور لأتمت اكتشاف النوى في الخلايا البيولوجية. ليس من قبيل المصادفة أن التجزئة أصبحت موضوعًا ساخنًا في معالجة الصور الطبية. وهكذا، يمكن للباحثين والأطباء فهم العمليات البيولوجية الأساسية وتسريع التشخيص الطبي.

تكمن الصعوبة المرتبطة بالبيانات في عدم تجانس الصور. في الواقع، الصور متنوعة للغاية، ولها تكبيرات مختلفة، وألوان مختلفة، وتحتوي على أنواع مختلفة من الخلايا. ولذلك ستكون هناك حاجة إلى عدة استراتيجيات للتغلب على هذه الاختلافات وحجم مجموعة البيانات. خصوصية الشبكة العصبية التلافيفية أو التعلم العميق، بشكل عام، تأتي من فرضية الثبات المكاني للخصائص المستخدمة في الصورة: نحن مهتمون بنفس الأنماط التي سيتم التعرف عليها في الأجزاء المختلفة من الصورة. من الناحية الفنية، يؤدي هذا إلى مشاركة جميع الخلايا العصبية في نفس الطبقة في أوزان مماثلة، مما يقلل بشكل كبير من عدد معلمات الشبكة. تجدر الإشارة إلى أن هذا الافتراض يمكن أن يحد من استغلال هيكل محددة للغاية في الصورة، مثل هندسة العضو (الوجه في القياسات الحيوية).

العمل الذي تم تطويره في هذه الأطروحة يجلب العديد من التطورات المبتكرة في هذا السياق. بدايةً، من أجل التصنيف، استخدمنا نموذج نقل التعلم المسمى VGG16 بعد استخراج المعلمات باستخدام أفضل مرشحات غابور التي تهدف إلى محاكاة أداء شبكية العين. بعد ذلك، قمنا بالتحقق من نهج جديد لتحسين الصورة من خلال استخراج الميزات العالمية والتركيبية بناءً على التعلم العميق (DL). أولاً، تمت معالجة صور الصدر بالأشعة السينية والأشعة المقطعية مسبقاً وتحسينها باستخدام معادلة الرسم البياني (HE). بعد ذلك، تم استخراج الميزات العالمية والمحلية باستخدام واصفات الميزات الهجينة مثل MobileNetV2 عبر نماذج النمط الثنائي المحلي (LBP) ومرشحات غابور. تم استخدام تسلسل أفضل النماذج لاستخراج الميزات الأمثل، وتم تطبيق أساليب التعلم العميق (DL) لاستخراج الميزات العميقة وتقليل البيانات العميقة من أجل التصنيف الأمثل. وللتحقق من صحة هذا العمل، اختبرنا هذا النهج أولاً على قاعدة بيانات كوفيد-19 (خلال فترة الوباء) وكذلك على أنواع مختلفة من الالتهابات الرئوية.

للتجزئة، استخدمنا خوارزمية Metaheuristic لتحسين سرب الجسيمات (PSO) من أجل تحسين أداء نظام التجزئة لدينا. تمت دراسة نوعين من التحسينات: تحسين التفسير التلقائي لتقليل تشويش الصورة قبل حلها في نموذج UNET، وتحسين الدقة UNET. والنتائج التي تم الحصول عليها مشجعة. تم إجراء جميع تجاربنا على مجموعات بيانات مختلفة (كوفيد-19، والالتهاب الرئوي الفيروسي، وسرطان الثدي، وسرطان الجلد، وقاعدة بيانات الشبكية الاصطناعية). تم الحصول على نتائج مرضية واعدة ويمكن تحسينها بشكل أكبر من خلال عمليات الكشف والمعالجة المسبقة بشكل أفضل.

الكلمات المفتاحية: التصنيف، التجزئة، التعلم العميق، الماسح الضوئي، التصوير الطبي.

List of Acronyms

AI Artificial Intelligence

DL Deep Learning

TL Transfer Learning

AUC Area Under Curve

CNN Convolutional Neural Networks

CT Computed Tomography

CV Computer Vision

DCNNs Deep Convolutional Neural Networks

HOG Histograms of Oriented Gradients

RGB Red, Green, and Blue

SVM Support Vector Machine

PSO Particule Swarm Optimization

IT Information Technology

MRI Magnetic Resonance Imaging

CX Chest X-ray

ML Machine Learning

LBP Local Binary Pattern

PET Positron Emission Tomography

AE Auto-Encoder

SE Stacked Encoder

SAE Stacked Auto Encoder

SDAE Stacked Denoising Auto Encoders

ANN Artificial Neural Network

SSDAE Sparse Stacked Denoising AE

Table of Contents

Acknowledgements.....	i
Dedication	ii
Publications & Scientific Communications-----	iii
Abstract.....	iv
Résumé.....	v
المخلص.....	vi
List of Acronyms	vii
List of Figures	xii
List of Tables.....	xiv
General Introduction.....	1
1 Chapter: Overview of Deep Learning in Medical Imaging.....	7
1.1 Introduction	7
1.2 Medical image processing chain	7
1.3 Medical Image Modalities	8
1.3.1 X-ray Imaging.....	9
1.3.2 Computed Tomography (CT)	10
1.3.3 Magnetic Resonance Imaging (MRI)	11
1.3.4 Ultrasound Imaging.....	12
1.4 Advancements in Deep Learning for Healthcare	15
1.5 Deep learning applied to medical imaging: Detection, Classification, and Segmentation of Diseases	15
1.5.1 Detection of Abnormalities	16
1.5.2 Classification of Diseases.....	16
1.5.3 Segmentation of Organs and Tissues	16
1.6 Conclusion.....	16
2 Chapter: AI Techniques for Medical Image Classification.	17
2.1 Introduction	17
2.2 State-Of-The-Art for COVID19 Classification.....	17
2.3 Transfer Learning effect for COVID19 classification for Diagnosis.....	20
2.3.1 Hardware.....	22

2.3.2	Performance Evaluation	22
2.3.3	An Approach Application.....	22
2.3.3.1	<i>Experience 1</i> :	23
2.3.3.2	<i>Experience 2</i> :	24
2.4	Gabor Descriptor and TL in COVID19 Classification for Diagnosis.....	27
2.4.1	Proposed Approach based Gabor Descriptor and TL	27
2.4.1.1	Preprocessing.....	27
2.4.2	Keras-Tensorflow with TL VGG16Implementation	29
2.5	Image Enhancement and Hybrid Deep Features for COVID19 Diagnosis	30
2.5.1	Proposed approach in COVID19 Classification.....	30
2.5.1.1	Preprocessing	31
2.5.1.1.1	Data augmentation	31
2.5.1.1.2	Histogram Equalization	31
2.5.1.1.3	Gabor Filter	32
2.5.1.1.4	Local Binary Pattern.....	32
2.5.1.1.5	Fine Tuning Model.....	33
2.5.2	Experimentation.....	33
2.5.2.1	Dataset	33
2.5.2.2	Proposed Model Implementaion in COVID19 classification	33
2.5.3	Comparison results with State of the Art.....	35
2.6	Conclusion.....	37

3 Chapter: Transfer Learning Fusion and Stacked Auto-Encoders for Viral Lung Disease Classification.....

3.1	Introduction	38
3.2	Contribution.....	44
3.3	Proposed TL Fusion and SAE Model.....	46
3.3.1	Preprocessing.....	48
3.3.2	Transfer Learning Fusion	49
3.3.3	Data Reduction and DFE.....	50
3.3.4	TLFSAE Model architecture	51
3.4	Experimentation and Results	53
3.4.1	Data Collection	53
3.4.2	Binary Classification.....	53
3.4.3	Concatenation models	55

3.4.4	Multi classification.....	57
3.4.4.1	<i>Experience 1</i> :.....	58
3.4.4.2	<i>Experience 2</i> :	58
3.4.4.3	<i>Experience 3</i> :	59
3.5	Comparison with state of the art methods.....	61
3.6	Conclusion.....	64
4	Chapter: Metaheuristic Optimization Deep Features for Pathologies Diagnostic	66
4.1	Introduction	66
4.2	Recent Work.....	66
4.2.1	Lung cancer classification based methods.....	69
4.2.2	Breast cancer classification based methods	71
4.2.3	Alzheimer disease classification based methods.....	72
4.3	Proposed Optimized Deep Features (ODF) Method.....	74
4.3.1	Preprocessing and Deep Features Extraction.....	75
4.3.2	Deep Features Extraction	75
4.3.3	Fine Tuning Model	76
4.3.4	Deep Features Optimization	77
4.3.4.1	Manta Ray Foraging Optimization	77
4.3.4.1.1	Chain Foraging	77
4.3.4.1.2	Cyclone Foraging.....	77
4.3.4.1.3	SomerSault Foraging.....	78
4.4	Classification	79
4.4.1	Binary classification.....	79
4.4.2	Multiclass Classification	80
4.5	Model Training with best solution and concatenation.....	81
4.5.1	Data Collection	82
4.6	ODF Results and Discussions.....	84
4.6.1	Binary Classification.....	85
4.6.1.1	CT Images Lung Cancer	85
4.6.1.2	Breast Cancer images.....	87
4.6.1.2.1	Breast Cancer Ultrasound Images	87
4.6.1.2.2	Breast Cancer Histological Images	89
4.6.1.3	Alzheimer Disease MRI images.....	92
4.6.2	Multiclass Classification	94

4.6.2.1	Lung Cancer	94
4.6.2.2	Breast Cancer Ultrasound Images.....	95
4.6.2.3	Alzheimer Disease.....	96
4.7	Comparison Results with State of the Art.....	96
4.8	Conclusion.....	98
5	Chapter: AI Techniques for Medical Image Segmentation and Optimization.	100
5.1	Introduction	100
5.2	The Evolution of Medical Image Segmentation.....	100
5.3	Overview of UNet and its variants architectures.....	101
5.4	OCAE and OUNET: Standard Automatic Optimization for Segmentation Medical Image	101
5.4.1	Recent Work of Segmentation.....	104
5.4.2	Contribution.....	107
5.4.3	Proposed OCAE+OUNET Approach.....	108
5.4.3.1	OCAE+OUNET Model description.....	109
5.4.3.1.1	Convolutional Auto Encoder	109
5.4.3.1.2	UNET	110
5.4.3.1.3	Particle Swarm Optimization (PSO)	111
5.4.4	OCAE+OUNET Experimentation	112
5.4.4.1	Datasets.....	112
5.4.4.2	Conversion, data augmentation, resizing and splitting.....	113
5.4.5	OCAE+OUNET Implementation.....	114
5.4.5.1	Optimized CAE for denoising	114
5.4.5.2	Optimized CAE+UNET for segmentation	116
5.4.5.3	Evaluation metrics	117
5.4.6	OCAE+OUNET Results and Discussion.....	118
5.4.6.1	Hyperparameters Selection	118
5.4.6.2	Monomodality.....	118
5.4.6.3	Multimodality.....	123
5.4.7	Application on different datasets	126
5.4.7.1	Dataset 3 Skin Cancer:.....	126
5.4.7.2	Dataset 4 Synthetic Retinal	127
5.4.7.3	Dataset 5 Stare Retinal dataset	128
5.5	State of the Art Comparison	131
5.6	Conclusion.....	133
	General Conclusion.....	135

List of Figures

Figure 1.1 Medical image processing chain	8
Figure 1.2 Medical Image Modalities [images are collected via the internet]	8
Figure 1.3 Example of X-ray Imaging	9
Figure 1.4 X-ray Radiography	10
Figure 1.5 Example of Computed Tomography (CT)	10
Figure 1.6 CT scan Radiography	11
Figure 1.7 Magnetic Resonance Imaging (MRI).....	11
Figure 1.8 MRI system.....	12
Figure 1.9 Example of Ultrasound Imaging	13
Figure 1.10 Ultrasound system	13
Figure 1.11 Design of the Based System.....	15
Figure 2.2 Accuracy, Sensitivity and Specificity with Tensorflow_Keras	23
Figure 2.3 Accuracy and loss curve with Tensorflow TL MobileNet_V2 (3000C/3000NC)	24
Figure 2.4 Accuracy and loss curve with Tensorflow TL MobileNet_V2 (4000C/4000NC)	24
Figure 2.6 Flowchart of proposed approach in COVID19 classification.....	31
Figure 2.7 Image Enhancement by HE.....	32
Figure 2.8 Gabor filter.....	32
Figure 2.9 LBP Descriptor	33
Figure 2.10 Results of different models on X-Rays and CT scan images.....	35
Figure 3.1 Block diagram of the proposed approach.....	47
Figure 3.3 Data preprocessing and splitting	49
Figure 3.4 Data Reduction and Deep Feature Extraction	51
Figure 3.5 Dataset with different size.....	51
Figure 3.6 Detailed parameters of VGG16 + Encoder model.....	52
Figure 3.7 Examples of CT scans and XR Lung images	53
Figure 3.8 Accuracy comparison of different TL models in CT and CXR image.....	55
Figure 3.9 Parameters number of different proposed approach	56
Figure 3.10 Comparing the validation and test accuracy, loss and time test of the different TL models.....	60
Figure 3.11 Val_Accuracy and Val_Loss for TL_SE /SAE	61
Figure 4.1 Proposed ODF Approach for Binary/Multiclass classification of different Diseases Classification ..	74
Figure 4.2 Used VGG16 Architecture.....	76
Figure 4.3 Fune Tuning VGG16 Model and Searching the Best Solution.....	76
Figure 4.4 Proposed approach with best solution.....	82
Figure 4.5 Different Dataset Collection	84
Figure 4.6 Accuracy and Loss curves of different experiences on Lung Cancer (CT Images) dataset.....	86

Figure 4.7 Confusion Matrixes for the VGG16+ODF with 2 classes on Lung Cancer CT dataset: (a) <i>Exp1</i> , (b) <i>Exp2</i> , (c) <i>Exp3</i> , (d) <i>Exp4</i>	87
Figure 4.8 Accuracy and Loss curves of different Binary Classification on Breast Cancer Ultrasound dataset ..	88
Figure 4.9 Confusion Matrixes for the VGG16+ODF with 2 classes on Breast Cancer Ultrasound dataset: (a) <i>Exp1</i> , (b) <i>Exp2</i> , (c) <i>Exp3</i>	89
Figure 4.10 Accuracy and Loss curves of different experiences on Breast Cancer histological and ultrasound dataset	90
Figure 4.11 Confusion Matrices for Binary Classification with VGG16+ODF on Breast Cancer datasets: (a) <i>Histologic and</i> (b) <i>Ultrasound</i>	91
Figure 4.12 Accuracy and Loss curves for Binary Classification on Alzheimer Disease MRI dataset	93
Figure 4.13 Confusion Matrixes for Binary Classification with VGG16+ODF on Alzheimer disease MRI dataset: (a): <i>Exp1</i> , (b): <i>Exp2</i> , (c): <i>Exp3</i> , (d): <i>Exp4</i>	94
Figure 4.14 Confusion Matrixes for the VGG16+ODF with:	95
Figure 4.15 Confusion Matrix for the VGG16+ODF with 4 classes on Alzheimer Disease dataset	96
Figure 5.1 Flowchart of the proposed approach OCAE+OUNET Segmentation.....	109
Figure 5.2 Different used Datasets.....	113
Figure 5.3 Optimization Segmentation based OCAE+OUNET Algorithm	114
Figure 5.5 Flowchart of the OUNET.....	116
Figure 5.6 Flowchart of the OCAE+OUNET algorithm	116
Figure 5.7 Images 2D Dataset1 and 3D Dataset2	120
Figure 5.9 Segmentation for original 3D dataset 2 COVID19 SLICE _Z (100 images/100 masks (256x256))	121
Figure 5.10 Segmentation for original 3D dataset 2 COVID19 SLICE _Z (2816 images/2816 masks (256x256))	122
Figure 5.11 Mono modality accuracy and Loss curves for COVID19 Lung datasets	122
Figure 5.12 Multimodality Segmentation for 3D dataset 2 (SLICE _Z, SLICE_Y) COVID19.....	124
Figure 5.13 Multimodality Segmentation for 3D Dataset 2 COVID-19 SLICE _Z, SLICE_Y, SLICE_X.....	125
Figure 5.14 Multimodality Accuracy and Loss curves	125
Figure 5.15 Segmentation for 2D Dataset 3 Skin Cancer (100 images/100 masks (256x256))	127
Figure 5.17 Segmentation for 2D Dataset 5 STARE Retinal (20 images/20 masks (256x256)).....	129
Figure 5.18 Accuracy and Loss curves for different datasets	130

List of Tables

Table 1-1 Comparison of the Imaging Modalities.....	14
Table 2-1 Data Distribution	21
Table 2-2 Performance metrics for classification.....	22
Table 2-3 Results achieved by our model (3000c/3000nc).....	23
Table 2-4 Results achieved by our model (4000c/4000nc).....	25
Table 2-5 Results of Different Experiments	25
Table 2-6 Comparison Results with State Of The Art.	26
Table 2-7 Results obtained with VGG16.....	29
Table 2-8 Results for HE and different descriptors with MobileNetV2 (CT scans images).	34
Table 2-9 Results for HE and different descriptors with MobileNetV2 (X-Rays images).....	34
Table 2-10 Comparison of Gabor Descriptor and TL with recent works.....	36
Table 3-1 Summary of the studies in literature	41
Table 3-2 Limitations of recent methods	43
Table 3-3 Data distribution.....	53
Table 3-4 Hyperparameters of different models	54
Table 3-5 Results obtained from different TL models.....	55
Table 3-6 Concatenation and Encoder effect for binary classification	56
Table 3-7 Best results for proposed multi classification	58
Table 3-8 Comparison results with state of the art methods	62
Table 4-1 Recent studies on lung cancer classification	70
Table 4-2 Recent studies on Breast cancer classification	72
Table 4-3 Recent studies on Alzheimer disease classification	73
Table 4-4 Hyperparameters Optimized by MRFO	84
Table 4-5 Protocol for Binary Classification	85
Table 4-6 Results with different Lung Cancer types for Binary Classification.....	85
Table 4-7 Results with different Breast Cancer types for Binary Classification	88
Table 4-8 Results with Breast Cancer Histological for Binary Classification.....	90
Table 4-9 Results with different Alzheimer Disease types for Binary Classification.....	92
Table 4-10 Protocol for Multiclass Classification	94
Table 4-11 Results with different Lung Cancer types for 4-classes Classification.....	95
Table 4-12 Results with different Breast Cancer types for 3-class Classification	95
Table 4-13 Results obtained with different Alzheimer disease types for 4-class classification	96
Table 4-14 Comparison with results and related studies.....	97
Table 5-1 Recent Work of Segmentation.....	105
Table 5-2 PSO parameters and Search Space	115
Table 5-3 Performance of different optimizers	118
Table 5-4 Effect of the data augmentation and size dataset CT D2 SLICE_ Z COVID19	119
Table 5-5 Effect of optimized preprocessing lung image on the results with small dataset.....	120
Table 5-6 Effect of the multi-view (multimodality) and data augmentation.....	123

Table 5-7 Application of the approach on Skin Cancer dataset	126
Table 5-8 Application of the approach on Synthetic Retinal dataset.....	127
Table 5-9 Application of the approach on Stare Retinal dataset	128
Table 5-10 Comparison with the state-of-the-art methods.....	131

General Introduction

This introduction provides a comprehensive overview of the work carried out throughout this thesis. We discuss the context in which the research was conducted, the underlying motivations, and how AI has been utilized to diagnose various diseases. We also highlight the primary challenges that need to be addressed and the contributions that our work has made towards their resolution.

Context

Nowadays, Computer Vision (CV) has become a rapidly expanding research field due to its importance in numerous applications. The goal of CV as a scientific discipline is to equip computers with the ability to process, analyze, and interpret images and videos similarly to how the human brain perceives visual stimuli [1]. It is an essential part of AI, enabling computers and systems to obtain useful information from digital visual inputs and execute decisions or make suggestions in response to that data. CV enables machines to see, interpret, and understand the visual environment whereas AI enables machines to think [2].

Human vision and CV both function similarly in the sense that they involve the process of perceiving, interpreting, and understanding visual information. However, when it comes to expertise and nuanced understanding, people often possess a higher level of knowledge and capability [3]. The human visual system learns object separation, distance estimation, motion detection, and anomaly detection through a lifetime of contextual learning. In contrast, CV trains machines to execute similar tasks using cameras, data, and algorithms. Deep Learning (DL) has impacted the area of CV through convolutional neural networks (CNNs). DL models can automatically extract hierarchical features from images, imitating the complex processing of the human visual system for use in various CV applications, such as image segmentation and classification.

In recent years, the field of Information Technology (IT) has experienced significant growth, leading to the development of various applications in medical imaging. This technological advancement has revolutionized the way medical professionals visualize and analyze diagnostic images [2]. Medical imaging involves analyzing scans such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Chest X-rays (CX) for diagnostic purposes. It plays a crucial role in identifying diseases and conditions by examining and interpreting the visual information captured in the images.

Medical imaging is essential to today's healthcare, providing non-invasive methods for diagnosing and monitoring a variety of diseases and conditions. However, analyzing medical images can be a challenging and time-consuming task, especially when it comes to segmenting and classifying different regions of interest.

Motivations

Radiology is a medical specialty that uses medical imaging techniques, such as X-rays, MRI, CT scans, MRI, and others, to diagnose and treat diseases. Predictive models in radiology refer to sophisticated computer-based tools and algorithms that analyze medical images to make predictions or assessments about various aspects of a patient's health. AI techniques, including Machine Learning (ML) and DL, are used to analyze the extracted radiomic data and make predictions. AI algorithms can learn patterns and relationships within the data that may not be obvious to human observers. Predictive models in radiology, enhanced by the combination of radiomics and AI, have revolutionized the field by providing highly accurate tools for medical image analysis.

In the field of medical imaging, professionals such as radiologists frequently rely on the manual examination and annotation of each individual image slice to pinpoint areas associated with specific diseases. This entails the meticulous process of outlining or demarcating affected regions within every slice of the 3D image volume. Manual disease classification and segmentation is notorious for being a time-intensive endeavor. Analyzing and accurately labeling each image slice demands a substantial amount of time. This task can become particularly arduous when dealing with a large number of slices within a single scan. Labeling image slices can become monotonous and mentally taxing. Experts are required to maintain a high level of concentration and precision throughout the process to ensure the accuracy of their annotations.

Objectives

The search for optimal and efficient solutions for diagnosing various diseases using AI is the primary objective for the majority of researchers. Our thesis aims to improve and enhance the performance of DL methods in the domain of medical image classification and segmentation of different diseases.

The main Steps of our work focus on the following aspects:

In this work, we aim to find an effective model to diagnose COVID19 by using a Transfer Learning (TL) model. We propose a DL based approach to detect COVID19 infection from chest CT Scans and X-rays images. We used Gabor Wavelet (GW) for feature extraction. Then, The VGG16 and MobileNetV2 was used in the classification of COVID19 and no-COVID19. Therefore, our goal is to find a robust framework to extract relevant features from images. Presenting a novel approach to image enhancement with global and textural features extraction based on DL. In this methodology, chest X-ray and CT scan images are initially preprocessed and enhanced using Histogram Equalization (HE). Then, global and local features are extracted using hybrid feature descriptors such as MobileNetV2 via the Local Binary Pattern models (LBP) and Gabor filters. Global features are computed using Gabor filters, while texture features are computed using the LBP descriptor.

Identifying an effective model based on features fusion using TL models for the classification of multiple viral respiratory diseases, encompassing COVID19.

Find a standard model to diagnose different pathologies using different sensors for different stages of diseases and above all detect the disease at its beginning for early diagnostic.

Discovering a rapid, effective, and, above all, standardized model capable of ensuring the diagnosis of various diseases affecting different organs. To achieve this, we undertake two distinct tasks: the first pertains to monomodality/multimodality, while the second focuses on optimizing the proposed CAE_UNET segmentation model by using PSO algorithm. These approaches are assessed across various pathologies and organ systems.

Contributions

In the classification part of the thesis, a detailed study on classification is conducted and utilizes the TL Fusion and Stacked Auto-Encoders (TLFSAE) methodology for the classification of viral lung diseases, bringing several important approaches.

- Using Transfer Learning: TLFSAE uses TL to leverage knowledge gained from previous tasks to improve the performance of classification models. This transfer acquired knowledge to related tasks such as image classification, which can be used to improve the accuracy of viral lung disease classification.
- Using Stacked Auto-Encoders: TLFSAE uses Stacked Auto-Encoders for unsupervised feature learning.

This helps to extract relevant features from the input data, which can improve the accuracy and robustness of classification models.

- Model fusion: TLFSAE uses model fusion to improve the robustness and efficiency of classification models. By merging information from multiple pre-trained models, a more complete and accurate representation of the characteristics of the input data can be obtained.
- Reduced data dimensionality: TLFSAE uses Stacked Auto-Encoders to reduce data dimensionality, allowing more complex data to be processed and reducing computation time.
- Utilize an effective and simple classification algorithm: TLFSAE employs metric classification algorithms to enhance model accuracy. By leveraging these algorithms, more reliable and precise predictions can be achieved.

The TLFSAE approach with model fusion has the potential to improve the accuracy of Viral Lung Disease classification. Further optimization of this approach can be achieved through exploring different pre-trained models, fusion techniques, and classification algorithms.

In the segmentation part, the proposed approach is based on optimized model using UNET architecture, applied to COVID19, skin cancer and retinal vessel segmentation. The framework contains three stages: the first stage is the preprocessing images, which contains the conversion of nifty to png file, the data augmentation and the resizing. Then, a proposed denoising of the images using CAE before segmentation is applied in the second stage. PSO algorithm is used in the third stage to optimize the parameters of CAE for properly denoising images and saving valuable information, after that UNET is used in the segmentation task, which ensures the segmentation of the infection region. To evaluate this model and testing its robustness and efficiency, lung datasets, skin cancer dataset, synthetic retinal dataset and retinal vessel segmentation (STARE) are used. A pipeline is proposed to generating synthetic medical images from a pair of Generative Adversarial Networks (GAN). The GAN trained with retinal vessel segmentations from the DRIVE dataset. Several tasks need to be explored to unlock a high-performance, efficient, and standardized segmentation method:

- Task 1: Study of monomodality (slice O_x , slice O_y , slice O_z), and choose which of the terms should be used;
- Task 2: Multimodality and the concatenation of slices to know which of the two contributes the most to the deep segmentation;

In the segmentation optimization part, the essential contributions are:

- Task 3: Study the effect of the denoising by OCAE on the segmentation model;

- Task 4: Selection of the optimal parameters OUNET model and OCAE filters number Optimization.

The contribution of this work focuses on implementing hybrid methodology, where the PSO algorithm is applied to find the optimal design of parameters for CAE architectures. The main contributions of this work are as follow:

- Datasets Collection: Collect various datasets for the comprehensive evaluation of the proposed segmentation models.
- CAE Parameter Optimization: PSO algorithm is used to fine-tune CAE parameters, ensuring optimal denoising performance.
- Architectural Approaches: UNET, CAE+UNET, and OCAE+OUNET architectures are explored for the segmentation of COVID19 and other pathologies.
- Monomodality and Multimodality Analysis: study of the impact of each modality and multimodal concatenation to know their relevance in segmentation tasks.

Thesis Organization

Our study has been organized into five primary chapters:

Chapter 1: The first chapter is dedicated to providing a comprehensive overview of medical images. Firstly, we explain the fundamental principles of various types of imaging and subsequently delineate the tools employed for evaluating their effectiveness. Our aim through this chapter is to contextualize the issues associated with each imaging modality and explicate their unique challenges and advantages vis-à-vis other techniques. Additionally, we analyze the key components of each imaging modality, namely detection, feature extraction, and recognition. Lastly, we underscore the challenges encountered in this domain that we intend to scrutinize in the course of this thesis.

Chapter 2: In this chapter, we will present a state of the art on classification techniques.

Chapter 3: In the classification part of the thesis, a detailed study on classification is conducted and utilizes the TL Fusion and Stacked Auto-Encoders (TLFSAE) methodology for the classification of viral lung diseases, bringing several important approaches.

Chapter 4: In the classification optimization part, we collect a larger dataset and focus on improving the quality of images in the preprocessing stage. We also plan to utilize metaheuristic algorithms to optimize our proposed model for the detection and diagnosis of various diseases

Chapter 5: In the segmentation part, the proposed approach optimizes the UNET architecture for application in COVID-19, skin cancer, and retinal vessel segmentation, where the PSO algorithm is applied to find the optimal design of parameters for CAE architectures. Optimization.

Conclusion: The final part presents a conclusion on our work and a vision for future work.

1 Chapter: Overview of Deep Learning in Medical Imaging

1.1 Introduction

In modern medicine, medical imaging plays a pivotal role in diagnosing various diseases across different medical disciplines. The exploitation of relevant information on the human brain has many useful clinical applications, which help the practitioner in the process of diagnosing many diseases [4]. Several medical imaging modalities have been employed in clinical settings for disease diagnosis. Techniques such as X-rays, CT scans, MRI, and Positron Emission Tomography (PET) scans play a crucial role in tumor identification, determining their size, location, and metastasis. Additionally, Chest X-rays and CT scans assist in diagnosing lung conditions such as pneumonia, tuberculosis, chronic obstructive pulmonary disease (COPD), and lung cancer. The ongoing advancement of imaging technologies broadens opportunities for accurate and early diagnoses across different medical specialties.

In this chapter, we explore the foundational aspects and broad scope of medical imaging. Our goal is to provide readers with a comprehensive understanding of the key concepts and principles that underpin the diverse field of medical image analysis. Next, we briefly introduce the principles behind various medical imaging techniques.

1.2 Medical image processing chain

Computed Tomography (CT) utilizes X-ray beams to capture 3D pixel intensities within the human body [5]. This process begins with a heated cathode emitting high-energy electrons, which subsequently release their energy in the form of X-ray radiation. These X-rays penetrate through the body's tissues and strike a detector situated on the opposing side. Notably, dense tissues, such as bones, absorb a greater amount of radiation compared to softer tissues like fat. Consequently, areas where X-rays are unabsorbed within the body, such as the air-filled regions within the lungs, appear black on the detector, resembling the appearance of a black film. In contrast, denser tissues are represented as white [6].

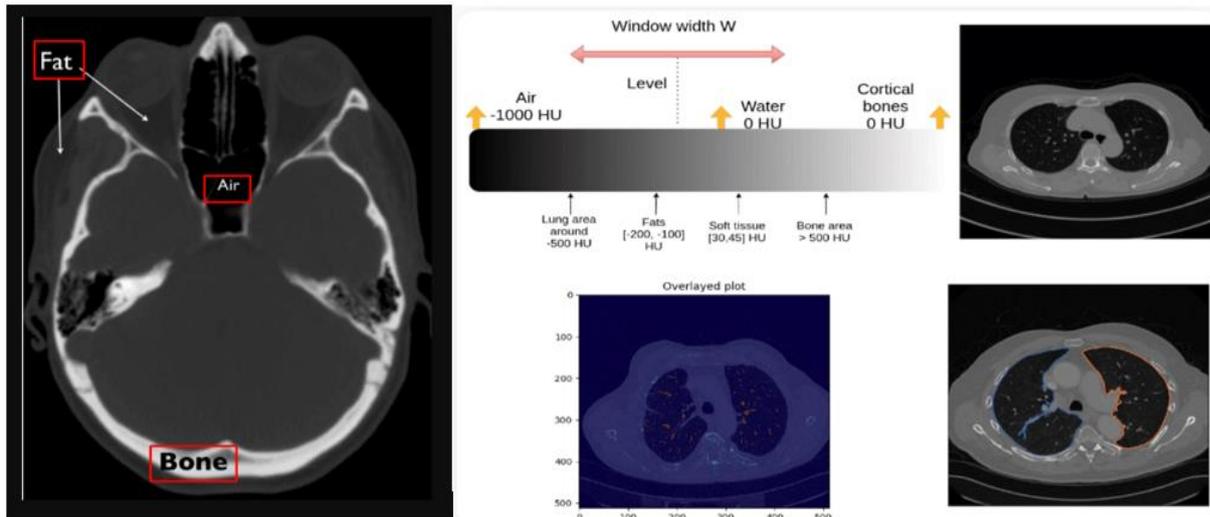


Figure 1.1 Medical image processing chain

1.3 Medical Image Modalities

Medical imaging modalities are various techniques employed to visualize the internal structures of the human body for diagnostic and research purposes [6]. The modality of imaging is a crucial aspect of the image for medical retrieval. Each modality possesses its strengths and limitations, and they are utilized based on the specific clinical needs of a patient. These modalities fulfill various purposes and offer valuable insights for diagnosis, treatment planning, and monitoring of medical conditions[7].



Imaging Modalities



Figure 1.2 Imaging Modalities [images are collected via the internet]

1.3.1 X-ray Imaging

A chest X-ray is a common medical imaging technique that uses a small dose of ionizing radiation to create pictures of the structures inside the chest, including the heart, lungs, blood vessels, ribs, and diaphragm. It's a valuable diagnostic tool used to detect and monitor a variety of conditions related to the chest area [8]. Chest X-rays can reveal a range of conditions. Some common findings include pneumonia (infection in the lungs), pleural effusion (accumulation of fluid around the lungs), pneumothorax (collapsed lung), lung tumors, congestive heart failure (fluid buildup in the lungs), and rib fractures. In addition, Depending on the findings of a chest X-ray, additional imaging tests might be recommended. For example, if an abnormality is detected, a more detailed imaging study like a CT scan or an MRI might be ordered to get a clearer view of the area in question.



Figure 1.3 Example of X-ray Imaging

Radiography is based on the use of X-rays, which possess the property of passing through tissues to varying degrees based on their density. Consequently, an X-ray source is positioned in front of the body to be radiographed, while a detector is positioned behind it. The reading must be carried out after checking the quality of the images according to various criteria. Frontal views in postero-anterior view and left profile should be taken with deep inspiration and breath-hold, with the shoulders and arms well clear. Emitted photons pass through the body,

experiencing different absorption levels depending on the density of the tissues they encounter. This differentiation allows for the visualization of bones and muscles in the resulting image [9].

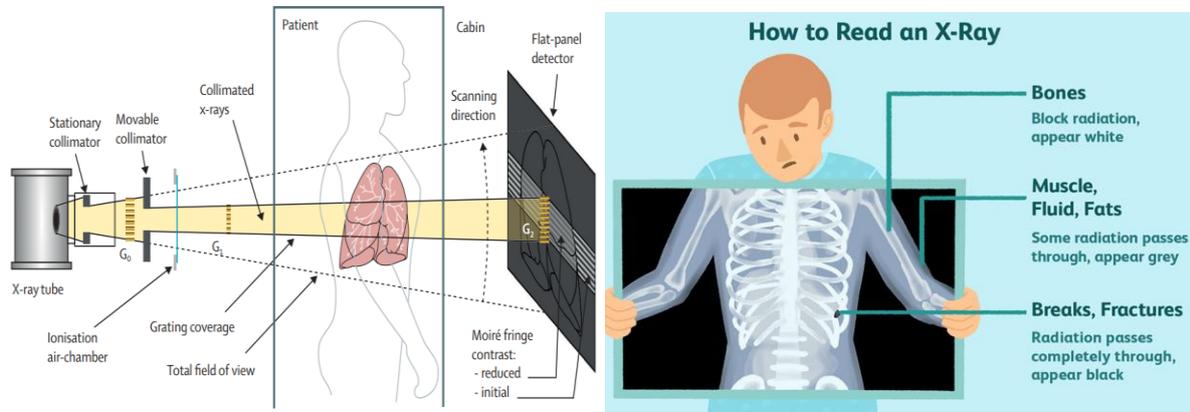


Figure 1.4 X-ray Radiography

1.3.2 Computed Tomography (CT)

The modality of CT images is highly radiative and harmful to the human body due to their high-frequency range from 3×10^{16} to 3×10^{19} Hz [4]. A CT scan, or Computed Tomography scan, is a medical imaging technique that uses X-rays and computer technology to create detailed cross-sectional images of various structures within the body [9]. It provides a comprehensive view of bones, organs, blood vessels, and tissues. During a CT scan, a rotating X-ray machine takes a series of images from different angles, and these images are processed by a computer to generate cross-sectional slices or 3D reconstructions of the area being examined. CT scans are valuable tools for diagnosing and monitoring various medical conditions, including injuries, tumors, infections, and internal diseases. They offer higher resolution and more detailed images compared to traditional X-rays and can provide essential information for healthcare professionals to make accurate diagnoses and treatment plans.

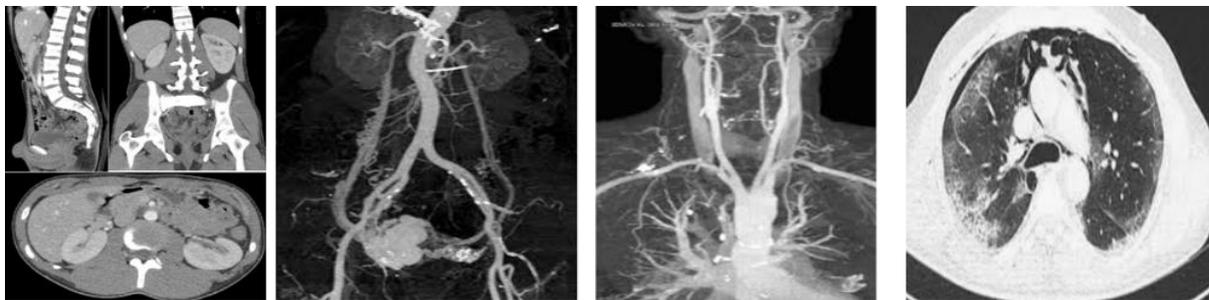


Figure 1.5 Example of Computed Tomography (CT)

The scanner operates on the same principle as radiology, employing an X-ray source and a detector positioned on opposite sides of the subject's body under examination. It enables the

acquisition of three-dimensional (3D) images by concurrently rotating both the X-ray emitting source and the detector around the body. Computer processing of the interim two-dimensional (2D) projections is then carried out to generate the final 3D images.

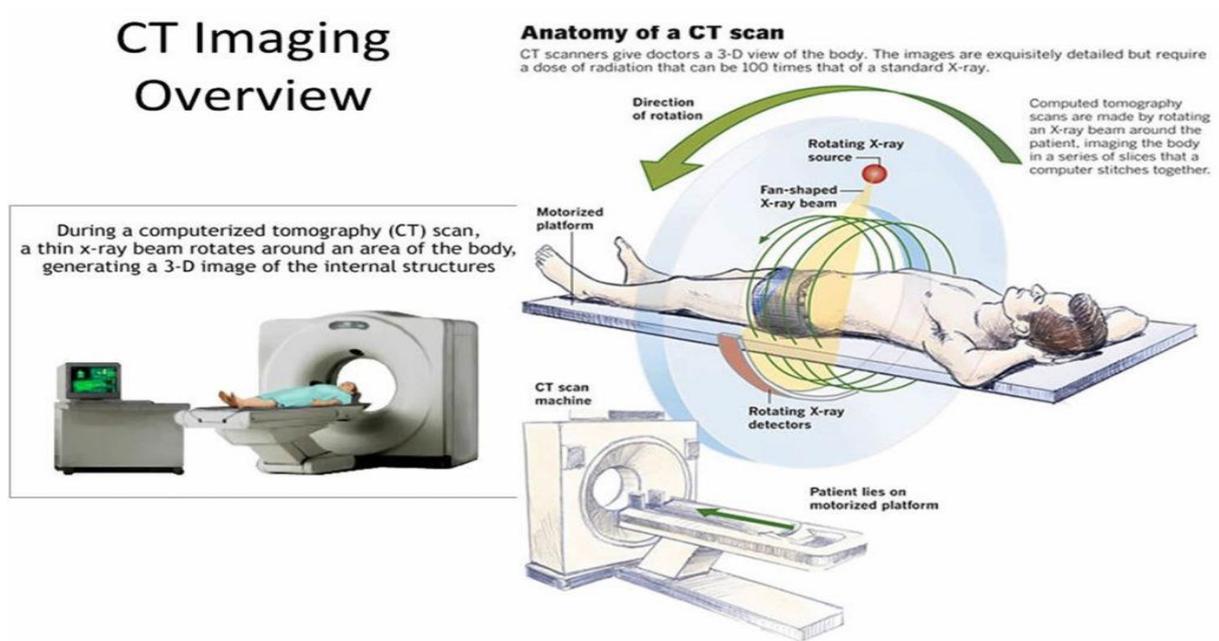


Figure 1.6 CT scan Radiography

1.3.3 Magnetic Resonance Imaging (MRI)

MRI is a sophisticated medical imaging technique that uses a strong magnetic field and radio waves to generate detailed and clear images of the internal structures of the body. Unlike X-rays or CT scans, MRI does not involve ionizing radiation, making it a safer option for repeated imaging [4]. MRI images are used to examine blood vessels, brain and breast tumors, abnormal tissues, spinal injuries, etc.

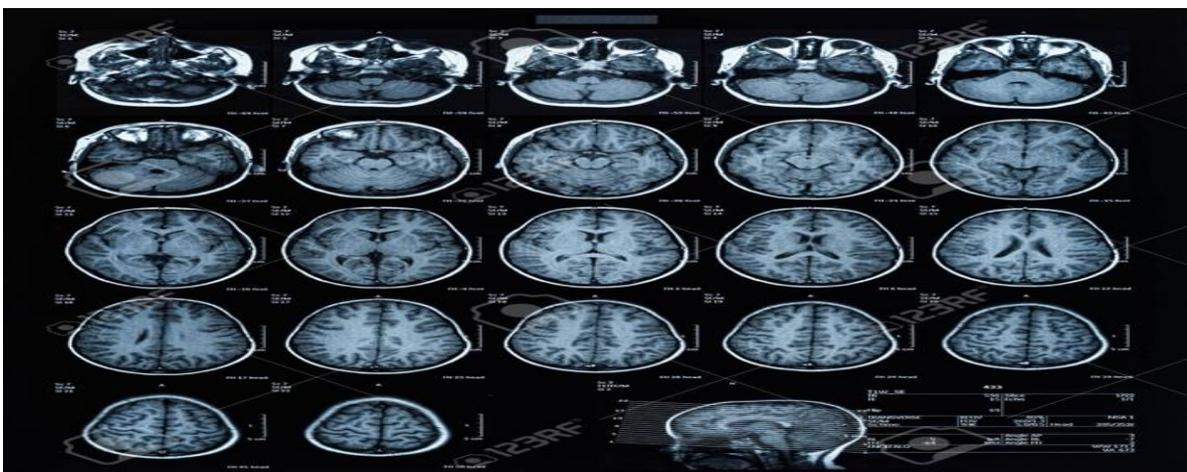


Figure 1.7 Magnetic Resonance Imaging (MRI)

MRI is based on the magnetic properties of water molecules, which constitute over 80% of the human body. Specifically, the hydrogen atoms within water molecules possess a "magnetic moment" or spin, akin to a miniature magnet. The MRI apparatus generates a robust magnetic field (B_0) using a coil. The patient is positioned at the center of this magnetic field, causing all water molecules within the body to align themselves along the B_0 field. An antenna, typically placed over the area of interest, facilitates the transmission and reception of specific frequencies. During emission, the induced frequency prompts the molecules to tilt in a plane perpendicular to the B_0 field. Subsequently, when the antenna ceases emission, the molecules return to their original alignment, thereby emitting a frequency. This emitted frequency is then captured by the antenna and processed as an electrical signal, which is subsequently analyzed by software. Notably, the signal varies based on the water content within the observed tissues, distinguishing tissues with differing water concentrations.

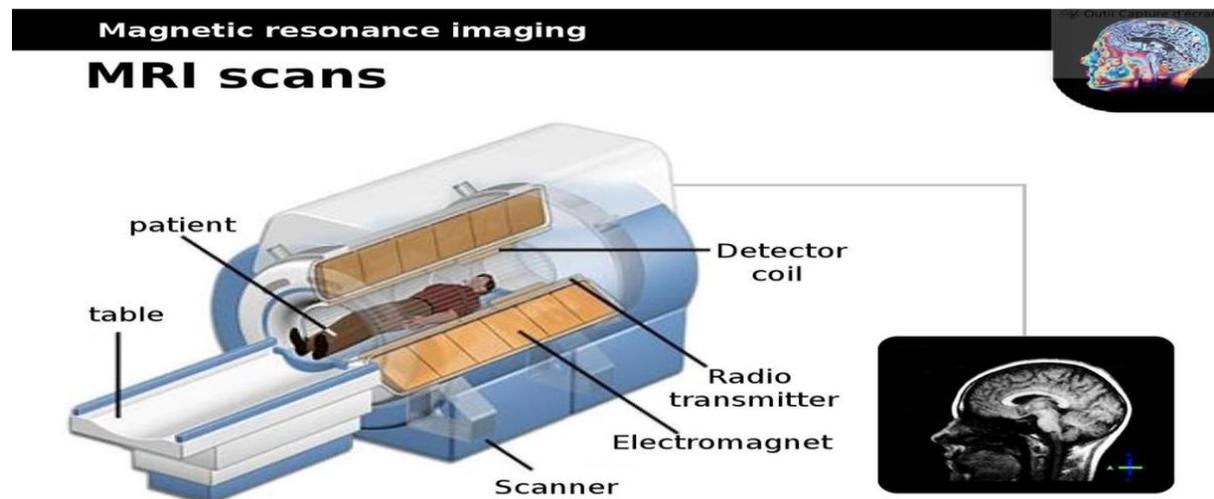


Figure 1.8 MRI system

1.3.4 Ultrasound Imaging

Ultrasound imaging, also known as sonography, is a medical imaging technique that uses high-frequency sound waves to create real-time images of the internal structures of the body. It is particularly useful for visualizing soft tissues, organs, blood flow, and the development of fetuses during pregnancy. Ultrasound imaging is non-invasive, safe, and does not involve ionizing radiation, making it a preferred choice for various medical applications.

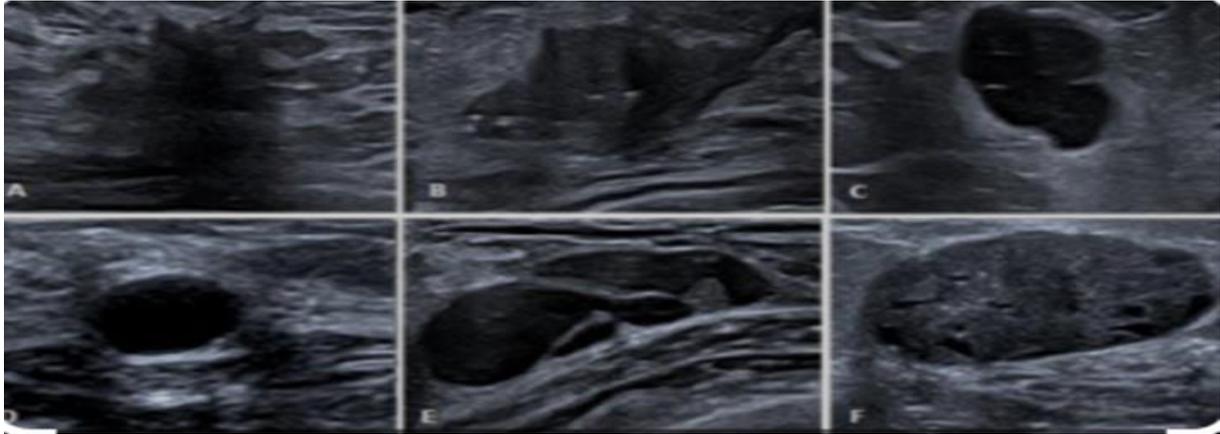


Figure 1.9 Example of Ultrasound Imaging

Ultrasound is a medical imaging modality that relies on exposing tissues to ultrasound waves and receiving their echo. The ultrasound system consists of a screen and a probe transmitting and receiving waves (called a transducer). When subjected to an electric current, the micro-ceramics on the surface of the probe vibrate and emit ultrasonic waves. These waves pass through the tissues and echo differently depending on their density: the denser the tissue, the greater the echo. The waves return to the probe, vibrating the ceramics, which induce an electrical current processed by a computer. Thus, during an ultrasound carried out in the case of pregnancy monitoring, it is possible to differentiate the skeleton of the fetus, its organs, the amniotic fluid, etc.

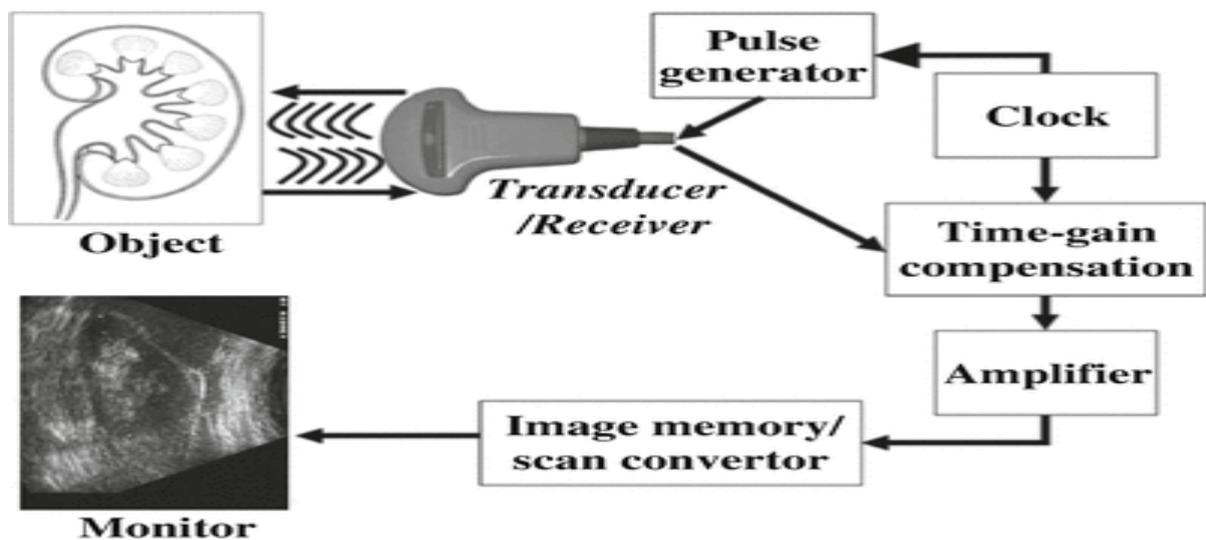


Figure 1.10 Ultrasound system

Table 1-1 Comparison of the Imaging Modalities

<i>Medical Modalities</i>	<i>Uses</i>	<i>Radiation</i>	<i>Image Quality</i>	<i>Soft Tissue Contrast</i>
X-Rays	Visualizing bones, detecting lung diseases, and identifying certain abnormalities.	Involves ionizing radiation, which poses some risk.	Provides 2D images with limited tissue differentiation	Limited ; better for visualizing dense structures like bones
CT	Detailed visualization of bones, organs, blood vessels, and tissues.	Involves ionizing radiation, which can carry some risk. Higher radiation dose compared to X-rays.	High-resolution cross-sectional images with excellent tissue differentiation.	Good, but not as detailed as MRI.
MRI	Detailed images of soft tissues, organs, and neurological structures	No ionizing radiation; considered safe for repeated use.	High-resolution images with excellent soft tissue contrast	Outstanding; superior for soft tissue evaluation compared to other modalities.
US	Real-time imaging of soft tissues, organs, blood flow, and fetal development.	No ionizing radiation; considered safe for all, including pregnant women	Real-time images with good tissue differentiation.	Adequate for most applications.

Table 1-2 Imaging Modalities (Advantages and Disadvantages)

<i>Imaging Modality</i>	<i>Advantages</i>	<i>Disadvantages</i>
X-ray Radiography	Fast and provide immediate results. Available in most healthcare facilities. Low Cost, and low dose ionizing radiation	Limited Soft Tissue Detail: less effective for evaluating organs and soft tissue abnormalities.
Computed Tomography (CT)	Provides detailed cross-sectional images of the body in multiple. Higher sensitivity for detecting nodes. Fast, and 3D Reconstruction	Higher radiation doses compared to some other modalities. More expensive than X-rays. Requires the use of contrast agents, which can have side effects in some patients.
Ultrasound	Provides real-time images and is useful for imaging moving structures, such as the heart or fetus. No Ionizing Radiation. Safe for Pregnancy.	Image quality can vary depending on the operator's skill. Limited Depth. Limited Bone Imaging.
MRI	Excellent Soft Tissue Contrast making it ideal for imaging the brain, spinal cord, and soft organs. No Ionizing Radiation. Can provide information about tissue composition, blood flow, and functional data.	Longer Exam Times. Cost: more expensive than X-rays and CT scans. Requires specific expertise.

1.4 Advancements in Deep Learning for Healthcare

Numerous papers have documented instances where DL has been employed with feature input in the realm of medical imaging. These applications encompass a wide range of areas, including the detection of lung nodules in chest X-rays and thoracic CT scans, the classification of lung nodules as benign or malignant in both chest X-rays and thoracic CT scans, the identification of micro calcifications in mammography, the discovery and differentiation of masses as benign or malignant in mammography, the recognition of polyps in CT colonography, and the spotting of aneurysms in brain MRI scans. Furthermore, DL has found utility in addressing regression problems, as exemplified by its use in determining subjective similarity measures for mammographic images [2].

1.5 Deep learning applied to medical imaging: Detection, Classification, and Segmentation of Diseases

DL advancements extend to image segmentation and classification. Disease classification models use CNNs to identify diseases from medical images accurately. This aids in early detection and intervention, improving patient outcomes.

DL has made significant strides in medical imaging, revolutionizing the way diseases are detected, classified, and segmented from medical images such as X-rays, MRI scans, CT scans, and ultrasound images. Here are some key applications.

