


Final thesis combined_21-10-24_except front pages.docx

 Delhi Technological University

Document Details

Submission ID

trn:oid:::27535:73667253

Submission Date

Dec 10, 2024, 9:29 PM GMT+5:30

Download Date

Dec 10, 2024, 9:36 PM GMT+5:30

File Name

Final thesis combined_21-10-24_except front pages.docx

File Size

4.9 MB

172 Pages

41,217 Words

236,387 Characters

8% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.





Filtered from the Report

- Bibliography
- Small Matches (less than 10 words)




Exclusions

- 3 Excluded Sources

Match Groups



-  **239** Not Cited or Quoted 8%
Matches with neither in-text citation nor quotation marks
-  **0** Missing Quotations 0%
Matches that are still very similar to source material
-  **0** Missing Citation 0%
Matches that have quotation marks, but no in-text citation
-  **0** Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

- 4%  Internet sources
- 5%  Publications
- 3%  Submitted works (Student Papers)

Integrity Flags

2 Integrity Flags for Review

-  **Replaced Characters**
79 suspect characters on 17 pages
Letters are swapped with similar characters from another alphabet.
-  **Hidden Text**
50 suspect characters on 1 page
Text is altered to blend into the white background of the document.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 239** Not Cited or Quoted 8%
Matches with neither in-text citation nor quotation marks
- 0** Missing Quotations 0%
Matches that are still very similar to source material
- 0** Missing Citation 0%
Matches that have quotation marks, but no in-text citation
- 0** Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

- 4% Internet sources
- 5% Publications
- 3% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	dokumen.pub	1%
2	Internet	link.springer.com	1%
3	Internet	prism.ucalgary.ca	0%
4	Internet	www.nature.com	0%
5	Internet	listens.online	0%
6	Internet	www.researchgate.net	0%
7	Publication	"Computational and Experimental Methods in Mechanical Engineering", Springer ...	0%
8	Internet	pmc.ncbi.nlm.nih.gov	0%
9	Publication	G. Neelima, Dhanunjaya Rao Chigurukota, Balajee Maram, B. Girirajan. "Optimal ...	0%
10	Publication	Priyanka Datta, Rajesh Rohilla. "An Introduction to Deep Learning Applications In...	0%

11	Publication	Maresh P Potadar, R.S Holambe. "Automatic Multi-Class Brain Tumor Classificatio...	0%
12	Publication	Debadyuti Mukherjee, Pritam Saha, Dmitry Kaplun, Aleksandr Sinitca, Ram Sark...	0%
13	Internet	deepai.org	0%
14	Internet	iris.polito.it	0%
15	Publication	Jiahuan Song, Xinjian Chen, Qianlong Zhu, Fei Shi, Dehui Xiang, Zhongyue Chen, Y...	0%
16	Publication	Shirin Kordnoori, Maliheh Sabeti, Mohammad Hossein Shakoor, Ehsan Moradi. "D...	0%
17	Internet	www.mdpi.com	0%
18	Submitted works	Higher Education Commission Pakistan on 2016-10-29	0%
19	Internet	doczz.net	0%
20	Submitted works	Anna University on 2024-04-16	0%
21	Submitted works	Liverpool John Moores University on 2024-04-21	0%
22	Publication	"ICICCT 2019 – System Reliability, Quality Control, Safety, Maintenance and Mana...	0%
23	Publication	"Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries", Spr...	0%
24	Publication	Sitendra Tamrakar, Shruti Bhargava Choubey, Abhishek Choubey. "Computational...	0%

25	Internet	pnrsolution.org	0%
26	Internet	www.hindawi.com	0%
27	Publication	Konttinen, Olivia R.. "The Kinetic Mechanism of DNA Strand Separation by High-fi...	0%
28	Publication	Sedra, Adel. "Microelectronic Circuits 7th Edition, International Edition", Oxford U...	0%
29	Internet	espace.curtin.edu.au	0%
30	Internet	ijisrt.com	0%
31	Submitted works	Liverpool John Moores University on 2024-03-15	0%
32	Submitted works	University of Leicester on 2023-09-11	0%
33	Internet	export.arxiv.org	0%
34	Internet	prr.hec.gov.pk	0%
35	Submitted works	rdia on 2024-02-18	0%
36	Publication	M. R. Maneesha Mohan, C. Helen Sulochana, T. Latha. "Medical image denoising u...	0%
37	Publication	Li, Maxwell P.. "Magnetic Domain Wall Skyrmions: Investigating Chiral Spin Textu...	0%
38	Submitted works	Anna University on 2024-04-18	0%

39	Publication	Vinod Kumar Dhakshnamurthy, Murali Govindan, Kannan Sreerangan, Manikand...	0%
40	Internet	openaccess.city.ac.uk	0%
41	Publication	Chaosheng Tang, Bin Li, Junding Sun, Shui-Hua Wang, Yu-Dong Zhang. "GAM-SpC...	0%
42	Publication	Gopal S. Tandel, Antonella Balestrieri, Tanay Jujaray, Narender N. Khanna, Luca S...	0%
43	Submitted works	Heriot-Watt University on 2024-11-28	0%
44	Publication	Mahmoud Khaled Abd-Allah, Ali Ismail Awad, Ashraf A.M. Khalaf, Hesham F.A. Ha...	0%
45	Publication	Muhammad Attique Khan, Imran Ashraf, Majed Alhaisoni, Robertas Damaševičiu...	0%
46	Publication	Vasanth SV, Srujana S, Swetha G, Manasa V. "Animated and Oil Painting Image G...	0%
47	Submitted works	CSU Northridge on 2023-11-17	0%
48	Publication	Cheng Peng, Andriy Myronenko, Ali Hatamizadeh, Vishwesh Nath et al. "HyperSe...	0%
49	Publication	Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio...	0%
50	Submitted works	Govind Ballabh Pant Engineering College, Pauri-Garhwal on 2021-01-28	0%
51	Publication	Kashfia Sailunaz, Sleiman Alhajj, Tansel Özyer, Jon Rokne, Reda Alhajj. "A survey o...	0%
52	Publication	Madhu M Nayak, Sumithra Devi K A. "Effective MRI based Brain Tumor Detection ...	0%

53	Submitted works	University of Essex on 2023-11-24	0%
54	Internet	ijircce.com	0%
55	Publication	Gaurav Garg. "Computer-aided Diagnosis Systems for Prostate Cancer: A Compre...	0%
56	Publication	Md Assaduzzaman, Monoronjon Dutta, Arpa Saha, Showmick Guha Paul. "ALSA-3:...	0%
57	Publication	Nikita Joshi, Sarika Jain, Amit Agarwal. "An improved approach for denoising MRI ...	0%
58	Publication	Shrestha, Nirjal. "An Active Set Method for a Neural Network (ASM-NN)", Universi...	0%
59	Publication	Sumod Sundar, Sumathy Subramanian, Mufti Mahmud. "Classification of Diabetic...	0%
60	Submitted works	University of Lancaster on 2008-01-17	0%
61	Submitted works	University of Northumbria at Newcastle on 2020-06-01	0%
62	Submitted works	University of Portsmouth on 2017-04-03	0%
63	Internet	healthdocbox.com	0%
64	Submitted works	iGroup on 2014-05-27	0%
65	Publication	"Neural Information Processing", Springer Science and Business Media LLC, 2017	0%
66	Submitted works	Bournemouth University on 2024-08-16	0%

67	Publication	Dimitrios Ntelitheos. "8 Negation", Walter de Gruyter GmbH, 2024	0%
68	Publication	Dongsheng Jiang, Weiqiang Dou, Luc Vosters, Xiayu Xu, Yue Sun, Tao Tan. "Denois...	0%
69	Publication	Jiasong Chen, Linchen Qian, Pengce Wang, Christina Sun et al. "A 3D Image Segm...	0%
70	Submitted works	Liverpool John Moores University on 2024-03-03	0%
71	Publication	Muhammed Mutlu Yapici, Rukiye Karakis, Kali Gurkahraman. "Improving Brain T...	0%
72	Submitted works	Nitte University on 2024-02-01	0%
73	Publication	Sachi Nandan Mohanty, Preethi Nanjundan, Tejaswini Kar. "Artificial Intelligence ...	0%
74	Submitted works	The University of the West of Scotland on 2023-12-22	0%
75	Submitted works	University of East London on 2023-05-07	0%
76	Publication	Yucheng Tang, Dong Yang, Wenqi Li, Holger R. Roth, Bennett Landman, Daguang ...	0%
77	Internet	uis.brage.unit.no	0%
78	Publication	"Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries", Spr...	0%
79	Publication	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2019", ...	0%
80	Publication	"Recent Trends in Image Processing and Pattern Recognition", Springer Science a...	0%

81	Publication	"Soft Computing: Theories and Applications", Springer Science and Business Medi...	0%
82	Publication	Ahmed Elhadad, A Ghareeb, Safia Abbas. "A blind and high-capacity data hiding o...	0%
83	Publication	Ankur Saxena, Nicolas Brault, Shazia Rashid. "Big Data and Artificial Intelligence f...	0%
84	Submitted works	Arts, Sciences & Technology University In Lebanon on 2024-01-20	0%
85	Publication	Asmita Dixit, Aparajita Nanda. "An improved whale optimization algorithm-based...	0%
86	Submitted works	Bournemouth University on 2023-08-31	0%
87	Publication	Bowen Jiang, Tau Yue, Xuemei Hu. "An improved attentive residue multi-dilated n...	0%
88	Publication	Hussein, Hagar Mohamed Wahid. "Stress Detection Using New Time-Frequency D...	0%
89	Submitted works	King's College on 2021-05-03	0%
90	Submitted works	Liverpool John Moores University on 2023-03-15	0%
91	Submitted works	Manchester Metropolitan University on 2024-11-01	0%
92	Publication	Maria Nazir, Sadia Shakil, Khurram Khurshid. "End-to-End Multi-task Learning Arc...	0%
93	Submitted works	Middlesex University on 2024-05-21	0%
94	Publication	Muhammed Talo, Ulas Baran Baloglu, Özal Yıldırım, U Rajendra Acharya. "Applica...	0%

