

**BRAIN TUMOR IDENTIFICATION USING LION SWARM
OPTIMIZATION BASED ENSEMBLE DEEP LEARNING
ALGORITHMS**

*The thesis submitted to the Bharathiar University in partial fulfilment of the
requirements for the degree of*

DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE

Submitted By

R.KALAIVANI

Under the Guidance and Supervision of

Dr. ANNA SARO VIJENDRAN, MCA., M.Phil., Ph.D.

Former Dean, Research

Sri Ramakrishna College of Arts & Science (Autonomous)

(Formerly S.N.R. Sons College), Coimbatore – 06



DEPARTMENT OF COMPUTER SCIENCE

SRI RAMAKRISHNA COLLEGE OF ARTS & SCIENCE (AUTONOMOUS)

(Formerly S.N.R. Sons College)

(Affiliated to Bharathiar University)

(Re-Accredited By NAAC with “A+” Grade)

(An ISO 9001:2015 Certified Institution)

COIMBATORE – 641 006

JUNE 2024

Certificate

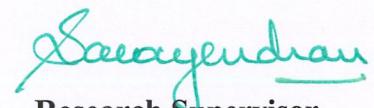
CERTIFICATE

This is to certify that the thesis, entitled "**BRAIN TUMOR IDENTIFICATION USING LION SWARM OPTIMIZATION BASED ENSEMBLE DEEP LEARNING ALGORITHMS**" submitted to the Bharathiar University, in Partial fulfilment of the requirements for the award of the Degree of **DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE** is a record of original research work done by **Ms. R.KALAIYANI** during the period **2017-2024** of her research in the Department of Computer Science at Sri Ramakrishna College of Arts & Science, under my supervision and guidance and the thesis has not formed the basis for the award of any Degree/ Diploma/ Associateship/ Fellowship or other similar title of any candidate of any University.

Place: Coimbatore

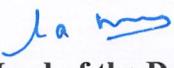
Date: 21/06/2024



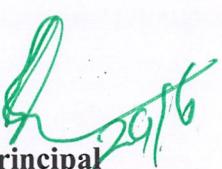

Research Supervisor

Dr. ANNA SARO VIJENDRAN MCA.,M.Phil.,Ph.D.,
Former Dean
Research
SRI RAMAKRISHNA
COLLEGE OF ARTS AND SCIENCE (Autonomous)
Avinashi Road, Nava India,
Coimbatore - 641 006.

Countersigned


Head of the Department

Dr. G. MARIA PRISCILLA
Associate Professor and Head
Department of Computer Science
Sri Ramakrishna College of Arts & Science
Coimbatore - 641 006.


Principal
PRINCIPAL
SRI RAMAKRISHNA
COLLEGE OF ARTS AND SCIENCE (Autonomous)
Avinashi Road, Nava India,
Coimbatore - 641 006.

Declaration

DECLARATION

I R.KALAIVANI hereby declare that the thesis entitled "**BRAIN TUMOR IDENTIFICATION USING LION SWARM OPTIMIZATION BASED ENSEMBLE DEEP LEARNING ALGORITHMS**" submitted to the Bharathiar University in partial fulfilment of the requirements for the award of the degree of **DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE** is a record of original research work done by me during the period **2017-2024** under the supervision and guidance of **Dr. ANNA SARO VIJENDRAN, MCA., M.Phil., Ph.D.**, Former Dean, Research, Sri Ramakrishna College of Arts & Science (Autonomous), Coimbatore and that it has not formed the basis for the award of any degree/ Diploma/Associate ship/Fellowship or other similar title to any candidate of any University.

Place: *Coimbatore*

Date: *21/06/2024*

R.kalaiyvani
Signature of the candidate

*Certificate of Genuineness of the
Publication*

CERTIFICATE OF GENUINENESS OF THE PUBLICATION

This is to certify that the Ph.D. candidate **KALAIVANI.R** working under my supervision has published the following research articles in the referred journals named:

1. "An Extensive Study on the Multimodal Medical Image Fusion Techniques", International Journal of Computer Science Trends and Technology, Vol. No. 6, Issue No. 4, Jul - Aug 2018, ISSN. 2347-8578 (UGC Approved Journal).
2. "M-UNet for Segmentation of Brain Images", International Journal of Life Science and Pharma Research, Vol. No. 12, Issue No. 03, MAY 2022, ISSN. 2250-0480 (Web of Science).
3. "Optimal segmentation and fusion of multi-modal brain images using clustering based deep learning algorithm", Measurement: Sensors, Vol. No. 27, June 2023, ISSN. 100691. (ELSEVIER).
4. "Ensemble Deep Learning Algorithm for Multi View Image Fusion", International Journal of Intelligent Systems and Applications in Engineering, Vol. No. 12, No. 8S (2024), December 2023, ISSN. 2147-6799. (Scopus).
5. "Comparative Study for Image Fusion Using Various Deep Learning Algorithms", Mapana – Journal of Sciences, Vol. No. 22, Issue No. 2, 313-333, December 2023, ISSN: 0975-3303. (UGC Approved Journal).

The Contents of the publication incorporate part of the results presented in her synopsis.




Signature of the Guide

Dr. ANNA SARO VIJENDRAN MCA.,M.Phil.,Ph.D.,
Former Dean
Research
SRI RAMAKRISHNA
COLLEGE OF ARTS AND SCIENCE (Autonomous)
Avinashi Road, Nava India,
Coimbatore - 641 006.


Head of the Department

Dr. G. MARIA PRISCILLA
Associate Professor and Head
Department of Computer Science
Sri Ramakrishna College of Arts & Science
Coimbatore - 641 006.


Principal
PRINCIPAL
SRI RAMAKRISHNA
COLLEGE OF ARTS AND SCIENCE (Autonomous)
Avinashi Road, Nava India,
Coimbatore - 641 006.

Certificate of Plagiarism Check



பாரதியர் பல்கலைக்கழகம்
BHARATHIAR UNIVERSITY
COIMBATORE - 641 046, TAMILNADU, INDIA

| State University | Accredited With A++ Grade - 3.63 CGPA by NAAC | 15th Rank among Indian Universities by MoE-NIRF |

CERTIFICATE OF PLAGIARISM CHECK

1	Name of the Research Scholar	R. KALATIVANT
2	Course of study	M.Phil., / Ph.D.,
3	Title of the Thesis / Dissertation	BRAIN TUMOR IDENTIFICATION USING LION SWARM OPTIMIZATION BASED ENSEMBLE DEEP LEARNING ALGORITHMS
4	Name of the Supervisor	DR. ANNA SARO VIJENDRAN, MCA., M.Phil., Ph.D.
5	Department / Institution/ Research Centre	DEPARTMENT OF COMPUTER SCIENCE
6	% of Similarity of content Identified	03 %
7	Acceptable Maximum Limit	10 %
8	Software Use	Drillbit
9	Date of verification	21/06/2024

Report on plagiarism check, items with % of similarity is attached

Signature of the Supervisor

(Seal)

Dr. ANNA SARO VIJENDRAN MCA., M.Phil., Ph.D.
Dean

Research

SRI RAMAKRISHNA
COLLEGE OF ARTS AND SCIENCE (Autonomous)
Avinashi Road, Nava India,
Coimbatore - 641 006.

Signature of the Researcher

(Seal)

Head of the Department

(Seal)

Dr. G. MARIA PRISCILLA

Associate Professor and Head
Department of Computer Science
Sri Ramakrishna College of Arts & Science
Coimbatore - 641 006.

Director i/c

Center for Research & Evaluation (BU)

(Seal)

University Librarian (BU)
University Librarian
(Seal)
Arignar Anna Central Library
Bharathiar University
Coimbatore - 641 046.



Bharathiar University

Certificate of Plagiarism Check for Thesis

Author Name	Kalaivani. R
Course of Study	Ph.D
Name of Guide	Dr. Anna Saro Vijendran
Department	Computer Science
Acceptable Maximum Limit	10%
Submitted By	buaacl.urbkund@gmail.com
Paper Title	BRAIN TUMOR IDENTIFICATION USING LION SWARM OPTIMIZATION BASED ENSEMBLE DEEP LEARNING ALGORITHMS
Similarity	3%
Paper ID	2027048
Submission Date	2024-06-21 11:21:26



University Librarian
University Librarian
Arignar Anna Central Library
Bharathiar University
Coimbatore - 641 046.

* This report has been generated by DrillBit Anti-Plagiarism Software

Acknowledgement

ACKNOWLEDGEMENT

“No one who achieves success, does so without acknowledging the help of others. The wise & confident acknowledge this help with gratitude”.

I thank the Almighty for his mercy and blessings showered upon me to complete the project.

At the outset, I express my sincere gratitude to Managing Trustee **Sri. D. Lakshminarayanaswamy**, SNR Sons Charitable Trust, Coimbatore for having given me an opportunity to pursue my research in their esteemed institution.

I am deeply grateful to **Thiru. R. Sundar**, Joint Managing Trustee, SNR Sons Charitable Trust, Coimbatore for having given me an opportunity to pursue my research in their esteemed institution.

I extend my sincere thanks to **Dr. B L. Shivakumar**, Principal and Secretary, Sri Ramakrishna College of Arts & Science for his encouragement in the successful completion of my research study.

My sincere gratitude to my guide **Dr. ANNA SARO VIJENDRAN, MCA., M.Phil., Ph.D.**, Former Dean, Research, Sri Ramakrishna College of Arts & Science, Coimbatore for her support and encouragement throughout the project. Her inspiring guidance and rich research experience enabled me to complete my research successfully. It was indeed a great privilege for me to work under her supervision.

I express my sincere thanks to my Doctoral Committee Member **Dr. P. Krishnapriya**, Dean Academics-Science, K G College of Arts and Science, Coimbatore for her instigation and guidance and consistent support.

I wish to express my gratitude to **Dr. G. Maria Priscilla**, Associate Professor & Head of the Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore for having provided timely assistance and facilities at all stages in this research.

I wish to express my gratitude to the Faculty Members, Research Scholars and non-teaching staff members of Sri Ramakrishna College of Arts & Science, Coimbatore, for having provided timely assistance and facilities at all stages in this research.

I express my thanks to all the members of Bharathiar University, Center for Research Evaluation and The Librarian, Bharathiar University, Coimbatore.

My heartfelt thanks to my Daughter **R.K Nivashini**, my parents, my brothers and his family and, other family members and friends for their support and concern in innumerable ways in the completion of the project.

R KALAIVANI

Table of Contents

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	LIST OF TABLES	x
	LIST OF FIGURES	xi
	LIST OF SYMBOLS	xiii
	LIST OF ABBREVIATIONS	xv
	ABSTRACT	xviii
1.	INTRODUCTION	
	1.1. General Background	1
	1.1.1. Image Fusion	2
	1.1.2. Medical Image Fusion	2
	1.2. Brain Tumor Overview	3
	1.2.1. Brain Tumor Development	5
	1.3. Significance of Image Fusion	7
	1.4. Image Pre-Processing Methods	9
	1.4.1. Noise reduction	9
	1.4.1.1. AMF	9
	1.4.2. Segmentation	10
	1.4.2.1. Region based segmentation	11
	1.4.2.2. Modified UNet (MUNet) algorithm	12
	1.4.3. Feature extraction	13
	1.4.3.1. Particle Swarm Optimization (PSO)	14
	1.4.3.2. Firefly algorithm	14
	1.4.3.3. Lion Swarm Optimization (LSO) algorithm	16
	1.4.4. Image fusion methods	17
	1.4.4.1. Convolutional Neural Network (CNN)	20
	1.4.4.2. Generative Adversarial Networks (GAN)	21

CHAPTER NO.	TITLE	PAGE NO.
	1.5. Problem Specification	23
	1.6. Objective of the Research	23
	1.7. Contribution of the Research	23
	1.7.1. Modified-UNet for Segmentation of Brain Images	25
	1.7.2. Multi-modal image fusion using deep learning algorithm	25
	1.7.3. Optimal segmentation and fusion of multi modal brain images using clustering based deep learning algorithm	26
	1.7.4. Ensemble deep learning algorithm for multi view image fusion	27
	1.7.5. Organization of the Thesis	28
2.	LITERATURE SURVEY	
	2.1. Introduction	30
	2.2. Review on Noise Reduction and Multimodal Image Fusion Methods	30
	2.3. Reviews on Segmentation Methods Over MRI Images	34
	2.4. Review on Feature Extraction and Optimization Techniques	38
	2.5. Reviews on Image Fusion Using Deep Learning and Ensemble Deep Learning Algorithms Over MRI Images	41
	2.6. Motivation for The Proposed Research and Key Difference Between Proposed Work and The Existing Works	48
	2.7. Summary	49
3.	MODIFIED-UNET FOR SEGMENTATION OF BRAIN IMAGES	
	3.1. Introduction	50
	3.2. Background Study	52
	3.2.1. U-net (U-Net)	52
	3.2.2. Drawbacks of UNet	55
	3.2.3. Solution	56
	3.3. Proposed Methodology	56

CHAPTER NO.	TITLE	PAGE NO.
	3.3.1.Modified Unet (MUnet)	57
	3.4. Result and Discussion	58
	3.5. Performance Metrics	59
	3.6. Summary	65
4.	MULTI-MODAL IMAGE FUSION USING DEEP LEARNING ALGORITHM	
	4.1. Introduction	67
	4.1.1. Deep Learning (DL)	67
	4.2. Proposed Methodology	69
	4.2.1.Feature Extraction Using CNN	71
	4.2.2.Image fusion of multi –model image	76
	4.2.3.Image Classification	78
	4.3. Results and Discussion	79
	4.4. Performance Metrics	79
	4.5. Summary	86
5.	OPTIMAL SEGMENTATION AND FUSION OF MULTI MODAL BRAIN IMAGES USING CLUSTERING BASED DEEP LEARNING ALGORITHM	
	5.1. Introduction	88
	5.2. Proposed Methodology	88
	5.2.1.Input image dataset	89
	5.2.2. Noise removal using Adaptive Median Filtering (AMF)	91
	5.2.3. Brain tumor segmentation via Region Growing based K-Means Clustering (RGKMC) algorithm	92
	5.2.4. Feature extraction using Adaptive FireFly Optimization based Convolutional Neural Network (AFFOCNN)	95
	5.2.5. Image fusion of multi modal image	104

CHAPTER NO.	TITLE	PAGE NO.
	5.3. Experimental Result	104
	5.4. Summary	113
6.	ENSEMBLE DEEP LEARNING ALGORITHM FOR MULTI VIEW IMAGE FUSION	
	6.1. Introduction	114
	6.2. Proposed Methodology	117
	6.2.1. Input image dataset	119
	6.2.2. Noise removal using Adaptive Median Filtering (AMF)	119
	6.2.3. Brain tumor segmentation via Region Growing based K-Means Clustering (RGKMC) algorithm	120
	6.2.4. Feature extraction using Lion Swarm Optimization based Convolutional Neural Network (LSOCNN)	120
	6.2.5. Image fusion of multi modal image	128
	6.2.5.1. DCNN algorithm	128
	6.2.5.2. Differentiable Fusion with Mutual Information-Network (DFMI-Net)	130
	6.2.5.3. Tissue-Aware conditional Generative Adversarial Network (TA-cGAN)	131
	6.2.5.4. Ensemble Deep Learning (EDL)	132
	6.3. Experimental Result	136
	6.4. Summary	143
7.	CONCLUSION AND FUTURE WORK	
	7.1. Conclusion	144
	7.2. Future Work	146
	REFERENCES	147
	PUBLICATIONS	

List of Tables

LIST OF TABLES

TABLE NO.	TITLE	PAGE NO.
4.1	Performance Analysis	82
5.1	Comparison values for given dataset	106

List of Figures

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE NO.
1.1.	Example of brain tumor	6
1.2.	Examples of multi modal brain medical image fusion	8
1.3.	a) MRI brain input image b) filtered image	9
1.4.	Example of MRI brain image segmentation	10
1.5.	Illustration of modified U-Net model	12
1.6.	Example of MRI image feature extraction	13
1.7.	General architecture diagram of the five proposed methodologies	24
3.1.	U -Net Architecture	52
3.2.	Overall block diagram of the proposed system	57
3.3.	Input image	60
3.4.	Segmentation image	61
3.5.	Output of the MUNet	61
3.6.	Accuracy	62
3.7.	Dice Co-efficient	63
3.8.	Jaccard Co-efficient	64
3.9.	Time Period	65
4.1.	Overall block diagram of the proposed system	70
4.2.	Image Fusion	81
4.3.	Accuracy	83
4.4.	Precision	84
4.5.	Recall	85
4.6.	Time Period	86
5.1.	Overall architecture diagram of the proposed AFFOCNN system	90

FIGURE NO.	TITLE	PAGE NO.
5.2.	Example of KMC algorithm	93
5.3.	Basic mechanism of the firefly algorithm	98
5.4.	a) Input image b) AMF image	105
5.5.	a) Extracted image b) Output images	105
5.6.	Accuracy	107
5.7.	Precision	108
5.8.	Recall	109
5.9.	F-measure	110
5.10.	MSE	111
5.11.	Execution time	112
6.1.	Overall block diagram of the proposed system	118
6.2.	Nature of LSO	122
6.3.	Network architecture of the Generator	133
6.4.	Network architecture of the discriminator	133
6.5.	EDL algorithm	135
6.6.	Accuracy	137
6.7.	Precision	138
6.8.	Recall	139
6.9.	F-measure	140
6.10.	MSE	141
6.11.	Execution time	142

List of Symbols

LIST OF SYMBOLS

SYMBOLS	DESCRIPTIONS
FMx	connection behind the residual path
FMy	previous to residual track
W	window matching function
M	optimal solution of W
p	pixel
q	neighbor pixel of p
N_P	pixels adjacent
I_{min}	minimum pixel value
I_{max}	maximum pixel value
W	current window size
W_{max}	maximum window size
I_{med}	median of the assigned window
x_i and y_i	two points in Euclidean n-space
$I(r)$	intensity at attractiveness r^2
I_0	absorption coefficient of the medium
β_0	attractiveness
$x_{i,k}$	k th factor of the spatial match
d	amount of dimensions
α	distance deviation degree
α_{max} and α_{min}	maximum and minimum features
x_{worst}	position of the worst individual

SYMBOLS	DESCRIPTIONS
I_d	number of dropped image
m_t	higher accuracy
I_p	pixel in image i.
P_{init}^i	initial image
e_E	execution time
e_M	maximum allowable delay
x_i and x_j	distance between two firefly nodes
n	<i>number of data points</i>
Y_i	<i>observed values</i>
\widehat{Y}_i	<i>predicted values</i>
$x_{N_{var}}$	generic image pixels
x_n	input images
$\widehat{x_n}$	corresponding reconstructed images
xf	fused image

List of Abbreviations

LIST OF ABBREVIATIONS

AFFOCNN	Adaptive Fire Fly Optimization based Convolutional Neural Network
AMF	Adaptive Median Filtering
ANN	Artificial Neural Network
BS	Base
BSE	Brain Surface Extraction
CAD	Computer-Aided Diagnosis
CBS	Cuckoo Based Search
CLFAHE	Contrast Limited Fuzzy Adaptive Histogram Equalization
CS	Coarse-Structure
CSF	Cerebrospinal Fluid
CT	Computed Tomography
DCNN	Deep Convolutional Neural Network
DDcGAN	Dual-Discriminator conditional Generative Adversarial Network
DFMI-Net	Differentiable Fusion with Mutual Information-Network
DIP	Digital Image Processing
DTI	Diffusion Tensor Imaging
DWI	Diffusion-Weighted Imaging
EDL	Ensemble Deep Learning
EEG	Electro Encephalography
ET	Echo Time
FKSRG	Fuzzy Knowledge-Based Seeded Region Growing
fMRI	Functional Magnetic Resonance Imaging

FS	Fine-Structure
GA	Genetic Algorithm
GAN	Generative Adversarial Networks
GBM	Glioblastoma multiforme
GMSF	Gradient Minimization Smoothing Filter
HGG	High-Grade Glioma
HH	high–high
HL	high–low
HWF	Hierarchical Wavelet Fusion
IGWT	Improved Gabor Wavelet Transform
ILSOA	Improved Lion Swarm Optimization Algorithm
ISWT	Inverse Stationary Wavelet Transform
ITRS	Information Theoretic Rough Sets
LBP	Local Binary Patterns
LGG	low-grade glioma
LH	low–high
LL	low–low
LSOCNN	Lion Swarm Optimization based Convolutional Neural Network
MEG	Magneto Encephalo Gram
MLEPF	Multi-Level Edge-Preserving Filtering
MPSO+MFCL	Modified Particle Swarm Optimization and Modified Fully Connected Layer
MR	Magnetic Resonance
MRI	Magnetic Resonance Image
MSE	Mean Square Error

MVFF	Multi-View Feature Fusion
NCIS	National Cancer Institute Statistics
OFFA	Oppositional fruit fly algorithm
OS	overall survival
PCA	Principal Component Analysis
PCNN	Pulse-Coupled Neural Network
PET	Positron Emission Tomography
PSNR	Peak Signal-to-Noise Ratio
RGKMC	Region Growing based K-Means Clustering
RMSE	Root Mean Square Error
RT	Repetition Time
SNR	Signal-to-Noise Ratio
SNR	Signal-Noise-Ratio
SPECT	Single Photon Emission Computed Tomography
SPF	Signed pressure force
SWT	Stationary <u>Wavelet Transform</u>
TA-cGAN	Tissue-Aware conditional Generative Adversarial Network
WHO	World Health Organization

Abstract

ABSTRACT

In recent years, multi-modal medical image fusion is a hot topic and attracts many researchers. The emergence of various types of medical images from different modalities leads to diverse research on this subject. Magnetic Resonance Image (MRI) displays the anatomical structure of internal parts of the human body. Brain tumors are solid masses that result from uncontrolled cell division, which leads to abnormal growth of brain cells. MRI images are used to detect the brain tumors. However, fusion of multi-modal images improves the accuracy of detection. Out of the various fusion methods, the deep learning based automatic segmentation methods have shown good performance in medical image analysis. However, the method has limitations in error rates and accuracy. Swarm-based optimization algorithms and deep learning algorithms have been proposed in this research work for improving the image fusion performance further.

In the first work, Modified-UNet is developed for Segmentation of Images. It is focused to avoid replication of least important attributes to mine attributes with more weight and of high determination data for the given brain images. The proposed system uses a residual way to filter replication of low-resolution attribute plot data. For the modified UNet, some of the proven alterations are implemented. Initially, residual connection was utilized in every block of encoder. 3 convolution layers are presented for every block of encoder. The outcome attained by previous encoder block is appended to the input of current block. Behind the concept of stack, UNet model was taken for the alteration. In case of modified UNet, in place of taking single UNet, dual UNet was considered where the output of the UNet was redirected back to its input. Making use of this repetitive design pattern consumes minimum parameters and also saves GPU memory.

In the second work, fusion of multi-modal images using deep learning algorithm is proposed. Initially, the database is collected which contains multi modal brain images. Brain tumor segmentation is performed via M-UNet algorithm efficiently. Then, feature extraction is done by using CNN algorithm which extracts more informative features. Finally, the image classification is done for obtaining more accurate results. CNNs signify a massive development in recognizing the images. This is helpful to analyse precisely classifying the images. From the result, it proves that the proposed Modified Particle Swarm Optimization and Modified Fully Connected Layer (MPSO+MFCL) provides greater image fusion performance in terms of precision, recall, accuracy and time period

The main contribution of this research is noise reduction, segmentation, feature extraction, image fusion and image classification. Initially, the database is collected which contains multi modal brain images. Then, noise removal is done by using Adaptive Median Filtering (AMF) which is focused to determine which pixels in an image have been affected by noise. Brain tumor segmentation is performed via Region Growing based K-Means Clustering (RGKMC) algorithm efficiently. In this work, tumor is identified in both images which provide earlier prediction of brain tumor. Then, feature extraction is done by using Adaptive Fire Fly Optimization based Convolutional Neural Network (AFFOCNN) algorithm which extracts more informative features via best fitness values. From the result, it concluded that the proposed AFFOCNN+MFCL algorithm provides better performance in terms of higher accuracy, precision, recall and lower Mean Square Error (MSE), execution time rather than the existing algorithms

In the final work, Ensemble deep learning algorithm for multi view image fusion is proposed. Initially noise removal and segmentation are done and then, feature extraction is applied through the Lion Swarm Optimization based Convolutional Neural Network (LSOCNN) algorithm which extracts the more informative image features. The ensemble DCNN, Differentiable Fusion with Mutual Information-Network (DFMI-Net) and Tissue-Aware conditional Generative Adversarial Network (TA-cGAN) algorithm is proposed to improve the image fusion performance. Here ANN is used for the detection of quality image features in cGAN layer. EDL algorithm integrates the best features of the above three algorithms for better image fusion results. From the result, it concluded that the proposed LSO based Ensemble Deep Learning (EDL) algorithm provides better performance in terms of higher accuracy, precision, recall and lower MSE, execution time compared to the existing algorithms.

Chapter 1

CHAPTER 1

INTRODUCTION

1.1 GENERAL BACKGROUND

The swift advancement of imaging technologies and computational capabilities enables the processing of diverse data from multiple sensing methods. This advancement is applied in numerous fields such as medical imaging, computer vision, and remote sensing. One of the key advantages of merging results from different modalities is the ability to harness complementary information from each, resulting in enhanced images. The technique commonly used to integrate information from multi-modal images is known as image fusion. A single image that captures the key aspects of the input images is the purpose of image fusion.

Image fusion involves combining complementary data from multiple images into a single composite image, which provides more detailed information than any individual source image. This technique is particularly useful for offering radiologists and physicians enhanced clinical information for medical diagnosis. Multimodal medical images, such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), and Single Photon Emission Computed Tomography (SPECT), are essential as they provide complementary information. Today, Medical imaging has become an essential component in many applications due to rapid advances in high technology and contemporary devices, including diagnosis, research, and treatment.

Early brain tumor identification may aid in the full recovery of the condition. So, the necessity of developing techniques to detect of brain tumor at an early stage is warranted. Different researchers have created and are currently working on a variety of approaches for the identification of brain tumors. This section gives an introduction of on image fusion and techniques.

1.1.1 Image Fusion

Image fusion combines many images of the same scene to create a better image for practical application and observation. One may see these several images of the same subject as a collection of connected images. Fusion can occur at the pixel, feature, or symbol level (Song et al., 2021) [94]. In the field of medical imaging, fusion typically utilizes pixel-level techniques. Additionally, several other methods are employed in the medical domain, including wavelet-based methods, the weighted average method, and contrast modulation techniques based on morphology.

There are several uses for image fusion in medical diagnostics. For instance, doctors physically combine the medical images to make a more accurate diagnosis. However, using the same type of image, decision and accuracy are inconsistent. Therefore, to increase diagnostic accuracy, an automated image fusion system must be developed.

1.1.2 Medical Image Fusion

Various clinical applications have made extensive use of multi-modality medical image fusion systems. This process generates images that provide both anatomical and physiological information, aiding specialists and reducing the need for multiple diagnostic procedures. The technique of merging features from many imaging modalities

into a single, clearer image is known as medical image fusion (Yadav and Yadav 2020) [113]. Maintaining the original distinctions between the input images and the merged image intact while avoiding distortions is the primary problem in image fusion. Multi-scale image fusion has great potential for information integration, according to recent research on medical image fusion.

Medical images come in various forms, including CT, MRI, PET, and SPECT, each serving different purposes (Liu et al., 2014) [67]. For example, PET and SPECT provide functional information, offering insights into visceral metabolism and blood circulation, albeit with relatively low spatial resolution. In contrast, CT, MRI, and B-mode ultrasound produce anatomical images with high spatial resolution.

1.2 BRAIN TUMOR OVERVIEW

In recent decades, analyzing human brain tumors has become a significant challenge in medical science. Typically, MRIs are used to diagnose brain tumors. If an MRI shows that a tumor is present, the type of brain tumor is usually identified through tissue sample analysis after a biopsy or surgery (Altaei and Kamil 2020) [7]. Brain scans provide detailed images of the brain, assisting doctors in detecting and diagnosing conditions such as tumors, the causes of strokes, or vascular dementia.

The body of a person consists of several types of cells, with the brain being the most important one (Shen et al., 2017) [91]. As the most critical part of the nervous system, the central nervous system's brain functions as its core. This intricate organ contains between 50 and 100 billion neurons. It comprises numerous cells, each fulfilling

a unique role. Most of these cells are produced within the body to create new ones, ensuring the proper functioning of the human body.

In the body, damaged or old cells decompose to make room for new ones to grow in their place. But sometimes, malfunctioning or aging cells fail to undergo programmed cell death, and new cells are created even when they are not required. This causes an overabundance of cells to accumulate and create a lump of tissue called a tumor. The delicate functioning of the brain is disrupted when a tumor grows. Due to its location and potential to spread, treating a brain tumor is extremely challenging and hazardous. Two primary forms of brain tumors are distinguished: benign tumors, which contain cancer-causing cells, and malignant tumors, which contain (Isensee et al., 2018) [44].

Brain tumors must be identified and detected as soon as possible. Currently, systematic and targeted identification of brain tumors is often accomplished via the use of Computer-Aided Diagnosis (CAD) systems. Brain tumors, though relatively rare, are a significant health challenge in India, with an incidence of about 5-10 cases per 100,000 individuals annually, accounting for approximately 2% of all malignancies. Gliomas, particularly glioblastomas, are the most prevalent types of cancerous brain tumors, followed by meningiomas, pituitary adenomas, and schwannomas. With a male-to-female ratio of 1.5:1 for malignant tumors, there is some male preponderance.

Age-specific prevalence is notable, with medulloblastomas more common in children and glioblastomas in adults. Urban areas report higher incidence rates due to better diagnostic facilities, while rural areas face significant challenges in terms of healthcare infrastructure, leading to delayed diagnoses. Symptoms such as persistent

headaches, seizures, and cognitive changes are often misattributed to less serious conditions, contributing to late diagnoses.

Survival rates for brain tumor patients in India are lower than in developed countries, with a five-year survival rate for malignant tumors estimated at 20-30%. This is primarily due to late-stage diagnosis and limited access to comprehensive treatment options.

To address these challenges, initiatives by the government and NGOs include establishing regional cancer centers, enhancing diagnostic facilities, and conducting awareness campaigns. The Indian Council of Medical Research (ICMR) and its National Cancer Registry Programme (NCRP) play a crucial role in data collection and analysis. Efforts to improve healthcare accessibility through telemedicine, mobile health units, and training of healthcare professionals are essential for better outcomes. On-going research aims to develop targeted therapies and improve treatment protocols. To preserve lives, advanced brain tumor detection is crucial. Furthermore, tumor identification procedures must be carried out with extreme precision and speed.

1.2.1 Brain Tumor Development

Brain tumors are caused by uncontrolled cell proliferation. Brain cells, meninges, glands, and nerves may cause these cancers. They can specifically damage brain cells by increasing pressure within the skull (Amin et al., 2020) [8]. Figure 1.1 provides a visual representation of a brain tumor. The most hazardous kind of tumors are malignant ones. The median overall survival (OS) for malignant glioma remains unchanged after decades of intensive investigation, specifically Glioblastoma multiforme (GBM), remains only fifteen months. Due to their severity, tumors are categorized into different grades. Grade

1 tumors are the least dangerous, typically associated with prolonged survival. They grow slowly and appear nearly normal under a microscope.

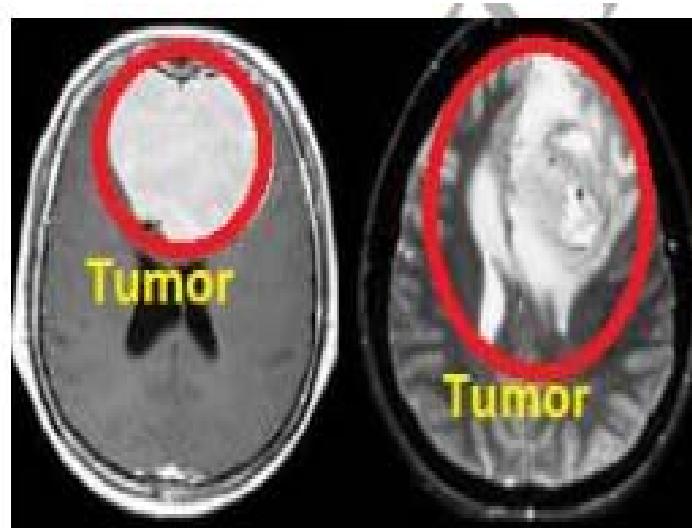


Fig 1.1 Example of brain tumor

Surgical treatment may be effective for this type of tumor grade. Examples of grade 1 brain tumors include pilocytic astrocytoma, ganglioglioma, and gangliocytoma. Grade 2 tumors grow slowly and appear abnormal under a microscope. Some of these tumors can invade nearby tissues and recur, as higher-grade malignancies on sometimes. Tumors in grade 3 are malignant, but there is typically little difference between grade 2 and grade 3 tumors. Tumors of grade 3 sometimes return as grade 4 types. Tumors of grade 4 are the most dangerous type. It grows rapidly, has an abnormal appearance under a microscope, and readily invades surrounding brain tissues, leading to the formation of new blood vessels. Dead cell regions are often seen in the center of these tumors. The grade 4 tumor GBM is one example.

1.3 SIGNIFICANCE OF IMAGE FUSION

The medical image fusion is an important part of the medical field to integrate two or more human body images into one to identify all possible diagnosis, diseases in the human body so that effective treatment can be applied (Hermessi et al., 2021) [37], (Abdulkareem 2018) [2]. Image fusion requires two fundamental things: input images must have pertinent information that the fused image needs and the fusion process must leave the merged image free of distortion and undesired elements. Image fusion is the technique which combines multiple sensor outputs into a single frame (Huang et al., 2014) [42]. The single sensor output does not provide entire information, while multiple sensors provide sufficient information.

The image fusion techniques consist of either single sensor or multiple sensors. When fusion of a single sensor, the images of the objects are collected during different time interval and can be combined into a single image frame. The image fusion approach combines a high-resolution panchromatic and low-resolution multi-spectral image (Zheng et al., 2018) [122]. With this technique, high-resolution multi-spectral images combine color information with panchromatic image spatial detail. The primary applications of image fusion algorithms are in the domains of medical imaging and remote sensing.

The technique of mapping many clinical characteristics into one image is known as medical image fusion. A single modality of same patients at different time interval is used to detect the various pathologies such as growth of tumor cells, at the same time single modality is limited with complementary feature capture. For example, CT imaging provides dense bone structure, MR imaging provides the soft tissues information, PET and SPECT imaging provides the functional details such as blood flow and metabolism.

The fused medical images have been used in various pathological identification and localization. The fused medical image frames not only support accurate diagnosis but also reduces the storage space and enhances the features of the image. Fig. 1.2 provides examples of many medical picture modalities that are often utilized for multimodal fusion applications.

The process of decomposing an input picture into its component frequency bands is known as decomposition (Tan et al., 2020) [100]. Transformation-based approaches are typically employed in this regard.

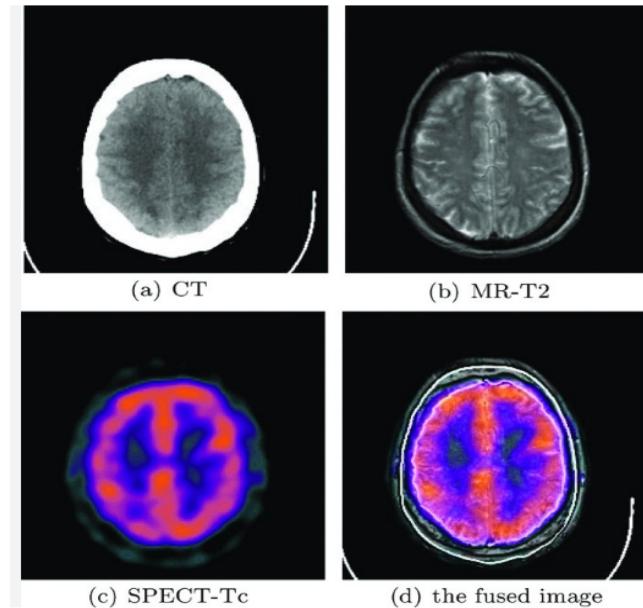


Fig 1.2 Multimodal brain medical image fusion examples

In this work, it is thought that MRI images enhance the overall performance of fusion. In particular, the study project aims to enhance the visual content via optimization by combining several complementing data sets. Additionally, it seeks to increase detection accuracy and efficiency.

1.4 IMAGE PRE-PROCESSING METHODS

1.4.1 Noise reduction

Improving image quality is the main objective of noise reduction through application of filters namely mean filters, median filters, and Adaptive Median Filters (AMF) (Suhas and Venugopal 2017) [97].

1.4.1.1 AMF

An example of a non-linear noise reduction technology is the AMF filter. By substituting the median intensity value from the surrounding pixels for the value of the center pixel, noise is removed from the transformed grayscale image (Meher and Singhawat 2014) [73]. As impulse noise appears as a scattering of black and the presence of white dots on the image is often known as salt and pepper noise. AMF filters are particularly good at reducing this type of noise. Figure 1.3 displays (a) the input MRI brain image and (b) the filtered image.

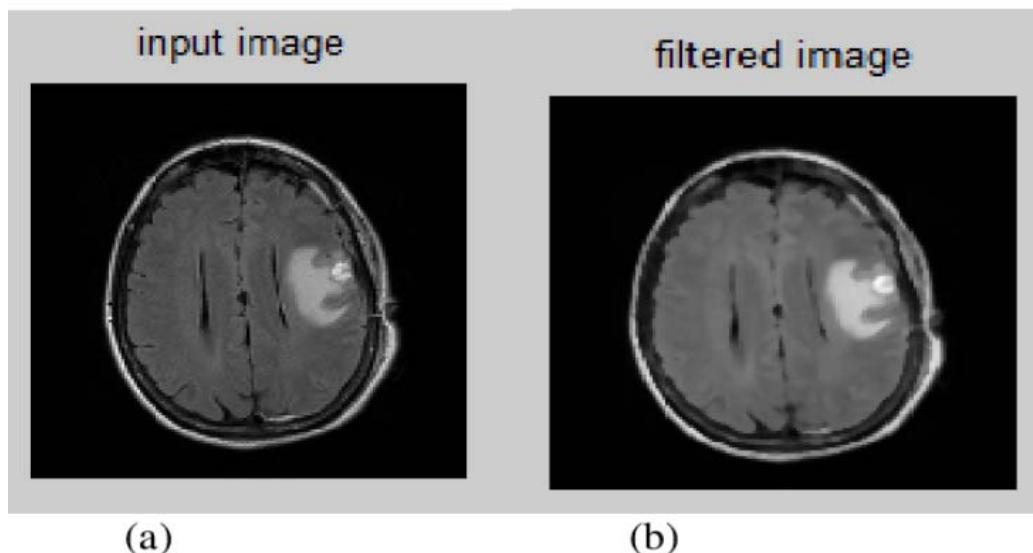


Fig 1.3 (a) MRI brain input image (b) filtered image

1.4.2 Segmentation

Segmentation divides an image into similar regions, grouping these similarities together (Çiçek et al., 2016) [21]. Segmentation aims to extract important features from images, allowing for easier interpretation of information. In medical imaging, separating brain tumors from MRI images is an intriguing but difficult issue. Figure 1.4 shows a case of MRI brain image segmentation.