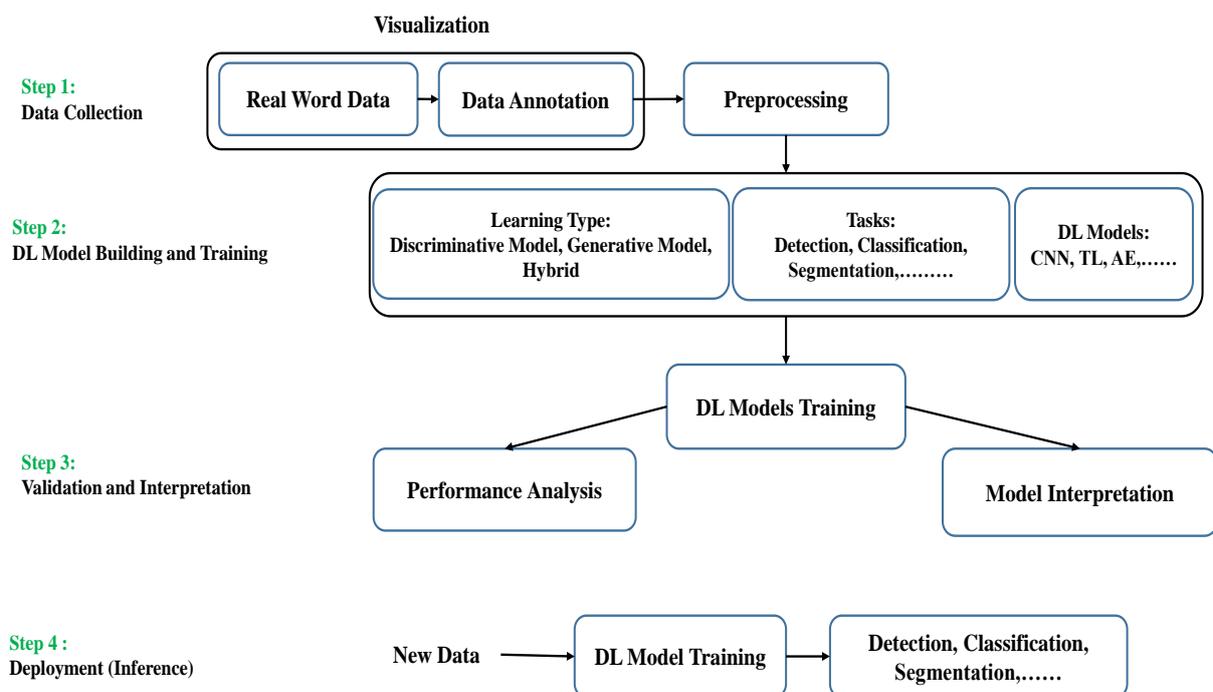


Figure 1.11 Design of the Based System

1.5.1 Detection of Abnormalities

DL models can be trained to detect abnormalities in medical images. For example, they can identify tumors, lesions, fractures, and other anomalies in X-rays, CT scans, and MRI images [10].

1.5.2 Classification of Diseases

DL models can classify medical images into different disease categories. For instance, they can classify skin lesions as benign or malignant, classify lung nodules as cancerous or non-cancerous, or classify retinal images to detect diabetic retinopathy [11].

1.5.3 Segmentation of Organs and Tissues

Segmentation involves partitioning an image into multiple segments to identify regions of interest. DL models can segment organs, tissues, and abnormalities in medical images. For example, they can segment the heart from cardiac MRI images, segment tumors from brain MRI images, or segment blood vessels from retinal images [12].

1.6 Conclusion

In this chapter, we have presented the essential concepts related to medical imaging, as well as the various modalities. Despite the advancements in medical imaging techniques, which are well-suited for observing the structures of different organs, practitioners still face challenges in analyzing the acquisition results. This is not only due to the artifacts specific to these types of imaging but also to the complexity of organ anatomy, which encompasses numerous structures.

To interpret medical images, doctors must accurately analyze the results to reach a final diagnosis. This operation is challenging and crucial for a reliable diagnosis. One of the fundamental operations in image processing is classification and segmentation. In the realm of medical imaging, classification involves categorizing specific elements within an image, such as tumors or abnormalities, into distinct groups based on their characteristics. On the other hand, segmentation is the process of delineating and outlining the boundaries of these elements, separating them from surrounding tissues. Through accurate classification and segmentation, healthcare professionals gain valuable insights into the nature and extent of diseases, which enables the identification, understanding, and monitoring of the progression of pathologies, aiding in precise diagnosis and tailored treatment plans. In this work, various methods for the classification and segmentation of medical images have been developed. These methods will be discussed in the next chapters.

2 Chapter: AI Techniques for Medical Image Classification

2.1 Introduction

In the context of the new pandemic, COVID19 classification using DL is crucial for early diagnosis. The healthcare field requires decision-making technologies to effectively manage the COVID19 virus and provide appropriate suggestions to prevent its spread. AI plays a vital role in suggesting the development of a system for early detection of COVID19. Preprocessing of medical images is a critical step in improving quality, removing noise, and extracting useful information. There are many techniques available to process images in the spatial and frequency domains. However, low contrast in medical images remains a significant challenge for detecting various diseases, including COVID19. Therefore, many researchers have focused on image preprocessing techniques to address this issue. In this chapter, we focus on COVID19, and subsequently, the studies have been expanded to other pathologies.

2.2 State-Of-The-Art for COVID19 Classification

Medical image classification is a critical aspect of image analysis within the realm of medical imaging. It involves assigning a label or category to an image based on the presence of specific features, patterns, or characteristics. The ultimate goal is to assist healthcare professionals in diagnosing diseases, identifying anomalies, and making informed decisions about patient care. Medical image classification aids in automating the diagnostic process, enhancing efficiency, reducing human error, and providing a quantitative basis for decision-making. It helps healthcare practitioners make accurate and timely assessments.

Various techniques are employed for medical image classification, including Traditional ML: such as Support Vector Machines (SVM), Random Forests, and k-Nearest Neighbors (k-NN) can be used with handcrafted features extracted from images, and DL such as Convolutional Neural Networks (CNNs). These neural networks are designed to automatically learn hierarchical features from raw image data, eliminating the need for manual feature engineering.

In this chapter, we will present a state of the art on classification techniques, we give a taxonomy of existing approaches, techniques, and algorithms. We will also expose disease diagnosis and detection methods. The history of disease diagnosis and detection methods spans

centuries and has evolved significantly with advancements in technology, scientific understanding, and medical knowledge. In this chapter, we will detail in particular the methods based on classification. Finally, we will end the chapter with a synthesis of the methods and techniques studied. The purpose of this chapter is to give an overview of the most significant methods of classification in the medical field.

Hussain, Lal, et al.[13] aim to develop an AI imaging analysis tool to classify COVID19 lung infection based on portable CXRs using Texture and morphological features extraction. They analyzed public datasets of COVID19, bacterial pneumonia, non-COVID19 viral pneumonia, and normal CXRs and used five supervised machine-learning algorithms to classify COVID19. The results show that AI classification of texture and morphological features of portable CXRs accurately distinguishes COVID19 lung infection in patients in multi-class datasets. The study concludes that DL methods have the potential to improve diagnostic efficiency and accuracy for portable CXRs. For binary classification, the accuracy, sensitivity and specificity were :100%, 100%, and 100% respectively, for COVID19 and normal; 96.34%, 95.35% and 97.44% for COVID19 and bacterial pneumonia; and 97.56%, 97.44% and 97.67% for COVID19 and non-COVID19 viral pneumonia. For multi-class classification, the combined accuracy and AUC were 79.52% and 0.87, respectively.

Wang, et al.[14] They introduced a deep convolutional neural network designed for detecting COVID19 called COVID-Net using a chest X-ray images dataset comprising 13,975 CXR images across 13,870 patient cases. COVID-Net employs an explainability method to identify critical factors associated with COVID cases and ensure decisions are based on relevant CXR image data.

Almalki, Yassir Edrees, et al.[15] proposed a new technique named CoVIRNet, which utilized Artificial Intelligence (AI) techniques like DL to detect COVID19. However, gathering a dataset for analysis is a challenge due to limited patients consenting to participate in the study. CoVIR-Net demonstrated a high level of accuracy, achieving 95.7%, and its feature extractor with a random-forest classifier obtained a 97.29% accuracy rate. The proposed method is anticipated to provide an automated approach for identifying COVID19 patients.

Mansour, Romany F., et al. [16] proposed a novel unsupervised DL-based model, the UDL-VAE, for COVID19 detection and classification. The model uses an adaptive Wiener filtering-based preprocessing technique to enhance image quality and Inception v4 with the Adagrad technique as a feature extractor. The unsupervised VAE model is applied for the

classification process. Experimental results show that the UDL-VAE model performs better than existing models with an accuracy of 0.987 and 0.992 for binary and multiple classes, respectively.

Fusco, Roberta, et al.[17] Presented an overview of using AI to combat COVID19 by utilizing chest CT and CXR images. The study analyzed several datasets from January 2019 to June 2021 and identified ML and DL Approaches for Diagnosis and Treatment. The results show that there is a high summarized accuracy of the selected papers, but with an important variability. Nonetheless, AI approaches could be used in various ways such as identifying disease clusters, monitoring cases, predicting future outbreaks, and managing COVID19 diagnosis and treatment.

Meraihi, Yassine, et al.[18] Provided an overview of 160 ML algorithms to detect, diagnose, and predict COVID19 with two categories : supervised learning and DL, and analyzed based on the ML algorithm used, parameters, problem type, analyzed data type, and evaluated metrics.

Saleem, Farrukh, et al.[19] Presented a Systematic Literature Review (SLR) that examines the latest advancements in the application of ML and DL algorithms for the analysis of COVID19 data. The study screened 218 articles and included 57 studies that met the PRISMA guidelines. The review discovered that CNN (DL) and SVM (ML) were the most widely used algorithms for forecasting, classification, and automatic detection. Compartmental models were found to be effective in determining the epidemiological characteristics of COVID19. The article recommends the utilization of ML, DL, and mathematical models as potential solutions to mitigate the healthcare impact of the epidemic.

Costa, Yandre MG, et al.[20] Have presented a review of the top 100 most cited papers in the field of using computer-based strategies to detect COVID19. While thousands of works have been published on this topic, not all have made significant contributions. In addition, the review aims to help researchers by evaluating the distribution of papers based on various aspects, such as the type of medical imaging, learning settings, segmentation strategy, explainable artificial intelligence (XAI), and the availability of dataset and code.

2.3 Transfer Learning effect for COVID19 classification for Diagnosis

This work aims to develop an effective model for diagnosing COVID-19 using a TL approach. The purpose is to classify COVID19 infected persons from chest X-Ray (XR) and Computed Tomography (CT) images. Several TL models have been studied to find the most efficient and effective among them. The proposed approach is based on Tensorflow and the architecture uses the MobileNet_V2 model.

The emergence and wide spread of the new coronavirus have infected many people around the world, making it necessary to combat and eradicate it. Early disease detection is crucial in preventing the spread of the pandemic due to its highly contagious nature and limited treatment options. Real-time reverse transcription-polymerase chain reaction (RT-PCR) is the strategy that has been used to discover the presence or absence of this type of virus. XR and CT images are the common tools used to detect COVID19, and the datasets are publicly available. The majority of research on image classification focuses on image feature extraction and classification methods since features are necessary for classification. Multiple network architectures are included in the benchmarking methodology for feature extraction and obtaining real-time performance. TL models have been used to aid in the detection of the severity and decision-making, such as ResNet50, MobileNet, ShuffleNet, VGG19, VGG16, GoogleNet, and AlexNet, DenseNet,....etc . Huang, G et al.[21] used four different TL models: AlexNet, ResNet18, and DenseNet201 to distinguish the new coronavirus and other pneumonia maladies. In this work, we will focus exclusively on utilizing the TL model known as MobileNet_V2 for the detection of COVID19. The MobileNet architecture leverages two key advantages: depth-wise and point-wise separable convolutions. In this context, it represents the pinnacle of the inception module, employing separate spatial convolutions for each channel, a technique referred to as depth-wise convolutions

In this proposed approach, we utilize the MobileNet_V2 TL model for image classification as follows:

Change of architectures for deep features extraction with :

- Tensorflow-MobileNet_V2.
- We add hidden layers and increase the size of dataset.

Improve the performance of the models by major steps can be summarized as follows :

- Realization of big dataset XR and CT images.

- Resizing the images in the dataset in 224x224 pixels.
- Data augmentation.
- Choice of parameters model of DL.

MobileNet_V2 model (see Fig 2.1) is used to reduce the number of parameters and computations.

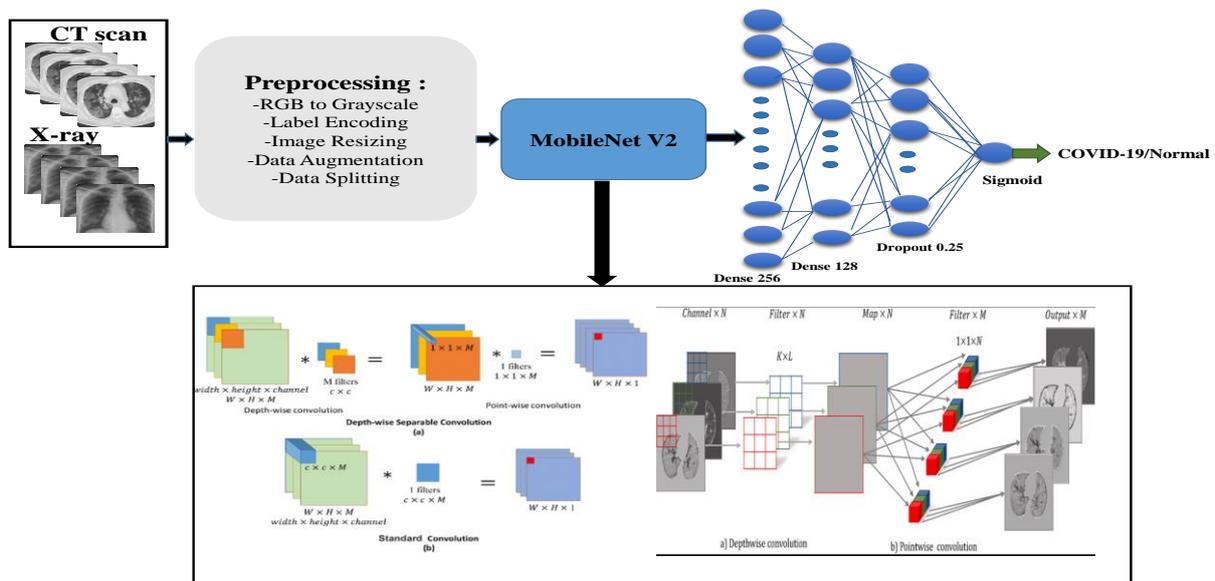


Figure 2.1 Proposed approach with MobileNet_V2

MobileNet is a CNN model that is based on an inverted residual structure used to classify images. The MobileNet model is advantageous in reducing the model size and computation, because it is a lightweight network, which uses depthwise separable and pointwise convolution to deepen the network. The MobileNet_V2 architecture included the residual block with a thin bottleneck layer of a stride of 1 and minimized the size of the layer with a stride of 2 using the ReLu. The ReLu6 activation function is used to assure the efficacy used in low-accuracy situations and enhance the randomness of the model [22]. Table 2-1 illustrate the distribution of the datasets that are used.

Table 2-1 Data Distribution

		Total Data		Training	Validation	Test
1	COVID- 19	CT	3300	2700	300	300
		XR	219	175	44	44
	Normal	CT	3300	2700	300	300
		XR	397	317	80	80
2	COVID- 19	CT	4800	3600	400	800
	Normal	CT	4800	3600	400	800

2.3.1 Hardware

The classification model proposed was deployed on the Google Colab platform, offering free GPU computing for up to 12 hours. We conducted the majority of our experiments utilizing Google Colab, which offers complimentary access to GPU resources and enables the direct writing and execution of Python code in our browser, eliminating the need for intricate configurations. Detailed descriptions of various experimental outcomes, including those from each stage of experimentation, are provided.

2.3.2 Performance Evaluation

For evaluating system, we used different performance measures such as accuracy, sensitivity, and specificity. Table 2-2 shows the different performance metrics.

The performance is assessed using the confusion matrix. Positive and negative patients were classified as COVID19 and normal infections, respectively, in this research. As a result, TP and TN refer to the number of COVID19 and normal infections correctly diagnosed, respectively. The numbers FP and FN represent the number of COVID19 infections that were misdiagnosed and the number of infections that were correctly diagnosed [23].

Table 2-2 Performance metrics for classification

Metrics	Accuracy	Sensitivity	Precision	Specificity
Value	$\frac{TN + TP}{TN + FN + TP + FP}$	$\frac{TP}{TP + FN}$	$\frac{TP}{TP + FP}$	$\frac{TN}{TN + FP}$

2.3.3 An Approach Application

Firstly, we employ data augmentation, a technique used to artificially increase the size of data by generating modified data points from existing ones, particularly when the initial dataset is limited for training purposes [24]. In this work, we augment the number of images to mitigate overfitting issues and enhance the model's performance. Data augmentation was conducted using the following parameters: rescale, shear_range=0.2, horizontal_flip=True, rotation_range=10, width_shift_range=0.2, and height_shift_range=0.2. As a result, the model achieved a testing accuracy of 99%, indicating improved performance.

The proposed model was trained with a variety of hyperparameters to choose the best ones. Table 2-3 shown different experiences. We train our models using TensorFlow.keras and use the standard RMSProp Optimizer, and the best hyperparameters, learning rate: LR= 0.001, loss = binary_crossentropy, batch size: BS= 64, and epochs = 50.

In this case, two experiments are conducted with samples of different sizes and distributions:

2.3.3.1 Experience 1

To train and evaluate our proposed model, we used 3000 confirmed CT scans images with COVID-19, and 3000 without COVID19. A proportion of 90% of dataset were applied for training the model and the rest of 10% were used for the validation set. For the test, 600 unknown images to the systems are added. Test images are presented in the proportions 300 COVID19/300 No-COVID19.

In the case of XR dataset, we used 219 COVID19 and 397 No-COVID19. The dataset is splitted into 80% for train and 20% for the test or validation. In the Table 2-3, we show the results achieved by our model with the two datasets.

Table 2-3 Results achieved by our model (3000c/3000nc)

Tensorflow_Keras_COVID-19							
Image	Acc(%)	Sens(%)	Spec (%)	Prec(%)	AUC(%)	Test (time/image)	Confusion matrix
XR	96.77	100	95.23	90.90	95	0.18 s	$\begin{bmatrix} 40 & 4 \\ 0 & 80 \end{bmatrix}$
CT	99.67	100	99.33	99.33	100	0.03 s	$\begin{bmatrix} 298 & 2 \\ 0 & 300 \end{bmatrix}$

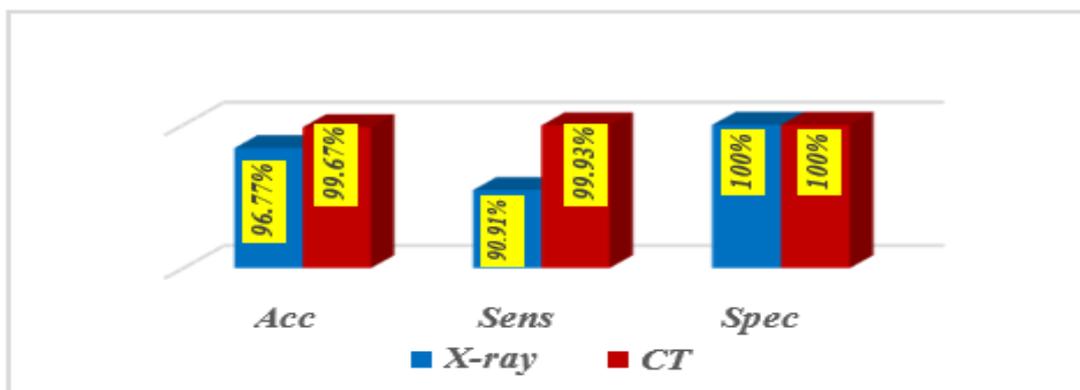
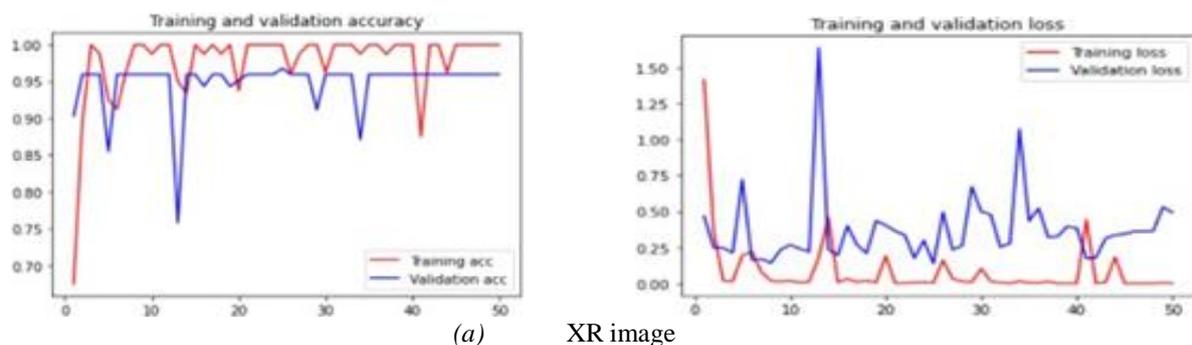


Figure 2.2 Accuracy, Sensitivity and Specificity with Tensorflow_Keras



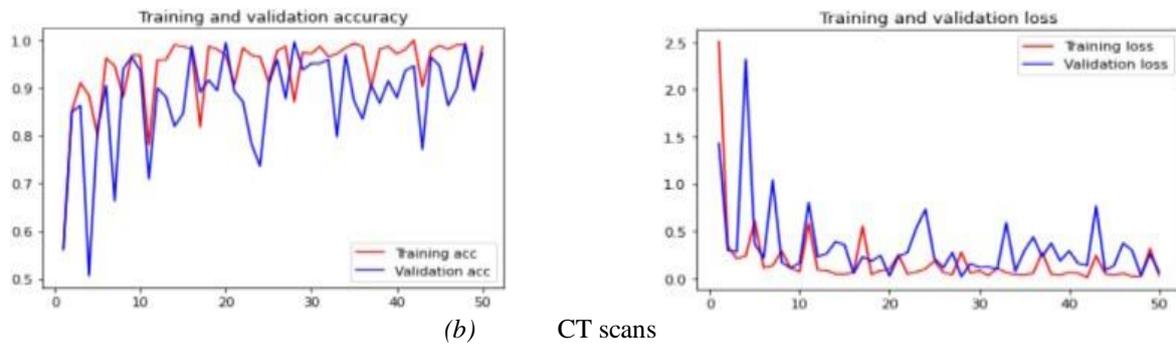


Figure 2.3 Accuracy and loss curve with Tensorflow TL MobileNet_V2 (3000C/3000NC)

Discussion:

The findings indicate that MobileNet_V2 outperforms other models, achieving an accuracy of 96.77% and 99.67%, sensitivity scores of 90.91% and 99.93% for XR and CT images respectively, and specificity scores of 100% for both XR and CT images. Additionally, we attained high sensitivity for CT images and high specificity for XR images.

2.3.3.2 Experience 2

Since Google Colab is free service and is faster than my GPU, we augmented the COVID19 CT dataset to train and evaluate the performance of MobileNet_V2. The dataset comprises 4000 COVID19 positive images and 4000 negative images. For training, 90% of the dataset was utilized, with the remaining 10% reserved for validation. During testing, a random sample of 3% from the training dataset was selected, to which we added 1600 unknown images. Test images were distributed in proportions of 800 COVID19 positive and 800 negative cases.

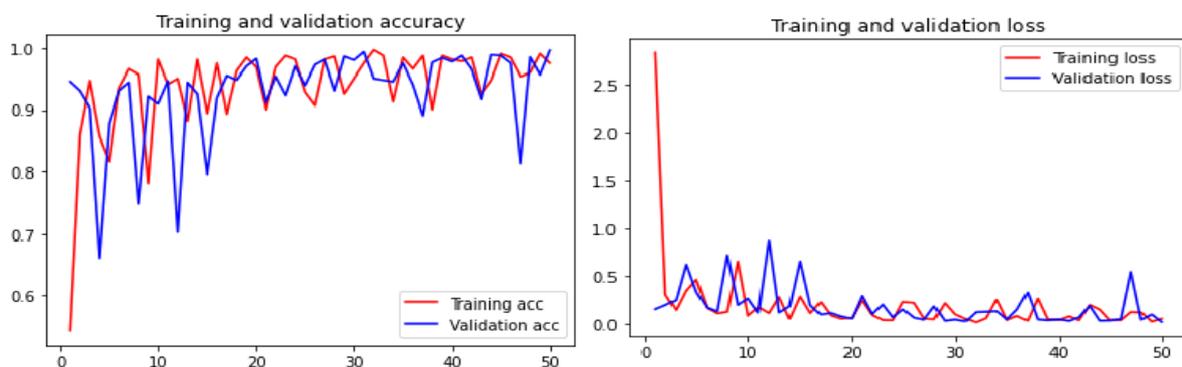


Figure 2.4 Accuracy and loss curve with Tensorflow TL MobileNet_V2 (4000C/4000NC)

Discussion:

The results presented in Table 2-3 indicate that MobileNet_V2 requires 0.18 seconds and 0.03 seconds per image for XR and CT images respectively. The most efficient performance in terms of test time per image is observed with the MobileNet_V2 model for the CT sample consisting of 3000 COVID19 positive and 3000 negative cases, along with 300 unknown images

for testing, resulting in a test time per image of 0.03 seconds (accuracy of 99.67%). However, with a larger sample of unknown images, specifically 800 COVID19 positive and 800 negative cases, the system's test time per image increases to 0.05 seconds (see Table 2-4). Despite this, our proposed framework remains swift for COVID19 monitoring. In scenarios involving individual examination where the number of unidentified images is substantial, the system requires an additional 0.02 seconds per image, which is deemed acceptable, especially considering the accuracy rate in this context is 99.62%.

Table 2-4 Results achieved by our model (4000c/4000nc)

<i>Tensorflow_keras_COVID19</i>							
<i>Image type</i>	<i>Acc (%)</i>	<i>Sen (%)</i>	<i>Spec (%)</i>	<i>Prec (%)</i>	<i>AUC (%)</i>	<i>Confusion matrix</i>	<i>Test time/image</i>
CT	99.62	99.25	100	100	100	$\begin{bmatrix} 400 & 0 \\ 3 & 397 \end{bmatrix}$	0.05 s

Table 2-5 Results of Different Experiments

Hyperparameters	Curves	Accuracy
BS=64, Epochs=50, LR=0.0001		98.73%
BS=64, Epochs=100, LR=0.0001		98.62%
BS=32, Epochs=100, LR=0.001		98.87%
BS=64, Epochs=30, LR=0.001		97.50%

While the new coronavirus has spread around the world, causing a large number of deaths. Numerous studies employing DL techniques have emerged for the detection and classification of COVID19 cases. Pre-trained CNN models have demonstrated superior performance by extracting relevant features to represent images and streamline data. These methods offer significant advantages in terms of performance and learning efficiency. Moreover, conducting experiments across various databases, especially with non-standardized images, may not facilitate direct overall comparisons. However, it contributes to deriving diverse insights and results from the analyses. In Table 2-6 below, we present a comparison of our approach with the recent state of the art.

Table 2-6 Comparison Results with State Of The Art.

Ref	Methods	Modality	Datasets	Accuracy (%)
[25]	CNN	XR	<i>Dataset A</i> : 125 COVID-19/500 Normal <i>Dataset B</i> : 300 COVID-19/300 Normal	<i>Dataset A</i> : 100 <i>Dataset B</i> : 96.06
[26]	Ensemble model (VGGNet, GoogleNet, DenseNet, NASNet)	XR	635 COVID-19/ 7081 Normal	98.58
[27]	VGG16 and ResNet50	XR	3616COVID19/ 1345ViralPneumonia/ 10192Normal	89.34
[28]	CNN and (GNB, SVM,DT, LR,RF)	CT	1252 COVID-19/1230 Normal	99.73
Ours	MobileNet_V2	CT & XR	CT : <i>Dataset 1 :</i> 3600 COVID19/3600 Normal <i>Dataset 2 :</i> 4800 COVID19/4800 Normal XR: 219 COVID-19/397 Normal	99.67 <i>Ttest/image=0.03s</i> 99.62 <i>Ttest/image=0.05s</i> 96.77 <i>Ttest/image=0.18s</i>

Results presented in Table 2-6 compare the study model with other recent works approaches. The proposed approach confirms that the MobileNetV_2 gives better results in terms of

performance than other state-of-the-art methods, with an accuracy score of 99.67% and $T_{\text{test}} = 0.03\text{s}$ for the best case.

2.4 Gabor Descriptor and TL in COVID19 Classification for Diagnosis.

Our work introduces a DL-based method for detecting COVID19 infection in chest CT scans and X-ray images. We employ the Keras-TensorFlow architecture, specifically utilizing VGG16 for the classification task. Feature extraction is performed using Gabor Wavelet (GW). The TL technique with VGG16 is employed in the classification process for COVID19 detection.

We are interested in a binary classification, a topic that has been widely explored by researchers. In this work, we limit our scope to methods that are both simple and effective.

2.4.1 Proposed Approach based Gabor Descriptor and TL

In the case of this approach with VGG16, we are interested in the contribution of the Gabor filter associated with the proposed architecture.

2.4.1.1 Preprocessing

Prior to feeding the data into neural networks, it is advisable to ensure uniformity in data size and the number of segments, which invariably enhances the accuracy and balance of the database. Therefore, the data is resized to 224×224 , maintaining consistency across original images (training, testing, and evaluation). Additionally, Gabor Wavelet filtering is applied as part of the preprocessing stage.

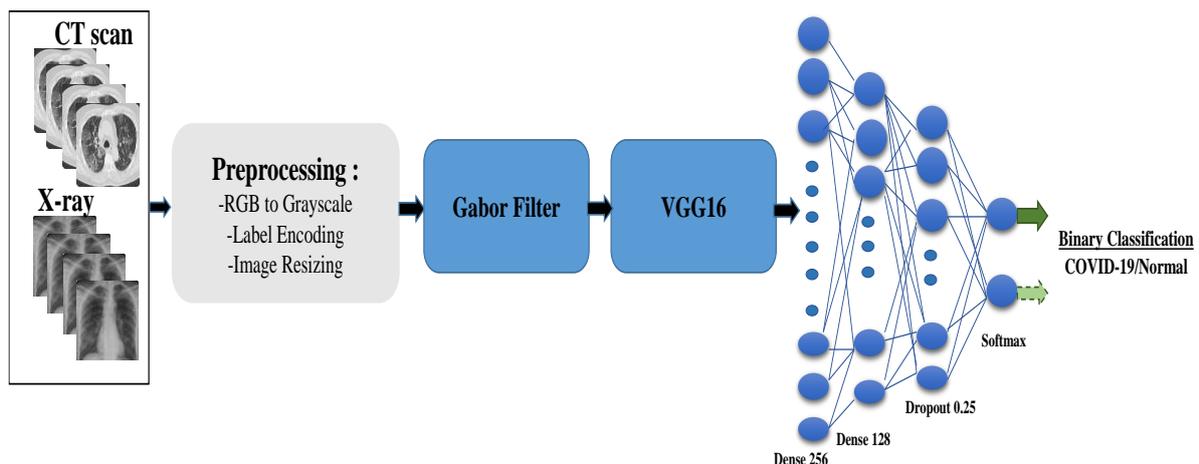


Figure 2.5 Proposed approach with VGG16

The Gabor wavelet [29] is a linear filter which impulse response is defined by a harmonic function multiplied by a Gaussian function. This filter is used to detect line endings and edge borders over multiple scales and with different orientations. The Gabor wavelet is defined as Eq.1 and Eq.2:

$$\psi(z) = \frac{P_{u,r}^2}{\sigma^2} \exp\left(\frac{P_{u,r}^2 z^2}{2\sigma^2}\right) \left[\exp(iP_{u,r}z) \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (2.1)$$

Where $z = (x, y)$, u and r define the orientation and scale of the Gabor wavelet, respectively.

$$P_{u,r} \text{ is defined as follows: } P_{u,r} = P_r e^{i\phi_r} \quad (2.2)$$

$$\text{Where } P_r = P_{max}/f^r \quad \text{and} \quad \phi_u = \frac{\pi u}{8} .$$

P_{max} is the maximum frequency, and f is the spacing factor between kernels in the frequency domain.

Gabor Filter and Parameter Analysis

The `cv2.getGaborKernel()` function generates a Gabor filter, a widely used tool in image processing for texture analysis, feature extraction, and edge detection.

Python code:

```
g_kernel = cv2.getGaborKernel((5, 5), 15.0, np.pi / 4, 3.0, 0.5, 0, ktype=cv2.CV_32F)
```

Each parameter in this function plays a critical role in defining the behavior of the Gabor filter:

Kernel Size (5, 5): Specifies the dimensions of the filter (width \times height), in this case, a 5×5 matrix. A larger kernel captures broader structures, while a smaller kernel focuses on fine details.

Sigma (15.0): Represents the standard deviation of the Gaussian envelope, controlling the spread of the filter. A higher sigma value results in a more dispersed filter response, capturing broader image features.

Theta ($\pi/4$ radians or 45°): Defines the orientation of the sinusoidal component. A value of $\pi/4$ radians means the filter is oriented diagonally, making it sensitive to edges along this direction.

Lambda (3.0): Represents the wavelength of the sinusoidal wave, measured in pixels. A smaller wavelength captures finer textures, while a larger wavelength captures coarser structures.

Gamma (0.5): Defines the spatial aspect ratio, determining the elongation of the Gaussian envelope. A gamma value of 0.5 results in an elliptical shape, influencing the directional sensitivity of the filter.

Psi (0 radians): Specifies the phase offset of the sinusoidal wave. A value of 0 ensures that the wave starts from zero phase.

Kernel Type (cv2.CV_32F): Defines the data type of the kernel, where cv2.CV_32F represents a 32-bit floating-point format, ensuring high precision in computations.

In our model, we perform data augmentation by setting the random image rotation setting to 15 degrees clockwise or counter clockwise. In this work, we use VGG16 technique for TL. VGG16 model is developed by the Visual Graphics Group (VGG) at Oxford and was described in the 2014 [30].

2.4.2 Keras-Tensorflow with TL VGG16 Implementation

Training methods: VGG16 TL is trained using Adaptive Moment learning rate (Adam) solver after applying data augmentation to increase the number of the dataset and to avoid overfitting. The *Keras* model is used in COVID19 detection method for X-ray and CT image. The model is trained with the following hyperparameters: *Adam* optimization method with *Learning Rate* (0.001), *Batch Size* (16, 8, 32, and 64) and *Epochs* (30, 50, 80, 100, 150, and 200), *Loss*: binary cross_entropy.

To train and evaluate the system, we collect the COVID19-CT and X-ray dataset:

COVID19-CT dataset contains **3000 positive CT scans** with clinical findings of COVID19, and **3000 negative images** without findings of COVID19. We have used 90% of CTscans images for training and the rest of the 10% for validation.

COVID19-X-ray dataset contains 219 COVID19 and 397 Non-COVID19. We have used 80% of X-ray images for training and the rest of the 20% for validation.

Table 2-7 Results obtained with VGG16.

Image s	VGG16 with Gabor				Conf. Matrix	VGG16 without Gabor				Conf. Matrix
	Acc (%)	Sens (%)	Spec (%)	Prec (%)		Acc (%)	Sens (%)	Prec	Spec (%)	
X-ray	95.28	91.11	97.56	91.11	$\begin{bmatrix} 41 & 4 \\ 2 & 80 \end{bmatrix}$	99.21	100	100	98.78	$\begin{bmatrix} 45 & 0 \\ 1 & 81 \end{bmatrix}$
CT	98.50	99.33	97.67	99.33	$\begin{bmatrix} 298 & 2 \\ 7 & 293 \end{bmatrix}$	99.83	99.67	99.66	100	$\begin{bmatrix} 299 & 1 \\ 0 & 300 \end{bmatrix}$

Discussion

The proposed method is applied to datasets comprising Chest X-ray and CT images. In the first scenario, VGG16 TL is utilized for the classification of COVID19 and nonCOVID19 cases. In the second scenario, Gabor filters are employed to extract features for COVID19

classification. Results indicate that the performance of the VGG16 without Gabor is the highest accuracy and sensitivity/specificity (COVID19/NoCOVID19 portion predicted correctly). From the results, we can see that the test time with Gabor gives better performance for X-ray images and shows no improvement for CT images since the test time remains the same in both cases. In addition, it is noticed that the obtained sensitivity and specificity for COVID19 class is remarkably higher (i.e., 100%) for both VGG16 models. From the obtained results, we can infer that proposed model can detect COVID19 patients (i.e. true positive) and non-covid19 (i.e. true negative) with high accuracy.

2.5 Image Enhancement and Hybrid Deep Features for COVID19 Diagnosis

In this work, we propose an effective method for detecting COVID19 using TL model called MobileNetV2. We apply histogram equalization (HE) for image enhancement, as well as Local Binary Pattern (LBP) and Gabor descriptor techniques on CT scans and chest X-ray images to improve image contrast and extract features. In the classification phase, we use a fine-tuning model to detect COVID19.

Our contribution is based on the effect of pre-processing the dataset to train the TL to accurately classification of COVID19 disease from a CT scans and Chest X-Rays images. We have used different pre-processing techniques to choose the appropriate and precise ones. The contributions of our work are presented as follows:

- ✓ Image enhancement using histogram equalization;
- ✓ Details information and image texture extraction using LBP and Gabor filters;
- ✓ Deep Features extraction using MobileNetV2 TL model;
- ✓ Fine Tuning model for COVID19 classification.

2.5.1 Proposed approach in COVID19 Classification

In this work, the MobileNetV2 is used for deep features extraction. The MobileNetV2 is based on depth wise separable and point wise convolution. The depth wise separable convolution produces the same output as the standard convolution, only it is more effective because it minimizes the number of parameters implicated in the operation. MobileNetV2 includes 19 inverted residual bottleneck layers after the first convolution layer. A residual block connects the first and ending of a convolution block with a skip connection with the objective of converting information to the deeper layer of the network [22].

2.5.1.1 Preprocessing

1) Data augmentation

To increase the size of dataset and avoid overfitting, we applied the data augmentation techniques such as, rotation, scale. The novel images can be created by using changes on the ancient tones.

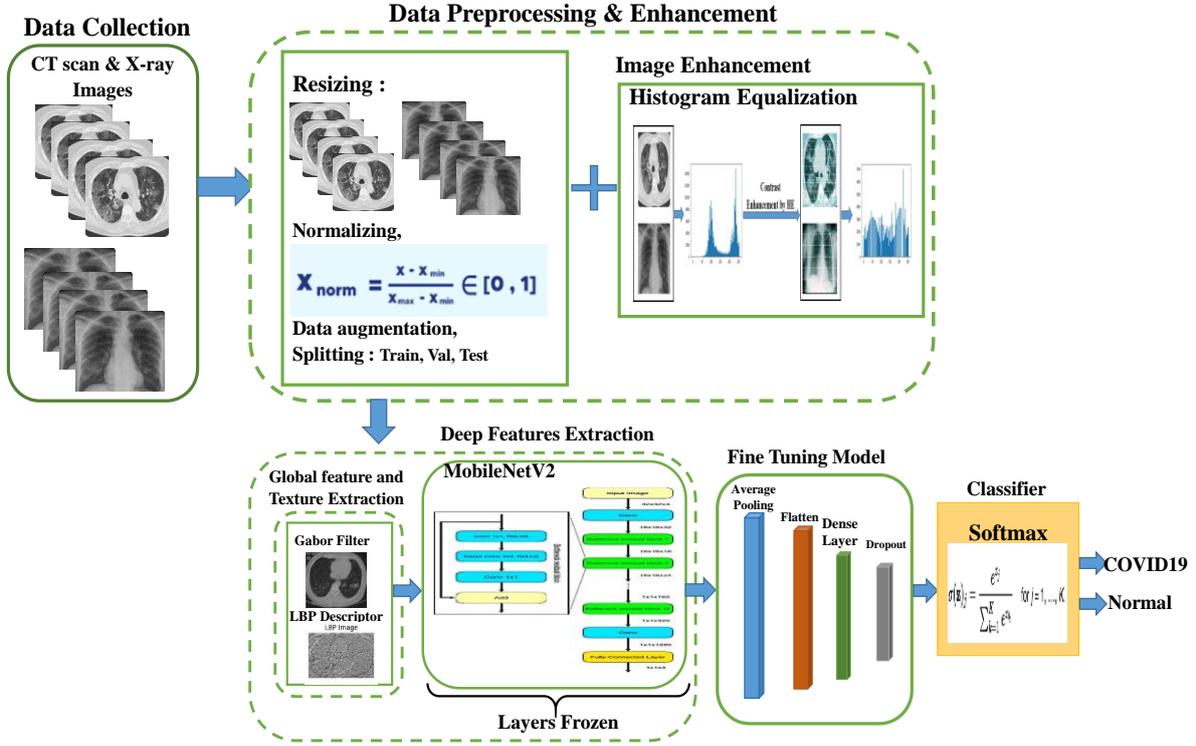


Figure 2.6 Flowchart of proposed approach in COVID19 classification

2) Histogram Equalization

To enhance the contrast of images, HE is used. The HE technique is important for image processing because it increments the contrast of the image to improve the goodness of the images [31]. HE is a no linear expansion that calculates the probability mass function in addition to the accumulative distribution function of image pixels, and maps new grey-level values for these pixels after distributing the intensities of the pixels across the histogram of the image. The intensity of the image is represented in range $[0, P - 1]$ by the function as follows:

$$g(s_i) = k_i \quad (2.3)$$

k_i is the count of pixels with i^{th} intensity value s_i

The histogram is normalized with the following equation:

$$f(s_i) = \frac{k_i}{H*W} \quad (2.4)$$

Where the H and W are the height and the width.

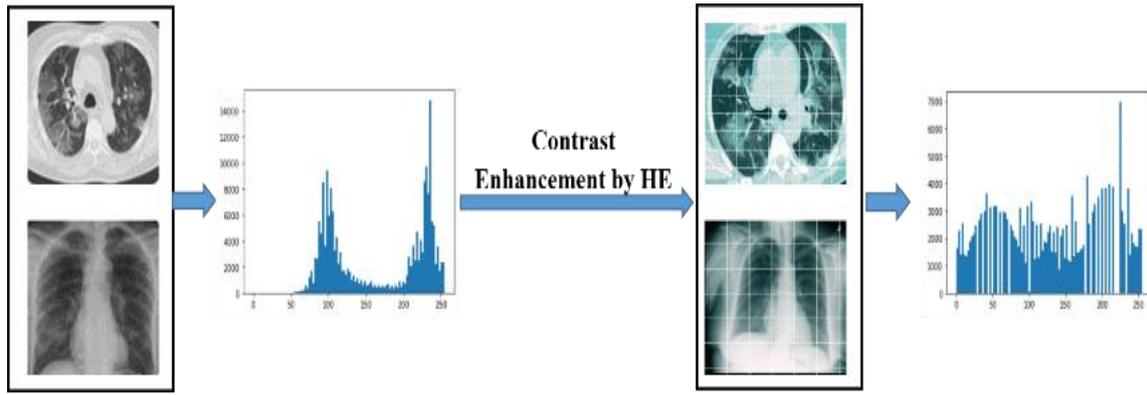


Figure 2.7 Image Enhancement by HE

3) Gabor Filter

The Gabor filter is a linear function that is applied on images for texture analysis and to find the different orientations around the region of interest. The Gabor filter is used to extract the features from analyzed images [32] (see Fig 2.8):

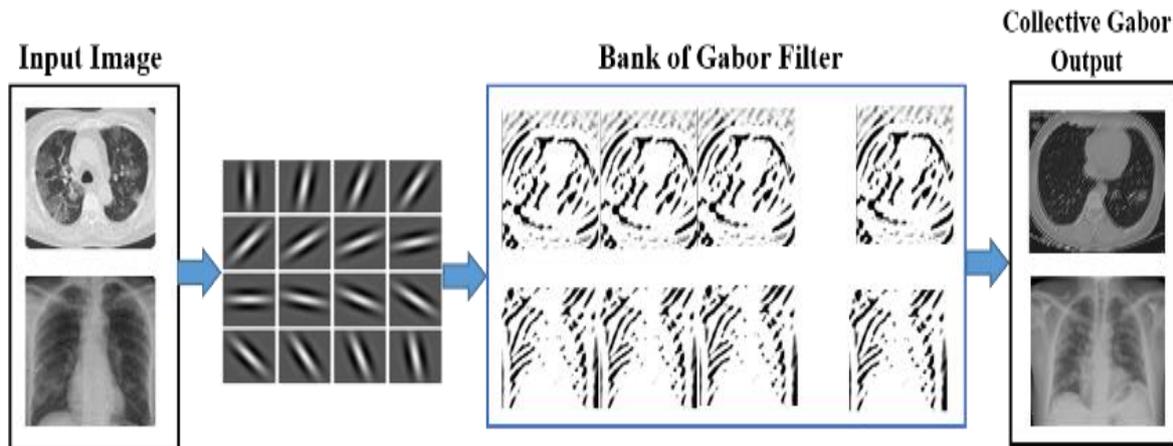


Figure 2.8 Gabor filter

4) Local Binary Pattern

The LBP is the one of several local image descriptors. It is method that is describe every pixel's neighborhood through a binary code that is gained by initial convolving the image with a specific group of linear filters and then binarizing the filter responses [33]. The proposed method is based on the LBP for COVID19 classification. The LBP creates binary labels to pixels using thresholding the neighbourhood pixels together with the centric value [34]. Furthermore, for the centre pixel p, I and its neighboring pixels N_i , a decimal value is converted to it (see Fig 2.9).

$$D = \sum_{i=1}^s 2^{i-1} I(P - N_i) \quad \text{where, } I(P, N_i) = \begin{cases} 1 & \text{if } P > N_i \\ 0 & \text{otherwise} \end{cases} \quad (2.5)$$

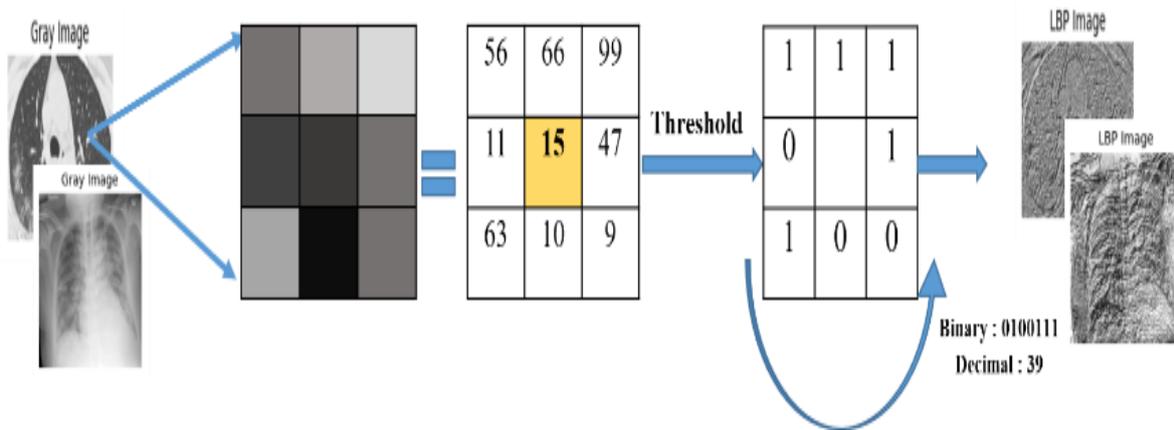


Figure 2.9 LBP Descriptor

5) Fine Tuning Model

Fine-tuning refers to the process of adjusting a model's parameters to improve its performance. In this work, fine-tuning involved the implementation of an average pooling layer, a flatten layer, a dense layer, a dropout layer, and the final layer which utilizes the softmax function for COVID19 classification.

2.5.2 Experimentation

2.5.2.1 Dataset

The dataset that was used is public available in [35] and [36]. They contain **2700 COVID19/2700 Normal CT scans** images and **3608 COVID19/3616 Normal Chest X-Rays images**. The ratio of 80% was for training the model and 20% for validation. All the images were resized to 128x128 to feed the preprocessing block before send it to the DL block.

2.5.2.2 Proposed Model Implementation in COVID19 classification

In this work, three tasks were implemented: the first task preprocessing using i) resizing, ii) HE, and iii) Data Augmentation. The LBP descriptor and Gabor filter were used for features extraction in second task. In the third task, deep features extraction using TL model called MobileNetV2 and the fine tuning, where the fully connected layers removed and replaced by new model composed by average pooling layer, Flatten, Dense layer, Dropout layer and the last layer is Softmax function to classify COVID19. In this section, we present the different experiments of the MobileNetV2 with the different preprocessing techniques and compare results with state-of-the-art studies. The proposed model is implemented on the COVID19 dataset using various hyperparameters as follows: *Adam optimizer*. *Learning Rate* = 0.001, *Batch Size* = 16, an *Epoch* = 50.

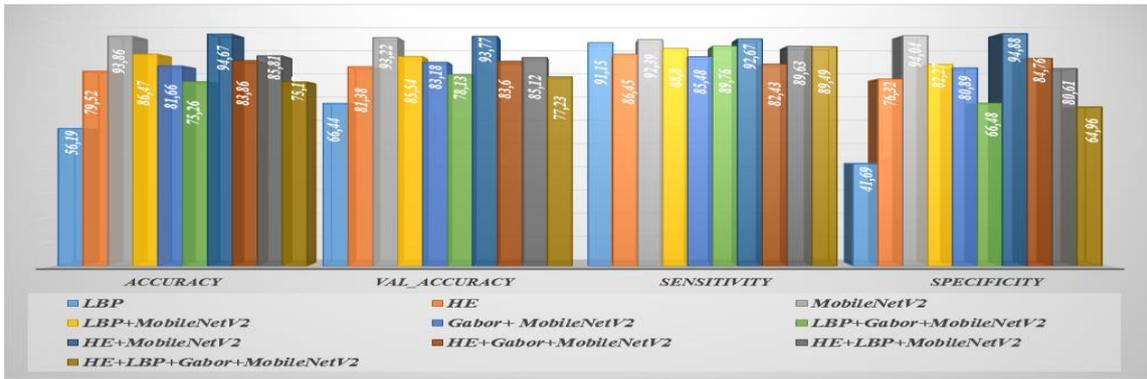
Table 2-8 Results for HE and different descriptors with MobileNetV2 (CT scans images).

<i>Methods</i>	<i>Acc. (%)</i>	<i>Val_Acc. (%)</i>	<i>Loss</i>	<i>Val_Loss</i>	<i>Sens. (%)</i>	<i>Spec. (%)</i>	<i>AUC. (%)</i>	<i>Confusion Matrix</i>
<i>MobileNetV2</i>	100	99.72	0.0003	0.009	99.44	100	100	$\begin{bmatrix} 537 & 3 \\ 0 & 540 \end{bmatrix}$
<i>Gabor + Fine Tuning</i>	99.19	99.72	0.037	0.019	99.63	99.81	100	$\begin{bmatrix} 538 & 2 \\ 1 & 539 \end{bmatrix}$
<i>HE + Fine Tuning</i>	98.40	99.44	0.054	0.032	99.63	99.26	99	$\begin{bmatrix} 538 & 2 \\ 4 & 536 \end{bmatrix}$
<i>LBP+ MobileNetV2</i>	95.53	91.48	0.119	0.239	90	92.96	91	$\begin{bmatrix} 486 & 54 \\ 38 & 502 \end{bmatrix}$
<i>Gabor+ MobileNetV2</i>	98.22	98.98	0.039	0.032	98.89	99.07	99	$\begin{bmatrix} 534 & 6 \\ 5 & 535 \end{bmatrix}$
<i>LBP+Gabor+ MobileNetV2</i>	83.17	87.04	0.364	0.306	84.26	89.81	87	$\begin{bmatrix} 455 & 85 \\ 55 & 485 \end{bmatrix}$
<i>HE+ MobileNetV2</i>	99.98	99.63	0.0009	0.008	99.26	100	100	$\begin{bmatrix} 536 & 4 \\ 0 & 540 \end{bmatrix}$
<i>HE+Gabor+ MobileNetV2</i>	98.59	98.24	0.037	0.043	99.26	97.22	98	$\begin{bmatrix} 536 & 4 \\ 15 & 525 \end{bmatrix}$
<i>HE+LBP+ MobileNetV2</i>	95.35	93.89	0.128	0.198	93.89	93.89	94	$\begin{bmatrix} 507 & 33 \\ 33 & 507 \end{bmatrix}$
<i>HE+LBP+ Gabor+ MobileNetV2</i>	84.49	86.30	0.345	0.300	79.26	93.33	86	$\begin{bmatrix} 428 & 112 \\ 36 & 504 \end{bmatrix}$

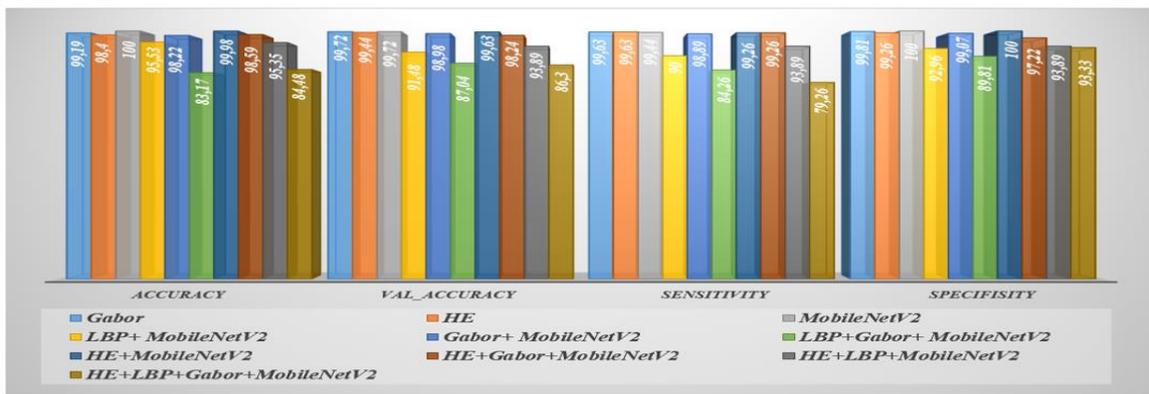
Discussion: The results demonstrate that the *MobileNetV2* model achieved high performance. HE enhancement techniques produced clearer images, resulting in good results as shown in the table 2-8. The combination of HE + Gabor for preprocessing images yielded better performance, while *HE+MobileNetV2* also resulted in good performance. Additionally, when used with Gabor or HE, the LBP descriptor demonstrated a good effect on the curves. Gabor is a competitive method and performs well on CT scan images where detail is significant.