95	Publication	Naveed Ilyas, Yoonguu Song, Aamir Raja, Boreom Lee. "Hybrid-DANet: An Encode...	0%
96	Publication	R. Gayathiri, Suganthi Santhanam. "C-SAN: Convolutional stacked autoencoder n...	0%
97	Publication	Ravikumar Sajjanar, Umesh D. Dixit, Vittalkumar K Vagga. "Advancements in hyb...	0%
98	Publication	Shalli Rani. "Emerging Technologies and the Application of WSN and IoT - Smart S...	0%
99	Publication	Shubhangi Solanki, Uday Pratap Singh, Siddharth Singh Chouhan, Sanjeev Jain. "A...	0%
100	Publication	Subathra Gunasekaran, Prabin Selvestar Mercy Bai, Sandeep Kumar Mathivanan, ...	0%
101	Submitted works	The Robert Gordon University on 2023-12-14	0%
102	Publication	Viet Tien Pham, Minh Hieu Ha, Bao V. Q. Bui, Truong Son Hy. "LightMed: A Light-w...	0%
103	Internet	arxiv.org	0%
104	Internet	e-space.mmu.ac.uk	0%
105	Internet	ebin.pub	0%
106	Internet	openaccess.altinbas.edu.tr	0%
107	Internet	www.jcboseust.ac.in	0%
108	Publication	"Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery", Sp...	0%

109	Publication	"Applied Intelligence and Informatics", Springer Science and Business Media LLC,...	0%
110	Publication	"Applied Intelligence and Informatics", Springer Science and Business Media LLC,...	0%
111	Publication	"Information Processing in Medical Imaging", Springer Science and Business Med...	0%
112	Publication	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2023", ...	0%
113	Publication	"Second International Conference on Image Processing and Capsule Networks", S...	0%
114	Internet	123dok.net	0%
115	Publication	Amita Dev, Arun Sharma, S. S. Agarwal. "Artificial Intelligence and Speech Techno...	0%
116	Submitted works	Anna University on 2024-06-14	0%
117	Submitted works	Anna University on 2024-08-17	0%
118	Publication	Bhanu Chander, Koppala Guravaiah, B. Anoop, G. Kumaravelan. "Handbook of AI-...	0%
119	Submitted works	Colorado Technical University Online on 2024-12-09	0%
120	Submitted works	Cornell University on 2020-06-01	0%
121	Submitted works	Cornell University on 2020-10-25	0%
122	Submitted works	De Montfort University on 2023-05-04	0%

123	Publication	Er-Yang Huan, Gui-Hua Wen. "Multilevel and Multiscale Feature Aggregation in D...	0%
124	Publication	Fabio Baselice, Giampaolo Ferraioli, Vito Pascazio, Antonietta Sorriso. "Denoising ...	0%
125	Submitted works	Guru Jambheshwar University of Science & Technology on 2021-09-25	0%
126	Publication	Khallaghi, Sam. "Advancing the Use of Deep Learning Techniques for Earth Obser...	0%
127	Submitted works	Liverpool John Moores University on 2022-03-08	0%
128	Submitted works	Manchester Metropolitan University on 2022-05-18	0%
129	Submitted works	Manchester Metropolitan University on 2023-05-23	0%
130	Publication	Nidhi Gupta, Pushpraj Bhatele, Pritee Khanna. "Glioma detection on brain MRIs u...	0%
131	Submitted works	Polytechnic of Turin on 2024-04-20	0%
132	Submitted works	Sikkim Manipal University on 2024-04-08	0%
133	Submitted works	Study Group Australia on 2021-02-12	0%
134	Submitted works	The Robert Gordon University on 2021-05-04	0%
135	Publication	Toufique A. Soomro, Lihong Zheng, Ahmed J. Afifi, Ahmed Ali, Shafiullah Soomro, ...	0%
136	Submitted works	University College London on 2023-11-09	0%

137	Submitted works	University College London on 2024-09-30	0%
138	Submitted works	University of Bradford on 2023-03-18	0%
139	Submitted works	University of Cumbria on 2020-03-20	0%
140	Submitted works	University of Greenwich on 2024-11-26	0%
141	Submitted works	University of Surrey on 2024-09-03	0%
142	Submitted works	University of Sydney on 2024-11-07	0%
143	Submitted works	University of Wales, Bangor on 2024-09-30	0%
144	Publication	Vivek S. Sharma, Shubham Mahajan, Anand Nayyar, Amit Kant Pandit. "Deep Lear...	0%
145	Submitted works	Xiamen University on 2018-01-30	0%
146	Internet	depercitywi.iqm2.com	0%
147	Internet	easychair.org	0%
148	Internet	journal.50sea.com	0%
149	Internet	thesai.org	0%
150	Internet	tnsroindia.org.in	0%

151	Internet	
www.frontiersin.org		0%
152	Internet	
www.fusfoundation.org		0%
153	Internet	
www.preprints.org		0%
154	Internet	
www.techscience.com		0%

CHAPTER 1

INTRODUCTION

The algorithms used by the Artificial Neural Network (ANN) relate similarly to the structure and functions of the brain, which allows the processing of the data by it in a similar manner to that of the human brain. The patterns are created through the replication, and this is intended to help in informing the decision-making. In the view of Artificial Intelligence (AI), Deep Learning (DL) is specified as the subset of the Machine Learning (ML) methodology. In order to manage the complex data representations and tasks, the AI makes use of the complex algorithms and neural network architectures [1]. Machine learning has the methods to drive data analysis, thereby allowing the computers to learn from the provided information without any dependency upon explicit programming. Fig.1.1 represents the relationship between artificial intelligence, machine learning, and deep learning, showing DL as a subset of ML, which is a subset of artificial intelligence.

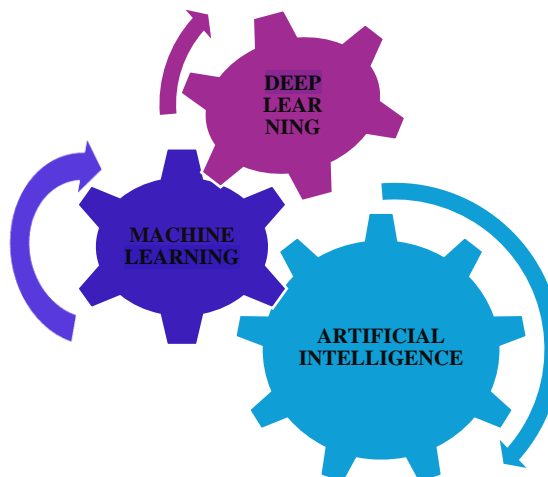


Fig.1.1. Relationship between artificial intelligence, machine learning, and deep learning

88 ML algorithms are further divided into two categories: supervised learning and unsupervised learning. Supervised learning differs from unsupervised learning as the former's algorithms get their knowledge from the labeled data, and the latter's algorithms extract their knowledge from the unlabeled data.

1 When there is an increased number of layers within the network, it reflects the characteristics of the DL, that normally denoted as the Deep Neural Network (DNN). Enabling fine-tuning by the backpropagation technique also permits the DNN to learn from the unstructured data. The DNN architecture consists of several levels or layers that work collaboratively to depict its features. In order to enable the sophisticated learning and representation of the intricate patterns in the data, the abstract information is provided by the higher levels in the network to a greater extent [2]. There has been a significant advancement in the DL due to the availability of abundant data in various forms worldwide, as witnessed by the digital era. The abundance of the data, often termed as big data, is derived from a variety of sources, such as e-commerce platforms, social media, internet searches, and business informatics. The presence of this abundant data has led to the sudden advancement in deep learning as it is using this data to automate applications.

Currently, there are a lot of DL-based computer vision applications that perform better as compared to human's performance in certain aspects. These aspects demonstrate the potential of DL technologies, offering a wider range from object recognition to tumor identification in Magnetic Resonance Imaging (MRI) scans and also as an indicator of blood cancer. Because of its exceptional performance across a variety of applications, deep learning is now identified as the most popular technique. To name a few such applications involving deep learning are object detection, voice recognition, face recognition, and medical imaging [3]. The unparalleled performance of deep learning in these domains has led to its further integration and growth as a promising technology in artificial intelligence and machine learning.

Medical imaging has emerged as the most essential technology, having been used for both diagnostic as well as therapeutic purposes. Even though the medical

professionals need further specialized training to analyze these medical images to help with decision-making about diagnosis and treatment based upon the results of these medical images. Further considering its disadvantages, medical image analysis is labor-intensive, time-consuming, costly, and error-prone process resulting in the surge/rise in the demand for computer-aided systems.

Recent advancements in DL have resulted in the enhancement in the clinical diagnosis techniques based upon the medical image analysis. To identify patterns, anomalies, and indicators in medical images, computer-aided systems can be trained by utilizing the DL algorithms. This will help medical professionals in reducing decision-making time and in making precise decisions, thereby making the use of medical image data for clinical purposes more effective and efficient.

133 There have been a number of reviews related to the use of DL in medical imaging written by various authors [[4], [5], [6], [7], [8], [9]]. However, only a very few number of researches have focused upon specific imaging modalities such as Electroencephalogram (EEG) [10], ultrasound [11], and MRI [12],[13]. These specific reviews and surveys give insight regarding the various developments and difficulties in each field, giving thorough information on how deep learning techniques are applied and optimized for specific medical imaging modalities. In recent years, DL has been considered as the de facto method in medical image analysis.

Fig. 1.2 depicts the tasks, application domains, and image modalities utilized in the area of medical image analysis.

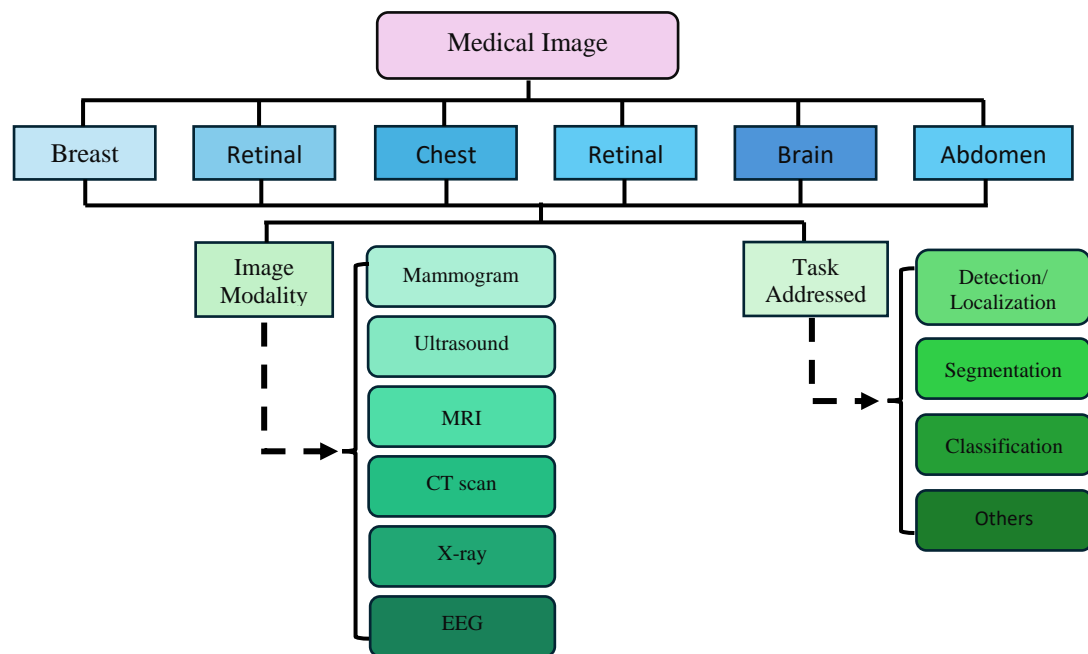


Fig.1.2. Demonstration of different tasks addressed, application areas, and image modalities used in medical imaging

The traditional ML methods encompass significant usage of domain knowledge along with human intervention for extracting the hand-engineered features. After that, these features are converted into a feature vector from raw data, which classifiers use as input to find and categorize patterns in images.