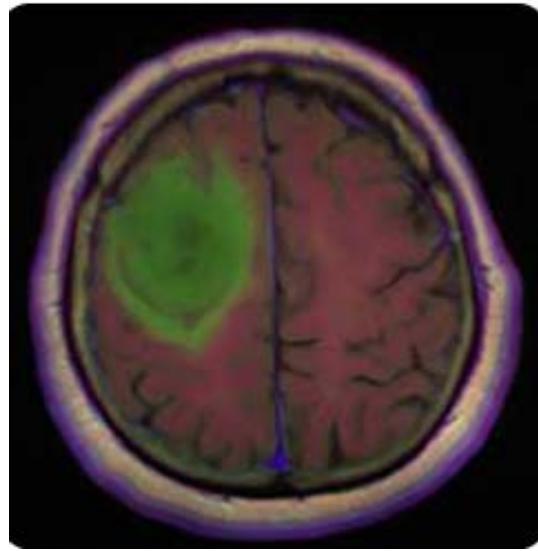


Fig 1.4 Example of MRI brain image segmentation

The medical image's representation is more understandable and straightforward for analysis is the aim of segmentation. The process of segmentation separates the image into sections, each of which is made up of linked pixels with comparable grayscale, color, or intensity values. The technique of splitting an image into many segments that collectively include the whole image is known as image segmentation (Du and Gao 2017) [23]. This process is helpful in determining whether each pixel belongs to a region of interest.

1.4.2.1 Region based segmentation

Region-based segmentation methods examine individual pixels in an image and combine homogeneous nearby pixels according to predetermined similarity criteria to form separate areas (Li et al., 2023) [54]. Methods such as region growing and watershed segmentation fall under this category and are commonly employed for brain tumor segmentation. Among these, the simplest and most popular approach is region growth, designed to isolate a contiguous region of similar pixels within an image. This method begins with one or more initial seeds located within the target structure. Neighboring pixels are then evaluated, and those that fit the requirements for similarity are included in the region. These criteria are defined by specific pixel intensity ranges or other image features. It is possible to choose seeds manually or automatically. Iteratively, the procedure continues until the area can no longer accommodate any more pixels.

The benefit of region growing is its ability to accurately segment regions with similar properties while producing connected regions. MR brain tumor segmentation research shows region expansion, particularly for homogenous tissues and areas, is a less computationally demanding and more successful way than other non-region-based approaches. However, MR brain image segmentation is less accurate due to the partial volume effect, is a major disadvantage of area growth. As a voxel might represent many tissue types, this effect softens the intensity differences at the boundaries between different tissue classes. Brain tumors may be automatically segmented using MRI data utilizing a fuzzy information fusion framework. Certain segmentation techniques improve their results by employing region growth.

1.4.2.2 Modified UNet (MUNet) algorithm

In (Mahasin et al., 2023) [69], the creation of a lightweight and precise image segmentation framework is explored. We want to examine the impacts of various MRI image input sizes in this work, from 512×512 to 16×16 , on evaluation measures like as the Dice Coefficient. Additionally, the study aims to analyze how these changes in image size influence the visual representation of the images (Chen et al., 2018) [20]. Reducing the input size decreases the number of computational steps, thereby enhancing processing speed. However, the sharpness of the image's visual representation reduces as the input size becomes smaller.

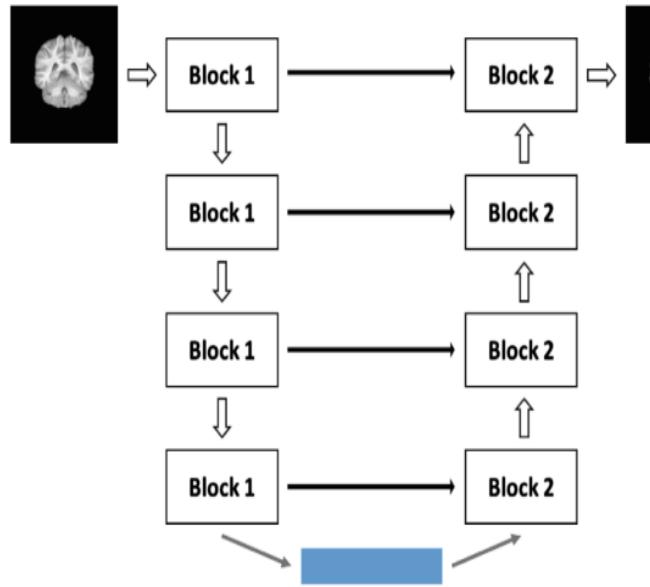


Fig 1.5 Illustration of modified U-Net model

The revised U-Net architecture is depicted in Figure 1.5, comprising two main sections: the contracting part and the expansive part, as described by (Chen et al., 2018) [20]. Each section includes multiple blocks. The primary modification is found in the

contracting section, when the input for the next layer is concatenated after 1×1 and 3×3 filters are employed inside the same layer. For each 2×2 max-pooling layer, the number of feature channels is increased to keep as much information as feasible since the output of the layer lowers the map size by half. The enlarged portion is still the same as the original U-Net. The outputs are transformed into a probability matrix for brain MR image segmentation in the last layer using four 1×1 filters.

1.4.3 Feature extraction

Evaluating to determine whether every area that was removed is a tumor is important once the regions of interest are extracted. This involves analyzing the features of each image to differentiate between actual tumors and false positive pixels. The feature extraction stage is crucial in image mining, as it identifies the most relevant features to reduce the complexity of the process. These features assist in identifying characteristics of brain tumors. Fig 1.6 shows example of MRI image feature extraction Optimization based feature extraction algorithms.

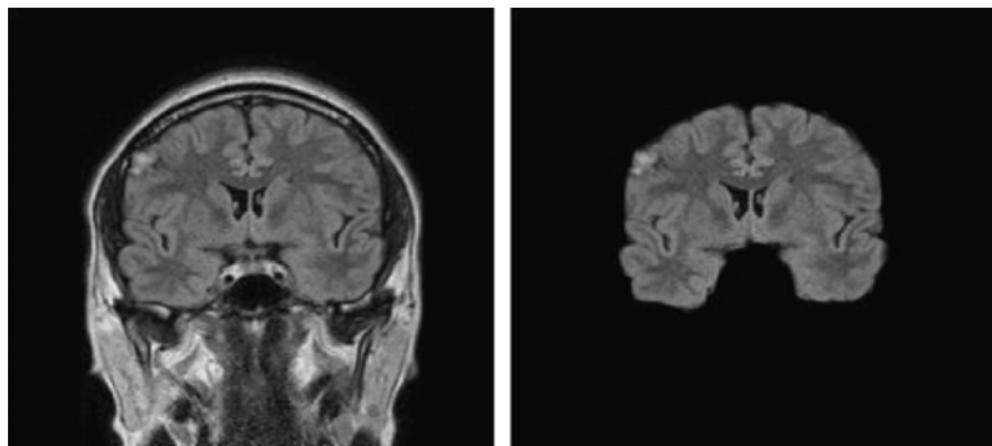


Fig 1.6 Example of MRI image feature extraction

1.4.3.1 Particle Swarm Optimization (PSO)

PSO accurately segment complete tumor region (Sharif et al., 2020) [89]. Initially, the skull is removed via Brain Surface Extraction (BSE). The image, now without the skull, is then processed with PSO to improve segmentation. Subsequently, a Genetic Algorithm (GA) selects the best Local Binary Patterns (LBP) and deep features from segmented images. The tumor grades are then categorized using an Artificial Neural Network (ANN) and other classifiers. With a maximum accuracy of 99%, the approach is assessed utilizing complicated brain datasets that are made publically accessible, including RIDER and the BRATS 2018 Challenge. The results, when compared to existing methods, demonstrate that this technique offers superior outcomes, proving its effectiveness and innovation.

1.4.3.2 Firefly algorithm

A firefly mating signal inspired the Firefly Algorithm (FA), as developed by (Bacanin et al., 2021) [14]. This metaheuristic algorithm addresses optimization problems by emulating use firefly flashing behavior to choose the optimal solution.

The Firefly Algorithm operates as follows, in brief:

Initialization: Generate a population of fireflies, each of which stands for a possible fix for the optimization issue. An object in the search space is represented by a firefly's position.

Objective Function Evaluation: Evaluate each firefly's fitness by comparing it to the optimization problem's objective function. Objective function determines firefly efficacy.

Movement of Fireflies: Fireflies exhibit positive phototaxis, meaning they are attracted to brighter fireflies in the search space. The allure of a firefly is contingent upon its luminosity, which is inversely correlated with its proximity to other fireflies. Fireflies are drawn to more luminous fireflies and navigate towards them as they explore the search area.

Light Intensity Update: The light intensity of each firefly is updated based on its distance to other fireflies. Fireflies closer to brighter fireflies increase their brightness, while those farther away decrease their brightness. This program replicates the intermittent illumination pattern seen by fireflies in their natural environment.

Update Positions: Update the positions of fireflies based on their movement towards brighter fireflies. This step ensures search space exploration and exploitation.

Termination Criteria: Once a termination requirement is satisfied for example, proceed with steps 2 through 5 until an acceptable solution or the maximum number of iterations is reached.

The Firefly Algorithm effectively balances exploration by simulating the attraction behavior of fireflies. In situations with complex objective functions and high-dimensional search spaces, it is very helpful in addressing optimization problems.

The algorithm's performance depends on several parameters, such as the attractiveness coefficient, the light absorption coefficient, and the population size. Proper parameter tuning and adjustments are essential for achieving good convergence and finding high-quality solutions.

Numerous optimization issues, such as those involving engineering design, scheduling, image processing, and machine learning, have been effectively solved using the Firefly Algorithm. Its simplicity, effectiveness, and ability to handle complex optimization landscapes make for solving optimization issues in a variety of domains, it is a popular option.

1.4.3.3 Lion Swarm Optimization (LSO) algorithm

The social behavior of lions in a pride serves as the model for the LSO algorithm, an optimization technique, as developed by (Sudha et al., 2019) [96]. Such as Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO), additional swarm intelligence algorithms, LSO addresses optimization challenges by mimicking natural animal behaviors.

In LSO, the lions in a pride represent the solutions to the optimization problem. Each lion's position in the search space corresponds to a potential solution, and the fitness of the lion represents how good that solution is. The lions cooperate and communicate with each other to explore the search space effectively and efficiently.

The algorithm involves several key steps:

Initialization: Within the search area, start a community of lions at random.

Objective Function Evaluation: Determine each lion's fitness by comparing it to the optimization problem's objective function.

Lion Movement: Based on personal experience, lions modify their positions and the collective information obtained from other lions.

Communication and Cooperation: Lions share information about their best solutions with other lions in the pride, allowing them to collectively explore promising regions of the search space.

Update: Update the positions of lions based on their movement and the shared information.

Termination: Until a termination condition (maximizing iterations or finding a suitable solution) is satisfied, repeat steps 2 through 5 again.

Regarding exploration, LSO seeks to achieve a balance between concentrating on the most promising areas discovered so far and scanning the whole solution space for potential regions. By mimicking the collective behavior of lions, the algorithm can efficiently navigate complex search spaces and find near-optimal solutions to optimization problems.

Numerous optimization issues, including feature selection, engineering design, scheduling, and many more, have been addressed by LSO. Its effectiveness depends on appropriate parameter tuning and problem-specific adjustments.

1.4.4 Image Fusion Methods

Image fusion combines numerous images of a scene to create a more detailed image. Image fusion extracts and integrates relevant information from each input image to generate a composite image with the most important attributes (Gholami et al., 2016) [30]. There are several methods used for image fusion, each with its own strengths and applications. Here's an explanation of some common image fusion methods:

Pixel-Level Fusion

Averaging: This method averages the input image pixel values to find the fused image's equivalent pixel value. This basic procedure may lose fine details but is efficient.

Weighted Sum: Similar to averaging, but with different weights assigned to each input image based on factors such as image quality or relevance to the scene. This method allows for more flexibility in preserving important information.

Maximum/Minimum Selection: The fused image is created by selecting the largest (or least) pixel value among related pixels in the input images. This approach preserves the most important input image attributes.

Transform-Based Fusion

Wavelet Transform: Wavelet transform divides input images into frequency bands. Fusion is performed by combining coefficients from different bands based on certain criteria, such as magnitude or phase consistency.

Discrete Cosine Transform (DCT): DCT decomposes the images into their frequency components, much like a wavelet transform. Fusion is performed by selecting coefficients from different frequency bands and combining them to form the fused image.

Principal Component Analysis (PCA): Through the use of statistics, PCA divides the images into orthogonal components. Fusion is performed by selecting components that capture the most variation in the input images.

Feature-Level Fusion

Feature Extraction and Combination: Features such as edges, textures, or key points are extracted from the input images using techniques like gradient analysis, Gabor filtering, or SIFT (Scale-Invariant Feature Transform). These characteristics are then combined to generate the fused image.

Segmentation-based Fusion: The input images are segmented into regions based on similarities in color, texture, or intensity. Fusion is performed by selecting regions from each input image and combining them to form the fused image.

Deep Learning-based Fusion: Deep neural networks are trained to learn feature representations from the input images and perform fusion based on learned features. CNNs and autoencoders are commonly used for this purpose.

Decision-Level Fusion

Rule-based Fusion: Fusion rules or decision-making algorithms are applied to determine the pixel values of the fused image based on predefined criteria or heuristics.

Fuzzy Logic Fusion: Fuzzy logic is used to model uncertainty and ambiguity in the fusion process. Fusion rules are defined using linguistic variables and fuzzy inference systems to make decisions based on imprecise or incomplete information.

Several standards, including the qualities of the input images, influence which image fusion approach is best. Each method has benefits and drawbacks, the application requirements, and computational constraints. Experimentation and evaluation are often necessary to determine the most suitable fusion method for a given task.

1.4.4.1 Convolutional Neural Network (CNN)

Image processing often makes use of CNN and other deep neural networks (Trivedi and Sanghvi 2022) [102]; (Shao and Cai 2018) [88]. When it comes to tasks like image identification, object detection, and classification, it is one of the most potent artificial intelligence algorithms.

Here's an overview of the key components and concepts of a CNN:

Convolutional Layers: These are the fundamental elements of CNNs, comprising learnable filters (kernels) that move across the input image to perform convolution operations, extracting features such as edges, textures, and patterns. With the help of these layers, the network may learn how to represent the input data hierarchically.

Activation Functions: Non-linearity from these functions lets the network capture complicated data connections. ReLU (Rectified Linear Unit), sigmoid, and tanh are the most used activation functions because of their simplicity and efficacy.

Pooling Layers: Pooling layers preserve critical information while reducing convolutional layer feature maps' spatial dimensions. Most CNNs employ max and average pooling.

Fully Connected Layers: In CNNs, convolutional and pooling layers are followed by one or more completely connected layers. To learn high-level characteristics and provide predictions, these layers connect each neuron in one layer to every other layer's neuron.

Softmax Layer: For classification tasks, CNNs usually conclude with a softmax layer, which converts the network's final outputs into probability distributions over the possible classes, enabling the network to forecast, given an input image, the chance of each class.

Training: Stochastic gradient descent (SGD) and its variations are examples of gradient-based optimization methods that are used to train CNNs. During training, the network minimizes a predefined loss function, such as categorical cross-entropy, by using backpropagation to change the connection weights.

Natural language processing, computer vision, and medical image analysis are just a few of the domains that CNNs have revolutionized. Their capacity to recognize hierarchical structures from unprocessed data automatically, along with their scalability for large datasets and complex problems, their importance as tools for contemporary applications of artificial intelligence and machine learning.

1.4.4.2 Generative Adversarial Networks (GAN)

(Guo et al., 2022) [34] presented GANs, a kind of deep learning framework. Two neural networks make up GANs: the discriminator and the generator, which are trained concurrently through a competitive process.

Here's an overview of how GANs function:

Generator: This network produces artificial data samples by using random noise as input. Initially, the outputs from the generator do not resemble the target data distribution.

Discriminator: Real data samples (from the genuine data distribution) and fictitious samples produced by the generator are distinguished by the discriminator

network via training. It gains the ability to give genuine samples a high probability and fake samples a low probability.

Adversarial Training: The generator's objective in this training procedure is to generate samples that the discriminator can't tell apart from genuine data. The discriminator's objective is to correctly discriminate between real and false samples. A min-max game is played between the networks in which the discriminator attempts to increase its accuracy and the generator wants to reduce it.

Loss Functions: The generator and discriminator have distinct loss functions. The generator's loss is based on the discriminator's feedback, aiming to minimize the likelihood that its generated samples are identified as fake. Conversely, the discriminator's loss focuses on maximizing its accuracy in classifying real versus fake samples.

Training Dynamics: As training continues, the generator improves at creating samples that closely match the real data distribution, while the discriminator becomes more adept at detecting fakes. Ideally, when training is complete, the discriminator finds it difficult to discern between authentic and fake samples, and the generator's outputs are identical to actual data.

GANs have achieved significant success in generating realistic images, audio, video, and other data types. Applications for which they are used include image synthesis, super-resolution, image-to-image translation, style transfer, and the production of artificial data for neural network training.

However, mode collapse in which the generator learns to generate just a limited number of sample types and instability during training may make training GANs

difficult. To address these obstacles, researchers have put forward several methods, such as architectural modifications, regularization methods, and different training strategies. Despite these challenges, GANs remain a powerful tool for generative modeling and have significantly advanced the field of deep learning.

1.5 PROBLEM SPECIFICATION

The obstacles encountered in image fusion methods affects the quality and efficiency. Restrictions in image sensors and signal transmission results in poor resolution in source images. Particularly brain tumor detection from such images tend to be imprecise. The existing fusion methods in brain tumour detection, due to these limitations have not guaranteed accuracy and reliability. The problems for such limitations have been identified in segmentation and fusion. The main objective of this research project is to make the aforementioned portions more efficient.

1.6 OBJECTIVE OF THE RESEARCH

The research work aims to achieve superior fused images of brain of high precision in lesser processing time. The primary focus is reducing the amount of data and producing images that are more acceptable for human and machine perception. The study aims to develop an ensemble-based deep learning system that will improve the performance of Multiview MRI brain image fusion.

1.7 CONTRIBUTION OF THE RESEARCH

1. Modified-UNet for Segmentation of Brain Images
2. Multimodal images fusion using Deep Learning Algorithms.

3. Employing a deep learning technique that relies on clustering, the best possible segmentation and integration of multimodal brain images
4. Deep learning approach for ensemble learning in multiview image fusion

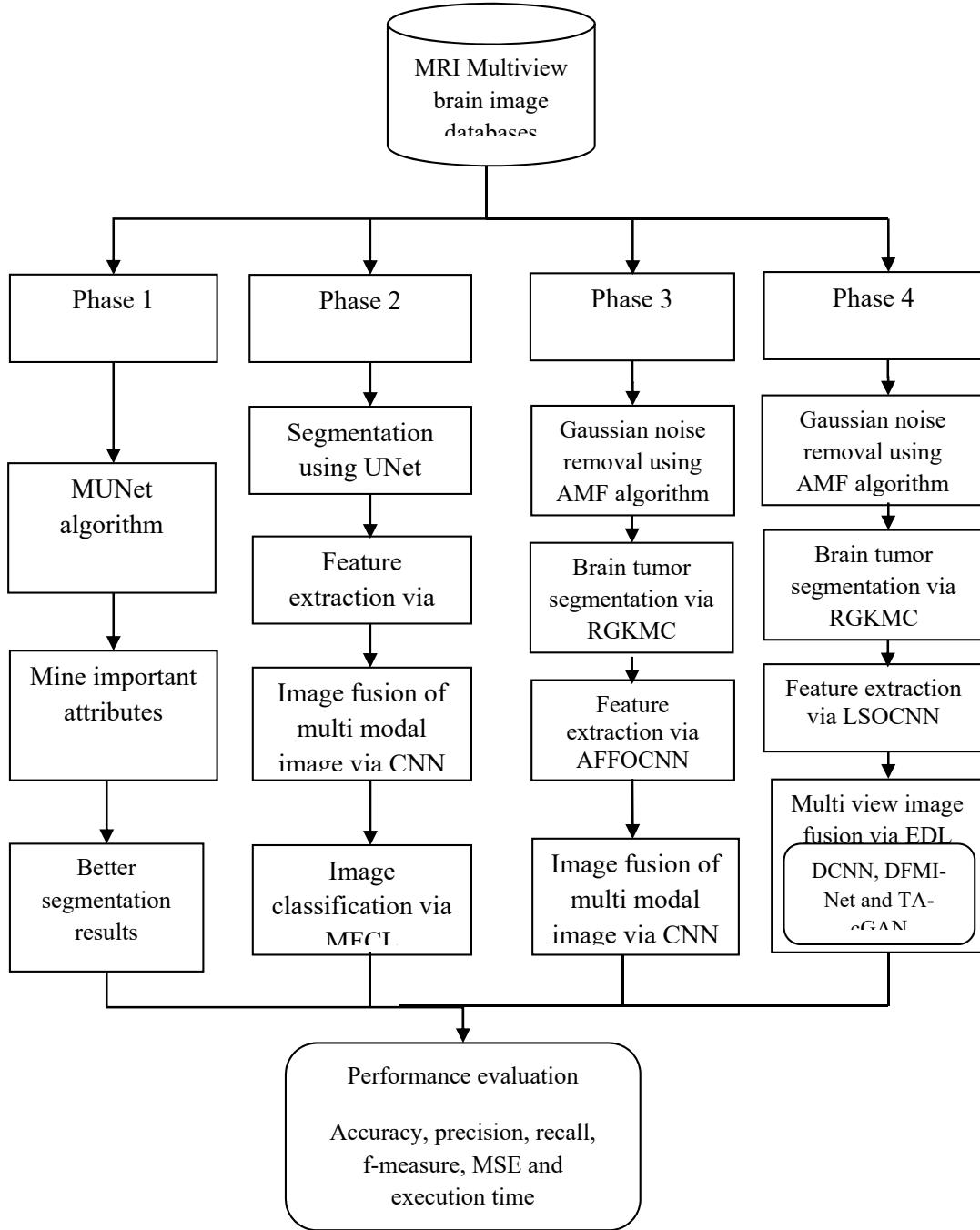


Fig. 1.7 General architecture diagram of the five proposed methodologies

1.7.1. Modified-Unet For Segmentation Of Brain Images

This study introduces the Modified UNet (MUNet) algorithm to address the issue of redundant replication of less significant attributes in processing brain images, prioritizing attributes of greater importance and higher data determination. The aim is to extract more refined overall attributes for insignificant objects by optimizing the assembly tasks. In comparison to conventional CNN and UNet approaches, the Modified UNet effectively retains edge and morphological data of the image elements. A residual method is employed to filter out redundant low-resolution attribute data. Several modifications are applied to the UNet framework: every encoder block has three convolution layers, residual connections are merged into each block, and the input and output of the preceding encoder block are concatenated. Additionally, update the max pooling layer by adjusting the final convolutional layer stride. The concept of Permeation Data (PD) is introduced, wherein the UNet model is modified to include dual UNets, redirecting the output of one UNet back to its input. Reduces parameters and saves GPU RAM with this repetitive design pattern. The cost function comprises the aggregated cost from the initial and successive iterations of the UNet. Incorporating intermediate outcomes in the cost estimation accelerates training efficiency, although computational complexity remains a challenge. Strategies to mitigate this complexity are discussed in the subsequent section.

1.7.2. Multi-Modal Image Fusion Using Deep Learning Algorithm

Modified Particle Swarm Optimization Modified Fully Connected Layer (MPSO_MFCL) has been presented in this study as a novel multi-modal medical image fusion technique. We specialize in segmentation, feature extraction, image fusion, and

classification. Initially, a database comprising multi-modal brain images is compiled. Efficient brain tumor segmentation is achieved using the MUNet algorithm. Subsequently, MRI brain images are processed to extract more useful characteristics using CNN-based feature extraction, considering images from various perspectives. Fusion of images from different views is conducted to produce a clearer composite image, leveraging the extracted features. Finally, image classification is performed to enhance accuracy. However, a challenge arises regarding inaccurate segmentation results, attributed to the exclusion of white images in prior studies. To address this limitation, the subsequent section proposes solutions to improve segmentation accuracy.

1.7.3. Optimal Segmentation And Fusion Of Multi Modal Brain Images Using Clustering Based Deep Learning Algorithm

In this study, we propose the Adaptive FireFly Optimization based Convolutional Neural Network (AFFOCNN) scheme to enhance image fusion performance. The primary focus of this research encompasses noise reduction, segmentation, feature extraction, image fusion, and image classification tasks. Initially, a database containing multi-modal brain images is compiled. Noise reduction is achieved using Adaptive Median Filtering (AMF), which identifies and replaces noise-affected pixels with the median value of neighbouring pixels that pass the noise labelling test. This approach significantly improves image quality by reducing noise interference. Efficient brain tumor segmentation is then accomplished through the RGKMC algorithm, facilitating the separation of tumors from normal brain tissues. Feature extraction is conducted using the AFFOCNN algorithm, this depending on ideal fitness values, retrieves more informative characteristics. Image fusion aims to preserve complementary information while

eliminating redundant data and distortions from individual source images. However, challenges arise in the early prediction stage of brain tumors. To address this issue, the subsequent section discusses the Enhanced Deep Learning (EDL) approach as a potential solution.

1.7.4. Ensemble Deep Learning Algorithm For Multi View Image Fusion

In this study, we provide a method for multi-view image fusion called Enhanced Deep Learning (EDL), which includes important phases such as feature extraction, segmentation, noise reduction, and image fusion. First, by removing noise interference, noise reduction is done to improve image quality. Subsequently, MRI images are segmented to subdivide them into constituent regions or objects, employing the RGKMC method, which segments images into black and white components based on Euclidean distance calculations for cluster centroids. This approach facilitates early brain tumor prediction by identifying tumors in both images. Following segmentation, Lion Swarm Optimization-based Convolutional Neural Network (LSOCNN) retrieves relevant visual characteristics. Multi-modal image fusion provides lower, medium, and higher-level image contents that may be examined and merged from many directions. An ensemble of techniques called Tissue-Aware conditional Generative Adversarial Network (TA-cGAN), Differentiable Fusion with Mutual Information-Network (DFMI-Net), and Deep Convolutional Neural Networks (DCNN) is suggested to improve image fusion performance. Artificial Neural Networks (ANN) are utilized for detecting quality image features in the cGAN layer. The EDL algorithm amalgamates the best features of these three algorithms for improved image fusion outcomes. Regarding mean square error (MSE) reduction, recall, accuracy, precision, and rapid execution, experimental findings show that the suggested

LSOC-based Ensemble Deep Learning (EDL) algorithm works better than current approaches. Future work will involve applying the EDL algorithm to MRI lung images.

1.7.5 Organization of the Thesis

In chapter 1, general introduction and image fusion analysis have been discussed. This section also provides overview of brain tumor in detailed way. In this chapter, additionally mentioned are imaging methods for brain tumor diagnosis. This section also explains significance of image fusion and fundamental steps of image fusion process. It includes discussion of noise reduction, segmentation, feature extraction and image fusion of MRI images. Finally, this chapter concludes with the problem statement, objective, and contribution of the overall thesis

In chapter 2, analysis of various research techniques that are introduced by different authors has been given. This chapter provides reviews of noise reduction and feature extraction methods, segmentation methods and image fusion techniques. The last consideration of the benefits and drawbacks of various research methodologies brought this part to a conclusion.

In chapter 3, the developed Modified-UNet technique for Brain Image Segmentation is explained. Experimental analysis adopting the proposed method has been explained.

In chapter 4, deep learning algorithms for fusion of images that has been developed is explained. Experimental analysis adopting the proposed method has been explained.

In chapter 5, the best possible segmentation and integration of multimodal brain images is addressed via the creation of a clustering-based deep learning algorithm. An explanation of experimental analysis using the suggested strategy has been provided.

In chapter 6, the ensemble deep learning approach that has been created for multitier image fusion is explained. Experimental analysis adopting the proposed method has been explained.

In chapter 7, overall conclusion from the research work based on the performance of the newly developed methods has been analysed. Suggestion for future development in this field has been given.

Chapter 2

CHAPTER 2

LITERATURE SURVEY

2.1 INTRODUCTION

The significance of medical image fusion in precisely identifying clinical images has gained an increase in interest. While numerous methods for medical image fusion have been developed over time, challenges persist, such as poor contrast, uneven illumination, noise, and inadequate fusion strategies, leading to insufficient representation of crucial features. Despite significant efforts by researchers, the research on medical image fusion continues to evolve. This chapter provides reviews on noise reduction and multimodal image fusion methods, reviews on segmentation methods, reviews on feature extraction and optimization techniques and reviews on image fusion using deep learning and ensemble deep learning algorithms over MRI brain images which are relevant to this research.

2.2 REVIEWS ON NOISE REDUCTION AND MULTIMODAL IMAGE FUSION METHODS

Suhas et al., [2017] [97] discuss noise reduction in medical imaging applications, and image processing methods for MRI medical images. This study evaluates MRI image de-noising algorithms, namely median, Gaussian, Max, Min, and Arithmetic Mean filters. MRI images of the brain and spinal cord are processed via these filters, and results are observed. By incorporating additional characteristics, a new technique is proposed which improves the conventional median filter. The experimental results of this novel approach are compared with the existing image-filtering methods. The statistical measures namely

Signal-to-Noise Ratio (SNR), Peak Signal-to-Noise Ratio (PSNR), and Root Mean Square Error (RMSE), are adopted to evaluate the output images' efficiency.

Tan et al., [2021] [99] propose a strategy for fusing multi-modal medical images using a multi-level edge-preserving filtering (MLEPF) decomposition model. Fine-structure (FS), coarse-structure (CS), and base (BS) layers are the three categories into which multi-modal medical images are first classified by the MLEPF model, based on weighted mean curvature filtering. Then, the FS, CS, and BS layers are fused using gradient-domain pulse-coupled neural network (PCNN) and energy attribute fusion, respectively. Fused images are created from these three sorts of fused layers. System assessment employs over 100 image pairs from six diseases and one normal dataset. In qualitative and quantitative evaluations, the algorithm outperforms existing state-of-the-art methods.

Pritika and Budhiraja [2016] [81] discuss on multimodal image Medical fusion by, integrating CT and MRI data in a single image. Existing techniques, such as those based on transform domains and guided filter-based spatial domains, often suffer from contrast reduction and halo artifacts. This study introduces guided filtered multi-scale decomposition for picture fusion. A guided filter separates input images into a base layer and numerous detail layers. The detail and foundation layers are then fused using numerous criteria. According to simulation studies, this fusion approach works better than conventional techniques since it preserves important characteristics in the merged image without causing distortion.

Tan et al., [2020] [100] propose a method for merging medical images across different modes which is, designed to address a variety of diagnostic challenges in healthcare. This method makes use of a novel fusion technique that combines energy

attribute fusion in the non-subsampled shearlet transform domain with boundary-measured pulse-coupled neural networks. To validate its effectiveness, the algorithm is tested on a dataset encompassing diverse diseases such as glioma, Alzheimer's, and metastatic bronchogenic carcinoma, comprising over 100 pairs of images. Both qualitative and quantitative assessments confirm its superior performance compared to existing methods, offering valuable insights for advancing medical diagnostics.

Maqsood and Javed [2020] [71] propose the use of sparse representation and two-scale image decomposition in a unique multimodal image fusion method is suggested. Initially, the source multimodal images undergo contrast enhancement to enhance their intensity distribution, thereby improving visualization. Subsequently, the contrast-enhanced images' edges are extracted using a spatial gradient-based edge detection algorithm. The final detail layer is then retrieved from the base and detail layers that are created during the decomposition of these improved multimodality images. Enhanced decision map and fusion create the merged image. The experimental findings, assessed both qualitatively and quantitatively, demonstrate that the multimodal image fusion system surpasses earlier methods.

Arif and Wang [2020] [10] suggested a fusion methodology that uses the Genetic Algorithm (GA) and curvelet transform for multimodal medical imaging. Integrating GA into the process helps mitigate uncertainties and diffuse inconsistencies present in the input image, while also optimizing fusion characteristics. Recent medical image fusion algorithms have been compared with this method, through extensive testing across numerous sets of medical images. The quantitative assessments and visual analyses

demonstrate that this method offers significant advantages in medical fusion image production compared to others, preserving data integrity and color fidelity in the base image.

Fenshia Singh and Magudeeswaran [2017] [27] explains that the anatomical structures of the brain can be analysed through brain MRI image segmentation, aiding in surgical planning, among other applications. However, the segmentation of the brain faces challenges due to noise and uncertainties present between different brain tissues in the images. This article addresses this issue through a two-stage approach. Initially, to increase the MRI brain image contrast, an approach called Contrast Limited Fuzzy Adaptive Histogram Equalization (CLFAHE) combines fuzzy enhancement with CLAHE. The contrast intensification operator known as Clip limit is used to control the image's contrast. The improved images go through segmentation in the second step. This is achieved using a novel level-set method possessing characteristics of both local and global segmentation. Additionally, a Signed Pressure Force (SPF) function is employed to effectively halt contours at weak and blurred edges.

Zhu et al., [2018] [124] introduced a novel technique for fusing images, based on breaking down images into their cartoon and texture constituents, followed by sparse representation. In this method, the source multi-modality images undergo decomposition into their cartoon and texture elements. A spatial-based technique is employed to preserve morphological structure within the cartoon components. The structure from each source picture is preserved via energy-based fusion. Use sparse representation for texture elements to train a robust dictionary. In the end, a texture enhancement fusion rule is used to combine cartoon and texture components. The visual and quantitative findings of this experiment show that it is superior to state-of-the-art procedures.

Liu et al., [2020] [63] introduced a novel image fusion technique which employs a bilevel optimization approach, specifically tailored for multi-modality image fusion assignments. Through focusing on certain elements that have been isolated from the initial images through alternation optimization, they achieve flexibility in fusion strategies across multi-modality images via adaptive integration weight maps. Utilizing this technique, it has been possible to effectively integrate visible and infrared imaging, magnetic resonance imaging combined with single-photon emission computed tomography. The quantitative and qualitative findings demonstrate this approach's efficiency and flexibility.

2.3 REVIEWS ON SEGMENTATION METHODS OVER MRI IMAGES

Alam et al., [2019] [5] propose the template-based K-means technique, which begins the segmentation process by selecting a template according to the gray-level intensity of the image. Subsequently, the fuzzy C-means (FCM) method determines the membership update by calculating distances between cluster data points and the cluster centroid until it finds the best option. Lastly, tumor sites are detected using the upgraded FCM clustering technique. Updates to the membership function are made in this way depending on several properties of the tumor images, including contrast, energy, dissimilarity, homogeneity, entropy, and correlation. Simulation outcomes demonstrate that this algorithm excels in detecting both abnormal and normal brain tissues, even with slight variations in gray-level intensity. Moreover, it remarkably detects human brain tumors within seconds, a significant improvement compared to minutes required by other algorithms.

Lim and Mandava [2018] [58] suggested a comprehensive approach in their study, incorporating both data-driven techniques and domain expertise for segmenting

multisequence MR images. Information modelling, information fusion, and visual object extraction are the three stages of the segmentation process. Initially, the random walks technique is modified to represent image data. To address challenges like ambiguous boundaries and intensity variations within the image, the weighting function of the random walk's method has extra elements about homogeneity and object attributes. The second step combines image sequences using weighted averaging. This method weights sequences for fusion using data and user skills. In the final phase, the concept of information theoretic rough sets (ITRS) is leveraged to tackle ambiguous boundaries between visual objects and their backgrounds during object extraction. The suggested methodology is assessed using the MICCAI brain tumor dataset for brain tumor extraction. Its performance is then compared to known methodologies to evaluate its effectiveness.

Forouzanfar et al., [2010] [28] explores the efficiency of GAs and PSO in determining the optimal degree of attraction. Although GAs are very proficient in achieving results that are close to ideal, they often encounter difficulties in locating solutions that are exact and accurate. Conversely, PSO's collective interactions improve the process of possible solutions. Therefore, combining the strengths of both methods in a hybrid approach is expected to yield significant improvements. To achieve this objective, a combined approach using GAs and PSO, known as "breeding swarms," is used to identify the optimal level of attraction. The segmentation process of existing FCM-based approaches are much improved as proved in quantitative and qualitative analyses done on brain MR images, both simulated and real, with various levels of noise.

Lin et al., [2012] [59] introduced Fuzzy Knowledge-Based Seeded Region Growing (FKSRG), an innovative image segmentation method for multispectral magnetic

resonance images, is presented. Fuzzy knowledge fuzzy edge, fuzzy similarity, and fuzzy distance derived from pixel correlations within multispectral magnetic resonance pictures are included in this technique. The adjusted seeded area developing process is then subjected to these fuzzy qualities. There can be an undetermined number of final regions when traditional areas merge. To ensure that the FKSRG approach neither oversegments nor undersegments images, target generation facilitates conventional area mergers. FKSRG effectiveness is evaluated using actual MR images and computer-generated phantom images. The FKSRG approach performs better in multispectral MR image segmentation than K-means, Support Vector Machine, and Functional MRI of the Brain Automated Segmentation Tool, according to experimental data.