Table 2-9 Results for HE and different descriptors with MobileNetV2 (X-Rays images).

<i>Methods</i>	<i>Acc. (%)</i>	<i>Val_Acc. (%)</i>	<i>Loss</i>	<i>Val_Loss</i>	<i>Sens. (%)</i>	<i>Spec. (%)</i>	<i>AUC. (%)</i>	<i>Confusion Matrix</i>
<i>MobileNetV2</i>	93.86	93.22	0.144	0.160	92.39	94.04	93	$\begin{bmatrix} 668 & 55 \\ 43 & 679 \end{bmatrix}$
<i>LBP</i>	56.19	66.44	0.686	0.684	91.15	41.69	66	$\begin{bmatrix} 659 & 64 \\ 421 & 301 \end{bmatrix}$
<i>HE</i>	79.52	81.38	0.435	0.390	86.45	76.32	81	$\begin{bmatrix} 625 & 98 \\ 171 & 551 \end{bmatrix}$
<i>LBP+ MobileNetV2</i>	86.47	85.54	0.306	0.325	88.80	82.27	86	$\begin{bmatrix} 642 & 81 \\ 128 & 594 \end{bmatrix}$
<i>Gabor+ MobileNetV2</i>	81.66	83.18	0.368	0.376	85.48	80.89	83	$\begin{bmatrix} 618 & 105 \\ 138 & 584 \end{bmatrix}$
<i>LBP+Gabor+ MobileNetV2</i>	75.26	78.13	0.473	0.441	89.76	66.48	78	$\begin{bmatrix} 649 & 74 \\ 242 & 480 \end{bmatrix}$
<i>HE+MobileNetV2</i>	94.67	93.77	0.128	0.193	92.67	94.88	94	$\begin{bmatrix} 670 & 53 \\ 37 & 685 \end{bmatrix}$
<i>HE+Gabor+ MobileNetV2</i>	83.86	83.60	0.349	0.353	82.43	84.76	84	$\begin{bmatrix} 596 & 127 \\ 110 & 612 \end{bmatrix}$
<i>HE+LBP+ MobileNetV2</i>	85.81	85.12	0.312	0.334	89.63	80.61	85	$\begin{bmatrix} 648 & 75 \\ 140 & 582 \end{bmatrix}$
<i>HE+LBP+Gabor+ MobileNetV2</i>	75.17	77.23	0.478	0.463	89.49	64.96	77	$\begin{bmatrix} 647 & 76 \\ 253 & 469 \end{bmatrix}$



a) X-ray



b) CT scan

Figure 2.10 Results of different models on X-Rays and CT scan images

For CT scan images, the LBP+MobileNetV2 model achieved satisfactory results compared to Chest X-ray images. *MobileNetV2* is a good candidate for this category of images which is noisy and where the texture information is the most requested. The best model to remember in the case of X-Rays images remains *HE+MobileNetV2*. The proposed method for COVID19 classification uses LBP for feature extraction, which captures the local texture information of images. LBP extracts relevant features from CT scans and Chest X-rays to differentiate COVID19 from other respiratory diseases. LBP is not used in the classification part; instead, a fine-tuning or TL model. LBP plays a significant role in feature extraction, but not in image classification.

2.5.3 Comparison results with State of the Art

The table 2-10 shows a comparison with different recent works on image classification and detection of COVID19.

Table 2-10 Comparison of Gabor Descriptor and TL with recent works

<i>Applications</i>	<i>References</i>	<i>Approachs</i>	<i>Size Datasets</i>	<i>N.C</i>	<i>Acc. (%)</i>
<i>Application 1</i> <i>Chest X-Rays</i>	<i>HE and different TL</i>				
	David et al. [37]	HE+ResNet-34	95 COVID19, NoCOVID19: (19BP,81N,41VP)	2	97.81
	Rahman, et al. [38]	HE+ ChexNet	3616 COVID19 8851 Normal 6012NoCOVID19	3	94.34
	<i>Proposed Method</i>	<i>HE+MobileNetV2</i>	<i>3616 COVID19</i> <i>3608 Normal</i>	2	94.67
	<i>Gabor and different TL</i>				
	Barshooi et al [32]	Gabor + VGG19	360 COVID19 4200 No COVID19	2	86
		Gabor + AlexNet		2	84.50
		Gabor +GoogLeNet		2	80.50
<i>Proposed Method</i>	<i>Gabor+ MobileNetV2</i>	<i>3616 COVID19</i> <i>3608 Normal</i>	2	81.66	
<i>Application 2</i> <i>CT Scan</i>	<i>HE and different TL</i>				
	Lawton et al [39]	HE +VGG19	1252 COVID19 1230NoCOVID19	2	90.90
		HE + MobileNetV2		2	91.31
		HE +DenseNet-201		2	93.93
	<i>Proposed Method</i>	<i>HE+MobileNetV2</i>	<i>2700 COVID19</i> <i>2700 Normal</i>	2	99.98
	<i>Gabor and different TL</i>				
	M. Ketfi et al [29]	Gabor + VGG16	3000 COVID19 3000 Normal	2	98.50
<i>Proposed Method</i>	<i>Gabor+ MobileNetV2</i> <i>Gabor+Fine Tuning</i>	<i>2700 COVID19</i> <i>2700 Normal</i>	2	98.22 99.19	
<i>Chest X-Rays</i> <i>& CT Scan</i>	<i>Hybrid Deep Descriptor</i>				
	M.Kaya et al. [40]	D ³ SENET DarkNet53, DarkNet19, DenseNet201, SqueezeNet, and EfficientNetb0	<u>X-Rays images</u> 1210 COVID19 1409 Normal 1345 VP	3	98.78
	El Shenbary et al.[41]	Fusion Classic Descriptor (DWT+PCA) with AlexNet or GoogLeNet	COVID, lung opacity, normal and VP	4	97.6
	<i>Proposed Method</i>	<i>HE+Gabor+MobileNetV2</i> <i>HE+MobileNetV2</i>	<u>X-Rays images</u> <i>2700 COVID19</i> <i>2700 Normal</i> <u>CT scans images</u>	2	98.59 99.98

The hybrid descriptors techniques used in this study can be applied to other medical imaging modalities. In fact, the techniques used in this study are general image processing and feature extraction techniques that can be applied to any medical imaging modality that requires the detection and classification of diseases.

Implementing a binary classification for COVID19 in a real-world clinical setting involves several steps, including data acquisition, model training, testing, and deployment. After creating and training the model, it can be utilized for real-time classification of lung images. To make the model easily accessible, the model can be integrated into a desktop or mobile application. If the model is to be implemented in a real-time application, it is crucial to ensure that it can process the images at an appropriate speed. However, it is important to keep

in mind that the processing time may vary depending on factors such as the image size, the model complexity, and the processing power of the device being used. Overall, it is important to collaborate with healthcare professionals, regulatory bodies, and other stakeholders to address the challenges and ensure the successful deployment of the model.

2.6 Conclusion

In this chapter, we delved into the technologies employed for COVID19 classification and evaluated the performance of our proposed model on chest XR and CT images. The process involves initial feature extraction followed by the meta-architecture for classification, utilizing the MobileNet_V2 CNN architecture for feature extraction.

Our investigation draws data from multiple sources for X-ray and CT images to implement an architecture for classifying COVID19-infected patients. We utilize the VGG16 CNN architecture, leveraging the Tensorflow and Keras library with Python. Despite Gabor filter usage not yielding optimal accuracy, sensitivity, and specificity for VGG16 architectures, its implementation notably reduces test time/image, enhancing patient comfort while maintaining acceptable accuracy parameters.

Our work demonstrates VGG16's robust performance on both X-ray and CT images, achieving accuracies of 99.21% and 99.83% respectively, with a time cost of 0.870 seconds per image for X-rays and 0.872 seconds for CT scans. Notably, our system maintained efficiency in patient follow-up scenarios but requires additional time for COVID19 screening when dealing with a larger volume of unknown images.

Furthermore, we presented an effective model for COVID19 classification, leveraging enhancement techniques such as HE and deep features extraction via Gabor filters and LBP descriptors, alongside TL MobileNetV2. Our approach achieved optimal accuracy, particularly with MobileNetV2 and HE+MobileNetV2 combinations. Future work involves exploring additional deep descriptors.

Despite its limitations, our proposed framework shows promise for the binary classification of lung images. We envision future endeavors aimed at developing a standardized system capable of classifying various diseases using diverse TL models, including multi-classification and datasets from different organs. Further studies and comparisons are warranted to comprehensively evaluate its performance. The techniques used at the different stages of disease classification are detailed in the next Chapters.

3 Chapter: Transfer Learning Fusion and Stacked Auto-Encoders for Viral Lung Disease Classification

3.1 Introduction

High dimensional data is found in a variety of smaller dimensions. So-called "intrinsic dimensional" data is very complicated. Typically, it is not easy to capture such structures using a simple parametric model, such as Principal Component Analysis (PCA). All these limitations of traditional MDS (Dimensionality Methods) can be solved by ML methods that can be the solution. We thus affirm that to be more effective in viral lung disease classification/detection, Auto Encoder (AE) based DL framework proves to be a good candidate and may be classification-specific in nature. In other words, the performance of recognition algorithms based on AE can be a solution if the learned representations integrate the class information [25]. All these reasons motivate us to push our interest in this track which is lung disease classification based on AEs.

The fields of application of AE are: AE denoising, reconstruction, classification and data reduction. One of the unsupervised methods is Stacked Denoising Auto Encoders (SDAE) introduced by Vincent et al [42]. The denoising AE learns from a distributed representation (encoding) of a set of data, then reconstructs the data from the encoder (decoding) by removing the noise from the image. The new output is a compact or fragmented representation serving as input to the next AE or other ML Automatic encoders are advantageous for dimensionality reduction and cluster analysis. The objective of this chapter endeavor is to identify an effective model for the classification of multiple viral respiratory diseases, encompassing COVID19. The feature extraction phase from medical images constitutes a formidable challenge in achieving optimal disease classification outcomes. In this work, a selection of the best models among several popular TL models is realized. The concatenation of the best models for better features extraction is used; the DL methods for deep features extraction and deep data reduction were applied for an optimal classification. This work includes two studies, the first was applied to binary classification (COVID19/Normal) and the second is concerned with multi-classification (COVID19/Normal/VPneumonia).

TL is mainly applied for two purposes: employing a pre-trained network as a feature extractor, where the last layer of the pre-trained model architecture is changed out with a fully connected layer with the same number of classes as the new dataset during performing TL, and fine-tuning a pre-trained network on medical data [43]. These models can be used directly in making predictions on new tasks or integrated into the process of training a new model. In addition, feature models have also been extracted using a CNN based approach. For feature learning and extraction, basic architectures such as DenseNet201, VGG16, VGG19, and GoogLeNet have been modified to detect no linearity in the lung and classify the different diseases such as COVID19 and Pneumonia. Several feature extraction techniques have been used to improve the accuracy of diseases detection. Concatenation method between TL models is the new techniques to enhance the performances of pre-trained models for classification tasks. In [44] a parallel deep feature extraction approach based on TL models and concatenation-based models were used to achieve a high accuracy.

For deep feature extraction, the SAEs are known for their power in these tasks. The SAE is a Deep Neural Network (DNN) architecture that has been used in varieties of applications. Benyelles et al. [45] presented a framework based on SAEs for the recognition of the disease characteristics in medical images to detect COVID19 disease from CXR and CT scan images. The paper referenced in citation [46] presented a proposal to utilize features extracted from both TL models and Convolutional Neural Networks (CNN) developed from scratch to classify and detect COVID19. The objective of our work is to extract profound features by utilizing a fusion approach involving the most effective TL models, subsequently associating them with either an Encoder (E) or Stacked Encoder (SE) to reduce data. Additionally, we employ Autoencoder (AE) or Stacked Autoencoder (SAE) methods to attain high performance. The experiments in this study involve an investigation of the impact of depth on unsupervised models, and an evaluation of data reduction outcomes. Our primary objective is to develop a robust and effective system that can diagnose COVID19 in binary classification and differentiate it from other pathologies. In a prior study [40], the VGG16 TL architecture was employed to classify images from CXR and CT scans. Initially, we employed Gabor filter for feature extraction in our study. However, the Gabor filter did not provide satisfactory accuracy levels. Consequently, we proposed a novel approach based on the fusion of TL selection with E/SE or AE/SAE architectures for the classification of CXR and CT scans images, aiming to detect pulmonary viral pathologies.

During the current pandemic, recent works have focused on the detection of COVID19 from chest CXR and CT images and several methods have applied based on DL technologies

using advanced CNN architectures[43]. Several studies have been conducted in the literature to develop DL-based models for COVID19 detection using chest X-ray and CT images. Different DL architectures, such as VGG16, ResNet-50, DenseNet121, MobileNet, NASNet, Xception, and EfficientNet, have been used for COVID19 detection. TL and MTSSL have been proposed as effective techniques for COVID-19 detection. Hybrid and deep feature extraction techniques such as SAE, AAE, EVAE, and CovNet have been implemented to improve classification accuracy. The reported accuracies for COVID19 detection in these studies range from 80.19% to 99.75%, with the highest accuracy achieved using ResNet-50 and DeepCovNet models. Overall, these studies show that DL-based models have the potential to aid in COVID19 diagnosis and screening.

Other recent research employed CT scans to discover the most critical virus features and correctly classify the infection[44].

Literature indicate that VGG16 architecture performed better compared to the other architectures. Additionally, deep transfer learning has been widely used in literature models to identify pneumonia from CXR and CT images. In spite of the numerous studies that have utilized DL for feature extraction and classification, there is still need for improvement in these stages. To address this issue, a promising technique called stacked autoencoder (SAE) has been proposed. SAE improves the gradient and enhances the training of convolution layers by employing separate encoding and decoding networks or separate training for each convolution layer. This method can reduce the pressure of classification by utilizing unsupervised DL. In addition, the autoencoder (AE) model, which learns about the dataset's features during unsupervised training, can be further enhanced by initializing the model with the learned weights during the decoding stage. Table 3-1 shows summary of the studies in literature.

Table 3-1 Summary of the studies in literature

Authors	Methods	Datasets		N C	Results (%)			
		CXR	CT		Acc.	Sens.	Spec.	F1-score.
DL (COVID-19(C-19)/Normal(N))								
Lahsaini et al [44]	DenseNet201 + GradCAM	–	1868 C-19 3118 N	2	98.8	98.54	99.22	99.02
S.H. Khan et al [47]	DHL & DBHL	3224 C-19, 3224 N	–	2	98.53	99	–	98
L. Mohamed et al.[48]	CNN Bayesian Optimizer	3616 C-19, 3616 N, 3616 P	–	3	96	–	–	–
Kaya et al. [49]	DarkNet53,DarkNet19, DenseNet201, SqueezeNet, and EfficientNetb0, SVM	–	3616C-19, 1345VP, 10191N	3	98.78	–	–	–
Benyelles et al. [50]	SAE	760 (CXR and CT): C-19 Other Pathologies	–	10	80	–	–	–
TL								
Nasir et al [51]	VGG16, ResNet50, InceptionResNetV2, MobileNetV2.	–	342 C-19, 193 NC-19	2	97.8	–	–	–
Madhavan et al.[52]	ResNet-50	150 C-19 Vs 1583 N, 150 C-19 Vs (2798 BP, 1480 VP)	–	2	98.4	–	–	98.1
				2	96.2	–	–	94.9
Khan, E, et al. [53]	EfficientNetB1, NasNetMobile, MobileNetV2	2473 C-19, 6012 Opacity, 10192 N, 1345 VP	–	4	96.13	–	–	–
A.Bhattacharya et al.[54]	VGG19 BRISK_RF	930 C-19, 1583N, 4273 P	–	3	96.60	95.0	97.4	–
M. Loey et al. [55]	AlexNet, VGGNet16, VGGNet19,GoogLeNet, ResNet50	–	345 C-19, 397 N	2	82.91	77.66	87.62	–
Wang, et al.[56]	Xception+SVM	565 N, 537 C-19	–	2	99.33	99.27	99.38	–
El Gannour, et al.[57]	<u>Concatenate</u> : 1) Xception, Inception V3 2) Resnet50 V2 MobileNet V2	2249 N, 3500 T, 1400 C-19,	–	–	99.80	99.71	99.93	–
				4				

		3250 VP,			99.71	99.67	99.88	–
Khan et al. [58]	Xception	284 C-19, 310 N, 330 BP, 327 VP	–	3 4	95 93	– –	– –	– –
DL+TL								
S.Dilshad et al. [59]	CNN_MobileNet	447 C-19, 447 N	–	2	96.33	–	–	93
Li, Daqiu, et al. [60]	SAE	–	275 C-19, 195 N	2	94.7	–	–	94.8
Shukla et al. [61]	CNN_GoogLeNet	1332 C-19, 1421 N	–	2	97.62	98.29	97.64	98.3

T: Tuberculosis

In [60] a Sparse Stacked Denoising AE (SSDAE) is proposed for feature extraction and an efficient TL approach is used to resolve the domain adaptation problem due to the diversity of the actual fault diagnosis distribution. Directly adopting features from the source domain’s pre-training phase and only adjusting the fine tuning step decrease the algorithm’s complexity. Madhavan et al. [52] proposed a TL model based on ResNet-50 architecture to classify COVID19 from various types of pneumonia using chest X-ray images. The dataset consisted of 5856 images, and the authors achieved an accuracy of 98.4% for COVID-19/Normal cases and 96.2% for COVID-19 against all other cases. Ullah et al. [62] proposed a framework based on multi-task semi-supervised learning (MTSSL) to detect COVID19. The authors utilized an adversarial autoencoder (AAE) for feature extraction and trained the model on a CXR dataset consisting of 8851 Normal, 1770 COVID19, 373 Negative, and 6069 Pneumonia, achieving an accuracy, sensitivity, and specificity of 96.95%, 91.07%, and 99.62%, respectively. Addo et al. [63] proposed the Ensemble Variational Autoencoder (EVAE) based on two encoders (ResNet50 and VGG16) for deep feature extraction, where the features are concatenated for effective classification of COVID19 using a CXR image dataset consisting of 3616 COVID19, 10192 Normal, 6012 lung opacity, and 1345 viral pneumonia.

The proposed framework achieved an accuracy of 99.19% for 4 classes and 98.66% for 3 classes. Reddy et al. [64] proposed a new framework called Multi-modal fusion of deep transfer learning (MMF-DTL) to detect COVID19. The authors utilized three TL models, namely VGG16, Inception v3, and ResNet50, for deep feature extraction from a CXR dataset containing 305 images with six classes. With the fusion model, the authors obtained an average sensitivity of 92.96%, specificity of 98.54%, and accuracy of 98.80%. Abdulkareem et al. [65] proposed CNN, SAE, and DNN to classify COVID19. The authors achieved an accuracy of 88.30% with the CNN model on the CT scan dataset consisting of 349 COVID19 and 397

Normal. Demir et al. [66] proposed an approach called DeepCovNet to classify COVID19 from normal and pneumonia cases. They used a convolutional autoencoder model for deep feature extraction and SVM optimized by SDAR algorithm for classification. The model achieved an accuracy of 99.75% on the CXR dataset consisting of 580 COVID19 images, 500 Pneumonia, and 1541 Normal. Khan et al. [58] proposed the Xception TL model for detecting COVID19 and other types of pneumonia using a CXR dataset containing 284 COVID19, 310 Normal, 330 Bacterial Pneumonia, and 327 Viral Pneumonia. The proposed model achieved an accuracy of 95% for 3 classes and 93% for 4 classes. Table 3-2 shows the limitation of recent studies.

Table 3-2 Limitations of recent methods

References	Recent works	Limitations
Madhavan et al. [52]	ResNet-50	This study's limitations become apparent as it only utilizes a single Transfer Learning model in conjunction with the CXR dataset, without comparing the results with those obtained by other models.
Ullah et al. [62]	Adversarial Auto Encoder (AAE)	Overall, while AAEs have shown potential for medical image generation and other applications, their limitations in terms of data, interpretability, generalization, and accuracy make them less suitable for diagnostic lung disease.
Addo et al. [63]	Variational Auto Encoder (EVAE) based on two encoders (ResNet50 and VGG16)	The limitation of this study is concentrated on chest X-ray images without using another modalities such as CT scan.
Reddy et al. [64]	Multi-modal fusion of deep transfer learning (VGG16, Inception v3, and ResNet 50)	As with AAEs, multi-modal fusion of deep transfer learning models requires large amounts of training data to generate meaningful results. There is also a limited interpret, generalization accuracy and complexity.
Abdulkareem et al. [65]	CNN, SAE and DNN	CNN and DNN are limited in interpretability, generalization and data. SAE is limited in interpretability, data and accuracy.
Demir et al [66]	DeepCovNet (Auto Encoder + SVM)	While DeepCovNet (Auto Encoder + SVM) has shown promising results for diagnosing lung diseases, it has limitations and should be used with caution in a medical context. It is important to carefully consider the limitations and potential biases of any ML algorithm and to supplement the algorithm's results with additional clinical information and expertise.
Khan et al [58]	Xception	The limitations of this work are shown in using one TL model with the CXR dataset and didn't use another model to compare the results.

Overall, while DL methods taken separately have shown promise in various image

classification tasks, their limitations in terms of interpretability, generalization, accuracy, and data make them less suitable for lung disease diagnosis. Therefore, more research is needed to determine the best approaches for using ML in medical diagnosis, including the use of alternative models or hybrid models that combine different techniques. From the studies reviewed, it can be concluded that neither the depth of the model, such as SAE, nor the fusion of TL alone can achieve optimal classification performance. This raises the question of whether the depth of the model is more effective than the fusion of TL or simply reducing the data. Further, it is worth investigating whether a combination of data reduction and fusion with depth can lead to more effective classification results.

3.2 Contribution

The TL Fusion and Stacked Auto-Encoders (TLFSAE) methodology for the classification of viral lung diseases brings several important approaches:

Using Transfer Learning: TLFSAE uses TL to leverage knowledge gained from previous tasks to improve the performance of classification models. This transfer acquired knowledge to related tasks such as image classification, which can be used to improve the accuracy of viral lung disease classification.

Using Stacked Auto-Encoders: TLFSAE uses Stacked Auto-Encoders for unsupervised feature learning. This helps to extract relevant features from the input data, which can improve the accuracy and robustness of classification models.

Model fusion: TLFSAE uses model fusion to improve the robustness and efficiency of classification models. By merging information from multiple pre-trained models, a more complete and accurate representation of the characteristics of the input data can be obtained.

Reduced data dimensionality: TLFSAE uses Stacked Auto-Encoders to reduce data dimensionality, allowing more complex data to be processed and reducing computation time.

Use an effective and simple classification algorithm: TLFSAE uses metric classification algorithms to improve model accuracy. By using metric classification algorithm, more reliable and accurate predictions can be obtained.

By combining these approaches, the TLFSAE methodology offers a novel approach for the classification of viral lung diseases, which can significantly improve the accuracy and reliability of classification models.

The primary objective of this work is to assess the effectiveness of the proposed classifier model in detecting various lung diseases, including COVID19. The primary contributions are outlined below:

- First application on CXR images: In this study, we introduce an efficient and reliable method for classifying lung infections, specifically COVID19, using chest X-ray (CXR) images. Deep TL models have been previously employed for pneumonia detection from CXR images; however, there is still scope for improvement in both feature extraction and advanced classification stages.
- Second application on CXR and CT scan images: The current study collected CXR and CT images and explored various TL methods, including E and Stacked Encoders (SE). The proposed classification method involves two stages for classifying different cases from CXR and CT images. The method involves performing TL-based SE and SAE on a large database, followed by concatenating the results to achieve improved classification performance.

Two studies and experiments have been conducted on two approaches: i) the first was applied to binary classification; ii) the second is concerned with multi classification (COVID-19/Normal/VPneumoni).

The main **contribution** of this work is centered on the selection of deep feature extraction methods. Specifically, this involves selecting the most effective models from a range of popular TL methods, including VGG16, VGG19, MobileNetV2, and examining their concatenation. Additionally, the work explores the best methods for deep data selection and reduction, such as SE and SAE, to enhance the accuracy of the classification task. The work also investigates the impact of model depth in conjunction with TL concatenation.

The **novelty** of this work stems from the deep data reduction and feature extraction techniques utilized, which enable the model to be applied to both CXR and CT modalities, thus selecting the most appropriate features for the dataset.

The research work consisted of the following steps:

- First, a dataset of CXR and CT images was collected from different sources. Next, various TL models were applied to the dataset and the best model was selected for lung image classification. Then, the performance of different models was compared to select the best one. The best TL models were concatenated to achieve high accuracy, and this concatenated model was applied to both CXR and CT image datasets.

- In the second stage, SE or SAE was added for dimensionality reduction or deep feature extraction to improve the performance of the best TL models. The effect of model depth associated with TL concatenation was also studied.

- Finally, the approach was applied to both binary and multiclass models to evaluate its effectiveness.

There are four scenarios in the experiments:

The **first scenario** shows the effectiveness of five TL models for classifying the chest CXR and CT cases and offers the importance of transfer features for the next stage. The **second scenario** is carried out to study data reduction using an E with the concatenation of the best TLs for binary classification and then a SE by association with the concatenation of the best TLs. The **third scenario** is designed for multi class classification to study the effect of the depth of the model and compare the E, SE, AE, and SAE algorithm for data reduction and feature extraction. The **fourth scenario** is conducted to test the proposed TL_E, TL_SE, TL_SAE or TL_SAE algorithm's ability as a classifier for improving the classification accuracy.

3.3 Proposed TL Fusion and SAE Model

TL Fusion and Stacked Auto-Encoders (TLFSAE) can be a powerful approach for Viral Lung Disease classification. TLFSAE is a technique that combines TL and Stacked Auto-Encoders to improve the accuracy of classification models. TL involves using the knowledge gained from one task to improve performance on another task. In the context of lung disease classification, TL can be used to leverage the large amount of pre-trained models available for image classification, such as ResNet, VGG, or Inception. Stacked Auto-Encoders are a type of neural network architecture that can be used for unsupervised feature learning. They consist of multiple layers of neurons that learn to encode and decode data, and can be used to reduce the dimensionality of data. By stacking multiple layers, Stacked Auto-Encoders can learn increasingly complex representations of the input data. The TLFSAE approach can be applied as follows:

- Pretrain a Stacked Auto-Encoder on unlabeled data, such as chest X-rays of healthy individuals, to learn a compressed representation of the input data.
- Fine-tune the pre-trained Stacked Auto-Encoder on labeled data of Viral Lung Disease and healthy individuals to learn a representation that is specific to the classification task.
- Use the learned representation as input to a metric classifier to predict the presence of Viral Lung Disease.

The TLFSAE approach can be further improved by fusing multiple pre-trained models using TL. By combining the knowledge from multiple models, the resulting representation is more robust and informative. Fusing can be done at different levels, such as at the feature level or decision level. For example, features from different pre-trained models can be concatenated and used as input to the Stacked Auto-Encoder, or the outputs from multiple classifiers can be combined using techniques such as majority voting or weighted averaging.

The architecture proposed in this work is composed of four stages. The first stage is dedicated to the preprocessing of the lung sequence. In the second stage, different TL models are studied and evaluated for their performance, and the best one is selected as shown in fig 3.1. The third stage consists of an encoder (E) and either SE for data reduction and deep data reduction or AE and SAE for feature extraction and deep features extraction. Finally, the fourth stage consists of different layers used for the classification of viral lung diseases.

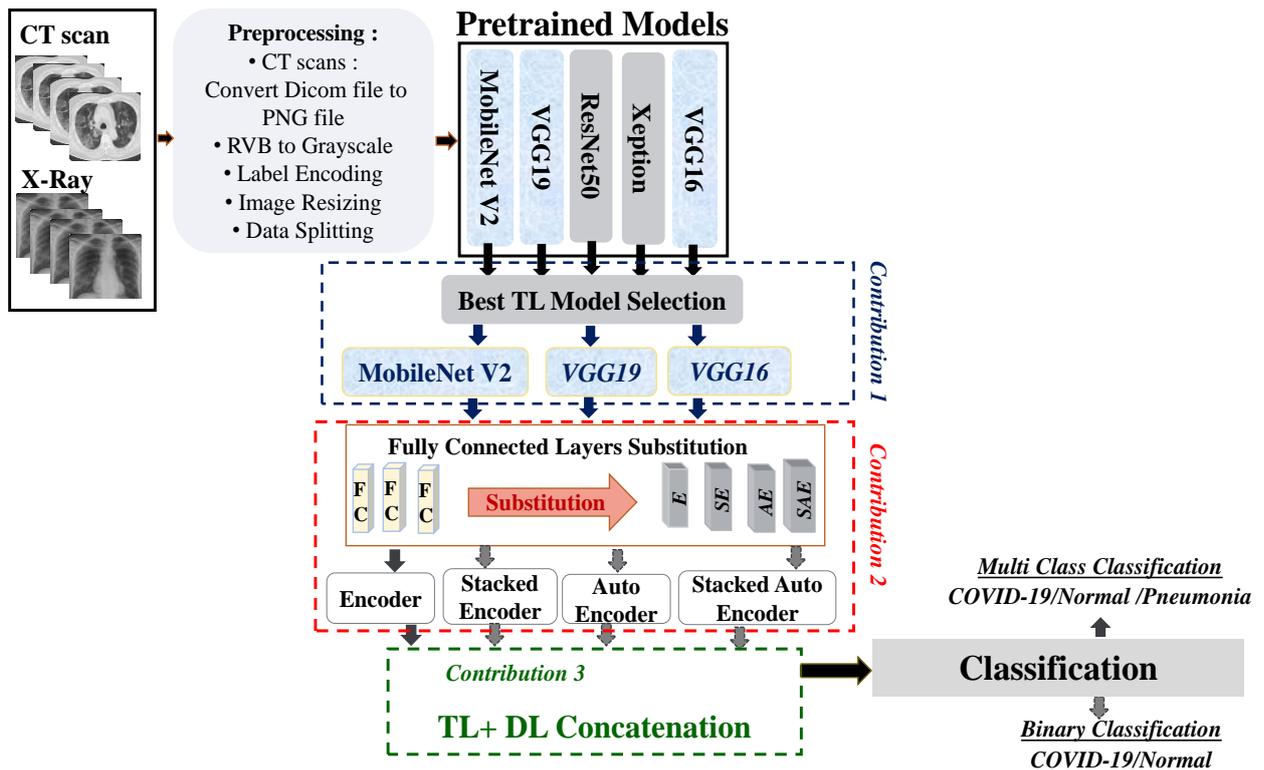


Figure 3.1 Block diagram of the proposed approach

The encoder extracts the pertinent features and reduces the number of data. It is generally used to discover non-linearity in the lungs, allowing for the detection of any lung problems. Several feature extraction methods have been used to enhance the accuracy of viral lung diagnosis. The E network is added to extract deep features from CT and CXR images, and the obtained features are utilized in the classification layers in the next stage. The Figure 3.2

Proposed approach for binary and multiclass classification

In the fig 3.2 , the contributions of the proposed method are depicted as follows:

- The first contribution (dotted blue line) is the selection of the best TL model for deep feature extraction.
- The second contribution (red dotted line) is the replacement of the Fully Connected (FC) layer with a more effective data reduction method.
- The third contribution (green dotted line) is the concatenation of three TL models with fine-tuning for improved classification results.

3.3.1 Preprocessing

Based on the characteristic signs of COVID19/VPneumonia, images were defined as inflammatory lesions and extracted by a CV model as per the following steps: CT scans Dicom files was converted to PNG file, image was converted to grayscale and resizing. Afterwards, label encoding has been done, and then dataset is splitted in train, validation and test (see Fig.3.3).

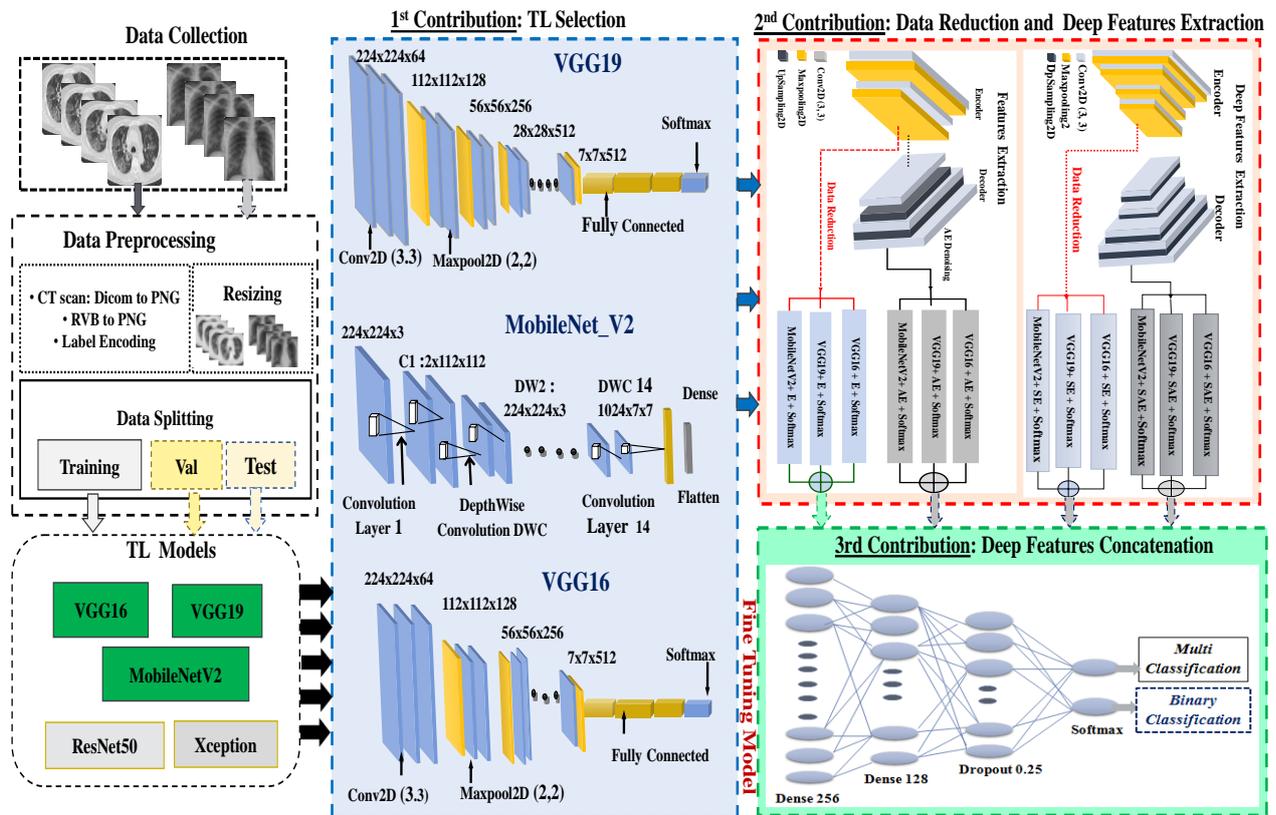


Figure 3.2 Proposed approach for binary and multiclass classification

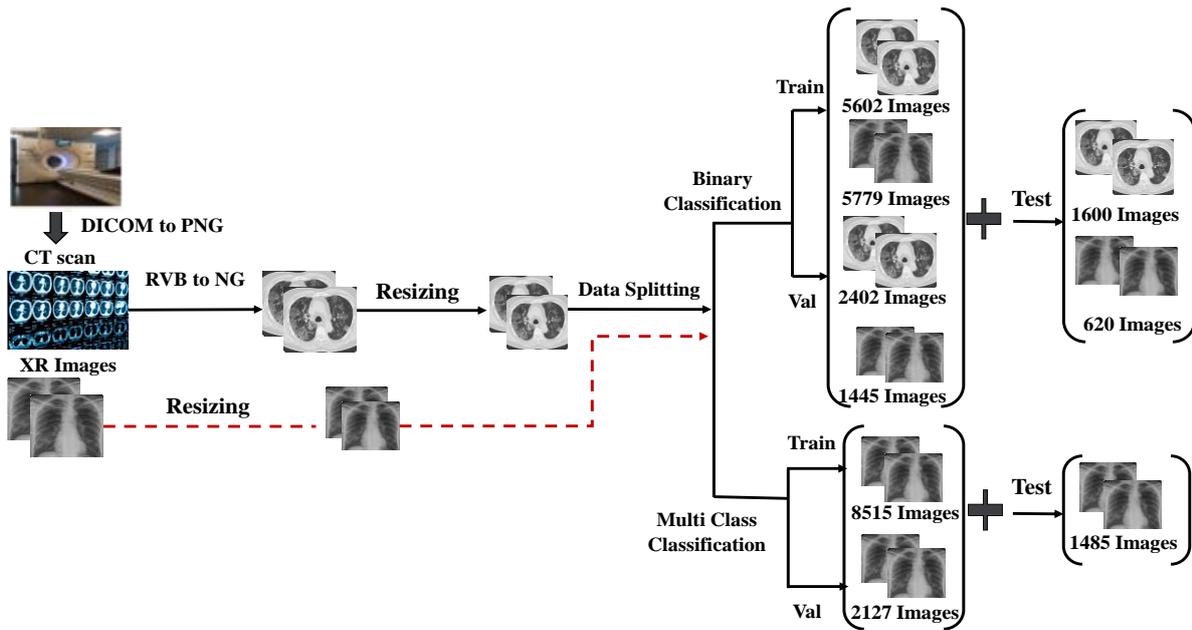


Figure 3.3 Data preprocessing and splitting

3.3.2 Transfer Learning Fusion

TL is a technique for re-using a previously trained model, and the knowledge obtained from the prior task is used in TL [43]. This technique compensates for the dataset's size limitation while additionally speeding up the training process. As a result, several DL models have been applied in the feature extraction stage. For COVID19 classification, pre-trained models such as MobileNet_V2, VGG16, VGG19, Xception, and ResNet50 have been applied to achieve high accuracy scores.

- **VGG16** is DL architecture consists of 13 Convolutional layers with 3 layers for fully connected layers. The final output layer consists of the 1000 classes of images for the input of 224 x 224 images, all the layers are followed by the convolutional layers and max-pooling layers. RELU is the activation function that is used in the hidden layers [67][30].
- **VGG19** is a pre-trained network that is trained on the ImageNet dataset, which achieved state-of-the-art performance on ILSVRC Challenge 2014. It is a DL architecture that consists of 16 Convolutional layers with 3 layers for fully connected layers [68].
- **MobileNetV2** is a streamlined architecture that constructs lightweight DCNN using depthwise separable convolutions and provides an efficient model for mobile and embedded vision applications. MobileNet model is depth-wise, and has shown to be efficient and accurate enough to run on lightweight computational device [23][69].

- **Xception** model includes depth-wise separable convolutions followed by point-wise convolutions in a DCNN architecture. The Xception module is equivalent to the Inception module, however the depth-wise separable convolutions replace the Inception modules [70].
- **ResNet50** (Residual Network) is an Artificial Neural Network (ANN) that uses skip connections to generate pyramidal cells and contains no linearity. The skip connections are used to get over the problem of decreasing gradients and simplify the network even more. ResNet-50 contains 50 layers, including 48 convolution layers[71].

3.3.3 Data Reduction and DFE

A decoder and an encoder constitute the generic AE model. The encoded coefficients are the learned features. A network with many more hidden neurons is commonly developed for the no linearly separable dataset (over complete representation) [50]. The basic AE can also be combined to create a deep model. A SAE neural network is a modeling method to use an AE neural network [60]. The loss of the AE network is the reconstruction loss, as shown in Eq. 1:

$$J = \frac{1}{N} \sum_{i=1}^N L(f_{encoding}(w, x), x) \quad (3.1)$$

Where N denotes the size of batch size, w is the parameter matrix of layer, x is the input image, and L is loss function Mean-Squared-Error (MSE) and cross entropy, as shown in Eq. 2 and Eq. 4:

$$((w, x), x) = (f_{encoding}(w, x) - x)^2 \quad (3.2)$$

Where J denotes the layer AE network loss function as the regularization item and h is the output of layer $f_{encoding}$, as shown in Eq. 3:

$$h = f_{encoding}(w, x) \quad (3.3)$$

$$L_{crossentropy}(f_{classification}(w, h), y) = -y \log(f_{classification}(w, h)) \quad (3.4)$$

The convolutional AE is a model that can combine between the convolution operation and the AE, where each layer of convolutional AE produce n feature maps, it can use as input in next layer [72]. In this case, the encoder was used for data reduction, and Auto Encoder was used for feature extraction. For deep feature extraction, we used the SE and SAE. The architecture of an AE and SAE models are shown in fig 3.4.

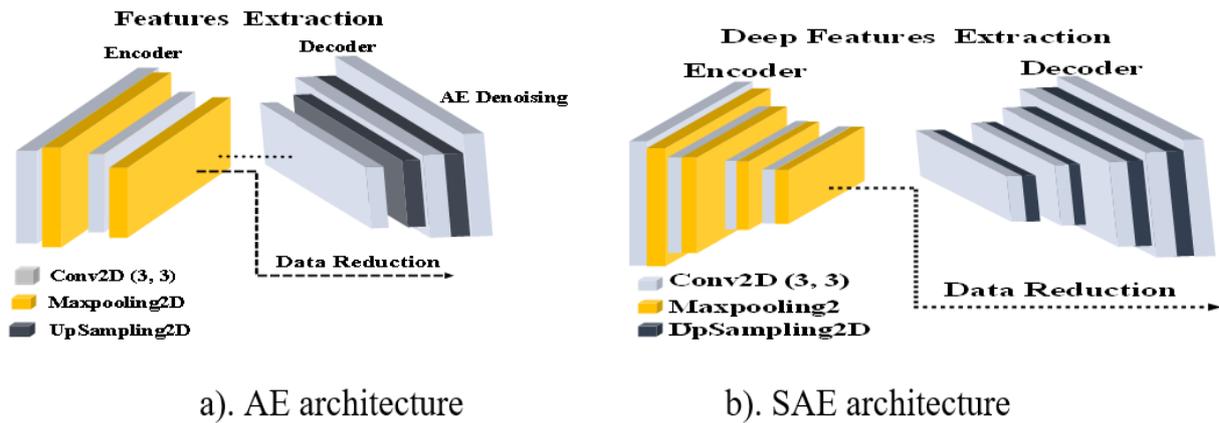


Figure 3.4 Data Reduction and Deep Feature Extraction

3.3.4 TLFSAE Model architecture

To evaluate our model, the images were preprocessed using different functions before the model was implemented on the input. Further, in the training phase, the dataset was resized by 50x50 and 128x128, as shown in fig 3.5, and the fig 3.6 shows detailed parameters of VGG16 + Encoder model with (50x50) and (128x128) size.

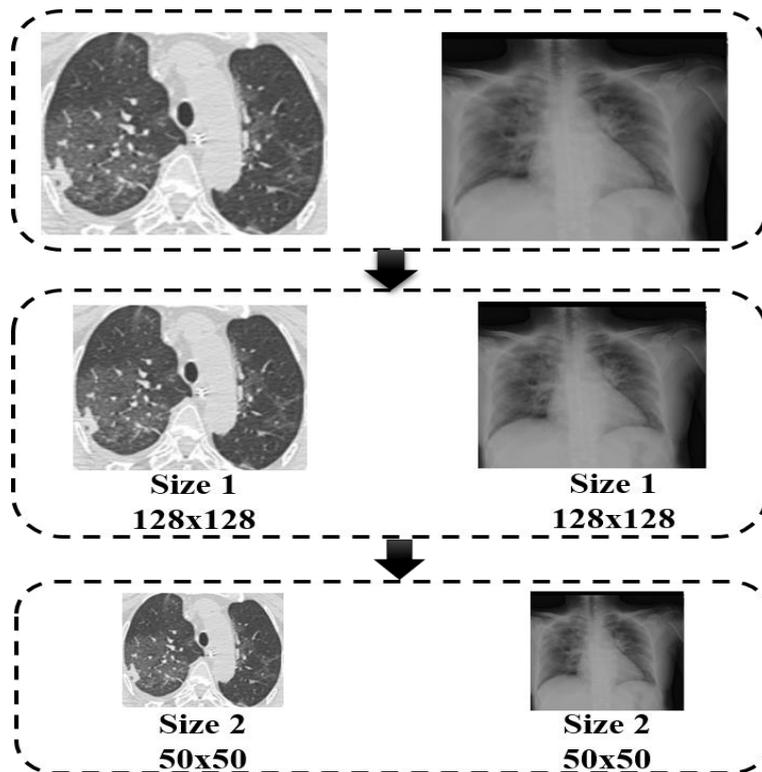


Figure 3.5 Dataset with different size

3.4 Experimentation and Results

3.4.1 Data Collection

Example of CT and CXR images are illustrated on fig 3.7. Before training the model, the first step is to partition the dataset. By the holdout method, the original dataset is divided into three mutually exclusive sets, which are divided into three sets: a training, verification, and test. Table 3-3 describes the dataset distribution.

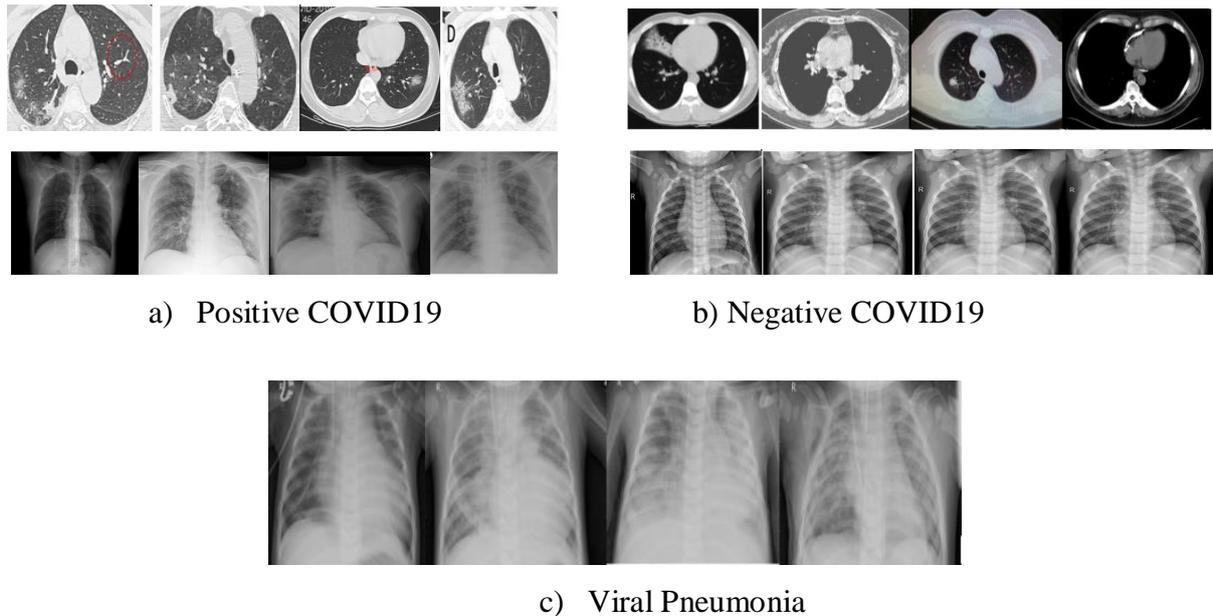


Figure 3.7 Examples of CT scans and XR Lung images

Table 3-3 Data distribution

Binary Classification						
	Training		Validation		Test	
	CT	CXR	CT	CXR	CT	CXR
COVID19	2800	2893	1200	723	800	315
Normal	2802	2886	1201	722	800	315
Multi Classification						
CXR						
COVID19	2893		723		315	
Normal	2886		722		315	
VPneumonia	2736		682		855	

3.4.2 Binary Classification

To train and assess the system on CXR images, we assemble a dataset comprising 3616 COVID19 and 3608 Normal images. 80% of the dataset is allocated for training, with the remaining 20% for validation. The test dataset consists of 315 COVID19 and 315 Normal images. For CT images, our dataset includes 4000 COVID19 and 4003 Normal images, with

70% earmarked for training and 30% for validation. The test dataset comprises 800 COVID-19 and 800 Normal images. We employ diverse hyperparameters during model training, detailed in Table 3-4 along with the functions used for binary and multiclass classification.

In this section of the work, the authors describe several experiments conducted for binary classification using a dataset of CXR and CT images.

- The first experiment involved using an Encoder with two layers, each consisting of Convolutional Layers, ReLU activation, and MaxPooling to extract the latent space from the images. Different hyperparameters were used during training, such as the number of epochs, batch size, learning rate, and image size, to evaluate the model's performance. Table 3-4 displays the best hyperparameters that produced the best results.
- The second experiment involved using TL models combined with a fine-tuning model consisting of a CNN (Convolutional Layers, ReLU activation, and MaxPooling) and two dense layers with dropout and Softmax activation to extract deep features from the CXR and CT images.
- The third experiment involved comparing the results of using Encoder alone, TL+CNN, and TL+Encoder models to select the most effective model. The proposed technique used TL combined with Encoder to achieve this goal.

Table 3-4 Hyperparameters of different models

	Optimizer	SI	LR	BS	Beta_1	Beta_2	Epochs	LF	AF
CT	Adam	128x128	10 ⁻⁵	32	0.9	0.999	10	Categorical_ Crossentropy	Softmax
XR	Adam	128x128	10 ⁻⁵	32	0.9	0.999	25 and 10	Categorical_ Crossentropy	Softmax

SI: Size Image; LR: Learning Rate; BS: Batch Size; LF: Loss Function; AF: Activation Function

Discussion: In fig 3.8 and table 3-6 we see experimentation results of different TL models with E or CNN applied to CXR and CT images. The reason why our model can achieve better detection performance is mainly because the encoder detector model that has the following advantages: firstly, each layer of the E detector model can be trained separately, which ensures the controllability of the dimensionality reduction of the CXR and CT scan images features. Secondly, the TL added in the encoder layer also plays an important role in improving the accuracy of the detection model. The TL models are used as a feature extractor; afterwards, the encoder (or CNN) is for deep feature extraction.

Table 3-5 Results obtained from different TL models

Methods	Val_Acc (%)	Sens. (%)	Spec. (%)	TP	FP	FN	TN
Encoder	CT-scan : 99.13	99	99.25	1188	12	9	1192
	CXR : 81.11	77.59	84.63	581	162	111	611
VGG19_Encoder	CT-scan : 96.25	98.51	94.18	1183	8	16	719
	CXR : 96.06	94.61	97.51	684	39	18	704
VGG19_CNN	CT-scan : 99	99.32	98.67	1192	17	73	718
	CXR : 97.02	96.82	97.23	700	23	20	702
VGG16_Encoder	CT-scan : 99	99.32	98.67	1192	17	73	718
	CXR : 97.02	96.82	97.23	700	23	20	702
VGG16_CNN	CT-scan : 99.71	99.75	99.67	1197	3	4	1197
	CXR : 91	90.18	91.83	652	71	59	663
MobileNetV2_Encoder	CT-scan : 99.79	99.66	99.91	1196	4	1	1200
	CXR : 96.26	96.96	95.57	701	22	32	690
MobileNetV2_CNN	CT-scan : 95.34	95.92	94.75	1151	49	63	1138
	CXR : 93.84	95.02	92.66	687	36	53	669
Xception_Encoder	CT-scan : 94.54	91.17	97.92	1094	106	25	1176
	CXR : 89.62	87.69	91.55	634	89	61	661
Resnet50_Encoder	CT-scan : 86.25	81.83	90.67	982	218	112	1089
	CXR : 75.22	64.87	85.60	469	254	104	618

From the results, we notice the strength of the encoder in the extraction of the optimal parameters and their reduction. It is significantly more efficient than the CNN model. We can say that the encoder used alone gives satisfactory and efficient results.