Deep learning algorithms, however, can directly map raw data—that is, pixels in the case of images—to desired outputs, which are typically labels or image classes. This represents a significant shift in medical image analysis, moving away from handcrafted features to non-handcrafted features where the system learns features directly from the data itself. This approach not only reduces the necessity of manually extracting features but also permits additional complex and nuanced patterns to be learned and detected, leading to improved accuracy and efficiency in medical image analysis tasks.

DL relies on well-specified training data to train deep networks effectively. In the context of medical images, a significant portion of medical data comprises images.

However, manually annotating these images by experts is tremendously time-consuming because of the large volume of data involved. Therefore, integrating DL methods with conventional machine learning techniques can lead to superior performance in medical image analysis tasks.

Rapid progress of DL techniques, particularly in the realm of medical imaging, has led to ongoing substantial advancements. Consequently, there is a continual stream of stimulating progress in the area of medical imaging, making it a successful area of study. Deep learning techniques and methods are constantly being explored by researchers and practitioners to improve medical image analysis and interpretation, which will ultimately result in improved patient care and more accurate diagnosis.

Because MRI is non-invasive and produces excellent contrast for soft tissues, it is regarded as a safe medical imaging modality. MRI does not change the construction, characteristics, or structure of particles inside the body like ionizing radiation-based methods do. The MRI setting offers possible hazards because of three magnetic fields that are, robust static magnetic fields, gradient-based magnetic fields, and pulsated radio-frequency fields which are used for the generation of 3-D images [14].

Notwithstanding possible risks connected to the MRI setting, it provides important details of tissue structures, including their location, dimension, and form. MRI is used in Computer-Aided Detection (CADe) systems and is highly respected in medical practice. MRI is a vital tool in contemporary medical imaging and diagnostic procedures because these systems use the detailed information from MRI scans to assist in the detection and diagnosis of numerous medical conditions.

The MRI utilizes radiofrequency signals in order to generate detailed images of the brain and other body organs. These images are the outcome after utilizing the radio waves, magnetic field gradients, as well as intensive magnetic fields. Magnetic Resonance Image is formed when the body's hydrogen nuclei or protons are aligned and spin in a phase being driven by the strong magnetic field. The remarkable ability

to distinguish between minute variations in tissue properties makes MRI a useful tool in medical diagnosis.

The difference between the MRI and other imaging techniques such as CT and PET scans is that the MRI doesn't use X-rays or ionizing radiation. The 2 subtypes of the MRI are structural imaging and functional imaging. The various types of structural imaging techniques are T1-weighted (T1w), T2-weighted (T2w), and Diffusion Tensor Imaging (DTI) which tends to offer details regarding the anatomy and tissue structure. However, the blood flow and the oxygenation level are measured conversely by employing functional imaging methods like functional MRI (fMRI), thereby recording alterations in brain activity [10].

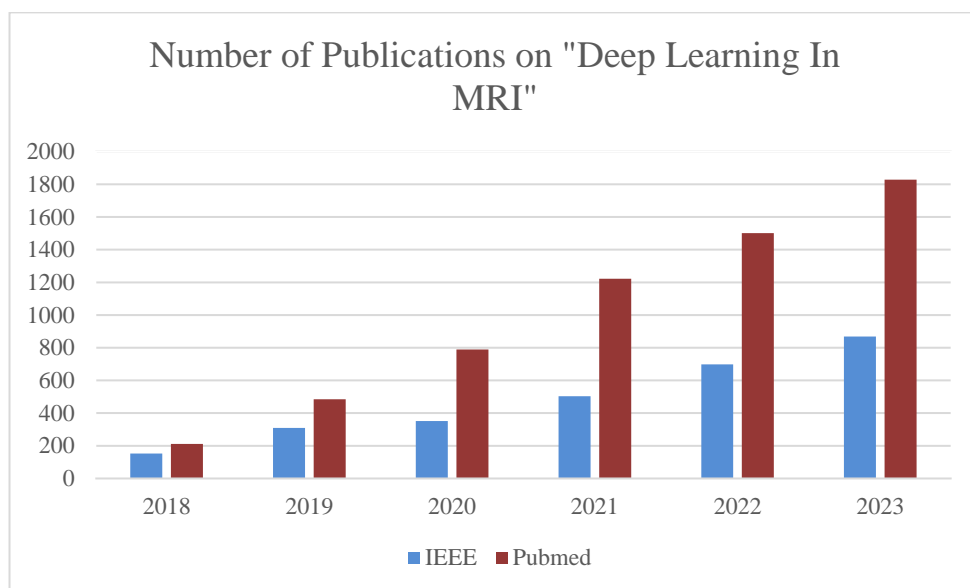


Fig.1.3. The quantity of articles published in "Deep Learning In MRI"

Fig. 1.3 is the annual depiction from 2018 to 2023 regarding the number of publications in the domain of DL in MRI. This tends to demonstrate the surge in interest and research efforts related to the usage of DL techniques to analyse, interpret, and improve the diagnostic capabilities of MRI technology. This demonstrates the increasing interest and research efforts focused on using DL techniques to analyse, interpret, and improve the diagnostic capabilities of MRI technology.

Image quality is an important component of Magnetic Resonance Imaging (MRI) analysis's efficacy. There could be a number of reasons for the deterioration of the MRI image quality, noise and distortion being the prime reasons. Below are the various reasons why this problem tends to occur:

- i. **Settings leading to Poor Imaging:** Improper selection or calibration of the imaging parameters leads to degraded image quality. This also includes parameters such as field strength, pulse sequence, and acquisition parameters. For example, the noisy or hazy images are the result of the wrong contrast or signal-to-noise ratio settings. (SNR).
- ii. **Transmission System Interference:** The presence of artifacts or interference during the transfer of the MRI signals can misrepresent the final images. The interference can be caused by factors such as RF (Radio Frequency) noise, Electromagnetic Interference (EMI), and also the patient's movement during the scanning procedure.
- iii. **Hardware Limitations:** MRI scanner hardware's problems such as broken coils, gradient coils, or receiver electronics, can also cause image distortions and noise.

Equipment maintenance, imaging protocols, and noise reduction techniques are some of the necessities upon which it is required to pay close attention during MRI scanning in order to reduce the above given factors. Image denoising factors and artifact correction techniques are the advanced signal processing algorithms to boost the accuracy of the MRI analysis and also improve its image quality.

The observed MRI image Z is described by the equation:

$$Z = S + N \quad (1.1)$$

where:

- S represents actual MRI image, which is a matrix of size $A \times B$ (where A and B are the dimensions of the image).
- Z represents the observed MRI image, which is also a matrix of size $A \times B$.

- N represents the noise added to the MRI image, which is also a matrix of size $A \times B$.

In Eqn. 1.1, the observed image, Z is the sum of actual MRI images, S and noise, N . There are various factors, such as system interference, patient motion, or other environmental factors, during the MRI scanning process that can cause the noise, N .

Rician distribution is often used to model the noise in an MRI scan [15]. The dynamics of this kind of noise depend on the strength of the signal. To be more precise, in high-density pixels, the Rician distribution is typically Gaussian distributed, whereas in low-intensity pixels, it is Rayleigh distributed. This property of Rician noise can cause low-contrast MRI images, which presents problems for MRI analysis noise removal techniques.

The Rician noise which is denoted as N having probability density function (PDF) expressed as:

$$P_R = \frac{R}{\sigma_N^2} \exp\left(-\frac{(R^2 + S^2)}{2\sigma_N^2} \mathbb{J}_0\left(\frac{RS}{\sigma_N^2}\right)\right) \quad (1.2)$$

In Eqn. 1.2, R is a random variable following the Rician distribution, while S represents the noise-free signal. $\mathbb{J}_0()$ refers to the zeroth order Bessel's function, and σ_N^2 is the noise variance. Eliminating this noise from MRI is crucial during the pre-processing stage. Various denoising algorithms for MRI images are utilized to minimize or eliminate noise (N) from the observed image (Z) and generate an estimated value (\widetilde{S}) that approximates the real value of S .

The brain supervises and coordinates a variety of body processes, acting as the body's central command center. It is a sophisticated organ that regulates not only the various bodily functions but also mental functions like memory, thinking, feeling, and

sense perception. The brain also controls breathing, body temperature, appetite, vision, and motor abilities.

142 The brain and its extending brain stem make up the central nervous system (CNS). Through chemical and electrical signals, this system makes it easier for the brain and the rest of the body to communicate. The regulation of particular body functions or processes is regularized by the brain by interpreting every signal it receives, thereby leading to the perseverance of general health and well-being. This system makes the body function to have seamless coordination.

The general well-being of an individual depends upon the brain's activity to function properly. Likewise, structural and functional abnormalities are marked as a resultant of brain abnormalities. The MRI tests tend to diagnose these aberrant brain conditions. Sometimes there are cases where the MRI results are found to be unreliable or equivocal, which makes it difficult for researchers and medical experts to diagnose particular brain disorders with accuracy.

MRI is identified as the potential diagnostic technique that provides fine-grained images of the anatomy and anomalies of the brain. These images are a combination of radio waves and magnetic fields and can be used for the diagnosis/detection of conditions such as tumors, strokes, and other neurological disorders. Occasionally, the MRI findings can be deceptive or imprecise. There could be several causes for it, such as technical issues with the scan, image artifacts, or difficulty in deciphering the findings, thereby leading to the inaccuracy in the MRI report, leading to the difficulty in diagnosis and thereby leading to failure in treating brain diseases. To arrive at a more precise diagnosis in such circumstances, medical teams frequently consult specialists, order additional tests, and conduct clinical evaluations. The dependability and accuracy of MRI scans are continually being improved by developments in imaging technology and interpretation methods, which aid medical practitioners in comprehending and treating brain abnormalities.