Alagarsamy et al., [2017] [4] created an automated hybrid algorithm utilizing Cuckoo Based Search (CBS) and interval type-2 fuzzy based clustering, with the goal of achieving effective segmentation of Magnetic Resonance (MR) brain images. Radiologists can rapidly analyze complicated tumor regions with inaccurate gray levels using automated MR brain image segmentation. This article identifies tumor locations with substantial intensity variability and inadequate borders to facilitate clinical identification. MSE, peak signal-to-noise ratio, processing time, Dice Overlap Index, and Jaccard Tanimoto Coefficient Index are used to evaluate the technique. CBS and interval type-2 fuzzy-based clustering may outperform kernel-, entropy-, graph-cut-, and self-organizing maps-based clustering with a sensitivity of 0.7143 and specificity of 0.9375. The article presents significant improvements in tumor region segmentation, enhancing clinical diagnosis. Two experienced radiologists have validated the methodology, endorsing its utility in clinical oncology processes.

Valindria et al., [2018] [104] explore the efficiency of leveraging multiple modalities for enhancing segmentation accuracy on each specific modality. They investigate the potential of employing a dual-stream encoder-decoder architecture to learn features that are independent of modality, thereby ensuring generalizability and robustness. Notably, the MRI and CT data used in the study are unpaired, originating from different subjects and not aligned with each other. Experimental findings indicate that multi-modal learning leads to enhanced overall accuracy compared to training solely on individual modalities. Moreover, the results highlight that leveraging information across modalities can notably boost performance, particularly for complex structures like the spleen.

Kumar et al., [2021] [53] employs MRI image segmentation to identify tumors via pre-processing, feature extraction, feature selection, classification, and segmentation. Before being subjected to further analysis, input images from an MRI go through pre-processing. This stage involves using the Improved Gabor Wavelet Transform (IGWT) to convert the images into a new domain. Next, GLCM-related characteristics are retrieved, and the Oppositional Fruit Fly Algorithm (OFFA) is used to identify noteworthy features. After that, a SVM classifier is trained using these chosen characteristics to determine if the image is normal or aberrant. After categorization, anomalous images are segmented using a crude k-means algorithm that works well. The approach's accuracy, specificity, and sensitivity are assessed. The strategy performs better than current methods, according to experimental data.

Agrawal et al., [2018] [3] introduces a hybrid method for segmenting brain lesions across various imaging modalities, employing a combination of median filtering, k-means clustering, Sobel edge detection, and morphological operations. The median

filter removes impulsive noise from brain images as the first step in preprocessing. The images are then subjected to morphological processing, Sobel edge detection, and k-means segmentation. On typical datasets, the automated system's performance is assessed using measures like execution time and segmentation accuracy. Compared to a skilled radiologist's hand delineation, the approach achieves an astonishing 94% accuracy. Furthermore, high significance values of 0.986 and 1 for ANOVA and correlation coefficient are obtained from statistical significance tests comparing lesion segmentation using the automated technique to expert delineation, respectively. The experimental findings are discussed in relation to recent studies in the field.

Liu and Guo [2015] [61] achieve brain MRI image segmentation using SVM and K-means clustering. The technique uses K-means clustering to partition the brain MRI image and create the class label. As training and test samples, brain tissue pixel feature vectors are selected. Final SVM segmentation of brain MRI. Experiments have shown that this segmentation strategy effectively reduces noise in brain images with a low signal-to-noise ratio.

2.4 REVIEWS ON FEATURE EXTRACTION AND OPTIMIZATION TECHNIQUES

Guo et al., [2020] [35] suggested a PCA-based multi-feature extraction approach to monitor nonlinear dynamic processes. Dynamic inner PCA (DiPCA), PCA, and kernel PCA are used sequentially to extract dynamic, linear, and nonlinear characteristics from process data. By splitting the data space into perpendicular subspaces, the approach can track abnormal fluctuations across features. The findings show that PCA outperforms other methods for process monitoring.

Zhang et al., [2017] [118] presented a straightforward, swift, effective method for combining visual and infrared images, with a focus on retaining visual information and extracting infrared features. Initially, the infrared background is reconstructed using Bezier interpolation and quadtree decomposition. Bright features may be recovered from the infrared image by removing the rebuilt background. Refinement is then used to eliminate superfluous background information. Before merging with the visible image, improved infrared characteristics are adaptively muted to avoid overexposure. By combining infrared bright characteristics, this method preserves useful visual quality by preserving a large portion of the original visual information, allowing the fusion image to highlight important but previously unseen infrared objects. The approach generally outperforms various sample image fusion algorithms, according to experimental findings on widely used image datasets.

Aymaz and Köse [2019] [12] introduce a novel method for fusing multi-focus images, blending a combination of techniques including super-resolution. Initially, the super-resolution process enhances the contrast and details of low-resolution source images, transforming them into high-resolution versions. Next, the source images are decomposed using Stationary Wavelet Transform (SWT) into four sub-bands: LL (approximation), LH (horizontal details), HL (vertical details), and HH (diagonal details). Principal Component Analysis (PCA) is applied to each sub-band to select the maximum eigenvector, which is then used to fuse the images. The fused sub-bands are reconstructed using Inverse Stationary Wavelet Transform (ISWT). The quality of the fusion is assessed objectively by resizing the fused image to match the original size of the source images and evaluating various metrics, both with and without a reference image.

The results show that the technique produces low distortion, huge clarity, good visual perception, and clean edges. This hybrid approach demonstrates improved quality in image fusion, as evidenced by both visual inspection and quantitative analysis.

Hrosik et al., [2019] [39] introduced a technique to segment brain images to identify various primary tumors. It highlights abnormalities such as gliomas, metastatic adenocarcinomas, metastatic bronchogenic carcinomas, and sarcomas by combining the firefly method with k-means clustering. Testing on standard benchmark images showed superior performance compared to existing methods. Future research could focus on automatically determining cluster numbers and refining the fitness function by integrating spatial information into the segmentation process.

Kaur et al., [2018] [49] proposed a novel technique for classifying MR brain tumor images for selecting the most effective features swiftly. This technique combines the Fisher algorithm with the parameter-free Bat (PFree Bat) optimization algorithm to achieve this goal. Unlike the traditional Bat algorithm, which struggles with exploration, this modified version, known as the PFree Bat algorithm, enhances guidance by considering pulse frequency, global best, and local best positions. By eliminating the velocity equation and directly updating the Bat position, this improved algorithm facilitates quicker optimization. The method, coupled with the Fisher criteria, effectively selects the optimal feature set for brain tumor classification. The least squares (LS), SVM classifier uses these characteristics to classify different brain areas as high or low grade. Tenfold cross-validation on 95 regions of interest (ROIs) from the BRATS 2012 dataset is used to evaluate this feature selection technique. After a thorough evaluation of

different hybrid techniques, this technique demonstrates superior recognition rates within minimal time frames.

Subbiah Parvathy et al., [2020] [95] propose Enhanced Monarch Butterfly Optimization (EMBO) method to find the best fusion thresholds within the shearlet transform framework in a unique fusion strategy that incorporates deep learning concepts. Subsequently, deep learning algorithms provide feature maps that are necessary for the merging of low- and high-frequency sub-bands. In particular, Multimodal Image Fusion (MMIF) is facilitated by the use of the Restricted Boltzmann Machine (RBM). Training and testing are performed on a standard benchmark dataset comprising publicly available CT and MR images. The final fused image is produced by fusing the high-frequency and low-frequency components. An assessment of the suggested approach indicates that it is efficacious concerning measures such as spatial frequency (SF), entropy, fusion factor (FF), mutual information (MI), edge quality (EQ), standard deviation (SD), and correlation factor (CF).

2.5 REVIEWS ON IMAGE FUSION USING DEEP LEARNING AND ENSEMBLE DEEP LEARNING ALGORITHMS OVER MRI IMAGES

Liu et al., [2017] [65] employs an advanced CNN was trained using both the blurred and high-quality image patches to provide an encoding framework. This method is unique in that it trains a CNN model to provide fusion rules and activity level measurements simultaneously. This addresses a challenge encountered by existing fusion techniques. This research presents a revolutionary multi-focus image fusion technique. Experiments show that the recommended technique yields outstanding fusion outcomes in objective evaluation and visual attractiveness. Additionally, leveraging parallel

computing, the method demonstrates swift computational speeds suitable for practical applications. Furthermore, the experiments briefly illustrate the potential of the trained CNN model for addressing other types of image fusion challenges.

Hsu et al., [2009] [40] presents an artificial neural network approach to address the challenge of multi-sensor image fusion. Enhancing the conventional IHS method, we introduce the concept of region-based fusion, aiming to utilize varying parameters across different regions, accounting for changes in time or weather conditions. Given the nonlinear relationship between environmental factors and fused parameters, we employ artificial neural networks to tackle this complexity. Moreover, the automated estimation of fused parameters enables adaptive adjustments to accommodate diverse environmental states. This architecture not only holds promise for numerous applications but also exhibits versatility across various fields.

Yan et al., [2020] [116] introduce a unique neural network-based supervised multi-view hash model that improves multi-view data. This strategy, which combines deep learning and multi-view approaches, is innovative. Their model actively investigates inter-view interactions by using an efficient view stability assessment approach, which impacts the optimization trajectory of the network in its entirety. Additionally, to make use of the advantages of both convolution and multi-view techniques, they have developed several multi-data fusion algorithms within the Hamming space. To prevent too much computational value when retrieval improvement is performed, they create an additional structure called a memory network that is trained concurrently. Their method surpasses existing single-view and multi-view hashing algorithms in extensive CIFAR-10, NUS-WIDE, and MS-COCO dataset studies.

Ma et al., [2020] [68] introduce Dual-Discriminator Conditional Generative Adversarial Network (DDcGAN) which is a new end-to-end model intended for the merging of visible and infrared images of different resolutions. This strategy puts a generator and two discriminators in conflict. Fused images should closely match genuine imagery, according to the generator, guided by a custom content loss, in order to deceive both discriminators. These discriminators, attempt to identify any structural differences as well as content loss between the two source photos and the merged image. The merged image must thus maintain both the heat radiation from the infrared image and the textural characteristics of the visible image. Additionally, can effectively combine sources with varying resolutions, for example, a high-resolution visible image and a low-resolution infrared image. The downsampled fused image and the infrared image are guaranteed to have similar quality by DDcGAN. In previous approaches, it is typical to have problems like obvious texture detail loss or blurring of heat radiation information. This strategy helps avoid these problems. In addition, they extend the use of DDcGAN to the integration of multi-modal medical images at varying resolutions. For example, it combines two images: one from low-resolution positron emission tomography and one from high-resolution magnetic resonance.

Wang et al., [2022] [108] suggest an attention-guided generative adversarial network (AMFNet) designed specifically for the merging of several models of images. The three components that make up the generator network within AMFNet are an information refinement network (which refines image feature maps), an attention network (which finds extensive dependencies within image representations), and a fusion network (which integrates the inform. To encourage focus on the most discriminative areas of the

multi-modal source images by the discriminator, they also introduce the convolutional block attention module. Experiments, both qualitative and quantitative, were out on a range of publicly available datasets showing that their technique outperforms other approaches in terms of visual quality and image detail retention.

El Boustani et al., [2020] [24] introduced a computerized method was created to extract and categorize brain tumors in medical images. This system demonstrates remarkable efficiency in processing Magnetic Resonance Images (MRI), boasting high detection rates. CNN is used for classification, while thresholding is used for image segmentation in the design. Additionally, a large-scale experiment was conducted to assess the efficacy of our strategy across several optimizers utilizing a huge dataset of MRI brain images. The results indicated that the RMSprop optimizer not only converges more rapidly but also achieves the highest accuracy when compared to other optimization methods.

Wen et al., [2020] [112] propose to create the clarity map from the source images, a dual-channel convolutional network is used. The clarity map is subsequently reduced using morphological filtering. Ultimately, the clear areas of the source images were combined to form the fusion image. Results from experiments show that their approach works better than many other fusion procedures available, both in terms of quantitative evaluations and visual quality.

Amin-Naji et al., [2019] [9] suggested a unique approach that uses ensemble learning methods and CNNs. Using many models and datasets is a better option than relying just on one. Enhancing model diversity and reducing overfitting problems on the training dataset are the objectives of ensemble learning. It's evident that the collective results of an ensemble of CNNs surpass those of a single CNN. Additionally, their

method introduces a straightforward yet effective type of multi-focus image dataset. By simply rearranging patches within multi-focus datasets, they enhance accuracy significantly. Thus, their approach incorporates a new network comprising three CNN models trained on three distinct datasets to construct the initial segmented decision map. Experiments employ numerous real multi-focus test images and compare results using both quantitative and qualitative metrics. The experimental outcomes demonstrate that the CNN-based network yields superior accuracy and decision mapping without the need for post-processing algorithms, outperforming existing state-of-the-art multi-focus fusion methods that rely heavily on such algorithms.

Zhang et al., [2021] [116] propose a novel multi-focus image fusion technique using generative adversarial networks (GANs) outfitted with gradient joint restrictions and adaptive features. Considering the variance of recurred blurring, assess the focus of source pixels, they use an adaptive decision block in this model. To ensure that the generator generates a fused output that matches the distribution of focused source images, this block dynamically directs optimization for a specially customized content loss. To improve texture details, they employ an adversarial game to match the joint gradient map from the source images and the fused output gradient map. Interestingly, their model can be trained without supervision, thus it doesn't need ground-truth fused images. They also provide a new dataset for benchmarking that consists of 120 high-quality multi-focus image pairings. The method's superiority over state-of-the-art alternatives is shown by experimental findings, which show improvements in both quantitative measurements and subjective visual impacts.

Zhao et al., [2023] [120] introduced the Multi-Discriminator Hierarchical Wavelet Generative Adversarial Network (MHW-GAN) for multimodal image fusion. Initially, they develop a Hierarchical Wavelet Fusion (HWF) module within MHW-GAN's generator. To reduce information loss across distinct modality layers, this module combines feature data at many scales and levels. Second, to reduce edge information loss, they integrate an Edge Perception Module (EPM) to aggregate edge information from several modalities. Thirdly, they use three discriminators and adversarial learning amongst them to control the creation of fusion images. To distinguish the fusion image from the joint edge image and the two source images, the three discriminators are used, respectively, while the generator attempts to deceive them by creating a fusion image. The resulting fusion image incorporates both structural and intensity information via adversarial learning. Experimentation on various multimodal image datasets, including public and self-collected ones, illustrates the algorithm's superiority over previous methods in both subjective and objective evaluations.

Khan et al., [2019] [50] proposed is a CADx system utilizing Multi-View Feature Fusion (MVFF) technique for mammogram classification, integrating features from four views. This comprehensive tool comprises three stages: initial classification of mammograms as abnormal or normal, followed by classification of mass or calcification in the second stage, and ultimately, distinguishing between malignant and benign cases in the final stage. Each view undergoes feature extraction via Convolutional Neural Network (CNN) models independently. The final layer is created by fusing these extracted characteristics, which provides the most accurate forecast.

Ting et al., [2020] [101] presented a fresh machine learning-driven fusion technique to integrate spatially distant ultrasound views, namely, apical and parasternal. This approach combines: 1) using a mutual information neural estimating network to maximize the mutual information between the source and fused images; and 2) an autoencoder framework for the creation of fused images. Experimental assessments demonstrate encouraging outcomes, compared to state-of-the-art approaches, the fused image improves signal-to-noise ratio by 18.23 dB and contrast-to-noise ratio by 21.76 dB.

Kang et al., [2020] [48] proposed a technique for merging brain PET and MRI images utilizing a Tissue-Aware conditional Generative Adversarial Network (TA-cGAN). In this method, the fusion process is conceptualized as a machine-driven adversarial task, balancing the retention of PET's color information and MRI's anatomical details. The tissue label map from the MRI image conditions the generator and discriminator in TA-cGAN, undergoing alternate training with a combined loss function. Through extensive experimentation, the method is shown to enhance the fused image's anatomical accuracy while effectively preserving the color information from the PET.

Hu et al., [2019] [41] introduced an innovative method for segmenting brain tumors, employing a Multi Cascaded Convolutional Neural Network (MCCNN) in conjunction with fully connected conditional random fields (CRFs). Two primary phases are involved in the segmentation process. First, a multi-cascaded network design was developed, which combines intermediate results from several linked components. This method uses multi-scale features to perform coarse segmentation while taking into consideration local label dependencies. After that, they used CRFs to include geographical contextual data, which helped to improve segmentation by removing false

positives. Three segmentation models were trained using axial, coronal, and sagittal image patches and then merged to get the final segmentation result.

2.6 MOTIVATION FOR THE PROPOSED RESEARCH AND KEY DIFFERENCE BETWEEN PROPOSED WORK AND THE EXISTING WORKS

Li et al., [2013] [55] merges various images to create a highly information-fused image rapidly and efficiently. Currently, a picture is divided into two scales: a detail layer for small-scale characteristics and a base layer for large-scale intensity variations. The suggested solution uses multi-scale decomposition.

Li et al., [2012] [56] has proposed a new method for dictionary learning, which they have named Dictionary Learning with Group Sparsity and Graph Regularization (DL-GSGR). At first, we used graph regularization to represent the geometric structure of atoms. Subsequently, DL-GSGR is a method that combines group sparsity with graph regularization. Dictionary update and alternating group sparse coding are used to solve it. This approach successfully maintains the learned dictionary's small group coherence, facilitating efficient group sparse coding of any signal. Finally, they apply group sparse representation with DL-GSGR to tasks such as 3-D medical image denoising and image fusion. However, the existing work has not ensured multi-view image fusion. Then proposed method is specifically aimed at enhancing the performance of multi-view image fusion.

Liu et al., [2016] [64], proposed a medical image fusion technique that uses a pulse-coupled neural network (PCNN) and the gradient minimization smoothing filter (GMSF) to enhance the transmission of detailed information from the source images to the fused image. The existing method has issue with noise rates and computational complexity. The proposed method has resolved the noise problem through Adaptive

Median Filter (AMF) and the proposed deep learning algorithm avoids the computational complexity.

Singh and Khare [2014] [93], describe a novel multimodal medical image fusion approach that utilizes a new multilevel Daubechies complex wavelet transform (DCxWT) and follows the notion of multiresolution. The existing method analyzed using CT bone images and the proposed method analyzed MRI brain images for improving the image fusion performance.

2.7 SUMMARY

This chapter examines research works conducted by various authors to explore the diverse techniques employed in image fusion. Numerous studies have sought to enhance image fusion performance through various means. Specifically focusing on MRI images, researchers have employed pre-processing, optimization, feature extraction, and classification techniques. The aim is to extract concealed insights from an MRI brain tumor database. Throughout the image fusion process, challenges arise particularly in segmentation and image fusion. This chapter delves into noise reduction, segmentation, feature extraction, and image fusion across various databases. These processes are referred and analysed from different authors' perspectives to gain deeper insights into specific directions. Additionally, the chapter provides a comprehensive overview of the developed approaches, detailing their characteristics and performance metrics. Each method is accompanied by its own set of demerits and advantages, and the results are analysed thoroughly examined and discussed.

Chapter 3

CHAPTER 3

MODIFIED-UNET FOR SEGMENTATION OF BRAIN IMAGES

3.1 INTRODUCTION

Brain tumors represent a prevalent affliction of the nervous system, posing significant risks to human health and often lead to fatal outcomes. Glioma, in particular, stands as one of the most lethal intracranial tumors, categorized into high-grade glioma (HGG) and low-grade glioma (LGG), with an average life expectancy of approximately two years for patients progressing to HGG (Goetz et al., 2015) [32]. Various imaging modalities, including Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), and Single-Photon Emission Computed Tomography (SPECT), have been utilized for brain tumor investigations. Among these options, MRI has become the main method for diagnosing and treating glioma due to its benefits, including better contrast for soft tissues, the ability to capture several parameters, the ability to image in any direction, and is non-invasive (Menze et al., 2010) [76]. Moreover, T1-weighted (T1), T1-weighted with contrast enhancement (T1c), T2-weighted (T2), and Fluid Attenuated Inversion Recovery (FLAIR) are among the various modalities that may be acquired with the use of MRI. These modalities provide different insights into the properties of brain tumors. For the purposes of medical diagnosis, surgical planning, and therapy design, brain tumor segmentation must be done accurately. It's critical to differentiate between normal brain tissues like gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) and tumor tissues like necrosis, edema, enhancing core, and non-enhancing core. However, achieving precise segmentation poses significant challenges due to several factors. Firstly, the shape,

location, appearance, and size of gliomas exhibit considerable variation among patients. Secondly, gliomas typically infiltrate surrounding tissues, leading to blurred boundaries (Menze et al., 2014) [75]. Thirdly, the segmentation problem is further exacerbated by image distortion and noise caused by many variables, such as imaging instruments or acquisition techniques.

Manual segmentation is a laborious and time-intensive process, and if the individual conducting the delineation of the region of interest (ROI) lacks adequate training, it often results in subpar segmentation outcomes (Chen et al., 2014) [19]. Thus, automatic or semi-automatic segmentation techniques have drawn the interest of researchers. Approaches based on generative models and discriminative models are used to widely classify existing automated or semi-automatic segmentation algorithms (Kamnitsas et al., 2017) [47], (Tustison et al., 2015) [103]. The learning process for generative models is somewhat difficult since it depends on past knowledge about brain architecture. While alternative techniques needing previous information are discussed, a probabilistic image atlas serves as the foundation for a typical generative model for MR brain images (Meier et al., 2014) [74], (Reza and Iftekharuddin 2014) [83]. In contrast, discriminative models are better suited for multi-category identification issues because they often utilize a discriminative classifier to convert local characteristics into class probability and frequently depend on low-level image information (Mamelak and Jacoby 200) [70]], (Krizhevsky et al., 2012) [52]. Typical discriminative models include conditional random fields (Saut et al., 2014) [85], (Girshick et al., 2014) [31], random forests, support vector machines, decision forests, etc.

3.2 Background Study

3.2.1 U-NET

The U-Net architecture is a recent method employed in medical image segmentation. It stands out alongside the Fully Convolutional Network (FCN) as a prominent deep learning approach in therapeutic image segmentation. Within this domain, U-Net has consistently demonstrated superior performance (Ronneberger et al., 2015) [84]. The U-Net design features a U-shaped arrangement, where the left segment serves as the encoder and the right segment handles the decoder tasks. An essential component of this design is that the encoder replicates the layers of the decoder. This architecture enables the resultant feature map to preserve characteristics from all layers. Moreover, the result is enhanced by including features from different phases while maintaining spatial information.

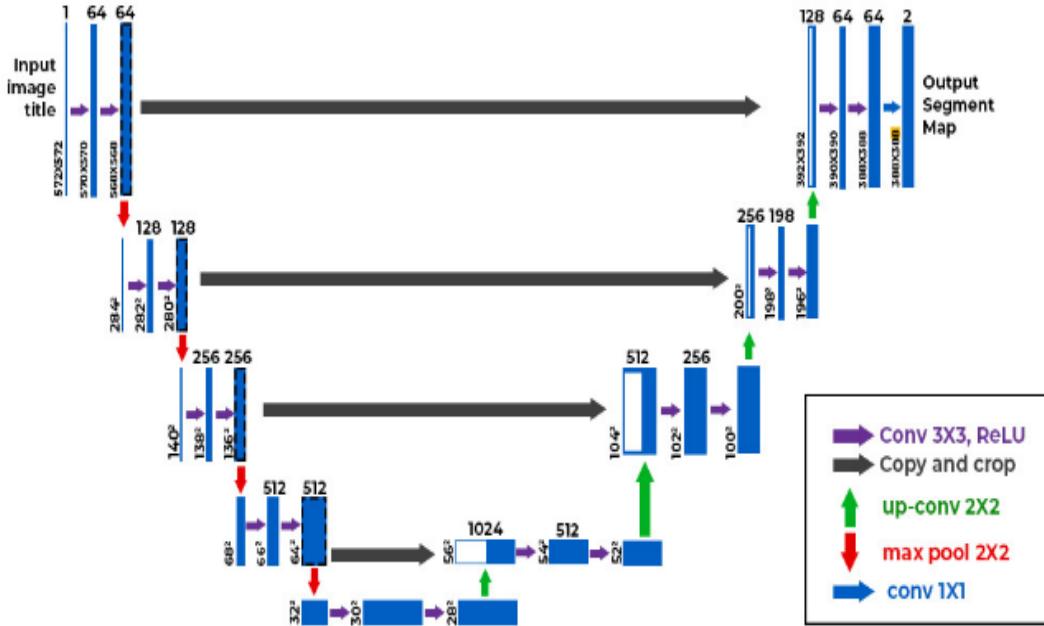


Fig 3.1 UNet architecture

The primary purpose of this design is medical image segmentation. The UNet architecture serves as an end-to-end fully convolutional neural network, featuring two paths known as the contraction path and the expansion path. Its structure resembles a 'U' shape. Within the contraction path, convolution operations are performed, followed by maximum pooling operations with a specific stride size. Conversely, the expansion path involves transposed convolution operations.

In the UNet architecture, two 3x3 convolutions are followed by ReLu and 2x2 maximum pooling operations with a stride of 2 for downsampling. A 2 x 2 transposed convolution method is utilized in the upsampling step to decrease the feature channels. Additionally, the UNet design incorporates convolutional route skip connections. These linkages facilitate the transfer of features from the contracting route to the expanding path, assisting in the restoration of spatial details that were lost during the downsampling processes. This design results in segmentation that is both rapid and more precise than other techniques of dividing or categorizing.

1. UNet stands out from other deep architectures due to its appealing features:
2. It incorporates a blend of convolutional, pooling, and up-sampling layers.
3. Rather than traditional activation functions like Tanh, logistic, arctan, or Sigmoid, it utilizes the ReLU function, mitigating the risk of the vanishing gradient problem.
4. It exhibits faster training compared to other deeper architectures.

Figure 3.1 illustrates the UNet architecture renowned for its capacity to process high-resolution images and generate precise segmentation maps. It is particularly effective for multi-class image segmentation tasks, accommodating numerous classes and

generating pixel-level segmentation maps for each class. Multi-channel feature maps have boxes indicating the number of channels at the top. Each box's bottom left corner shows the feature maps' x and y dimensions. Duplicated feature maps are white boxes. Convolution and pooling are shown by blue arrows, while up-sampling operations are indicated by green arrows.

The primary methods for dissecting brain cancer are based on MRI, including the 3D separation approach, and on slice-centered 2D separation. These techniques have been highlighted in studies by (Kamnitsas et al., 2017) [47] and (Wang et al., 2018) [106]. In the context of MRI-based 3D separation, there is a scarcity of annotated training data, posing challenges in increasing the number of objects. Particularly, issues with large grid boundaries and storage constraints make the training of 3D models difficult, as noted by (Havaei et al., 2017) [36].

A detailed multipath CNN has been proposed to delineate the brain cancer region within the 2D segmented objects of MRI images, as discussed by (Shen et al., 2017) [92] and (Baid et al., 2019) [15]. Additionally, dual preprocessing stages are considered to manage the variability in input data sessions, as highlighted by (Wang et al., 2019) [107] and (Myronenko 2019) [78]. An approach utilizing limit-oriented FCN is suggested to advance the segmentation process, as recommended by (McKinley et al., 2019) [72]. Subsequently, a 3D system named Deep Medic has been developed, using a two-path architecture to combine multi-scale feature maps and contextually.

A novel approach, termed as modified UNet (mUNet), is introduced in this study, aimed at restructuring the residual path and mitigating the limitations of the conventional UNet architecture, as proposed by (Abd Khalid et al., 2020) [1]. The mUNet framework

incorporates features from the residual path into the skip connections, enabling (1) the avoidance of redundant feature duplication, (2) the extraction of more significant features with higher weights and greater data relevance for large datasets, and (3) the extraction of comprehensive features for minor objects through optimal aggregation strategies, as outlined by (Bennai et al., 2020) [16]. Comparative analysis with CNN and traditional UNet demonstrates that the modified UNet concept effectively preserves edge and morphological information of the data entities, as observed by (Daimary et al., 2020) [22].

Performing image segmentation is a common technique, particularly in clinical image analysis, where it accurately partitions images with limited training data, as highlighted by (Yousef et al., 2023) [115]. This capability has positioned UNet as highly effective in clinical imaging, it has become the standard medical imaging segmentation tool, as noted by (Khan et al., 2019) [50]. UNet's success is evident in its widespread usage across major imaging modalities such as CT scans, MRI, X-rays, and microscopy. While UNet is primarily a segmentation tool, there have been instances of its application in other domains, as observed by (Ting et al., 2020) [101]. The potential of UNet continues to grow with advancements in image modalities, allowing for enhanced accuracy. However, the computational complexity of UNet remains unchanged.

3.2.2 *Drawbacks of UNet*

The existing UNet method has problem with loss of spatial resolution and some details of the input images. UNet affects training and accuracy of the MRI images (Kang et al., 2020) [48]. Another limitation is the requirement of manual segmentation for MRI images, which is time-consuming. Also, in previous work, segmentation accuracy and

multi view image fusion results are not ensured effectively. To overcome the above mentioned problems, M-UNet is proposed in this work

3.2.3 *Solution*

Achieving accurate segmentation using the MUNet algorithm is the primary objective of this work. While numerous research endeavors and methodologies have been introduced, they often fall short in ensuring segmentation accuracy. The key contribution of this research lies in attribute mining and segmentation refinement. The proposed method aims to enhance segmentation accuracy by employing a robust algorithm tailored for the provided MRI image dataset.

3.3 PROPOSED METHODOLOGY

This section elaborates on the Modified UNet (MUNet) approach for segmentation. The MUNet algorithm proposed in this study aims to prevent redundancy in less crucial attributes and extract weightier and more determinative data from the provided brain images. Additionally, it endeavors to capture comprehensive attributes for minor objects through optimal assembly strategies. Comparative analysis with CNN and the traditional UNet demonstrates that the Modified UNet concept adeptly retains edge and morphological data of the data items, as observed by (Hu et al., 2019) [41].

Figure 3.2 shows the proposed system's detailed block diagram.

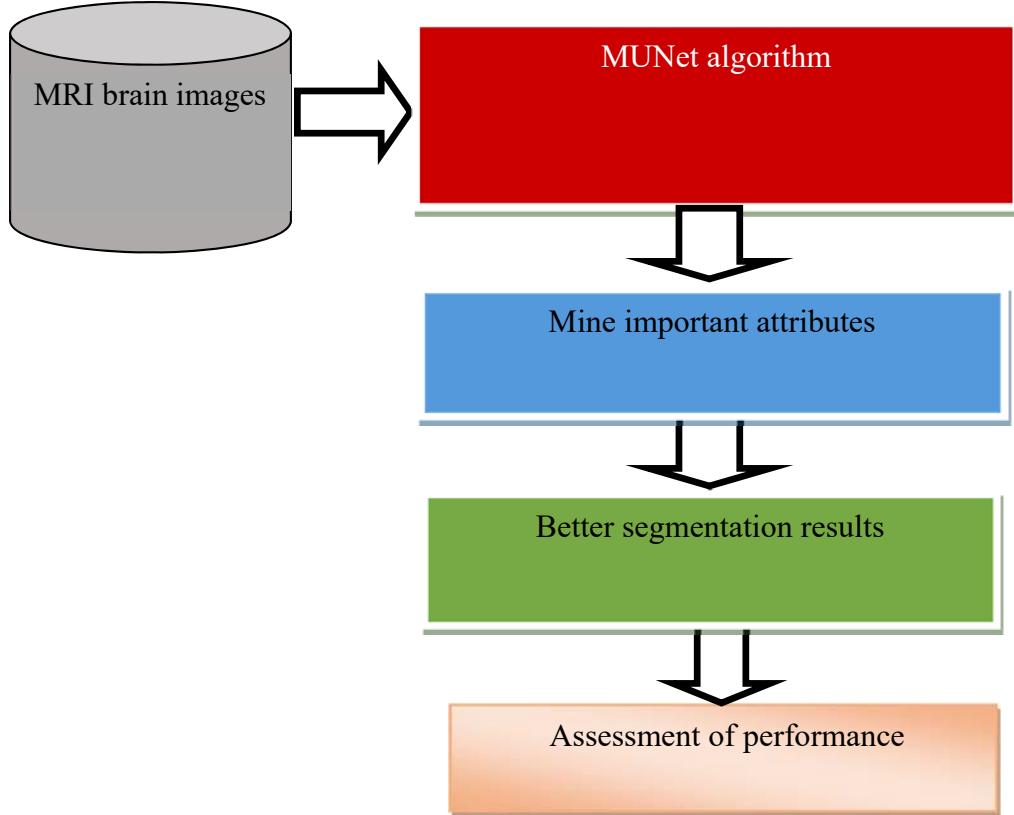


Fig 3.2 The proposed system's overall block diagram

3.3.1 Modified Unet (MUNet)

The system being proposed employs a residual pathway to prevent the duplication of low-resolution attribute plot data. This suggested network, in contrast to current techniques, keeps the leftover route right after assembly. This approach dynamically constrains and incorporates extra convolution layers into the skip association using high-resolution edge data from the attribute plots that traverse across the skip connections. To evaluate this adaptive filtering's effectiveness, permeation data (PD) is defined.

$$PD = \begin{cases} -0.5, & \text{if } FMy(p, q) < 0.01 \\ \sum \frac{FMx(p, q)}{FMy(p, q)}, & p, q \in \text{object tag}(p, q), \text{Otherwise} \end{cases} \quad (3.1)$$

Here,

FMx – The skip link behind the residual route has a standardized attribute plot.

FMy- The residual track is shown after the standardized attribute plot.

All standard attribute plots have a range of [0, 1].

When $FMy(p,q) < 0.01$, the permeability rate may be set to -0.5, the skip link does not seem to have any important properties.

Several proven modifications were applied to the modified UNet. Initially, a residual connection was integrated into every encoder block, featuring three convolution layers per block. After receiving the output from the preceding encoder block, the max pooling layer in the final convolutional layer was updated using a stride of 2. The UNet model was modified via stacking. A dual UNet technique looped the UNet's output to its input. The repeating design style saves GPU memory and parameters. In the first and subsequent iterations, the cost function included the cumulative UNet cost. By incorporating intermediate outcomes in the cost estimation, training efficiency was significantly improved.

3.4 RESULTS AND DISCUSSION

The Brain Tumor Segmentation (BRATS) datasets were employed for experimentation. These datasets' training sets include 3D MRI images from 285 individuals diagnosed with brain tumors, with 210 having High-grade Glioma (HCG) and 75 with Low-grade Glioma (LCG). In BRATS 2017 and 2018, 46 and 66 people with unclear tumor grades provided 3D MRI images for validation. Brain data from each person is in four modalities: R1, R1c (after contrast), R2, and Aptitude. These modalities have undergone skull-stripping, resampling, and Coregistration. The records consist of

four distinct labels representing different tumor types: enhancing tumor, edema, necrosis, and background. Observations are classified into three categories, namely the overall tumor region, tumor core area, and increasing tumor area, for assessment. Experts in the field have determined accurate and definitive information by manually dividing it into segments. The public cannot access the segmentation labels for the validation sets. To get measurable evaluations such as the Dice and Jaccard coefficients, contributors are required to submit their grid-based findings to the BRATS online evaluation platform (Arora et al., 2021) [11].

3.5 PERFORMANCE METRICS

Accuracy

Accuracy serves as a crucial metric in assessing the effectiveness of classification models, offering insights into the proportion of accurate forecasts the model made. The computation of accuracy is typically conducted through the following mathematical expression.

$$\text{Accuracy} = \frac{\text{(number of correct prediction)}}{\text{(Total number of predictions)}} \quad (3.2)$$

Dice Similarity Coefficient (DSC)

The precise ratio between the actual tumor and nontumor pixels that are accessible and the predicted tumor and nontumor pixels is ascertained using DSC.

$$\text{DSC} = \frac{\text{(2TP)}}{\text{(FP + 2TP + FN)}} \times 100 \quad (3.3)$$

Jaccard similarity index (JSI)

JSI is estimated using the usual equation to determine the percentage of similarity between the number of expected tumor pixels and the actual tumor pixels in the area of interest.

$$\text{JSI} = \frac{(\text{TP})}{(\text{TP} + \text{FN} + \text{FP})} \times 100 \quad (3.4)$$

Time Period

Elapsed time is a measure of how long it takes to finish the segmentation process, expressed in milliseconds or seconds.

Figures 3.3 and 3.4 show the input and segmented images, respectively. The suggested MUNet's output is seen in Fig. 3.6.