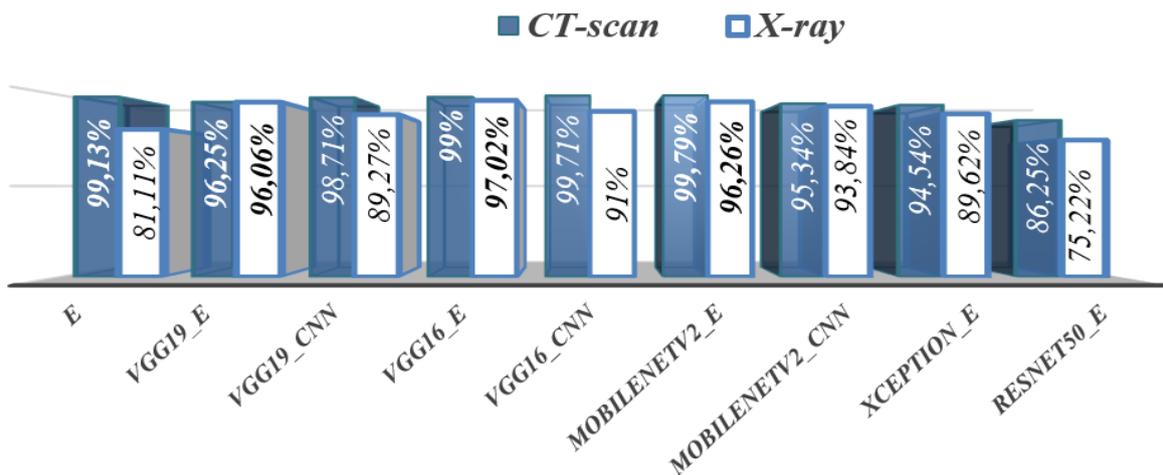


Figure 3.8 Accuracy comparison of different TL models in CT and CXR image

3.4.3 Concatenation models

The initial block is the TL models stacked with an E. The output of this block is used as feature vector input to the softmax function. Next, the three softmax are concatenated together

and passed to the next stage. We used flatten, dense layers, and softmax function for COVID19 classification.

The results show in table 3-6 that the concatenation TL-E model is indeed effective. Our proposed model has about 37M parameters (see Fig 3.9), which is very large. This is mainly due to the E neural network having strong feature expression ability and the advantages of TL concatenation. It can usually obtain the hierarchical grouping structure feature or the partial whole structure feature of the input, and parameters in general, are weights that are learned during training.

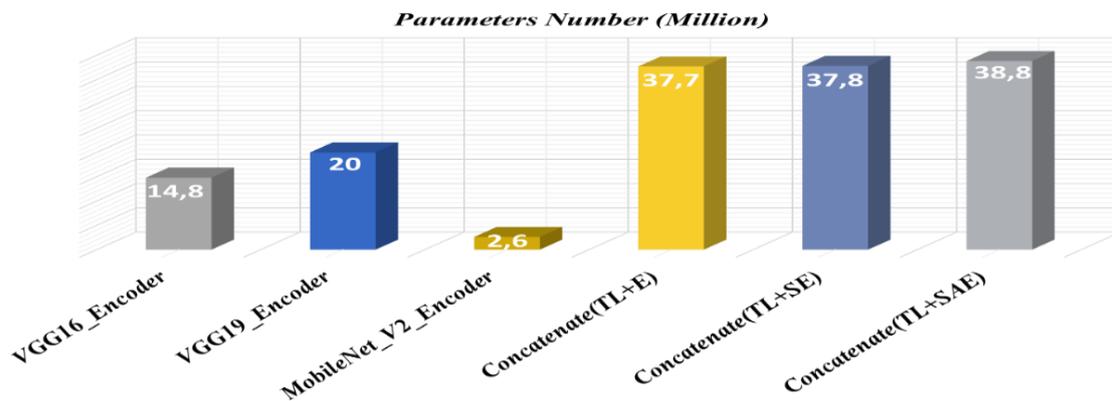


Figure 3.9 Parameters number of different proposed approach

Error! Reference source not found. shows that the best reduction in the number of parameters is given by the MobileNetV2_Encoder model. This model is a good candidate and remains to be improved.

Table 3-6 Concatenation and Encoder effect for binary classification

CXR										
Methods	Pn	Acc. (%)	Spec. (%)	Sens. (%)	Acc_test. (%)	T/image.(s)	TP	FP	FN	TN
1.VGG16_Encoder	14.866.482	97.02	96.82	97.23	98.44	0.01	700	23	20	702
2.VGG19_Encoder	20.176.178	96.06	94.61	97.51	96.87	0.01	684	39	18	704
3.MobileNetV2_Encoder	2.630.962	96.26	96.96	95.57	98.44	0.006	701	22	32	690
Concatenation (1+2+3)	37.708.824	98.41	98.07	98.33	99.21	0.008	711	12	14	708
CT_scan										
1.VGG16_Encoder	14.866.482	99	99.32	98.67	99.21	0.001	1192	8	16	1185
2.VGG19_Encoder	20.176.178	96.25	98.51	94.18	97.65	0.002	1183	17	73	1128
3.MobileNetV2_Encoder	2.630.962	99.79	99.66	99.91	98.43	0.001	1196	4	1	1200
Concatenation (1+2+3)	37.708.824	99.87	99.75	100	100	0.002	1196	3	0	1201

Pn: parameters number

From table 3-7, we can say that the encoder used here for the data reduction associated with MobileNetV2 has a good test time performance.

The concatenation is the best only its disadvantage is the number of parameters hence the slower test time. In the case of the COVID19/Normal binary classification, table 3-6 shows that the best results are obtained in the case of the concatenation associated with the E a one layer that plays the role of data reduction in this case it is similar to PCA. The difference between a PCA and an AE is that the AE performs analysis on the data with a non-linear activation function on the hidden layers.

The proposed method achieved the highest performance with remarkable accuracy, specificity, and sensitivity. Specifically, for the CXR dataset, the method achieved an accuracy of 98.41%, a specificity of 98.01%, and a sensitivity of 98.33%. Meanwhile, for the CT scan dataset, the proposed method yielded an accuracy of 99.87%, a specificity of 99.75%, and a sensitivity of 100%. Moreover, the concatenation models demonstrated improvements in true positives/true negatives and reductions in false positives/false negatives for both CXR and CT scan datasets. In what follows, we are interested in the depth of the model and the effect of the concatenation of the best TL on the performance of the multi-classification.

3.4.4 Multi classification

To train and evaluate the system, we collect a CXR dataset containing 3616 COVID19, 3608 Normal, and 3418 Pneumonia. All the images were resized to 128x128 pixels and converted to grayscale images.

In this part of the work on multi-classification as shown in table 3-8, several experiments were conducted with a dataset on CXR images and which look like this:

3.4.4.1 Experience 1

SAE: In this study, we employed a Stacked Encoder-Decoder architecture. The Encoder is composed by four layers, with each layer containing Convolutional (Conv) operations with Rectified Linear Unit (ReLU) activation functions, and Maxpooling. The Decoder also has four layers, with each layer containing Conv operations with ReLU activation functions and UpSampling.

Table 3-7 Best results for proposed multi classification

Methods	Val_Acc (%)	Val_Loss	Test_Acc (%)	Test_Loss	TestTime (s)	Sens(%)	Spec (%)	Confusion Matrix
Stacked Auto-Encoders 2 Layers								
SAE	94.26	0.177	92.96	0.166	10.594	32.95	66.20	$\begin{bmatrix} 231 & 267 & 239 \\ 233 & 255 & 238 \\ 235 & 238 & 209 \end{bmatrix}$
VGG19+ Stacked Auto-Encoders 2 Layers								
VGG19+ SAE	92.82	0.207	94.53	0.186	17.483	35.88	67.94	$\begin{bmatrix} 249 & 251 & 237 \\ 222 & \mathbf{280} & 224 \\ 204 & 237 & \mathbf{241} \end{bmatrix}$
Concatenation : TL_Encoder (TL_E) 1 Layer								
VGG16+ MobileNetV2 +VGG19_E	97.48	0.099	88.28	0.383	5.828	35.91	67.95	$\begin{bmatrix} \mathbf{259} & 267 & 211 \\ 214 & \mathbf{277} & 235 \\ 230 & 217 & \mathbf{235} \end{bmatrix}$
Concatenation : TL_Encoder (TL_E) 2 Layer								
VGG16+ MobileNetV2 +VGG19_SE	97.34	0.112	84.25	0.306	43.967	31.05	65.53	$\begin{bmatrix} 229 & 264 & 244 \\ 240 & 241 & 245 \\ 229 & 256 & 197 \end{bmatrix}$
Concatenation : TL_Stacked Auto-Encoders (TL_AE) 1 Layers								
VGG16+ MobileNetV2 +VGG19_AE	97.76	0.099	84.37	0.365	8.969	33.07	66.56	$\begin{bmatrix} 228 & 267 & 242 \\ 240 & 268 & 218 \\ 238 & 229 & 215 \end{bmatrix}$
Concatenation : TL_Stacked Auto-Encoders (TL_SAE) 2 Layers								
VGG16+ MobileNetV2 +VGG19_SAE	97.11	0.108	88.28	0.399	8.744	34.26	67.13	$\begin{bmatrix} 258 & 243 & 236 \\ 230 & 262 & 234 \\ 210 & 251 & 213 \end{bmatrix}$

3.4.4.2 Experience 2

VGG19+SAE: In this case, fusion VGG19 with SAE are used to deepen the model and extract the discriminate features. All the VGG19 layers were freeze and the fully connected layers were removed and replaced by SAE.

3.4.4.3 Experience 3

Concatenation three TL model and E, or SE or SAE: Each of the TL models was combined with E, SE, or SAE and the fine-tuning model was added on the top of each model. The concatenation of the softmax of the three models was applied and the two dense layers and dropout layer with softmax activation function were added to classify different types of pneumonia.

The model is trained with a variety of hyperparameters. In addition, we got the best-SAE detection model through these parameters. Different models results for multi-class classification are described in table 3-8.

The algorithm of the TL_SE or SAE approach is given by algorithm 2 below:

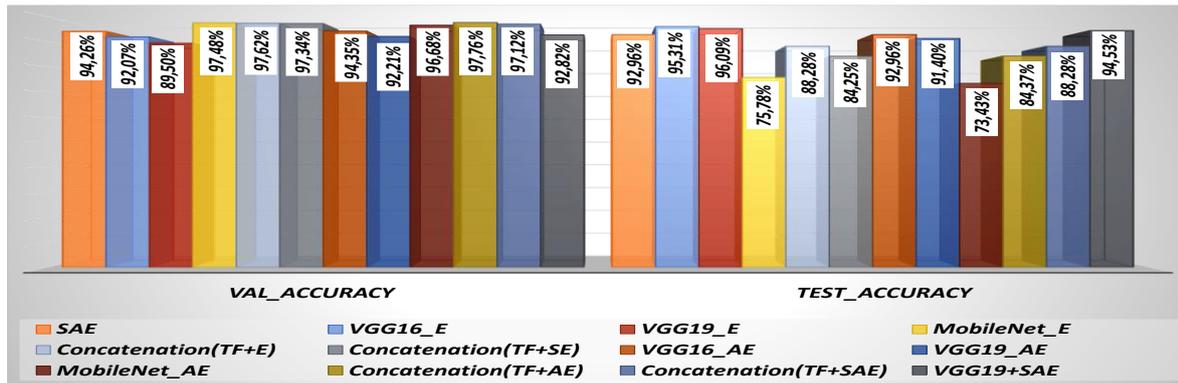
Algorithm 2: *Concatenation Transfer Learning SE or SAE for CT and CXR classification*

```

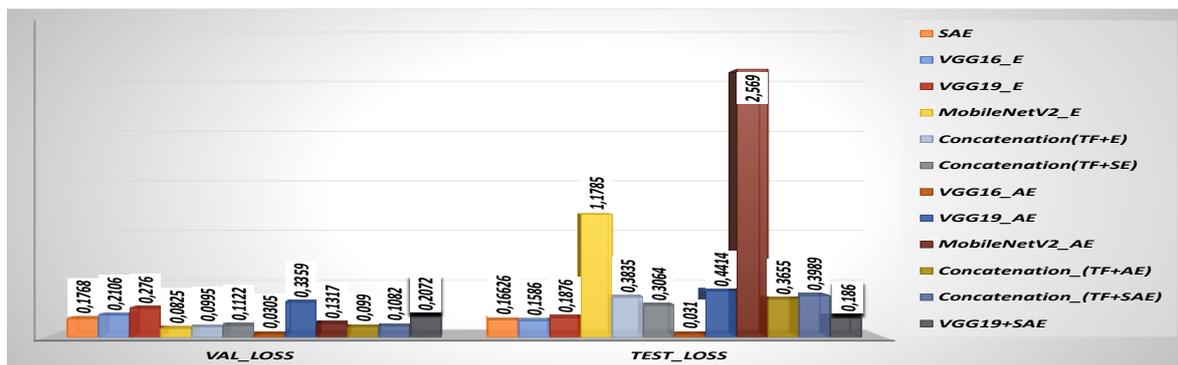
Input: (128x128x3) dimensional CXR and CT images.
Output: predicted class label:
    • 0: COVID-19, 1: Normal (CXR and CT images).
    • 0: COVID-19, 1: Normal, 2: VPneumonia (CXR images).
for i=1 to N do (N: number of epochs)
    Train(VGG16_Encoder, train_img, img_label)
end
Softmax1 = save (VGG16_E, SE or SAE)
for i=1 to N do
    Train(VGG19_Encoder, train_img, img_label)
end
Softmax2 = save(VGG19_E, SE or SAE)
for i=1 to N do
    Train (MobileNetV2_Encoder, train_img, img_label)
end
Softmax3 = save(MobileNetV2_E, SE or SAE)
Concatenate = concatenation(softmax[softmax1, softmax2, softmax3])
Dense layers (for features softmax vector)
Formed features vector is used as input to softmax function to perform classification.
    
```

According to the results, it is clear that the fusion of the TL is prevalent over the depth of the model. The more layers that are added to the models slower the test.

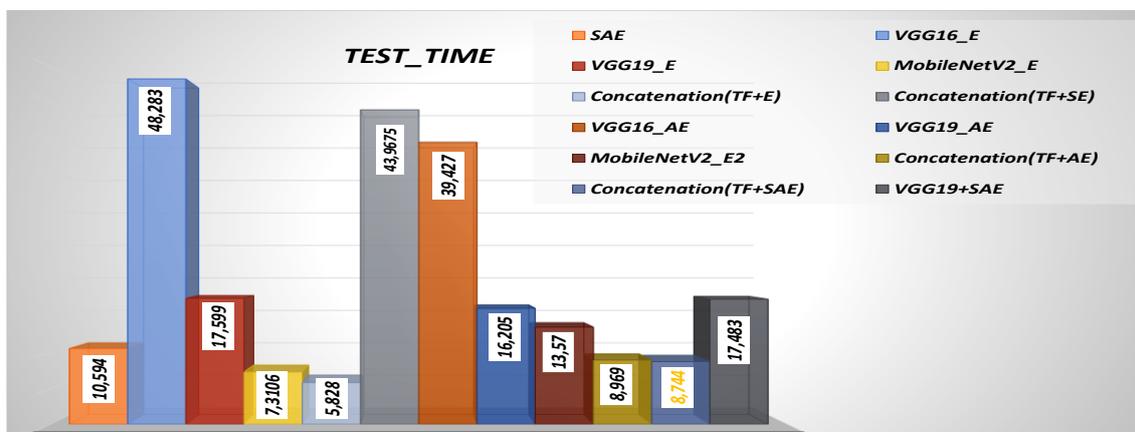
For the multi class classification, The results in table 3-8 show that the concatenation of the TLs associated with the reduction by the encoder (VGG16+MobileNetV2+VGG19_E) is more efficient than the depth of the model (VGG16+MobileNetV2+VGG19_SAE) in the case of the multi-classification COVID19/VPneumonia/Normal.



a) Validation and Test Accuracy



b) Validation and Test Loss



c) Test Time

Figure 3.10 Comparing the validation and test accuracy, loss and time test of the different TL models

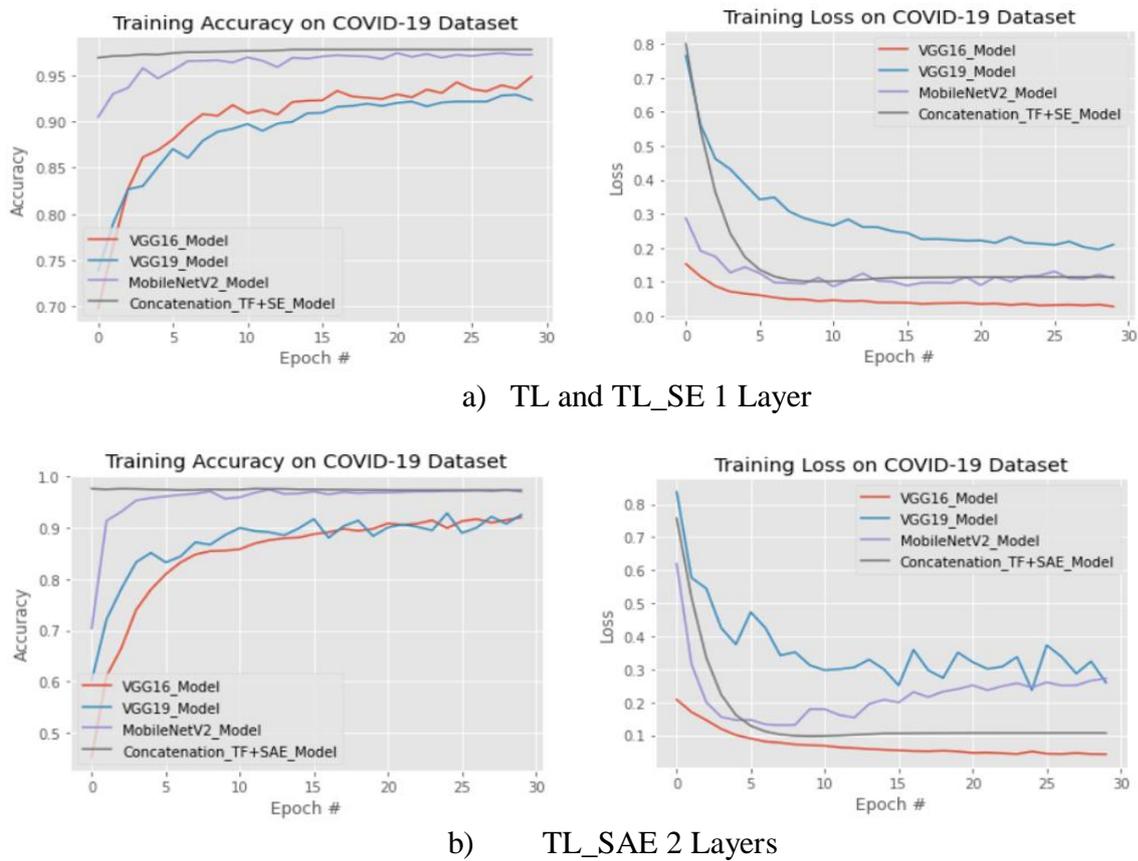


Figure 3.11 Val_Accuracy and Val_Loss for TL_SE /SAE

Discussion: For multi classification, the concatenation between three TL_E, TL_SE, and TL_SAE is used for deep feature extraction and data reduction. The TL_E 1 layer is the fastest of all methods. This is explained by the fact that the image coming out of the AE is denoised while the one coming out of the E is reduced and coded only there is a loss of information, which explains the results.

Through results shown in table 3-8, SAE without concatenation gives better performance results than concatenation of different TLs with Test_Acc=92.96% and Test_Loss = 0.166 against Test_Acc = 88.28% , Test_Loss = 0.399, but it is slower with Test_{time} = 10.594s against Test_{time}=8.744s for concatenation. We can say that concatenation plays an important role in the test speed but it does not improve the performance of the classification of the multiclass model.

3.5 Comparison with state of the art methods

The multiclass gives good results when the pathologies to be classified are not viral, while classification difficulties are linked to the viral aspect of COVID19 and Viral Pneumonia.

Table 3-8 Comparison results with state of the art methods

Authors	Methods	Datasets	Class	Results (%)					
				Acc	Sens	Spec	Test time(s)	P _n (M)	
Binary Classification (COVID19(C-19)/Normal(N)) : DL									
Rashid, et al. [67]	AE	CT	CXR						
		–	816 N 408C-19		90.13	–	–	–	–
Ullah et al [62]	AE	–	8851 N, 1770 C-19, 373 Neg, 6069 P	N, C-19, Neg, P	96.95	91.07	99.6	–	–
Rashid et al.[67]	AutoCovNet	–	408 C-19, 816 NC-19	C-19 NC-19	99.39	99.39	100	–	–
Proposed Model	SE	4000 C-19 4003 N	3616 C-19 3608 N	C-19 N	CT				
					99.13	99	99.25	–	92 Th
					CXR				
					81.11	77.59	84.63	–	
Binary Classification (COVID19(C-19)/Normal(N)) : DL and TL									
Lahsaini et al.[44]	DenseNet201+ GradCam	1868 C-19 3118 N	–	C-19 N	98.8	98.54	99.02	–	–
Narin et ai. [70]	ResNet50	–	2800 N, 341 C-19, 2772 BP, 1493 VP	1. N, C-19, 2. C-19, VP 3. C-19, BP	96.1 99.5 99.7	–	–	–	–
Madhavan et al.[52]	ResNet-50	–	150 C-19, 1583 N, 2798BP, 1480VP	1. N, C-19 2. C-19,(BP, VP)	98.1 96.2	–	–	–	–
Ketfi et al [75]	MobileNetV2	1,3600 C-19 3600 N 2,4800 C-19 4800 N	219 C-19 397 N	C-19 N	CXR				
					96.77	90.91	100	0.18	–
					CT				
					99.67 99.62	99.93 100	100 99.25	0.03 0.05	– –
Li, Daqiu, et al [60]	SAE	275 C-19, 195 N	–	–	94.7	–	–	–	–
	MobileNetV2_E	4000	3616		CT				

Proposed Model		<i>C-19,</i> <i>4003 N</i>	<i>C-19,</i> <i>3608 N</i>	<i>C-19</i> <i>N</i>	99.79	99.66	99.91	0.001	2.63
					CXR				
					96.26	96.96	95.57	0.006	
		MobileNetV2_ CNN	<i>4000</i> <i>C-19,</i> <i>4003 N</i>	<i>3616</i> <i>C-19</i> <i>3608 N</i>	<i>C-19</i> <i>N</i>	CT			
				95.34	95.92	94.75	–		
				CXR					
				93.84	95.02	92.66	–		
Binary Classification (COVID19(C-19)/Normal(N)) : DL and TL Concatenation									
Kong et al.[71]	Fusion(DenseNet+VGG16)+GAB + CAB	–	576 C-19 1583 N	C-19 N	98.0	–	–	–	–
Reddy et al. [64]	Concatenation (VGG16, Inception v3, and ResNet 50)	–	305 images with six classes	–	98.80	92.96	98.54	–	–
Proposed Model	Concatenation VGG16_E +VGG19_E +MobileNetV2_E	<i>4000</i>	<i>3616</i>	<i>C-19</i> <i>N</i>	CT				37.7
		<i>C-19,</i> <i>4003 N</i>	<i>C-19</i> <i>3608 N</i>		99.87	99.75	100	0.002	
					CXR				
					98.41	98.07	98.33	0.008	
Multi Classification (COVID19(C-19)/VPneumonia(P)/Normal(N)) : DL and TL									
L. Mohamed et al. [48]	CNN Bayesian Optimizer	3616 C-19, 3616 N, 3616VP		3	96	–	–	–	–
Kong et al. [71]	Fusion(DenseNet+VGG16) + GAB + CAB	–	576 C-19 1583 N 4273 CP	3	97.3	–	–	–	–
Khan, E, et al.[53]	EfficientNetB1, NasNetMobile, MobileNetV2	2473 C-19, 6012 Opacity, 10192N, 1345VP	–	4	96.13	–	–	–	–
Xu et al. [73]	UNet : ResNet + CNN model	1840 N, 433 C-19, 2780BP, 1345VP 394 T	–	N, C-19, BP, VP, T	96.32	–	–	2h37mn	23.5
Proposed Model	Concatenation VGG16_SE + VGG19_SE + MobileNetV2_	CXR <i>12127</i>	<i>3931</i> <i>C-19,</i> <i>3923 N,</i> <i>4273 VP</i>	<i>C-19</i> <i>N</i> <i>VP</i>	97.48	–	–	0.004	37.7

	<i>SE</i>								
	<i>Concatenation</i> <i>VGG16_SAE +</i> <i>VGG19_SAE +</i> <i>MobileNetV2_</i> <i>SAE</i>				<i>97.12</i>	<i>_</i>	<i>_</i>	<i>0.006</i>	<i>38.8</i>

The recently published models are compared and discussed the differences between the proposed models. Obviously, the accuracies obtained varied, but the advantage of our model is that it distinguishes COVID19 from other typical types of pneumonia. For example, despite the high accuracy of some models, they were conducted on limited datasets, applied to a single modality, and sometimes could not distinguish between COVID19 and other SARS. Some models obtained quite similar accuracies; however, they compared COVID19 and other lung diseases. Table 3-8 shows comparative results of our approach with the existing models, it also indicates that compared to other models, our proposed concatenation model achieves the highest accuracy for binary classification of lung CT images when compared to existing models.

3.6 Conclusion

In conclusion, the TLFSAE approach with model fusion has the potential to improve the accuracy of Viral Lung Disease classification. This approach can be further optimized by exploring different pre-trained models, fusion techniques, and classification algorithms.

In this work, different TL models are used. We have proposed the concatenation method between the three best TL models used with SE and SAE. They are linked together for obtained better results. As a result, an automatic recognition system is desperately needed to cut down on false positives. VGG16, VGG19, ResNet50, MobileNetV2, and Xception are just some of the many pre-trained CNN models. This work introduces new concatenation models between three TL models and SE or SAE. The model is suitable for diagnosing COVID19 by applying different pre-trained models. For binary classification, the approach achieves the highest performance with an accuracy of 99.87% for CT scans and 98.41% for a CXR. The test time per image for our approach spent $t_{CXR} = 0.008s$, $t_{CT} = 0.002s$. For multi class classification, the proposed model obtained an accuracy of 97.48% for concatenation TF and SE, and 97.10% for concatenation TL and SAE. TL_SAE is the fastest with $Time_{test} = 8.744s$ and VGG19 is the best candidate for the confusion matrix. In this case, we can say that the concatenation plays an important role in the test speed but it does not improve the performance of the classification of

the multiclass model. Our proposed model outperformed existing methods and demonstrated the best performance.

Therefore, more research is needed to determine the best approaches for using ML in medical diagnosis, including the use of alternative models or hybrid models that combine different techniques. There are several perspectives for the TL Fusion and Stacked Auto-Encoders (TLFSAE) approach in the context of Viral Lung Disease classification :i) Improved accuracy by leveraging the knowledge gained from pretrained models and learning a representation that is specific to the classification task. The use of Stacked AutoEncoders for feature learning can also help in identifying relevant features that are important for classification. It can also be applied to other areas of medical imaging to improve diagnostic accuracy and reliability.

As future work, we aim to collect a larger dataset and focus on improving the quality of images in the preprocessing stage. We also plan to utilize metaheuristic algorithms to optimize our proposed model for the detection and diagnosis of various diseases. Additionally, we aim to improve the test time and standardize the model while also exploring novel measures for evaluation of the disease evolution.

4 Chapter: Metaheuristic Optimization Deep Features for Pathologies Diagnostic

4.1 Introduction

In the domain of medical imaging, the task of pinpointing infected regions within MRI, CT scans, and Ultrasounds remains a daunting challenge, even for seasoned experts. The power of DL has proven to be a groundbreaking approach in the realm of medical pathology diagnostic. The objective is to find a standard model to diagnose different pathologies using different sensors for different stages of diseases and above all detect the disease at its beginning for early diagnostic.

This work introduces a framework to detect and classify not one, but three critical diseases: lung and breast cancer, as well as Alzheimer disease (AD). Proposed approach lies in a hybrid model fine-tuned through the application of Manta-Ray Foraging Optimization (MRFO), an advanced metaheuristic optimization technique. This optimization process is instrumental in enhancing the feature extractor, primarily based on the VGG16 model. Within this fine-tuned model, a series of enhancements have been implemented. A global average pooling layer and a residual connection block, both optimized using MRFO, constitute the core elements of this model. To further elevate its performance, a strategically placed skip connection links the first dense layer to the subsequent layer. This integration is effectively utilized in the final layer of our model.

The aim of this work is to build a fine-tuning model optimized that achieves high accuracy with low computational power. The proposed model achieves better accuracy than other Transfer Learning (TL) models when applied to different pathologies in Binary and Multiclass Classification

4.2 Recent Work

In recent years, Computer-Aided Diagnosis (CAD) systems have been developed for diagnosing various illnesses. In this work, we are interested in three types of pathologies. Our motivation is based on different reasons presented in the following. Each disease has different subtypes, such as lung cancer with four types, breast cancer with three types, and Alzheimer disease with four types. Lung cancer is a deadly and severe disease that can go undetected until it has spread. Symptoms of lung cancer include coughing up blood, chest pain, hoarseness, loss

of appetite, shortness of breath, and fatigue. Early detection of lung cancer is absolutely vital for increasing the chances of survival and improving overall outcomes for patients. When lung cancer is diagnosed in its early stages, it is often localized, meaning it has not spread to other parts of the body. Early detection plays a crucial role in the successful treatment of lung cancer. Patients diagnosed in the early stages have a higher chance of successful treatment outcomes [76]. In Algeria, a National Cancer Plan for 2015-2019 was implemented to address the national priority of lung cancer. The most common types of lung cancer in Algeria are adenocarcinoma and squamous cell carcinoma. These malignancies predominantly afflict individuals with a history of smoking, whether current or former smokers. Adenocarcinoma, a cancer that originates in the glandular cells lining the airways, and squamous cell carcinoma, arising from the thin, flat cells covering the airway surfaces, pose significant health challenges in the nation [77]. Breast cancer is a leading cause of death among women, characterized by the uncontrolled growth of cells in the breast. Early detection is key to reducing mortality rates. AD is a type of dementia that is characterized by mild memory loss, which gradually progresses to a loss of the ability to carry on a conversation and respond to the environment. AD affects parts of the brain that control thought, memory, and language.

DL, through the use of convolutional neural networks (CNNs) and other techniques, has demonstrated the ability to recognize the underlying structure of data. These networks are adept at automatically learning features from raw data, enabling them to discern subtle patterns, textures, and shapes within images. By employing multiple layers of convolutions and pooling operations, CNNs can extract hierarchical representations of visual information, leading to superior accuracy in tasks like object recognition, image segmentation, and classification [78]. With the rapid advancement of technology, ML has emerged as a powerful force in the field of radiology, transforming the way medical imaging data is analyzed and interpreted. The integration of ML algorithms into radiological practices has significantly enhanced the accuracy, efficiency, and speed of diagnosing various medical conditions, including complex diseases and abnormalities[5]. TL models have shown hopeful performance in detecting different diseases, using medical images.

In the field of medical research and disease detection, TL has been widely utilized to leverage pre-trained models and enhance the accuracy and efficiency of disease detection algorithms. TL allows the knowledge learned from one task (often on a large dataset) to be applied to a different but related task, even with limited data availability. While numerous techniques have been proposed to improve disease detection using TL models, there remains a

notable gap in comprehensive experimental studies exploring the impact of metaheuristic algorithms on fine-tuning models. Metaheuristic algorithms are powerful optimization techniques that can effectively search through complex solution spaces to find optimal or near-optimal solutions for specific problems. In the context of fine-tuning TL models for disease detection, the application of metaheuristic algorithms holds significant potential to optimize hyperparameters and improve the overall performance of the models. By carefully fine-tuning the TL models using metaheuristic algorithms, researchers can unlock hidden patterns and relationships within medical datasets, leading to more accurate and reliable disease detection systems.

In this context, we propose a collection of experiments to compare the performance of VGG16 with a fine-tuning model optimized by the MRFO algorithm. Additionally, the MRFO algorithm optimizes parameters and hyperparameters such as data augmentation parameters, batch size, optimizer, TL layers trainable ratio, and dropout ratio. This work also targets to rummage deep into effective studies optimizing the fine-tuning model techniques that can increase the performance of diagnostic systems based on DL.

This work proposes a novel framework to detect lung cancer, breast cancer, and alzheimer's disease from CT scan, ultrasound scan and MRI images. Our contribution is inspired by the work of reference [79], which utilized eight pre-trained CNN architectures with TL and metaheuristic optimization techniques, specifically the Manta Ray Foraging Optimization (MRFO) approach, to optimize TL parameters and hyperparameters.

The main contributions of the proposed method are as follows:

The contribution acts directly on the architecture of the parameter extraction phase by adding several layers based on a data reduction and a metaheuristic optimization algorithm in order to select the optimal parameters for a better classification. The important phases of the contribution are listed as follows:

- Features extraction from different types of disease using VGG16 TL network ;
- Data reduction : Flatten or Global average pooling layer was used after VGG16 for local and global features extraction to find which one is best ;
- Hyperparameters optimization based metaheuristic MRFO algorithm on different databases ;
- VGG16 features optimization based MRFO algorithm and added residual block to obtain high accuracy;
- Dense layers addition and their MRFO optimization ;

- Concatenation of optimized dense layers (added) ;
- Selection of best features concatenated by the final dense layer optimized ;
- Add dropout layer and optimize it ;

DL models are categorized into two main types: non-pre-trained and pre-trained models: Pre-trained models tend to exhibit higher performance compared to non-pre-trained models because they have been initially trained on large datasets, enabling them to capture rich and generalizable features. By leveraging this pre-existing knowledge, pre-trained models can avoid overfitting and demonstrate better performance on specific tasks even with limited data. In contrast, non-pre-trained models start from random initialization and require significant amounts of data to learn meaningful representations, which may lead to challenges in achieving comparable performance. Thus, the use of pre-trained models with fine-tuning is a common practice to boost the efficiency and accuracy of DL models across various applications.

4.2.1 Lung cancer classification based methods

Lung cancer classification refers to the task of distinguishing between different types of lung cancer or identifying whether a patient's lung tissue is cancerous or non-cancerous based on medical imaging or other relevant data. Over the years, various methods and techniques have been developed for lung cancer classification. Lakshmanaprabu, S. K., et al. [80] proposed a framework called Optimal Deep Neural Network (ODNN) optimized by Modified Gravitational Search Algorithm (MGSA) and Linear Discriminate Analysis (LDA) to analyze and classify the CT scan lung nodules images. Their results showed that the proposed classifier gives a sensitivity of 96.2%, specificity of 94.2%, and accuracy of 94.56%. Hammad, Mohamed, et al. [81] achieved high accuracy of 99.99% with a new deep CNN for detecting normal and abnormal cancer images database. Civit-Masot, et al. [82] proposed a diagnostic aid system (DAS) to detect healthy and non-healthy images from lung cancer tissue. They obtained an accuracy of 97.11 and 99.96%, and a value of the area under ROC curve of 99.77 and 99.94% depending on the number of classes classified. Ajai, et al. [83] proposed a method called Shuffled Social Sky Optimizer-based Multi-Object Rectified Attention Network (SSSO-based MORAN) to classify lung cancer. They employed the Gaussian filtering to pre-process the image, and for feature extraction, they used Local Gabor XOR Pattern (LGXP), Gray-Level Co-occurrence Matrix (GLCM) features, Global Binary Pattern (GBP), Tetrolet transform, and statistical features. Their algorithm showed an accuracy, Mean Absolute Error (MAE), sensitivity, and specificity of 0.896, 0.104, 0.8969, and 0.845, respectively. Xuan, Ping, et al [84] proposed a new convolutional bidirectional gated recurrent unit based module technique

to improve lung tumours' segmentation. Dodia, Shubham, et al [85] presented an overview of recent research performed in medical image analysis of lung cancer using DL algorithms, and Tomassini, Selene, et al. [86] presented a survey to show the contribution of convolutional neural networks for identifying malignant lung nodules and classifying lung cancer from computed tomography data. Tian, Qingji, et al. [87] proposed a new method to achieve high accuracy for lung cancer classification. they used a new version of a metaheuristic, called the Converged Search and Rescue (CSAR) algorithm to Enhance Capsule Networks (ECN) for the diagnosis. They showed that the suggested method achieved 96.35 % precision, 96.07 % recall, 96.41 % F1-score and 96.65 % accuracy. Siddiqui, et al [88] proposed an IGF-EDBN with SVM classifier approach for identifying and classifying lung CT images into normal, malignant and benign categories. The suggested method outperforms state-of-the-art techniques in terms of accuracy, sensitivity, specificity, F-1 score, false positive rate (FPR), and false negative rate (FNR). The IGF-EDBN approach achieves impressive performance metrics, including an F1 score of 99.37%, accuracy of 99.42%, sensitivity of 98.49%, and specificity of 98.31%. El Hamdi, et al [89] proposed a novel approach using PET/CT images and a multi-output CNN for lung cancer classification. The VGG16 network extracts relevant features, fed to a three-branch classifier for TN staging and histologic subtypes. Experimental results show good performance in TN staging and histology classification, achieving high accuracy (0.94) and AUC (0.97) for tumor size classification on Lung-PET-CT-Dx dataset. Table 4-1 presents recent studies on lung cancer classification.

Table 4-1 Recent studies on lung cancer classification

Authors	Years	Methodology	Results (%)
Lakshmanaprabu, S. K., et al. [80]	2019	ODNN Optimized by MGSA and LDA for CT scan Lung Nodules	Sensitivity= 96.2 Specificity= 94.2 Accuracy= 94.56
Tian, Qingji, et al.[87]	2021	CSAR Algorithm to Enhance Capsule Networks for Diagnosis	Precision = 96.35 Recall = 96.07 F1-score = 96.41 Accuracy = 96.65
Hammad, Mohamed, et al. [81]	2022	New CNN for Detecting Normal and Abnormal Cancer Images	Accuracy = 99.99
Civit-Masot, et al. [82]	2022	Diagnostic Aid System (DAS) for Healthy and Non-Healthy Images	Accuracy = 97.11/99.96 AUC = 99.77/99.94
Ajai, et al. [83]	2022	SSSO-based MORAN for Lung Cancer Classification	Accuracy = 89.6 MAE = 10.4 Sensitivity = 89.69 Specificity = 84.5

Siddiqui, et al [88]	2023	IGF-EDBN with SVM Classifier for Lung CT Image Classification	F1 score = 99.37 Accuracy = 99.42 Sensitivity = 98.49 Specificity = 98.31
El Hamdi, et al [89]	2023	Multi-Output CNN for Lung Cancer Classification using PET/CT	Accuracy = 94 AUC = 97

4.2.2 Breast cancer classification based methods

Various types of deep neural network architecture have been utilized for breast cancer classification using images. Murtaza, Ghulam, et al [90] reviewed different deep neural network models that were used for this purpose. M, Nawaz, et al [91] proposed a DL approach based on a CNN model for multi-class breast cancer classification. The DenseNet CNN model achieved high performance with an accuracy of 95.4% in the multi-class breast cancer classification task. Khan, SanaUllah, et al [92] proposed a framework based on CNN architectures of GoogLeNet, VGGNet, and ResNet. The combined features were fed into a fully connected layer for breast cancer classification. Han, Zhongyi, et al [93] proposed a new model called class structure-based deep convolutional neural network (CSDCNN) for breast cancer multi-classification. The structured DL model achieved high performance with an average accuracy of 93.2%. Omondiagbe, et al [94] presented three ML algorithms: Support Vector Machine, Artificial Neural Networks, and Naïve Bayes, and a hybrid approach using linear discriminant analysis (LDA) for breast cancer classification. The proposed approach achieved an accuracy of 98.82%. Baghdadi, et al. [79] proposed a framework based on convolutional neural networks, TL, and the MRFO for parameters and hyperparameter optimization using histopathological and ultrasound breast cancer datasets to classify breast cancer automatically with high performance. They achieved a score of 97.73% on the histopathological dataset and 99.01% on the ultrasound dataset in terms of accuracy. Boulenger, et al. [95] presented VGG19 architecture to predict breast cancer from ultrasound images. The model performs well with an accuracy of 85%, a sensitivity of 86%, and a specificity of 86%. Ali, Muhammad Danish, et al [96] presented a novel breast cancer classification approach with 90% accuracy on the BUSI dataset. They used multiple CNN models (Inception V3, ResNet50, DenseNet121) in a meta-learning framework for improved generalization and accuracy, especially in detecting malignant tumors. The study demonstrates the potential of meta-learning and ensemble techniques to enhance breast cancer diagnosis. Table 4-2 shows recent studies on breast cancer classification.

Table 4-2 Recent studies on Breast cancer classification

Authors	Years	Methodology	Results (%)
Han, Zhongyi, et al. [93]	2017	CSDCNN Model for Breast Cancer Multi-Classification	Average Accuracy = 93.2
M,Nawaz, et al. [91]	2018	CNN Model for Multi-Class Breast Cancer Classification	Accuracy = 95.4
Omondiagbe, et al. [94]	2019	SVM, ANN, Naïve Bayes, and hybrid Approach for Breast Cancer Classification	Accuracy = 98.82
Baghdadi, et al. [79]	2022	Framework with CNNs, TL, and MRFO for Breast Cancer Classification	HD: Accuracy = 97.73 US: Accuracy = 99.01
Boulenger, et al. [95]	2023	VGG19 Architecture for Predicting Breast Cancer from Ultrasound Images	Accuracy = 85 Sensitivity = 86 Specificity = 86
Ali, Muhammad Danish, et al. [96]	2023	Meta-Learning Framework with Multiple CNN models for Breast Cancer Classification	Accuracy = 90

4.2.3 Alzheimer disease classification based methods

MRI is a tool that is used to analyze the anatomical structures of the brain due to its high spatial resolution. The classification of Alzheimer disease using DL has gained significant attention in recent decades. ML has become a popular area of study in the health sciences, particularly in improving AD diagnosis and prognosis. This section presents a review of recent studies that utilize DL for AD diagnosis and prediction. Yamanakkanavar, et al. [97] provided an overview of current DL and ML approaches for brain MRI segmentation and classification of AD for high accuracy in detecting AD. Babu, et al. [98] introduced a novel AD diagnosis model with two main phases: proposed feature extraction and classification using Deep Convolutional Neural Network (DCNN) optimized by a new hybrid model termed as Combined Gray Wolf and Dragon Updating (CG-DU). They achieved an average accuracy of 98.79%, sensitivity of 98.67%, and specificity of 99.42%.

Islam, et al. [99] developed a DCNN model for four-class classification of AD based on MRI scans. The model was trained and evaluated using the OASIS dataset, achieving an accuracy of 73.75%. However, due to limited data, the model's accuracy was inadequate. Wen, Junhao, et al. [100] provided an overview of various DL methods used for AD classification and compares their performance on a publicly available dataset. The authors also provide a reproducible evaluation framework for benchmarking the performance of new models. Zhang, et al. [101] presented an extreme learning machine (ELM) classification model for binary AD, utilizing images of 627 patients from the ADNI database. In addition, a technique for multi-

class detection for AD and cognitive impairment stage was proposed. Ismail, et al.[102] proposed a novel technique combines MRI and PET images to diagnose AD into three groups. DNN models such as InceptionV3, AlexNet, and ResNet-18 were used for binary and multi-class classification tasks. The hyperparameters were optimized with a multi-objective algorithm to learn the distinguishing characteristics of AD. The features were classified using Softmax, SVM, and RF classifiers. The proposed method shows promise in the early detection of AD. They collected 1617 DICOM images of brain tissue from the Alzheimer Disease Neuroimaging Initiative (ADNI) database, which included 511 AD, 571 early MCI, and 535 NC subjects acquired from MRI and PET scans. The MultiAz-Net model combined with softmax achieved an accuracy of 90.13%, while MultiAz-Net combined with SVM achieved a specificity of 92.3%, and MultiAz-Net combined with RF achieved a sensitivity of 89.0%. Leela. M, et al.[103] proposed HEMRDTL, a hybrid EEG and fused CT-MRI model that uses TL, VGG19 techniques, and RPCA for automatic and early detection of AD. The model extracts features from both EEG and fused CT-MRI signals for classification, representing the brain's functional and structural properties. Their approach outperformed several state-of-the-art methods for AD detection on a large dataset, showing promise for accurately detecting categories of AD from both fused CT-MRI and EEG signals. Table 4-3 presents recent studies on Alzheimer disease classification.

Table 4-3 Recent studies on Alzheimer disease classification

Authors	Years	Methodology	Results (%)
Islam, et al. [99]	2017	DCNN Model for Four-Class Classification of AD based on MRI Scans	Accuracy = 73.75
Zhang, et al.[101]	2019	ELM Classification Model for Binary AD and Multi-Class Detection using Images from the ADNI Database	–
Babu, et al.[98]	2022	AD Diagnosis Model with Feature Extraction and DCNN Optimized by CG-DU Model	Average Accuracy = 98.79 Sensitivity = 98.67 Specificity = 99.42
Ismail, et al. [102]	2023	Novel Technique Combining MRI and PET Images for AD Diagnosis into Three Groups; use of DNN Models (InceptionV3, AlexNet, ResNet-18)	Accuracy = 90.13 Specificity = 92.3 Sensitivity = 89.0

4.3 Proposed Optimized Deep Features (ODF) Method

The proposed method draws from TL principles, widely acknowledged for its efficacy across diverse medical applications. TL involves feature extraction and representation from a pre-trained model, eliminating the need to train anew. In this proposed approach, the VGG16 TL model is employed, where we modify the architecture by replacing fully connected layers with a new model featuring global average pooling. This alteration reduces computations and facilitates the extraction of global features. In Figure 4.1, the proposed TL model showcases the significance of pooling in dimension reduction, enhancing feature relevance while discarding unnecessary details [104]. Additionally, a residual connection block with two optimized dense layers is incorporated and concatenated to select optimal features. The last layer is a dropout layer optimized for preventing overfitting. This configuration is designed for the multi-class classification of various cancer types in the lung, breast, and AD.

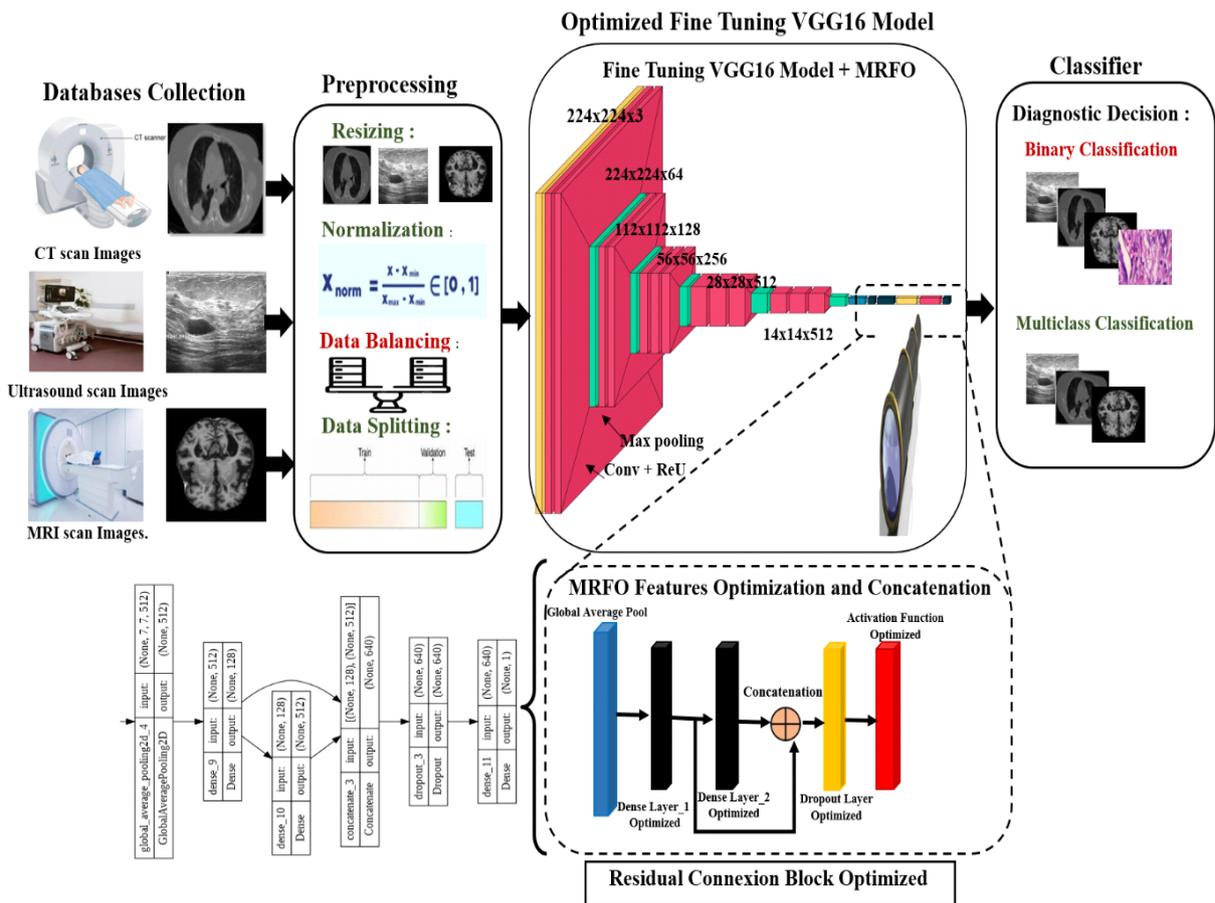


Figure 4.1 Proposed ODF Approach for Binary/Multiclass classification of different Diseases Classification

4.3.1 Preprocessing and Deep Features Extraction

The preprocessing of medical images and extracting the important characteristics using AI techniques are great challenges.

- **Resizing** : Resizing refers to the process of changing the dimensions of an image. In medical image preprocessing, resizing is often used to standardize the dimensions of images in a dataset. This step is crucial when dealing with images of different sizes, resolutions, or aspect ratios. In this work, all images were reshaped to (224x224x3).

- **Normalization** : Normalization is a preprocessing technique used to standardize the pixel values of images. In the context of medical image analysis, normalization typically involves scaling the pixel values to a specific range, such as [0, 1] or [-1, 1]. By ensuring that pixel values are within a consistent and standardized range, normalization prepares the images for more effective feature extraction and model training.

- **Data Splitting** : In the context of medical image analysis and ML, the dataset is usually divided into three sets: the training set, which is used to train the model; the validation set, which is used to fine-tune hyperparameters and prevent overfitting; and the test set, which is used to evaluate the model's performance.

4.3.2 Deep Features Extraction

Different methods for detecting and extracting the relevant information from lung cancer, breast cancer, and alzheimer's disease were used, for example, traditional ML like SVM [88], ANN [94], and DL techniques like CNN [89] architecture. Today, TL techniques are being broadly used in the classification problem. The assistance of Optimal Deep Neural Network (ODNN) and Linear Discriminate Analysis (LDA) has been used for extracting the deep features from CT scan of lung images. In [105] a review of recent state-of-the-art DL algorithms and architectures proposed as CAD systems for lung cancer detection was presented.

In this work, the VGG16 model is used for feature extraction. The architecture of VGG16, depicted in Figure 4.2, comprising 16 layers, including 13 convolutional layers and three fully connected layers, VGG16 employs small 3x3 filters with a stride of 1 in its convolutional layers. The use of max-pooling layers with 2x2 filters and a stride of 2 contribute to down-sampling, while the fully connected layers are equipped with ReLU activation functions.

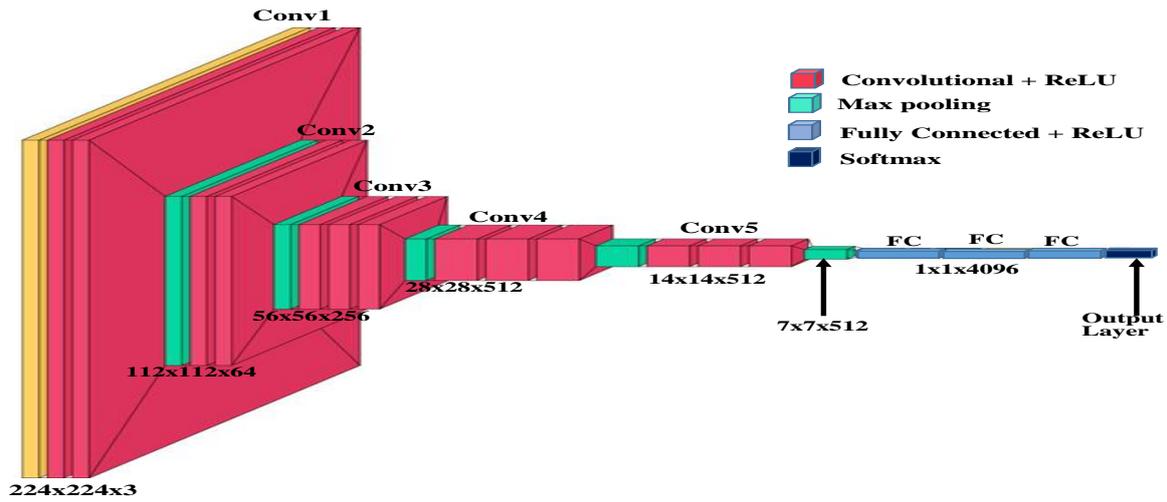


Figure 4.2 Used VGG16 Architecture

4.3.3 Fine Tuning Model

The proposed model is built upon the VGG16 architecture, with a specific focus on fine-tuning and optimization using the MRFO algorithm. To enhance the model's adaptability to our specific task, we have removed the fully connected layers of the original VGG16 (See Fig 4.3). In place of the removed fully connected layers, we have introduced a novel architecture. This new structure comprises a global average pooling layer and a residual block, which, in turn, consists of two dense layers. These dense layers are carefully optimized and concatenated to ensure effective feature extraction and representation. To address potential overfitting concerns, a dropout layer has been incorporated at the end of the architecture. This strategic addition helps mitigate the risk of overfitting, thereby enhancing the model's generalization capabilities. Through this modification and optimization process, we aim to achieve superior performance and accuracy in our specific application

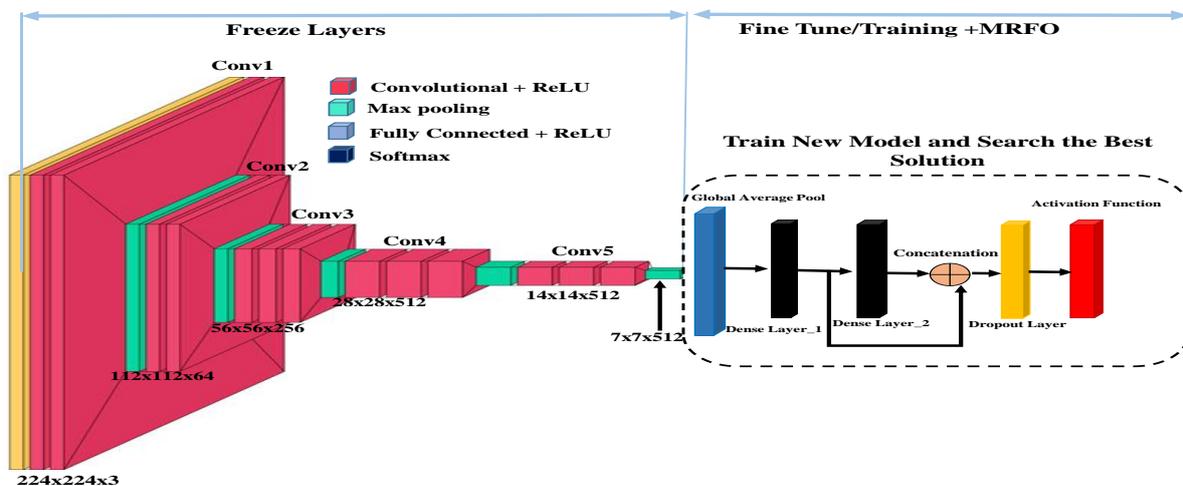


Figure 4.3 Fine Tuning VGG16 Model and Searching the Best Solution

4.3.4 Deep Features Optimization

There are diverse types of optimizers that can be used in DL. Some common optimizers include Stochastic Gradient Descent (SGD), Adam, RMSprop, Nadam, and Adagrad, each with its own strengths and weaknesses. In this work, for each experience we use the MRFO algorithm to choose the best between them. The algorithm uses a combination of cyclone, random, chain, and somersault foraging strategies for exploration and exploitation in the search space, aiming to find the best solution for the optimization problem. The selected solution represents the optimal set of hyperparameters for the VGG16 TL + ODF model, as determined by the Manta Ray Foraging Optimization process.