For a very precise diagnosis in such circumstances, the medical teams have very frequent consultations with the specialists, placing orders for additional tests and also conducting clinical evaluations. Due to the continuous development in imaging technology and interpretation methods, there is increased accuracy and dependability on MRI scans, which enhances the medical practitioners in treating brain abnormalities. The human brain is exposed to an extensive range of abnormalities, and its correct categorization presents a very substantial difficulty for medical professionals. One of the main obstacles is dealing with inappropriate or inaccurate MRI reports, leading to difficulty in the diagnosis process.

Brain tumors are portrayed among the main concerns when we talk about abnormalities in the brain. Brain tumor is designated as the aberrant development of cells within the brain tissue. Even though tumor refers to the unchecked cell growth in any part of the body [16], these can either be benign (non-cancerous) or malignant (cancerous) as per their nature. These brain tumors have different characteristics and related treatment options; therefore, it is important to understand how to differentiate between them. While malignant tumors necessitate rapid and intensive treatment, benign tumors may not be life-threatening right away and can be monitored over time.

The challenge lies in accurately identifying and categorizing brain abnormalities, especially tumors, using diagnostic tools such as MRI scans. Inaccuracies in MRI reports can lead to difficulties in distinguishing between various brain abnormalities, impacting the treatment approach and patient outcomes.

Brain tumor MRI scans provide essential information to doctors regarding the tumor's characteristics, such as its location, dimension, shape, irregularities, and internal structure. This qualitative or quantitative helps assess the efficacy of treatments and identify the tumor's growth stage [17]. It is crucial to remember that significant volumes of data are required for accurate medical image analysis, especially for training machine learning or deep learning models.

Traditional methods for survival prediction rely on clinical data and manually extracted features from MRI scans, which can be labor-intensive, time-consuming, and subjective [18]. This presents a problem in automated evaluation. However, advancements in MRI technology have resulted in the progression of automatic techniques to localize and classify tumors. In medical image analysis, transfer learning (TL)—an approach that makes use of previously trained models for image representations—is widely used [19].

The current techniques used in MRI analysis face several challenges:

- i. Intensity inhomogeneity is a significant issue in MRI images, occurring when there are variations in low-frequency intensity and non-uniformity. During image processing, this can lead to the loss of certain image characteristics, resulting in incorrect identification of tumor regions.
- ii. Brain MRI segmentation is a complex task, particularly due to challenges in segmenting non-uniform regions, which consumes a considerable amount of time and computational resources.
- iii. The quality of MRI analysis heavily relies on image quality, which is often impacted by noise and distortions. These issues can stem from inappropriate imaging settings or disruptions in the transmission system. Thus, it is crucial to remove these distortions as a preprocessing step to ensure accurate and reliable analysis.

1.1 MOTIVATION

The healthcare sector is a high-priority area, where medical experts interpret medical data most of the time. Due to their complexity, variability in the parameters, and also, fundamental understanding of the subject, medical image interpretation is largely restricted to specially qualified experts. Over time, the difficulty of collecting data has decreased due to advancements in image acquisition devices. Healthcare professionals have an increasing burden when there is a rise in data. Consequently, there may be a higher chance of human error. The brain and the spinal cord are the most significant parts of the human nervous system, and comprise the human CNS.

The brain manages the bulk of physical processes, such as analyzing, integrating, organizing, making decisions, and delivering directions to the body's other systems. The anatomical structure of the human brain is extremely complicated.

A brain tumor is an accumulation of aberrant cells that grow in a hard skull that surrounds the brain [20]. Any kind of extension in the interior of this kind of limited space may cause problems. Any kind of tumor expanding within the skull causes brain damage and poses a major threat to the brain [21]. Brain tumors are the most common reason for death among both adults as well as children. Tumors are of numerous forms, and depending on their texture, position, and shape, each one has a very poor chance of survival [22]. Brain tumors impact approximately 250,000 persons each year, with 2% of those instances being verified as malignant [23].

According to one study, brain tumors are approximately 85 and 90 percent of all significant CNS tumors [24]. Brain tumors should be detected early to significantly lower the rate of death. Medical professionals have extensively utilized medical imaging for tumor detection. MRI is a prominent approach for the early identification of brain malignancies. The radiologists employ manual methods to detect brain tumors.

Tumor grading time is determined by the radiologist's skills and expertise. On the other hand, tumor detection is an expensive and inaccurate process. Brain tumor misdiagnoses can have serious consequences and significantly decrease a patient's chances of survival.

The following reasons motivate me to work in this area:

- i. As DL techniques and AI models can improve and automate crucial medical imaging procedures, they offer significant benefits for brain tumor researchers. Convolutional Neural Networks (CNNs) are eminent DL frameworks that are effective at extracting complicated patterns from extensive brain datasets. For the purpose of medical decision-making and therapeutic planning, this facility significantly rises the accuracy of tumor identification, segmentation, as well

as classification. DL methods also encourage investigation and innovation because of their capacity for expansion and flexibility. In order to address issues like early identification of disease, researchers are continually improving these predictive models by exploring novel architectures and techniques. When it comes to the challenging area of brain tumor analysis, DL subsequently motivates researchers with its effective tools for knowledge advancement, enhanced diagnostic capabilities, and better treatment for patients.

- ii. Since brain tumors diverge widely in size, location, growing patterns, and responsiveness to treatment, it is critical to evaluate different features of the disease and its effects on patients using a variety of metrics. Accurate differentiation among tumor tissues and healthy brain structures is essential for brain tumor detection. Area Under Curve (AUC) of ROC curves, sensitivity, specificity, accuracy, and other metrics assess the degree to which diagnostic techniques or algorithms diagnose tumors accurately while lowering false positives. High specificity lowers the possibility of misdiagnosing healthy tissue as tumors, which is essential for timely intervention as well as for accurate diagnosis/detection. High sensitivity assures that true positive cases are identified. Accurate location and classification of brain tumors are essential for managing treatment, directing surgery, and tracking the advancement of the disease. The spatial correspondence between automatic or manual segmentations and ground truth regions of tumors is measured using metrics such as the Dice coefficient, the Hausdorff distance, and the Jaccard index (IoU). These parameters evaluate the segmentation algorithm's accuracy and reliability in defining tumor boundaries.

1.2 ORGANIZATION OF THESIS

The whole thesis is arranged into the following nine chapters: Chapter 1 covers the introduction of the work done. Chapter 2 describes a literature survey related to the previous work done in the field of abnormalities/tumors of brain images. It acts as a foundation, providing a critical review of existing research. By examining previous

studies, the researchers can identify any gaps in knowledge and establish clear objectives for their own investigation.

Denoising of magnetic resonance images (MRI) is done in Chapter 3. In MRI, denoising is vital for the improvement of the quality of images because it minimizes noise artifacts, which may cover essential information used for detection. This crucial process guarantees the clarity as well as accuracy of the data before further analysis.

Chapter 4 then shifts gears, focusing on the application of transfer learning techniques for diagnosing brain tumors. Transfer learning uses DL models that have previously been trained to identify the MRI images as either normal or cancerous. This technique improves diagnostic accuracy through the adaptation of large dataset knowledge to new, smaller datasets dedicated to brain tumor analysis.

Detection of brain tumors using a triplanar multi-layer morphological convolution neural network is discussed in Chapter 5. This technique focuses on the aspects of reducing the number of computations and time complexity so as to make a cost-effective system. This innovative technique likely involves analysing MRIs from different planes (axial, coronal, sagittal) and extracting intricate features using a specifically designed CNN architecture.

Chapter 6 describes the detection of brain tumors by amalgamation of segmentation and rendering technique. In order to aid clinicians, segmentation splits MRI scans into useful regions, and rendering makes such separated areas apparent to locate and identify tumors,

Chapter 7 shows the segmentation and detection of pixels in brain tumor image by the aggregation of GAN (Generative Adversarial Network) models with a vision transformer. GANs create synthetic images to augment training data, while the vision transformer improves feature representation, enhancing the accuracy of tumor segmentation and detection tasks.

Chapter 8 introduces an autonomous and intelligent system for classifying various brain abnormalities. This potentially involves a hybrid model combining CNNs, recurrent neural networks (RNNs), and Long Short-Term Memory (LSTM) networks to handle complex temporal dependencies present in brain data.

Finally, the conclusion of the thesis provides an overview of findings and suggestions of some prospective studies for future scope in Chapter 9. The conclusion emphasizes the established approaches potential for clinical applications and emphasizes their significance in enhancing the detection of brain tumors. The thesis's social implications for society are also addressed in this chapter.

In summary, the thesis offers a thorough investigation of novel techniques for MRI analysis used in the identification of brain tumors/abnormalities. With its unique contributions to the area, each chapter addresses significant issues related to image denoising, detection, segmentation, and classification.

CHAPTER 2

LITERATURE SURVEY

Radiology, also termed as medical imaging, is the medical specialty that involves the creation of images of the body parts for the purpose of diagnostics or treatment. The procedures involved therefore help the medical professionals in diagnosing the diseases and injuries. The medical image analysis of the brain is considered as a major area of interest because of its complexity and significance, and the automation of the same can be done using various tools and techniques. There are a variety of image processing techniques used for brain image analysis, to name a few are the DL, ML, hybrid models, etc. There are a variety of reasons, such as the shape, dimension, textures, and other related features, due to which the analysis of brain tumors tends to become complicated. Henceforth, this chapter will give a comprehensive review of the brain tumor image analysis, with the inclusion of topics such as the fundamentals of brain tumors, brain imaging, actions involved in brain image analysis, models utilized, characteristics of brain tumor images, metrics for model evaluation, and datasets of brain tumor and medical images that are available.

2.1 INTRODUCTION

As the human brain is one of the utmost essential components of the human nervous system, therefore researchers have focused a lot on carrying out medical image analysis on brains in order to identify neurological disorders. This analysis is being used extensively and collaboratively by researchers in the medical and technology domain to identify and assess diseases like Parkinson's disease, Alzheimer's disease, and brain tumors/cancers. As per the stats, over 200,000 people die each year due to brain cancers, that initially start as brain tumors. Hence, it is very significant to develop novel technologies and methods for diagnosing and treating these conditions effectively [25].

Tumors are seen as the 2nd leading cause of death. Tumors are basically the abnormal proliferation of tissues. The non-controlled division of the cells gives rise to these growths, and these are irregular in shape, and overabundance of the cells is generally the root cause for the same, leading to the deviation from the normal cell life cycle formation, growth, and death.