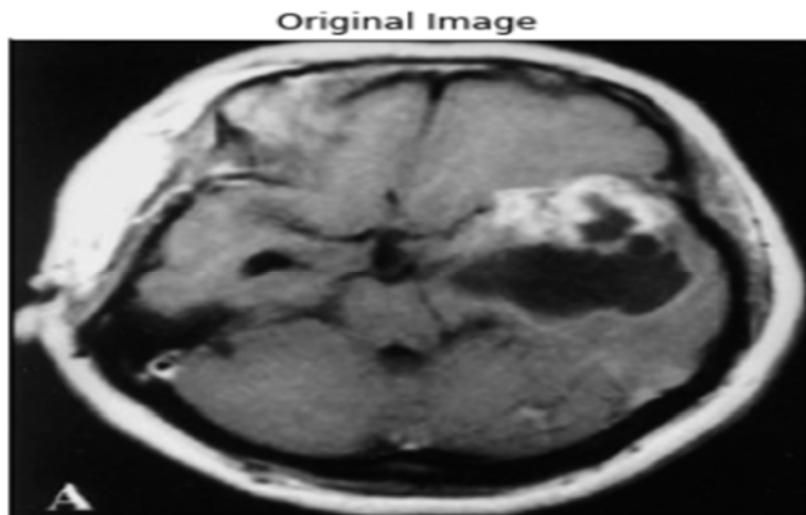


Fig 3.3 Input image

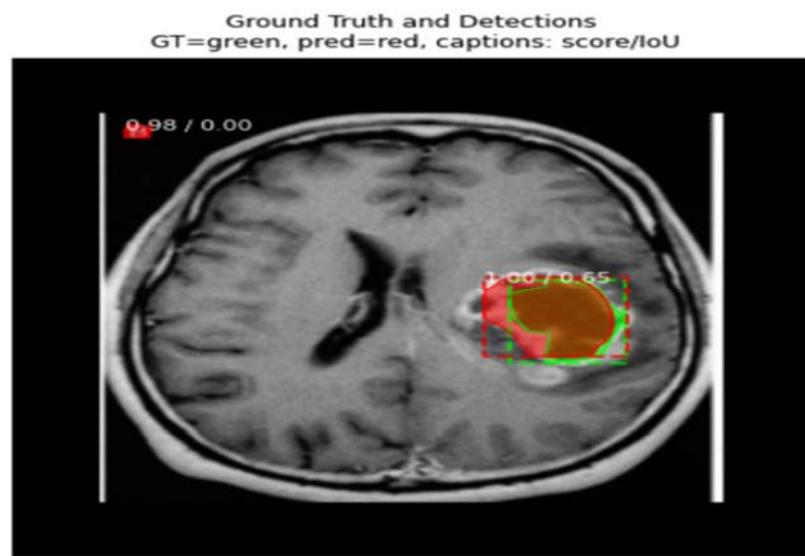


Fig 3.4 Segmentation image

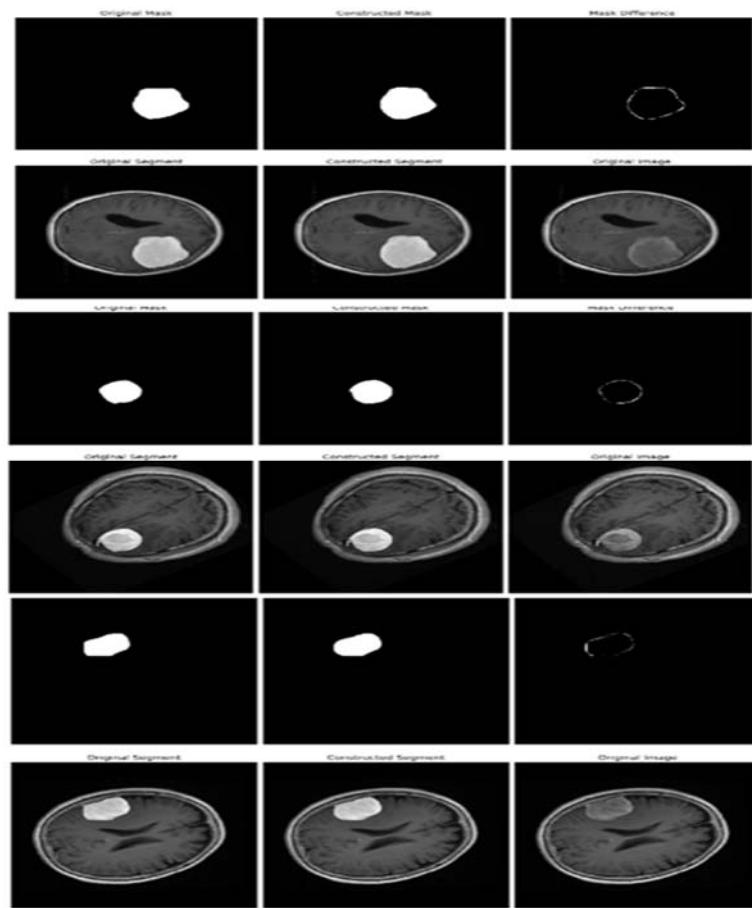


Fig 3.5 Output of the MUNet

Fig 3.5 shows how the MUNet provides more accurate segmentation results. MUNet is used for dividing the brain MRI image into distinct regions and each representing a specific tissue type or structure. Because no padding is being utilized, the output size is less than the input size. The combination of the two paths enables MU-net to learn both global and local features and to achieve high accuracy in segmentation tasks.

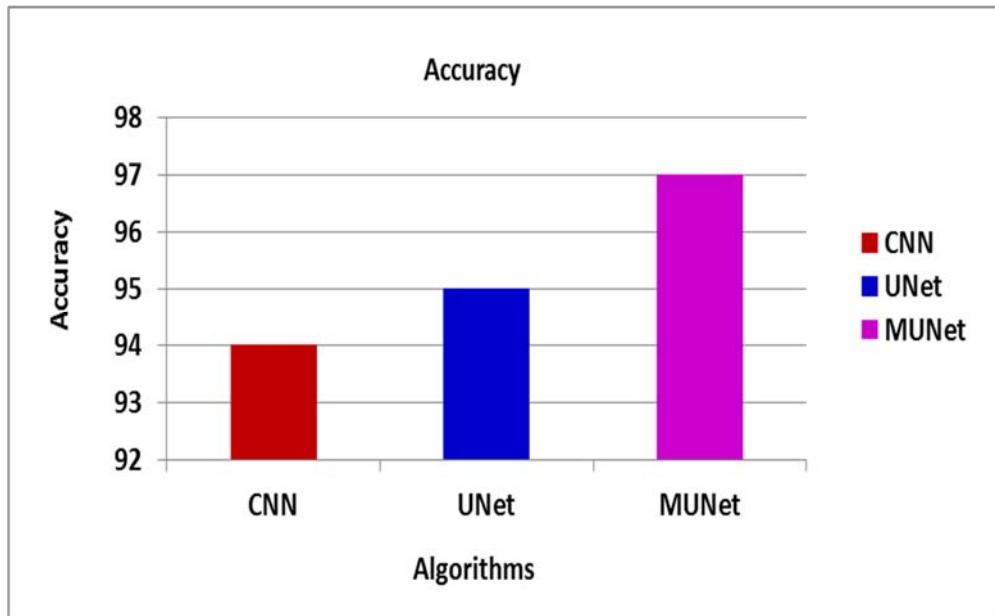


Fig 3.6 Accuracy

Figure 3.6 illustrates a comparison of accuracy metrics between established and proposed methodologies. While conventional approaches like CNN and the UNet algorithm demonstrate lower accuracy, the proposed MUNet algorithm achieves higher accuracy in segmenting MRI brain images. This outcome underscores the enhancement in segmentation facilitated by the MUNet algorithm through its optimal feature utilization.

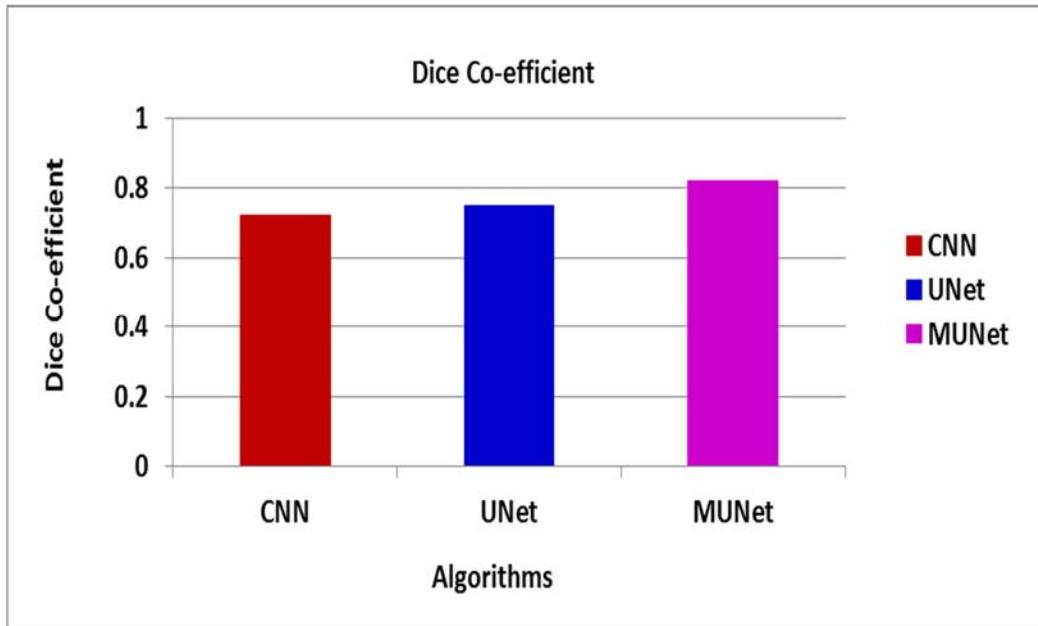


Fig 3.7 Dice Co-efficient

Fig 3.7, shows the comparison of Dice Co-efficient metric for existing and proposed methods. The current methods, such as the CNN and UNet algorithms, yield lower Dice Coefficients compared to the proposed MUNet algorithm, which achieves higher Dice Coefficients for the provided MRI brain images. Consequently, these results indicate that the proposed MUNet algorithm enhances the segmentation process by leveraging optimal features.

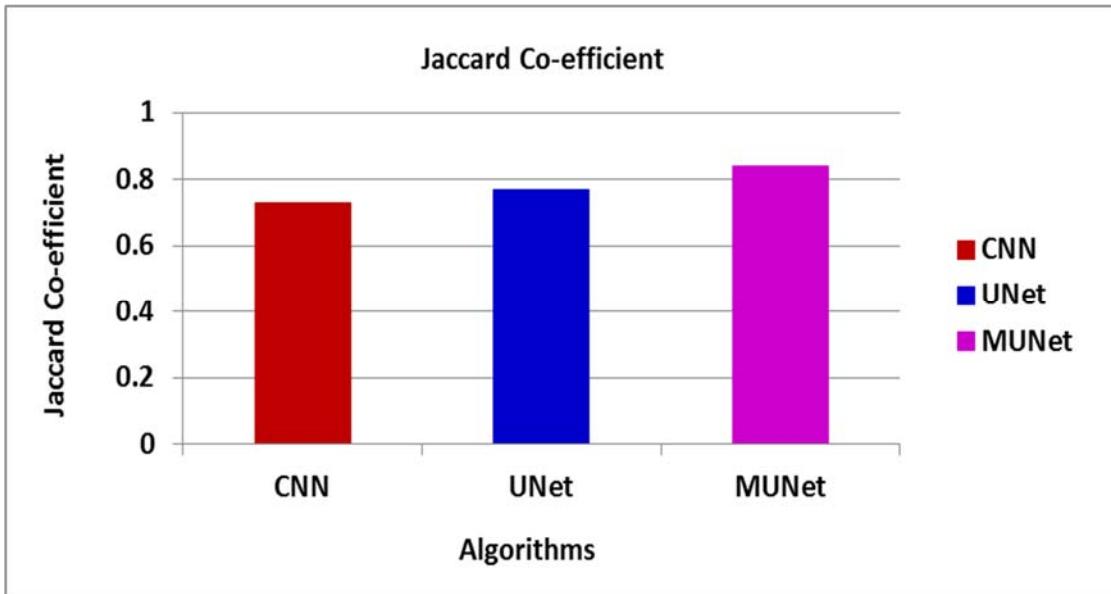


Fig 3.8 Jaccard Co-efficient

Fig 3.8, shows the comparison of Jaccard Co-efficient metric for existing and proposed methods. The current methods, like the CNN and UNet algorithms, yield lower Jaccard Coefficients compared to the proposed MUNet algorithm, which achieves higher Jaccard Coefficients for the provided MRI brain images. Consequently, these findings suggest that the proposed MUNet algorithm enhances the segmentation process by utilizing optimal features.

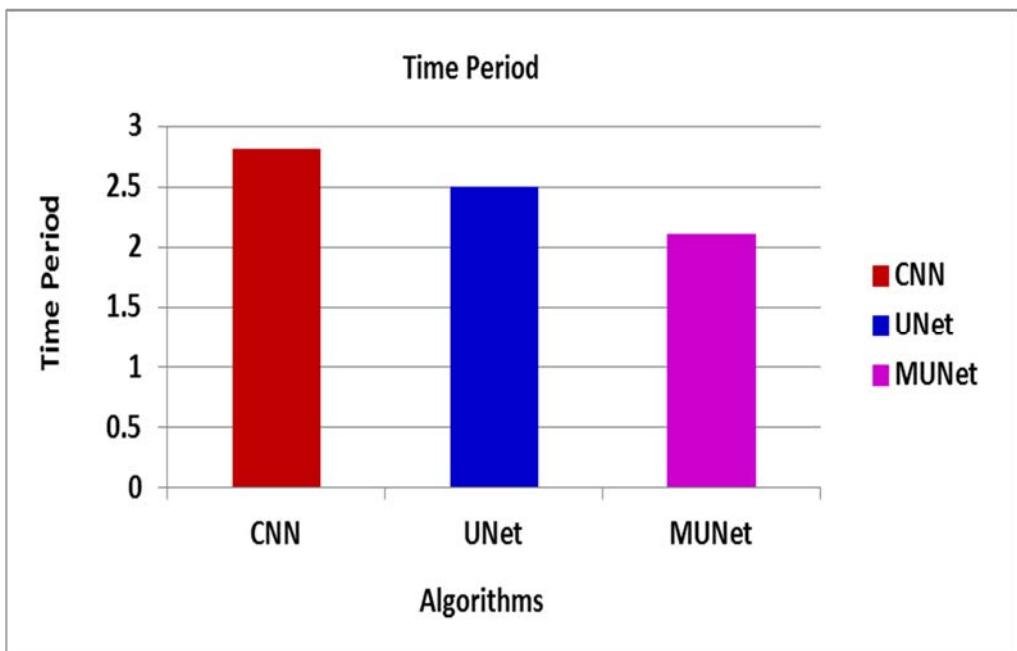


Fig 3.9 Time Period

Figure 3.9 illustrates a comparison between the time period metrics of current and suggested approaches. While established methods like CNN and UNet algorithms exhibit greater time complexity, the proposed MUNet algorithm demonstrates reduced time complexity in segmenting MRI brain images. This outcome suggests that the MUNet algorithm enhances the segmentation process by leveraging optimal features.

3.6 SUMMARY

In the first work, modified UNet is proposed for more accurate segmentation results. The MUNet algorithm is introduced to improve attribute mining in brain images by avoiding redundant replication of less significant attributes. It focuses on identifying attributes with greater importance and higher data determination, utilizing assembly tasks optimally. By combining residual pathways and an object-dependent up-sampling prototype, the proposed system effectively avoids duplicating low-resolution data, learns

to extract ever more sophisticated global properties from small object inputs and produces higher-level feature maps that highlight high-resolution edge features of larger data. The Dice Coefficient, Jaccard Coefficient, and accuracy are used to evaluate both current and suggested approaches. The outcomes demonstrate that the modified UNet approach yields the most promising results. Moreover, the proposed technique doesn't require any preprocessing steps, making it easily applicable to images of other organs with poor contrast. Overall, the results indicate that MUNet significantly enhances image segmentation performance.

Chapter 4

CHAPTER 4

MULTI-MODAL IMAGE FUSION USING DEEP LEARNING

ALGORITHM

4.1 INTRODUCTION

The main aim of this research work is the multi modal image fusion using MRI brain images. There are several research and methodologies introduced but feature extraction is not done effectively. The existing approaches has drawback with extraction of relevant and important features (Liu et al., 2014) [67]. To address the aforementioned challenges, this study introduces the Modified Particle Swarm Optimization-Convolutional Neural Network (MPSOCNN) approach aimed at enhancing image quality holistically. The fundamental contributions of this study are mostly focused on image fusion, segmentation, feature extraction, and classification. By employing advanced algorithms tailored to the image dataset at hand, the proposed methodology aims to yield heightened precision in its outcomes.

4.1.1. Deep Learning (DL)

Recently, in a variety of computer vision and image processing applications, deep learning (DL) has achieved notable breakthroughs. Over the last three years, research on integrating DL approaches into pixel-level image fusion has been more active. The incorporation of DL models into image fusion, as proposed by (Li et al., 2021) [57], aims to introduce novel approaches to the process. By establishing a training database with successful fusion outcomes, this method aims to generate a new model capable of batch processing multiple medical images. This approach is particularly applicable to

multi-modal image fusion, offering enhancements in both efficiency and accuracy of image processing.

(Tan et al., 2021) [99] propose the MLEPF decomposition approach for fusing multi-modal medical images while preserving the edges at several levels. They propose employing an MLEPF model that utilizes weighted mean curvature filtering to separate a multi-modal medical picture into three layers: fine (FS), coarse (CS), and base (BS). After that, an energy attribute fusion technique fuses the BS layers and a gradient-domain Pulse-Coupled Neural Network (PCNN) merges the FS and CS layers. Creating the fused image from three fused layers.

A novel image fusion framework, termed IFCNN introduced by (Zhang et al., 2020) [119], is proposed based on convolutional neural networks. Researchers use two convolutional layers to extract important characteristics from numerous input images, inspired by transform-domain image fusion methods. Depending on the input images, the convolutional features from these images are merged using a suitable fusion rule (e.g., element-wise max, min, or mean). A useful fusion image is produced by reconstructing the merged features using two convolutional layers. Since the model is completely convolutional, it can be trained without post-processing. Using an extended RGB-D dataset, a massive multi-focus picture dataset was created to train the model. The ground-truth fusion images in this collection are more diversified and bigger than those in previous image fusion datasets.

A two-stage convolutional neural network fusion approach is employed to achieve precise decision mapping and address the challenge of indistinct fusion boundaries (Fan et al., 2017 [26]; Gai et al., 2020) [29]. In the initial stage, an optimized decision

map is generated by training an upgraded dense network to categorize whether the image patch is in focus or defocused, then applying the fusion rule. Additionally, a diverse dataset containing multiple versions of blurred images is created to enhance the network's generalization capabilities. Subsequently, EDGAN refines boundaries in the second step. Boundary-deblurred images are refined using five loss functions. These two processing steps provide a detailed image with well-defined fusion boundaries.

4.2. PROPOSED METHODOLOGY

Initially, the database is collected which contains multi modal brain images. Efficient brain tumor segmentation is conducted through the MUNet algorithm, followed by feature extraction utilizing the CNN algorithm to derive more comprehensive features. MRI brain scans are analyzed from various perspectives, and images from different views are fused to enhance clarity. The fusion process combines the features extracted from these images, resulting in a clearer composite image. Finally, the image classification is done for obtaining more accurate results.

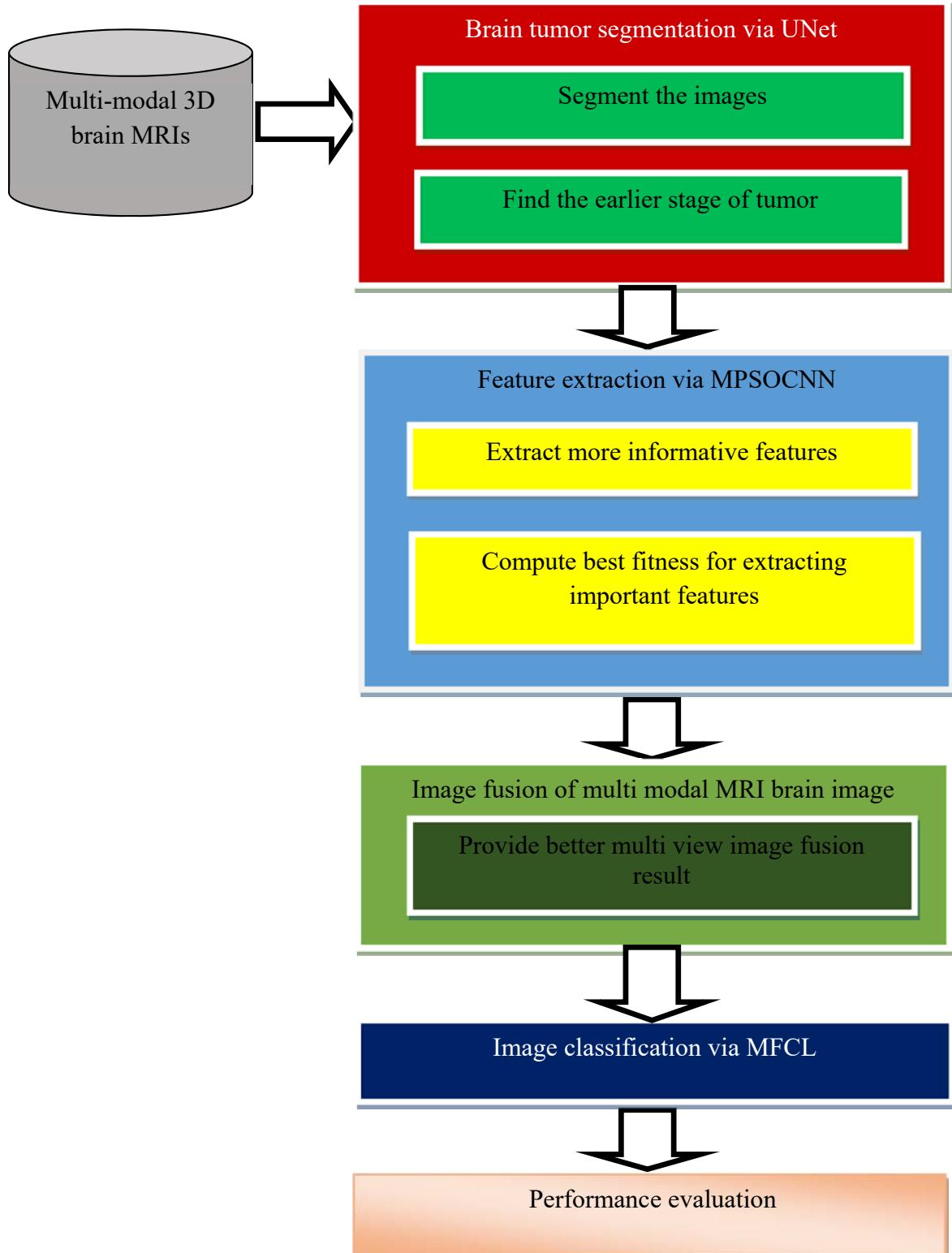


Fig 4.1 Overall block diagram of the proposed system

4.2.1. Feature Extraction Using CNN

Three convolutional layers collect picture properties from source images during feature extraction. This feature mining stage is critical to image fusion.

In CNN applications, employing efficient initialization and convolution kernel training is paramount and widely regarded as the best technique. Consequently, it's recommended to utilize the convolutional layer of GoogLeNet, which is well-equipped for ImageNet. Specifically, the main Convolutional Layer (CONV1) within GoogLeNet comprises 64 convolution cores of 7×7 dimensions, adept at extracting efficient image features. While CONV1 covers a wide cognitive range, additional illustrations can complement its function. Thus, not all image properties must be presented during fusion. The second and third Convolutional Layers, CONV2 and CONV3, enhance CONV1's attribute screening and provide an attribute plot for integration. In cases where input images are inadequately sampled, there's a risk of missing portions of image data, which could impact the outcome during the feature extraction phase. Therefore, it's essential to regularly update the value of the consistent pace and padding for the kernel of each convolution layer, as highlighted by (Liu 2018) [66] and (Scarpa et al., 2018) [86].

PSO is a repetitive optimization procedure. Its impression arises from the opinion of the searching activities of birds. The fundamental concept is that every distinct in the population repeatedly modernizes itself as per the primary value. To find the ideal result, individuals share data with each other to update the populace (Ali et al., 2022) [6]. PSO is considered to be more efficient because of benefits like search speed, extensive exploration range, simple arrangement, and easy execution.

This study employs PSO (Particle Swarm Optimization) to continually update parameters and systematically identify the most suitable deep learning configuration for optimal outcomes. The primary objective is to select the most influential factors that impact achieving high-performance CNN (Convolutional Neural Network) results and then apply the PSO procedure to determine the best configurations.

The selection of factors to be enhanced was based on evaluating CNN performance through experimental training, where factors are adjustable. Various values of CNN factors yield diverse outcomes for the same task; hence, the goal is to discover optimal configurations. The chosen factors for optimization include:

- The number of convolutional layers.
- The size of filters used in each layer.
- The number of filters to extract feature maps.
- The batch size: indicating the number of data samples fed into CNN during each training phase.

In this approach, CNN integrates factor optimization by incorporating the PSO algorithm. During this phase, PSO initialization is conducted based on the selected factors, generating particles. The training process iterates until all PSO-generated elements are evaluated for each generation. Computational complexity increases with database size, element dimension, PSO iterations, and phase element counts. The completely connected layer, PSO optimizes the layer to produce superior image quality results.

For example, if the PSO is implemented with 10 elements and 10 iterations, the CNN training procedure is performed 100 times.

The steps for enhancing CNN using the PSO procedure are outlined below:

- **Input Database Feeding:** This phase involves selecting and categorizing the database to be utilized for the CNN. It's crucial to meticulously note all the features.
- **Particle Population Generation for PSO:** Setting up the PSO involves defining parameters such as the number of iterations, the number of particles, inertia weight, cognitive and social constants.
- **Configuring the CNN:** The CNN is designed by including the variables acquired from PSO, including the number of convolutional layers, filter dimensions, number of convolution filters, and batch size. These factors are then merged with other parameters to ensure that the CNN is well-suited to the input data.
- **CNN Training and Validation:** Image processing is used to train, verify, and test the CNN utilizing input databases. Recognition rate and AIC value are calculated at this stage.
- **Objective Function Evaluation:** The PSO method finds the optimal objective function value.
- **Updating PSO Parameters:** Every particle modifies its location and velocity in each cycle according to the best-known position (Gbest) and its own best-known position (Pbest) in the search space.
- **Iterative Process:** The method continues by repeatedly assessing all particles until the specified termination conditions are satisfied.

- **Determination of Optimal Solution:** Finally, the most efficient option has been determined. This entails choosing the particle that is described by Gbest, which represents the most optimal configuration for the CNN model.

In (Shehanaz et al., 2021) [90], the formula for each particle was provided in its developing state as follows:

$$v_i^{(t+1)} = wv_i^{(t)} + c_1r(p b_i^{(t)} - x_i^{(t)}) + c_2R(g b_i^{(t)} - x_i^{(t)}) \quad (4.1)$$

$$x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)} \quad (4.2)$$

$v_i^{(t)}$ and $x_i^{(t)}$ represent particle i in the t^{th} iteration speed and position $p b_i^{(t)}$ and $g b_i^{(t)}$ respectively, particle i personal best and the swarm's global best. r and R are random values in $[0,1]$, and w, c_1 and c_2 are weight parameters

The MPSO can be achieved by changing all the coefficients (w, c and α) at the same time. Based on this setting, the values of V_c, ρ_1 and F through

$$V_c^{(t)} = \begin{cases} V_{\max} & t < t_1 \\ \frac{(t - t_1)(V_{\min} - V_{\max})}{t_2 - t_1} + V_{\max} & t_1 < t < t_2 \\ V_{\min} & t > t_2 \end{cases} \quad (4.3)$$

This function ensures that, from iteration t_1 to t_2 , the value of V_c decreases linearly, having the largest at the beginning and the smallest at the end of the search.

$$F^{(t)} = \begin{cases} F_{\min} & t < t_1 \\ 1 & t_1 < t < t_2 \\ F_{\max} & t > t_2 \end{cases} \quad (4.4)$$

More attention is focused on the individual best thanks to this equation, which also assures that the search is balanced between the individual and global best and ultimately centers on the present (Zhang et al., 2022) [117]. The values of coefficients are changed during the time of optimization.

Algorithm 4.1: MPSOCNN

1. For each image pixels
2. For every pixels \in input layers
3. Perform convolution layer
4. Convert the input into sub layers
5. Initialize the particles (pixel) with velocity
6. Initialize the position
7. Do
8. For each pixel
9. Compute fitness values
10. If compared to the greatest fitness in history, if fitness of value
11. Set current value as new pbest
12. Set current value as new gbest
13. End
14. End
15. Decide which pixel is the best overall in terms of fitness value

16. For each pixel
17. Compute the pixel with velocity
18. Compute the pixel position
19. Update pixels with velocity and position using (4.1) & (4.2)
20. Adaptively change the coefficients using (4.3) and (4.4)
21. Select best pixel
22. Return optimal solution

The above algorithm 4.1 describes about the feature extraction using MPSOCNN algorithm. Initially, all the pixels are obtained in CNN's input layer (Liu et al., 2014) [67]. Then convolution layer is performed. This component forms the fundamental unit of a CNN, housing a collection of filters (or kernels) whose parameters are gradually adjusted during training. Typically, the size of these filters is smaller than the original image. As the filters convolve with the image, they generate activation maps. This process enables the extraction of more informative features from the input image. In this context, the PSO algorithm is utilized to optimize the neurons within this layer. Its function is to reduce the number of neurons in the convolutional layer while preserving essential information. Through the generation of global best and local best features, the neurons are fine-tuned to extract the most optimal features.

4.2.2. Image fusion of multi –model image

Feature extraction is a broad term encompassing methods aimed at generating diverse combinations to effectively address challenges while accurately representing data (Schnabel et al., 2016) [87]. Many computer scientists believe that well-optimized feature

extraction is the key to successful model development, and it can be achieved through various levels of image analysis:

- Low-level content: This pertains to the visual aspects of an image, including shape, texture, and color features.
- Middle-level content: This entails recognizing and positioning various elements, sceneries, and objects inside the image.
- High-level content: This encompasses the emotional and conceptual meaning derived from the integration of perceptual features.

In this proposed system, the CNN algorithm is employed during the training phase. Following the introduction of training samples at the input layer, the input, output, hidden layers, and anticipated output matrix are evaluated and compared with the fusion image outputs. Initially, each input image is broken down into pixels, which are interpreted as 2D arrays for black and white images (Azam et al., 2021) [13]. Each pixel holds a value ranging from 0 to 255, allowing for the extraction of high-frequency details from the two source images.

Zero indicates black and 255 for white. In these numerical values lies the grayscale, which serves as a basis for the machine's subsequent operations. For a color image, this data forms a 3D array comprising blue, green, and red layers, with each color's values spanning from 0 to 255. By combining the values from each layer, colors are formed. At this stage, the PSNR is computed to gauge the error and is then propagated backward in time (Yadav and Yadav 2020) [113]. To improve model efficiency, adjustments are performed to the feature indicators and weights. The output

layer consists of a node configured as an image fusion display. This process involves the utilization of multi-modal images, particularly MRI scans of the brain taken from various perspectives. Images from different angles are fused to enhance clarity, with the final image being a result of feature-based fusion.

4.2.3. Image Classification

CNNs are an important development in image recognition; they are a subset of deep learning neural networks (Pereira et al., 2016) [80]. They greatly aid in accurately analyzing and classifying images. The process of accepting an input and classifying it includes image classification.

Convolutional, ReLU, pooling, and fully linked layers are some of the layers found within a CNN. When it comes to classification algorithms, CNNs usually need less preprocessing than other methods. Convolutional layers in a CNN are responsible for convolving the input data so that information may flow to the next layer. In the next layer, pooling creates a single neuron by combining the outputs of neuron groups. Neurons in one layer are coupled to neurons in another by fully connected layers. CNNs function by taking characteristics out of images, which removes the requirement of manual feature extraction. This makes deep learning models, in particular, CNNs extremely useful for tasks involving computer vision. As a classifier in the output layer, the fully connected layer functions similarly to a traditional Multi-Layer Perceptron. A Softmax activation function is usually used by the classifier.

Every neuron in a completely connected layer is connected to every other neuron in the layer below it, enabling the utilization of attributes extracted from the previous layer's output for image classification based on training data. MFCL (Multi-Feature

Convolutional Layer) is a specialized type of CNN consisting solely of convolutional layers. By utilizing only convolutional layers, FCNs (Fully Convolutional Networks) can handle input images of any size. The MFCL layer connects the information extracted from previous steps, such as convolutional layers and sub-layers, to the output layer, ultimately classifying the input into the desired label.

4.3. RESULTS AND DISCUSSION

A dataset termed BraTS 2018 includes professional medical diagnoses of brain cancer and also multimodal 3D brain MRIs. The four MRI modalities included in each dataset instance are as follows: T1, FLAIR, T1c, and T2. The findings comprise three tumor subregions: peritumoral edema, necrotic and non-enhancing tumor core, and enhancing tumor. These data include three nested sub-regions: the total tumor, the central tumor, and the enhancing tumor. The dataset was compiled from data collected at 19 institutes, utilizing various MRI scanners.

4.4. PERFORMANCE METRICS

To assess the effectiveness and make a comparison of our findings with previous studies, we use different evaluation metrics including, accuracy, precision, recall and time period. These metrics are calculated as follows:

Accuracy

Accuracy pertains to how closely a measured value aligns with a standard or true value. Essentially, it denotes the tool's capability to precisely measure the exact value, which is quantifiable.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.5)$$

Precision

In combining true and false instances, real cases are divided to get the accuracy,

$$Precision = \frac{TP}{TP + FP} \quad (4.6)$$

Recall

To find recall, divide the count of true items by the entire count of items in the same class.

$$Recall = \frac{TP}{TP + FN} \quad (4.7)$$

Time Period

Calculated in milliseconds or seconds, the duration required to finish the segmentation process is shown as elapsed time.

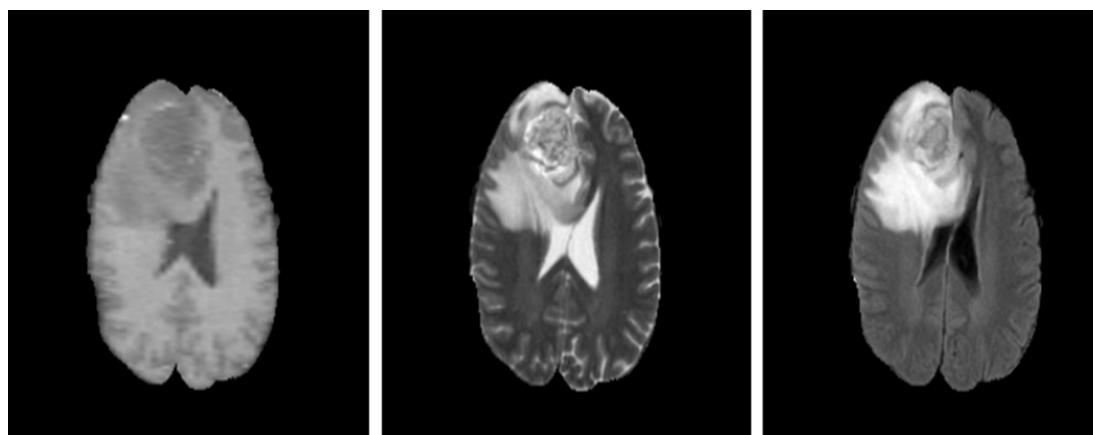
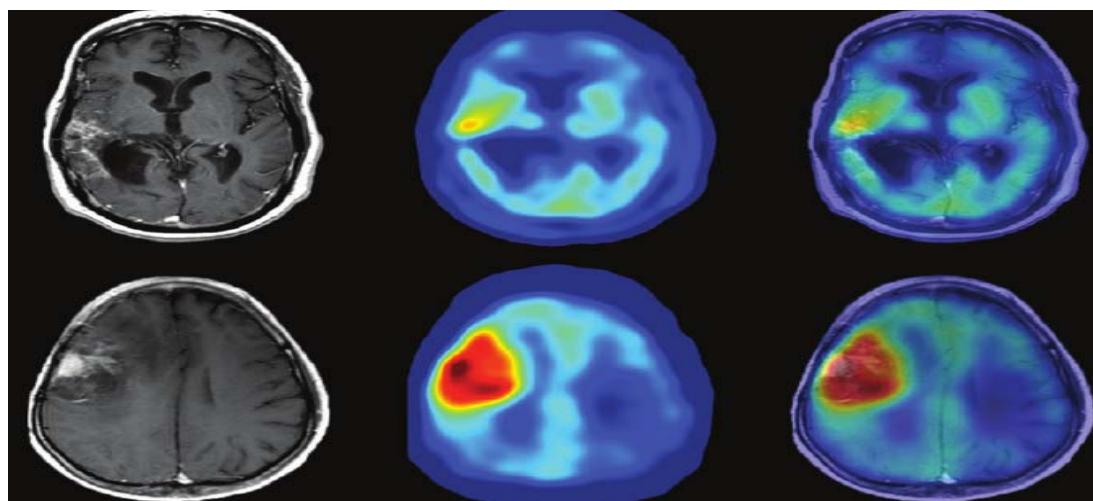
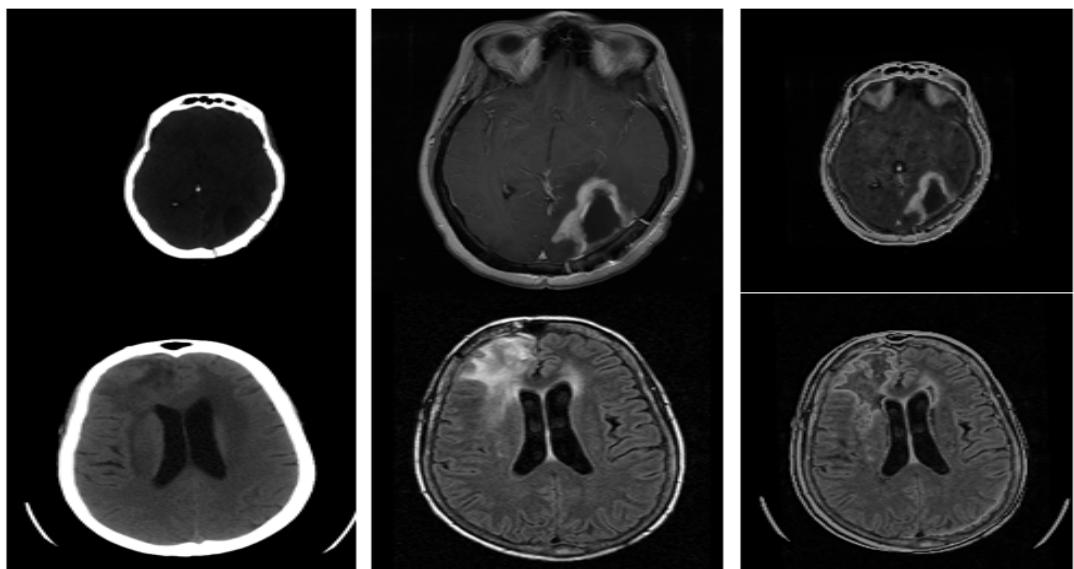


Fig 4.2 Image Fusion

From the above Fig 4.2, initially, input image is obtained and segmentation is applied on the given images using MUNet algorithm. Then, MPSOCNN algorithm is utilized for feature extraction. It provides more informative features with best images which is highlighted in blue color. In CNN layer, the fusion is performed. Finally, classified images are provided.

Table 4.1: Performance Analysis

Algorithm	Accuracy	Precision	Recall	Time Period
PSO_FCL	94	90	91	5.3
MPSO_FCL	95	92	93	4.6
CPSO_MFCL	96	95	96	4.2
MPSO_MFCL	98	97	98	3.8

Table 4.1 displays the schemes' performance analysis. MPSO_MFCL offers 4%, 3% and 2% better Accuracy when compared to PSO_FCL, MPSO_FCL and CPSO_MFCL respectively (Fig 4.3).