4.3.4.1 Manta Ray Foraging Optimization

Metaheuristic algorithms are an effective approach for finding accurate solutions to real-world problems. One such algorithm that is commonly used in DL to optimize hyperparameters and improve performance is the MRFO algorithm. Inspired by the intelligent behaviors of Manta Rays, this algorithm has become increasingly popular in DL research. Its effectiveness in optimizing hyperparameters has been demonstrated in several studies, making it a valuable tool for improving DL performance. Overall, the MRFO algorithm is a promising addition to the field of metaheuristics and has the potential to be utilized in a wide range of real-world applications [106].

- **Chain Foraging**

This technique is based on rays of Manta to find the plankton and move toward them. can be presented mathematically as follows [107]:

$$x_i^d(t+1) = x_i^d(t) + r * (x_{best}^d(t) - x_i^d(t)) + \alpha * (x_{best}^d(t) - x_i^d(t)), \quad i = 1 \quad (4.1)$$

$$x_i^d(t+1) = x_i^d(t) + r * (x_{i-1}^d(t) - x_i^d(t)) + \alpha * (x_{best}^d(t) - x_i^d(t)), \quad i = 2, \dots, N \quad (4.2)$$

$$\alpha = 2 * r * \sqrt{|\log(r)|} \quad (4.3)$$

x_i^d is the placement of i^{th} single in the d^{th} dimension at time t , α is a weight coefficient, r is a random vector in $[0, 1]$. x_{best}^d is the position of high-concentration planktons.

- **Cyclone Foraging**

In this case, every Manta swims towards the one in front of it pointing spirally moving to the food, wherever this activity behavior may be expanded for n-D space. The mathematical equation can be expressed as:

$$x_i^d(t+1) = x_{best}^d(t) + r * (x_{best}^d(t) - x_i^d(t)) + \beta * (x_{best}^d(t) - x_i^d(t)), i = 1 \quad (4.4)$$

$$x_i^d(t+1) = x_{best}^d(t) + r * (x_{i-1}^d(t) - x_i^d(t)) + \beta * (x_{best}^d(t) - x_i^d(t)), i = 2, \dots, N \quad (4.5)$$

$$\beta = 2e^{r_1 \frac{T-t+1}{T}} * \sin(2\pi r_1)$$

T represents the max iteration, β is the weight coefficient, and r_1 a rand number in the range [0, 1]. Different technique based on making each search for a new position from the current best position by assigning a random position to detect the search space:

$$x_{rand}^d = Lb^d + r * (Ub^d - Lb^d) \quad (6)$$

$$x_i^d(t+1) = x_{rand}^d(t) + r * (x_{bestrand}^d(t) - x_i^d(t)) + \beta * (x_{rand}^d(t) - x_i^d(t)), i = 1 \quad (4.7)$$

$$x_i^d(t+1) = x_{rand}^d(t) + r * (x_{i-1}^d(t) - x_i^d(t)) + \beta * (x_{bestrand}^d(t) - x_i^d(t)), i = 2, \dots, N \quad (4.8)$$

- **SomerSault Foraging**

In this situation, Manta swims in an axis way and updates their positions. The mathematical equation [107] can be presented as follow:

$$x_i^d(t+1) = x_i^d(t) + S * (r_2 * x_b^d - r_3 * x_i^d(t)), i = 1, \dots, n \quad (4.9)$$

S represents the somersault factor, r_2 and r_3 represent two random numbers in [0, 1].

Algorithm 3 outlines the various steps of MRFO optimization

Algorithm 3 : MRFO Optimization*Initial Setup and Population Initialization***Input:**

- Number of solutions in the population (Population Size)
- Maximum iterations for MRFO (Max Iterations)
- Training, validation, and test datasets
- Problem-specific parameters and constraints

Output:

- Best hyperparameters obtained from MRFO process

Initialize Population:

- Randomly initialize a population of solutions based on problem-specific parameters and constraints;
- Each solution represents a set of hyperparameters for the VGG16 TL + ODF model.

Define Fitness Function:

- Implement a fitness function to evaluate the performance of each solution;
- The fitness function evaluates the model's performance on the validation dataset.

MRFO Optimization:*For $t=1$ to Max Iterations:*

- Evaluate fitness for each solution in the population using the Fitness Function;
- Apply MRFO foraging strategies to update solutions in the population;
- Continue optimizing solutions within the specified maximum iterations.

Select Best Solution:

- After completing the iterations, select the solution with the highest fitness score as the best hyperparameters configuration.

Output:

- Return the best solution, which represents the optimal hyperparameters obtained from the MRFO process.

4.4 Classification

4.4.1 Binary classification

To evaluate the efficacy of the proposed model, a series of experiments are conducted across three distinct datasets, specifically targeting Lung Cancer, Breast Cancer, and Alzheimer Disease. The focus of our assessment centered on binary classification, wherein the model's performance was rigorously tested and analyzed against each dataset. This method allowed us to assess the model's ability to accurately distinguish between positive and negative instances within the context of Lung Cancer, Breast Cancer, and Alzheimer Disease datasets, providing a thorough and nuanced evaluation of its classification capabilities.

Data balancing: In the case of binary classification when the program searches the hyperparameters optimization, dataset balancing is used to enhance the results. This technique is used for Lung and Breast Cancer datasets:

For Lung Cancer dataset: The original dataset [108] has 159 normal images, 326 Adenocarcinoma images, 163 Large Cell Carcinoma, and 252 Squamous Cell. After applying data augmentation, the total number of any classes were increased to 326. Therefore, the final image dataset was (326 normal images, 326 Adenocarcinoma images, 326 Large cell carcinoma, and 326 Squamous Cell), and in case of large cell carcinoma and squamous cell, the total number of images was (260 Large Cell Carcinoma, and 260 Squamous Cell).

For Breast Cancer dataset: The database has 437 benign images, 210 malignant images, 133 normal. After applying data augmentation, the total number of any classes were increased to 437. Therefore, the final image dataset was (437 benign images, 437 malignant images, 437 normal).

4.4.2 Multiclass Classification

To assess the model, diverse experiments are conducted across three datasets (Lung Cancer, Breast Cancer, and Alzheimer Disease), employing Multiclass Classification. Our objective is to evaluate the model's performance across various scenarios. Algorithm 4 outlines the distinct steps for both binary and Multiclass Classification.

Algorithm 4: Binary/Multiclass Classification

Input: (224x224x3) dimensional Chest CT, Ultrasound or IMR images.

- Split dataset into Training, Validation, and Test sets;
- Population size N, Maximum iterations.

VGG16 TL+ ODF (Optimized Dense Features) model.

Output:

- Best solution found through MRFO Algorithm.

Manta Ray Foraging Optimization (MRFO) Algorithm:

Initialization:

- Initialize a population of N solutions (each represented as a search agent);
- Set maximum number of iterations;
- Define the VGG16 TL + ODF model architecture;
- Initialize constants

Fitness Evaluation:

- Evaluate the fitness score for each search agent using the defined model and the fitness function.

MRFO Algorithm:

For t=1 to Maximum Iterations:

For i=1 to N

Randomly generate a threshold *rand*

If (*rand* < 0.5): (cyclone foraging)

If $t/T_{\max} < rand$: //Update agent $x_i^d(t+1)$ using cyclone foraging formula

$$x_i^d(t+1) = x_{best}^d(t) + r * (x_{best}^d(t) - x_i^d(t)) + \beta * (x_{best}^d(t) - x_i^d(t)), i = 1$$

$$x_i^d(t+1) = x_{best}^d(t) + r * (x_{i-1}^d(t) - x_i^d(t)) + \beta * (x_{best}^d(t) - x_i^d(t)), i = 2, \dots, N$$

Else // Update agent $x_i^d(t+1)$ using random foraging formula.

$$x_i^d(t+1) = x_{rand}^d(t) + r * (x_{bestrand}^d(t) - x_i^d(t)) + \beta * (x_{rand}^d(t) - x_i^d(t)), i = 1$$

$$x_i^d(t+1) = x_{rand}^d(t) + r * (x_{i-1}^d(t) - x_i^d(t)) + \beta * (x_{bestrand}^d(t) - x_i^d(t)), i = 2, \dots, N$$

End If

Else // (chain foraging): Update agent $x_i^d(t+1)$ using chain foraging formula.

$$x_i^d(t+1) = x_i^d(t) + r * (x_{best}^d(t) - x_i^d(t)) + \alpha * (x_{best}^d(t) - x_i^d(t)), i = 1$$

$$x_i^d(t+1) = x_i^d(t) + r * (x_{i-1}^d(t) - x_i^d(t)) + \alpha * (x_{best}^d(t) - x_i^d(t)), i = 2, \dots, N$$

End If

Compute fitness for the updated individual

Perform somersault foraging for each agent:

Update agent using somersault foraging formula.

Compute fitness for the updated individual

End For

End For

Selection: Select the best solution from all the search agents based on their fitness scores.

Output: Return the best solution found through MRFO algorithm

4.5 Model Training with best solution and concatenation

Following the meticulous selection of optimal solutions through MRFO, our model underwent training across three datasets (Lung Cancer, Breast Cancer, and Alzheimer Disease). During this training phase, the finest parameters and hyperparameters are employed, meticulously derived through the application of MRFO to our model. This iterative

optimization process allowed us to fine-tune the model's configuration to achieve optimal performance for each dataset. Furthermore, the model is subjected to a comprehensive evaluation, employing a diverse array of metrics.

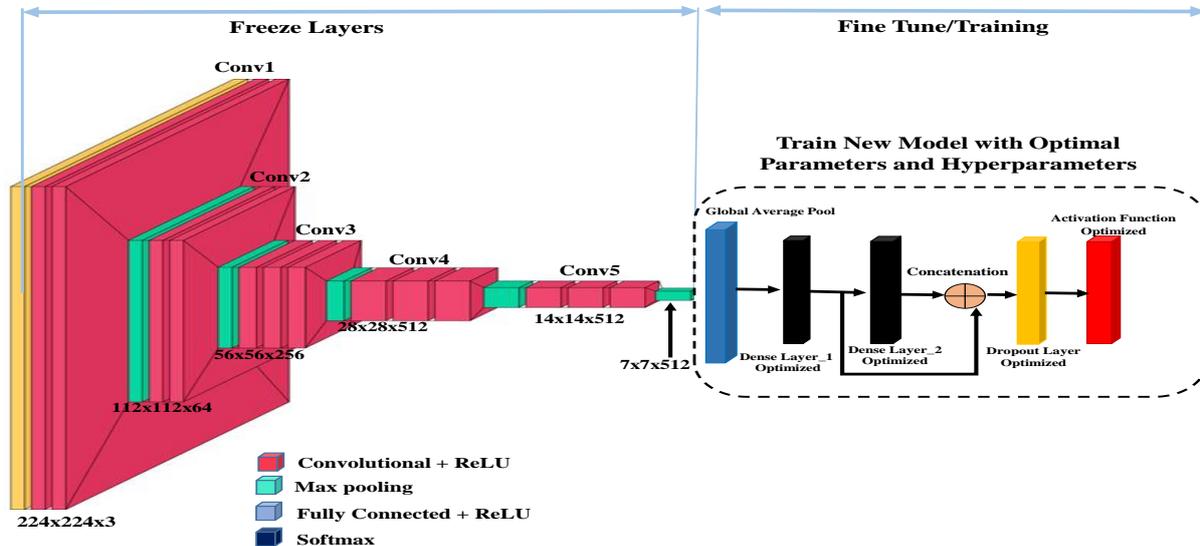


Figure 4.4 Proposed approach with best solution

4.5.1 Data Collection

In this study, four public datasets representing various pathologies are utilized.

- **Public Dataset for Lung Cancer**

The dataset contains three types of Chest Cancer (See Fig 4.5 (a)) Adenocarcinoma, Large Cell Carcinoma, and Squamous Cell Carcinoma. Additionally, there is one folder dedicated to Normal Cells. This dataset is publicly available on Kaggle repository [108].

- **Public Dataset for Breast Cancer**

Dataset_1 consists of ultrasound scans for three types of Breast Cancer: normal, benign, and malignant (See Fig 4.5(b)). This dataset is publicly available on Kaggle repository [109] and from reference [110]. The dataset, collected in 2018, includes breast ultrasound images from 600 women aged between 25 and 75 years. It contains 780 images with an average size of 500x500 pixels, saved in PNG format. The ground truth images are presented alongside the original images.

Dataset_2 is a Breast Cancer histological images contains two types of cancer: benign and malignant (See Fig 4.5(c)). Dataset is public available in Kaggle repository [111]. The dataset consists of 2479 images benign and 5304 images malignant.

▪ *Public Dataset for Alzheimer Disease*

The dataset contains four types of Alzheimer Disease (See Fig 4.5(d)): Mild Demented, Moderate Demented, Very Mild Demented and Non Demented. The dataset is public available in Kaggle repository [112]. Fig 4.5 shows the different dataset. Algorithm 5 shows the model training with best solution initialization.

Algorithm 5: Model Training with Best Solution

Input:

- *BestSolution:* Optimal hyperparameters obtained from the MRFO algorithm
- *Training dataset:* Chest CT, Ultrasound, or IMR images of size 224×224×3
- Split dataset into Training, Validation, and Test sets.

Output:

- Trained and Optimized VGG16 TL + ODF Model

Configure Data:

• Preprocess the Training, Validation, and Test datasets based on the chosen hyperparameters from the best solution.

- Apply normalization to the image data.
- Implement data augmentation techniques based on the specified augmentation parameters.
- Prepare the datasets as input for model training.

Build the Model:

• Initialize the VGG16 TL model with pre-trained ImageNet weights.

• Remove the top (classification) layers of the VGG16 model.

• Add additional dense layers (ODF) to the model architecture based on the best solution's hyperparameters.

- Add an output layer with appropriate activation function for the specific task (e.g., sigmoid for binary classification and softmax for multiclass).

Compile the Model:

• Compile the model with binary cross-entropy loss for binary classification or appropriate loss function for the task.

- Use the optimizer specified in the best solution along with its learning rate.
- Choose appropriate evaluation metrics

Train the Model:

- Train the model using the preprocessed training dataset.
- Train the model for a specified number of epochs based on the best solution's hyperparameters.
- Use the specified batch size during training.

Evaluate the Model:

- Evaluate the trained model's performance using the test dataset to obtain final metrics.

Output: The trained VGG16 TL + ODF model optimized using the best solution from the MRFO process.

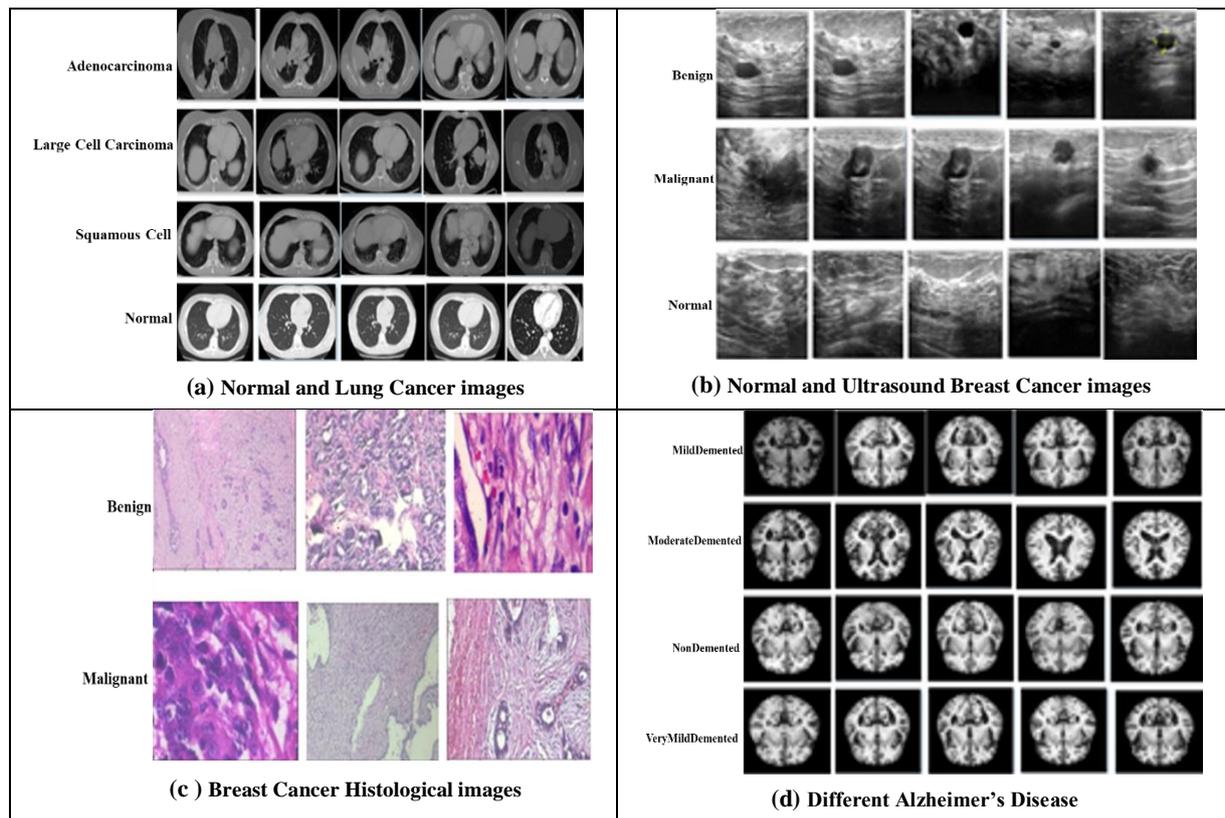


Figure 4.5 Different Dataset Collection

4.6 ODF Results and Discussions

In this section, we present a series of experiments conducted with diverse datasets. In each experiment, hyperparameters undergo optimization to enhance accuracy. To achieve heightened precision in performance, the MRFO algorithm is employed to search for the best hyperparameters as detailed in Table 4-4.

Table 4-4 Hyperparameters Optimized by MRFO

Hyper Parameters	Search Range	Results											
		Lung Cancer				Breast Cancer			Alzheimer Disease				
Optimizer	Adam, Nadam, RMSProp, Adadelata, SGD	Exp1	Exp2	Exp3	Exp4	Exp1	Exp2	Exp3	Exp1	Exp2	Exp3	Exp4	
			Nadam	Adam	RMS prop	RMS prop	Adam	Adam	Adam	Adam	SGD	Adam	SGD
batch size	8, 16, 32, 64	8	8	8	8	16	8	8	8	8	8	8	
dense_1	1024, 512, 256, 128	512	1024	1024	1024	1024	1024	1024	128	512	1024	512	
dense_2	512, 256, 128, 64	512	512	512	512	128	512	512	128	512	512	512	
dense_3	512, 256, 128, 64	256	512	256	512	128	512	512	64	512	512	512	
dropout	0→5 (step = 1)	0.1	0.1	0.1	0.1	0.06	0.1	0.1	0.1	0.06	0.1	0.06	

The protocol utilized for the 2-class classification is outlined in Table 4-5, along with details about the various datasets employed in this study.

Table 4-5 Protocol for Binary Classification

Datasets	Classes	Total Images	Training Set	Testing Set
Lung Cancer	'adenocarcinoma', 'normal'	659	527	132
	'adenocarcinoma', 'large_cell_carcinoma'	660	528	132
	'adenocarcinoma', 'squamous_cell_carcinoma'	796	636	160
	'large_cell_carcinoma', 'squamous_cell_carcinoma'	564	451	113
Breast Cancer	'benign', 'malignant'	863	690	173
	'benign', 'normal'	786	628	158
	'malignant', 'normal'	563	450	113
Alzheimer Disease	'MildDemented', 'ModerateDemented'	607	485	122
	'MildDemented', 'NonDemented'	1097	877	220
	'MildDemented', 'VeryMildDemented'	1108	886	222
	'ModerateDemented', 'NonDemented'	618	494	124

4.6.1 Binary Classification

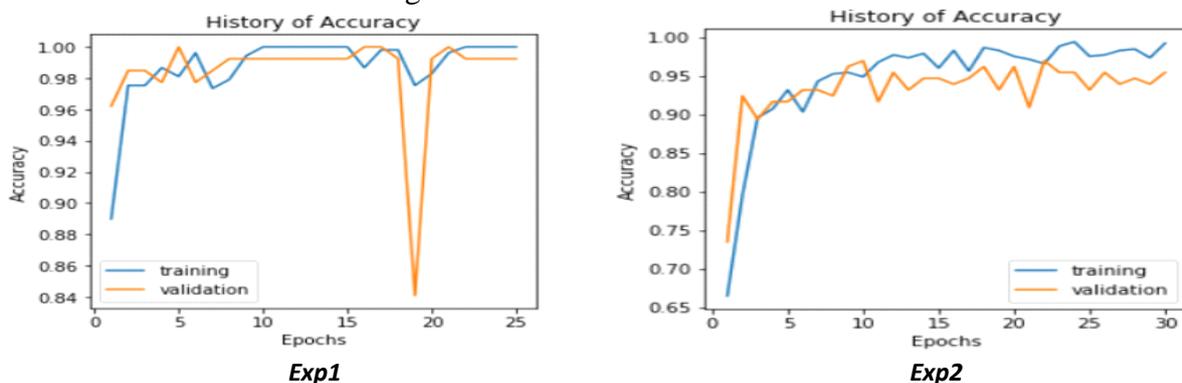
4.6.1.1 CT Images Lung Cancer

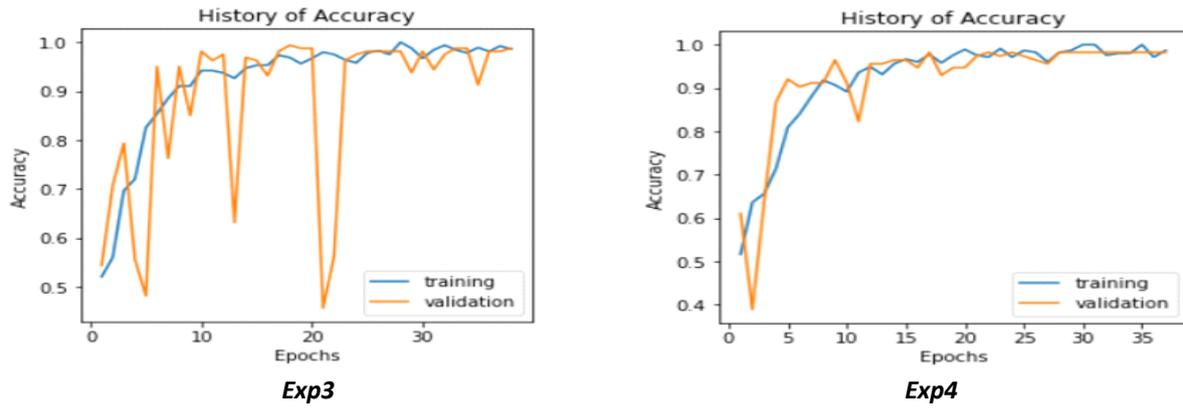
Table 4-6 present different results related to the binary dataset of Lung Cancer, with the optimal solutions derived from the proposed model after the training and optimization process.

Table 4-6 Results with different Lung Cancer types for Binary Classification

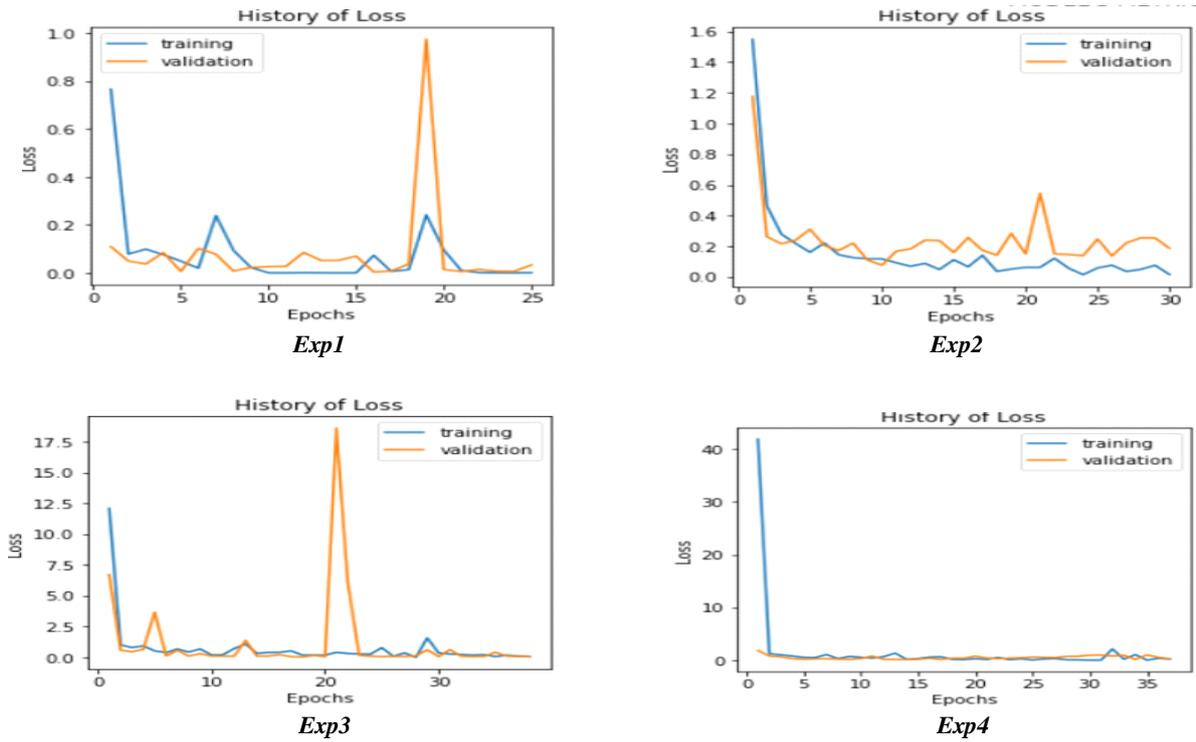
	Acc (%)	Loss (%)	Precision (%)	AUC (%)	Val_Acc (%)	Val_Loss (%)	Val_Precision (%)	Val_AUC (%)	Test_Acc (%)
<i>Exp1</i> Adenocarcinoma/Normal	100	0.0008	100	100	99.24	0.032	99.24	99.98	99.24
<i>Exp2</i> Adenocarcinoma/ Large_Cell_Carcinoma	99.24	0.016	99.24	99.99	96.97	0.076	96.97	99.62	77.27
<i>Exp3</i> Adenocarcinoma/ Squamous_Cell_Carcinoma	98.58	0.048	98.58	99.66	98.75	0.069	98.75	99.34	80.62
<i>Exp4</i> Large_Cell_Carcinoma/ Squamous_Cell_Carcinoma	98.67	0.209	98.67	98.80	98.23	0.122	98.23	98.78	82.30

The best curves are shown in Figure 4.6.





(a) Accuracy curves of different experiences



(b) Loss curves of different experiences

Exp1: Adenocarcinoma/Normal *Exp2: Adenocarcinoma/Large_cell_carcinoma* *Exp3: Adenocarcinoma/Squamous_cell_carcinoma* *Exp4 Large_cell_carcinoma/Squamous_cell_carcinoma*

Figure 4.6 Accuracy and Loss curves of different experiences on Lung Cancer (CT Images) dataset

In this work, we apply the optimized VGG16+ODF model to classify Lung Cancer images into two classes. The detailed metrics are outlined in Table 4-6, providing a comprehensive overview of the model's performance. Additionally, Figure 4.7 visually represents the model predictions through the confusion matrix. Upon analysis, the model demonstrates exceptional performance in classifying Exp1 (Adenocarcinoma/Normal) images, achieving 100% accuracy. However, challenges arise in more complex scenarios like Exp2 (Adenocarcinoma /Large_Cell_Carcinoma), Exp3 (Adenocarcinoma/Squamous_Cell_Carcinoma), and Exp4

(Large_Cell_Carcinoma /Squamous_Cell_Carcinoma). In these cases, the model exhibits a notable number of false positives, leading to slightly lower accuracy rates of 99.24%, 98.85%, and 98.67%, respectively. These findings indicate that while the proposed model effectively distinguishes Cancerous and Normal Lung images, it encounters difficulties in classifying between different types of Lung Cancers.

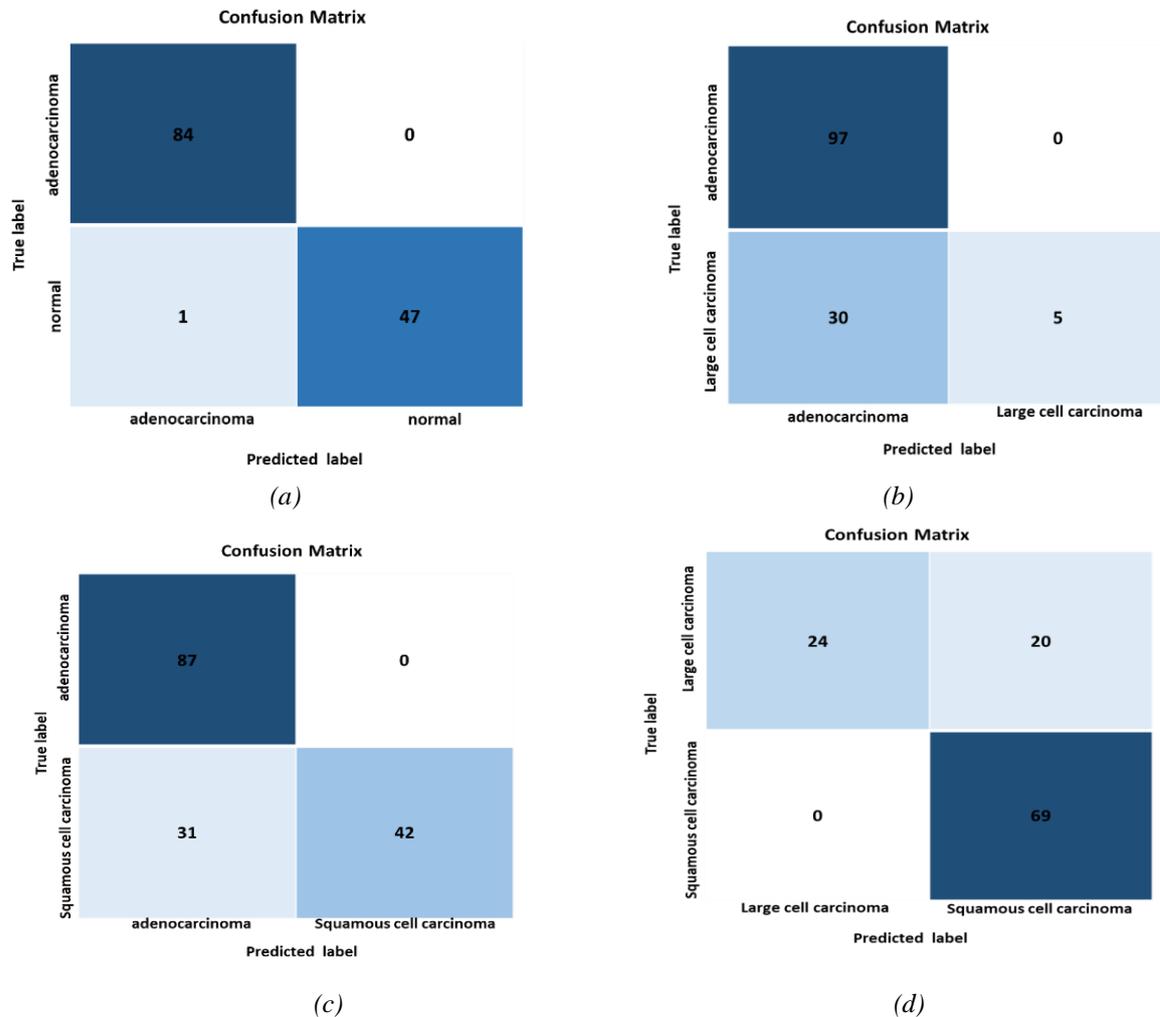


Figure 4.7 Confusion Matrixes for the VGG16+ODF with 2 classes on Lung Cancer CT dataset: (a) Exp1, (b) Exp2, (c) Exp3, (d) Exp4

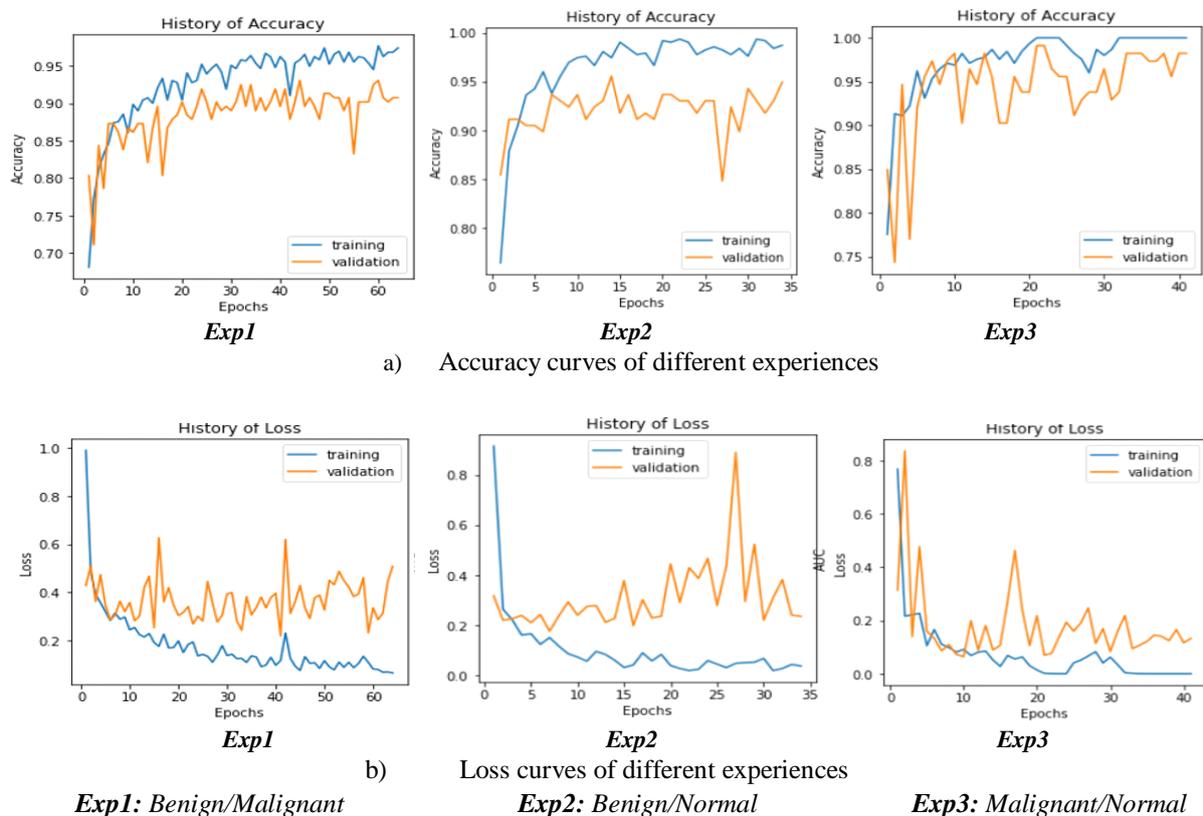
4.6.1.2 Breast Cancer images

- **Breast Cancer Ultrasound Images**

Table 4-7, Figures 4.8 and 4.9 present different results related to the binary dataset of Breast Cancer ultrasound images with the optimal solutions derived from the proposed model after the training and optimization process.

Table 4-7 Results with different Breast Cancer types for Binary Classification

	<i>Acc (%)</i>	<i>Loss (%)</i>	<i>Precision (%)</i>	<i>AUC (%)</i>	<i>Val_Acc (%)</i>	<i>Val_Loss (%)</i>	<i>Val_Precision (%)</i>	<i>Val_AUC (%)</i>	<i>Test_Acc (%)</i>
<i>Exp1 : Benign/malignant</i>	97.39	0.065	97.39	99.74	93.06	0.362	93.06	96.34	71.10
<i>Exp2 : Benign/normal</i>	98.73	0.037	98.73	99.90	95.57	0.227	95.57	97.52	79.11
<i>Exp3 : Malignant/normal</i>	100	0.00009	100	100	99.12	0.069	99.12	99.05	84.96

**Figure 4.8** Accuracy and Loss curves of different Binary Classification on Breast Cancer Ultrasound dataset

In the case of Breast Cancer dataset, a better accuracy was obtained with different experiences using the optimized VGG16+ODF with two classes; the metrics results are elaborated in Table 4-7. The confusion matrix obtained is shown in Fig 4.9.

According to the results, the model obtained a Val_accuracy of 93.06% and Test_accuracy of 71.10% and has a large number of false positives (in the case of benign and malignant). On the other hand, the model makes the lowest error in the case of (malignant and normal) with a Val_accuracy of 99.12% and Test_accuracy of 84.96% and has a best confusion matrix with lowest number of false positives.

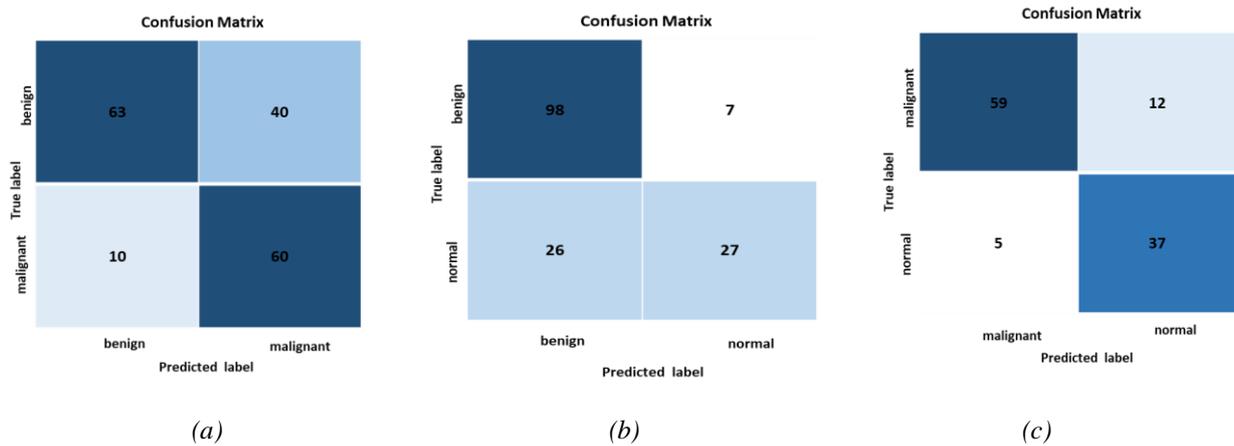


Figure 4.9 Confusion Matrixes for the VGG16+ODF with 2 classes on Breast Cancer Ultrasound dataset: (a) Exp1, (b) Exp2, (c) Exp3

At this stage of the study, a question arises: *Could histological data for breast cancer yield better results?*

In the following section, we attempt to address this question through the experiment outlined below.

- **Breast Cancer Histological Images**

Dataset: *Breast Cancer Histological dataset (texture color different)*

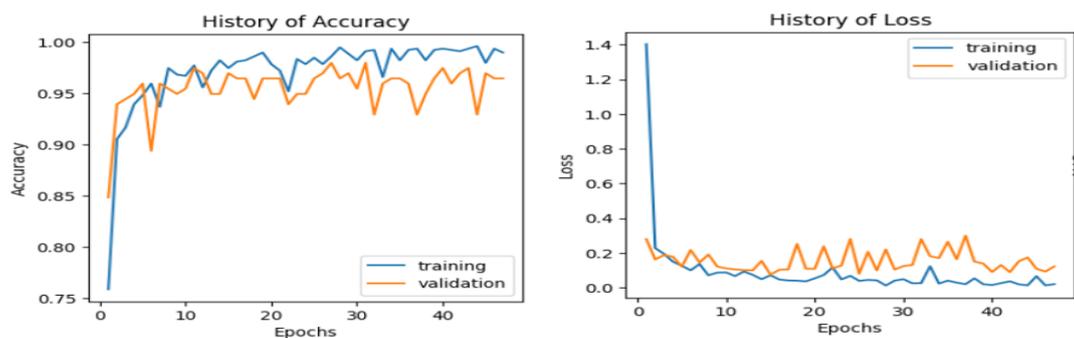
The dataset comprises 7909 Breast Cancer Histopathology images has been meticulously compiled from examinations conducted on 82 patients. These images are made publicly available and can be accessed from Kaggle repository [111]. This comprehensive dataset offers a diverse array of images, featuring cases from both benign and malignant categories. Specifically, there are 2479 images representing benign cases and 5304 images representing malignant cases.

In this work, the dataset was divided according to a specific protocol to ensure a balanced and fair evaluation of the proposed methods. The protocol stipulated a clear distribution of the dataset into two distinct categories: benign and malignant cases. There were 495 images representing benign cases and an equal number of 495 images representing malignant cases.

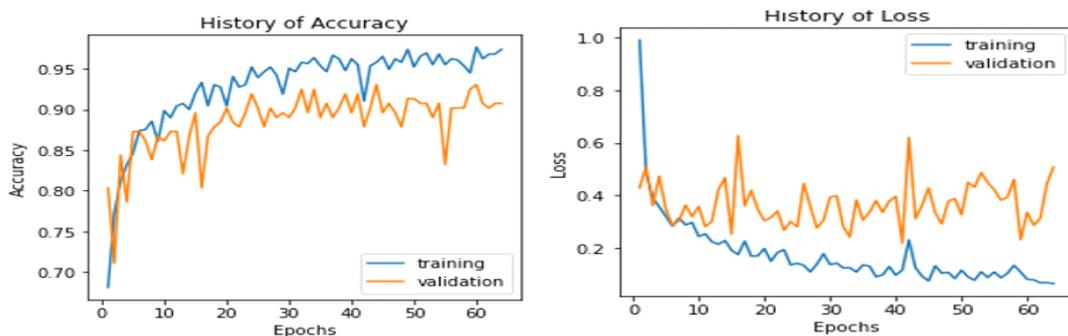
For the training phase, 80% of the data from both benign and malignant categories. The remaining 20% of the dataset was reserved for the testing phase. This protocol, with its balanced representation of both classes and clear division between training and testing sets, aimed to provide a robust and reliable framework for evaluating the effectiveness of the proposed methodologies in distinguishing between benign and malignant Breast Cancer cases. Table 4-8 and Figures 4.10 and 4.11 show the different results obtained through the implemented methods.

Table 4-8 Results with Breast Cancer Histological for Binary Classification

Classes	Acc (%)	Loss	Precision (%)	AUC (%)	Val_Acc (%)	Val_Loss	Val_Precision (%)	Val_AUC (%)	Test_Acc (%)
<u>Histologic</u> Benign/Malignant	98.99	0.0195	98.99	98.99	96.46	0.121	96.46	98.73	85.35
<u>Ultrasound</u> Benign/Malignant	97.39	0.065	97.39	99.74	93.06	0.362	93.06	96.34	71.10



(a) Breast Cancer histological images



(b) Breast Cancer ultrasound images

Figure 4.10 Accuracy and Loss curves of different experiences on Breast Cancer histological and ultrasound dataset

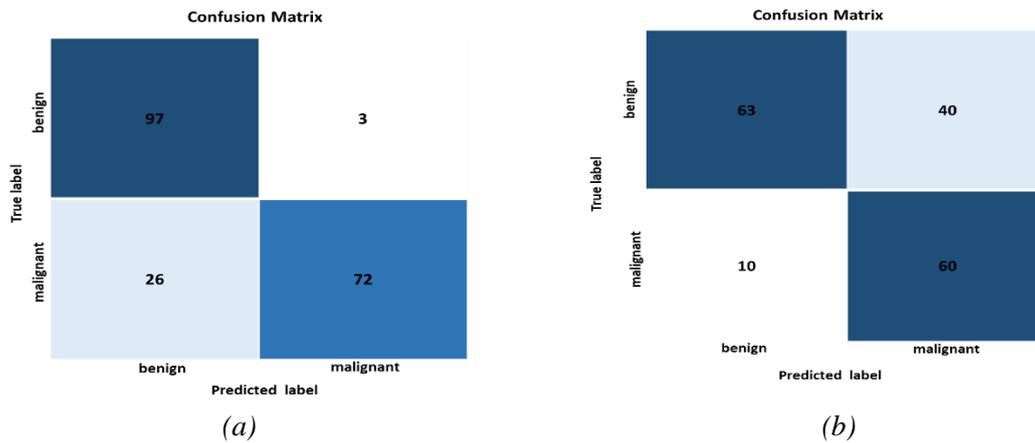


Figure 4.11 Confusion Matrices for Binary Classification with VGG16+ODF on Breast Cancer datasets: (a) *Histologic* and (b) *Ultrasound*

The results derived from the evaluation of both ultrasound and Histological Breast Cancer images, including the analysis of curves and confusion matrices, offer insights into the performance of our proposed methodology. In response to the inquiry regarding the efficacy of Breast Cancer Histological data, the exceptionally high accuracy values of 98.99% during both the training and validation phases serve as compelling evidence. These results affirm the superior performance of our approach in accurately classifying Breast Cancer Histopathology images when compared to Ultrasound images.

The Val accuracy for histologic images remains commendable at 96.46%, showcasing the robustness of the proposed approach. While slightly lower than the training and validation accuracies, this discrepancy is well within an acceptable range. The detailed breakdown of metrics, including precision, AUC, and other performance indicators, provides a comprehensive view of our model's efficacy in distinguishing between benign and malignant cases in histologic images. In contrast, ultrasound images exhibit a lower Val accuracy of 93.06%, indicating a noticeable performance gap compared to histologic images. Despite this disparity, it's crucial to emphasize that the ultrasound model still demonstrates a respectable level of accuracy and precision in classifying benign and malignant cases. Given the nearly identical sizes of the two datasets, a meaningful comparison can be drawn by assessing the occurrences of false positives and false negatives. Notably, this proposed model demonstrates consistent performance with comparable percentages of false positives and false negatives in both datasets.

In the histological dataset, the proposed model exhibits an ability to identify benign cancer, characterized by a minimal number of false positives in contrast to the ultrasound

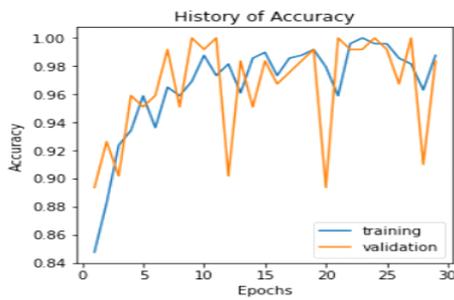
dataset. Furthermore, when it comes to detecting malignant cancer, the proposed model shows nearly identical percentages of false negatives across both datasets.

4.6.1.3 Alzheimer Disease MRI images

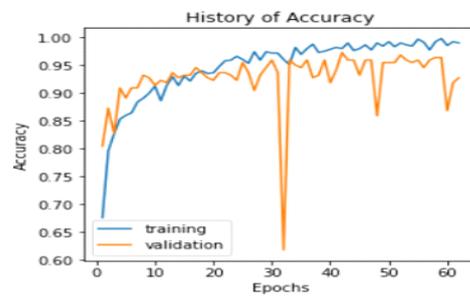
In our investigation of Alzheimer disease types, we observed varied outcomes with the proposed optimized model. The detailed results are presented in Table 4-9, Figures 4.12 and 4.13.

Table 4-9 Results with different Alzheimer Disease types for Binary Classification

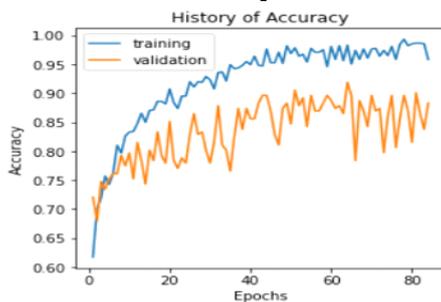
Classes	Acc (%)	Loss	Precision (%)	AUC (%)	Val_Acc (%)	Val_Loss	Val_Precision (%)	Val_AUC (%)	Test_Acc (%)
Exp1 MildDemented/ ModerateDemented	98.76	0.034	98.76	99.93	98.36	0.042	98.36	99.88	89.34
Exp2 MildDemented/ NonDemented	98.97	0.024	98.97	99.97	97.27	0.163	97.27	98.52	93.64
Exp3 MildDemented/ VeryMildDemented	95.82	0.09	95.82	99.52	91.89	0.407	91.89	95.30	72.07
Exp4 ModerateDemented/ NonDemented	98.99	0.017	98.99	99.98	99.19	0.016	99.19	99.98	90.32



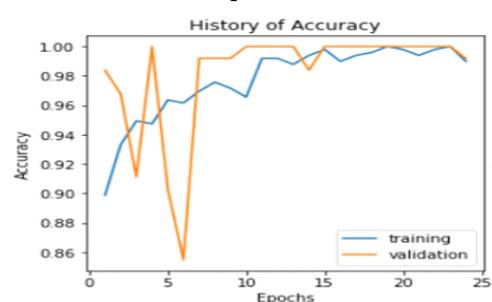
Exp1



Exp2



Exp3



Exp4

a) Accuracy curves of different experiences

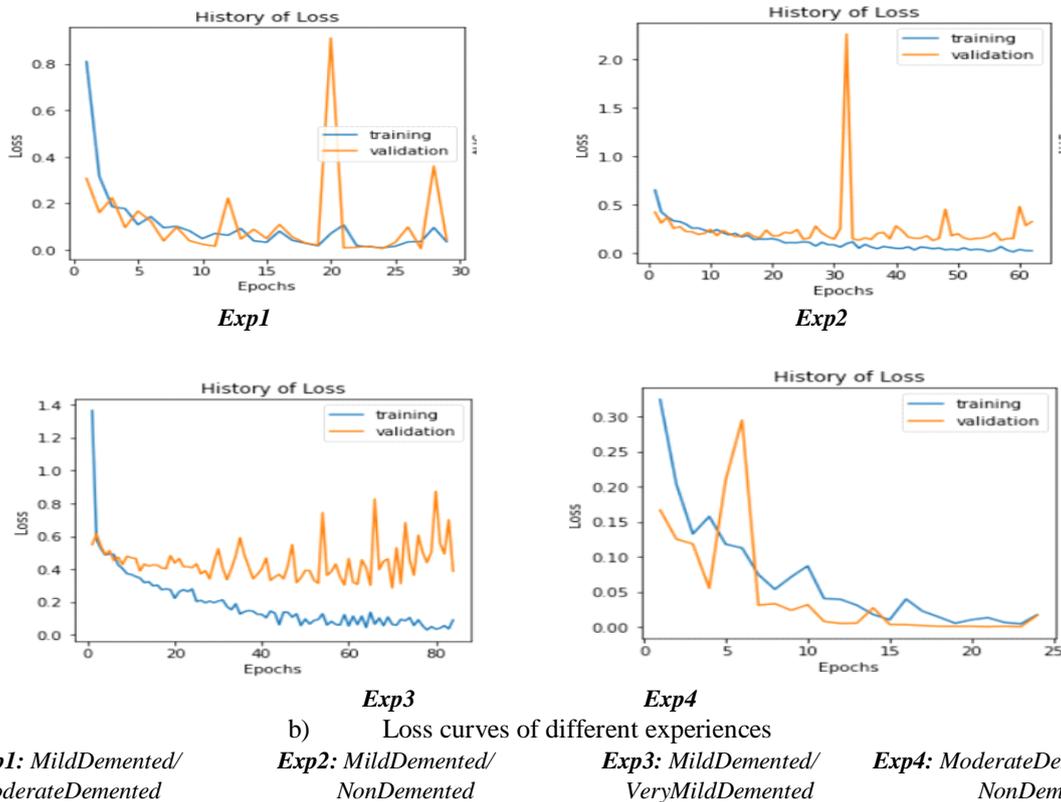


Figure 4.12 Accuracy and Loss curves for Binary Classification on Alzheimer Disease MRI dataset

In the case of the Alzheimer disease dataset, we achieved superior accuracy through various experiments utilizing the optimized VGG16+ODF model with two classes. The detailed metrics outlining the findings are comprehensively presented in Table 4-9, providing a comprehensive overview of the model's performance. Additionally, the confusion matrix, illustrating the model's predictions is depicted in Figure 4.13.