Usually, the organ or body part from which a tumor originates is used to name the tumor. The tumors that initiate from any part of the brain or the skull that also includes the bones, tissues, membranes, and brain nerves are referred to as the brain tumors. The high death rate associated with these tumors—roughly 70%—highlights how seriously they can affect health outcomes. In order to effectively address the challenges presented by brain tumors, early detection, and effective treatment strategies are essential. [26]

Based on their precise location within the brain, brain tumors can be classified into more than 150 different types. These growths fall into one of 2 categories: benign or malignant. Examples of benign brain tumor comprise meningiomas, chordomas, gangliocytomas, and others. These are benign tumors that usually grow slowly and have little chance of spreading to other areas of the body. Malignant brain tumors, on the other hand, are cancerous and represent a greater risk to health. Glioblastomas, medulloblastomas, and ependymomas are a few types of malignant brain tumors. These tumors have a worse prognosis and are harder to treat because of their rapid growth and ability to invade neighboring tissues. Making the distinction amongst benign and malignant brain tumors is vital for choosing the best course of action and projecting the possible outcomes for patients. Depending on tumor's kind, location, and stage, treatment for brain tumors may involve radiation therapy, chemotherapy, or surgery [27].

The brain tumors are classified into 4 classes: Categories 1 and 2 are referred to as the "low-grade" tumors, as they tend to grow slowly and are usually benign [28]; Categories 3 and 4 are "high-grade" tumors, which are considered to be malign; also have a tendency to reoccur [29].

Category I. These tumors, like pilocytic astrocytoma, grow slowly without metastasizing, and because of their well-defined nature and low propensity for aggressive behaviour, they are frequently completely surgically removed.

Category II. These tumors, such as oligodendroglioma, grow more slowly and may eventually spread to higher grades by invading surrounding tissues. These tumors may not go away and may even get bigger.

Category III. These tumors, such as Aden squamous astrocytoma which grow more rapidly as compared to Grade II tumors and have the ability to spread to neighboring tissues. After surgery, these tumors usually need additional treatment, such as chemotherapy or radiation therapy, because surgery is frequently insufficient to treat them.

Category IV. Glioblastoma multiforme is a Grade IV tumor and is considered to be extremely aggressive and having its likeliness to be spread. The most dangerous and aggressive kind of tumors are said to be those that use blood vessels to speed up their growth.[12]

To improve the patient's probability of recovery, appropriate and effective treatment needs to be provided, and for that, the brain tumors need to be accurately detected and classified. However, due to diverse shapes, sizes, appearances, positions, and scanning parameters of the tumors, it becomes quite a difficult task to identify the tumors.

There are various sophisticated as well as conventional methods that are being used in identifying a brain tumor. To name a few, conventional techniques such as the Leksell Gamma Knife (GK) and radioactive beams are quite efficient but have some disadvantages of being labor intensive and needing significant human intervention. The modalities such as the Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI), and computer tomography (CT) are being used frequently for brain tumor detection. Chemical Exchange Saturation Transfer (CEST) is considered as a special technique of the MRI which promotes the possibility of imaging substances that are at very low concentrations and may not be visible using the standard MRI methods. Of these modalities, MRI is a non-invasive technique that

employs microwave pulses and magnetization to provide exhaustive internal body structure imaging.

The ability of the human eye to discern minute variations in intricate imaging methods is restricted. In order to help radiologists make accurate diagnoses, researchers have created Computer-Aided Diagnosis (CAD) systems. Therefore, applying ML is crucial to solving this problem. Although MRI images can be used to accurately identify tumor regions using traditional ML techniques, DL models are becoming more and more popular because of their superior performance and capacity to handle large, complex datasets and high-performance computing devices.

2.2 SURVEY METHODOLOGY

The PRISMA methodology is considered to provide a guaranteed methodical exacting approach and is used extensively in brain tumor detection. PRISMA offers an improved review structure by offering precise instructions for choosing, locating, and assessing pertinent research. The PRISMA is being used due to its ability to preserve the rigour and transparency in their review process, that tends to facilitate the interpretation and validation of the study results. PRISMA [30] contributes to a thorough and trustworthy analysis of brain tumor detection techniques by mitigating bias and guaranteeing that all pertinent literature is taken into account.

There are a few databases that were used in this study: IEEE Xplore, Web of Science, Academic Search, Springer, SPIE Digital Library, ACM Digital Library, Elsevier, MDPI, Google Scholar, Scopus, and PubMed. In addition, searches were conducted for pertinent papers on four preprint platforms: ArXiv, TechRxiv, MedRxiv, and ChemRxiv. Additionally, a manual search was done. During the first phase, 8356 research publications were found in total. Fig. 2.1 shows the search criteria used for finding research articles.

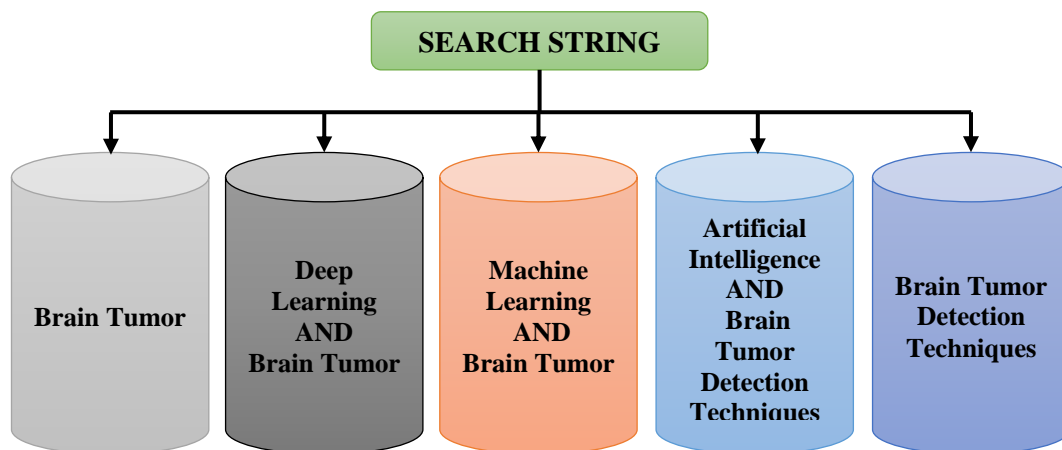


Fig.2.1. The search criteria used for finding research articles

Following the removal of duplicates and unnecessary articles, 1976 research articles were carefully chosen for further review. On the basis of title, abstract, and introduction, 1224 of these were disqualified. After additional analysis, the remaining 651 were not included. That left 101 articles for the following stage. 19 articles were removed during the eligibility phase following a careful assessment.

The criteria to either select or remove the different research papers are given in Table 2.1. The diverse PRISMA stages that were adhered to in this review are depicted in Fig. 2.2.

Table 2.1 Criteria for Selecting and Removing Various Research Articles

S. No.	Parameter	Criteria of Selection	Criteria of Removal
1	Time Frame	Research articles that are published from 2014-2024	Research articles that are published prior to 2014
2	Analysis	Research articles encompassing various brain tumor detection	Research articles encompassing diverse cancer detection
3	Comparison	The research article focuses on DL methods utilized for brain tumor detection	The research article focuses on diverse methods used for detecting brain tumor
4	Study	Research incorporating experimental findings and mathematical foundation	Research incorporating case studies and articles in languages other than English

139

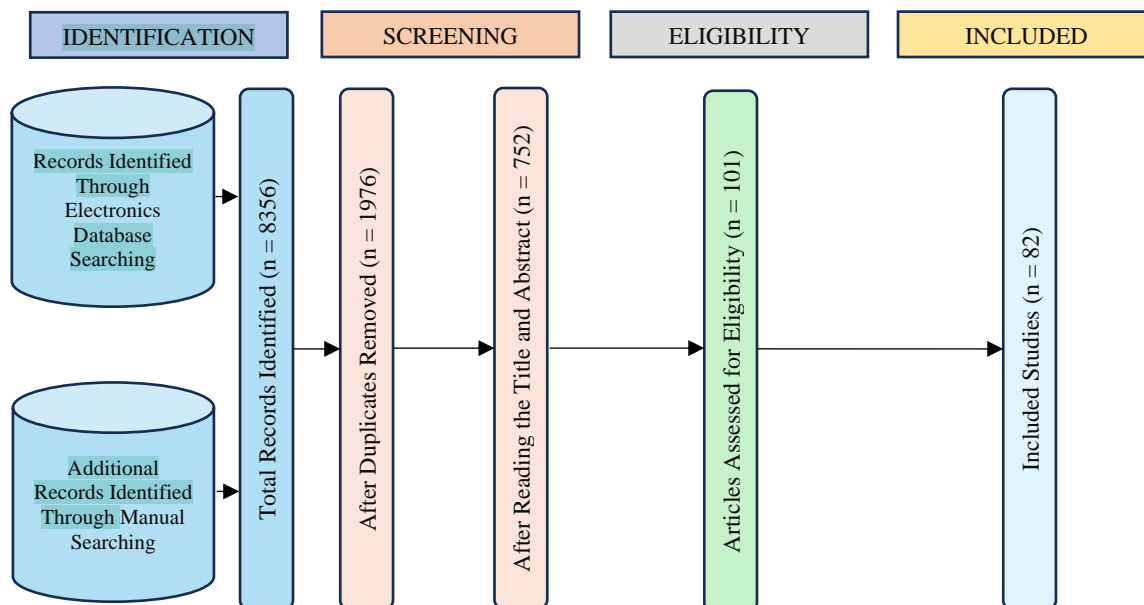


Fig.2.2. PRISMA flow diagram on brain tumor review strategy

2

2.2.1 Research Questions Asked by Researchers

The primary goal of this survey is to furnish a wealth of detailed information to enlighten and motivate emerging scientists in the domain of brain tumor detection. Table 2.2 outlines various questionnaires pertaining to brain tumor detection that have been explored in this study. This compilation aims to aid forward-thinking scientists in grasping the foundational aspects of detecting and diagnosing brain tumors, along with highlighting the unresolved challenges within the domain.

19

Table 2.2 Research questions correlated to brain tumor detection

Questions	Research Questions
Q1	What is a brain tumor?
Q2	Describe the detection methods for brain tumors?
Q3	What are the risk factors linked to brain tumors?
Q4	What are the various screening techniques used in detecting brain tumors?
Q5	What types of DL methods are employed in brain tumor detection?
Q6	How are DL-based brain tumor detection techniques evaluated for performance?
Q7	What challenges are associated with using DL for brain tumor detection?
Q8	What are the Future researches being conducted to explore the detection of brain tumors using DL techniques?
Q9	What role does deep learning play in detecting brain tumors, and how does it differ from other approaches?