Accuracy

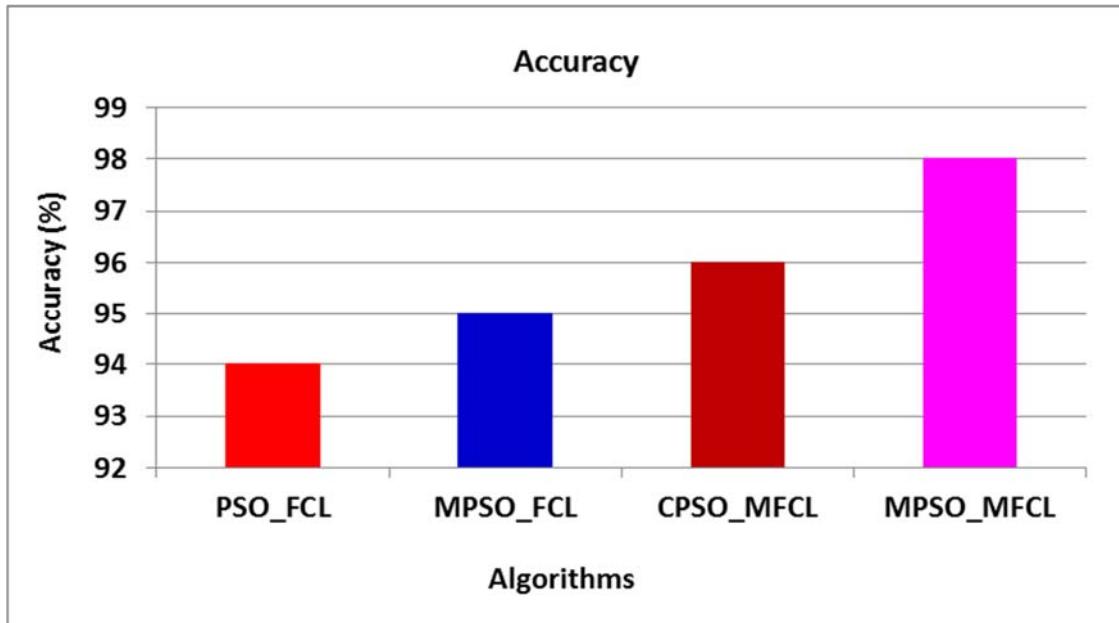


Figure 4.3 Accuracy

A comparison of accuracy metrics between the suggested and current approaches is shown in Figure 4.3. Existing methods, namely PSO_FCL, MPSO_FCL, and CPSO_MFCL algorithms, exhibit lower accuracy levels, whereas the proposed MPSO_MFCL algorithm demonstrates higher accuracy when applied to MRI brain images. Consequently, the findings suggest that the proposed MPSO_MFCL algorithm enhances the classification process by leveraging optimal features.

Precision

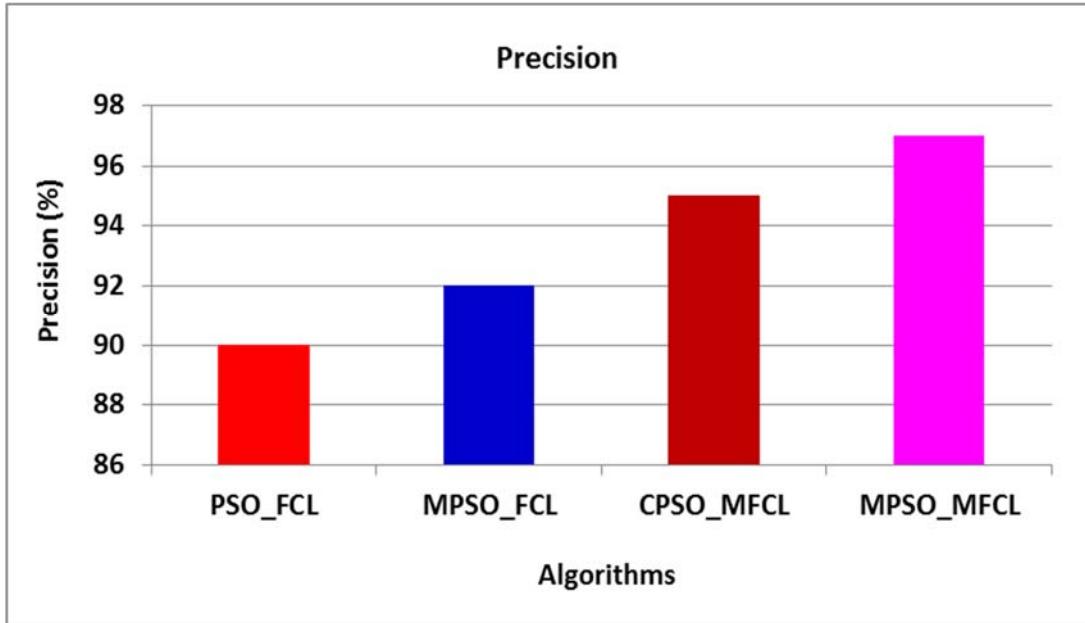


Figure 4.4: Precision

Figure 4.4 illustrates the comparison of the accuracy metric between the current and suggested approaches. The proposed MPSO_MFCL algorithm improves MRI brain image precision over PSO_FCL, MPSO_FCL, and CPSO_MFCL. The suggested MPSO_MFCL method enhances classification using optimum features.

Recall

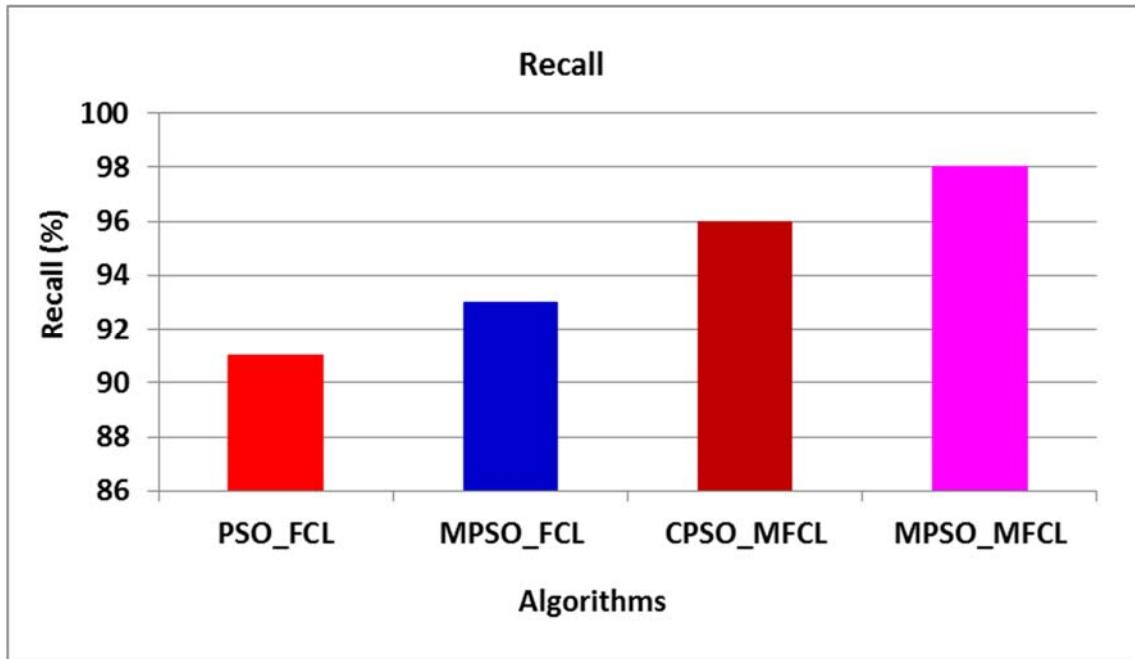


Figure 4.5: Recall

Figure 4.5 illustrates the comparison of the recall metric between the present and suggested techniques. For MRI brain images, the suggested MPSO_MFCL algorithm has a higher recall than PSO_FCL, MPSO_FCL, and CPSO_MFCL. The MPSO_MFCL method enhances classification using optimum features.

Time complexity

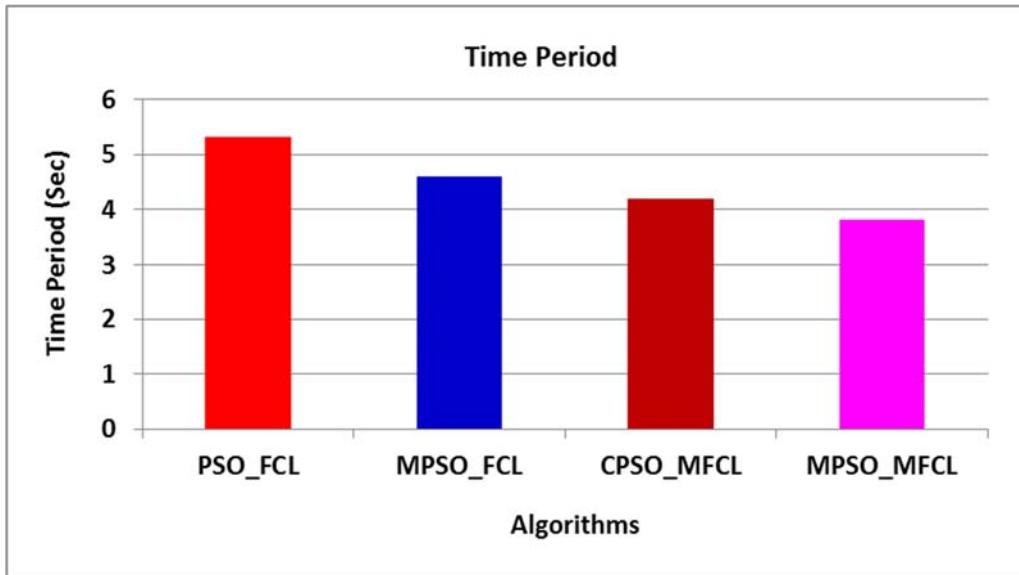


Figure 4.6: Time Period

Fig 4.6, shows the comparison of time period metric for existing and proposed methods. The existing methods are such as PSO_FCL, MPSO_FCL and CPSO_MFCL algorithm provides higher time complexity whereas Lower time frame complexity for the provided MRI brain images is offered by the suggested MPSO_MFCL method. Consequently, the findings indicated that the suggested MPSO_MFCL algorithm enhances the classification procedure by using the best characteristics.

4.5.SUMMARY

In this work, a multi-modal medical image fusion technique related to MPSO_MFCL is proposed. The proposed technique could routinely regulate the optimal parameters for source images, and deal with medicinal imageries with dissimilar modalities. Also, the variety of images boosts the strength of proposed algorithm. Because of lightweight and first-class factors of proposed algorithm, its potential for use

in intelligent medicine is extensive. The model's performance can be improved if it is extended to include more image data sets. In all aspects the proposed MPSO_MFCL outperforms the existing methodologies. Additionally, the suggested technique is very helpful for the radiologists in medical application. Brain tumor segmentation is efficiently conducted using the MUNet algorithm, followed by feature extraction via the CNN algorithm. MRI brain images, captured from various perspectives, are then fused to enhance clarity. This fusion process integrates features from different views, resulting in a clearer image. Subsequently, image classification is performed to achieve more precise results. The outcomes demonstrate that the proposed MPSO_MFCL approach delivers superior image fusion performance, as evidenced by improvements in precision, recall, accuracy, and processing time.

Chapter 5

CHAPTER 5

OPTIMAL SEGMENTATION AND FUSION OF MULTI MODAL BRAIN IMAGES USING CLUSTERING BASED DEEP LEARNING ALGORITHM

5.1. INTRODUCTION

This research aims to enhance the efficiency of MRI image fusion. Despite various methodologies introduced previously, achieving satisfactory image quality has been elusive. Existing approaches suffer from drawbacks such as time consumption and inaccurate results (Koley et al., 2016) [51]. To address these issues, this study proposes the Adaptive Firefly Optimization based Convolutional Neural Network (AFFOCNN) scheme to enhance overall system performance. The primary contributions of this study are noise reduction, segmentation, feature extraction, and image fusion. By leveraging effective algorithms, the proposed method aims to deliver more precise results using the provided MRI image dataset.

5.2 PROPOSED METHODOLOGY

This study introduces the AFFOCNN scheme to enhance the efficiency of image fusion. The primary objective of this system is to achieve efficient image fusion using MRI images. The comprehensive block diagram of the proposed system is depicted below in Figure 5.1.

5.2.1. Input image dataset

Multi-modal 3D brain MRIs and expert-interpreted brain cancer segmentations are available in the BraTS 2018 dataset. Four MRI modalities are included in each instance in the dataset: T1, T1c, T2, and FLAIR. Three tumor subregions are included in the dataset's observations: the necrotic and non-enhancing tumor core, the enhancing tumor, and the peritumoral edema. The overall tumor, central tumor, and enhancing tumor are the three nested sub-regions into which these findings are further divided. The data were collected from 19 institutes, utilizing a variety of MRI scanners.

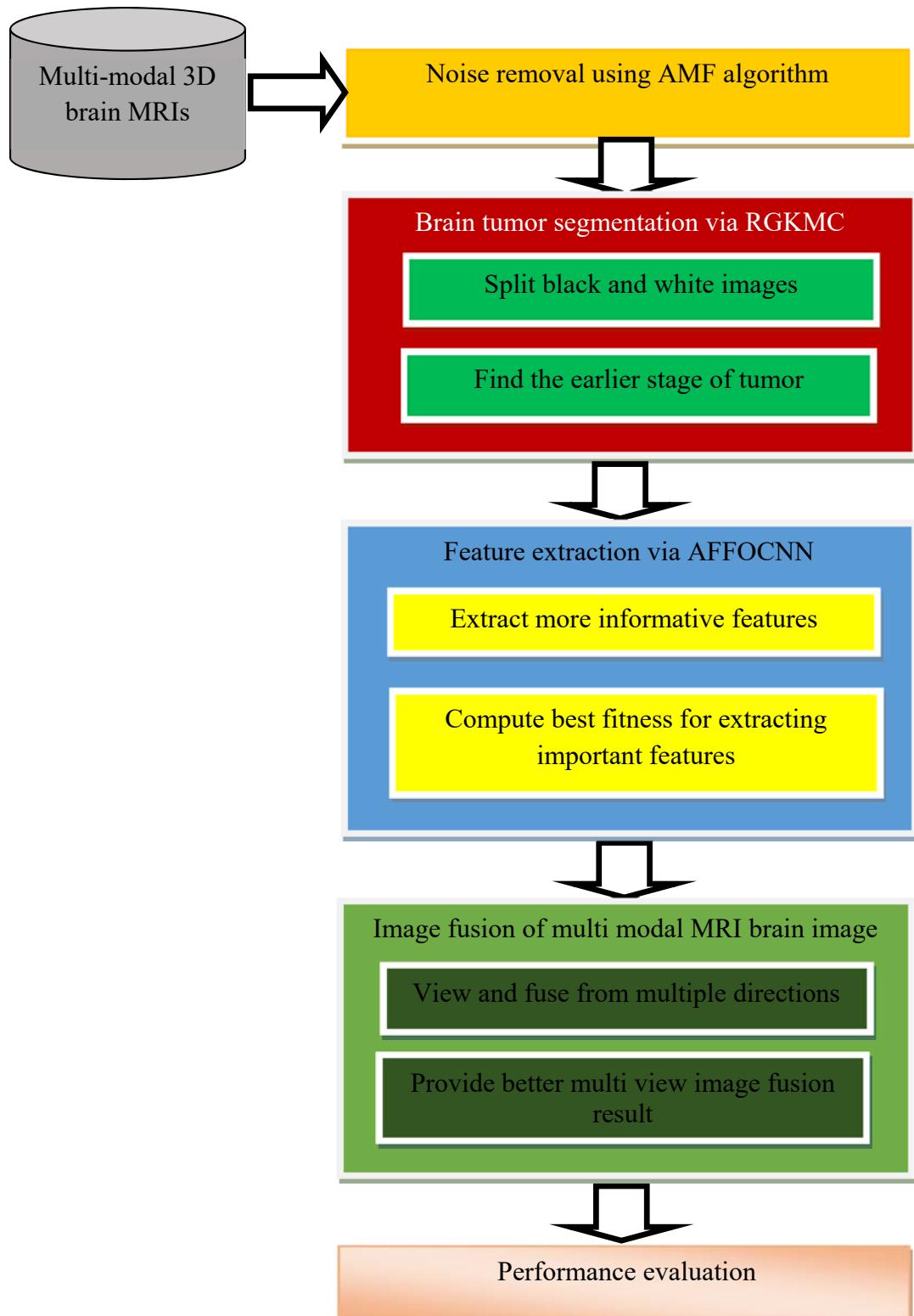


Fig 5.1 Overall architecture diagram of the proposed AFFOCNN system

5.2.2. Noise removal using Adaptive Median Filtering (AMF)

This study utilizes the AMF algorithm for noise removal, which is specifically designed to eliminate unnecessary noise from the input images. The primary objective is to contribute to a robust and effective model for enhancing image findings, especially in situations when noise is high. To ascertain which of an image's pixels have been affected by noise, the AMF method uses spatial processing. Pixels are classified as noise by the AMF by comparing each pixel to its neighbours. Two modifiable characteristics are the neighborhood's size and the comparison threshold. Noise is defined as pixels that are physically unaligned with comparable pixels and vastly different from most of their neighbours. The median pixel value of the nearby pixels that passed the noise labelling test is then used to replace these noisy pixels. Consequently, by reducing noise levels in the picture, the AMF algorithm improves image quality.

The adaptive filter's window size expands based on the quantity of identified noise candidates within a particular region. Equation (5.1) utilizes the quantity of pixel noise in every region to determine an appropriate window size. Through extensive simulations aimed at achieving optimal results, MSE is used to define thresholds for changing based on the number of noisy pixels in a region, the window size.

$$W(M) = \sum_p C(p, D_p) + \sum_{q \in N_p} T[|M_p - M_q| = 1] + \sum_{q \in N_p} T[|M_p - M_q| > 1] \quad (5.1)$$

W represents the window matching function, and M is W ideal solution. In the shown image, p is the real pixel and q is p neighboring pixel. N_p stands for the collection of pixels that are next to p . The variables M_p and M_q , respectively, represent the

estimated matched windows of p and q . The cost of a certain matched window is indicated by C . $T[\cdot]$ is a logic function that returns 1 in the case if the statements it contains are true and 0 if they are false. When there are no noise-free pixels in the window or when a window's median pixel is noisy, expand the window's size in equation (5.1) above. If the requirements are not satisfied, the adaptive window size used to filter image pixels expands. The window median filters the pixel if the criteria are fulfilled. Let I_{ij} is the pixel of the corrupted image, I_{min} be the minimum pixel value, I_{max} be the maximum pixel value of the window, W be the current window size, W_{max} be the maximum window size that is obtained and I_{med} be the median of the assigned window. $T[\cdot]$

5.2.3. Brain tumor segmentation via Region Growing based K-Means Clustering (RGKMC) algorithm

This study effectively uses the RGKMC algorithm to segment brain tumors. In the existing system, MUNet is applied for segmentation. But it has issue with inaccurate localization of nonstandard shapes, which can affect the precision of segmentation results. Additionally, MUNet not effectively capture long-distance spatial relations in medical images, which can limit its performance in brain tumor diagnosis

To address the aforementioned issue, this study begins with preprocessing, followed by proposing the RGKMC algorithm-based segmentation. KMC, a robust clustering technique outlined by (Sulaiman and Isa 2010) [98], effectively segregates similar data into groups using initial cluster centroids. It utilizes the Euclidean distance concept to compute the centroids of these clusters. The images are further divided into black and white photographs by RGKMC. Tumors are detected in both images in this study, allowing for early brain tumor prediction. Data is first partitioned randomly using

the method, and After that, (i) each data item is repeatedly reallocated to the cluster whose center is closest to it (ii) based on the current cluster centres, which are determined by calculating each cluster's average vector in the data space. This process continues until no further reassessments occur. Consequently, intra-cluster variance, or the sum of squares of the differences between the features of the data and the cluster centers corresponding to those features, is minimized locally. A demonstrative example of the KMC method may be seen in Figure 5.2.

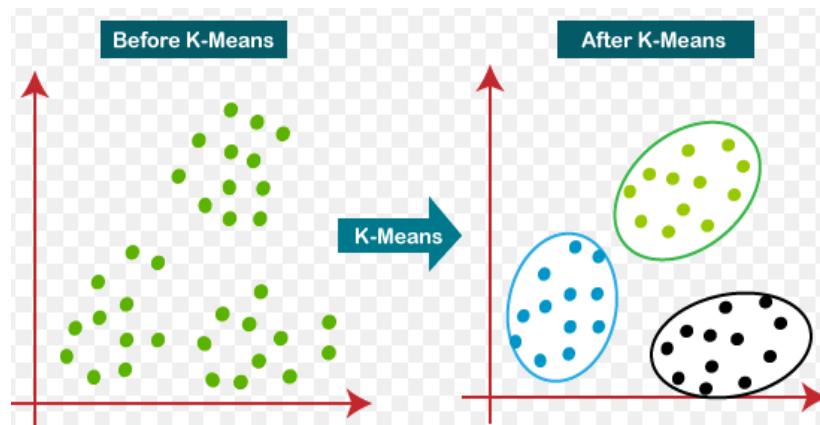


Fig 5.2 Example of KMC algorithm

RGKMC identifies the groups pixels in affected regions based on similarity criteria through centroids. The process initiates by selecting a seed pixel within a designated region and progressively enlarges this seed by integrating neighboring pixels within regions that satisfy particular similarity criteria. It clusters or partitions given image into multiple segments or regions based on specific common attributes.

A category of image segmentation techniques that use pixels and include choosing beginning seed points of interest. To determine if nearby pixels of the original seed points should be included in the area, this segmentation technique assesses those

pixels. Region splitting and merging are iterative processes. Typically, initial region splitting is performed on an image to divide it into numerous regions, followed by merging these regions to produce a well-segmented version of the original image. It postulates that neighboring pixels within the same region have similar intensity values

The simplicity of implementation and linear runtime of K-means are two of its strongest points. The cluster number and the number of classes are maintained equal in this work. Find the cluster centroids by calculating the Euclidean distance using the formula below.

$$d(i,j) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5.2)$$

Where x_i and y_i are two points in Euclidean n-space

Algorithm 5.1: RGKMC algorithm

1. Select a number of clusters k from the MRI database (D).
2. Initialize cluster centers μ_1, \dots, μ_k and segment black and white images.
3. Choose k data points and set cluster centers to these points, then extract initial seeds.
4. Clusters should be randomly assigned points, their means should be determined, and nearby pixels should be identified.
5. Find the cluster center that each data point is closest to, then use equation (5.2) to get the distance measure.
6. Select the pixels with the minimum distance.

7. Following evaluating the degree of similarity between the pixels, the data point is assigned to the cluster that is closest to the seed and updated.
8. Re-calculate cluster centers (mean of data points in the cluster) and detect tumors in black images.
9. Estimate the likelihood of early brain tumor prediction in white images.
10. Stop the process when there are no new reassessments.

The RGKMC algorithm is utilized to cluster complete pixels from a database. It processes pixels individually, assigning region-based values to them. Primarily designed to detect early-stage brain tumors in black and white images, it segments these images based on tumor indications. Notably, this algorithm effectively considers the probability of early tumors in white areas of the image. When applied to MRI databases, it verifies the clustering of newly added pixels, ensuring they are correctly classified. If a pixel is in the correct cluster, its value is made permanent; otherwise, alternative values are assigned until the correct cluster is identified. Consequently, this approach enhances the accuracy of early brain tumor detection while also improving image quality.

5.2.4. Feature extraction using Adaptive FireFly Optimization based Convolutional Neural Network (AFFOCNN)

This research utilizes the AFFOCNN algorithm, which identifies more useful features by optimizing fitness values, is used to extract features. Three convolutional layers are used in the feature extraction process to extract the properties from the original images. An essential phase in the image fusion process is extracting these characteristics or properties.

For CNNs, the best methods are regarded as effective initialization and convolution kernel training. GoogLeNet's convolutional layer is hence advised to be used. This layer is the primary Convolutional Layer (CONV1) and is well-suited for ImageNet. The device is equipped with 64 convolution cores, each with dimensions of 7×7 , which effectively extract information from images. Despite the broad recognition range of CONV1, other examples may be used to show that not every aspect of the image has to be included in the fusion. To examine the characteristics acquired by CONV1 and finally create a combined attribute plot, CONV2 or CONV3, the second and third Convolutional Layers, respectively, are added. An improper sample of the input images might result in the deletion of some image data, which would affect the feature extraction process. To mitigate this, the consistent pace and padding for each convolution layer's kernel should be updated. However, standard CNNs require substantial amounts of training data and computational power to be trained effectively. Hence, it has drawback with time consumption during implementation. To overcome this problem, firefly algorithm is integrated to fine tune the CNN layers. Firefly algorithm provides global and local optimal solutions to the given image dataset. It reduced the time complexity by generating best fitness features.

This work uses AFFO to automatically determine the optimal deep learning setup for achieving efficient results via parameter updates. The primary goal is to select the most relevant factors that significantly impact the performance of CNNs and then use the AFFO method to identify the optimal factors.

The selection of the elements to be optimized is based on results derived from the actual training of a CNN, where these factors can be varied. Different values of CNN

factors produce a range of outcomes for the same task; thus, the aim is to find the optimal configurations. The factors presented are specifically chosen for optimization.

- The number of convolutional layers
- The size of the filter used in each cycle
- The number of filters to extract the subsequent feature maps
- The batch size: this value represents the number of samples fed into the CNN during each training step.

In this approach, the CNN integrates factor optimization using the AFFO algorithm. Initially, AFFO is set up according to the factors selected for computation, generating fireflies. The training process is iterative and continues until all elements produced by AFFO are evaluated in each generation. It is important to note that the computational load is significant, depending on the database size, the number of elements, the repetitions of AFFO, and the number of elements at each stage.

The Firefly Algorithm (FA) (Liu et al., 2020) [62] inspires the biochemical and social behaviors shown by actual fireflies. These insects produce short, regular bursts of light that help them attract mates and function as a warning system for protection. FA models this flashing behavior as an objective function to be optimized. The algorithm operates based on the principle of fireflies' flashing lights, where the intensity of the light guides a group of fireflies toward brighter and more attractive positions, ultimately leading to the optimal solution for the problem at hand.

This mechanism standardizes certain characteristics of fireflies, as illustrated by (Liu et al., 2013) [60]:

- i. Each firefly is attracted to others regardless of their sex.
- ii. A firefly's brightness is closely correlated with its attraction; when two fireflies are nearby, the brighter of the two the less brilliant one. In the absence of a brighter firefly nearby, a firefly will migrate at random.

In terms of mathematics, the objective function governs the luminosity of a firefly. Figure 5.3 illustrates the basic mechanisms of the firefly algorithm.

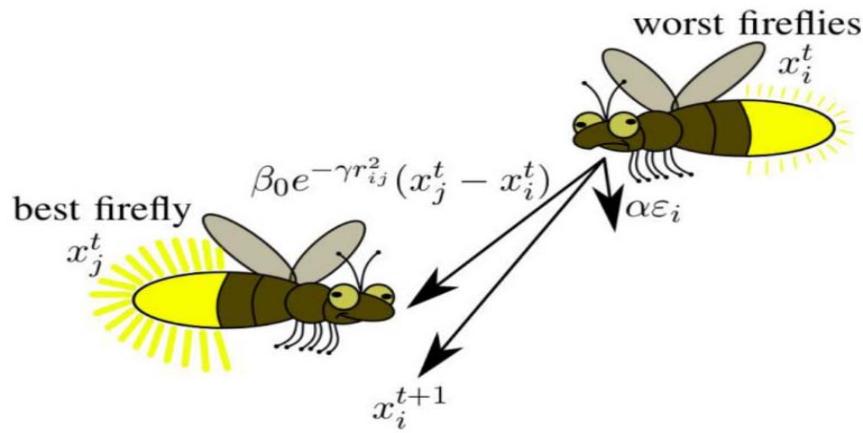


Fig 5.3 Basic mechanism of the firefly algorithm

The firefly method was selected because it may provide the best results when used for multi-objective tasks. The brightness and the objective function are intimately correlated in a maximizing situation. For simplicity, the brightness or light intensity of a firefly, which is linked to the encoded objective function, is often assumed to influence its attractiveness. a) The source's attractiveness and light intensity: According to the inverse square law, the light intensity changes in proportion to the inverse of the square of the distance.

$$I(r) = \frac{I_0}{r^2} \quad (5.3)$$

Where $I(r)$ is the brightness of the light appealing r^2

Appealing produced by allocating pixels at random

b) The following is the light intensity, and an intermediate is available:

$$I(r) = I_0 \exp(-\gamma r) \quad (5.4)$$

Where I_0 is the medium's absorption coefficient.

c) To prevent singularities, the following Gaussian approximation form is taken into consideration.

$$I(r) = I_0 \exp(-\gamma r^2) \quad (5.5)$$

The light intensity that other fireflies in the area notice is closely correlated with how appealing a firefly. A new solution is derived by factoring in possible variations and randomly adjusting the allocation. Some characteristics deemed desirable must be applied to each pixel in a batch. Thus, the attractiveness β of a firefly is determined as follows:

$$\beta = \beta_0 \exp(-\gamma r^m) \quad (5.6)$$

Where β_0 is the attractiveness at $r=0$.

The distance between any two fireflies may be calculated using the formula below (pixels), i and j:

$$r_{i,j} = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2} \quad (5.7)$$

In the equation, $x_{i,k}$ represents the k^{th} factor of the spatial match x_i for the k^{th} firefly, with d denoting the number of dimensions. An adaptation parameter is incorporated for both absorption and random parameters, leading to the development of an Adaptive FireFly Optimization (AFFO) algorithm. These modifications enhance both the global and local search capabilities by adjusting the parameter linearly throughout the iteration process (Liu et al., 2013) [60]. AFFO is utilized to generate optimal image features for the specified database.

Calculate the parameter α as follow:

$$\alpha(t+1) = \left(1 - \frac{t}{MaxG}\right) \alpha(t) \quad (5.8)$$

To improve the solution's precision and rate of convergence, α adjusts the value based on the optimization's degree of distance divergence. Simultaneously, it is rewritten as follows to improve population flexibility:

$$\alpha = \alpha_{min} + (\alpha_{max} - \alpha_{min}) \times ||x_i - x_{best}|| / L_{max} \quad (5.9)$$

Where

$$L_{max} = (x_{worst} - x_{best}) \quad (5.10)$$

The maximum and lowest characteristics are α_{max} and α_{min} . The distance between the worst individual x_{worst} and the global optimum individual x_{best} is L_{max} in Eq.(5.10). The firefly people are spread around the region in the algorithm's initial phase, with most distant from the globally perfect persons. At this point, the value of $\|x_i - x_{best}\|$ is larger, L_{max} and $(\alpha_{max} - \alpha_{min})$ are fixed values. Eq. (5.9) shows that a higher α value in the early stage leads to better global optimization. The algorithm's implementation draws me to fireflies with near-global optimal characteristics and brighter than themselves. Later on, firefly individuals would congregate around the world's most ideal individuals, the value of $\|x_i - x_{best}\|$ is now smaller, making it favourable to enhance the finding of ideal characteristics. During each iteration, the parameter α is adjusted based on the position of the optimum, thereby enhancing the convergence rate of the algorithm. As per the aforementioned analysis, the step size factor α undergoes adaptive and dynamic changes in relation to the distance between individual fireflies, ensuring a balance between algorithmic development and search capability.

This research uses a unique fitness function that takes execution time and accuracy into account, represented as:

$$f(x) = \frac{\left(I_d / I_t \right) \times \left(I_p / P_{init} \right)}{exp^{-e_E / e_M + H_{accuracy}}} \quad (5.11)$$

where I_d is the number of dropped image features. m_t is the total number of features sent with higher accuracy

I_p is the pixel in image i.

P_{init}^i is the initial image.

e_E is the execution time and e_M is the maximum allowable delay.

$$x_i = x_i + \beta_0 e^{-\gamma r^2} (x_j - x_i) + \alpha (rand - \frac{1}{2}) \quad (5.12)$$

Where x_i and x_j is distance between two firefly nodes

Every trait's fitness within the population is determined. In the beginning, in the first-generation, each firefly's fitness is assessed based on the randomization of pixel allocation within a batch. Subsequently, a selection process is employed to choose two fireflies. The firefly with the greater brightness, indicative of higher fitness, is chosen for advancement to the next generation.

The steps for enhancing the CNN through the AFFO procedure are outlined as follows:

Algorithm 5.2: AFFOCNN

1. Feed database for CNN's equipment. It is important to note all the aspects in this step, the process involves choosing the database to be managed and organized for CNN.
2. Objective function (x), $x = (x_1, \dots, x_n)^T$ the aim function is to minimize the MSE, reduce implementation time, and maximize accuracy.
3. Produce the fireflies for the AFFO procedure. Produce initial population of fireflies x_i ($i = 1, 2, \dots, n$), light intensity I_i at x_i is found via $f(x_i)$, describe light absorption coefficient γ
4. Configure the CNN architecture, incorporating factors obtained through AFFO such as the number of convolution layers, filter dimensions, quantity of

convolution filters, and batch size. These parameters are adjusted in tandem with additional factors to ensure the CNN is appropriately tailored to shape the input data.

5. CNN setup and validation - For testing, preparation, and authentication, the CNN analyzes image-containing input databases. This stage produces some recognition.
6. Determine the objective of the function. The objective function is determined by the AFFO process, which determines the optimal value.
7. while ($t < \text{MaxGeneration}$)
8. for $i=1:n$ all n fireflies
9. for $j=1:i$ all n fireflies
10. if ($Ij > Ii$), Move firefly i towards j in d -dimension;
11. end if
12. Attractiveness changes along with distance r via $\exp[-\gamma r]$
13. Compute fitness function using (5.11) and (5.12)
14. Compute objective model using (5.7)
15. Update light intensity and estimate new solutions using (5.4)
16. Update the optimal features using (5.9)
17. end for j
18. end for i
19. Find the best fireflies right now by ranking
20. end while

21. An appealing firefly change to a The process is iterative and involves evaluating every component until the stop condition is satisfied.
22. Finally, the ideal resolution is chosen. The firefly chooses the best feature for the CNN model throughout this procedure.

5.2.5. Image fusion of multi modal image

The objective of image fusion is to remove distortion and extraneous information from individual source images while preserving complementary information. Therefore, performance measurements are essential for evaluating fusion approaches' efficacy and comparing results from different algorithms (Jain et al., 2021) [45]. To identify and treat brain cancer, it is essential to estimate brain tumors kind, size, location, and extent with accuracy. When brain cancers are detected and diagnosed quickly, fused MRI brain images provide superior results than separate MRI images. The remaining content is already discussed in chapter 4.

5.3 EXPERIMENTAL RESULT

BraTS 2018 includes a collection of multi-modal 3D brain MRIs annotated by medical professionals with information on brain cancer. Four MRI modalities are used in each case: T1, FLAIR, T1c, and T2. The dataset indicates three tumor sub-regions: the necrotic and non-enhancing tumor core, the enhancing tumor, and the peritumoral edema. The whole tumor, central tumor, and enhancing tumor are the three nested sub-regions into which these findings are further divided. Data collection spans 19 institutes, utilizing a variety of MRI scanners.

For assessment, several performance parameters are taken into account, including execution time, precision, recall, F-measure, accuracy, and mean squared error (MSE). The novel AFFOCNN algorithm as well as the well-known UNet, MUNet, and MPSO+MFCL algorithms are used in this evaluation. The AMF filtering image and the input MRI images are shown in Figures 5.4(a) through 5.4(b).

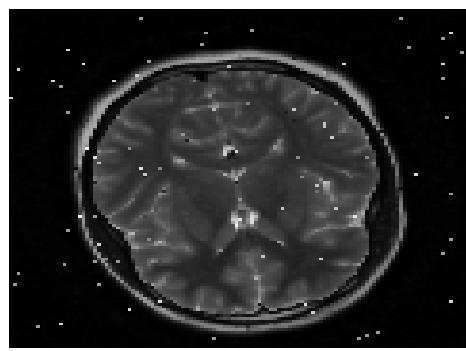
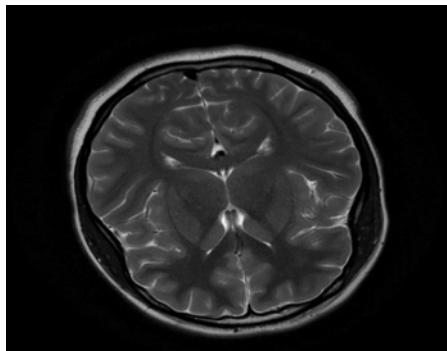


Fig 5.4 (a) Input image Fig 5.4 (b) AMF image

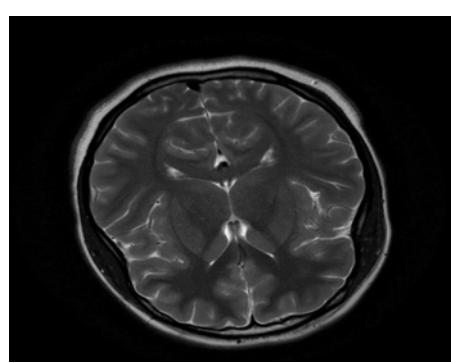
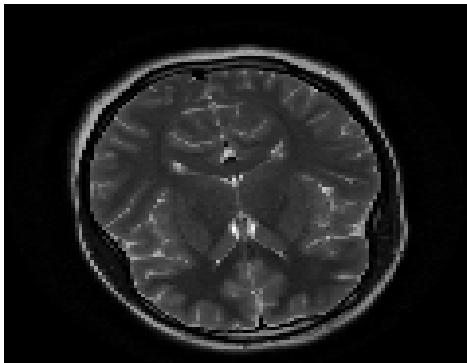


Fig 5.5 (a) Extracted image

Fig 5.65(b) Output image

Figures 5.5(a) through 5.5(b) show the MRI image extraction and output findings.

Table 5.1 Comparison values for given dataset

Methods/Metrics	CNN	UNet	MUNet	MPSO+MFCL	Proposed AFFOCNN
Accuracy (%)	92	95	97	98	99
Precision (%)	86	90	93	97	98
Recall (%)	82	87	91	94	96
F-measure (%)	81	89	92	95	97
MSE (%)	49	45	34	23	12
Time complexity (sec)	5.1	4.6	3.8	3.1	2.4

The comparative values for the image fusion dataset utilizing the suggested and current approaches are shown in Table 5.1.