Upon scrutinizing the results, the model demonstrated a commendably low number of false positives, particularly in the cases of MildDemented and NonDemented classifications. This indicates the model's robust ability to accurately identify these categories. However, challenges arose when detecting ModerateDemented disease, as evidenced by higher error rates in experiments 2 and 4. This signifies a potential area for improvement, suggesting that the model may benefit from further optimization to enhance its accuracy in identifying this particular disease subtype. Furthermore, in the case of MildDemented and VeryMildDemented classifications, the model faced difficulties, leading to a considerable number of false negatives and false positives. This highlights the intricacies of distinguishing between these closely related categories, suggesting the need for fine-tuning and refinement to address these specific challenges.

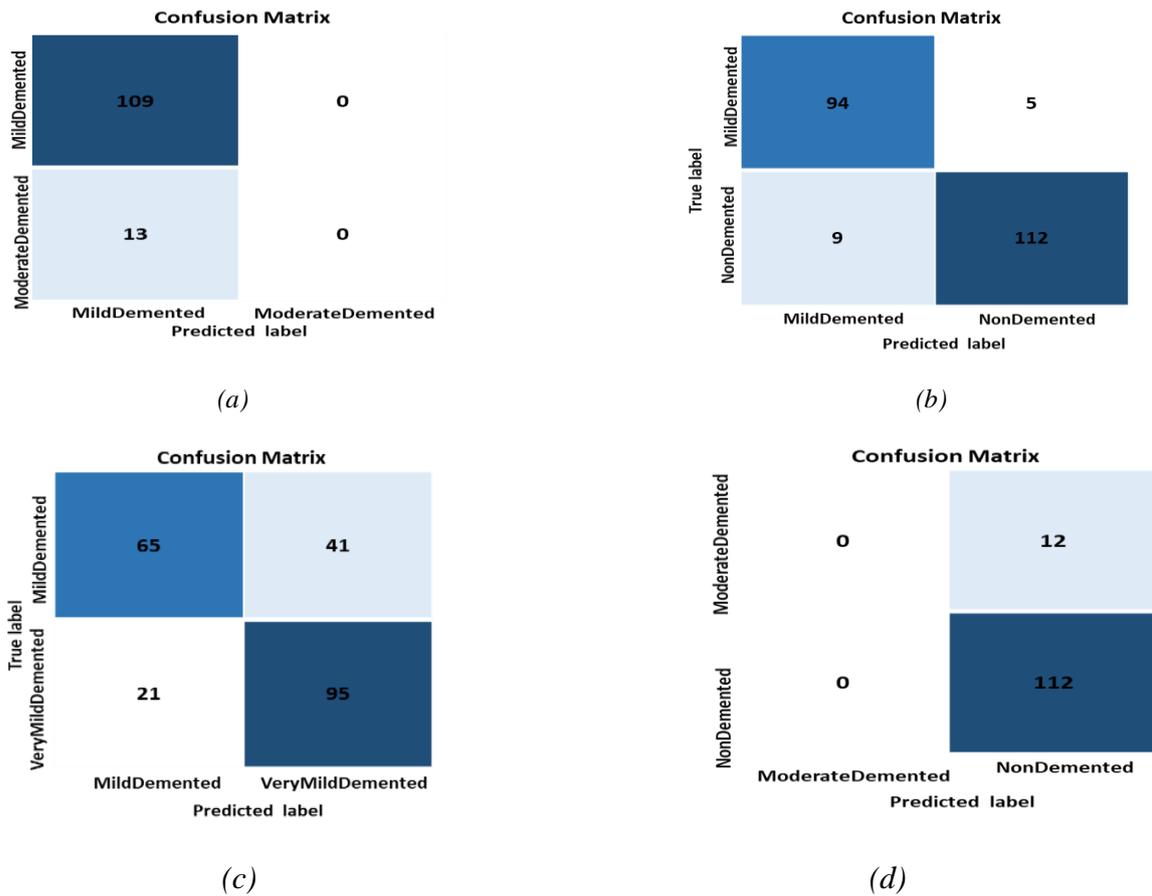


Figure 4.13 Confusion Matrices for Binary Classification with VGG16+ODF on Alzheimer disease MRI dataset: (a): *Exp1*, (b): *Exp2*, (c): *Exp3*, (d): *Exp4*

4.6.2 Multiclass Classification

In the case of Multiclass Classification, the hyperparameters of the model are chosen from the best results obtained with the binary classification of different experiences of each dataset. The protocol used for each dataset is shown in table 4-10.

Table 4-10 Protocol for Multiclass Classification

<i>Datasets</i>	<i>Total Images</i>	<i>Training Set</i>	<i>Testing Set</i>
<i>Lung Cancer</i>	1223	987	245
<i>Breast Cancer</i>	1106	884	222
<i>Alzheimer Disease</i>	1726	1380	346

4.6.2.1 Lung Cancer

Table 4-11 presents the different results obtained with Multiclass Classification in this case.

Table 4-11 Results with different Lung Cancer types for 4-classes Classification

Classes	Acc (%)	Loss	Precision (%)	AUC(%)	Val_Acc(%)	Val_Loss	Val_Precision(%)	Val_AUC(%)	Test_Acc (%)
4_classes	99.05	0.059	98.16	99.75	97.96	0.456	95.92	98.54	72.24

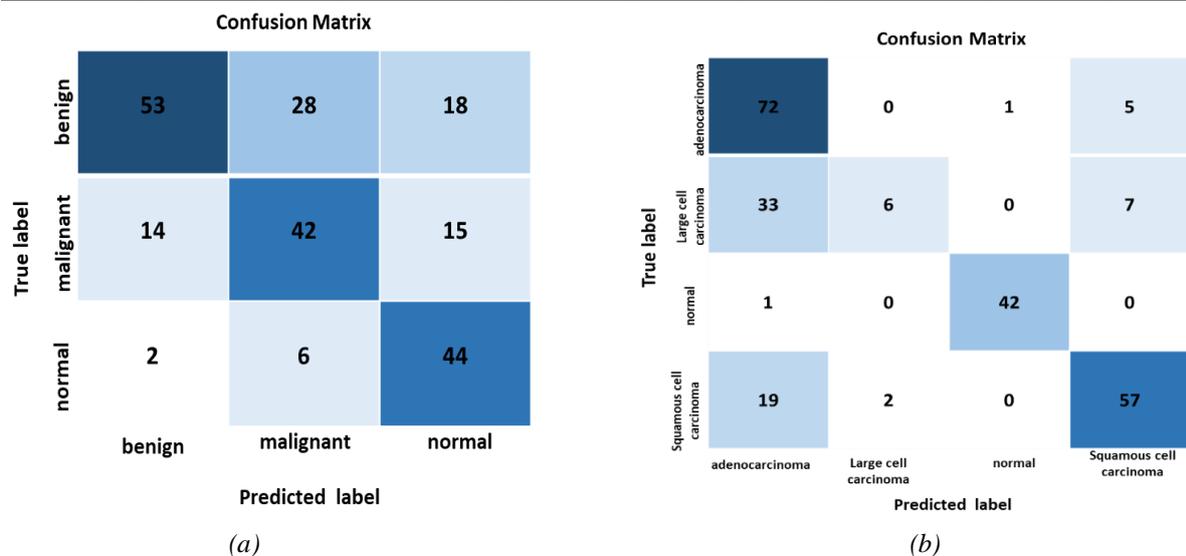
The challenges observed in accurately classifying images, particularly those depicting *Adenocarcinoma* and *Large Cell Carcinoma*, stem from the convergence of infected areas in the images. When the infected area of *Adenocarcinoma* images is large and overlaps with the small infected area of *Large Cell Carcinoma* images, the system encounters difficulties and makes errors in classification (See Fig 4.14(a)). This overlap creates ambiguity in the features, making it challenging the model to distinguish between the two classes accurately. Moreover, the confusion matrix highlights a significant number of false positives, especially in cases involving *Large Cell Carcinoma* and *Squamous Cell Carcinoma*, indicating a tendency of the model to misclassify these instances.

4.6.2.2 Breast Cancer Ultrasound Images

For Breast Cancer Ultrasound images, Table 4-12 presents the different results obtained with multiclass classification.

Table 4-12 Results with different Breast Cancer types for 3-class Classification

Classes	Acc(%)	Loss	Precision(%)	AUC(%)	Val_Acc(%)	Val_Loss	Val_Precision(%)	Val_AUC(%)	Test_Acc (%)
3_Classes	96.83	0.136	95.56	99.35	92.64	0.570	89.14	95.27	62.61

**Figure 4.14** Confusion Matrices for the VGG16+ODF with:

(a) 4 classes on Lung Cancer and (b) 3 classes on Breast Cancer datasets

For Multiclass Classification of the Breast Cancer Ultrasound images dataset, the model achieved a Val_accuracy of 92.64% and Test_accuracy of 62.61% and has a large number of false positives. The ultrasound is not the most suitable detector for this model (28, 14 and 15). For this kind of acquisition denoising is necessary.

4.6.2.3 Alzheimer Disease

For Alzheimer's Disease, Table 4-13 and Figure 4-15 present the different results obtained with multiclass classification.

Table 4-13 Results obtained with different Alzheimer disease types for 4-class classification

Classes	Acc(%)	Loss	Precision(%)	AUC(%)	Val_Acc(%)	Val_Loss	Val_Precision(%)	Val_AUC(%)	Test_Acc(%)
4_Classes	99.24	0.005	98.48	99.91	96.46	0.052	93.04	99.08	74.57

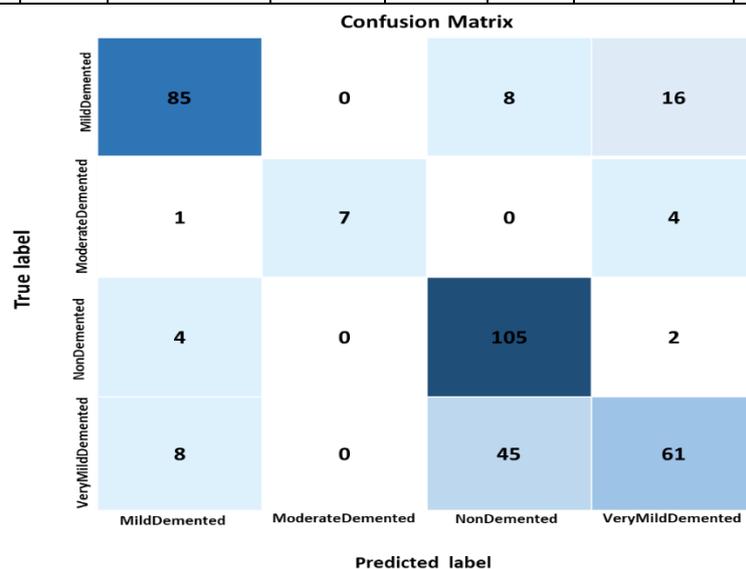


Figure 4.15 Confusion Matrix for the VGG16+ODF with 4 classes on Alzheimer Disease dataset

While our model achieves an impressive overall accuracy of 99.24%, it exhibits a significant challenge in correctly identifying cases of *VeryMildDemented* individuals. Despite the low number of false positives in other categories, the model consistently misclassifies a substantial portion of *VeryMildDemented* cases.

4.7 Comparison Results with State of the Art

Numerous studies have delved into the potential of CNN models for tasks such as feature extraction and classification. These models have revolutionized the way we interpret complex data. Building upon this foundation, our research aims to extend the boundaries of knowledge

in this area. In this work, we conduct a comparative analysis, pitting our proposed framework against state-of-the-art methods.

Table 4-14 Comparison with results and related studies.

<i>Authors</i>	<i>Methods</i>	<i>Datasets</i>	<i>N.C</i>	<i>Acc (%)</i>	<i>Precision (%)</i>	<i>AUC (%)</i>
Breast Cancer Dataset						
Mukhlif et al. [113]	Dual TL (DTL) (Xception MODEL)	ICIA2018 Breast Cancer : 100 Benign, 100 InSituCarcinoma, 100 InvasiveCarcinoma, 100 Normal	4	99	99.003	–
Chakravarthy et al. [114]	ResNet18+CSOA-wKNN	WDBC Mammogram Breast Cancer : 357 Benign, 212 Malignant	3	97.33	–	–
Baghdadi et al. [79]	TL + MRFO	Breast Cancer Histological Images : 2479 Benign, 5304 Malignant.	2	97.72	97.72	99.57
		Ultrasound Images : 437 Benign, 210 Malignant, 133 Normal	3	99.01	99.01	99.77
Atban et al.[115]	ResNet18-EO+SVM	Histological Images : 587 Benign, 587 Malignant	2	97.73	–	–
Boulenger et al. [95]	adaptive Histogram Equalization+VGG19	Ultrasound Images : 831 images	2	85	–	86
Ours	MRFO+VGG16+ODF	Ultrasound Images : 437 Benign, 210 Malignant, 133 Normal.	Benign/ Malignant	97.39	97.39	99.74
			Benign/ Normal	98.73	98.73	99.90
			Malignant/ Normal	100	100	100
		Histological Images : 495 Benign/ 495 Malignant	Benign/ Malignant	98.99	98.99	98.99
		Ultrasound images	3	96.83	95.56	99.35
Lung Cancer Dataset						
Xu, Yeguo et al. [116]	MSBO+Alexnet	CT scan 4682 images (Healthy and Cancerous)	2	95.96	–	–
Elnakib et al. [117]	GA+VGG19+SVM	160 Normal 160 Cancerous	2	96.25	–	–
Civit-Masot et al. [82]	Colour CNN Greyscale CNN	5000 Healthy	2	99.7	99.7	–
		5000 Adenocarcinoma 5000 Squamous-cell carcinoma	3	97.11	96.1	–
Lakshmana prabu et al. [80]	ODNN	CT scan images	3	94.56	–	–
Salama et al. [118]	ResNet50	CXR Images	2	98.91	97.72	98.85
Ours	MRFO+VGG16+ODF	CT scan 159 Normal, 326 Adenocarcinoma, 163 Large_Cell_Carcinoma, 252 Squamous_Cell_Carcinoma'	Adenocarcinoma/ Normal	100	100	100
			Adenocarcinoma/ Large_cell_carcinoma	99.24	99.24	99.99
			Adenocarcinoma/ Squamous_cell_carcinoma	98.58	98.85	99.66

			Large_cell_carcinoma / Squamous_cell_carcinoma	98.67	98.67	98.80
			4	99.05	98.16	99.75
Alzheimer Disease Dataset						
Ismail et al. [102]	MOGO + MultiAz-Net + softmax	DICOM : 1617 images MRI and PET scans : 511 AD, 571 Early MCI, 535 NC.	AD/NC	87.7 94.4 92.8	–	–
	MultiAz-Net + SVM		MCI/NC	91.5 93.2 89	–	–
	MultiAz-Net + RF		AD/MCI	89.4 90.0 83.0	–	–
	MultiAz-Net : InceptionV3, AlexNet, and ResNet-18		AD/ NC/ MCI	90.1 92 89.0.3	–	–
Baghdadi et al. [79]	A ³ C-TL-GTO	MRI dataset : 6400 images	4	96.25	–	–
Babu et al. [98]	DCNN +CG-DU	MRI (Not clear)	AD/ NC/ MCI	98.79	98.70	–
Mahendran et al. [119]	mRmR-WPSO-AEand IDBN	161 samples (AD and normal) and 54675 features	AD/Normal	96.78	–	–
Ours	MRFO+VGG16+ODF	MRI dataset : 6400 images	MildDemented/ ModerateDemented	98.76	98.76	99.93
			MildDemented/ NonDemented	98.97	98.97	99.97
			MildDemented/ VeryMildDemented	95.82	95.82	99.52
			ModerateDemented/ NonDemented	98.99	98.99	99.98
			4	99.24	98.48	99.91

Table 4-14 compares the outcomes of the proposed MRFO+VGG16+ODF framework with those from various state-of-the-art studies in the realms of Breast Cancer, Lung Cancer and Alzheimer Disease classification. Our work endeavors to redefine the landscape of medical image analysis. By not only outperforming existing methods but also highlighting a level of adaptability previously unseen, MRFO+VGG16+ODF framework represents a significant leap forward. It not only demonstrates the potency of advanced ML techniques but also highlights the importance of flexibility in handling diverse and complex medical datasets.

4.8 Conclusion

In conclusion, this work underscores the pivotal role of AI in revolutionizing healthcare, especially in the early detection of various cancers through medical image classification. The primary objective has been to enhance the accuracy of disease classification models across a spectrum of illnesses, encompassing Alzheimer Disease, Lung Cancer and Breast Cancer. Employing the MRFO algorithm for optimal feature subset selection, our study aims to

significantly improve the overall precision of disease classification models. This, in turn, contributes to the advancement of more effective diagnostic and treatment strategies. In this study, we introduce a novel approach to disease classification utilizing a modified TL model. Specifically, the VGG16 model serves as the foundation for feature extraction, with global average pooling enhancing the extraction of global features. The proposed model demonstrates notable accuracy rates of 99.24%, 100%, and 99.25% for Alzheimer Disease, Lung Cancer, and Breast Cancer respectively. These results signify the potential of the proposed approach in advancing the field of medical image classification for more accurate and timely disease diagnosis.

Despite the promising outcomes achieved in this research, it is essential to acknowledge certain limitations. Firstly, the performance of the proposed model may be influenced by the quality and quantity of available data. Additionally, the generalizability of the model to diverse diseases should be further investigated. Future research could explore the scalability and robustness of the proposed technique across larger and more varied datasets.

5 Chapter: AI Techniques for Medical Image Segmentation and Optimization

5.1 Introduction

Medical image segmentation is a fundamental task in image analysis and computer vision, particularly in the context of medical imaging. It involves dividing an image into distinct and meaningful regions, each corresponding to a specific anatomical structure, tissue type, or object of interest within the image. The primary goal of segmentation is to facilitate accurate quantitative analysis, visualization, and interpretation of the structures present in the image. There are various techniques used for medical image segmentation, including Thresholding, Edge-based Methods, Active Contours, and DL.

Amidst the backdrop of advancing medical imaging technology, the segmentation of diseases within medical images has emerged as a critical frontier in diagnosis and treatment. As healthcare systems confront the challenges and the increasing incidence of various diseases, the need for precise and efficient disease segmentation methodologies has become paramount. DL and other advanced computational techniques have revolutionized disease segmentation in medical imaging. These technologies offer unprecedented capabilities to delineate and analyze pathological features within images, enabling clinicians to identify and quantify disease manifestations with greater accuracy and speed.

5.2 The Evolution of Medical Image Segmentation

Medical image segmentation refers to the process of partitioning medical images into different regions of interest, which can be used for analysis, diagnosis, and treatment planning. This involves identifying and separating different structures or tissues within an image, such as organs, tumors, blood vessels, or bones. Medical image segmentation is important in healthcare because it can provide clinicians and researchers with valuable information about the structure and function of organs and tissues, as well as help guide treatment decisions. For example, in cancer treatment, accurate segmentation of tumors can help determine the extent of the disease, plan radiation therapy, and assess treatment response.

5.3 Overview of UNet and its variants architectures

Medical image segmentation plays a pivotal role in the field of medicine and successively a vital component in the computer-aided diagnosis paradigm. UNet is the most widespread image segmentation architecture due to its flexibility, optimized modular design, and success in all medical image modalities [120]. Several extensions of UNet have been proposed to address the scale and complexity created by medical tasks.

The original UNet is a convolutional neural network architecture designed for image segmentation tasks, particularly in the field of medical image analysis. The UNet architecture is known for its effectiveness in producing accurate pixel-wise segmentation masks from input images [121]. The UNet architecture is an advanced iteration of the fully convolutional network (FCN). It enhances the network's capabilities by incorporating larger feature channels during the upsampling phase, facilitating the seamless transmission of contextual information to higher-resolution layers. Additionally, UNet adopts a stack of convolutional layers instead of dense layers for improved performance[122]. The UNet network is designed to learn input images in an end-to-end, pixel-to-pixel manner. It achieves this by amalgamating three distinct functional components: the encoder, the decoder, and the integration of short and long connections, all of which collectively form a U-shaped structure. Skipping connections facilitate the integration of the encoder's output with the decoder's input at multiple resolutions. This mechanism aids in recovering spatial information that might be lost during the down sampling process.

5.4 OCAE and OUNET: Standard Automatic Optimization for Segmentation Medical Image

Medical imaging technology plays a vital role in modern healthcare by generating detailed images of various human body organs. Through modalities such as X-Rays (XR), Computed Tomography (CT) scans, MRI, Ultra Sound, and Positron Emission Tomography (PET) scans, medical professionals gain access to detailed images of various organs [123]. For instance, the lung area can be affected by infections caused by bacteria and viruses, resulting in conditions such as pneumonia. One notable example is COVID19, which can cause severe pneumonia affecting both lungs. The inflammation triggered by lung infections can significantly impair a person's breathing, making it difficult for them to breathe properly. Additionally, infections can impact other organs, such as retinal vessels or even lead to

conditions like skin cancer. Skin cancer, particularly melanoma poses a significant public health challenge. Melanoma is recognized as the most fatal type of skin cancer and accounts for a majority of skin cancer-related deaths. Between the years 1990 and 2019, there were approximately 289950 reported cases of melanoma globally, with 62840 deaths attributed to the disease [124]. These images are pivotal in the meticulous process of disease identification, where doctors analyze them with utmost care to make accurate diagnoses. However, these tasks are challenging and time-consuming. In response, techniques like segmentation and quantitative analysis have become instrumental, enabling precise delineation of structures within images and providing in-depth understanding of a patient's condition. Moreover, the integration of advanced technologies, particularly DL, has revolutionized the field. DL has achieved unremarkable advancements in image analysis, classification, and segmentation, enabling early detection of anomalies and timely prevention of pathologies.

Segmentation is essential in the field of medical image processing, as it involves the division of a digital image into distinct regions. This technique is commonly used to detect and separate various objects within an image. In the context of medical images, accurate segmentation is crucial for identifying boundaries of anatomical structures and tissues. These boundaries are essential for automated analysis and diagnosis of medical images, enabling medical professionals to extract valuable information and make informed decisions regarding patient care. [12] conducted a comprehensive thematic survey on medical image segmentation using DL techniques, with a focus on supervised learning. Their analysis delved into backbone networks, network block designs, and loss function enhancements. Additionally, the study explored weakly supervised learning methods, including data augmentation, Transfer Learning (TL), and interactive segmentation approaches. This research enhances our understanding of the rationale and potential improvements in medical image segmentation using DL. In [125] authors explored the increasing importance of DL in medical image segmentation, covering fundamental concepts applicable to various medical imaging tasks such as categorization, recognition, segmentation, and registration. The paper introduced basic DL ideas, applications, and frameworks, emphasizing the necessity for future research to address these challenges and advance the field of medical image segmentation. The reference [126] presented a review that discusses the significance of medical image segmentation in the field of biomedical image processing, emphasizing its growing importance in sustainable healthcare and its integration into Computer Vision (CV) research. The focus of the work lies in exploring medical image segmentation using DL techniques, particularly deep convolutional neural networks. The work

introduced the basic concepts and features of medical image segmentation based on DL, outlining the current research status and its limitations.

Auto Encoders constructed using convolutional layers have proven to be successful in various medical image analysis tasks, such as denoising, and are well suited for efficient segmentation of medical images. Compared to other types of AE, CAEs are particularly advantageous for medical image preprocessing due to their ability to fully leverage the power of CNNs and exploit the inherent structure of the images [127].

The UNET model is widely recognized as one of the most effective DL networks for medical image segmentation. It employs an encoder-decoder architecture that facilitates the extraction of relevant features from the input image and enables the reconstruction of the segmented image to its original size. The encoder-decoder structure is a key component of the UNET model, has demonstrated impressive performance and is often used as baseline architecture for various medical image segmentation networks [128]. In addition to the encoder-decoder structure, the UNET model incorporates skip connections. These connections play a crucial role in joining low-level features captured by the encoder block with high-level features obtained from the decoder block. This allows for the seamless integration of detailed information from different levels of the network, enabling accurate and precise segmentation of medical images. The skip connections contribute to preserving important spatial information and improving the overall performance of the UNET model in medical image segmentation tasks.

Recently, there has been a growing interest in DL techniques for multi-modal medical image segmentation. This is because multi modal imaging modalities can provide complementary and diverse information about a specific target or region of interest. By incorporating multi modalities, DL models can leverage the rich and varied data to enhance the segmentation accuracy and performance [129]. Due to the challenges faced by medical image segmentation methods based on DL, several researchers continue to explore and improve these techniques. DL model optimization is widely employed to enhance the challenging task of image segmentation in the medical field [130].

The objective of this work is to find a fast, efficient and above all standard model that can ensure the segmentation of several diseases on different organs. For this, we carry out two different works; the first deals with monomodality/multimodality and the second with the optimization of the proposed CAE+UNET segmentation model. The proposed technique was evaluated on

datasets with various pathologies and organs.

Different algorithms have been developed using DL to assist clinicians. The extensive utilization of normal radiography and other modalities can significantly augment the information contained in the used dataset for training DL algorithms, thereby introducing challenges in achieving accurate diagnoses for critical diseases. Our objective is to identify an efficient and rapid model that adheres to established standards, enabling its widespread application across various pathologies and organs.

Image noise removal is a crucial preprocessing step in medical image analysis. The first stage of the proposed work involves improvements to the Convolutional Auto Encoder (CAE) that can potentially serve as a solution to the problem, among various other artificial intelligence (AI) methods. Then, the optimization of the DL model is proposed, this is a promising candidate for detecting anomalies and segmenting them accurately in medical images. In this study, we introduce a standardized approach for organ segmentation utilizing an optimized fully convolutional network (OUNET) and an Optimized CAE (OCAE). To enhance the dataset and improve image denoising, the Particle Swarm Optimization (PSO) algorithm is used, preventing information loss during the denoising process conducted by OCAE.

5.4.1 Recent Work of Segmentation

AI applications rely on big datasets to assist doctors so that they can diagnose and predict accurately and rapidly the risk of diseases and prevent them in time. However, the availability of annotated data is not easily possible as compared to other imaging areas [131]. One of the most common tasks in medical imaging is semantic segmentation. Achieving this segmentation automatically has been an important research field, but the process has been proven very complex due to the large variation of anatomy between different patients [126]. The segmentation of organs or lesions from medical images plays a vital role in many clinical applications. Nevertheless, obtaining the annotated data is big challenge in medical images and generating annotations requires expertise and time. Liu et al. [132] provided overall ideas for medical image segmentation using DL techniques including UNET, VNET, attention UNET, and nnUNET. In addition, there are several techniques such as Support Vector Machines (SVM) [133], multi-task, multi-instance M2UNET [134], Global Feature network (GFNet) [135], and UNET [136] have been proposed for medical image segmentation. All of the above techniques focused on the precision and robustness of segmentation. A comparative study between two modules: target region selection and single-instance segmentation for COVID19 lesion detection is presented in [137].

Table 5-1 Recent Work of Segmentation

<i>Authors</i>	<i>Years</i>	<i>Methods</i>	<i>Databases</i>	<i>Results</i>
He et al [134]	2021	A multi-task multi-instance deep network (M² UNet)	666 chest CT images	DSC : 0.759 ± 0.055
Fan et al [138]	2020	Inf-Net	100 axial CT images from different COVID19 patients	Inf-Net :Dice = 0.682 Semi-Inf-Net : Dice = 0.725
Saeedizadehet al.[139]	2021	TV-Unet	around 900 images	mIoU rate of over 99%, Dice score = 86%.
Amyar et al. [140]	2020	multitask deep learning model (classification, segmentation, reconstruction)	347 COVID19, 397 non-COVID 100 COVID19 CT scan with ground truths 425 CT scans of normal, 98 of lung cancer.	Dice-coef = 80%
Chen et al.[141]	2020	modifiedU-Net	100COVID19 with masks	DSC = 0.94
Yan et al.[142]	2020	Deep CNN	21,658 annotated chest CT images	Dice similarity coefficients = 0.987 for lung and = 0.726 for COVID19.
Elharrouss et al.[143]	2020	encoder-decoder network	100 COVID-19	Dice = 0.786
Müller et al.[144]	2020	3D U-Net	20 annotated COVID19 chest CT volumes	Dice similarity coefficients of 0.956 for lungs and 0.761 for infection
Abdel-Basset et al.[145]	2021	FSS-2019-nCov	110 axial CT slices, nine CT volumes, 20 CT volumes	DSC = 0.798
Shan et al.[146]	2021	VB-Net	549 CT images COVID19 infection regions	DSC: $91.6\% \pm 10.0\%$
Chaganti et al.[147]	2020	Dense UNet	9749 chest CT volumes	Abnormality Segmentation PO (P < .001) = 0.92, PHO (P < .001) = 0.97 LSS (P < .001) = 0.91, LHOS (P < .001) = 0.9
Negi, A et al.[148]	2020	RDA-UNET-WGAN	1062 images obtained from the Breast Ultrasound Lesions Dataset	DSC = 0.8841

More recently, the introduction of DL techniques and the development of segmentation algorithms have helped researchers to optimize classification and segmentation problems in the medical field. Indeed, various methods have been proposed for the segmentation of medical images across different domains. Skin cancer segmentation, brain segmentation, retinal vessel segmentation, and lung segmentation are among the many areas of focus in medical image segmentation. Recently, there has been significant attention directed towards the detection and

segmentation of COVID19 infections at the lung level using DL methods. With the emergence of the COVID19 pandemic, there has been an urgent need to develop automated and accurate tools for diagnosing and monitoring the disease. For example, He et al. [134] presented a multi-task, multi-instance deep network (M²UNET) to appreciate the gravity of COVID19 infections and lung lobe segmentation. Fan et al. [138] proposed novel COVID19 lung Infection segmentation Deep Network called Inf-Net to automatically identify infected regions from chest CT slices. Amyar et al. [140] presented an automatic classification and segmentation tools for detecting COVID19 pneumonia using chest CT imaging. The model was designed for three learning tasks: segmentation, classification, and reconstruction. The model obtained a Dice Similarity Coefficient (DSC) of 0.88 and an area under the ROC curve of 97%. Saha et al. [149] proposed a novel DL algorithm by using Deep Neural Network (DNN) composed of an attention-based dense UNET with deep supervision for COVID19 lung lesion segmentation from chest CT images. The model achieved a DSC of 0.86 with the dataset containing 100 CT images and 100 masks. Chaganti et al. [136] have used a Dense UNET method for segmentation and quantification of abnormal CT_COVID19. They validated their work on a dataset with 9749 CT volumes for lung segmentation, lobe segmentation, and COVID19 lesions.

For skin cancer segmentation, Mirikharaji et al. [150] presented a comprehensive review of 177 research papers related to DL based segmentation. The analysis covered multiple aspects, including input data like datasets, preprocessing methods, and synthetic data generation. Additionally, they delved into model design, including architecture, modules, and losses, as well as evaluation criteria such as data annotation requirements and segmentation performance metrics. Adegun et al. [151] presented a comprehensive survey of techniques that have been used for detecting skin cancer from skin lesion images. Kumar et al. [152] presented an optimization based model to discover skin cancer. They proposed a U-RP-Net model, which is obtained by integrating UNET and RP-Net for segmentation, and Aquila Whale Optimization (AWO) based SqueezeNet for skin cancer detection. Dash et al. [153] proposed an automated psoriasis lesion segmentation method based on a modified UNET architecture. They achieved a DSC of 0.9303 and an accuracy of 94.8% for retinal vessel segmentation. Badar et al. [154] presented a review of DL retinal image application analysis for automated classification of retinal landmarks, pathology, and disease classification. Zhang et al. [155] utilized a UNET with residual to detect vessels. They obtained a high performance with AUC of 97.99% on DRIVE dataset. Hu Kai et al. [156] have proposed a novel retinal vessel segmentation method of the eye fundus images based on CNN and fully connected conditional random fields (CRFs).

Two public datasets were used DRIVE and STARE. They achieved an accuracy of 95.33% and 96.32% for the DRIVE and STARE datasets respectively. Iqbal et al. [157] proposed a new Generative Adversarial Network for Medical Imaging (MI-GAN). The MI-GAN generates synthetic medical images and their segmented masks. The proposed model achieved a DSC of 0.837 on STARE dataset and 0.832 on DRIVE dataset.

In the recent studies [134-136, 138] examined, there is a noticeable gap in the optimization techniques employed, particularly concerning the widely used UNET architecture. Many recent works have not prioritized optimizing UNET, a foundational model in medical image segmentation. It is imperative for future research endeavors to address this gap by focusing on the integration of metaheuristic algorithms. By optimizing models like UNET using these advanced techniques, researchers can unlock more robust and refined medical image segmentation solutions. This enhancement is essential for advancing the field and, ultimately, improving patient care.

The innovation in this proposed work lies in the utilization of a metaheuristic approach for optimizing both CAE and UNET. Among the array of methods available, PSO stands out due to its superior calculation accuracy and its ability to perform global and local search simultaneously. Our focus is on strategically integrating the PSO algorithm. Firstly, it is used to select the optimal number of filters for the CAE model, enhancing image denoising for preprocessing. Then, in a second step, we optimize UNET for improved segmentation. This enhances the performance and precision of the UNET model

5.4.2 Contribution

The proposed approach is based on an optimized model using UNET architecture, applied to COVID-19, skin cancer and retinal vessel segmentation. The framework contains three stages: the first stage is the preprocessing images, which contains the conversion of *nifty* to *png* file, the data augmentation and the resizing. Then, a proposed denoising of the images using CAE before segmentation is applied in the second stage. PSO algorithm is used in the third stage to optimize the parameters of CAE for properly denoising images and saving valuable information, after that UNET is used in the segmentation task, which ensures the segmentation of the infection region. To evaluate this model and test its robustness and efficiency, lung datasets, skin cancer dataset, synthetic retinal dataset and retinal vessel segmentation (STARE) are used. A pipeline is proposed for generating synthetic medical images using Generative Adversarial Networks (GANs). The GAN trained with retinal vessel segmentations from the

DRIVE dataset. Several tasks to be explored to find the lock to a high-performance, efficient and standard segmentation:

- **Task 1:** Study of monomodality (slice Ox, slice Oy, slice Oz), and choose which of the terms should be used;
- **Task 2:** Multimodality and the concatenation of slices to know which of the two contributes the most to the deep segmentation;
- **Task 3:** Study the effect of the denoising by OCAE on the segmentation model;
- **Task 4:** Selection of the optimal parameters OUNET model and OCAE filters number Optimization.

The contribution of this work focuses on implementing hybrid methodology, where the PSO algorithm is applied to find the optimal design of parameters for CAE architectures. The main contributions of this work are as follow :

- **Datasets Collection:** Collect various datasets for the comprehensive evaluation of the proposed segmentation models.
- **CAE Parameter Optimization:** PSO algorithm is used to fine-tune CAE parameters, ensuring optimal denoising performance.
- **Architectural Approaches:** UNET, CAE+UNET, and OCAE+OUNET architectures are explored for the segmentation of COVID19 and other pathologies.
- **Monomodality and Multimodality Analysis:** study of the impact of each modality and multimodal concatenation to know their relevance in segmentation tasks.
- **Robustness Evaluation:** the methodology is rigorously evaluated on datasets representing four distinct pathologies, demonstrating its robustness and applicability in various medical contexts.

5.4.3 Proposed OCAE+OUNET Approach

Early detection of diseases has become necessary, for the prevention of certain diseases and to avoid their growth. For this reason, and with the advancements in computational capacity and the availability of massive datasets, researchers have developed various DL based algorithms that outperform the specialists in that field, such as medical image disease detection, classification, and segmentation.

The high diversity in infection features and poor intensity contrast between infections and normal tissues make segmenting infected regions from CT slices difficult [133]. Different DL methods can give good segmentations, but they are likely to make different mistakes. However, in this study, we tried to develop novel optimizer DL-derived segmentation to extract features

of the ROI to segment the infected area.

5.4.3.1 OCAE+OUNET Model description

In this work, the proposed approach and optimization with PSO algorithm is described (see Fig. 5.1). The main idea and mathematical background are explained. Furthermore, proposed approach implementation and its optimization with the PSO algorithm are detailed as well as the presentation of the method in the context of the CAE optimization problem. We have also described the improvements that were needed to adapt the basic PSO algorithm for the optimization number filters of CAE model.

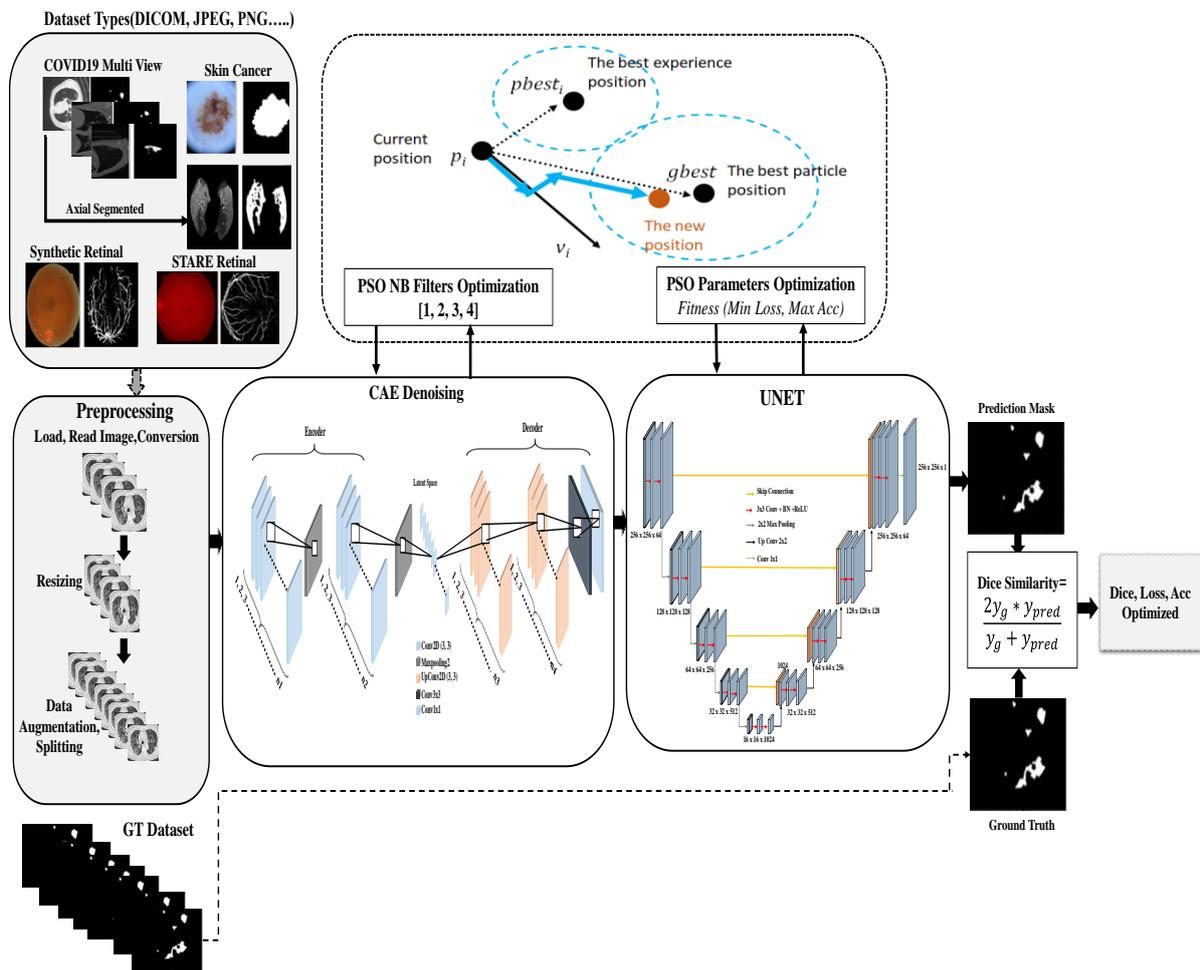


Figure 5.1 Flowchart of the proposed approach OCAE+OUNET Segmentation

- **Convolutional Auto Encoder**

Medical images usually contain noise and show blurred boundaries. The presence of noise may confuse the identification and analysis of diseases, which may result in unnecessary deaths. Hence, denoising of medical images is a mandatory and essential preprocessing technique. For this reason, CAE is used before UNET to reduce the noise in medical images. In another hand,

to obtain an optimal CAE, a method of optimization called PSO is used to optimize the number of filters for each layer in CAE.

The quality of the dataset and an effective labeling strategy are also essential. Denoising the datasets using CAE [157] plays an important role in increasing the accuracy and facilitating diagnostic models. Denoising CAE does two things:

- Encode the input with preserving the information about the data. The encoder takes high dimensional input data and maps to a low dimensional space (latent-space) using Convolution Layer with Rectified Linear Unit (ReLU) activation function and Max Pooling Layer, where, the ReLU activation function is :

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \quad (5.1)$$

Suppose, if convolution is performed on an image with more than one kernel, then dimensions of output also increases by ' n ' dimensions.

- Decoder takes data in the latent-space and try to reconstruct this data in the original high dimensional space

Denoising CAE forces the hidden layer to extract more robust features and restrict it from merely learning the identity.

- **UNET**

For medical image segmentation, the UNET architecture [121] has been found to be highly effective, where the UNET efficiently concatenates low-level and high-level image features through skip connections, which is a typical solution for medical image segmentation tasks.

The first part is the Encoder, which uses the convolutional network. It consists of the refined application of convolutions, ReLU, Dropout, and Maxpooling. After, the spatial dimensions of the feature map are transformed to the size of 2x2, and then the second part starts (Decoder) to increase the spatial dimensions by reducing the Encoder feature map. At every downsampling stride, the decoder consists of: i) an upsampling of the feature map pursued by a 2x2 convolution, which halves the number of feature channels, ii) a fusion with the coinciding cropped feature map from the Encoder, and iii) two 3x3 convolutions, each followed by a ReLU. At the final 1x1 convolution layer is used for the coveted number of classes.

- **Particle Swarm Optimization (PSO)**

Particle Swarm Optimization (PSO) is a search-based optimization technique that starts with a population of particles in a d-dimensional vector space [158]. It is proposed to imitate birds searching for food, the movement of fish's shoals, etc. It can be classified as a swarm intelligence algorithm like Ant Colony Algorithm, Artificial Bee Colony Algorithm, and Bacterial Foraging.

PSO algorithm is proposed in 1995 by Kennedy and Eberhart [159] where the article "Particle Swarm Optimization" became very popular due to his continue optimization process allowing variations to multi-targets. Consisting in the constant search for the best solution, the method moves the particles with a certain velocity calculated in every iteration. Each particle's movement has the influence of this own the best-recognized position and the best-recognized position in the space search. The result expected is that the particle swarm converges to the best solution [160].

The velocity is updated by using the equation (5.4) and the position is calculated using the equation (5.2):

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (5.2)$$

$x_i(t)$: position of particle i in a time t , $v_i(t)$: velocity can modify the position of the particle, and all particles are randomly grouped within pre-defined ranges x_{min} and x_{max} , where :

$$x_i(0) = x_{min} + rand(x_{max} - x_{min}) \quad (5.3)$$

$$v_i(t+1) = v_i w + c_1 r_1 [x_{pbest_i} - x_i] + c_2 r_2 [x_{gbest} - x_i(t)] \quad (5.4)$$

v : velocity, i : particle. c_1 : cognitive, c_2 : social factors, r_1 and $r_2 \in [0, 1]$, and w : inertia weight. The best position of the particle x_{pbest} is updated after each iteration if better than previous, and the best global position x_{gbest} is among all particles. This parameter affects the movement propagation given by the last velocity value. The swarm consists of N particles, so an objective function f calculates particle fitness. The personal and global best values are updated using equations (5.5) and (5.6), respectively at a time t . Thus, $i \in 1 \dots N$

$$x_{pbest_i}(t+1) = \begin{cases} x_{pbest_i}(t) & \text{if } f(x_{pbest_i}(t)) \leq f(x(t+1)) \\ x_i(t+1) & \text{if } f(x_{pbest_i}(t)) > f(x_i(t+1)) \end{cases} \quad (5.5)$$

$$x_{gbest}(t+1) = \max\{f(x_{pbest}), f(x_{gbest})\} \quad (5.6)$$

Where:

$$x_{pbest} \in x_{pbest_0}(t), x_{pbest_1}(t), \dots, \dots, \dots, x_{pbest_N}(t) \quad (5.7)$$

5.4.4 OCAE+OUNET Experimentation

To demonstrate the effectiveness of the proposed model, this section discusses the different experiments segmentation tests on several datasets. Experiments are conducted with diversified classes and modality of medical images.

5.4.4.1 Datasets

Four kinds of CT datasets are used in the experiments to test the proposed framework. The dataset of COVID19 was collected between May 2020 and December 2020 at the LI3C, Mohamed Khider University, Biskra, Algeria using different public datasets sources. The first dataset contains 100 lungs images with 100 infected masks [161]. In this dataset, a preprocessing step was performed to reduce the area to be analyzed, and remove unimportant regions. While the second dataset [162] contains 20 labeled (Nifti dataset) COVID19 CT scans (COVID19-CT-Seg_20 cases, Infection Mask, Lung_and_Infection_Mask and Lung Mask). Left lung, right lung, and infections are labeled by two radiologists and verified by an experienced radiologist. In this study, COVID19 and infected Masks 2D CT slices are extracted from the twenty nine 3D volumes of the CT imaging having no-uniform or varying dimensions. Each of these slices is annotated carefully by expert radiologists to generate the segmentation mask. CT scan is a set of images, which contains black slices at the beginning of each volume and other images that do not contain any part of the lung, for this reason, these images are removed. In experiment, 100 images and 100 masks from each slice were selected to tune the parameters of the proposed model. In the third skin cancer dataset, 100 images and 100 masks were used to evaluate the proposed model [163]. In the fourth dataset, the Synthetic Retinal dataset contains 91 images and 91 masks [164] STARE set contains 20 images, and 20 masks are used [165]. Figure 5.2 shows the different datasets with the ground truth segmentation mask.

(SR:

STARE

Retinal)

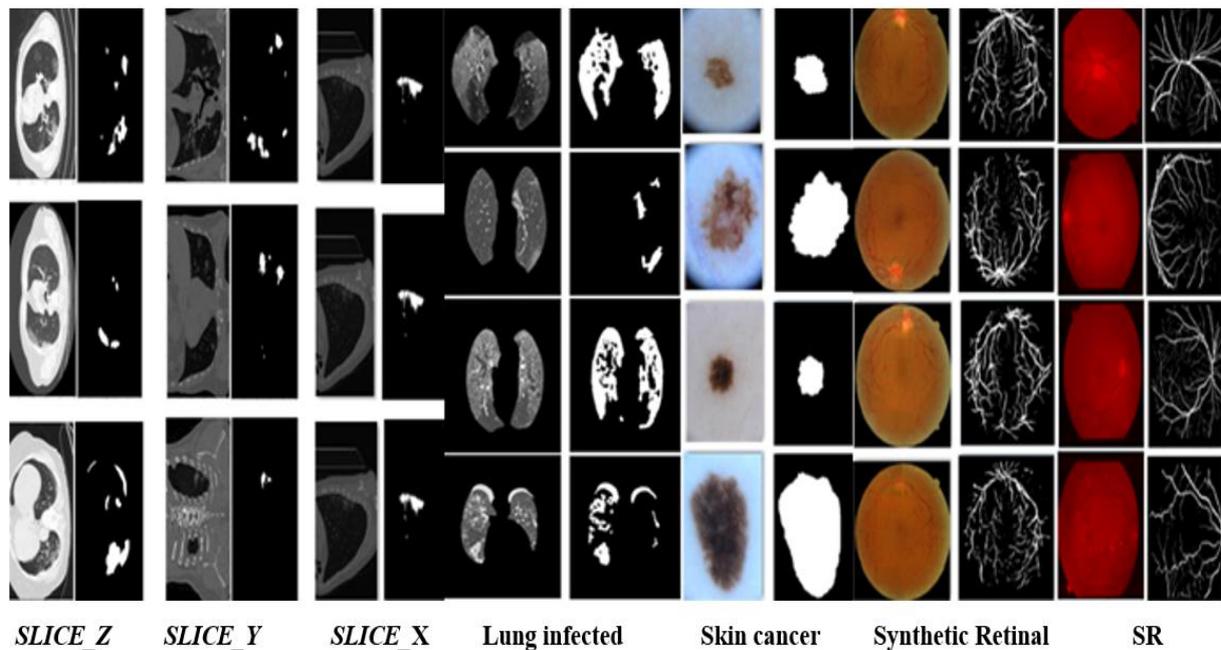


Figure 5.2 Different used Datasets

The objective is to select the most relevant number of filters that have influence to obtain good performance of AE and then implement the PSO algorithm to find these optimal numbers. Figure 4 illustrates the organizational structure of our OCAE+OUNET algorithm. The PSO algorithm, through its iterative optimization process, strategically fine-tunes the parameters of both the CAE and UNET components. This meticulous tuning ensures that the model converges to an optimal solution, maximizing its segmentation Dice, Accuracy and Loss.

5.4.4.2 Conversion, data augmentation, resizing and splitting

Image file types may be converted from one type to another. In this case, for COVID19 dataset, the *Dicom* images have been converted to *png* image file format, and for the skin cancer dataset, the *jpg* file have been converted to *png* file.

The data augmentation technique is applied in this work on the four datasets to create new training data artificially. Data augmentation techniques can be used to generate additional data, thereby producing different versions of the original data. In this work, we used data augmentation to avoid overfitting and increase the accuracy of the model. First, a model is run without any augmented images. Next, a model using augmentations is created like zooming, rotating, flipping, and cropping images. In addition, all the dataset are resized to have the same size to 256×256 and while the dataset contains 2816 images and 2816 masks, all images and masks have been resized to 128×128 . All the datasets are split into training and validating sets with a ratio of 0.9:0.1, and resized them into 256×256 size images, where the data augmentation

is applied to the training set (a total number of training images is increased by 6 times) to increase the accuracy of the model.

5.4.5 OCAE+OUNET Implementation

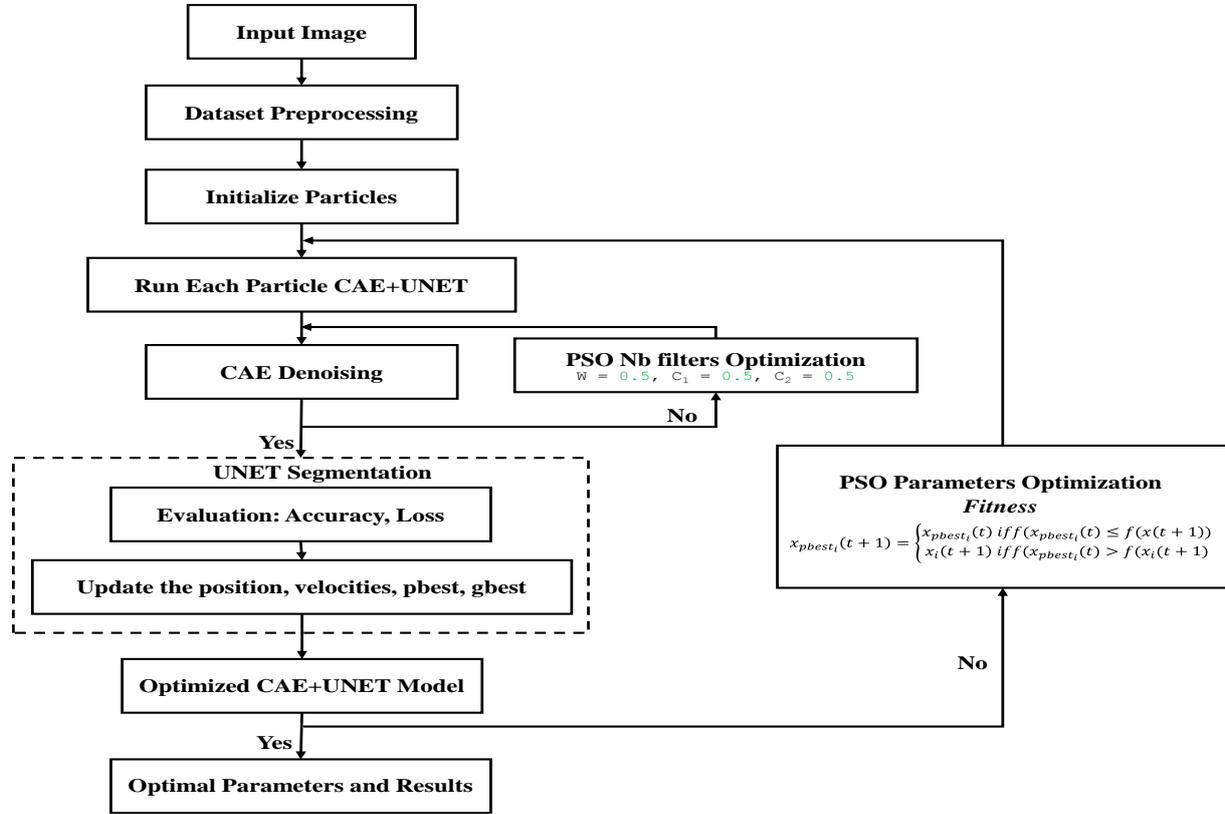


Figure 5.3 Optimization Segmentation based OCAE+OUNET Algorithm

5.4.5.1 Optimized CAE for denoising

The first step in enhancing image quality and improving contrast is preprocessing, which plays a major role in obtaining accurate results. Different classical methods for enhancing image were used to achieve high quality and contrast of images. In this work, the CAE denoising techniques are implemented, where the PSO algorithm is applied to optimize the number of filters for the different layers of encoder and decoder. Table 1 presents the detail of the particle, and the search space used and illustrates an example of a particle generated by PSO for different datasets.