In Table 2.3, a checkmark (✓) showcases towards the fact that the study has addressed the given research queries, while a cross (✗) showcases that it has not been addressed.

Table 2.3 Research question recent survey related to brain tumor answered

Survey	Review Year	Research Questions								
		Q 1	Q 2	Q 3	Q 4	Q 5	Q 6	Q 7	Q 8	Q 9
[31]	2019	✓	✓	✓	✓	✓	✗	✓	✗	✓
[20]	2019	✓	✗	✓	✓	✓	✗	✓	✗	✗
[32]	2020	✓	✗	✓	✗	✓	✗	✓	✓	✓
[33]	2023	✓	✗	✓	✓	✓	✓	✓	✗	✓
[34]	2023	✓	✗	✗	✗	✓	✓	✓	✓	✗
Proposed study	2024	✓	✓	✓	✓	✓	✓	✓	✓	✓

2.3 BRAIN TUMOR ANALYSIS

As illustrated in Fig. 2.3, tasks related to analyzing brain tumors can be categorized into three primary areas: (i) detection of brain tumors versus non-tumors, (ii) segmentation of brain tumors, and (iii) classification of brain tumors. Researchers have approached these tasks using various methodologies. Different types of brain medical images are utilized by the AI based systems as their input data. The analysis of the image to detect whether the tumor is present or not is carried out by the tumor detection task, with the output being a decision (whether the tumor is present or not). The tumor segmentation task's aim is to identify the precise region consisting of tumors, thereby extracting that region from the remaining images.

The tumor classification task can include actions such as assessing the tumor's severity (such as benign/malignant) or identifying various tissues within the tumor (like the core tumor, edema, enhancing tumor, non-enhancing tumor, etc.).

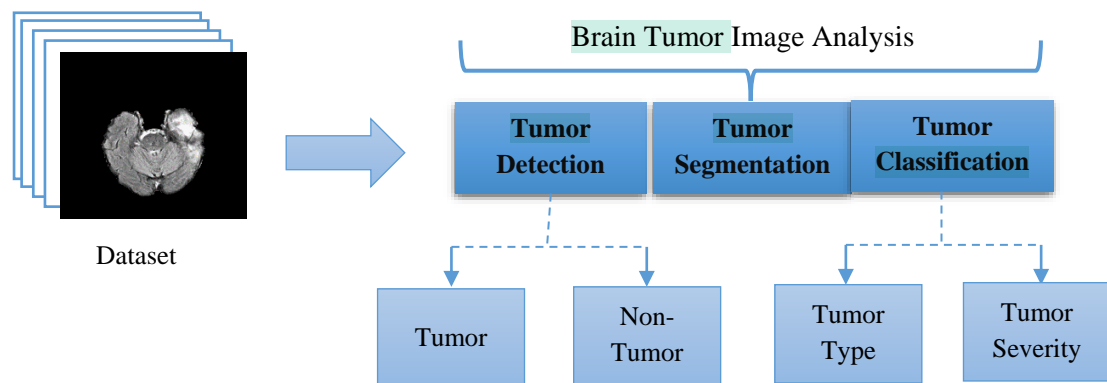


Fig.2.3. AI-driven system designed for tasks related to analyzing brain tumors

Brain tumor detection involves classifying brain MRIs as either tumor or non-tumor cases, determining whether the brain is healthy or exhibits abnormalities. A range of techniques have been deployed by the researchers, which include ML models, DL model approach, and hybrid models for carrying out the classification task. [35]. This classification depends upon the datasets that comprise healthy brain MRIs as well as those showing tumors, with labeled data used to train various models such as supervised, unsupervised, and deep learning models. These models extract distinctive features from both healthy and tumor tissues to facilitate accurate classification.

In certain studies, tumor classification is also utilized to identify specific tumor types, such as distinguishing between benign and malignant tumors or identifying particular types of tumors [36]. However, these tasks are often not examined independently; rather, they're integrated into the brain tumor segmentation task. In analyzing brain MRIs, recent works have predominantly focused on utilizing deep learning models for automatic feature extraction, known as deep feature extraction. The extraction of various intensity features, shape features, and texture features from MRIs. are some alternative approaches upon which the research has focused considerably. These features are either combined with the images or utilized solely as inputs for the system, leading to effective results in terms of detection and segmentation [28].

5
3
102
The brain tumor image segmentation task involves splitting images to extract a specified area, known as the region of interest (ROI), depending on specific conditions relevant to the problem. Brain tumors can comprise various classifications of tissues, such as core tumor tissues, edema tissues, enhancing tumor tissues, non-enhancing tumor tissues, etc. In recent years, image segmentation has gained significant popularity amongst researchers from the medical field to analyse medical images like MRIs, CT scans, etc. This technique is crucial for accurately identifying and analyzing different components within medical images, especially in complex conditions like brain tumors.

Image segmentation methods are valuable for identifying distinct regions or clusters of pixels that encompass regions of interest (ROIs), organs, or lesions within organs. This area of study forms the basis of medical image segmentation [37]. Noninvasive brain scans are being utilized into brain tumor image segmentation as explored by researchers across diverse fields. Their aim is to develop effective and precise automated systems for brain tumor segmentation that are capable of segmenting tumors, classifying tumor types, and assessing severity. These systems are pivotal in detecting abnormalities promptly and facilitating timely treatment [27].

107
The majority of existing research is concentrated on the brain tumor image segmentation and classification task, as accurately delineating the tumor area from the surrounding image is critical for diagnosis. It's worth noting that only a few researchers have addressed multiple tasks simultaneously, primarily accounting for the lack of suitable datasets and the complexity associated with each task. Consequently, there's a significant future scope in developing a comprehensive and precise system capable of autonomously handling all tasks without human intervention, which remains a key objective in the realm of brain tumor image analysis.

2.4 DATASET USED

As seen in the different research papers, researchers used number of publicly available datasets of MRI modality to test the methods suggested by them. Table 2.4

shows a list of some well-known publicly available datasets of MRI modality used in the research.

Table.2.4 List of some publicly available datasets of MRI modality used in the research work

Dataset Name	Dataset Size	Description	Link
BRATS2019 [38][39][40]	259 high-grade glioblastomas (HGG) and 76 low-grade gliomas (LGG) MRI volumes	Contain T1, T1-W(T1Gd), T2- W(T2), and T2-FLAIR of Glioblastoma (GBM)/HG glioma and LG glioma brain tumor	https://www.med.upenn.edu/cbica/brats2019/data.html
BRATS2020 [38][39][40]	55 brain slices where each slice is of dimension 240 x240 pixels.	T1, T1-weighted, post-contrast T1-weighted (T1gd), T2-weighted, and T2-flair volumes annotated as necrotic and non-enhancing tumor core (NET), enhancing tumor (ET), peritumoral edema (ED).	https://www.med.upenn.edu/cbica/brats2020/data.html
The Cancer Imaging Archive (TCIA) Publicly Accessible Repository [41]	20 Patients	DICOM Format. contain T1w (pre- and post-contrast agent), FLAIR, T2w.	https://www.cancerimagingarchive.net/collection/brain-tumor-progression/
Harvard Medical School Data (AANLIB) [42]	400 Patients	Contain normal and abnormal images includes cerebrovascular (stroke or "brain attack"), neoplastic (brain tumor), degenerative, and inflammatory diseases.	http://www.med.harvard.edu/AANLIB/
Masoud2021 dataset [43]	7023 images	4 classes: glioma, meningioma, no tumor and pituitary	https://doi.org/10.34740/KAGGLE/DSV/2645886
Figshare dataset [44]	3064 images	Consists of 3064 T1-weighted contrast-enhanced images from 233 patients with three kinds of brain tumor: meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices).	http://dx.doi.org/10.6084/m9.figshare.1512427

2.5 RELATED WORK

There has been a number of research that has been carried out for analyzing medical images and studying brain tumors in the last few decades. Many studies have looked at and summarized existing works from different angles and time periods. In [45], there is a detailed discussion of medical image analysis utilizing semi-supervised machine learning, multi-instance learning, and TL techniques for unique organs, including references to relevant systems. Certain survey papers, such as [37], have explored current methods relying upon DL models for the analysis of medical images, giving an outline of their problems and proposed alternatives.

Tiwari et al. [27] conducted a comprehensive review of current methodologies for brain tumor segmentation and classification in the MRIs of the brain. They made an analysis of 62 research papers published between 2014 and 2019. Their review includes details on MRIs and various imaging techniques related to the brain and their outcomes. It gives insights regarding the differences and gains of utilizing MRIs to derive gray matter, white matter, and cerebrospinal fluid separately. Additionally, their review also focused on the super classes and subclasses for brain tumor classifications, including a comprehensive categorization of gliomas and tumor grades depending on the degree of severity and their location within the brain. The review includes different methods for segmenting brain tumor images. This involves techniques like thresholding-based methods, supervised machine learning, and unsupervised methods. The review additionally covers tumor classification methods, both supervised and unsupervised. The various aspects, such as the technique, measurement of performance, true and false positive rates, and data sets, have been summed up in a tabular format along with the textual summary from 62 articles. The review provided a comprehensive list of 19 datasets that have been utilized in previous studies on brain tumor segmentation. The authors gave references and usage frequencies for each dataset, giving a clear overview of the benchmark datasets and their availability. They highlighted that BRATS 2013 and BRATS 2015 are most frequently utilized in recent research. Moreover, the researchers presented statistical results regarding the reviewed articles country of origin, publishers, citations, publication years, datasets utilized, and

methodologies deployed. They compared methods and found that CNN and particle swarm optimization (PSO) have been dominant in most current years, with support vector machine (SVM), Fuzzy C -means (FCM), and k-means ranked 2nd from 2014 to 2019. Various performance metrics along with the appropriate equations utilized in the analysis of the brain tumors have also been explained by the authors. There is the provision of network visualization maps for the terms and their related links to give a basic understanding of the terminologies utilized in the related articles. The authors concluded that DL and PSO methodologies are the existing and future focus of research in this study, and their review of the most current findings on brain tumor segmentation and classification has validated this conclusion. The review paper gave a detailed overview of recent research, covering all 62 articles. It included tables and statistical analysis for easy comparison of different methods, datasets, publications, and performance metrics. Some summaries were concise because of the similarities with different works, and the paper identified the methods utilized in different studies by name. It listed benchmark datasets with their names and references. It would be helpful to incorporate additional information about the data sets, like the no. of items, data type, public availability, and ground truth availability, for selecting a suitable dataset for future research.