Accuracy

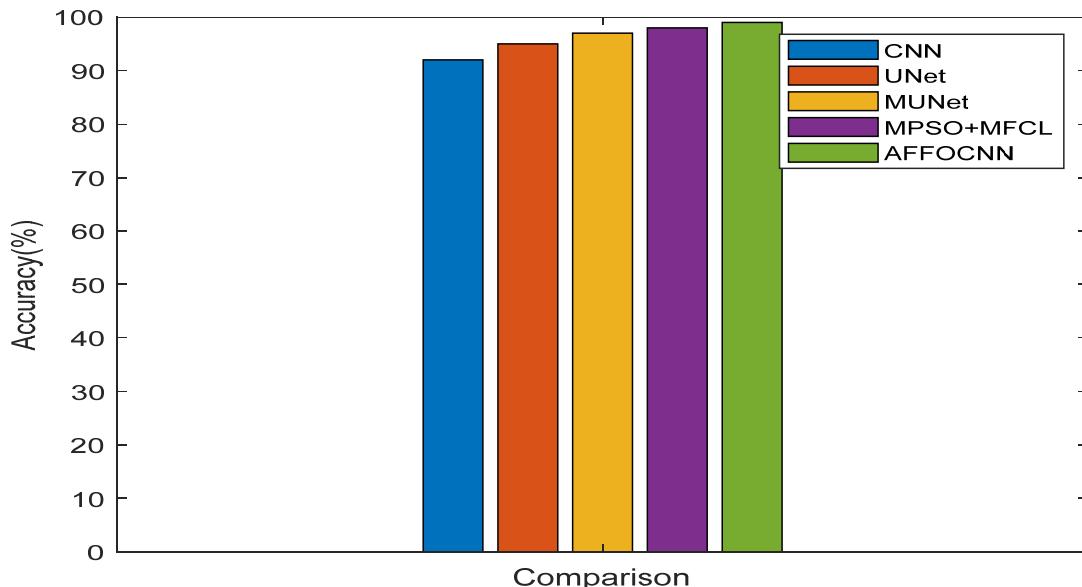


Fig 5.6 Accuracy

A comparison of accuracy metrics between previously suggested and established approaches is shown in Figure 5.6. Conventional approaches such as CNN, UNet, MUNet, and MPSO+MFCL algorithms exhibit lower accuracy, while the proposed AFFOCNN algorithm demonstrates higher accuracy with the provided MRI datasets. Additionally, by taking neighboring pixels into account, the suggested region-growing-based segmentation method improves the quality of images. To increase the accuracy of image fusion, pre-processing is performed using the AMF method. In general, the results indicate that by extracting more relevant features, the suggested AFFOCNN algorithm improves the image fusion process.

Precision

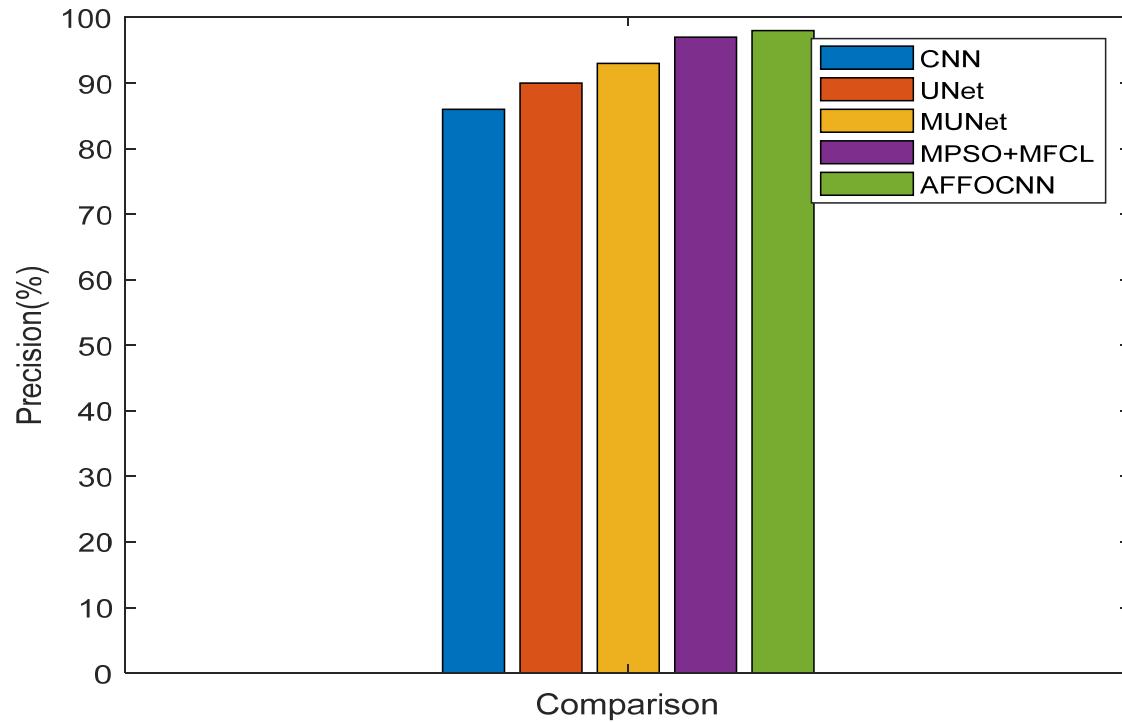


Fig 5.7 Precision

A comparison of accuracy measures amongst accepted approaches is shown in Figure 5.7. Compared to other approaches that are currently in use, such as CNN, UNet, MUNet, and MPSO+MFCL, which have lower accuracy, the suggested AFFOCNN methodology shows greater precision. The AMF filter-based image fusion methods demonstrate improved handling of spatial domain distortion. To provide a more useful image, the suggested image fusion combines multi-view MRI scans of the same scene. Consequently, the results imply that the AFFOCNN method improves image feature accuracy for image fusion.

Recall

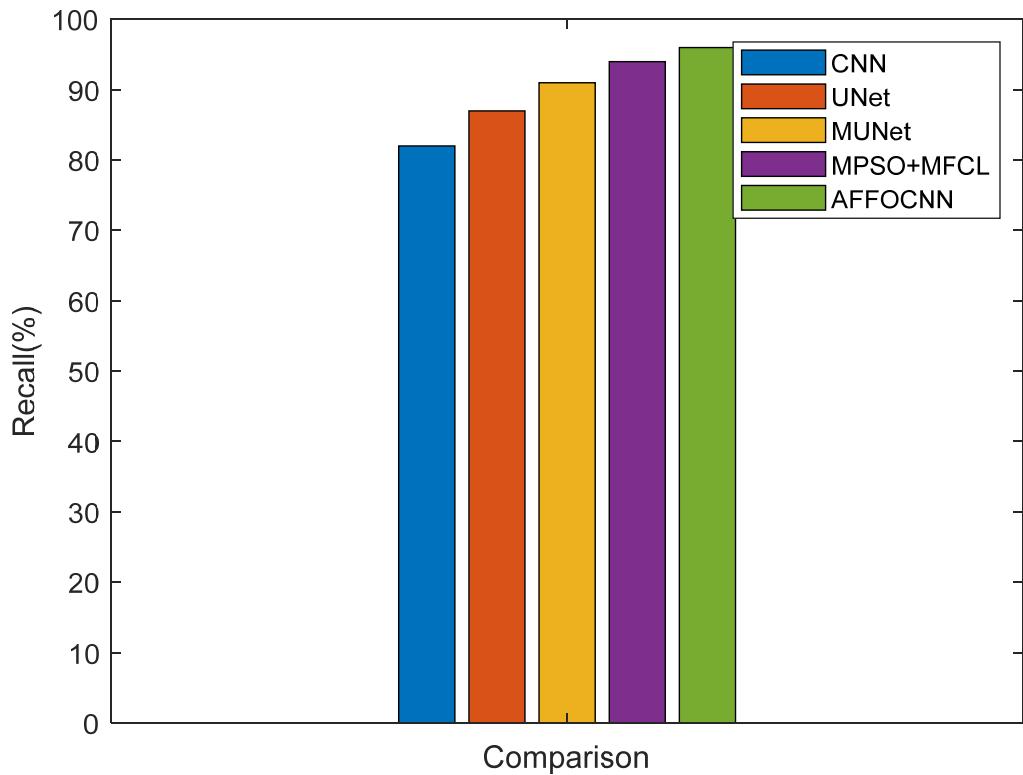


Fig 5.8 Recall

A comparison of the recall metric between the current techniques is shown in Figure 5.8. Recall metrics are less for CNN, UNet, MUNet, and MPSO+MFCL, whereas the suggested AFFOCNN approach shows greater recall. Best-correlated pixels and informative features are used by both RGKMC and AFFOCNN to enhance image quality. The proposed image fusion method creates a single, more informative image by combining pertinent data from several perspective MRI images of the same scene. Consequently, the results show that the AFFOCNN method improves image feature accuracy for image fusion.

F-measure

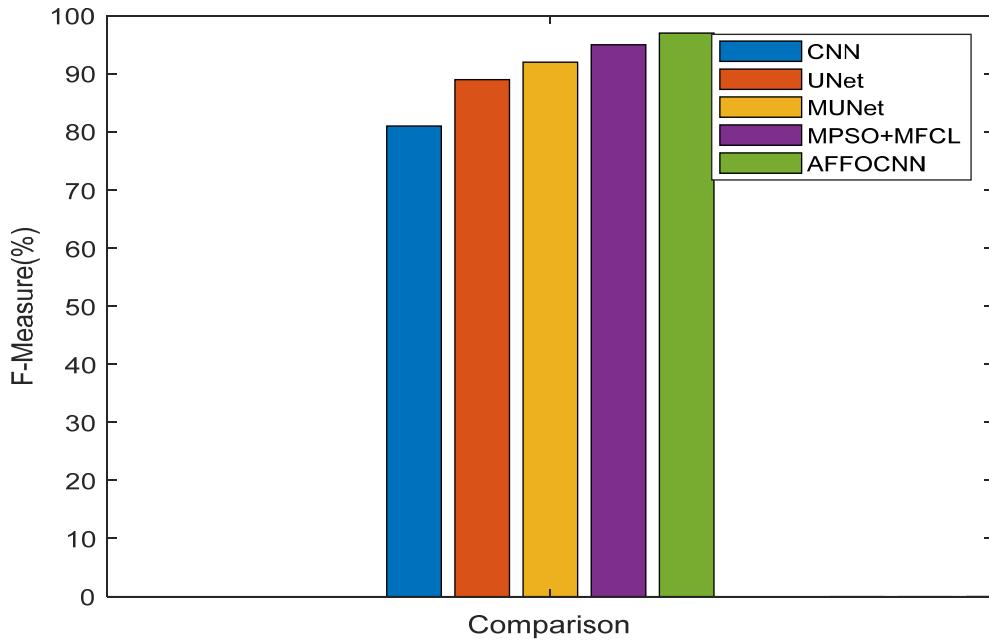


Fig 5.9 F-measure

In Figure 5.9, the comparison of F-measure metrics between existing and proposed algorithms is assessed. Accuracy and recall are both accurately measured by the F1 score, evaluates the validation's overall performance by equally weighing both precision and recall. For the given MRI database, the suggested AFFOCNN algorithm shows greater F-measure values than standard techniques like CNN, UNet, MUNet, and MPSO+MFCL. Specifically, the proposed AFFOCNN algorithm achieves a remarkable F1 score of 97% in prediction, without any misidentified features. By extracting more informative features, the AFFOCNN model effectively distinguishes between affected and non-affected features. Consequently, for MRI datasets, the suggested AFFOCNN method provides better accuracy and improved image fusion performance.

MSE

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (5.13)$$

n = number of data points, Y_i is observed values and \hat{Y}_i is predicted values

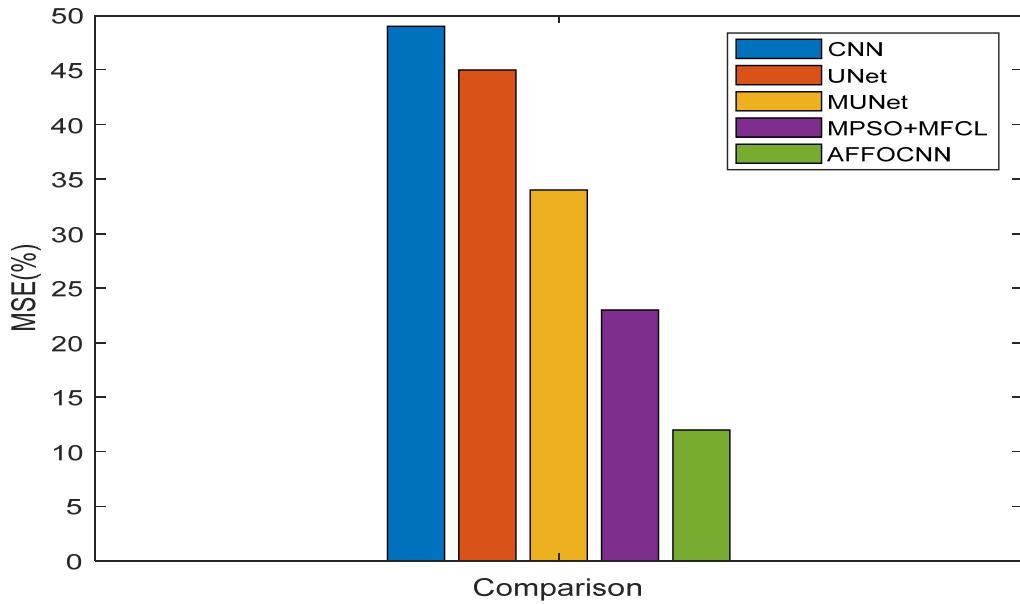


Fig 5.10 MSE

Fig 5.10, shows the comparison of MSE metric for existing methods. The AFFOCNN method proposed exhibits reduced Mean Squared Error (MSE), whereas conventional methods like CNN, UNet, MUNet, and MPSO+MFCL yield higher MSE values. MFCL specifically aims to minimize errors in multimodal medical image fusion and effectively addresses spatial errors. Consequently, the results indicate that the AFFOCNN strategy that has been suggested improves the image fusion process considerably.

Execution time

When the suggested method runs in less time, the system performs better.

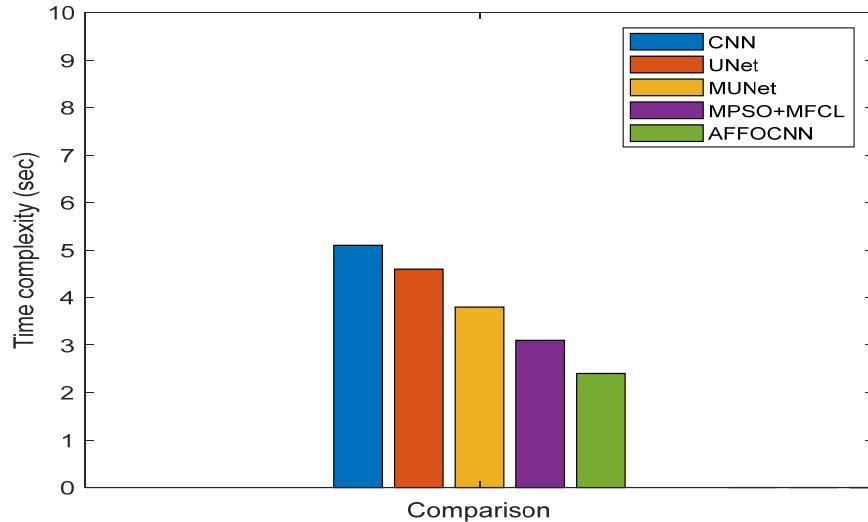


Fig 5.11 Execution time

A comparison of the suggested and current approaches' execution time metrics is shown in Figure 5.11. The suggested AFFOCNN algorithm shows reduced temporal complexity compared to conventional techniques like CNN, UNet, MUNet, and MPSO+MFCL algorithms. The RGKMC approach is used in this study to choose the optimal pixels, and the best fitness function values obtained from the AFFO algorithm are employed to extract more relevant features. Therefore, the proposed strategy optimizes the overall performance of image fusion while simultaneously increasing speed. Thus, the findings show that efficiency is increased by the proposed AFFOCNN algorithm.

5.4 SUMMARY

In this work, AFFOCNN algorithm is proposed to improve the image fusion performance for the given MRI images. Noise reduction, segmentation, feature extraction, and image fusion are the four main elements of this research. The noise removal module aims to enhance image quality, followed by MRI image segmentation using the RGKMC algorithm, which segments black and white images to identify brain tumors in both modalities for early prediction. Feature extraction utilizes the AFFOCNN algorithm to efficiently extract more informative features. Subsequently, for real-time applications, image fusion integrates pertinent and helpful image information. According to experimental findings, when compared to current algorithms, the suggested AFFOCNN method obtains better accuracy, precision, recall, and reduced MSE and execution time.

Chapter 6

CHAPTER 6

ENSEMBLE DEEP LEARNING ALGORITHM FOR MULTI VIEW IMAGE FUSION

6.1. INTRODUCTION

Rapid developments in sensor and computer science technologies have made medical imaging an increasingly crucial component of many clinical applications, including medical diagnosis, surgical navigation, and treatment planning. This technology serves as a critical tool for doctors in accurately diagnosing diseases (Wang et al., 2020) [109]. Brain tumors are solid masses that develop when brain cells grow abnormally due to unchecked cell division. They may be aggressive and invasive, disrupting normal brain functioning and presenting a danger to life. They can also be benign (non-cancerous) or malignant (cancerous). Additionally, fluid buildup around the tumor, known as edema, further compromises healthy tissue function by exerting pressure. Edema is important for diagnosis since radiologists may use it to determine the size and development of tumors. MRI and CT are the most helpful brain imaging modalities for assessing brain malignancies due to their superior soft tissue contrast. In this work, MRI image database are taken for implementation

Medical image fusion is a technique used to enhance the precision of clinical diagnosis by creating a merged image that preserves the crucial characteristics and information from the original images. The spatial domain or the transform domain may be used for multimodal medical image fusion. Fusion rules are used in the spatial domain to merge matching spatial pixels from images like CT or MRI (Zhu et al., 2020) [123].

However, because of insufficient edge and contour recognition, which causes blurring, fused images created in the spatial domain often display inferior quality. One major disadvantage of the spatial domain method is this constraint. In general, accurate representation of edges and contours is crucial in medical image analysis as they define factors such as tumor size or tumor outline.

Typically, medical imaging employs various techniques, each tailored to capture specific tissue or organ details, such as X-ray, CT, and MRI. CT scans excel in accurately locating dense structures like bones and implants, while MRI scans offer detailed soft tissue information alongside high-resolution anatomical details (Hou et al., 2019) [38]. Creating a unified image that captures the unique characteristics of multimodal medical imaging is the primary objective of image fusion. This consolidated image aids doctors in making precise diagnostic decisions across various medical conditions (Ray et al., 2018) [82].

Feature extraction is used as a technique for dimension reduction in the image fusion process. It's recognized as an effective strategy for both reducing computational complexity and enhancing accuracy. Feature extraction provides an additional efficient way to reduce the dimensionality of data in addition to feature selection. A hyperspectral image's feature space is altered by first applying a linear transformation to make it more accurate. Then, for categorization, only the most relevant elements are kept (Zhao et al., 2018) [121]. Principal component analysis (PCA) and independent component analysis (ICA) are two unsupervised and supervised dimensionality reduction approaches that are used. Another supervised methodology is linear discriminant analysis. Among these techniques, PCA stands out for its ability to preserve the majority of information from the hyperspectral image within a small set of significant Principal Components (PCs).

In today's age, numerous methods for selecting features are employed on medical databases to extract the most pertinent information. These methods are essential for making a variety of disease predictions based on medical information. Heuristic and evolutionary approaches are widely used in feature selection strategies to minimize computational complexity. These strategies effectively tackle optimization problems with high dimensions, yielding satisfactory outcomes within reasonable time frames (Bhat and Karki 2017) [17]. Among nature-inspired metaheuristic techniques, swarm-based algorithms are widely accepted. Swarm Intelligence (SI) stands out as an AI-based method that fosters collective behaviors in self-organized and decentralized systems. It comprises a population of simple actors that communicate locally, restricted to their immediate surroundings.

The image fusion method aims to enhance the visualization of both the structure and abnormalities in meningioma brain tumors. It makes use of fuzzy statistics and the contourlet transform. This method mainly highlights the tumor and the high-intensity region around it, resulting in enhanced imaging for radiologists. The approach involves contourlet transform for breaking down images into multiple scales and directions, fuzzy energy based on regions to combine detailed coefficients from two input images with comparable orientations, and fuzzy entropy to combine approximation coefficients. The technique creates two fusion rules to combine the images matching higher and lower frequency sub-bands (Chen et al., 2020) [18].

The primary objective of this study is to enhance image fusion with MRI images, addressing previous limitations in brain tumor detection and image quality. To tackle these challenges, the research proposes the Ensemble Deep Learning (DCNN+DFMI+cGAN)

algorithm, aiming to enhance system performance comprehensively. The key contributions include noise reduction through AMF, segmentation utilizing RGKMC, feature extraction via LSOCNN algorithm, and image fusion employing the EDL approach. This method aims to yield more precise results by leveraging efficient algorithms on the provided MRI image dataset.

6.2 PROPOSED METHODOLOGY

In this work, EDL is presented to considerably enhance image fusion. Fig. 6.1 shows the proposed system's overall block diagram.

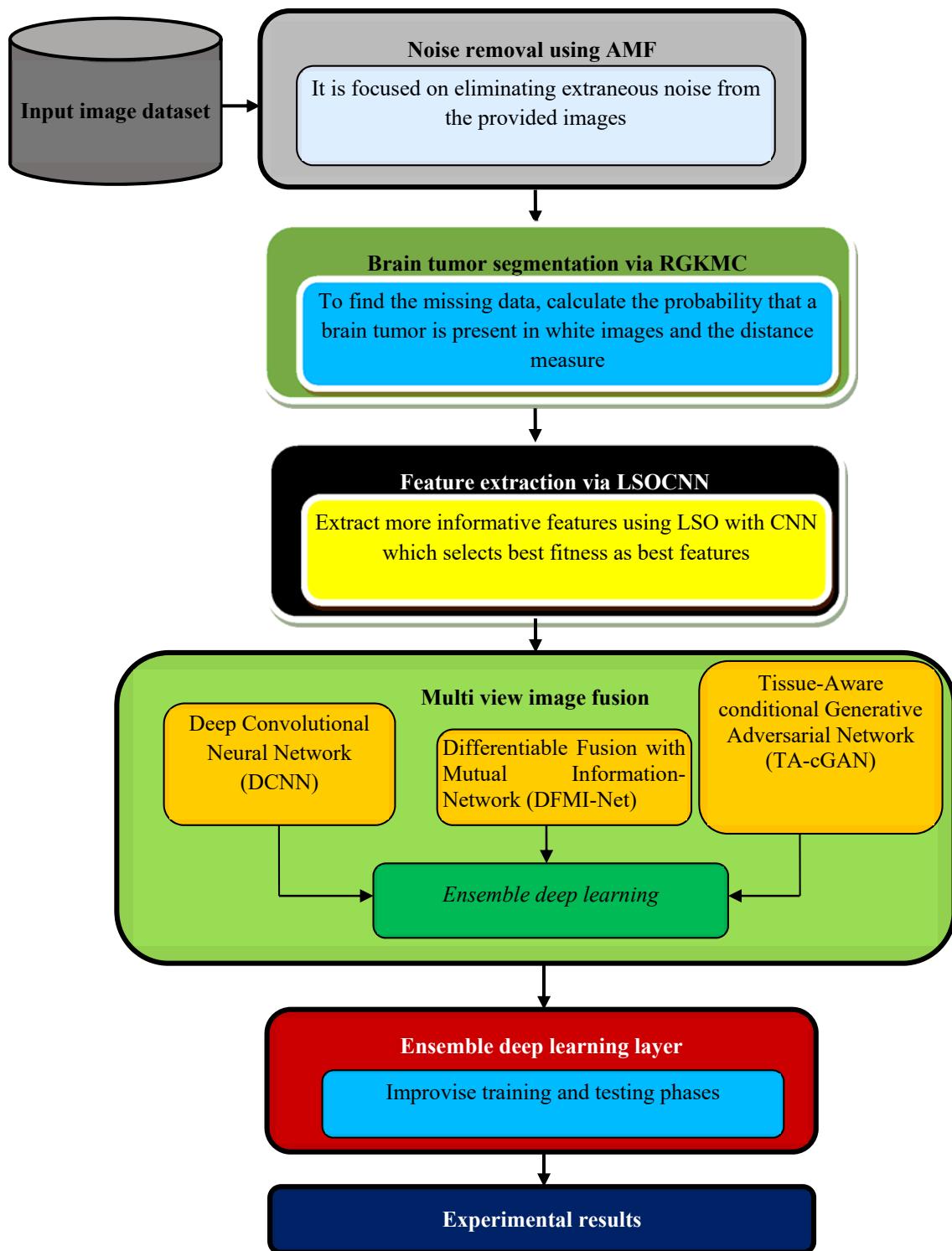


Fig 6.1 The proposed system's overall block diagram

6.2.1. Input image dataset

A set of 3D brain MRIs and expert medical evaluations of brain cancer are available in the BraTS 2018 dataset. The four MRI modalities employed are T1, T1c, T2, and FLAIR. Necrotic and non-enhancing tumor core, enhancing tumor, and peritumoral edema are the dataset's tumor subregions. Three nested sub-regions total tumor, central tumor, and augmenting tumor are additional classifications for these findings. The data originate from 19 institutes and were generated using a variety of MRI scanners.

6.2.2 Noise removal using Adaptive Median Filtering (AMF)

This study employs the Adaptive Median Filter (AMF) algorithm to eliminate unnecessary noise from images, aiming to enhance image quality in the presence of high levels of noise. Spatial processing is used by AMF to detect noisy pixels in an image, classifying them based on a comparison with neighboring pixels. The algorithm allows for adjustable parameters such as neighborhood size and threshold values for the comparison process. Pixels deemed as noise are those that significantly differ from their neighbors and lack structural alignment with similar pixels. These noisy pixels are subsequently replaced with the median value of neighboring pixels that have passed the noise identification criteria (Verma et al., 2015) [105]; (Ibrahim et al., 2008) [43]. Hence, by reducing noise in the image, it is used to enhance image quality. The required content is already discussed in chapter 5

6.2.3 Brain tumor segmentation via Region Growing based K-Means Clustering (RGKMC) algorithm

This study utilizes the RGKMC algorithm for efficient brain tumor segmentation. This process involves distinguishing the tumor from normal brain tissue, crucial for diagnosis and treatment planning in clinical settings. However, due to tumors' irregular shapes and unclear boundaries, segmentation remains challenging. To address this, the RGKMC algorithm is introduced. KMC, a clustering technique, is employed to group similar data using initial centroids, as outlined by (Sulaiman and Isa 2010) [98]. It calculates centroids based on Euclidean distance. RGKMC divides images into black and white representations, enabling tumor identification in both. The algorithm iteratively computes current cluster centers and assigns data items to the cluster with the closest center until no further reassessments occur. This process minimizes intra-cluster variance, enhancing segmentation accuracy. The required content is already discussed in chapter 5

6.2.4 Feature extraction using Lion Swarm Optimization based Convolutional Neural Network (LSOCNN)

To extract more relevant features based on optimum fitness values, this research uses the LSOCNN method. The feature extraction phase involves the utilization of three convolutional layers to delve into the characteristics of source images. In the image fusion approach, this mining of features or features is considered to be essential.

For CNNs, employing efficient initialization and training techniques for convolutional kernels is essential for optimal performance. Thus, utilizing the convolutional layer of GoogLeNet is advisable, particularly because it is well-optimized for ImageNet. The primary convolutional layer (CONV1) within GoogLeNet is robust,

featuring 64 convolution cores of 7×7 dimensions, effectively extracting key image features. Although CONV1 covers a broad recognition range, additional examples can complement its findings, proving that the fusion process does not need every aspect of the image to be completely represented. CONV2 and CONV3 serve as subsequent layers, building upon the features identified by CONV1, ultimately forming a comprehensive feature map for integration. When input images are inadequately sampled, there's a risk of missing important image data, which could adversely impact the feature extraction phase. Hence, it's crucial to adjust the stride and padding values consistently for the kernels of each convolutional layer to ensure accurate processing.

The Lion Swarm Optimization (LSO) algorithm is a meta-heuristic approach known for its ability to produce multiple solutions during each iteration. In the animal kingdom, lions, being the most robust mammals, typically reside in groups led by females. These groups defend their social territory with the dominant male lions. When a wandering lion intrudes upon another's territory, a confrontation occurs between the intruder and the resident lion. The winner of this battle claims the territory, while the loser is driven out. If the intruder wins, they also eliminate the cubs of the resident lion and force mating with the female. As the resident cubs mature, they too may challenge the dominant lion for control of the territory. This natural behavior serves as inspiration for the Lion Swarm Optimization algorithm adapted from (Wang et al., 2019) [110].

The social behavior of lions can be categorized as either resident or nomadic, each with its distinct characteristics. LSO seeks to find the most effective solutions to problems by examining these behaviors, which can manifest as either defensive actions or takeovers. Figure 6.2 illustrates the essence of LSO.

- Territory defense is the process by which local lions and cubs battle against males that travel around to claim territory. The Lion Search Optimization (LSO) method assesses the effectiveness of the current territorial solution against a new nomadic solution. If the nomadic approach proves superior, it supplants the territorial lions or the existing solution.
- Territorial Takeover, a behavior demonstrated by LSO, include removing everything but the best male and female options and keeping just selected.

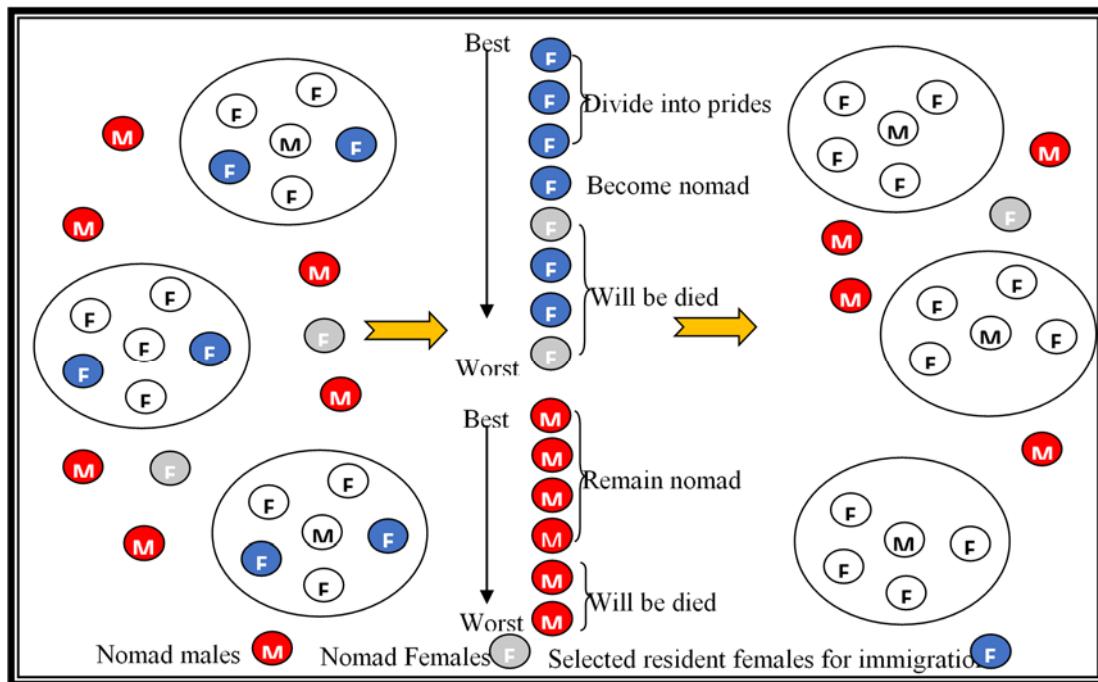


Fig 6.2 Nature of LSO

Initialization: In solutions, lions are symbolized by: Starting at random, a population of lions is created for the LSO initialization procedure. This population is then stored as the solution space in matrix form.

$$Lion = [x_1, x_2, \dots, x_{N_{var}}] \quad (6.1)$$

Where $x_{N_{var}}$ are the pixels in the generic image selected. The remainder are resident lions, and N is the randomly produced nomad lion ratio. The LSO solves the problem of selecting pixels from inputs and finds hidden connections between network components.

Fitness computations

Lions in the sorted and saved matrix have their fitness levels evaluated by an objective function.

$$f(Lion) = f(x_1, x_2, \dots, x_{N_{var}}) \quad (6.2)$$

Higher classification accuracy serves as the basis for fitness values. The majority of the characteristics needed for this fitness function include increased precision, as shown in

$$Fitness = Max \frac{\sum_{i=1}^N (P_{neighborhood\ pixel\ with\ higher\ accuracy}^i + P_{quality\ pixels}^i)}{2} \quad (6.3)$$

Where, $P_{neighborhood\ pixel\ with\ higher\ accuracy}^i$ – closer neighborhood pixels and $P_{quality\ pixels}^i$ – quality image pixels

When computing fitness values, they undergo updates through operations such as hunting, mating, roaming, and defense. Female lions focus on hunting prey within their territory, thereby housing the optimal solution within it. The selection of feature extraction in this study relies on the fitness function values and updates outlined above.

Hunting Procedure: This process involves three categories of hunters, with the one possessing the highest fitness value positioned at the center, while the other two flank it. Hunters are randomly selected to pursue dummy prey, which evade capture as the hunter's fitness improves and prompts relocation to a new area.

Safety-Oriented Movement: Only specific female lions engage in hunting, while others remain within secure territory. Within each region, the ideal locations are determined and noted. Lower success rates imply opportunities for improvement, while higher success rates show that lions have migrated away from the optimal position. Thus, evaluating competition outcomes serves as a gauge of achievement.

Lion Roaming: Roaming poses a challenging task and limits the lions' search efforts. Nonetheless, it is utilized by the LSO to explore the search space and uncover improved solutions. Lions advance by n units towards their preferred territory during this phase.

$$n \sim U(0, 2 * d) \quad (6.4)$$

Where ' n ' represents a randomly generated number with a uniform distribution, and ' d ' signifies the distance of a male lion from the selected territory. Within the search area, nomadic lions also move at random.

Mating: Lions reproduce by mating, which results in the birth of new generations. Cubs are born when the right male and female lions are found. Through crossings and mutations, this process creates new and better solutions from the ones that already exist. Taking out the weaker lions guarantees that the optimal answers are found.

Defense: Lions respond differently when they are on defense. Mature male lions engage in combat with one another. The victorious lions claim the territory of the

defeated nomadic lion in these competitions. The LSO utilizes two protective measures to safeguard lions from both inexperienced local males and migratory males. That being said, the LSO decides who is the strongest lion in the group.

Migration: Fitness values affect the rearrangement of randomly chosen females throughout the movement process known as migration, which results in their becoming nomads. The fittest females replace the positions of those who have become nomads.

Population Equilibrium of Lions: After every iteration, the population of lions is assessed to determine the maximum number of individuals based on gender. This assessment aims to identify an equilibrium point or stable position for the lion population. Lions with a low fitness value are eliminated from the population.

Termination Criteria: The proposed algorithm terminates when a higher number of iterations achieve the best fitness values. The feature with the highest fitness value is deemed the best solution and is extracted to enhance classification accuracy and image fusion performance. Lions with lower fitness ratings are replaced by those with higher fitness values during territorial conquest. Thus, only male-female solutions are kept and others are rejected.

Algorithm 6.1: LSOCNN

Input: MRI image features

Output: More informative features

1. Initiate the algorithm.
2. Organize the Database.

3. This stage involves selecting and categorizing the database to be managed and classified for the CNN. It is crucial to annotate all the features.
4. Define the objective function (x), where $x = (x_1, \dots, x_T)$, aiming for lower Mean Squared Error (MSE), reduced execution time, and higher accuracy as objective functions.
5. Initialize the lions and assign them random places.
6. Configure the CNN representation, incorporating factors determined by the LSO (such as the quantity of convolution layers, filter dimensions, number of convolution filters, and batch dimensions). This setup equips the CNN to adapt to the input data efficiently.
7. CNN preparation and validation - In the preparation, authentication, and testing phases of processing the input databases, CNN makes use of images. This phase determines the degree of acknowledgment.
8. Evaluate the function of the aim. To get the optimal value, the LSO technique computes the objective function.
9. Utilizing equation (6.3), determine each feature's fitness.
10. To assess quality, choose the first aspect and contrast it with others.
11. Choose the image feature pixels with the highest accuracy and lowest MSE & execution time.
12. Record the best image pixels.
13. For each feature:

14. Identify the best solutions.
15. End loop.
16. Reorder features and highlight the best ones.
17. Integrate the new best feature with the old one.
18. Arrange all features based on their fitness.
19. Discard weaker features, retaining those that are optimal.
20. If termination criteria are met,
21. Return the best selection of features.
22. Else,
23. $t=t+1$
24. Until $t > \text{Max_Iteration}$.
25. Return the optimal image features.
26. Select the ideal resolution last. The CNN model's optimal attributes are chosen by the lions in this stage.
27. Stop the algorithm.

Algorithm 6.1 outlined above details the LSO process. In summary, lions, renowned as the strongest mammals, exhibit social conduct that falls into one of two categories: domestic or nomadic. Lions can switch between these behaviors. Residential lions occupy a specific zone where they engage in mating for reproduction. On the other

hand, nomadic lion movements are irregular, with individuals or pairs traveling, particularly when outmatched males band together.

6.2.5 Image fusion of multi modal image

The objective of image fusion is to retain valuable complementary information while eliminating redundant data and distortions found in individual source images. Therefore, performance measures are essential for assessing how well fusion strategies work and comparing results from various algorithms. Precise identification of the kind, dimensions, position, and scope of a brain tumor is crucial for the diagnosis and treatment of brain cancer. The fusion of MRI brain imaging enables rapid tumor detection and diagnosis, resulting in improved results compared to standalone MRI scans. In certain scenarios, telemedicine and diverse telehealth applications are revolutionizing various healthcare sectors by enabling remote diagnosis and prompt first aid administration. Feature extraction serves as a broad term encompassing methods aimed at generating diverse combinations to address these challenges while accurately representing the data.