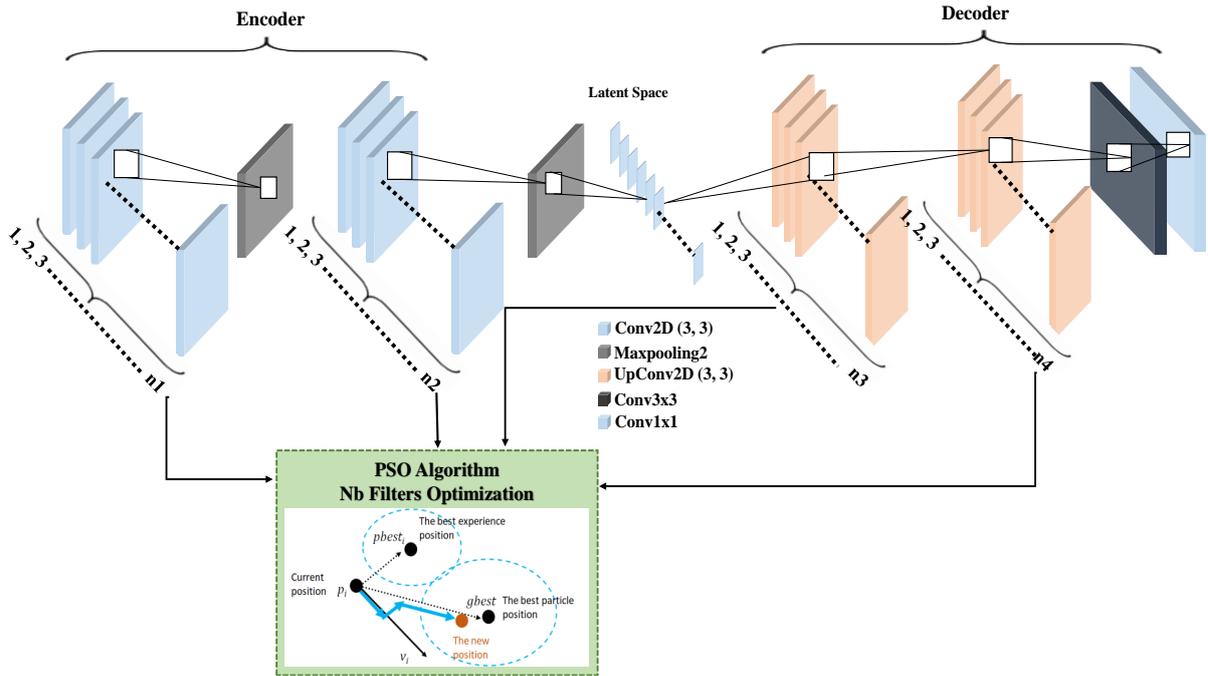


Figure 5.4 Flowchart of the OCAE algorithm ($n1$: Filters Number in Conv1, $n2$: Filters Number in Conv2; $n3$: Filters Number in UpConv1; $n4$: Filters Number in UpConv2)

Table 5-2 PSO parameters and Search Space

Parameters	Description	Search Space	Datasets					
			D1	D2			D3	D4
				SLICE Z	SLICE Z, Y	SLICE Z, Y, X		
$Particles\ i$	Particles Number	–	4	4	4	4	4	4
$Maxiter$	Max of iteration	–	10	10	10	10	10	10
C_1	Cognitive factors	–	0.5	0.5	0.5	0.5	0.5	0.5
C_2	Social factors	–	0.5	0.5	0.5	0.5	0.5	0.5
W	Inertia weight	–	0.5	0.5	0.5	0.5	0.5	0.5
E_1	Filters Number in 1 st layer Encoder	[8- 32]	28	29	25	32	18	15
E_2	Filters Number in 2 nd layer Encoder	[8- 32]	17	20	29	21	22	32
D_1	Filters Number in 1 st layer Decoder	[8- 32]	22	15	20	32	21	14
D_2	Filters Number in 2 nd layer Decoder	[8- 32]	19	25	29	19	26	23

The fitness function evaluation uses the OCAE+OUNET model on different datasets. It reports the required number of filters and returns the fitness function score. The fitness function is generated to achieve the best segmentation measure by maximizing its accuracy and minimizing Loss. It is evaluated based on the below expression:

$$f_{fitness} = \text{Max}(accuracy) \tag{5.8}$$

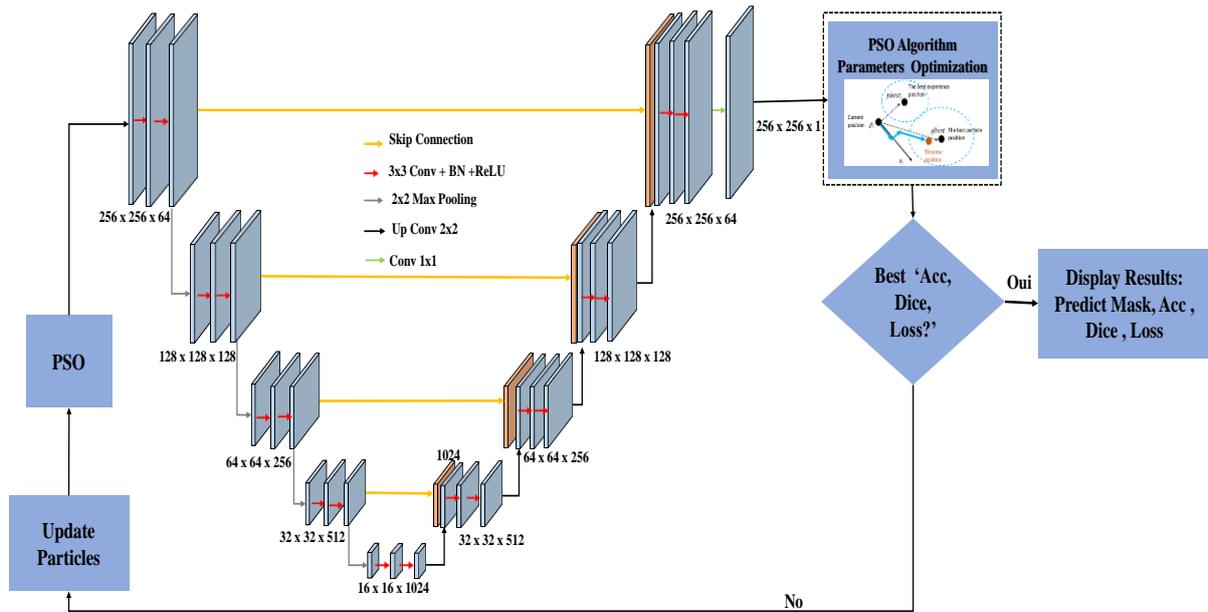


Figure 5.5 Flowchart of the OUNET

5.4.5.2 Optimized CAE+UNET for segmentation

The segmentation method is presented in this section; it begins by optimizing the AE to denoise the images without losing any information. As each image may contain details about the contour of an object as well as details about its texture that may be important in computer vision tasks.

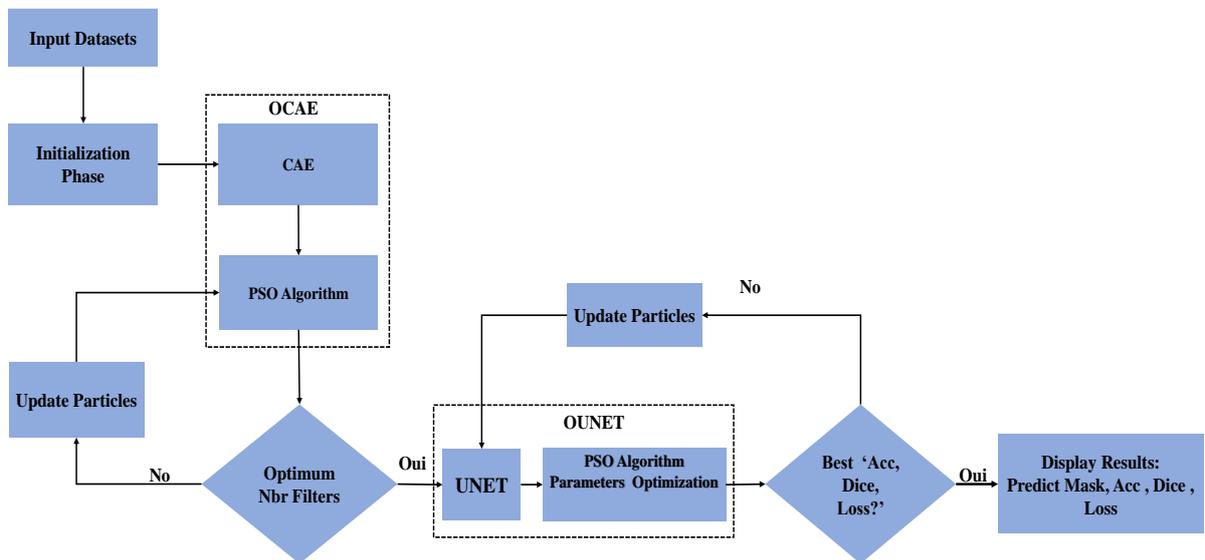


Figure 5.6 Flowchart of the OCAE+OUNET algorithm

In this case, denoising techniques are employed to achieve high-quality contrast in images, thereby enhancing the performance of the UNET model. To optimize the number of filters for the various convolution layers of both the Encoder and Decoder while maintaining the size kernel of Maxpooling layers, the PSO algorithm (see Fig. 5.6) is utilized.

In the OCAE+OUNET model parameter configuration, certain static parameters are utilized, including the Adam optimizer, Sigmoid activation function in the final layer, and ReLU as the non-linear activation function. A fixed epoch number of 100 is set for all datasets, with "early stopping" implemented. The model is iteratively implemented 10 times with a PSO iteration number of 10. In addition, the datasets are divided into training and validation sets. The training set is utilized to train the models, while the validation set is used to assess the performance of the trained models.

5.4.5.3 Evaluation metrics

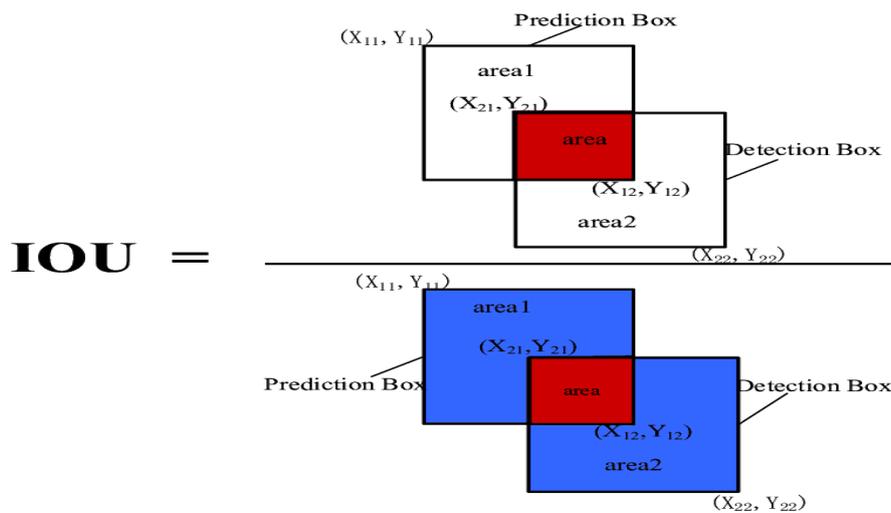
The performance of segmentation models is evaluated using a several metrics [139]. In this work, the following metrics are used to evaluate the model: Intersection over Union (IoU) and Dice Similarity Coefficient (DSC).

- **Dice Similarity Coefficient:** The concept is intuitive we count the similar pixels (taking intersection, present in both the images) in the both images we are comparing and multiple it by 2, and divide it by the aggregate number of pixels in both the images.

$$DICE_{Coef} = \frac{2|Y_g \cap Y_{pred}|}{|Y_g| + |Y_{pred}|}$$

Where Y_g and Y_{pred} denote the ground truth and predicted masks.

- **Intersection over Union:** also known as the Jaccard index. In this case, the intersection means is divided by the number of similar pixels by the union total number of unique pixels in both the images.



$$IoU(Y_g, Y_{pred}) = \frac{|Y_g \cap Y_{pred}|}{|Y_g \cup Y_{pred}|} = \frac{|Y_g \cap Y_{pred}|}{|Y_g| + |Y_{pred}| - |Y_g \cap Y_{pred}|}$$

Where **area1:** Y_g and **area2:** Y_{pred}

5.4.6 OCAE+OUNET Results and Discussion

5.4.6.1 Hyperparameters Selection

In this work, different combinations hyperparameters are evaluated and the best combination is selected. To simplify the tuning process, we choose the number of epochs= 50, 100, 200, and 300, and fixed it to 300 with early stopping; whenever the validation loss does not improve for 15 continuous epochs, learning rate =0.001, and the batch size = 16. Binary cross-entropy is used for loss functions and different optimizers (such as RMSprop, Adamax, Adagrad, and Adam) were used to choose the best optimizer. In this case, Adam gives high Accuracy and Dice similarity. Table 5-3 shows the performance of different optimizers where Adam optimizer gives the best performance.

Table 5-3 Performance of different optimizers

<i>Optimizer</i>	<i>DSC (%)</i>	<i>Accuracy (%)</i>
Adam	81.13	97.83
RMSprop	78.50	97.38
Adamax	78.89	97.47
Adagrad	65.40	96.52

5.4.6.2 Monomodality

In this work, four different datasets are used to demonstrate the effectiveness of our method for other diseases or other organs.

Experience 1: Effect of the data augmentation and size dataset CT D2 SLICE_ Z COVID19

Because chest CT scans are volumetric images, they can be viewed by scrolling through three different planes: the Coronal, Axial, and Sagittal. In this section, we trained a mono-view model using the proposed network to screen patients with COVID19 using CT images of the lung regions in axial view.

In this case, our question: Is data augmentation or dataset size better?

For this reason, two experiments are conducted: i) the first with a 100 images and 100 masks (256x256) with data augmentation, ii) the second with a 2816 images and 2816 masks (128x128) without data augmentation. The results obtained are shown in Table 5-4.

Table 5-4 Effect of the data augmentation and size dataset CT D2 SLICE_ Z COVID19

Model	100 images/100 masks (256x256)				2816 images/2816 masks (128x128)			
	With Augmentation (Small Size Dataset)				Without Augmentation (Big Size Dataset)			
	Acc (%)	Loss	DSC(%)	IoU(%)	Acc(%)	Loss	DSC(%)	IoU(%)
<i>UNET</i>	99.16	0.0172	88.56	79.47	98.92	0.004	84.55	74.54
<i>CAE+UNET</i>	99.20	0.1615	89.05	80.27	98.92	0.004	84.77	69.74
<i>OCAE+OUNET</i>	99.21	0.0151	89.46	80.93	98.78	0.005	83.37	71.48

The information given by the data augmentation with small dataset is better than the big dataset size as shown in Table 5-4 except for the loss curve and image predict.

The model trained on a small dataset of 100 images and 100 masks with data augmentation demonstrated superior performance across key metrics when compared to models trained on larger datasets without augmentation. Specifically, the augmented dataset model achieved an accuracy of 99.21% and a Dice Similarity Coefficient (DSC) of 89.46%, whereas the dataset without augmentation attained an accuracy of 98.78% and a DSC of 83.37%. This clearly confirms the effectiveness of the proposed model for small samples, which is often the case in reality and meets the intended objective. This significant performance gap underscores the impact of data augmentation techniques in enhancing the accuracy and precision of the proposed model. By augmenting the limited dataset, we effectively expanded its diversity, enabling the model to recognize a broader range of patterns and variations within the images. Consequently, the augmented dataset model exhibited a higher accuracy rate, ensuring more precise identification of COVID19 cases. Moreover, the Dice Similarity Coefficient, a metric assessing the spatial overlap between the predicted and ground truth masks, demonstrated a substantial improvement with data augmentation. This emphasizes the model's enhanced ability to delineate the exact boundaries of infected areas within the lung images.

In all that follows, we will limit our study to the case 100 images/100 masks (256x256) with data augmentation.

Experience 2: Effect of the optimized preprocessing

In this case, the question asked: *Are augmentation or/and denoising effective for small dataset?*

100 images/100 masks (256x256) are chosen for this experiment. Sample 2D Dataset1 and 3D Dataset2(see fig5.7)

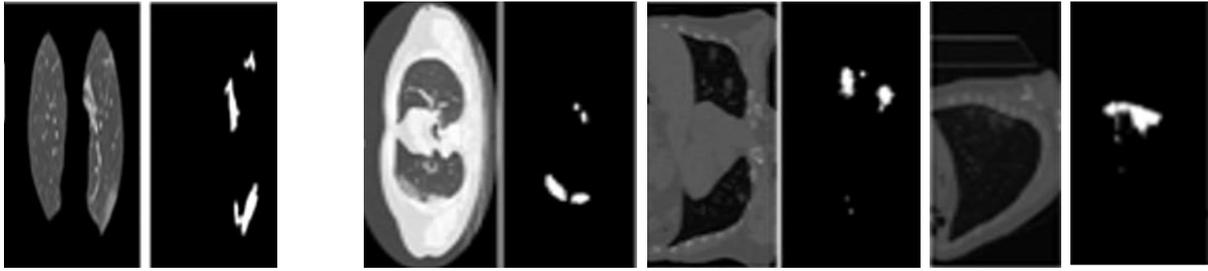


Figure 5.7 Images 2D Dataset1 and 3D Dataset2

Table 5-5 presents comparison results between preprocessed lung images and original lung images. The first experiment is applied on dataset D1. As shown table 5-5, there are very clear improvements going from the impact of PSO on CAE+UNET. We also note that the accuracy and dice score are increased.

Table 5-5 Effect of optimized preprocessing lung image on the results with small dataset

	<i>MONOMODALITY</i>					
	2D Dataset1: Optimized Preprocessing Lung Image (COVID-19)					
	SLICE_Z: Nbr Filters Optimized : [28, 17, 22, 19]					
	<i>Without Augmentation</i>			<i>With Augmentation</i>		
	<i>UNET</i>	<i>CAE+UNET</i>	<i>OCAE+OUNET</i>	<i>UNET</i>	<i>CAE+UNET</i>	<i>OCAE+OUNET</i>
<i>Acc (%)</i>	93.11	93	92.82	97.42	97.73	97.83
<i>Loss</i>	0.175	0.185	0.170	0.118	0.062	0.066
<i>DSC (%)</i>	52.81	61.96	64.62	73.21	80.41	81.13
<i>IoU (%)</i>	38.37	26.41	44.87	63.70	63.05	63.85
<i>Time(s)</i>	0.612	0.679	0.681	2.785	2.995	3.060
	3D Dataset2: Original Lung Image (COVID-19)					
	SLICE_Z: Nbr Filters Optimized : [19, 8, 25, 12]					
<i>Acc (%)</i>	97.72	98.39	98.39	99.16	99.20	99.21
<i>Loss</i>	0.033	0.012	0.013	0.017	0.016	0.016
<i>DSC (%)</i>	65.15	87.95	88.81	88.56	89.05	89.46
<i>IoU (%)</i>	48.32	78.49	79.87	79.47	80.27	80.93
<i>Time(s)</i>	1.046	0.672	0.661	2.827	3.776	3.110

IoU : Intersection Over Union

- COVID-19 Lung Image preprocessed (Augmentation, Resizing and OCAE denoising)

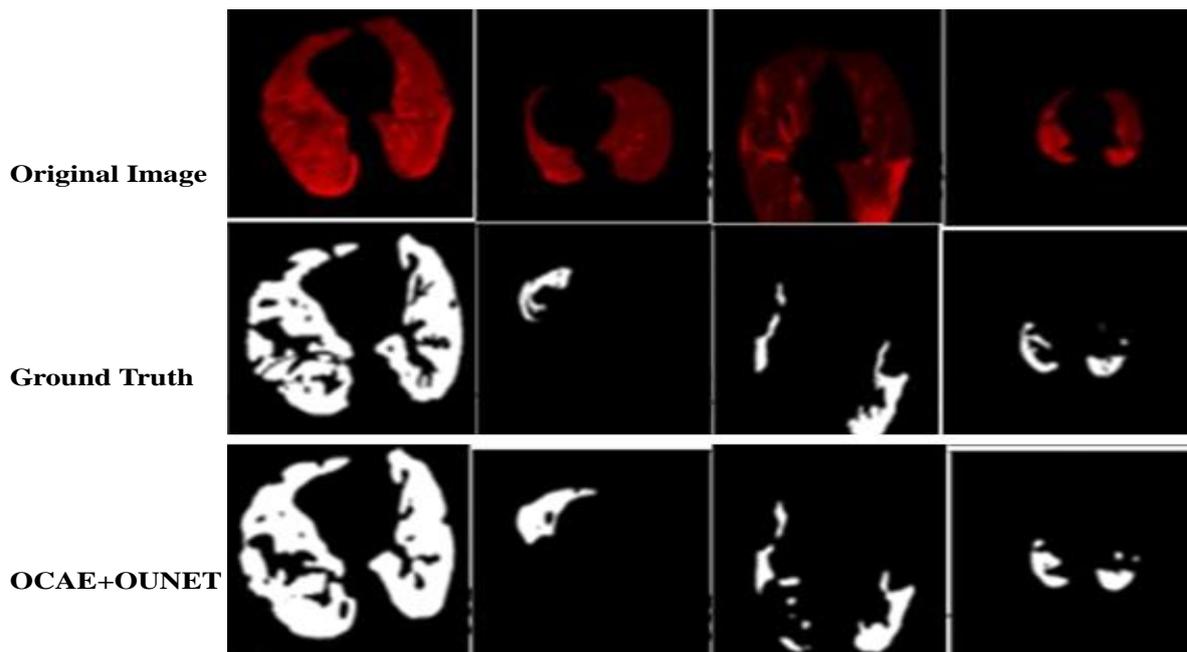


Figure 5.8 Dataset D1 100 images/100 masks with augmentation

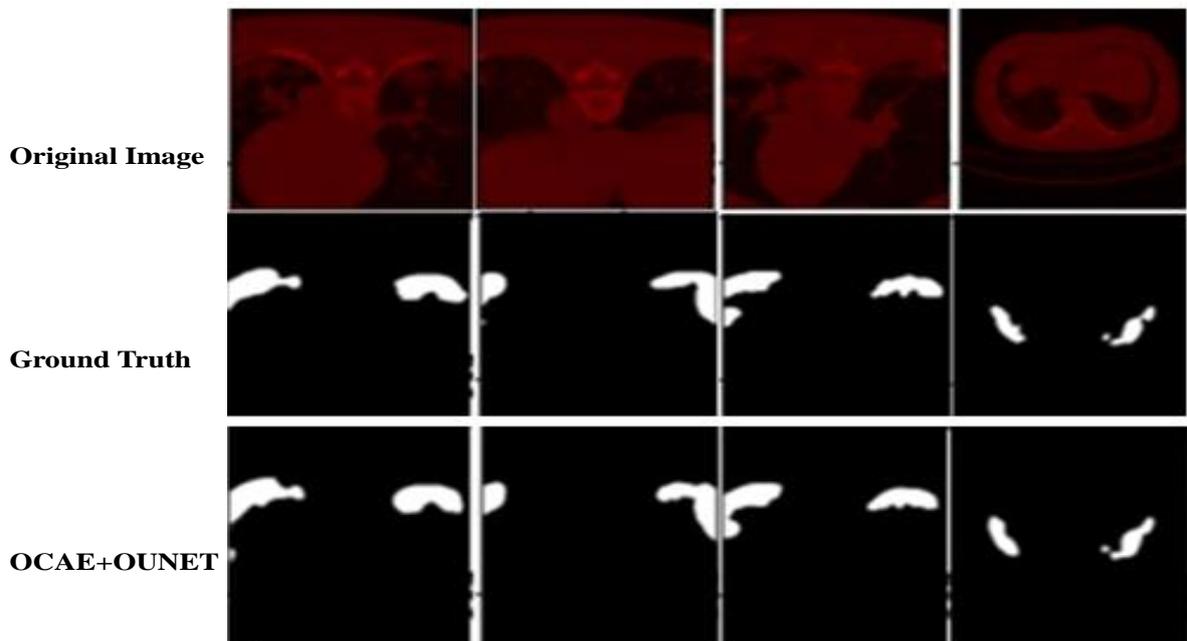


Figure 5.9 Segmentation for original 3D dataset 2 COVID19 SLICE _Z (100 images/100 masks (256x256))

For the case (100 images/100 masks (256x256)), the different models achieve a good segmentation, but the OCAE+OUNET are closer to the Ground Truth.

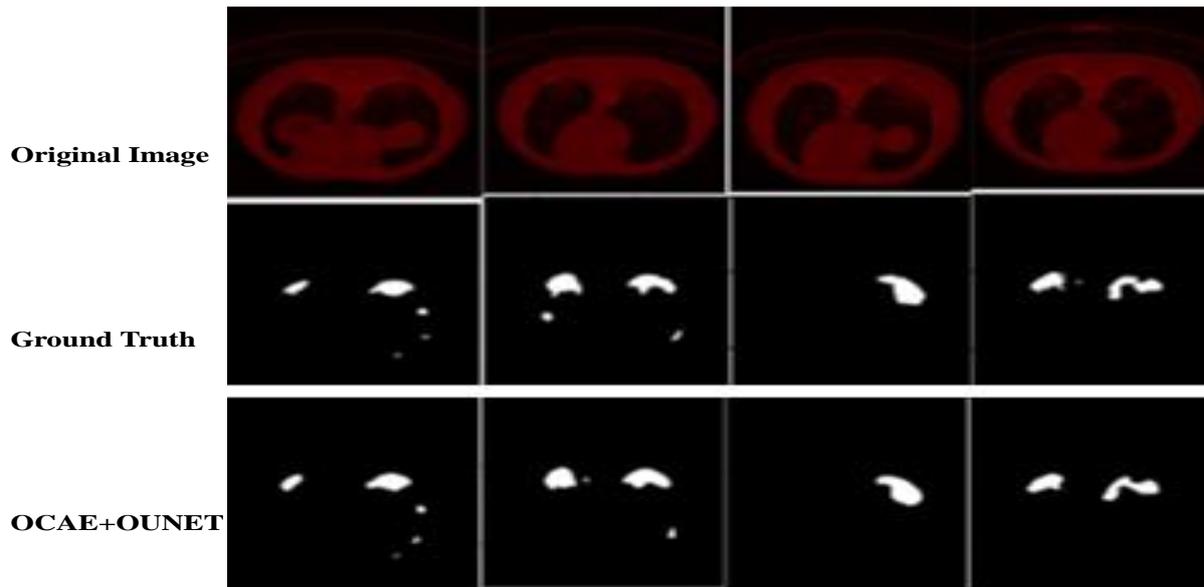
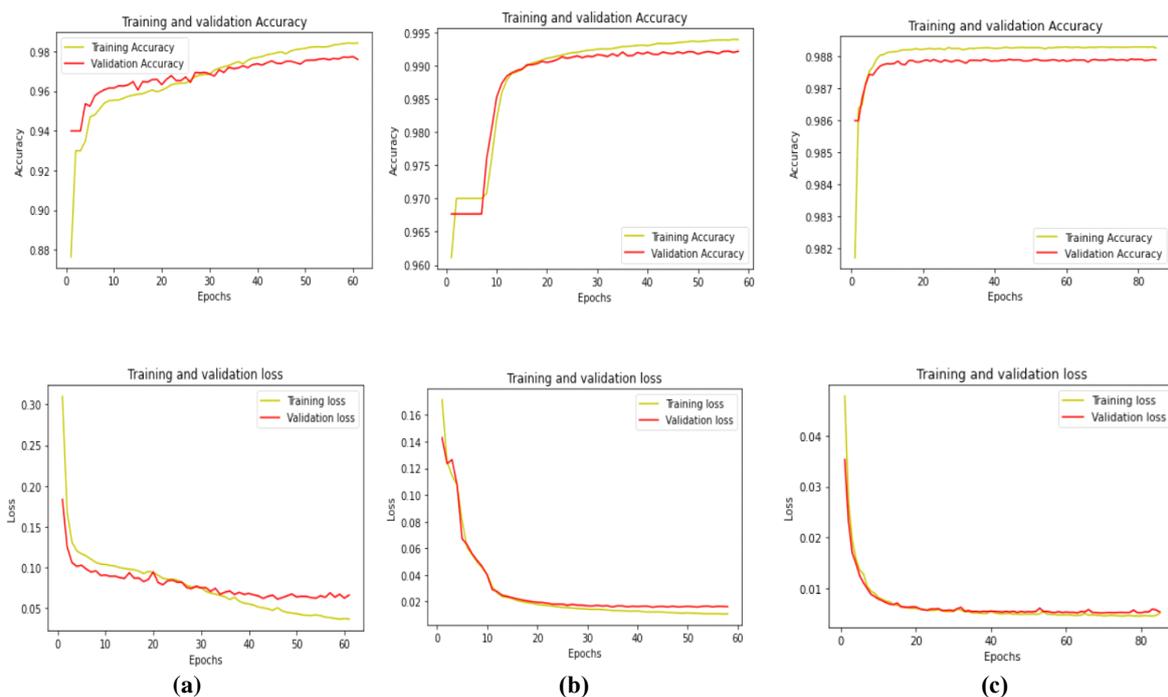


Figure 5.10 Segmentation for original 3D dataset 2 COVID19 SLICE_Z (2816 images/2816 masks (256x256))

COVID19 Lung Image preprocessed and OCAE+OUNET Augmentation + Resizing + OCAE
OCAE + OUNET



(a) Optimized Preprocessing Dataset D1 COVID-19 (256x256) (100 images/100 masks), (b) 3D dataset 2 COVID-19 SLICE_Z (100 images/100 masks), (c) 3D dataset 2 COVID-19 SLICE_Z (2816 images/2816 masks)

Figure 5.11 Mono modality accuracy and Loss curves for COVID19 Lung datasets

In the case of 2D and 3D mono modality, we notice that:

- The data augmentation plays a positive role;
- The optimization of the number of filters of the CAE model by PSO clearly improves the results;

- The original image is better suited for ML with a view to segmentation because it is a *Nifti* sequence; this is not the case for dataset 1 which is in the form of preprocessed images;
- The best time is obtained on the dataset without preprocessing and without data augmentation. While the time decreases on the dataset with preprocessing and data augmentation compared to the original lung image dataset.
- The accuracy is efficient on datasets with data augmentation despite the time being long.

Figure 10 confirms that the optimization of the number of CAE denoising model filters and the optimization of the Loss and the Acc of the UNET model used for the segmentation clearly improves the accuracy and loss performance curves. It also notices that the OCAE+OUNET model gives the best segmentation results because it manages to predict the smallest details and anomalies in the original image.

5.4.6.3 Multimodality

Experience 3: Multimodality with augmentation

In this study, 100 images/100 masks (256x256) from SLICE_Z, SLICE_Y and SLICE_X to demonstrate the effectiveness of our method.

The second experiment is applied on dataset D2 as shown table 5-6. In this case, different slices from axial, sagittal, coronal view of lung and masks infected by COVID19 are tested.

Table 5-6 Effect of the multi-view (multimodality) and data augmentation

MULTIMODALITY							
Dataset2: Original Lung Image (COVID19)							
SLICE_Z and SLICE_Y Nbr Filters Optimized: [25, 29, 20, 29]							
	Without Augmentation				With Augmentation		
	UNET	CAE+UNET	OCAE+OUNET		UNET	CAE+UNET	OCAE+OUNET
<i>Acc (%)</i>	97.97	98.57	98.58		98.89	98.90	99.04
<i>Loss</i>	0.0272	0.0075	0.0072		0.016	0.016	0.014
<i>DSC (%)</i>	66.41	93.35	93.57		90.28	90.73	90.72
<i>IoU (%)</i>	49.71	87.54	87.92		82.28	83.03	83.02
<i>Time(s)</i>	1.291	1.756	1.058		6.088	5.579	8.881
SLICE_Z, SLICE_Y and SLICE_X Nbr Filters Optimized : [32, 21, 32, 19]							
<i>Acc (%)</i>	98.79	98.82	98.81		99.09	99.24	99.25
<i>Loss</i>	0.008	0.006	0.006		0.015	0.011	0.011
<i>DSC (%)</i>	90.23	92.27	92.57		86.82	89.62	90.11
<i>IoU (%)</i>	82.20	85.65	86.17		76.71	81.20	81.99
<i>Time(s)</i>	1.859	1.764	1.891		8.381	10.216	9.015

The observation that accuracy tends to increase with increasing amount of data is

consistent with the common understanding of M. A large data size generally leads to better performance, because the model has more diverse examples, which ensures good learning and can be generalized to unseen data. However, the fact that some experiments without data augmentation present good results and short processing times we can deduce that data augmentation requires slower calculation and learning times. Additionally, some metrics may be more sensitive to changes introduced by increased data. Therefore, a choice depending on the progression and severity of the pathology is necessary because even if one measure improves with increase, another could deteriorate.

- **Dataset 2 SLICE_Z, SLICE_Y**

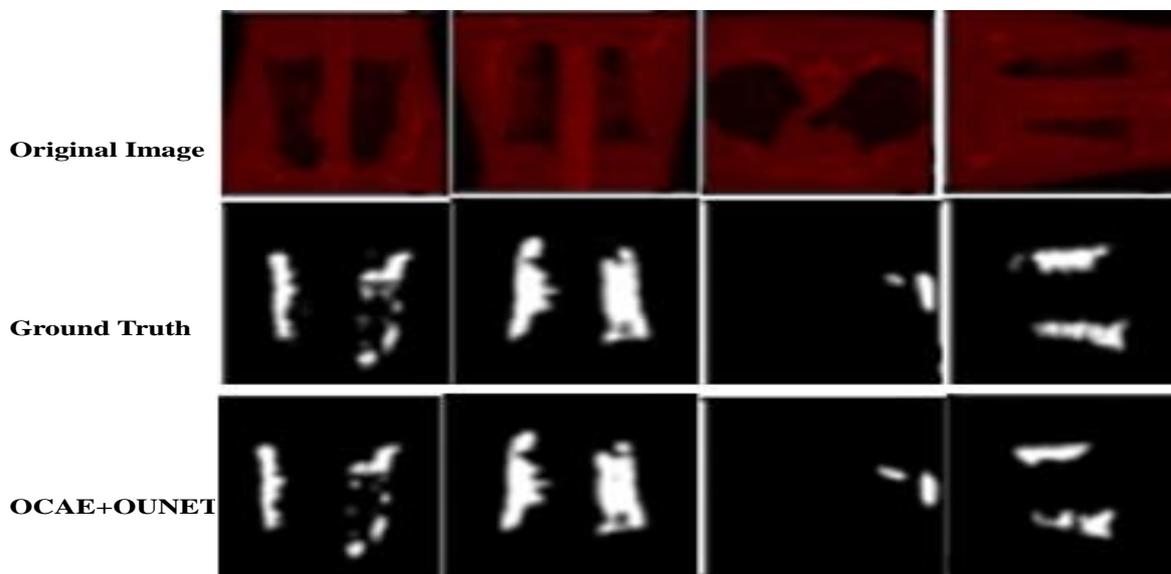


Figure 5.12 Multimodality Segmentation for 3D dataset 2 (SLICE_Z, SLICE_Y) COVID19

- **Dataset 2 SLICE_Z, SLICE_Y, SLICE_X with augmentation**

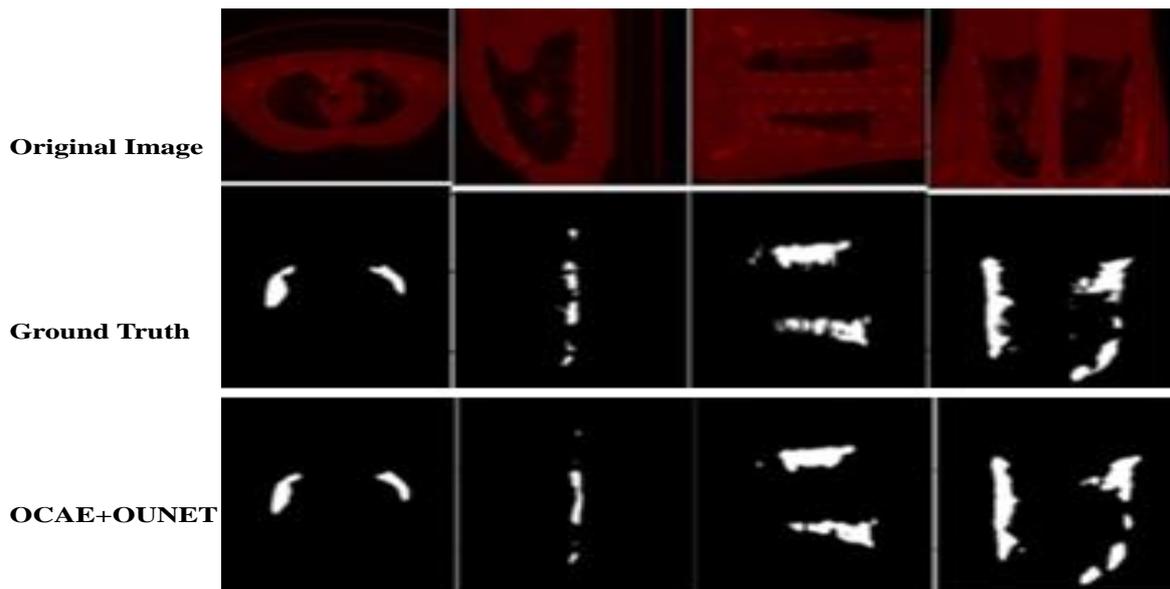
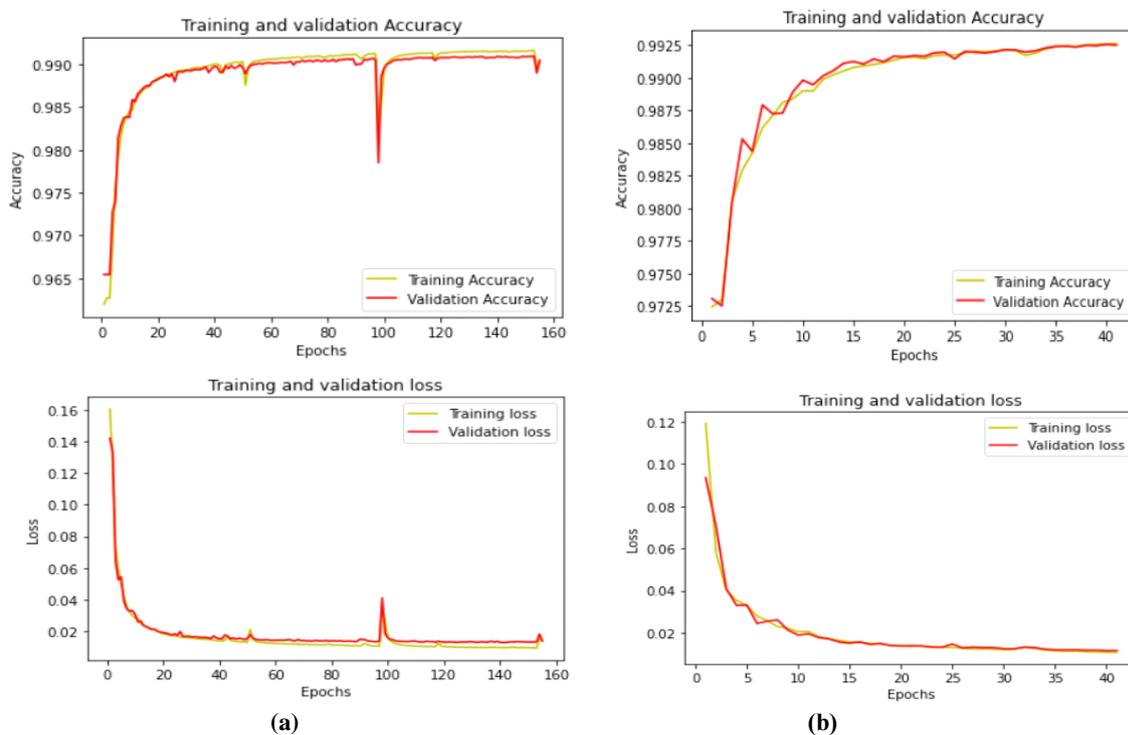


Figure 5.13 Multimodality Segmentation for 3D Dataset 2 COVID-19 SLICE_Z, SLICE_Y, SLICE_X

In the case of multimodality, the data augmentation played its role as well as the optimization of PSO and the use of all three modalities (SLICE_Z, SLICE_Y and SLICE_X). This is explained by the information provided by each of the modalities.

OCAE+OUNET



(a) Dataset 2 (SLICE_Z, SLICE_Y) COVID-19 (256x256) (images/100 masks), (b) Dataset 2 (SLICE_Z, SLICE_Y, SLICE_X) COVID-19 (100 images/100) masks with augmentation

Figure 5.14 Multimodality Accuracy and Loss curves

Analyzing slices from different views (axial, sagittal, and coronal) provides a more

comprehensive understanding of the 3D structure of the lungs and the spread of COVID19 infection. This multi-view approach ensures that the model learns from various perspectives, enhancing its ability to detect and segment the infected region.

The results obtained from SLICE_Z and SLICE_Y with and without data augmentation outperformed those from the three combined slices because each Different view (axial, sagittal, coronal) represent the anatomical structure differently. It's possible that certain orientations emphasize key features related to COVID19 infection more prominently (SLICE_Z and SLICE_Y).

Individual slices (especially SLICE_Z and SLICE_Y) provide higher resolution and more localized information compared to combined slices. This higher level of detail might be crucial in accurately identifying and segmenting small or intricate COVID19 infected regions within the lungs.

5.4.7 Application on different datasets

To validate the effectiveness and robustness of the developed model, we conducted experiments on diverse datasets for specific medical imaging tasks. The primary focus is on skin cancer analysis, for which we utilized a curate dataset consisting of 100 images with corresponding masks.

5.4.7.1 Dataset 3 Skin Cancer

In this case, 100 images/100 masks (256x256) are used to evaluate our proposed model.

Table 5-7 Application of the approach on Skin Cancer dataset

MONOMODALITY 2D Dataset 3: Skin Cancer							
Nbr Filters Optimized : [18, 22, 22, 26]							
	Without Augmentation				With Augmentation		
	UNET	CAE+UNET	OCAE+OUNET		UNET	CAE+UNET	OCAE+OUNET
<i>Acc (%)</i>	92.50	93.54	93.31		93.74	94.82	94.24
<i>Loss</i>	0.195	0.114	0.138		0.248	0.156	0.167
<i>DSC (%)</i>	89.76	92.03	91.48		92.38	93.81	93.12
<i>IoU (%)</i>	81.43	85.24	84.30		85.85	88.34	87.13
<i>Time(s)</i>	0.615	0.954	0.758		3.386	3.088	2.895

Afterwards, experience is validating with third dataset D3 that contains skin cancer and the masks.

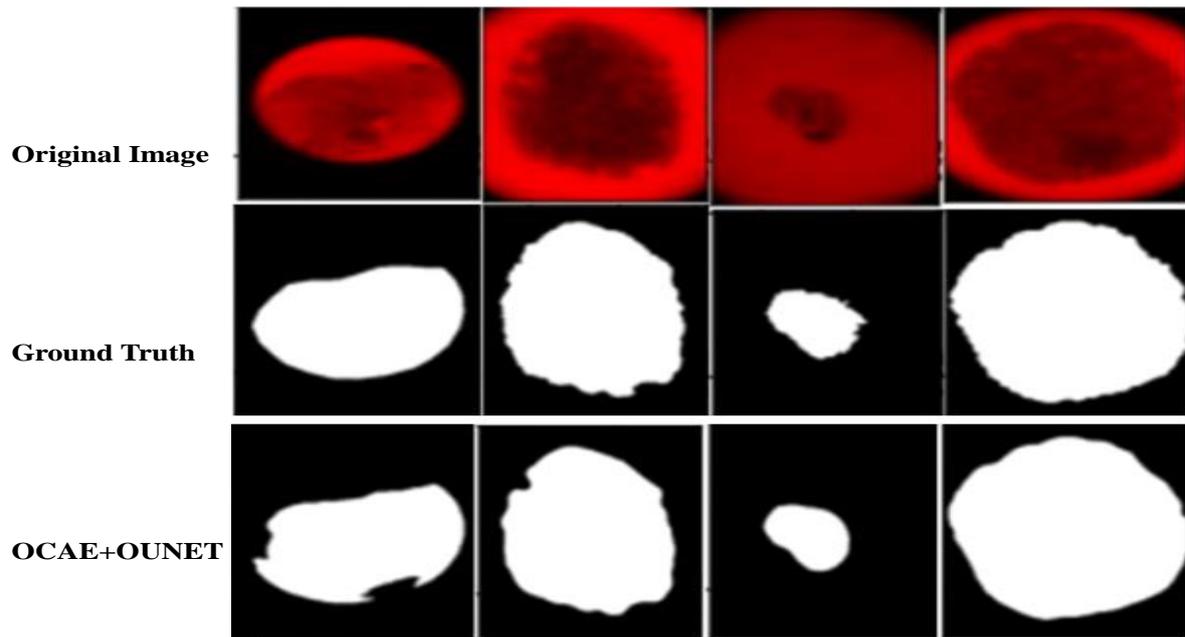


Figure 5.15 Segmentation for 2D Dataset 3 Skin Cancer (100 images/100 masks (256x256))

The effective performance of the proposed model, with high dice score of 91% and 93% with and without data augmentation respectively (as shown in table 5-7) on a completely different dataset (D3) indicates its robustness and ability to generalize across diverse medical imaging data. This is a critical factor in the practical application of any segmentation model.

5.4.7.2 Dataset 4 Synthetic Retinal

Furthermore, we expanded our validation to include Synthetic Retinal images, consisting of 91 images and 91 corresponding masks, as well as STARE Retinal datasets comprising 20 images and 20 masks. This extension broadened the scope of our analysis, allowing us to comprehensively assess the model's performance across varied datasets.

Table 5-8 Application of the approach on Synthetic Retinal dataset

MONOMODALITY							
2D Dataset 4: Synthetic Retinal							
Nbr Filters Optimized : [15, 32, 14, 23]							
	Without Augmentation				With Augmentation		
	UNET	CAE+UNET	OCAE+OUNET		UNET	CAE+UNET	OCAE+OUNET
<i>Acc (%)</i>	70.45	70.45	70.44		83.80	83.75	83.92
<i>Loss</i>	0.129	0.125	0.128		0.127	0.129	0.127
<i>DSC (%)</i>	65.74	66.11	66.86		67.36	66.82	66.60
<i>IoU (%)</i>	48.97	49.38	50.22		50.78	50.17	49.92
<i>Time(s)</i>	0.634	0.655	0.735		3.417	2.811	2.658

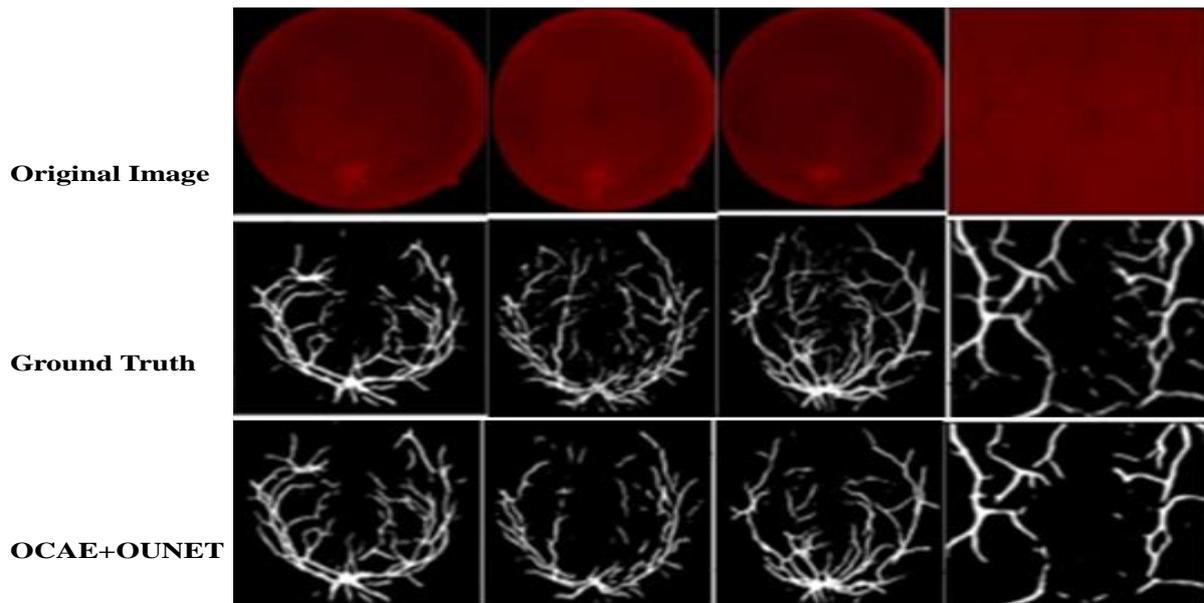


Figure 5.16 Segmentation for 2D Dataset 4 Synthetic Retinal (91 images/91 masks (256x256))

5.4.7.3 Dataset 5 Stare Retinal dataset

Unlike synthetic retinal datasets, which may lack the authenticity and realism of real clinical images, datasets derived from actual clinical cases such as STARE offer genuine, diverse, and varied retinal images that effectively capture the complexity of retinal conditions. For this reason, we chose to evaluate our model using the STARE retinal dataset. We assessed our model's performance on authentic and diverse retinal images, providing a more reliable assessment of its capabilities in real-world scenarios. Due to the limited size of the STARE Retinal dataset, which consists of only 20 images and 20 masks, we applied data augmentation techniques to enhance our model's performance. Table 5-9 displays the various results obtained.

Table 5-9 Application of the approach on Stare Retinal dataset

MONOMODALITY						
2D Dataset 5 STARE Retinal (20 images/20 masks)						
Nb Filters Optimized : [8, 29, 15, 22]						
	<i>Without Augmentation</i>			<i>With Augmentation</i>		
	UNET	CAE+UNET	PSO_CAE+UNET	UNET	CAE+UNET	OCAE+OUNET
<i>Acc (%)</i>	–	–	–	92.66	93.26	93.24
<i>Loss</i>	–	–	–	0.141	0.109	0.113
<i>DSC (%)</i>	–	–	–	63.13	71.87	72.07
<i>IoU (%)</i>	–	–	–	46.11	56.09	56.34
<i>Time(s)</i>	–	–	–	0.854	0.758	1.833

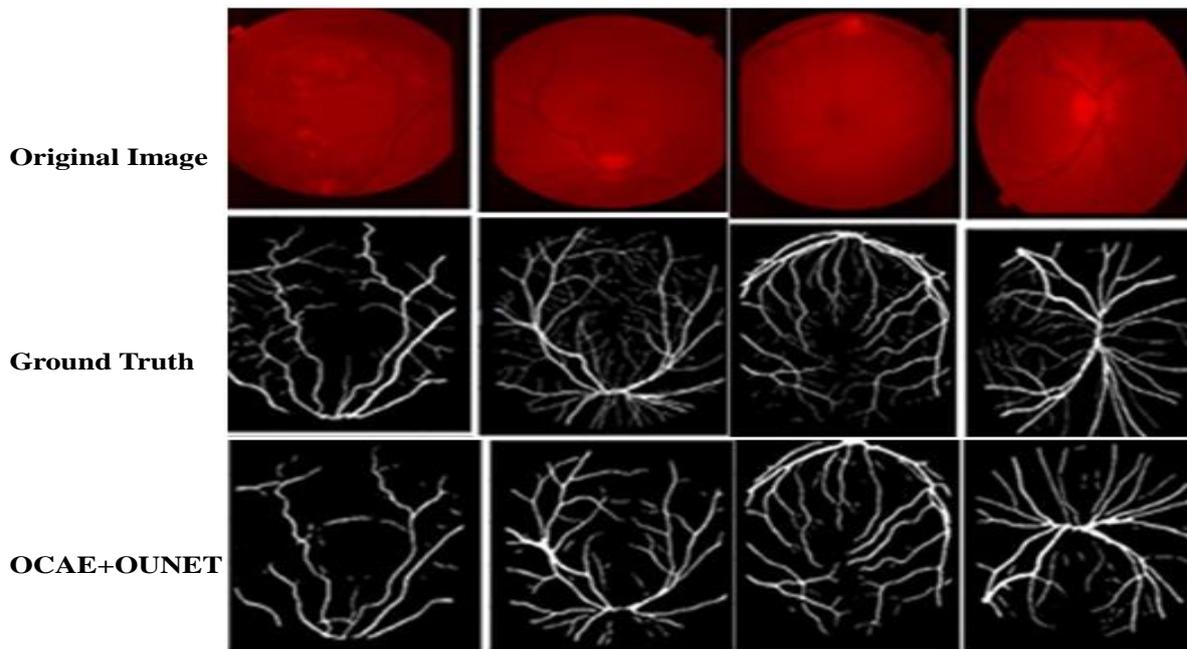


Figure 5.17 Segmentation for 2D Dataset 5 STARE Retinal (20 images/20 masks (256x256))

Finally, experiment with database D4 that contains Synthetic and database D5 contains STARE Retinal images and masks; Synthetic images, especially in medical contexts, are computer-generated and lack the organic complexity of real-world images. These subtle differences, such as variations in tissue texture, color gradients, and lighting conditions, are critical for medical professionals. In real medical scenarios, these variations often carry diagnostic significance, aiding doctors in identifying specific conditions or abnormalities. For this reason, the proposed model demonstrates superior performance when applied to STARE Retinal images in comparison to Synthetic Retinal images (as shown in table 5-8 and 5-9). The disparities in image realism and complexity between these datasets significantly influenced the outcomes, highlighting the challenges associated with working with synthetic data.