A study on tumor segmentation and feature extraction techniques related to the brain was carried out by Saman et al. [28]. The study discussed the magnetic resonance imaging systems, technology for magnetic resonance imaging processing, and magnetic resonance imaging sequencing, providing explanations and instances. It also detailed different levels of brain tumors, their region, symptoms, and curing methodologies. The brain tumor detection studies from the past 10 years were reviewed by the paper, with primary attention on research from the past 3 years. The authors covered the pre-processing of MRIs in detail, discussing steps such as stripping of the skulls, correction of the bias field, removal of noise, object detection, and segmentation utilizing various tools and algorithms. They provided an overview regarding the current research over the segmentation of brain tumors, categorizing it into conventional methods, intensity-based methods, atlas-based methods, deformable model-based methods, as well as hybrid and deep learning methods. The authors also

presented pros as well as cons of using various studies deploying these methodologies in a table format. They parallelly analyzed and reviewed the most current study related to feature derivation methodology for tumor analysis related to the brain. The paper defined and described several kinds of features, inclusive of texture-based features, intensity-based features, and shape-based features, along with their corresponding equations. Additionally, the authors discussed DL-based automation feature derivation techniques.

5 The authors discussed the recent emphasis on different CNNs and U-Nets along with their modifications. Their comprehensive analysis of the features used in brain tumor segmentation was informative for understanding and comparing feature extraction and processing in similar systems. The paper extensively referenced and mentioned over 130 research works from the past decade, contributing to the analysis of renowned methods and their performance related to the research on brain tumors. However, there were a significant number of pertinent articles were present in the references, and many of which were inadequately summarized, thereby making it challenging to draw conclusions based on comparison without individually reviewing the cited papers.

Certain reviews upon brain image analysis have focused on specific aspects of image processing of brains, like skull stripping. They utilized available work and data sets as references [46], [47]. These methods were occasionally used as former or later-processing steps for analysis related to brain tumors or were connected with studies related to brain tumors. The datasets reviewed in this study were also utilized in the analysis method of the brain tumor. Table 2.5 summarizes these survey papers.

Table 2.5 Recent Survey Related to Brain Tumor Analysis

Ref	Year	Approach	Technique Used	Details
[26]	2021	Classification and Segmentation	Deep Learning	Covering publications from 2018 to February 2020, this review provides a comprehensive overview of several topics.

Continued on Page-29

Read Table 2.5 (continued)				
				These include the datasets used, the deep learning methods applied, the performance metrics assessed, a detailed explanation of CNN architecture, the popular frameworks of CNNs used in medical image analysis, and the challenges faced in brain tumor analysis.
[48]	2021	Segmentation	Deep Learning	The review provides a thorough overview of image segmentation, covering different deep learning models, architectures, datasets, performance metrics, and software tools used. It also thoroughly examines publications focusing on specific architectures, including U-Net, dual-pathway, single-pathway, ensemble, and cascaded architectures.
[27]	2020	Classification Segmentation	Conventional ML	Publications from 2014 to 2019 analyze brain tumor prediction using MRI, presenting insights into methods, datasets, performance metrics, and future challenges.
[47]	2020	Skull-stripping	Conventional ML and DL	An overview of datasets with visualizations, along with a review of histogram thresholding, morphology, deformable surface-based, atlas-based, region-growing, meta-heuristics, and 2D/3D CNN methods
[37]	2019	Segmentation	Deep Learning	The review offers a comprehensive look at CNN architectures, covering 2D CNN, 2.5D CNN, and 3D CNN, as well as variations like FCNNs, U-Nets, CRNs, and RNNs. It also delves into different network training techniques and transfer learning methods. Additionally, the paper addresses challenges in medical imaging and proposes potential solutions to tackle these challenges effectively.
[31]	2019	Segmentation Classification	Conventional ML and DL	A comprehensive discussion of CAD for brain MRI and datasets, including a comparative analysis of detection/classification methods, feature extraction, performance metrics, and datasets, as well as the limitations in tumor detection, classification, and segmentation.
[49]	2019	Segmentation	Conventional ML	A detailed comparison of conventional and machine learning methods, analyzing MRI modalities, time, data, abnormality types, and performance in publications on BRATS 2015 and BRATS 2013.

2.6 FEATURES EXTRACTION METHODS FOR BRAIN MRI

Brain image analysis requires the capability to extract features from brain MRIs, which are then utilized by various models and algorithms to produce outputs for analyzing brain functions. A wide array of features and models have been employed in brain image analysis studies, and they are outlined below:

- i. **Extraction of Features:** Image analysis involves the extraction of meaningful features from brain MRIs.
- ii. **Model Utilization:** The diverse models and algorithms utilized such features subsequently to create the outputs for the tasks related to brain analysis.
- iii. **Features Types:** Several Features are explored within the study that relate to the analysis of the brain image utilizing aspects such as the intensity, shape, texture, frequency, and various deep learning-based features.
- iv. **Types of Models:** The models such as inclusion of the ML algorithms and DL architectures, are applied in the study of brain image analysis in order to process the extracted features and get valuable insights.

These models and features combined give advancements in our understanding of the brain structure, functions, and abnormalities, thereby fostering progress in neuroimaging research.

Brain MRI's derived features are categorized into 4 classes and are identified in [28]. The classes involve intensity-based, texture-based, shape-based, and deep learning-based features, as illustrated in Fig 2.4.

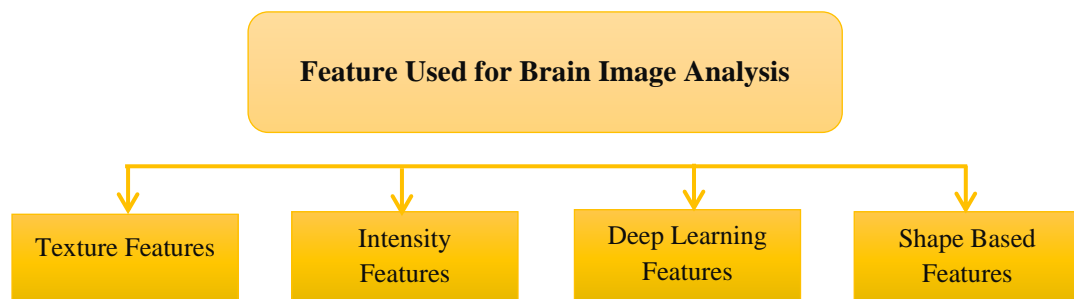


Fig.2.4. Types of features used in brain tumor analysis

2.6.1 Texture-Based Features

The collection of measurements used in image analysis for evaluation of the apparent texture of an image is referred to as the Image Texture. It provides insights into how colors or intensities are arranged within an image or a specific part of an

image. Key texture-based features commonly used in analyzing brain images include gray-level co-occurrence matrix (GLCM), autocorrelation, correlation, homogeneity, energy, entropy, local binary pattern, dissimilarity, contrast, cluster shade, and prominence [50].

2.6.2 Intensity-Based Features

Derivation of the Intensity features is from the order one histogram, which encompasses 4 types of statistical measures: variance (Var), mean (\bar{X}), kurtosis (k) and skewness (S_k). These features are computed based on the pixel values of grayscale present within the image. Let p_i be denoted as the pixel value for a specific pixel referred to as ' i ' lying within the grayscale image boundary. Moreover, assuming the overall pixel count is t , then the statistical mean indicates the central tendency related to the image's histogram. Variance quantifies the spread or dispersion of pixel values, skewness reflects the histogram's asymmetry, and kurtosis gauges the histogram's peakedness or flatness. Eqns 2.1, 2.2, 2.3, and 2.4 are typically employed to calculate these four features.

$$Var = \frac{1}{t} \sum_{i=1}^t (p_i - \bar{X})^2 \quad (2.1)$$

$$\bar{X} = \frac{1}{t} \sum_{i=1}^t p_i \quad (2.2)$$

$$k = \frac{1}{t} \sum_{i=1}^t (p_i - \bar{X})^4 \quad (2.3)$$

$$S_k = \frac{1}{t} \sum_{i=1}^t (p_i - \bar{X})^3 \quad (2.4)$$

2.6.3 Features Based on Deep Learning

DL features are the combination of simple as well as intricate features that are automatically gathered by deep learning models on each level related to CNN. As the network progresses through its layers, it discovers new features and generates feature maps. For instance, early layers might detect edges, followed by corners in subsequent layers, and progressively more complex shapes and patterns. Each layer captures increasingly detailed feature information from the input image, creating feature maps or vectors that are then processed via the corresponding neural net. These features are consolidated and applied in tasks like detection related to brain tumors, segmentation, and classification.

2.6.4 Features Based on Shapes

Shape-based features describe the form of objects and include metrics such as area, volume, perimeter, convex area, polygon shape, and circularity/irregularity, among others. These features can be derived from a brain MRI as part of shape-based analysis.

2.7 PERFORMANCE EVALUATION METRICS

For the evaluation of the performance related to the brain cancer prediction models, there are various performance metrics utilized. These performance methods are further classified also referred to as prediction and classification metrics. Fig. 2.5 represents the performance measures classification.

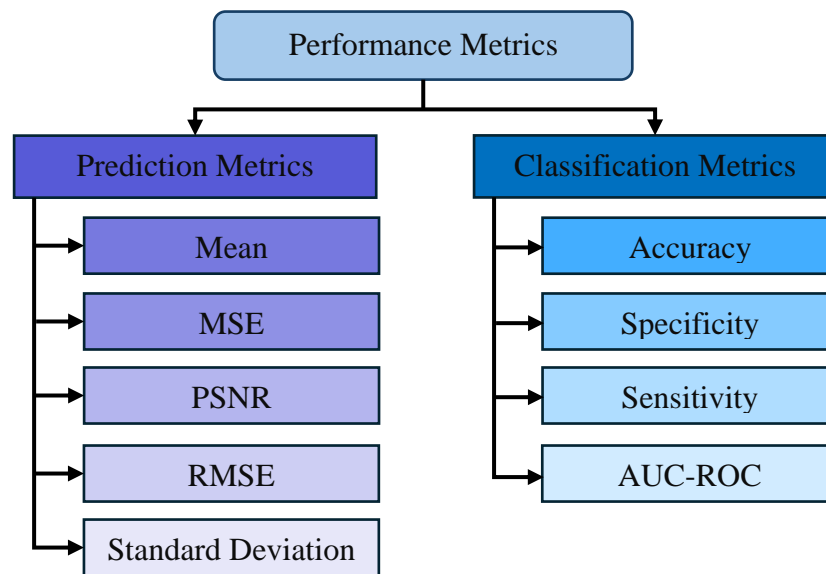


Fig.2.5. Classification of performance measures

2.7.1 Prediction Metrics

Well-known prediction metrics include Mean, standard deviation, mean square error, root mean square error, and peak signal-to-noise ratio. Subsequent sub-sections provide definitions and mathematical formulation of such metrics.