6.2.5.1 DCNN algorithm

For each patient, four different mammography images were used to train a deep convolutional neural network architecture. The use of the mini-MIAS and CBIS-DDSM datasets facilitated the completion of this task. The CNN model analyzes many attributes from each of the four mammogram representations. During training, early fusion combines data from several perspectives. The early fusion process combines numerous feature vectors into one. The CADx system, which combines features from the four views, employs two strategies as described by (Khan et al. 2019) [50]. Previous to finalizing the network, extensive experiments were conducted, varying hyperparameters

to select the optimal classification network. Various hyperparameter combinations were explored to determine the best-performing classification network.

1. It increased the number of convolutional layers from three to seven during the first round of feature extraction. After all convolutional layers, we carried out experimental evaluations using max pooling and average pooling.
2. Across convolutional layers and the last two fully connected layers, we explored ReLU and Leaky ReLU activation functions. In the final classification layer, we evaluated sigmoid and softmax functions.
3. As for optimization algorithms, we experimented with stochastic gradient descent using various initial learning rates ranging from 1e-2 to 1e-5. We adjusted the momentum value within the range of 0.7 to 0.9.
4. Dropout was implemented after different convolutional and fully connected layers, exploring dropout ratios ranging from 15% to 50%.

An input layer, numerous hidden layers, and an output layer comprise a DCNN. Convolutional, pooling, and completely linked hidden layers are common. Convolutional layers process the input by simulating how individual neurons respond to visual input. Pooling layers, whether local or global, amalgamate neuron outputs from one layer into a single neuron in the next layer. For example, mean pooling determines the average value from every cluster of neurons in the layer above. All of the neurons in one layer are linked to all of the neurons in another layer via fully connected layers. Conceptually, the DCNN mirrors the traditional multi-layer perceptron neural network. It has layers for categorization, convolution, and input. There are clear benefits to using this method for

examining complicated data by implementing parameter sharing within convolutional layers to manage and diminish parameter count.

6.2.5.2 Differentiable Fusion with Mutual Information-Network (DFMI-Net)

The architecture of the auto encoder is influenced by the Fusion W-Net (FW-Net) introduced by (Fan et al., 2019) [25]. Using its local connection structure, FW-Net expands upon the U-Net architecture (Ronneberger et al., 2015) [84] by aligning output vectors and producing an aesthetically pleasing image fusion. Additionally, FW-Net can capture semantic information from source images, such as variations in brightness, and map them to a unified semantic space across different views. This process also integrates structural details, edge information, and crucial diagnostic features into the same semantic space. The encoder ($E\phi$) creates a fused picture that is the same size as the source images by using a pair of spatially and temporally aligned images as input. Subsequently, the decoder ($D\phi$) reconstructs two images from the fused output. Include the loss function if provided:

$$Loss = \sum_{n=1}^N [\alpha_n \cdot \mathbb{E}[|x_n - \hat{x}_n|] + \beta_n \cdot MINE(x_n, x_f)] \quad (6.5)$$

Where x_n denotes the input images, \hat{x}_n denotes the corresponding reconstructed images, and x_f denotes the fused image. The Mean Absolute Error (MAE) loss, which measures the pixel-by-pixel difference between the input and reconstructed images, is represented by the first part of the equation (6.5). Reducing this loss helps the encoder maintain important characteristics from the source images in the combined image, as mentioned by (Ting et al., 2020] [101]. The subsequent term pertains to the Mutual

Information (MI) loss, computed using MINE-Net to assess the mutual information exchanged between the input and fused images. Maximizing this MI loss facilitates the synthesis of a superior fused image.

6.2.5.3 Tissue-Aware conditional Generative Adversarial Network (TA-cGAN)

The TA-cGAN comprises two sub-networks: the generator and the discriminator. Unlike the traditional GAN, which primarily focuses on translating one image to another, TA-cGAN is intended for translating multiple images into a single output image. Specifically, it takes multiple input images, such as an MRI image (IM) and its corresponding label map (IL), and produces a single fused image (IF). The detailed architecture of the TA-cGAN network is described below.

Generator

This method's generator is based on the U-Net design, which combines high-level data from the deeper levels of the decoder with low-level characteristics from the encoder's shallow layers via the use of skip connections. This technique helps mitigate the issue of gradient vanishing. U-Net is effective in image synthesis because of its skip connections (Wang et al., 2018) [111] G, the generator network in our investigation, has an encoder and a decoder (Figure 6.3). The network generates a fused image (IF) as output after receiving an MRI image (IM) and the label map (IL) that corresponds to it as inputs. In particular, there are a total of 12 convolutional layers in the generator. The encoder uses 6 down-sampling layers, each with 3x3 filters, BN, and ReLU activation with a negative slope of 0.2. To retain feature map spatial resolution and MRI image detail, pooling techniques are avoided. Each down-sampling layer also employs 1×1 zero padding. The decoder has six up-sampling stages: the last layer convolutionally

employs a 1×1 filter, in the first five layers, convolution-BN-ReLU procedures are performed. The feature maps in the encoder and decoder are concatenated in the decoder via skip connections, as shown by the dotted arrows in Figure 6.3.

Discriminator

The discriminator D, in contrast to the generator G, is mainly designed to handle classification jobs. The main objective of the discriminator in this work is to distinguish between the MRI image pair (the MRI image IM and label map IL) and the fused image pair (the fused image IF and label map IL). The discriminator D architecture used in this study is shown in Figure 6.4 (Goodfellow et al., 2020) [33], (Kang et al., 2020) [48]. According to Figure 6.4, the discriminator outputs a class label after receiving as inputs either the fused image pair or the MRI image pair, indicating whether the input is distinguished (labeled as 1) or not (labeled as 0).

6.2.5.4 Ensemble Deep Learning (EDL)

This section couples DCNN, DDFMI-Net, and TA-cGAN to enhance multi-view image fusion. The TA-cGAN method is improved using ANN. The general layout of the discriminator is seen in Figure 6.4. One fully connected layer, five convolutional layers, and a sigmoid activation function make up the discriminator D, a basic convolutional neural network. Comparable to the encoder structure of the generator G, these convolutional layers perform convolution-BN-ReLU processes.

In this approach, a conditional generative adversarial network (cGAN) is utilized to achieve MRI image fusion. The MRI fusion process is conceptualized as a two-player adversarial game involving a generator and a discriminator. A tissue label map obtained

from the MRI scans serves as the basis for both the discriminator and the generator. High spatial resolution and comprehensive structural details about soft tissues are provided by MRI technology, along with color data that reveals the functional characteristics of the tissues. This technique aids doctors in diagnosing diseases more effectively. Specifically, the TA-cGAN method is employed for fusing brain MRI images.

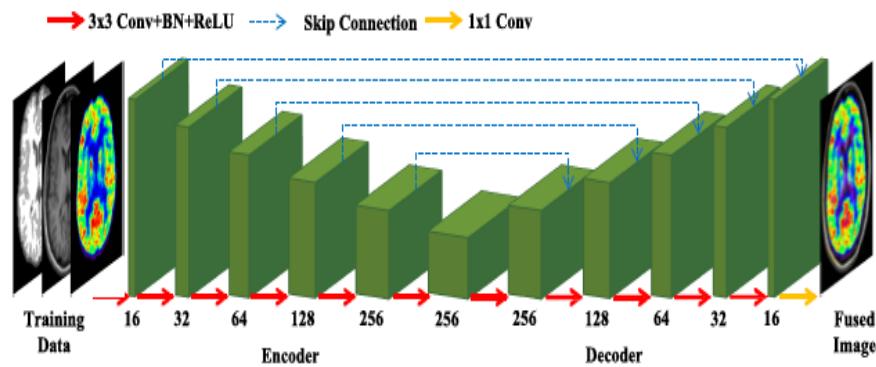


Fig 6.3Network architecture of the Generator

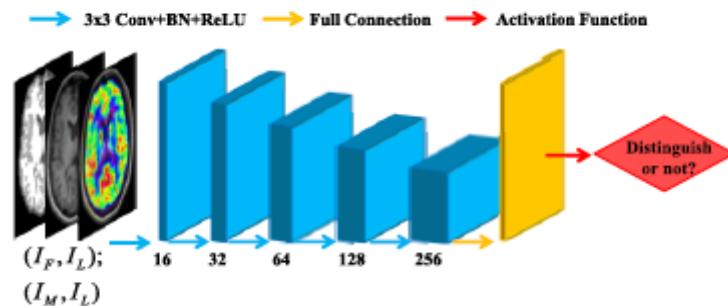


Fig 6.4 Network architecture of the discriminator

The three phases of Artificial Neural Network (ANN) operation are the input layer, the hidden layer, and the output layer. ANNs learn via experience. After input images are gathered and analyzed, a predetermined number of inputs are produced by the input layer. This processing is guided by certain weights, which contain the information necessary for problem-solving in neural networks (Mohamed and El-Bhrawy 2016) [77].

The hidden layer extracts concealed information from the input layer and relays it to the output layer after performing some valuable extraction. ANNs are utilized for detecting quality image features, where an MRI image dataset is trained, and during testing, these features are classified.

Training of the Discriminators

Every T epochs, for a fixed φ , the generator generates $2n$ batches $\{B_{a1}, \dots, B_{an}, B_{g1}, \dots, B_{gn}\}$ quality points that are not obtained from the real datasets, or b pixel points. Additionally, from each dataset D_i that is accessible, the discriminator at dataset i samples a batch B_{r_i} of b points. The following loss value is computed by each dataset once each created batch from the generator is sent to it.

$$L_i(\theta_i) = \frac{1}{b} \left[\sum_{x \in B_{r_i}} \log D_{\theta_i}(x) + \sum_{x \in B_{a_i}} \log(1 - D_{\theta_i}(x)) \right] \quad (6.6)$$

$L_i(\theta_i)$ explains how each dataset discriminator in (6.6) has its value function approximated. Next, by using the Adam optimizer a gradient descent technique (Parwez et al., 2019) [79], the dataset updates its own weights θ_i

Training of the Central Generator

Every T epochs, every database i uses B_{gi} to calculate the following loss:

$$L_i^g = \frac{1}{b} \left[\sum_{x \in B_{gi}} \log(1 - D_{\theta_i}(x)) \right] \quad (6.7)$$

which, for each database generator in (6.7), is an estimate of the value function.

TA-cGAN utilizes joint losses including spectral loss L_{Spec} , structural loss L_{Str} and adversarial loss L_{Adv} . This is the mathematical expression for the joint loss that was employed in the paper.

$$L_{Joint} = \lambda_1 L_{Spec} + \lambda_2 L_{Str} + \lambda_3 L_{Adv} \quad (6.8)$$

where spectral deficit L_{Spec} insists the merged image has the same color information (MRI pixel intensities); Lost structure The L_{Str} attempts to make the fused image have comparable structural information (MRI gradients); Enemy loss L_{Adv} adds detail to the merged image; Weights for spectral, structural, and adversarial losses are defined as λ_1, λ_2 and λ_3 .

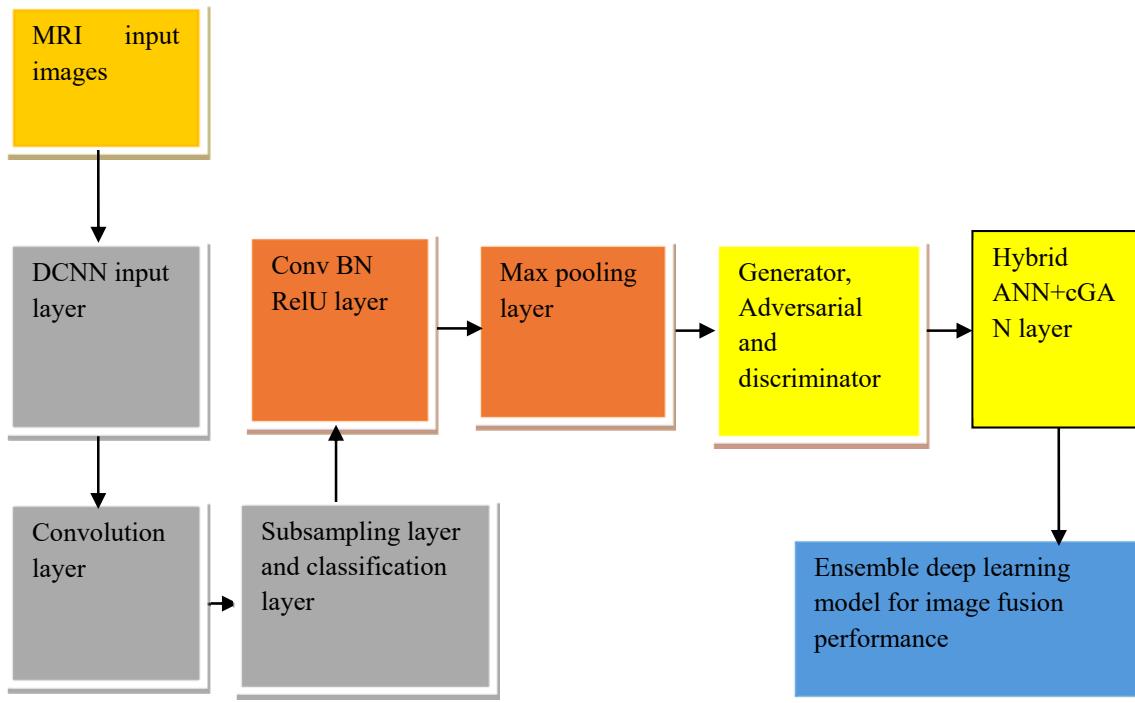


Fig 6.5EDL algorithm

The discussed ensemble framework describes the input MRI feature-extracted image dataset is sent into the input layer of the DCNN. In this context, significant characteristics are produced by the convolutional layers. These features then pass through the subsampling and classification layers, producing the output results. Subsequently, Conv BN ReLU and max pooling process characteristics. To improve multi-view image fusion, the GAN layer uses an ANN model to anticipate and extract more useful features. Figure 6.5 illustrates the EDL algorithm.

6.3. EXPERIMENTAL RESULT

The BraTS 2018 dataset offers segmentations of brain cancer that are analyzed by medical professionals in addition to multi-modal 3D brain MRIs. T1, T1c, T2, and FLAIR are the four MRI modalities used in each instance. Three tumor subregions are included in the dataset: the necrotic and non-enhancing tumor core, the enhancing tumor, and peritumoral edema. Existing algorithms like DCNN, DFMI-Net, and TA-cGAN are taken into consideration and assessed together with the proposed LSO-based EDL method for performance metrics including accuracy, MSE, precision, recall, F-measure, and execution time.

Accuracy

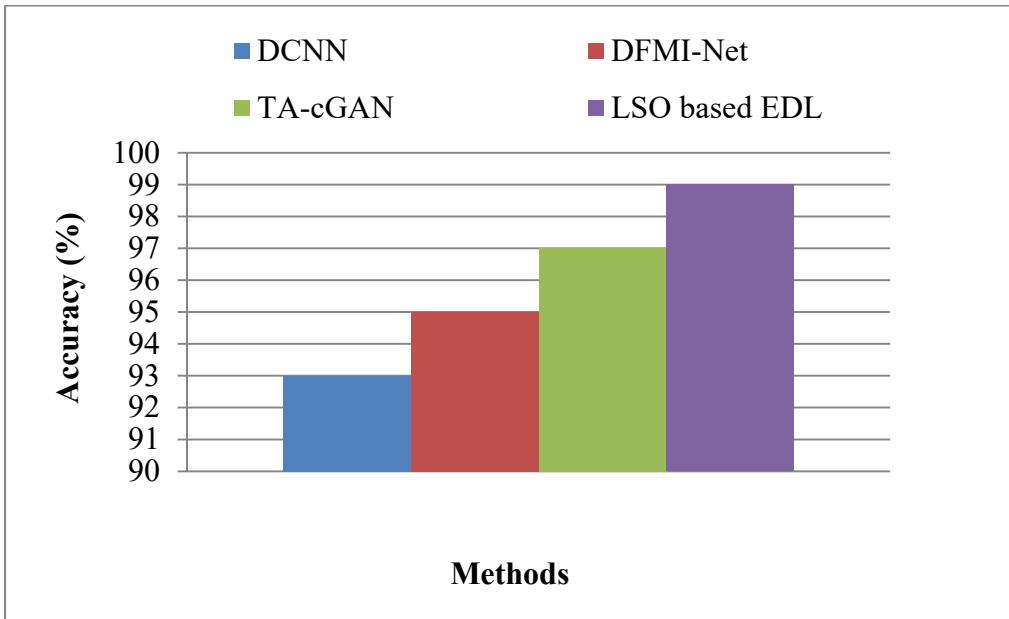


Fig 6.6 Accuracy

From Figure 6.6, it is evident that the comparison metric is assessed based on accuracy for both existing and proposed methods. The x-axis denotes the various procedures, while the y-axis indicates the accuracy levels. The LSO-based EDL algorithm outperforms existing approaches such as DCNN, DFMI-Net, and TA-cGAN algorithms in terms of accuracy, which achieves higher accuracy for the given MRI datasets. The LSOCNN algorithm enhances the quality of image fusion by extracting more meaningful features. Additionally, the recommended segmentation approach, based on region growth, increases image quality by taking into account the surrounding pixels. Consequently, the results indicate that the proposed LSO-based EDL algorithm significantly enhances overall performance.

Precision

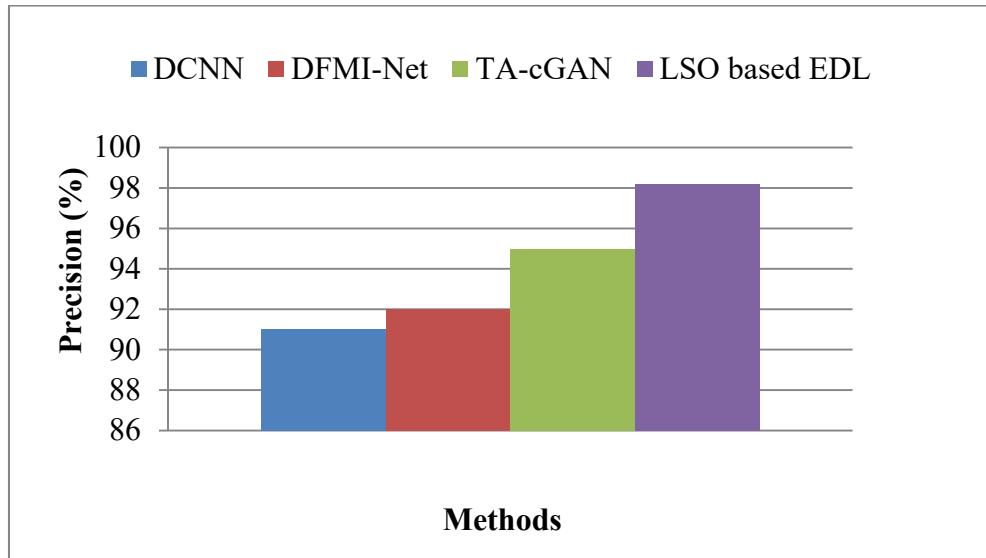


Fig 6.7 Precision

Figure 6.7 shows that precision is the comparison metric evaluated for various methods. The accuracy values are shown on the y-axis, while the techniques are listed on the x-axis. While current approaches like DCNN, DFMI-Net, and TA-cGAN demonstrate lesser accuracy, the proposed LSO-based EDL method achieves greater precision. The proposed image fusion technique aims to provide a consolidated and enhanced image by integrating relevant data from many prospective MRI scans of the same scene. Therefore, the findings suggest that the suggested LSO-based EDL method greatly improves the precision of image features in the image fusion procedure. Furthermore, the ANN with the cGAN technique specifically aims to enhance the quality of multi-view image fusion by optimizing its characteristics.

Recall

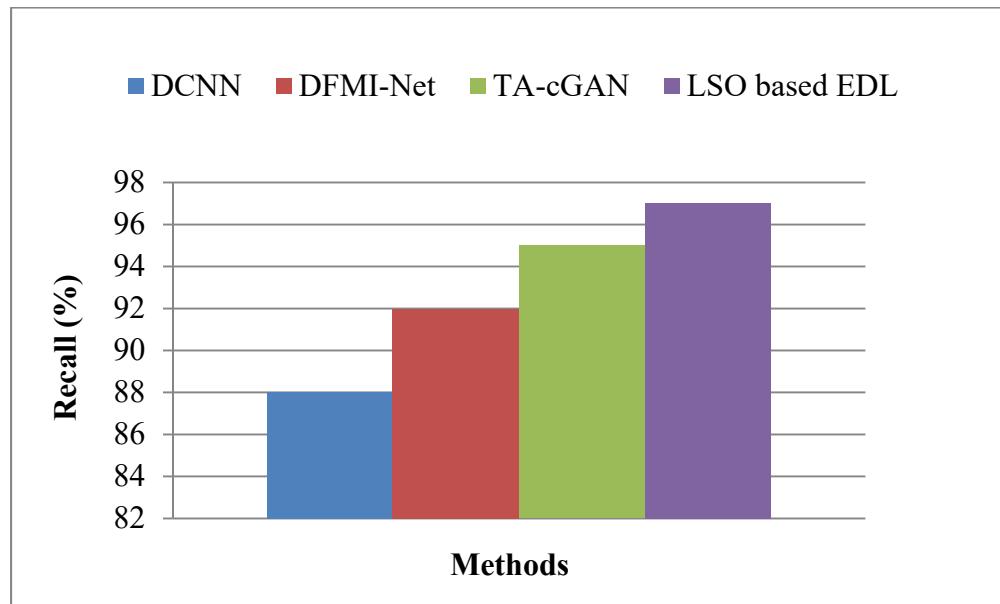


Fig 6.8 Recall

Based on Figure 6.8, it's evident that the comparison metric assesses existing methods concerning recall. The x-axis represents the methodologies, while the y-axis shows the recall values. The suggested LSO-based EDL technique delivers a greater level of recall compared to current methods such as DCNN, DFMI-Net, and TA-cGAN, which exhibit a lower level of recall. The RGKMC and LSOCNN algorithms aim to enhance image quality by utilizing the most correlated pixels and extracting informative features, respectively. When combining relevant data from several view MRI images of the same scene, the suggested image fusion approach aims to create a single, more useful image. Consequently, the findings suggest that the proposed LSO-based EDL approach effectively enhances image features for the image fusion process.

F-measure

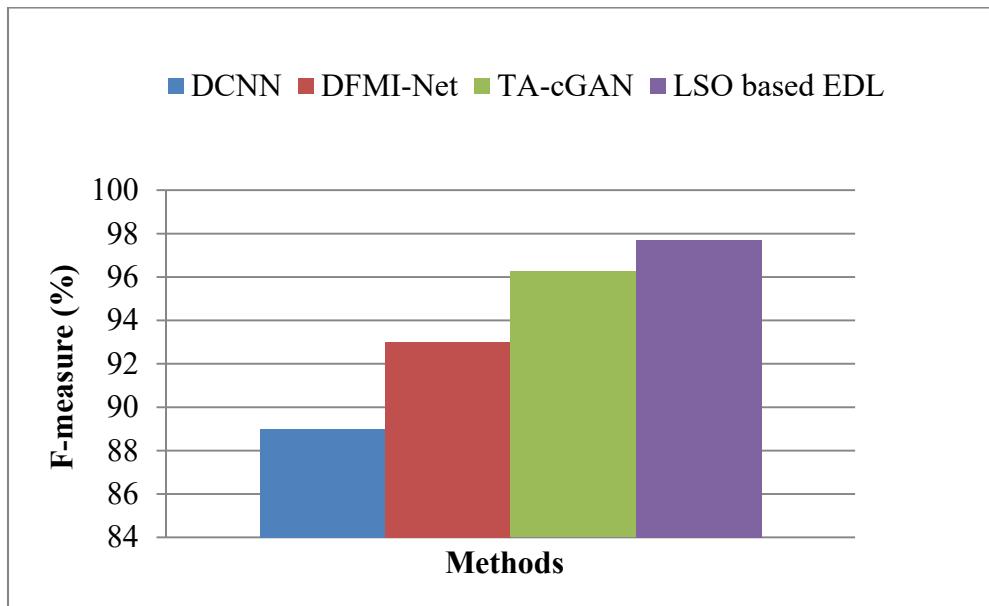


Fig 6.9 F-measure

From Figure 6.9, the F-measure metric is utilized to compare existing and proposed algorithms. Existing methods like DCNN, DFMI-Net, and TA-cGAN show lower F-measure values for the provided MRI database. Conversely, the proposed LSO-based EDL classifier achieves an impressive F1 score of 97.7% in prediction, without any incorrectly identified features. Finding more informative features using the LSOCNN model helps identify impacted and non-affected features. In conclusion, the ensemble deep learning technique improves MRI database multi-view image fusion.

MSE

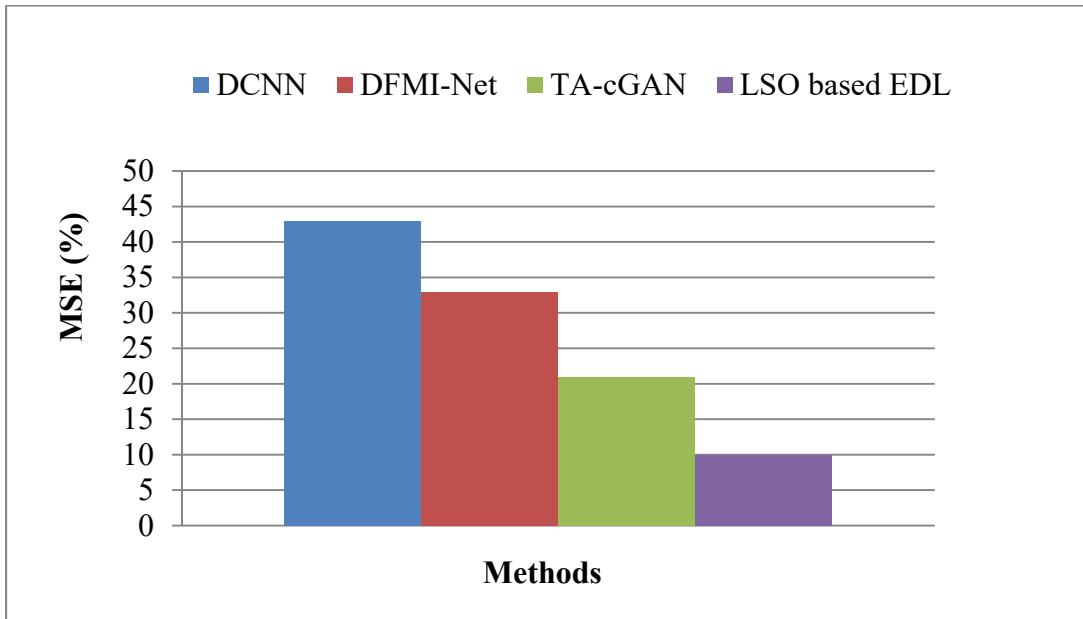


Fig 6.10 MSE

Based on Figure 6.10, it's evident that existing methods are evaluated in terms of Mean Squared Error (MSE). The x-axis represents the methods, while the y-axis displays the MSE values. The proposed LSO-based EDL method achieves lower MSE, while existing methods such as DCNN, DFMI-Net, and TA-cGAN demonstrate higher MSE. The LSOCNN focuses on effectively reducing errors in multimodal medical image fusion. Consequently, the results indicate that multi-view image fusion performance is greatly improved by the suggested ensemble deep learning technique.

Execution time

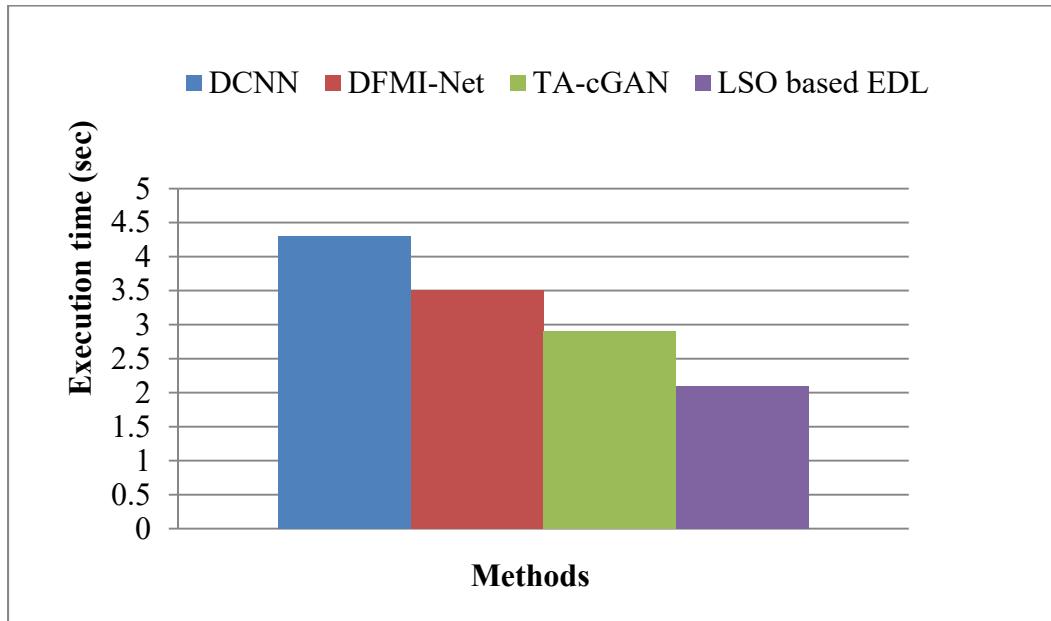


Fig 6.11 Execution time

Figure 6.11 demonstrates that the comparison measure evaluates the execution times of both the recommended and present methodologies. The x-axis represents the methods, while the y-axis represents the execution time values. The suggested LSO-based EDL approach shows reduced temporal complexity in comparison to existing techniques such as DCNN, DFMI-Net, and TA-cGAN. In this study, the optimal pixels are chosen using the RGKMC method, and the best fitness function values of the LSO algorithm are used to extract more relevant features. The suggested method maximizes total picture fusion performance by increasing speed. As a result, the results show that the suggested ensemble deep learning method increases effectiveness.

6.4.SUMMARY

This study proposes the use of the EDL algorithm to enhance the performance of multi-view image fusion for MRI images. The research comprises four primary modules: noise removal, segmentation, feature extraction, and image fusion. Noise removal aims to enhance image quality, followed by segmentation using the RGKMC algorithm, which identifies brain tumors in black and white images to enable early prediction. Feature extraction is conducted through the LSOCNN algorithm to efficiently extract informative features. Subsequently, image fusion is the process of combining important and pertinent picture characteristics for applications that need immediate processing, made possible by the EDL algorithm. The EDL method surpasses previous algorithms in accuracy, precision, recall, MSE, and execution time, according to experiments.

Chapter 7

CHAPTER 7

CONCLUSION AND FUTURE WORK

7.1. CONCLUSION

Accurate diagnosis and treatment of brain tumors rely on detailed information from medical imaging modalities like Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). Each modality provides partial and sometimes ambiguous information about the disease. Therefore, the fusion of images from different modalities becomes crucial to enhance overall image quality. Particularly in medical imaging, for accurate clinical diagnosis, multi-modality medical images must be combined to improve anatomical and spectral information.

In this research work, early and accurate detection of brain tumor from medical images is the objective. As earlier said, fusion of images in the required specific information is the criteria. For achieving this task, the information needs to be collected from the images through segmentation and feature extraction. Some images require pre-processing like noise removal before the above processes. All the above factors have been considered to improve the overall performance of fusion. The task has been achieved through modification of existing methods as well as by developing new algorithms in four phases.

Modified U Net algorithm has been developed in the first phase. It is designed to optimize attribute selection in images for more precise segmentation. Another advantage of this method is that the image does not require any pre-processing while using the images segmented by M U-Net, the outcome of the fused images gave promising results

in terms of accuracy, precise location of diseased part in lesser time. Compared to the existing methods, while testing in single modal image viz. MRI of brain.

In the second phase, the task has to improve the quality of fused images obtained from different modalities, through modifications in Deep Learning method, which requires Segmentation, feature extraction and classification to produce a more precise training data set. The M U-Net is used to segment the images. Modified Particle Swarm Optimization is developed to improve the feature extraction of CNN method which consequently improves the accuracy of training data set. The experimental analysis with multi-modal images for fusion has shown better accuracy and precision compared to the existing methods.

The Deep Learning algorithm is improved in the third phase through Adaptive Fire fly optimization to extract more informative and important features from the images. Region Growing based K-Means Clustering algorithm is used for segmentation which precisely locate any non-standard shapes compared to M U-Net. However, noise removal is required before segmentation. The advantage is this method performs better in real time application.

In the final phase, Multi view image fusion is improved. For extracting specific features, Lion Swarm optimization CNN is used apart from the methods for noise removal and segmentation in the previous phase. To improve the accuracy of the training dataset, three Deep Learning algorithms namely DCNN, DFMI- Net and TA-cGAN are ensembled. When ensembled, the above methods improve the precision of the Deep Learning layer in the Fusion algorithm particularly in fusion of multi- view images.

The general conclusion is that the Deep Learning algorithm used in clinical diagnosis of brain tumour has been improved through the various phases of this research work. The fact has been substantiated and provided through experimental analysis using standard image sets.

7.2. FUTURE WORK

The ensembled Deep Learning algorithm developed and implemented in this research work can be extended for detection and diagnosis of other diseases through more advanced training models for extracting the required and relevant features.

References

REFERENCES

1. Abd Khalid, N.E, Ismail, M.F, Manaf, M.A.A, Fadzil, A.F.A and Ibrahim, S, (2020). MRI brain tumor segmentation: A forthright image processing approach. *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 3, pp. 1024-1031.
2. Abdulkareem, M.B, (2018). Design and development of multimodal medical image fusion using discrete wavelet transform. *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp. 1629-1633.
3. Agrawal, R, Sharma, M and Singh, B.K, (2018). Segmentation of brain lesions in MRI and CT scan images: a hybrid approach using k-means clustering and image morphology. *Journal of the Institution of Engineers (India): Series B*, vol. 99, pp. 173-180.
4. Alagarsamy, S, Kamatchi, K, Govindaraj, V and Thiagarajan, A, (2017). A fully automated hybrid methodology using Cuckoo-based fuzzy clustering technique for magnetic resonance brain image segmentation. *International journal of Imaging systems and technology*, vol. 27, no. 4, pp. 317-332.
5. Alam, M.S, Rahman, M.M, Hossain, M.A, Islam, M.K, Ahmed, K.M, Ahmed, K.T, Singh, B.C and Miah, M.S, (2019). Automatic human brain tumor detection in MRI image using template-based K means and improved fuzzy C means clustering algorithm. *Big Data and Cognitive Computing*, vol. 3, no. 2, pp. 1-18.

6. Ali, M, Shah, J.H, Khan, M.A, Alhaisoni, M, Tariq, U, Akram, T, Kim, Y.J and Chang, B, (2022). Brain tumor detection and classification using pso and convolutional neural network. *Comput. Mater. Continua*, vol. 73, no. 3 pp. 4501-4518.
7. Altaei, M.S.M and Kamil, S.Y, (2020). Brain tumor detection and classification using SIFT in MRI images. In *AIP Conference Proceedings*, vol. 2292, no. 1.
8. Amin, J, Sharif, M, Yasmin, M and Fernandes, S.L, (2020). A distinctive approach in brain tumor detection and classification using MRI. *Pattern Recognition Letters*, vol. 139, pp. 118-127.
9. Amin-Naji, M, Aghagolzadeh, A and Ezoji, M, (2019). Ensemble of CNN for multi-focus image fusion. *Information fusion*, vol. 51, pp. 201-214.
10. Arif, M and Wang, G, (2020). Fast curvelet transform through genetic algorithm for multimodal medical image fusion. *Soft Computing*, vol. 24, no. 3, pp. 1815-1836.
11. Arora, A, Jayal, A, Gupta, M, Mittal, P and Satapathy, S.C, (2021). Brain tumor segmentation of MRI images using processed image driven u-net architecture. *Computers*, vol. 10, no. 11, pp. 1-17.
12. Aymaz, S and Köse, C, (2019). A novel image decomposition-based hybrid technique with super-resolution method for multi-focus image fusion. *Information Fusion*, vol. 45, pp. 113-127.
13. Azam, M.A, Khan, K.B, Ahmad, M and Mazzara, M, (2021). Multimodal Medical Image Registration and Fusion for Quality Enhancement. *CMC-COMPUTERS MATERIALS & CONTINUA*, vol. 68, no. 1, pp. 821-840.

14. Bacanin, N, Bezdan, T, Venkatachalam, K and Al-Turjman, F, (2021). Optimized convolutional neural network by firefly algorithm for magnetic resonance image classification of glioma brain tumor grade. *Journal of Real-Time Image Processing*, vol. 18, no. 4, pp. 1085-1098.
15. Baid, U, Talbar, S, Rane, S, Gupta, S, Thakur, M.H, Moiyadi, A, Thakur, S and Mahajan, A, (2019). Deep learning radiomics algorithm for gliomas (drag) model: A novel approach using 3d unet based deep convolutional neural network for predicting survival in gliomas. In *International MICCAI Brainlesion Workshop*, pp. 369– 379.
16. Bennai, M.T, Guessoum, Z, Mazouzi, S, Cormier, S and Mezghiche, M, (2020). A stochastic multi-agent approach for medical-image segmentation: application to tumor segmentation in brain MR images. *Artificial Intelligence in Medicine*, vol. 110, pp. 1-41.
17. Bhat, M and Karki, M.V, (2017). Feature selection based on PCA and PSO for multimodal medical image fusion using DTCWT, pp. 1-8.
18. Chen, J, Li, X, Luo, L, Mei, X and Ma, J, (2020). Infrared and visible image fusion based on target-enhanced multiscale transform decomposition. *Information Sciences*, vol. 508, pp. 64-78.
19. Chen, L.C, Papandreou, G, Kokkinos, I, Murphy, K and Yuille, A.L, (2014). Semantic image segmentation with deep convolutional nets and fully connected CRFs, pp. 1-14.