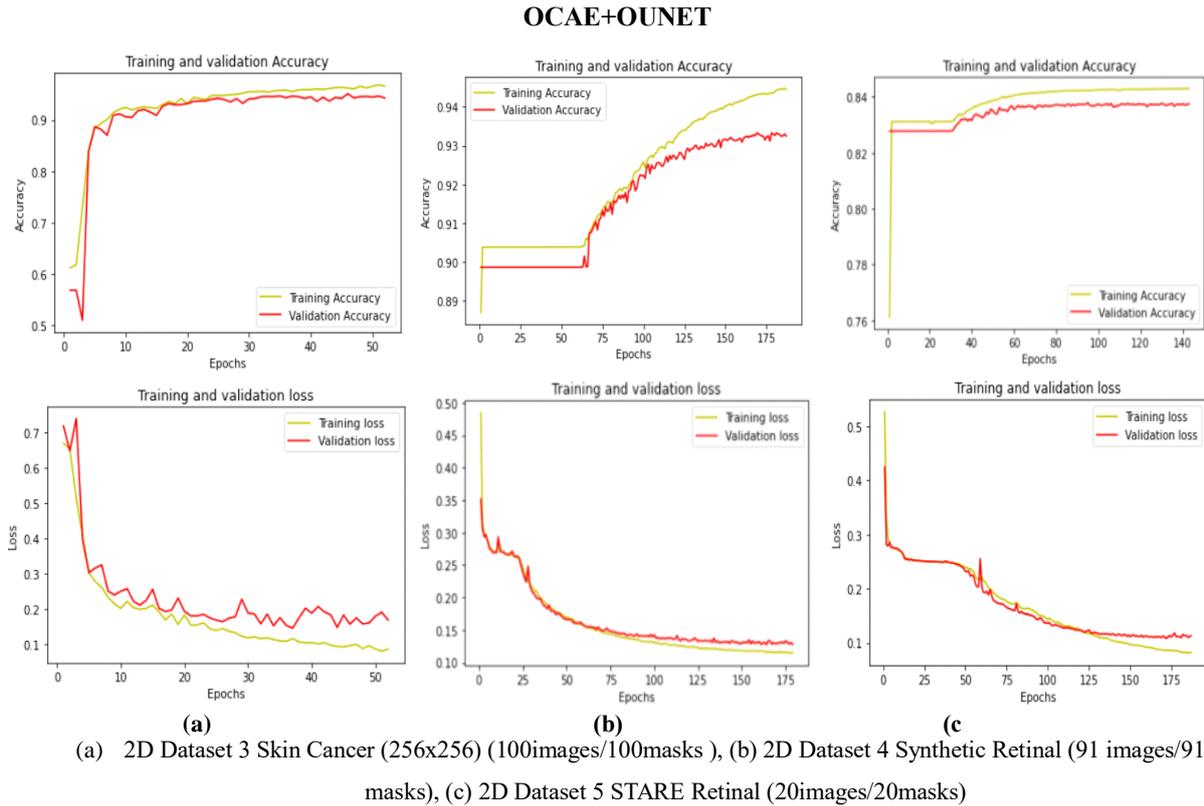


Figure 5.18 Accuracy and Loss curves for different datasets

Based on the obtained results, it is evident that the OCAE+OUNET model displayed suboptimal parameter values. Notably, the performance curves failed to converge or stabilize when applied to the 2D Dataset 4 Synthetic Retinal (see Fig 5.18 (b)). This lack of convergence indicates a potential limitation or challenge in the model's adaptability to this specific dataset. In addition, when the same model was employed on the 2D Dataset 5 STARE Retinal, a noteworthy improvement was observed (see Fig 5.18 (c)). The performance curves exhibited significant enhancement and successful convergence, suggesting a more favorable interaction between the model architecture and the characteristics of the STARE Retinal dataset. This divergence in performance between the two retinal datasets highlights the importance of dataset-specific considerations in model evaluation and selection.

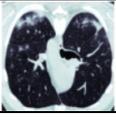
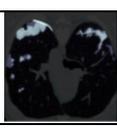
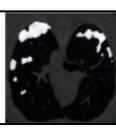
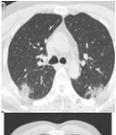
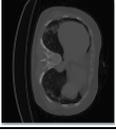
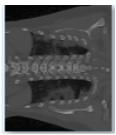
Furthermore, in the case of the 2D Dataset 3 Skin Cancer, the OCAE+OUNET model demonstrated superior performance. The performance curves for this dataset exhibited clear optimization, indicating that the model effectively adapted to the features present in the Skin Cancer dataset (see Fig 5.18 (a)). This positive outcome emphasizes the model's potential and effectiveness in handling diverse datasets, but it also underscores the need for a nuanced understanding of its performance across various domains. These findings underscore the importance of considering dataset-specific nuances and potential limitations when assessing the

performance of complex models, providing valuable insights for further refinement and optimization.

5.5 State of the Art Comparison

In this work, we will compare our proposed method with different state of the art methods.

Table 5-10 Comparison with the state-of-the-art methods

<i>Authors</i>	<i>Methods</i>	<i>Datasets</i>	<i>DSC (%)</i>	<i>Time (s)</i>	<i>IoU (%)</i>	<i>Acc (%)</i>	<i>Original images</i>	<i>Ground Truths</i>	<i>Segmented Images</i>
COVID19									
Amyar et al. [140]	Encoder-Decoder	CT COVID-19 100 images/ 100 masks	88	–	–	94.67			
Fan et al. [138]	Inf-Net	CT axial COVID19 100 images/ 100 masks	Inf-Net = 68.2 Semi-Inf-Net = 72.5	–	–	–			
Saha et al. [149]	ADU+ Net	CT COVID-19 100 images/ 100 masks	86	–	–	–			
Elghamrawy et al. [166]	AIMDP	432 Images/ 432 masks	–	Classifier : SVM: 98.94 NB: 101.3 DA: 99.2	–	90.4		–	
	AIMDP+ FSWOA	432 Images/ 432 masks	–		–	Classifier : SVM 97.14 NB 94.99 DA 94.71			
Agarwal et al. [167]	FCN+ PSO	lungs X-Ray dataset	–	–	–	–			
Ours	OCAE+ OUNET	Preprocessing CT COVID-19 SLICE_Z 100 images/ 100 masks	81.13	3.06	63.85	97.83			
		Original images CT COVID-19 SLICE_Z 100 images/ 100 masks	89.46	3.11	80.93	99.21			
		Original images COVID-19 SLICE_Z, SLICE_Y, 100 images/ 100 masks	93.57	1.058	87.92	99.04			

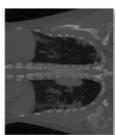
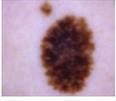
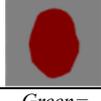
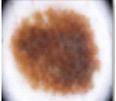
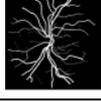
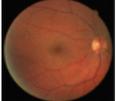
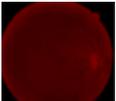
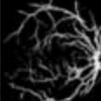
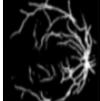
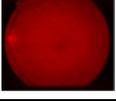
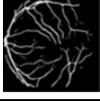
		Original images COVID-19 SLICE_Z, SLICE_Y, SLICE_X 100 images/ 100 masks	92.57	1.89	86.17	99.25			
Skin cancer									
Kumar et al. [152]	AWO+based Squeeze Net	Skin Cancer 1800 images / 1800 masks	–	–	–	92.5		–	
Attia et al. [168]	CNN+RNN	Skin Cancer 900 images/ 900 masks	93	–	–	98			
Vesal et al. [169]	SkinNet	Skin Cancer 2000 images/ 2000 masks	85.1	–	76.67	–		Green= ground truth Blue= SkinNet output	
Ahmed et al. [170]	ITSU+Efficient Net	Skin cancer dataset (ISIC 2018 challenge dataset:2594 images	94.36			92.9			
Ours	OCAE+OUNET	Skin Cancer 100 images/ 100 masks	93.12	2.89	88.34	94.24			
Retinal									
L.Yan et al. [171]	AA+UNET	Retinal STARE 20 images/ 20 masks	–	–	96.24	96.40			
Yanhonget al. [172]	ResDO+UNET	Retinal STARE 20 images/ 20 masks	–	–	–	95.67			
Yan et al. [173]	OTSU+PSO	Retinal STARE 20 images/ 20 masks	–	–	–	95.79			
Popat et al. [174]	GA+UNET	Segmenting blood vessels in retinal images (stare retinal)	–	–	–	0.9668	–	–	–
Ours	OCAE+OUNET	Synthetic Retinal 100 images/ 100 masks	66.86	0.73	50.22	70.44			
		Retinal STARE 20 images/ 20 masks	72.07	1.83	56.34	93.24			

Table 5-10 provides a comprehensive comparative overview of the proposed approach against various state-of-the-art methods across three diverse datasets: COVID19, skin cancer, and retinal images. The purpose of this analysis is to evaluate the effectiveness of the proposed

method in comparison to existing techniques for image segmentation. The table illustrates that the proposed method consistently outperforms most of the state-of-the-art segmentation techniques. Different metrics, accuracy, and Dice Similarity Coefficient (DSC), have been employed to gauge the segmentation quality. The accuracy metric reflects the precision of the segmentation process, while DSC quantifies the spatial overlap between the predicted and ground truth segmentations. In the context of the COVID19 dataset, the proposed method showcases remarkable accuracy of 99.04% and DSC scores of 93.57% with the two slices. This suggests the method's proficiency in accurately delineating infected regions within lung images, a critical task in COVID-19 diagnosis and monitoring.

Moving to the skin cancer dataset, the proposed method continues to demonstrate superior segmentation results. Achieving high DSC scores of 93.12% in skin cancer segmentation is pivotal for precise identification of cancerous lesions, aiding in early detection and subsequent treatment planning. Furthermore, within the retinal dataset, the proposed method maintains its superiority, indicating its effectiveness in segmenting complex structures within retinal images. Accurate segmentation with an accuracy of 93.24% and DSC of 72.07% in retinal imaging is vital for diagnosing various eye conditions, making these results particularly promising for ophthalmic applications.

While our work shows promising results, there are certain limitations that should be acknowledged. Firstly, our approach was tested on small datasets, and its performance on larger datasets remains to be explored. Additionally, the utilization of hybrid metaheuristic algorithms could further enhance the optimization process, and this avenue warrants further investigation. Moreover, the real-world applicability and generalizability of proposed model in diverse clinical settings need to be thoroughly validated before practical implementation. Addressing these limitations will be crucial for the broader adoption and effectiveness of our proposed methodology in the field of medical image segmentation and disease diagnosis.

5.6 Conclusion

New advancements in AI have made a huge contribution to enhancing our lives, and AI has become widely used in our daily lives in a multitude of situations. As a result, there is a strong belief that indeed designed AI research will fully utilize AI's potential in supporting and helping humans in combating rapidly and diagnosing different diseases. The framework presented in this work is based on DL optimized by a metaheuristics algorithm called PSO. We

have proposed to apply the PSO optimization algorithm in the CAE neural network to optimize its network parameters to denoising the datasets of images without losing information included in it before the segmentation stage. The UNET was optimized to achieve high accuracy, low loss; low-test time and robustness. The results indicated that the proposed approach gives well dice scores with different pathologies. The current study proposed image denoising with CAE optimized by PSO before segmentation task, which was trained with different datasets images. The approach has given encouraging results on different small datasets. The results remain stable and efficient from one dataset to another.

In the future, we plan to evaluate our model using large datasets and explore the integration of hybrid metaheuristic algorithms. Adapting and validating our techniques on extensive datasets will be crucial for ensuring the robustness and reliability of our models in real-world, large-scale medical applications.

General Conclusion

This thesis has presented a robust and multifaceted approach to medical image analysis, focusing on the early detection and diagnosis of COVID-19 while extending its applicability to other critical diseases. By combining advanced DL architectures, innovative preprocessing techniques, and optimization strategies, we have developed a framework capable of addressing complex challenges in medical imaging.

Our work began by tackling the problem of COVID-19 detection using chest X-ray and CT images. MobileNet_V2 was employed as the backbone for feature extraction due to its efficiency and accuracy. For enhanced feature extraction, we integrated VGG16 with Gabor filters, enabling the capture of detailed texture and structural patterns within medical images. A significant contribution of this study was the use of a CAE for denoising images, ensuring cleaner inputs for segmentation tasks. To further enhance the CAE's effectiveness, we applied the PSO algorithm to optimize the filter parameters, significantly improving the segmentation accuracy of the U-Net model.

A key strength of this thesis lies in the systematic comparison of monomodality and multimodality approaches, which provided valuable insights into their relative effectiveness. By employing techniques such as HE for image enhancement and hybrid descriptors combining Gabor filters with LBP, we achieved higher classification accuracy and demonstrated the versatility of the proposed model. The inclusion of a curated dataset focusing on pneumonia and other viral lung diseases, given their visual similarity to COVID-19, allowed us to expand the model's capability to multiclass classification.

Moreover, this research extended beyond COVID-19 to explore its applicability to other diseases. The model demonstrated adaptability by effectively segmenting and classifying conditions such as skin cancer and retinal abnormalities. For additional classification tasks, the MRFO algorithm combined with VGG16 enabled accurate detection of diseases like lung cancer, breast cancer, and Alzheimer's disease, showcasing the framework's scalability and versatility in addressing diverse medical challenges.

While the proposed methodologies achieved commendable results, this work also acknowledged certain limitations. The effectiveness of the model is influenced by the quality and quantity of the dataset, which underscores the need for larger, more diverse datasets to enhance training and evaluation. Additionally, further research is required to evaluate the

model's generalizability to broader medical imaging domains and ensure its robustness across varied clinical environments.

The utilization of AI-driven technologies in the field of healthcare is still mostly a research area at this time. However, the general understanding is that in the near future, these revolutionary technologies will completely change the field of medical practice. The healthcare sector typically takes new innovations at a more controlled pace than certain industries, which rapidly incorporate technology. Furthermore, without the active engagement, collaboration, and support of significant clients, such as healthcare professionals and patients, the successful integration of AI into the current medical workflow faces tremendous obstacles. For this developmental journey to be successful, a strong legal and regulatory framework is essential.

Limitation and Perspectives: The advancements in this thesis open up numerous opportunities for further research and improvement. Future work should address the following areas:

The current model's performance is highly dependent on the dataset's quality and variety. Incorporating larger and more diverse datasets, including multi-institutional data, will improve model generalization and applicability to broader patient populations. Additionally, datasets that capture rare conditions or anomalies can further expand the model's diagnostic capabilities.

While traditional augmentation methods were used in this work, future studies could explore generative approaches like Generative Adversarial Networks (GANs) to synthesize realistic medical images. This would enhance data diversity and help mitigate data scarcity issues in rare diseases.

To improve the clinical adoption of AI models, integrating explainable AI techniques is essential. Future work could focus on developing interpretable models that provide insights into the decision-making process, highlighting the critical regions or features used for diagnosis.

While PSO and MRFO have been applied successfully, exploring other optimization algorithms like Ant Colony Optimization (ACO), Genetic Algorithms (GA), or hybrid optimization methods may yield further performance improvements. These techniques could also help reduce computational overhead.

The contributions of this research have significant implications for the future of medical image analysis. By leveraging hybrid models, optimization techniques, and TL, we have demonstrated a pathway toward the development of diagnostic tools that are not only accurate

but also adaptable and efficient. These advancements hold the potential to revolutionize medical imaging by enabling early detection and precise diagnosis of diseases, thereby improving patient outcomes and supporting clinical decision-making.

In conclusion, this thesis underscores the transformative potential of DL and optimization techniques in advancing medical image analysis. By addressing critical challenges and proposing future directions, the methodologies developed in this research pave the way for innovative diagnostic solutions with wide-ranging applications across medical imaging domains. The findings presented here lay a strong foundation for further exploration, aiming to bridge the gap between cutting-edge technology and practical healthcare delivery, ultimately contributing to better patient care and clinical outcomes.

References

- [1] A. A. Khan, A. A. Laghari, and S. A. Awan, "Machine learning in computer vision: a review," *EAI Endorsed Trans. Scalable Inf. Syst.*, vol. 8, no. 32, pp. e4–e4, 2021.
- [2] K. Suzuki, "Overview of deep learning in medical imaging," *Radiol. Phys. Technol.*, vol. 10, no. 3, pp. 257–273, 2017.
- [3] A. Chaddad, J. Peng, J. Xu, and A. Bouridane, "Survey of explainable AI techniques in healthcare," *Sensors*, vol. 23, no. 2, p. 634, 2023.
- [4] M. A. Azam *et al.*, "A review on multimodal medical image fusion: Compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics," *Comput. Biol. Med.*, vol. 144, p. 105253, 2022.
- [5] K. Chockley and E. Emanuel, "The end of radiology? Three threats to the future practice of radiology," *J. Am. Coll. Radiol.*, vol. 13, no. 12, pp. 1415–1420, 2016.
- [6] K. Iniewski, *Medical imaging: principles, detectors, and electronics*. John Wiley & Sons, 2009.
- [7] P. Suetens, *Fundamentals of medical imaging*. Cambridge university press, 2017.
- [8] S. T. H. Kieu, A. Bade, M. H. A. Hijazi, and H. Kolivand, "A survey of deep learning for lung disease detection on medical images: state-of-the-art, taxonomy, issues and future directions," *J. imaging*, vol. 6, no. 12, p. 131, 2020.
- [9] A. Khalil, *Imagerie thoracique de l'adulte et de l'enfant*. Elsevier Health Sciences, 2013.
- [10] S. Bulusu, B. Kailkhura, B. Li, P. K. Varshney, and D. Song, "Anomalous example detection in deep learning: A survey," *IEEE Access*, vol. 8, pp. 132330–132347, 2020.
- [11] I. Wahlang *et al.*, "Deep Learning methods for classification of certain abnormalities in Echocardiography," *Electronics*, vol. 10, no. 4, p. 495, 2021.
- [12] R. Wang, T. Lei, R. Cui, B. Zhang, H. Meng, and A. K. Nandi, "Medical image segmentation using deep learning: A survey," *IET Image Process.*, vol. 16, no. 5, pp. 1243–1267, 2022.
- [13] L. Hussain *et al.*, "Machine-learning classification of texture features of portable chest X-ray accurately classifies COVID-19 lung infection," *Biomed. Eng. Online*, vol. 19, pp. 1–18, 2020.
- [14] L. Wang, Z. Q. Lin, and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," *Sci. Rep.*, vol. 10, no. 1, p. 19549, 2020.
- [15] Y. E. Almalki *et al.*, "A novel method for COVID-19 diagnosis using artificial intelligence in chest X-ray images," in *Healthcare*, 2021, vol. 9, no. 5, p. 522.
- [16] R. F. Mansour, J. Escorcia-Gutierrez, M. Gamarra, D. Gupta, O. Castillo, and S. Kumar, "Unsupervised deep learning based variational autoencoder model for COVID-19 diagnosis and classification," *Pattern Recognit. Lett.*, vol. 151, pp. 267–274, 2021.
- [17] R. Fusco *et al.*, "Artificial intelligence and COVID-19 using chest CT scan and chest X-ray images: machine learning and deep learning approaches for diagnosis and treatment," *J. Pers. Med.*, vol. 11, no. 10, p. 993, 2021.
- [18] Y. Meraihi, A. B. Gabis, S. Mirjalili, A. Ramdane-Cherif, and F. E. Alsaadi, "Machine learning-based research for covid-19 detection, diagnosis, and prediction: A survey," *SN Comput. Sci.*, vol. 3, no. 4, p. 286, 2022.
- [19] F. Saleem, A. S. A.-M. Al-Ghamdi, M. O. Alassafi, and S. A. AlGhamdi, "Machine learning, deep learning, and mathematical models to analyze forecasting and epidemiology of COVID-19: A systematic literature review," *Int. J. Environ. Res. Public Health*, vol. 19, no. 9, p. 5099, 2022.
- [20] Y. M. G. Costa *et al.*, "Covid-19 detection on chest x-ray and ct scan: a review of the top-100 most cited papers," *Sensors*, vol. 22, no. 19, p. 7303, 2022.

- [21] 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.
- [22] R. Indraswari, R. Rokhana, and W. Herulambang, "Melanoma image classification based on MobileNetV2 network," *Procedia Comput. Sci.*, vol. 197, pp. 198–207, 2022.
- [23] M. Rahimzadeh and A. Attar, "A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2," *Informatics Med. unlocked*, vol. 19, p. 100360, 2020.
- [24] A. Bhattacharyya, D. Bhaik, S. Kumar, P. Thakur, R. Sharma, and R. B. Pachori, "A deep learning based approach for automatic detection of COVID-19 cases using chest X-ray images," *Biomed. Signal Process. Control*, vol. 71, no. PB, p. 103182, 2022, doi: 10.1016/j.bspc.2021.103182.
- [25] N. Muralidharan, S. Gupta, M. R. Prusty, and R. K. Tripathy, "Detection of COVID19 from X-ray images using multiscale Deep Convolutional Neural Network," *Appl. Soft Comput.*, vol. 119, p. 108610, 2022.
- [26] S. D. Deb, R. K. Jha, K. Jha, and P. S. Tripathi, "A multi model ensemble based deep convolution neural network structure for detection of COVID19," *Biomed. Signal Process. Control*, vol. 71, p. 103126, 2022.
- [27] N. S. Kavya, N. Veeranjanyulu, and D. D. Priya, "Detecting Covid19 and Pneumonia from chest X-Ray images using Deep Convolutional Neural Networks," *Mater. Today Proc.*, 2022.
- [28] M. R. Islam and M. Nahiduzzaman, "Complex features extraction with deep learning model for the detection of COVID19 from CT scan images using ensemble based machine learning approach," *Expert Syst. Appl.*, vol. 195, p. 116554, 2022.
- [29] M. Ketfi, M. Belahcene, and S. Bourenane, "Efficient Diagnosis COVID-19 using Gabor and Transfer Learning," in *2021 1st International Conference On Cyber Management And Engineering (CyMaEn)*, 2021, pp. 1–6.
- [30] C. Sitaula and M. B. Hossain, "Attention-based VGG-16 model for COVID-19 chest X-ray image classification," *Appl. Intell.*, vol. 51, no. 5, pp. 2850–2863, 2021.
- [31] A. Maity, T. R. Nair, and A. Chandra, "Image Pre-processing techniques comparison: COVID-19 detection through Chest X-Rays via Deep Learning," *Int. J. Sci. Res. Sci. Technol.*, vol. 7, pp. 113–123, 2020.
- [32] A. H. Barshooi and A. Amirkhani, "A novel data augmentation based on Gabor filter and convolutional deep learning for improving the classification of COVID-19 chest X-Ray images," *Biomed. Signal Process. Control*, vol. 72, p. 103326, 2022.
- [33] A. Hadid, J. Ylioinas, M. Bengherabi, M. Ghahramani, and A. Taleb-Ahmed, "Gender and texture classification: A comparative analysis using 13 variants of local binary patterns," *Pattern Recognit. Lett.*, vol. 68, pp. 231–238, 2015.
- [34] A. Dhall, R. Goecke, S. Lucey, and T. Gedeon, "Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark," in *2011 IEEE international conference on computer vision workshops (ICCV workshops)*, 2011, pp. 2106–2112.
- [35] "iCTCF - CT images and clinical features for COVID-19." <http://ictcf.biocuckoo.cn/Resource.php> (accessed Sep. 22, 2022).
- [36] "Chest X-ray (Covid-19 & Pneumonia) | Kaggle." <https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia> (accessed Sep. 22, 2022).
- [37] J. David Freire *et al.*, "The Impact of Histogram Equalization and Color Mapping on ResNet-34's Overall Performance for COVID-19 Detection," in *2021 4th International Conference on Data Storage and Data Engineering*, 2021, pp. 45–51.
- [38] T. Rahman *et al.*, "Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images," *Comput. Biol. Med.*, vol. 132, p. 104319, 2021.
- [39] S. Lawton and S. Viriri, "Detection of COVID-19 from CT lung scans using transfer learning," *Comput. Intell. Neurosci.*, vol. 2021, 2021.

- [40] M. Kaya and M. Eris, "D3SENet: A hybrid deep feature extraction network for Covid-19 classification using chest X-ray images," *Biomed. Signal Process. Control*, p. 104559, 2023.
- [41] H. A. El Shenbary, E. A. Ebeid, and D. Baleanu, "COVID-19 classification using hybrid deep learning and standard feature extraction techniques," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 29, no. 3, pp. 1780–1791, 2023.
- [42] P. Aggarwal, N. K. Mishra, B. Fatimah, P. Singh, A. Gupta, and S. D. Joshi, "COVID-19 image classification using deep learning: Advances, challenges and opportunities," *Comput. Biol. Med.*, p. 105350, 2022.
- [43] A. K. Das, S. Ghosh, S. Thunder, R. Dutta, S. Agarwal, and A. Chakrabarti, "Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network," *Pattern Anal. Appl.*, vol. 24, no. 3, pp. 1111–1124, 2021.
- [44] I. Lahsaini, M. E. H. Daho, and M. A. Chikh, "Deep transfer learning based classification model for covid-19 using chest CT-scans," *Pattern Recognit. Lett.*, vol. 152, pp. 122–128, 2021.
- [45] N. Khanday and S. Sofi, "Deep Insight: Convolutional Neural Network and its Applications for COVID-19 Prognosis," *Biomed. Signal Process. Control*, vol. 69, p. 102814, May 2021, doi: 10.1016/j.bspc.2021.102814.
- [46] M. K. Hasan *et al.*, "Challenges of deep learning methods for COVID-19 detection using public datasets," *Informatics Med. Unlocked*, vol. 30, p. 100945, 2022.
- [47] S. H. Khan *et al.*, "COVID-19 detection in chest X-ray images using deep boosted hybrid learning," *Comput. Biol. Med.*, vol. 137, p. 104816, 2021.
- [48] M. Loey, S. El-Sappagh, and S. Mirjalili, "Bayesian-based optimized deep learning model to detect COVID-19 patients using chest X-ray image data," *Comput. Biol. Med.*, vol. 142, p. 105213, 2022.
- [49] X. He *et al.*, "Sample-efficient deep learning for COVID-19 diagnosis based on CT scans. medrxiv 2020," *Google Sch.*, 2020.
- [50] F. Z. Benyelles, A. Sekkal, and N. Settouti, "Content based COVID-19 chest X-Ray and CT images retrieval framework using stacked auto-encoders," in *2020 2nd international workshop on human-centric smart environments for health and well-being (IHSH)*, 2021, pp. 119–124.
- [51] S. Wang *et al.*, "A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)," *Eur. Radiol.*, vol. 31, no. 8, pp. 6096–6104, 2021.
- [52] M. V. Madhavan, A. Khamparia, D. Gupta, S. Pande, P. Tiwari, and M. S. Hossain, "Res-CovNet: an internet of medical health things driven COVID-19 framework using transfer learning," *Neural Comput. Appl.*, pp. 1–14, 2021.
- [53] E. Khan, M. Z. U. Rehman, F. Ahmed, F. A. Alfouzan, N. M. Alzahrani, and J. Ahmad, "Chest X-ray classification for the detection of COVID-19 using deep learning techniques," *Sensors*, vol. 22, no. 3, p. 1211, 2022.
- [54] A. Bhattacharyya, D. Bhaik, S. Kumar, P. Thakur, R. Sharma, and R. B. Pachori, "A deep learning based approach for automatic detection of COVID-19 cases using chest X-ray images," *Biomed. Signal Process. Control*, vol. 71, p. 103182, 2022.
- [55] M. Loey, G. Manogaran, and N. E. M. Khalifa, "A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images," *Neural Comput. Appl.*, pp. 1–13, 2020.
- [56] D. Wang, J. Mo, G. Zhou, L. Xu, and Y. Liu, "An efficient mixture of deep and machine learning models for COVID-19 diagnosis in chest X-ray images," *PLoS One*, vol. 15, no. 11, p. e0242535, 2020.
- [57] O. El Gannour *et al.*, "Concatenation of Pre-Trained Convolutional Neural Networks for Enhanced COVID-19 Screening Using Transfer Learning Technique," *Electronics*, vol. 11, no. 1, p. 103, 2021.
- [58] A. I. Khan, J. L. Shah, and M. M. Bhat, "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images," *Comput. Methods Programs Biomed.*, vol. 196, p. 105581, 2020.

- [59] S. Dilshad *et al.*, “Automated image classification of chest X-rays of COVID-19 using deep transfer learning,” *Results Phys.*, vol. 28, p. 104529, 2021.
- [60] D. Li, Z. Fu, and J. Xu, “Stacked-autoencoder-based model for COVID-19 diagnosis on CT images,” *Appl. Intell.*, vol. 51, no. 5, pp. 2805–2817, 2021.
- [61] P. K. Shukla, J. K. Sandhu, A. Ahirwar, D. Ghai, P. Maheshwary, and P. K. Shukla, “Multiobjective genetic algorithm and convolutional neural network based COVID-19 identification in chest X-ray images,” *Math. Probl. Eng.*, vol. 2021, 2021.
- [62] Z. Ullah, M. Usman, and J. Gwak, “MTSS-AAE: Multi-task semi-supervised adversarial autoencoding for COVID-19 detection based on chest X-ray images,” *Expert Syst. Appl.*, p. 119475, 2023.
- [63] D. Addo *et al.*, “EVAE-Net: An Ensemble Variational Autoencoder Deep Learning Network for COVID-19 Classification Based on Chest X-ray Images,” *Diagnostics*, vol. 12, no. 11, p. 2569, 2022.
- [64] A. S. K. Reddy *et al.*, “Multi-modal fusion of deep transfer learning based COVID-19 diagnosis and classification using chest x-ray images,” *Multimed. Tools Appl.*, pp. 1–25, 2022.
- [65] K. H. Abdulkareem *et al.*, “[Retracted] Automated System for Identifying COVID-19 Infections in Computed Tomography Images Using Deep Learning Models,” *J. Healthc. Eng.*, vol. 2022, no. 1, p. 5329014, 2022.
- [66] F. Demir, K. Demir, and A. Şengür, “DeepCov19Net: automated COVID-19 disease detection with a robust and effective technique deep learning approach,” *New Gener. Comput.*, pp. 1–23, 2022.
- [67] N. Rashid, M. A. F. Hossain, M. Ali, M. I. Sukanya, T. Mahmud, and S. A. Fattah, “AutoCovNet: Unsupervised feature learning using autoencoder and feature merging for detection of COVID-19 from chest X-ray images,” *Biocybern. Biomed. Eng.*, vol. 41, no. 4, pp. 1685–1701, 2021.
- [68] N. Agarwal, S. N. Mohanty, S. Sankhwar, and J. K. Dash, “A Novel Model to Predict the Effects of Enhanced Students’ Computer Interaction on Their Health in COVID-19 Pandemics,” *New Gener. Comput.*, vol. 41, no. 3, pp. 635–668, 2023, doi: 10.1007/s00354-023-00224-3.
- [69] W. Wang, Y. Li, T. Zou, X. Wang, J. You, and Y. Luo, “A novel image classification approach via dense-MobileNet models,” *Mob. Inf. Syst.*, vol. 2020, 2020.
- [70] A. Narin, C. Kaya, and Z. Pamuk, “Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks,” *Pattern Anal. Appl.*, vol. 24, no. 3, pp. 1207–1220, 2021.
- [71] L. Kong and J. Cheng, “Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion,” *Biomed. Signal Process. Control*, vol. 77, p. 103772, 2022.
- [72] M. Chen, X. Shi, Y. Zhang, D. Wu, and M. Guizani, “Deep feature learning for medical image analysis with convolutional autoencoder neural network,” *IEEE Trans. Big Data*, vol. 7, no. 4, pp. 750–758, 2017.
- [73] Y. Xu, H.-K. Lam, and G. Jia, “MANet: A two-stage deep learning method for classification of COVID-19 from Chest X-ray images,” *Neurocomputing*, vol. 443, pp. 96–105, 2021.
- [74] O. F. Layode and M. Rahman, “A chest X-ray image retrieval system for COVID-19 detection using deep transfer learning and denoising auto encoder,” in *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2020, pp. 1635–1640.
- [75] M. Ketfi, M. Belahcene, and S. Bourennane, “Efficient Transfer Learning for COVID-19 Diagnosis using X-Ray and Computed Tomography Images,” in *2022 International Conference of Advanced Technology in Electronic and Electrical Engineering (ICATEEE)*, 2022, pp. 1–5.
- [76] D. S. Hong *et al.*, “KRAS G12C Inhibition with Sotorasib in Advanced Solid Tumors,” *N. Engl. J. Med.*, vol. 383, no. 13, pp. 1207–1217, Sep. 2020, doi: 10.1056/NEJM0A1917239.
- [77] A. Bounedjar *et al.*, “Incidence of lung cancer in males and females in Algeria: The lung cancer registry in Algeria (LuCaReAl),” *Cancer Epidemiol.*, vol. 69, p. 101799, 2020.
- [78] S. M. Salaken, A. Khosravi, A. Khatami, S. Nahavandi, and M. A. Hosen, “Lung cancer classification using deep learned features on low population dataset,” in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2017, pp. 1–5.

- [79] N. A. Baghdadi, A. Malki, H. M. Balaha, Y. AbdulAzeem, M. Badawy, and M. Elhosseini, "Classification of breast cancer using a manta-ray foraging optimized transfer learning framework," *PeerJ Comput. Sci.*, vol. 8, p. e1054, 2022.
- [80] S. K. Lakshmanaprabu, S. N. Mohanty, K. Shankar, N. Arunkumar, and G. Ramirez, "Optimal deep learning model for classification of lung cancer on CT images," *Futur. Gener. Comput. Syst.*, vol. 92, pp. 374–382, 2019.
- [81] M. Hammad, M. Bakrey, A. Bakhiet, R. Tadeusiewicz, A. A. Abd El-Latif, and P. Pławiak, "A novel end-to-end deep learning approach for cancer detection based on microscopic medical images," *Biocybern. Biomed. Eng.*, 2022.
- [82] J. Civit-Masot, A. Bañuls-Beaterio, M. Domínguez-Morales, M. Rivas-Pérez, L. Muñoz-Saavedra, and J. M. R. Corral, "Non-small cell lung cancer diagnosis aid with histopathological images using Explainable Deep Learning techniques," *Comput. Methods Programs Biomed.*, vol. 226, p. 107108, 2022.
- [83] A. K. Ajai and A. Anitha, "Clustering based lung lobe segmentation and optimization based lung cancer classification using CT images," *Biomed. Signal Process. Control*, vol. 78, p. 103986, 2022.
- [84] P. Xuan *et al.*, "Convolutional bi-directional learning and spatial enhanced attentions for lung tumor segmentation," *Comput. Methods Programs Biomed.*, p. 107147, 2022.
- [85] S. Dodia, B. Annappa, and P. A. Mahesh, "Recent advancements in deep learning based lung cancer detection: A systematic review," *Eng. Appl. Artif. Intell.*, vol. 116, p. 105490, 2022.
- [86] S. Tomassini, N. Falcionelli, P. Sernani, L. Burattini, and A. F. Dragoni, "Lung nodule diagnosis and cancer histology classification from computed tomography data by convolutional neural networks: A survey," *Comput. Biol. Med.*, p. 105691, 2022.
- [87] Q. Tian, Y. Wu, X. Ren, and N. Razmjoooy, "A new optimized sequential method for lung tumor diagnosis based on deep learning and converged search and rescue algorithm," *Biomed. Signal Process. Control*, vol. 68, p. 102761, 2021.
- [88] E. A. Siddiqui, V. Chaurasia, and M. Shandilya, "Detection and classification of lung cancer computed tomography images using a novel improved deep belief network with Gabor filters," *Chemom. Intell. Lab. Syst.*, vol. 235, p. 104763, 2023.
- [89] D. El Hamdi, I. Elouedi, and I. Slim, "Computer-Aided Classification of Cell Lung Cancer Via PET/CT Images Using Convolutional Neural Network," *Int. J. Image Graph.*, p. 2450040, 2023.
- [90] G. Murtaza *et al.*, "Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges," *Artif. Intell. Rev.*, vol. 53, no. 3, pp. 1655–1720, 2020.
- [91] M. Nawaz, A. A. Sewissy, and T. H. A. Soliman, "Multi-class breast cancer classification using deep learning convolutional neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 6, pp. 316–332, 2018.
- [92] S. Khan, N. Islam, Z. Jan, I. U. Din, and J. J. P. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," *Pattern Recognit. Lett.*, vol. 125, pp. 1–6, 2019.
- [93] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, and S. Li, "Breast cancer multi-classification from histopathological images with structured deep learning model," *Sci. Rep.*, vol. 7, no. 1, pp. 1–10, 2017.
- [94] D. A. Omondiagbe, S. Veeramani, and A. S. Sidhu, "Machine learning classification techniques for breast cancer diagnosis," in *IOP Conference Series: Materials Science and Engineering*, 2019, vol. 495, no. 1, p. 12033.
- [95] A. Boulenger *et al.*, "Deep learning-based system for automatic prediction of triple-negative breast cancer from ultrasound images," *Med. Biol. Eng. Comput.*, vol. 61, no. 2, pp. 567–578, 2023.
- [96] M. D. Ali *et al.*, "Breast Cancer Classification through Meta-Learning Ensemble Technique Using Convolution Neural Networks," *Diagnostics*, vol. 13, no. 13, p. 2242, 2023.
- [97] N. Yamanakkanavar, J. Y. Choi, and B. Lee, "MRI segmentation and classification of human brain using deep learning for diagnosis of Alzheimer's disease: a survey," *Sensors*, vol. 20, no. 11, p. 3243, 2020.

- [98] G. S. Babu, S. N. T. Rao, and R. R. Rao, "Automated assessment for Alzheimer's disease diagnosis from MRI images: Meta-heuristic assisted deep learning model," *Int. J. Imaging Syst. Technol.*, vol. 32, no. 2, pp. 544–563, 2022.
- [99] J. Islam and Y. Zhang, "A novel deep learning based multi-class classification method for Alzheimer's disease detection using brain MRI data," in *Brain Informatics: International Conference, BI 2017, Beijing, China, November 16-18, 2017, Proceedings*, 2017, pp. 213–222.
- [100] J. Wen *et al.*, "Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation," *Med. Image Anal.*, vol. 63, p. 101694, 2020.
- [101] F. Zhang, Z. Li, B. Zhang, H. Du, B. Wang, and X. Zhang, "Multi-modal deep learning model for auxiliary diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 361, pp. 185–195, 2019.
- [102] W. N. Ismail, F. R. PP, and M. A. S. Ali, "A Meta-Heuristic Multi-Objective Optimization Method for Alzheimer's Disease Detection Based on Multi-Modal Data," *Mathematics*, vol. 11, no. 4, p. 957, 2023.
- [103] M. Leela, K. Helenprabha, and L. Sharmila, "Prediction and classification of Alzheimer Disease categories using Integrated Deep Transfer Learning Approach," *Meas. Sensors*, p. 100749, 2023.
- [104] A. Zafar *et al.*, "A Comparison of Pooling Methods for Convolutional Neural Networks," *Appl. Sci.*, vol. 12, no. 17, p. 8643, 2022.
- [105] D. Riquelme and M. A. Akhloufi, "Deep learning for lung cancer nodules detection and classification in CT scans," *Ai*, vol. 1, no. 1, pp. 28–67, 2020.
- [106] W. Zhao, Z. Zhang, and L. Wang, "Manta ray foraging optimization: An effective bio-inspired optimizer for engineering applications," *Eng. Appl. Artif. Intell.*, vol. 87, p. 103300, 2020.
- [107] M. A. Khodeir, J. I. Ababneh, and B. S. Alamoush, "Manta Ray Foraging Optimization (MRFO)-Based Energy-Efficient Cluster Head Selection Algorithm for Wireless Sensor Networks," *J. Electr. Comput. Eng.*, vol. 2022, 2022.
- [108] "Chest CT-Scan images Dataset | Kaggle." <https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images> (accessed Oct. 08, 2022).
- [109] "Breast Ultrasound Images Dataset | Kaggle." <https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset> (accessed Apr. 24, 2023).
- [110] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, "Dataset of breast ultrasound images. Data Brief 28, 104863 (2020)." 2019.
- [111] "Breast Cancer Dataset | Kaggle." <https://www.kaggle.com/datasets/anasmasy/breast-cancer-dataset> (accessed Apr. 24, 2023).
- [112] "Alzheimer's Dataset (4 class of Images) | Kaggle." <https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images> (accessed Apr. 24, 2023).
- [113] A. A. Mukhlif, B. Al-Khateeb, and M. A. Mohammed, "Incorporating a Novel Dual Transfer Learning Approach for Medical Images," *Sensors*, vol. 23, no. 2, p. 570, 2023.
- [114] S. R. S. Chakravarthy, N. Bharanidharan, and H. Rajaguru, "Deep Learning-based Metaheuristic Weighted K-Nearest Neighbor Algorithm for the Severity Classification of Breast Cancer," *IRBM*, p. 100749, 2023.
- [115] F. Atban, E. Ekinici, and Z. Garip, "Traditional machine learning algorithms for breast cancer image classification with optimized deep features," *Biomed. Signal Process. Control*, vol. 81, p. 104534, 2023.
- [116] Y. Xu, Y. Wang, and N. Razmjoooy, "Lung cancer diagnosis in CT images based on Alexnet optimized by modified Bowerbird optimization algorithm," *Biomed. Signal Process. Control*, vol. 77, p. 103791, 2022.
- [117] A. Elnakib, H. M. Amer, and F. E. Z. Abou-Chadi, "Early lung cancer detection using deep learning optimization," 2020.
- [118] W. M. Salama, A. Shokry, and M. H. Aly, "A generalized framework for lung cancer classification based on deep generative models," *Multimed. Tools Appl.*, vol. 81, no. 23, pp. 32705–32722, 2022.

- [119] N. Mahendran, P. M. D. R. Vincent, K. Srinivasan, and C.-Y. Chang, “Improving the classification of alzheimer’s disease using hybrid gene selection pipeline and deep learning,” *Front. Genet.*, vol. 12, p. 784814, 2021.
- [120] X.-X. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang, “U-Net-Based medical image segmentation,” *J. Healthc. Eng.*, vol. 2022, 2022.
- [121] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234–241.
- [122] U. T. Salim, F. Ali, and S. A. Dawwd, “U-Net Convolutional Networks Performance Based on Software-Hardware Cooperation Parameters: A Review,” *Int. J. Comput. Digit. Syst.*, 2021.
- [123] Y. Bahendwar and D. D. Thankachan, “A Review on Medical Image Segmentation Techniques,” *Available SSRN 4607754*, 2023.
- [124] Z. Li *et al.*, “Spatiotemporal trends of the global burden of melanoma in 204 countries and territories from 1990 to 2019: Results from the 2019 global burden of disease study,” *Neoplasia*, vol. 24, no. 1, pp. 12–21, 2022.
- [125] J. Moorthy and U. D. Gandhi, “A Survey on Medical Image Segmentation Based on Deep Learning Techniques,” *Big Data Cogn. Comput.*, vol. 6, no. 4, p. 117, 2022.
- [126] X. Liu, L. Song, S. Liu, and Y. Zhang, “A review of deep-learning-based medical image segmentation methods,” *Sustainability*, vol. 13, no. 3, p. 1224, 2021.
- [127] Y. Sun, J. Cong, K. Zhang, M. Jian, and B. Wei, “Unsupervised medical image feature learning by using de-melting reduction auto-encoder,” *Neurocomputing*, vol. 523, pp. 145–156, 2023.
- [128] M. Krithika alias AnbuDevi and K. Suganthi, “Review of Semantic Segmentation of Medical Images Using Modified Architectures of UNET,” *Diagnostics*, vol. 12, no. 12, p. 3064, 2022.
- [129] S. Zhou *et al.*, “Medical image segmentation using deep semantic-based methods: A review of techniques, applications and emerging trends,” *Med. Image Anal.*, vol. 82, 2022.
- [130] F.-P. An and J. Liu, “Medical image segmentation algorithm based on optimized convolutional neural network-adaptive dropout depth calculation,” *Complexity*, vol. 2020, 2020.
- [131] M. I. Razzak, S. Naz, and A. Zaib, “Deep learning for medical image processing: Overview, challenges and the future,” *Classif. BioApps*, pp. 323–350, 2018.
- [132] H. Liu, D. Hu, H. Li, and I. Oguz, “Medical image segmentation using deep learning,” *Mach. Learn. Brain Disord.*, pp. 391–434, 2023.
- [133] J. Dolz, L. Massotier, and M. Vermandel, “Segmentation algorithms of subcortical brain structures on MRI for radiotherapy and radiosurgery: a survey,” *Irbm*, vol. 36, no. 4, pp. 200–212, 2015.
- [134] K. He *et al.*, “Synergistic learning of lung lobe segmentation and hierarchical multi-instance classification for automated severity assessment of COVID-19 in CT images,” *Pattern Recognit.*, vol. 113, p. 107828, 2021.
- [135] C. Fan, Z. Zeng, L. Xiao, and X. Qu, “GFNet: Automatic segmentation of COVID-19 lung infection regions using CT images based on boundary features,” *Pattern Recognit.*, vol. 132, p. 108963, 2022.
- [136] S. Chaganti *et al.*, “Automated quantification of CT patterns associated with COVID-19 from chest CT,” *Radiol. Artif. Intell.*, vol. 2, no. 4, p. e200048, 2020.
- [137] L. Jingxin *et al.*, “COVID-19 lesion detection and segmentation—A deep learning method,” *Methods*, vol. 202, pp. 62–69, 2022.
- [138] D.-P. Fan *et al.*, “Inf-net: Automatic covid-19 lung infection segmentation from ct images,” *IEEE Trans. Med. Imaging*, vol. 39, no. 8, pp. 2626–2637, 2020.
- [139] N. Saeedzadeh, S. Minaee, R. Kafieh, S. Yazdani, and M. Sonka, “COVID TV-Unet: Segmenting COVID-19 chest CT images using connectivity imposed Unet,” *Comput. Methods Programs Biomed.*

Updat., vol. 1, p. 100007, 2021.

- [140] A. Amyar, R. Modzelewski, H. Li, and S. Ruan, "Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation," *Comput. Biol. Med.*, vol. 126, p. 104037, 2020.
- [141] X. Chen, L. Yao, and Y. Zhang, "Residual attention u-net for automated multi-class segmentation of covid-19 chest ct images," *arXiv Prepr. arXiv2004.05645*, 2020.
- [142] Q. Yan *et al.*, "COVID-19 chest CT image segmentation--a deep convolutional neural network solution," *arXiv Prepr. arXiv2004.10987*, 2020.
- [143] O. Elharrouss, N. Subramanian, and S. Al-Maadeed, "An encoder-decoder-based method for COVID-19 lung infection segmentation," *arXiv Prepr. arXiv2007.00861*, 2020.
- [144] D. Müller, I. S. Rey, and F. Kramer, "Automated chest ct image segmentation of covid-19 lung infection based on 3d u-net," *arXiv Prepr. arXiv2007.04774*, 2020.
- [145] M. Abdel-Basset, V. Chang, H. Hawash, R. K. Chakraborty, and M. Ryan, "FSS-2019-nCov: A deep learning architecture for semi-supervised few-shot segmentation of COVID-19 infection," *Knowledge-Based Syst.*, vol. 212, p. 106647, 2021.
- [146] F. Shan *et al.*, "Lung infection quantification of COVID-19 in CT images with deep learning," *arXiv Prepr. arXiv2003.04655*, 2020.
- [147] S. Chaganti *et al.*, "Quantification of tomographic patterns associated with COVID-19 from chest CT," *ArXiv*, 2020.
- [148] A. Negi, A. N. J. Raj, R. Nersisson, Z. Zhuang, and M. Murugappan, "RDA-UNET-WGAN: an accurate breast ultrasound lesion segmentation using wasserstein generative adversarial networks," *Arab. J. Sci. Eng.*, vol. 45, no. 8, pp. 6399–6410, 2020.
- [149] S. Saha, S. Dutta, B. Goswami, and D. Nandi, "ADU-Net: An Attention Dense U-Net based deep supervised DNN for automated lesion segmentation of COVID-19 from chest CT images," *Biomed. Signal Process. Control*, vol. 85, p. 104974, 2023.
- [150] Z. Mirikharaji *et al.*, "A survey on deep learning for skin lesion segmentation," *Med. Image Anal.*, p. 102863, 2023.
- [151] A. Adegun and S. Viriri, "Deep learning techniques for skin lesion analysis and melanoma cancer detection: a survey of state-of-the-art," *Artif. Intell. Rev.*, vol. 54, no. 2, pp. 811–841, 2021.
- [152] K. A. Kumar and C. Vanmathi, "Optimization driven model and segmentation network for skin cancer detection," *Comput. Electr. Eng.*, vol. 103, p. 108359, 2022.
- [153] M. Dash, N. D. Londhe, S. Ghosh, A. Semwal, and R. S. Sonawane, "PsLSNet: Automated psoriasis skin lesion segmentation using modified U-Net-based fully convolutional network," *Biomed. Signal Process. Control*, vol. 52, pp. 226–237, 2019.
- [154] M. Badar, M. Haris, and A. Fatima, "Application of deep learning for retinal image analysis: A review," *Comput. Sci. Rev.*, vol. 35, p. 100203, 2020.
- [155] Y. Zhang and A. Chung, "Deep supervision with additional labels for retinal vessel segmentation task," in *International conference on medical image computing and computer-assisted intervention*, 2018, pp. 83–91.
- [156] K. Hu *et al.*, "Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function," *Neurocomputing*, vol. 309, pp. 179–191, 2018.
- [157] T. Iqbal and H. Ali, "Generative adversarial network for medical images (MI-GAN)," *J. Med. Syst.*, vol. 42, no. 11, pp. 1–11, 2018.
- [158] J. J. Liang, A. K. Qin, P. N. Suganthan, and S. Baskar, "Comprehensive learning particle swarm optimizer for global optimization of multimodal functions," *IEEE Trans. Evol. Comput.*, vol. 10, no. 3, pp. 281–295, 2006.
- [159] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *MHS'95. Proceedings of*

the sixth international symposium on micro machine and human science, 1995, pp. 39–43.

- [160] P. Erdogmus, “Introductory chapter: swarm intelligence and particle swarm optimization,” *Part. Swarm Optim. with Appl.*, vol. 10, 2018.
- [161] “adnan-saood/COVID19-DL.” <https://github.com/adnan-saood/COVID19-DL> (accessed Aug. 31, 2022).
- [162] “COVID-19 CT Lung and Infection Segmentation Dataset | Zenodo.” <https://zenodo.org/record/3757476#.Yw93qXbMJpb> (accessed Aug. 31, 2022).
- [163] “ISIC Challenge.” <https://challenge.isic-archive.com/landing/2017/42/> (accessed Sep. 08, 2022).
- [164] “harsha-20/Synthetic-Medical-Images: Generating Medical Images using GANs.” <https://github.com/harsha-20/Synthetic-Medical-Images> (accessed Sep. 05, 2022).
- [165] J. T. Guibas, T. S. Virdi, P. S. Li, S. Henry, and M. Gunn, “Synthetic Medical Images from Dual Generative Adversarial Networks.” [Online]. Available: <https://synthmed.github.io>
- [166] S. Elghamrawy and A. E. Hassanien, “Diagnosis and prediction model for COVID-19 patient’s response to treatment based on convolutional neural networks and whale optimization algorithm using CT images,” *MedRxiv*, 2020.
- [167] M. Agarwal, S. K. Gupta, and K. K. Biswas, “Development of a compressed FCN architecture for semantic segmentation using Particle Swarm Optimization,” *Neural Comput. Appl.*, vol. 35, no. 16, pp. 11833–11846, 2023.
- [168] M. Attia, M. Hossny, S. Nahavandi, and A. Yazdabadi, “Skin melanoma segmentation using recurrent and convolutional neural networks,” in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017, pp. 292–296.
- [169] S. Vesal, N. Ravikumar, and A. Maier, “SkinNet: A deep learning framework for skin lesion segmentation,” in *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*, 2018, pp. 1–3.
- [170] R. A. Khaja, S. Z. A. Jalil, and S. Usman, “Improved Tuna Swarm-based U-EfficientNet: Skin Lesion Image Segmentation by Improved Tuna Swarm Optimization,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 5, 2023.
- [171] Y. Lv, H. Ma, J. Li, and S. Liu, “Attention guided U-Net with atrous convolution for accurate retinal vessels segmentation,” *IEEE Access*, vol. 8, pp. 32826–32839, 2020.
- [172] Y. Liu, J. Shen, L. Yang, G. Bian, and H. Yu, “ResDO-UNet: A deep residual network for accurate retinal vessel segmentation from fundus images,” *Biomed. Signal Process. Control*, vol. 79, p. 104087, 2023.
- [173] M. Yan, J. Zhou, C. Luo, T. Xu, and X. Xing, “Multiscale Joint Optimization Strategy for Retinal Vascular Segmentation,” *Sensors*, vol. 22, no. 3, p. 1258, 2022.
- [174] V. Popat, M. Mahdinejad, O. S. D. Cedeno, E. Naredo, and C. Ryan, “GA-based U-Net architecture optimization applied to retina blood vessel segmentation,” 2020.