20. Chen, Y, Cao, Z, Cao, C, Yang, J and Zhang, J, (2018).A modified U-Net for brain Mr image segmentation. Cloud Computing and Security: 4th International Conference, ICCCS 2018, Haikou, China, June 8-10, 2018, Revised Selected Papers, Part VI, vol. 4. pp. 233-242.
21. Çiçek, Ö, Abdulkadir, A, Lienkamp, S.S, Brox, T and Ronneberger, O, (2016). 3D UNet: learning dense volumetric segmentation from sparse annotation. In International conference on medical image computing and computer-assisted intervention, pp. 424–432.
22. Daimary, D, Bora, M.B, Amitab, K and Kandar, D, (2020). Brain tumor segmentation from MRI images using hybrid convolutional neural networks. Procedia Computer Science, vol. 167, pp. 2419-2428.
23. Du, C and Gao, S, (2017). Image segmentation-based multi-focus image fusion through multi-scale convolutional neural network.IEEE access, vol. 5, pp. 15750-15761.
24. El Boustani, A, Aatila, M, El Bachari, E and El Oirrak, A, (2020). MRI brain images classification using convolutional neural networks. International Conference on Advanced Intelligent Systems for Sustainable Development, pp. 308-320.
25. Fan, F, Huang, Y, Wang, L, Xiong, X, Jiang, Z, Zhang, Z and Zhan, J, (2019). A semantic-based medical image fusion approach, pp. 1-6.

26. Fan, H., Liu, D., Xiong, Z. and Wu, F., 2017, September. Two-stage convolutional neural network for light field super-resolution. In 2017 IEEE International Conference on Image Processing (ICIP) (pp. 1167-1171). IEEE.
27. Fenshia Singh, J and Magudeeswaran, V, (2017). A machine learning approach for brain image enhancement and segmentation. International Journal of Imaging Systems and Technology, vol. 27, no. 4, pp. 311-316.
28. Forouzanfar, M, Forghani, N and Teshneshlab, M, (2010). Parameter optimization of improved fuzzy c-means clustering algorithm for brain MR image segmentation. Engineering Applications of Artificial Intelligence, vol. 23, no. 2, pp. 160-168.
29. Gai, Di, et al. "Multi-focus image fusion method based on two stage of convolutional neural network." *Signal Processing* 176 (2020): 107681.
30. Gholami, A, Mang, A and Biros, G, (2016). An inverse problem formulation for parameter estimation of a reaction–diffusion model of low-grade gliomas. *Journal of mathematical biology*, vol. 72, no. 1-2, pp. 409–433.
31. Girshick, R, Donahue, J, Darrell, T and Malik, J, (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580–587.
32. Goetz, M, Weber, C, Binczyk, F, Polanska, J, Tarnawski, R, Bobek-Billewicz, B, Koethe, U, Kleesiek, J, Stieltjes, B and Maier-Hein, K.H, (2015). DALSA: Domain adaptation for supervised learning from sparsely annotated MR images. *IEEE Trans. Med. Imag.*, vol. 35, no. 1, pp. 184–196.

33. Goodfellow, I, Pouget-Abadie, J, Mirza, M, Xu, B, Warde-Farley, D, Ozair, S, Courville, A and Bengio, Y, (2020). Generative adversarial networks. Communications of the ACM, vol. 63, no. 11, pp. 139-144.
34. Guo, K, Hu, X and Li, X, (2022). MMFGAN: A novel multimodal brain medical image fusion based on the improvement of generative adversarial network. Multimedia Tools and Applications, pp. 1-39.
35. Guo, L, Wu, P, Lou, S, Gao, J and Liu, Y, (2020). A multi-feature extraction technique based on principal component analysis for nonlinear dynamic process monitoring. Journal of Process Control, vol. 85, pp. 159-172.
36. Havaei, M, Davy, A, Warde-Farley, D, Biard, A, Courville, A, Bengio, Y, Pal, C, Jodoin, P.M and Larochelle, H, (2017). Brain tumor segmentation with deep neural networks. Medical image analysis, vol. 35, pp. 18–31.
37. Hermessi, H, Mourali, O and Zagrouba, E, (2021). Multimodal medical image fusion review: Theoretical background and recent advances. Signal Processing, vol. 183.
38. Hou, R, Zhou, D, Nie, R, Liu, D and Ruan, X, (2019). Brain CT and MRI medical image fusion using convolutional neural networks and a dual-channel spiking cortical model. Medical & biological engineering & computing, vol. 57, no. 4, pp. 887-900.
39. Hrosik, R.C, Tuba, E, Dolicanin, E, Jovanovic, R and Tuba, M, (2019). Brain image segmentation based on firefly algorithm combined with k-means clustering. Stud. Inform. Control, vol. 28, no. 2, pp. 167-176.

40. Hsu, S.L, Gau, P.W, Wu, I.L and Jeng, J.H, (2009). Region-based image fusion with artificial neural network. International Journal of Computer and Information Engineering, vol. 3, no. 5, pp. 1262-1265.
41. Hu, K, Gan, Q, Zhang, Y, Deng, S, Xiao, F, Huang, W, Cao, C and Gao, X, (2019). Brain tumor segmentation using multi-cascaded convolutional neural networks and conditional random field. IEEE Access. Vol. 7, pp. 92615-92629.
42. Huang, P.W, Chen, C.I, Chen, P, Lin, P.L and Hsu, L.P, (2014). PET and MRI brain image fusion using wavelet transform with structural information adjustment and spectral information patching. 2014 IEEE International Symposium on Bioelectronics and Bioinformatics (IEEE ISBB 2014), pp. 1-4.
43. Ibrahim, H, Kong, N.S.P and Ng, T.F, (2008). Simple adaptive median filter for the removal of impulse noise from highly corrupted images. IEEE Transactions on Consumer Electronics, vol. 54, no. 4, pp. 1920-1927.
44. Isensee, F, Kickingereder, P, Wick, W, Bendszus, M and Maier-Hein, K.H, (2018). Brain tumor segmentation and radiomics survival prediction: contribution to the brats 2017 challenge. In International MICCAI Brainlesion Workshop, pp. 287–297.
45. Jain, K, Aji, A and Krishnan, P, (2021). Medical Image Encryption Scheme Using Multiple Chaotic Maps. Pattern Recognition Letters, vol. 152, pp. 356-364.
46. Javadpour, A and Mohammadi, A, (2016). Improving brain magnetic resonance image (MRI) segmentation via a novel algorithm based on genetic and regional growth. Journal of biomedical physics & engineering, vol. 6, no. 2, pp. 95-108.

47. Kamnitsas, K, Ledig, C, Newcombe, V.F, Simpson, J.P, Kane, A.D, Menon, D.K, Rueckert, D and Glocker, B, (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical image analysis*, vol. 36, pp. 61–78.
48. Kang, J, Lu, W and Zhang, W, (2020). Fusion of brain PET and MRI images using tissue-aware conditional generative adversarial network with joint loss. *IEEE Access*, vol. 8, pp. 6368-6378.
49. Kaur, T, Saini, B.S and Gupta, S, (2018). A novel feature selection method for brain tumor MR image classification based on the Fisher criterion and parameter-free Bat optimization. *Neural Computing and Applications*, vol. 29, pp. 193-206.
50. Khan, H.N, Shahid, A.R, Raza, B, Dar, A.H and Alquhayz, H, (2019). Multi-view feature fusion based four views model for mammogram classification using convolutional neural network. *IEEE Access*, vol. 7, pp. 165724-165733.
51. Koley, S, Galande, A, Kelkar, B, Sadhu, A.K, Sarkar, D and Chakraborty, C, (2016). Multispectral MRI image fusion for enhanced visualization of meningioma brain tumors and edema using contourlet transform and fuzzy statistics. *Journal of Medical and Biological Engineering*, vol. 36, no. 4, pp. 470-484.
52. Krizhevsky, A, Sutskever, I and Hinton, G.E, (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pp. 1097–1105.

53. Kumar, D.M, Satyanarayana, D and Prasad, M.G, (2021). An improved Gabor wavelet transform and rough K-means clustering algorithm for MRI brain tumor image segmentation. *Multimedia Tools and Applications*, vol. 80, pp. 6939-6957.
54. Li, H, Nan, Y, Del Ser, J and Yang, G, (2023). Region-based evidential deep learning to quantify uncertainty and improve robustness of brain tumor segmentation. *Neural Computing and Applications*, pp. 1-15.
55. Li, S, Kang, X and Hu, J, (2013). Image fusion with guided filtering. *IEEE Transactions on Image processing*, vol. 22, no. 7, pp. 2864-2875.
56. Li, S, Yin, H and Fang, L, (2012). Group-sparse representation with dictionary learning for medical image denoising and fusion. *IEEE Transactions on biomedical engineering*, vol. 59, no. 12, pp. 3450-3459.
57. Li, Y, Zhao, J, Lv, Z and Li, J, (2021). Medical image fusion method by deep learning. *International Journal of Cognitive Computing in Engineering*, vol. 2, pp. 21-29.
58. Lim, K.Y and Mandava, R, (2018). A multi-phase semi-automatic approach for multisequence brain tumor image segmentation. *Expert systems with applications*, vol. 112, pp. 288-300.
59. Lin, G.C, Wang, W.J, Kang, C.C and Wang, C.M, (2012). Multispectral MR images segmentation based on fuzzy knowledge and modified seeded region growing. *Magnetic resonance imaging*, vol. 30, no. 2, pp. 230-246.

60. Liu, C, Tian, Y, Zhang, Q, Yuan, J and Xue, B, (2013). Adaptive firefly optimization algorithm based on stochastic inertia weight. 2013 Sixth International Symposium on Computational Intelligence and Design, vol. 1, pp. 334-337.
61. Liu, J and Guo, L, (2015). A new brain MRI image segmentation strategy based on K-means clustering and SVM. 2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics, vol. 2, pp. 270-273.
62. Liu, J, Mao, Y, Liu, X and Li, Y, (2020). A dynamic adaptive firefly algorithm with globally orientation. Mathematics and Computers in Simulation, vol. 174, pp. 76-101.
63. Liu, R, Liu, J, Jiang, Z, Fan, X and Luo, Z, (2020). A bilevel integrated model with data-driven layer ensemble for multi-modality image fusion. IEEE Transactions on Image Processing, vol. 30, pp. 1261-1274.
64. Liu, X, Mei, W and Du, H, (2016). Multimodality medical image fusion algorithm based on gradient minimization smoothing filter and pulse coupled neural network. Biomedical Signal Processing and Control, vol. 30, pp. 140-148.
65. Liu, Y, Chen, X, Peng, H and Wang, Z, (2017). Multi-focus image fusion with a deep convolutional neural network. Information Fusion, vol. 36, pp. 191-207.
66. Liu, Y.H., (2018). Feature extraction and image recognition with convolutional neural networks. Journal of Physics: Conference Series, vol. 1087, pp. 1-7.

67. Liu, Z, Yin, H, Chai, Y and Yang, S.X, (2014). A novel approach for multimodal medical image fusion. Expert systems with applications, vol. 41, no. 16, pp. 7425-7435.
68. Ma, J, Xu, H, Jiang, J, Mei, X and Zhang, X.P, (2020).DDcGAN: A dual-discriminator conditional generative adversarial network for multi-resolution image fusion. IEEE Transactions on Image Processing, vol. 29, pp. 4980-4995.
69. Mahasin, M.M, Naba, A, Widodo, C.S and Yueniwati, Y, (2023). Development of a Modified UNet-Based Image Segmentation Architecture for Brain Tumor MRI Segmentation. International Conference of Medical and Life Science (ICoMELISA 2021), pp. 37-43.
70. Mamelak, A.N and Jacoby, D.B, (2007), Targeted delivery of antitumoral therapy to glioma and other malignancies with synthetic chlorotoxin (tm-601). Expert opinion on drug delivery, vol. 4, no. 2, pp. 175–186.
71. Maqsood, S and Javed, U, (2020). Multi-modal medical image fusion based on two-scale image decomposition and sparse representation. Biomedical Signal Processing and Control, vol. 57.
72. McKinley, R, Meier, R and Wiest, R, (2019). Ensembles of densely-connected cnns with label-uncertainty for brain tumor segmentation. In International MICCAI Brain lesion Workshop, pp. 456–465.
73. Meher, S.K and Singhawat, B, (2014). An improved recursive and adaptive median filter for high density impulse noise. AEU-International Journal of Electronics and Communications, vol. 68, no. 12, pp. 1173-1179.

74. Meier, R, Bauer, S, Slotboom, J, Wiest, R and Reyes, M, (2014). Appearance-and context-sensitive features for brain tumor segmentation. In Proc. MICCAI BRATS Challenge, pp. 20–26.
75. Menze, B.H, Jakab, A, Bauer, S, Kalpathy-Cramer, J, Farahani, K, Kirby, J, Burren, Y, Porz, N, Slotboom, J, Wiest, R and Lanczi, L, (2014). The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans. Med. Imag, vol. 34, no. 10, pp. 1993–2024.
76. Menze, B.H, Van Leemput, K, Lashkari, D, Weber, M.A, Ayache, N and Golland, P, (2010). A generative model for brain tumor segmentation in multi- modal images. In Proc. Int. Conf. Med. Image Comput. Comput. -Assist. Intervent. Beijing, China: Springer, pp. 151–159.
77. Mohamed, N.E.K and El-Bhrawy, A.S.M, (2016). Artificial Neural Networks in Data Mining. IOSR Journal of Computer Engineering, vol. 18, pp. 55-59.
78. Myronenko, A, (2019). 3d MRI brain tumor segmentation using autoencoder regularization. In International MICCAI Brainlesion Workshop, pp. 311–320.
79. Parwez, M.A and Abulaish, M, (2019). Multi-label classification of microblogging texts using convolution neural network. IEEE Access, vol. 7, pp. 68678-6869.
80. Pereira, S, Pinto, A, Alves, V and Silva, C.A, (2016). Brain tumor segmentation using convolutional neural networks in MRI images. IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1240-1251.

81. Pritika and Budhiraja, S, (2016). Multimodal medical image fusion based on guided filtered multi-scale decomposition. International Journal of Biomedical Engineering and Technology, vol. 20, no. 4, pp. 285-301.
82. Ray, S, Kumar, V, Ahuja, C and Khandelwal, N, (2018). An automatic method for complete brain matter segmentation from multislice CT scan, pp. 1-17.
83. Reza, S and Iftekharuddin, K.M, (2014). Improved brain tumor tissue segmentation using texture features.In Proc. MICCAI BraTS (Brain Tumor Segmentation Challenge), pp.27–30.
84. Ronneberger, O, Fischer, P and Brox, T, (2015). UNet: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pp. 234–241.
85. Saut, O, Lagaert, J.B, Colin, T and Fathallah-Shaykh, H.M, (2014). A multilayer grow-or-go model for gbm: effects of invasive cells and anti-angiogenesis on growth. Bulletin of mathematical biology, vol. 76, no. 9, pp. 2306–2333.
86. Scarpa, G, Gargiulo, M, Mazza, A and Gaetano, R, (2018). A CNN-based fusion method for feature extraction from sentinel data. Remote Sensing, vol. 10, no. 2, pp. 1-20.
87. Schnabel, J.A, Heinrich, M.P, Papież, B.W and Brady, J.M, (2016). Advances and challenges in deformable image registration: from image fusion to complex motion modelling. Medical image analysis, vol. 33,pp. 145-148

88. Shao, Z and Cai, J, (2018). Remote sensing image fusion with deep convolutional neural network. *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 11, no. 5, pp. 1656-1669.
89. Sharif, M, Amin, J, Raza, M, Yasmin, M and Satapathy, S.C, (2020). An integrated design of particle swarm optimization (PSO) with fusion of features for detection of brain tumor. *Pattern Recognition Letters*, vol. 129, pp. 150-157.
90. Shehanaz, S., Daniel, E., Guntur, S.R. and Satrasupalli, S., 2021. Optimum weighted multimodal medical image fusion using particle swarm optimization. *Optik*, 231, p.166413.
91. Shen, D, Wu, G and Suk, H.I, (2017). Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng*, vol. 19, pp. 221–248
92. Shen, H, Wang, R, Zhang, J and McKenna, S.J, (2017). Boundary-aware fully convolutional network for brain tumor segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 433–441.
93. Singh, R and Khare, A, (2014). Fusion of multimodal medical images using Daubechies complex wavelet transform—A multiresolution approach. *Information fusion*, vol. 19, pp. 49-60.
94. Song, J, Zheng, J, Li, P, Lu, X, Zhu, G and Shen, P, (2021). An effective multimodal image fusion method using MRI and PET for Alzheimer's disease diagnosis. *Frontiers in digital health*, vol. 3, pp. 1-12.

95. Subbiah Parvathy, V, Pothiraj, S and Sampson, J, (2020). A novel approach in multimodality medical image fusion using optimal shearlet and deep learning. International Journal of Imaging Systems and Technology, vol. 30, no. 4, pp. 847-859.
96. Sudha, M.N, Selvarajan, S and Suganthi, M, (2019). Feature selection using improved lion optimization algorithm for breast cancer classification. International Journal of Bio-Inspired Computation, vol. 14, no. 4, pp. 237-246.
97. Suhas, S and Venugopal, C.R, (2017). MRI image preprocessing and noise removal technique using linear and nonlinear filters. 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), pp. 1-4.
98. Sulaiman, S.N and Isa, N.A.M, (2010). Adaptive fuzzy-K-means clustering algorithm for image segmentation. IEEE Transactions on Consumer Electronics, vol. 56, no. 4, pp. 2661-2668.
99. Tan, W, Thitøn, W, Xiang, P and Zhou, H, (2021). Multi-modal brain image fusion based on multi-level edge-preserving filtering. Biomedical Signal Processing and Control, vol. 64.
100. Tan, W, Tiwari, P, Pandey, H.M, Moreira, C and Jaiswal, A.K, (2020). Multimodal medical image fusion algorithm in the era of big data. Neural Computing and Applications, pp. 1-21

101. Ting, J, Punithakumar, K and Ray, N, (2020). Multiview 3-d echocardiography image fusion with mutual information neural estimation. 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 765-771.
102. Trivedi, G.J and Sanghvi, R, (2022). Medical Image Fusion Using CNN with Automated Pooling. Indian Journal of Science and Technology, vol. 15, no. 42, pp. 2267-2274.
103. Tustison, N.J, Shrinidhi, K.L, Wintermark, M, Durst, C.R, Kandel, B.M, Gee, J.C, Grossman, M.C and Avants, B.B, (2015). Optimal symmetric multi-modal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTsR. Neuroinformatics, vol. 13, no. 2, pp. 209–225.
104. Valindria, V.V, Pawlowski, N, Rajchl, M, Lavdas, I, Aboagye, E.O, Rockall, A.G, Rueckert, D and Glocker, B, (2018). Multi-modal learning from unpaired images: Application to multi-organ segmentation in CT and MRI. 2018 IEEE winter conference on applications of computer vision (WACV), pp. 547-556.
105. Verma, K, Singh, B.K and Thoke, A.S, (2015). An enhancement in adaptive median filter for edge preservation. Procedia Computer Science, vol. 48, pp. 29-36.
106. Wang, G, Li, W, Ourselin, S and Vercauteren, T, (2018). Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In International MICCAI Brainlesion Workshop, pp. 178–190.

107. Wang, G, Li, W, Ourselin, S and Vercauteren, T, (2019). Automatic brain tumor segmentation using convolutional neural networks with testtime augmentation. In International MICCAI Brainlesion Workshop. Springer, pp. 61–72.
108. Wang, J, Yu, L, Tian, S, Wu, W and Zhang, D, (2022).AMFNet: An attention-guided generative adversarial network for multi-model image fusion. Biomedical Signal Processing and Control, vol. 78.
109. Wang, K, Zheng, M, Wei, H, Qi, G and Li, Y, (2020). Multi-modality medical image fusion using convolutional neural network and contrast pyramid. Sensors, vol. 20, no. 8, pp. 1-17.
110. Wang, M, Wu, C, Wang, L, Xiang, D and Huang, X, (2019). A feature selection approach for hyperspectral image based on modified ant lion optimizer. Knowledge-Based Systems, vol. 168, pp. 39-48.
111. Wang, Y, Yu, B, Wang, L, Zu, C, Lalush, D.S, Lin, W, Wu, X, Zhou, J, Shen, D and Zhou, L, (2018). 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. Neuroimage, vol. 174, pp. 550-562.
112. Wen, Y, Yang, X, Celik, T, Sushkova, O and Albertini, M.K, (2020). Multifocus image fusion using convolutional neural network. Multimedia Tools and Applications, vol. 79, pp. 34531-34543.
113. Yadav, S.P and Yadav, S, (2020). "Image fusion using hybrid methods in multimodality medical images. Medical & Biological Engineering & Computing, vol. 58 pp. 669-687.

114. Yan, C, Gong, B, Wei, Y and Gao, Y, (2020). Deep multi-view enhancement hashing for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 4, pp. 1445-1451.
115. Yousef, R, Khan, S, Gupta, G, Siddiqui, T, Albahlal, B.M, Alajlan, S.A and Haq, M.A, (2023). U-Net-based models towards optimal MR brain image segmentation. *Diagnostics*, vol. 13, no. 9, pp. 1-27.
116. Zhang, H, Le, Z, Shao, Z, Xu, H and Ma, J, (2021). MFF-GAN: An unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion. *Information Fusion*, vol. 66, pp. 40-53.
117. Zhang, Y, Xiang, W, Zhang, S, Shen, J, Wei, R, Bai, X, Zhang, L and Zhang, Q, (2022). Local extreme map guided multi-modal brain image fusion. *Frontiers in Neuroscience*, vol. 16, pp. 1055451.
118. Zhang, Y, Zhang, L, Bai, X and Zhang, L, (2017). Infrared and visual image fusion through infrared feature extraction and visual information preservation. *Infrared Physics & Technology*, vol. 83, pp. 227-237.
119. Zhang, Y., Liu, Y., Sun, P., Yan, H., Zhao, X. and Zhang, L., 2020. IFCNN: A general image fusion framework based on convolutional neural network. *Information Fusion*, 54, pp.99-118.
120. Zhao, C, Yang, P, Zhou, F, Yue, G, Wang, S, Wu, H, Chen, G, Wang, T and Lei, B, (2023). MHW-GAN: Multidiscriminator Hierarchical Wavelet Generative Adversarial Network for Multimodal Image Fusion. *IEEE Transactions on Neural Networks and Learning Systems*.

121. Zhao, W, Wang, D and Lu, H, (2018). Multi-focus image fusion with a natural enhancement via a joint multi-level deeply supervised convolutional neural network. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 4, pp. 1102-1115.
122. Zheng, Y, Blasch, E and Liu, Z, (2018). Multispectral image fusion and colorization, vol. 481.
123. Zhu, R, Li, X, Zhang, X and Ma, M, (2020). Mri and ct medical image fusion based on synchronized-anisotropic diffusion model. *IEEE Access*, vol. 8, pp. 91336-91350.
124. Zhu, Z, Yin, H, Chai, Y, Li, Y and Qi, G, (2018). A novel multi-modality image fusion method based on image decomposition and sparse representation. *Information Sciences*, vol. 432, pp. 516-529.

Publications

An Extensive Study on the Multimodal Medical Image Fusion Techniques

Kalaivani Ramasamy, Dr. Anna Saro Vijendran

Research scholar in Computer Science

Dean School of computing

Sri Ramakrishna College of Arts and Science

Coimbatore-641 006

Tamil Nadu, India

ABSTRACT

Medical image fusion is the method of registering and combining multiple images from single or multiple imaging modalities. Multimodal medical image fusion has become a powerful tool for medical diagnoses. In order to obtain a more absolute and more consistent medical image, this paper presents a novel approach for multimodal medical image fusion using an improved pulse-coupled neural network (PCNN) in non sub sampled contour let transform (NSCT) domain. First, the image is split into sub-bands with different scales and different directions by NSP and NSDFB. Next, local area singular value is introduced to determine the structural information factor which will be the linking strength parameter of PCNN. Finally, inverse NSCT is used for fused images. Using the ‘max selection’ rule low frequency sub bands (LFSs) are fused. A PCNN model is utilized for the fusion of high-frequency sub bands (HFSs). Modified Spatial Frequency (MSF) in DRT domain is input to motivate the PCNN and coefficients in DRT domain with large firing times are selected as coefficients of the fused image. Then inverse DRT (IDRT) is applied to the fused coefficients to get the fused image. On study of Multi-modal medical image fusion algorithms and devices, it improves clinical accuracy of decisions based on medical images. We characterize the medical image fusion research based on (1) the widely used image fusion methods, (2) imaging modalities and (3) imaging of organs that are under study. Our proposed algorithm in multimodal medical image fusion is proved to perform better in robustness and reliability over the existing methods, meeting the requirement of human vision.

Keywords:- Medical Image Fusion, Multimodal Image fusion, Principle Component analysis, MRI Images, PET Images, CT Images, PCNN, NSCT

I. INTRODUCTION

Medical image fusion is an important task to retrieve an image which provides as much as information of the same organ at the same time it also helps to reduce the storage capacity to a single image. For instance, X-ray and Computed Tomography (CT) are able to display the bones and other hard tissues structures according to the differences in density and thickness whereas MRI can display soft tissues like blood vessels. And Positron Emission Computed Tomography (PET) is able to show physiological and pathological contents of human organs [1]. In practical clinical applications, a single modality of image is often unable to provide sufficient information. Multimodal medical image fusion technique can combine several source images to provide more comprehensive and more reliable information aiming to assist clinicians to diagnose and treat diseases accurately [1]. It is also used to obtain a more complete and accurate description of the same object, which provides an easy access for

image-guided medical diagnostic and treatment [2]. To fuse the image, primary step as image denoising methods has to be applied in order to smooth and obtain the details of the features in the image in detail.

In data integration approaches, data from different modalities are usually analyzed through separate pipelines and the results combined at the interpretation level to yield decision level fusion [3]. Therefore, the quality of image fusion highly depends on the performance of image feature extraction. The multivariate model using structure and sparse component on image fusion techniques to the multimodal medical image has been defined as proposed model. Through imposed sparsity, parsimonious multivariate methods increase the interpretability of the output and potentially improve the generalizability of the produced model.



M-UNet for Segmentation of Brain Images

Dr. Anna Saro Vijendran¹ and Kalaivani Ramasamy^{2*}

¹Dean-Research, Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore – 641044, Tamil Nadu, India

²*Research Scholar, Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore – 641044, Tamil Nadu, India

Abstract: A mass or progress of unusual cells in the brain is termed as brain tumor. Several categories of brain tumors occur in human. Certain types of brain tumors are non-cancerous which is indicated as benign, whereas certain brain tumors are cancerous, called malignant. In this paper, images are segmented using Modified- Universal Education and Training, Ltd. (M-UNet). The main aim is to investigate network architectures (M-UNet) based on deep learning which is used for enhanced classification and segmentation of brain tumor images. Segmentation of brain cancer images is the procedure of splitting the tumor from usual brain muscles; in medical routine, it offers valuable information for analysis and treatment planning. It is still a complex job due to the asymmetrical arrangement and perplexing borders of tumors. The Convolutional Neural Network (CNN) and Universal Education and Training, Ltd.(UNet) are considered to be notable techniques in segmentation of images. The concept of CNN is a dominant technique for recognition of images and forecasts. CNN is typically utilized for brain cancer separation, classification, and estimate of existence period for infected people. UNet is a familiar image separation method established mainly for analyzing clinical images that can exactly divide images using an unusual quantity of preparation facts. These qualities make UNet efficient in clinical imaging forum and support wide-ranging implementation of UNet in performing separation jobs in therapeutic imaging. M-UNet is recommended in this paper to slice the given input images in a well-defined manner. Experimental results have shown that the proposed M-UNet achieves accuracy of 97% which is notably better when compared to the existing CNN and UNet techniques. The results are also compared based on Dice Coefficient, Jaccard Coefficient and time period. It is evident that the M-UNet outperforms the existing techniques on all assessment parameters. A novel frame work using M-UNet that includes extraction of both global and local features is proposed to increase the segmentation accuracy. The outcomes show better performance in segmenting the 5 tumor areas on the huge BRATS 2018 dataset. The performance of the network is assessed by comparing the forecast segmentation of tumor areas to the ground truth offered by the dataset. Dice Similarity Coefficient (DSC) and Jaccard Coefficient (JC) give the like nessamid the anticipated tumor area and ground truth by associating the overlay areas. In this paper, brain image segmentation is performed using UNet and M-UNet methods and the proposed method efficiently predicts the border of then segmentation pixel.

Keywords - Brain Tumor, Convolutional Neural Network, Image Recognition, UNet, M-UNet

***Corresponding Author**

Kalaivani Ramasamy , Research Scholar,
Department of Computer Science, Sri Ramakrishna
College of Arts & Science, Coimbatore – 641044,
Tamil Nadu, India

Received On 18 October, 2021

Revised On 22 February, 2022

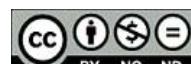
Accepted On 1 April, 2022

Published On 1 May, 2022

Funding This research did not receive any specific grant from any funding agencies in the public, commercial or not for profit sectors.

Citation Dr.Anna Saro Vijendran and Kalaivani Ramasamy , M-UNet for Segmentation of Brain Images.(2022).Int. J. Life Sci. Pharma Res.12(3), P36-43 http://dx.doi.org/10.22376/ijpbs/lpr.2022.12.3.P36-43

This article is under the CC BY- NC-ND Licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)



Copyright @ International Journal of Life Science and Pharma Research, available at www.ijlpr.com



Optimal segmentation and fusion of multi-modal brain images using clustering based deep learning algorithm

Anna Saro Vijendran ^a, Kalaivani Ramasamy ^{b,*}

^a Department of Computer Application, Sri Ramakrishna College of Arts & Science, Coimbatore – 641044, Tamil Nadu, India

^b Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore – 641044, Tamil Nadu, India

ARTICLE INFO

Keywords:

Multimodal images
Adaptive FireFly Optimization based
Convolutional Neural Network (AFFOCNN)
Modified Fully Connected Layer (MFCL)
Computed Tomography (CT)
clustering
Segmentation
Magnetic Resonance Imaging (MRI)

ABSTRACT

In the last several years, the world of medical technology has seen a boom in multimodal picture fusion. Information is constrained since a single medical instrument can only acquire single modal pictures. Doctors often need a large number of multimodal pictures to get the complete information necessary for disease diagnosis. The burden associated with illness diagnosis will significantly rise when multimodal pictures are employed directly, and errors and interference are likely to occur. Fusion algorithms, which have been extensively employed in the medical industry, may very effectively combine a lot of information in multimodal pictures. However, the existing method has an issue with the earlier stages of brain tumor prediction in white images and inaccuracy image results. To overcome the above-mentioned problems, in this work, Adaptive Firefly Optimization based Convolutional Neural Network (AFFOCNN) and Modified Fully Connected Layer (MFCL) scheme is proposed. This work contains main steps such as noise removal, segmentation, feature extraction, image fusion, and image classification process. Initially, noise removal is done for improving the image quality by removing the noise. Then the modality MRI images are segmented and it is used for subdividing an image into its constituent regions or object. It segments the image into black and white images. After that, feature extraction is applied through the AFFOCNN algorithm which extracts the more informative image features. Image fusion of multi-modal images derived the lower-level, middle-level, and higher-level image contents. It can be viewed in multiple directions and fused in all directions. Finally, image classification is performed by using a Modified Fully Connected Layer (MFCL) which improves the training and testing features efficiently. It was determined from the results that the suggested combination of AFFOCNN and MFCL algorithm performs improved than the current algorithms with the increased accuracy, precision, recall, and mean square error (MSE), as well as execution time with the values of 99.00%, 98.00%, 96.00%, 12.00% and 2.40 seconds respectively.

1. Introduction

In order to get better picture quality while preserving the details of a single image, medical image fusion involves integrating numerous images taken using different modalities. The fusion process improves the visual information and clarity, which helps doctors identify and evaluate the ailment. Medical imaging is now an essential part of a huge series of applications, together with diagnosis, research, and therapy, thanks to the fast growth in high-technology and contemporary instruments [15]. Medical picture combination is the concept of merging pictures from several imaging technologies, such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), and solo PET, in order to enhance the image content (SPECT).

Computed Tomography (CT) offers the finest information on denser tissue with less distortion for medical diagnosis. Better information on soft tissue is provided by magnetic resonance imaging (MRI), which has higher distortion [1]. Fig. 1 (a) shows the different medical brain imaging modalities and Fig. 1 (b) shows the MRI brain image. In this work, MRI brain image is considered for image fusion evaluation.

Image segmentation, which contracts with the explanation and identification of relevant objects in 2D and 3D vision-data pictures, and figure registration [18], which deals with the arrangement of data from several multimodal or time bordered images, are used in the prevalent practice of image fusion. As a result, image fusion produces an output picture that is enriched and improved over the input images [3]. As a result, the main goal of picture fusion is to produce classifiable,

* Corresponding author.

E-mail addresses: saroviji@rediffmail.com (A.S. Vijendran), tokalai.2011@gmail.com (K. Ramasamy).

Ensemble Deep Learning Algorithm for Multi View Image Fusion

Anna Saro Vijendran¹, Kalaivani Ramasamy^{*2}

Submitted: 09/10/2023

Revised: 30/11/2023

Accepted: 09/12/2023

Abstract: In order to create a synthetic image with more valuable data than just single particular image could ever provide, image fusion attempts to combine many image graphs of the similar subject. The clarity of the majority of original images is restricted by imaging sensors' limitations and broadband signal transmission. An innovative multi-modality clinical image fusing technique is presented in this study to enhance the image quality and earlier brain tumor detection performance. Hence, ensemble based deep learning algorithm is proposed in this study to increase the Magnetic Resonance Imaging (MRI) brain image fusion performances by minimizing noises, executing segmentations, extracting features, fusing images in primary phases. Initial noise reduction improves image quality. Segmentation process MRI scans divides an image into its parts. Black-and-white images are created. Following that, Lion Swarm Optimization based Convolutional Neural Network (LSOCNN) extracts image characteristics with most information. Eventually, multi-modal image fusion generated lower, intermediate, and upper level images. It can be seen from all angles and fused. To increase image fusion performance, ensemble CNN, DFMI-Net, TA-cGAN algorithms are presented. From the result, proposed ensemble DCNN+DFMI+Cgan compared to previous methods, this approach performs improved in respect of correctness, clarity, memory, and mean square error (MSE).

Keywords: *Differentiable Fusion with Mutual Information-Network (DFMI-Net), Image fusion, Lion Swarm Optimization based Convolutional Neural Network (LSOCNN), MRI images, Tissue-Aware Conditional Generative Adversarial Network (TA-cGAN).*

1. Introduction

Medical imaging is crucial to clinical diagnostics, surgery guidance, and treatment planning because to the fast growth of sensor and computer science technologies [1]. Irregular cell reproduction causes aberrant brain cell proliferation, which results in solid masses known as brain tumors. They are separated into benign and malignant groups. They are potentially fatal because of how intrusive and persistent they are, which interferes with the brain's regular functions. Furthermore, the edema, or fluid buildup around the tumor, puts stress on normal tissue, causing them to malfunction. Due to radiologists' ability to link swelling to the volume and progression of a tumor, it is crucial in the diagnosing process. The remarkable soft tissue contrast of MRI and CT makes it their more useful modalities for assessing brain malignancies amongst the many brain imaging methods. This study uses an application of an MRI image storage.

Because increasing the clinically diagnostic reliability is the goal of medical image fusion, the resulting fused image must be formed by carefully retaining the most important characteristics and characteristics of the original images. Utilizing either the spatially region or the transformation

region to carry out the execution of multifunctional clinical image merging. Through the use of fusing criteria, matching spatial pixels from CT or MRI images merge in the spatial domain [2]. Unfortunately, because to weak edge and contour recognition and distortion, the grade of the combined images is mediocre. The primary issue using spatial domains is that. Because these describe the tumor size or shape, edge and contour representations is often important for clinical image interpretation.

X-ray, CT, and MRI are a few examples of imaging techniques that are often utilized to produce clinical images. These techniques are targeted on revealing data about certain tissues or organs. While CT images assist in precisely identifying solid structures like implants and bones, MRI scans characterize soft tissue in their high-resolution anatomy [3]. The primary goal of image fusing is to create a single complete image encompassing the distinctive qualities of multidimensional clinical imaging, that may assist clinicians in making correct diagnosis across a variety of conditions [4].

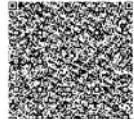
In image fusion process, feature extraction is a method of dimension reduction. It is well acknowledged to be a successful method for both lowering computing difficulty and elevating reliability. A powerful method for reducing the dimensions of the information is feature extraction, which is in addition to feature selection. Before projecting a hyperspectral image onto another feature space, a linear transition is used. Following that, only the most important components are maintained for classification [5]. Unsupervised approaches like PCA, ICA, as well as

¹Dean-Research, Department of Computer Application, Sri Ramakrishna College of Arts & Science, Coimbatore, Tamil Nadu, India.

Email: saroviji@rediffmail.com

²Research Scholar, Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore, Tamil Nadu, India.

* Corresponding Author Email: tokalai.2011@gmail.com



Comparative Study for Image Fusion Using Various Deep Learning Algorithms

Anna Saro Vijendran* & Kalaivani Ramasamy*

Abstract

Image fusions are that join medical images from many modalities like CTS (Computed Tomography Scans) and MRI (Magnetic Resonance Imaging) with the aim of presenting better clinical content to clinicians and doctors for planning treatments or therapies. Prior studies based on white images have showed issues in early predictions of brain tumours like inaccurate image. This study attempts to overcome this issue by comparing MUNets (Modified-UNets), MCNNs (Multi-Cascaded Convolution Neural Networks) with fully connected CRFs (Conditional Random Fields), MFCLs (Modified Fully Connected Layers), TA-cGANs (Tissue-Aware conditional Generative Adversarial Networks) and EL (Ensemble Learning) algorithms. Furthermore, in order to enhance image quality and enable the early diagnosis of brain tumours, novel multimodal medical image fusion techniques are being examined. Furthermore, it is suggested that the EL algorithm improve MRI brain image fusion performance. The four primary phases of the suggested system are segmentation, image synthesis, feature extraction, and noise reduction. To eliminate noises, AMFs (Adaptive Median Filters) are used for reducing noises in MRI images and thus assist in enhancing classification accuracies. These features are taken into segmentation process using RGKMC (Region Growing based K-Means Clustering). Feature extractions are performed using AFFOCNNs (Adaptive FireFly Optimization based Convolution Neural Networks) algorithm which computes necessary and prominent

* Department of Computer Application, Sri Ramakrishna College of Arts & Science, Coimbatore - 641044, Tamil Nadu, India; E-mail: saroviji@rediffmail.com, tokalai2011@gmail.com