A Mean (μ)

Mean is defined as the average brightness of an image. The mathematical formulation of the mean (μ) is defined as [51]:

$$\mu = \frac{1}{mn} \sum_{a=1}^m \sum_{b=1}^n pA(a, b) \quad (2.5)$$

where m indicates the no. of rows and n indicates the no. of columns when an image is being considered. pA indicating the coefficient of approximation. The mean value must be on the higher side in order to achieve good outcomes[51].

B Standard Deviation (σ)

Standard deviation indicates by what value the contrast intensity should increase when there is an increase in irregularity of the textures [52] and is defined as:

$$\sigma = \sqrt{\frac{1}{mn} \sum_{a=1}^m \sum_{b=1}^n (pA(a, b) - \mu)^2} \quad (2.6)$$

C Mean Square Error (MSE)

MSE is the parameter used to measure the differential among the observed and the predicted values [53]. The mathematical formula for the MSE is as given below:

$$MSE = \frac{1}{mn} \sum_{a=1}^m \sum_{b=1}^n (I_o - I_p)^2 \quad (2.7)$$

where I_o represents the observed value and I_p represent the predicted value.

D Root Mean Square Error (RMSE)

RMSE is the parameter calculated by the square root of the second moment of difference between the values that were observed and that were predicted [54]. It can be explained as the standard deviation of prediction errors.

$$RMSE = \sqrt{\frac{1}{mn} \sum_{a=1}^m \sum_{b=1}^n (I_o - I_p)^2} \quad (2.8)$$

The accuracy and the preciseness obtained from the prediction models lies upon the RMSE which is a reliable model. It consistently maintains a non-negative value at every instance proportional to scale and is notably sensitive to outliers.

E Peak Signal-to-Noise Ratio (PSNR)

The ratio of the max possible power of the image that can be obtained to the power of noise that deteriorates and lowers its quality of representation is referred to as the Peak Signal to noise ratio [55].

$$PSNR = 10\log_{10} \frac{(l-1)^2}{MSE} \quad (2.9)$$

Here, l refers to the number of image's maximum permissible intensity levels.

2.7.2 Classification Metrics

Classification metrics are used to evaluate the brain tumor prediction model's performance. Well-known classification metrics include positive predictive value, sensitivity, accuracy, specificity, and area under receiver operating characteristics. Subsequent sub-sections provide definitions and mathematical formulation of such metrics.

A Positive Predictive Value (PPV)

The proportion of appropriate instances amongst the retrieved instances is referred to as Positive predictive value (PPV) [56]. The same is also referred to as the precision and is defined as:

$$PPV = \frac{TP}{TP + FP} \quad (2.10)$$

Here, TP denotes the true positives and FP denotes the false positives, respectively. The positives that the prediction models precisely predict are known as the true positive tuples and the negative tuples are the ones which the prediction models predict falsely/incorrectly. The PPV values fall under the range [0 and 1].

B Sensitivity (S_n)

A parameter for evaluating the effectiveness of brain tumor prediction models is called Sensitivity. S_n is sometimes referred to as the rate of recognition [56]. It indicates the percentage of positive tuples that the prediction model efficaciously predicts.

$$S_n = \frac{TP}{TP + FN} \quad (2.11)$$

Here, FN is also referred to as False negatives and the FN describes the positive tuples that the prediction models predict incorrectly.

C Accuracy (Acc)

Accuracy is defined as the percentage of projected synergy scores that deviate from observed outcomes while staying in a permissible error range [56]. It is defined as:

$$Acc = \frac{(TP+TN)}{(TP+TN+FN+FP)} \times 100 \quad (2.12)$$

Here, TN which is also referred to as the True Negatives is the negative tuples correctly predicted by the prediction models.

D Specificity (S_p)

The specificity is described by the true negative rate. The percentage of negative tuples that the prediction models predicted is referred to by the same [57].

$$S_p = \frac{TN}{TN + FP} \quad (2.13)$$

E Area Under Receiver Operating Characteristics Curve (AUC-ROC)

The tradeoff between the true positive rate (TP_r) and the false positive rate (FP_r) is represented by Receiver Operating Characteristic (ROC) [57]. The x-axis and y-axis of the ROC curve signify the false positive rate and the true positive rate. In

order to compute model accuracy, the area under the ROC curve (AUC-ROC) is the responsible metric. The AUC-ROC value lies in the ranges of [0.5, 1].

$$TP_r = \frac{TP}{TP + FN} \quad (2.14)$$

$$FP_r = \frac{FP}{FP + TN} \quad (2.15)$$

2.8 METHODS

2.8.1 Machine Learning Techniques

Machine learning techniques, including supervised, unsupervised, and hybrid approaches, have been utilized to detect brain tumors, their classification, and segmentation. While conventional machine learning approaches and their variants keep on showing outcomes comparable to those of deep neural networks, recent research has focused primarily on deep learning methods because of their superior performance metrics.

The process of diagnosing brain tumors using machine learning techniques typically involves four main phases: pre-processing, segmentation, feature extraction, and categorization. A technique for segmenting and predicting brain tumors using adaptive SVM and modified region growing was presented by Reddy et al. [58]. To cover various methodologies such as SVM, kernel clustering, fuzzy clustering, thresholding, and Convolutional Neural Networks related to several medical image analysis, they summarized suggested procedures, pros, and cons for the eighteen associated researches initiated in the year 2001 to the year 2019.

A different model that used a random forest classifier and the histogram of oriented gradients (HoG) and discrete wavelet transform (DWT) for feature extraction achieved 98.37% accuracy [59]. Moreover, an ensemble learning framework processing gradient and Gabor wavelet features performed better than other approaches with overall Dice Score scores ranging from 0.86 to 0.88 [25].

Alam et al. [60] achieved 97.5% accuracy using a hybrid approach based on template-based K means and improved fuzzy C means techniques for the detection of brain tumors. A Simple Linear Iterative Clustering (SLIC)-based method utilizing Walsh Hadamard Transform (WHT) texture features for the segmentation of brain tumors was suggested by Angulakshmi et al. [61], greatly enhancing performance.

In order to achieve 97.56% accuracy, Devanathan et al. [62] proposed a framework that combined bilateral filtering, artificial bee colony-based thresholding, and SVM classification. The accuracy and performance metrics of other models that used different hybrid approaches also demonstrated notable improvements.

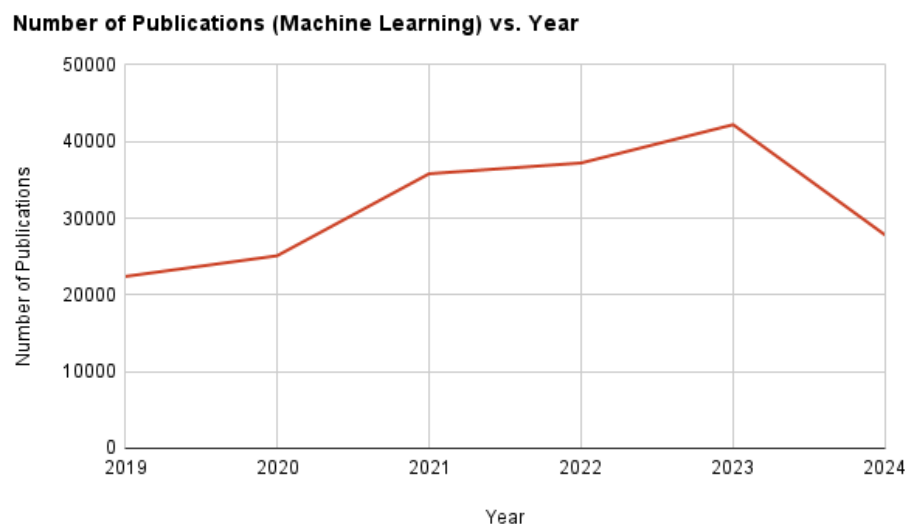


Fig.2.6. Last 5 years Publications on Machine Learning Based Brain Tumor Detection

K-means, K-Nearest Neighbor (KNN), SVM, Random Forest (RF), FCM, and artificial neural network (ANN) are among the machine learning models that show notable benefits for brain tumor analysis, despite their high computational complexity and reliance on large datasets. While clustering techniques work well for segmentation, supervised models are appropriate for detection and classification. ANN models perform well in brain tumor image analysis and are adaptable to all tasks. Fig.2.6 draws data from reputable publishers such as Scopus, SCI, etc., and presents

an overview of the growth in publications on machine learning-based brain tumor detection from 2019 to 2024. It draws attention to a growing body of research that uses sophisticated methods—like SVMs—and is frequently tested on datasets like BRATS. Important metrics are sensitivity and accuracy, and problems are computational complexity and data scarcity, which will push future research toward more complex algorithms and a wider range of datasets for better detection techniques.

2.8.2 Deep Learning Techniques

Deep learning is a powerful tool for analyzing brain tumors. U-Net, a specific deep learning framework, is particularly useful. It's fast and accurate because it combines shrinking and expanding sections with special filters (convolution layers) and data manipulation techniques. U-Net can find general and specific details in images, improving its results [63]. While the original U-Net paper didn't go into much detail about data, some variations, like path aggregation U-Net, have improved accuracy even further. However, these variations can take longer to run because they're more complex [64].

Building upon the success of U-Net for brain tumor segmentation, researchers have proposed exciting variations. Hybrid Two-Track U-Net (HTTU-Net) utilizes dual pathways and a combined error function for improved accuracy but suffers from slow execution times [65]. Triple intersecting U-Net (TIU-Net) tackles this by combining binary and multi-class segmentation U-Nets with a specialized loss function, achieving superior performance at a higher computational cost [66]. Alternatively, Sharif et al.'s [67] the method leverages pre-trained models for efficiency but faces challenges in generalizing to new datasets. These advancements highlight the ongoing effort to refine U-Net for brain tumor analysis, striving for the optimal balance between accuracy, efficiency, and generalizability.

Beyond U-Net variations, researchers are exploring entirely new deep-learning architectures. Sun et al. built a 3D FCN with unique features like multi-pathway extraction and random slice processing for improved brain tumor segmentation [68].

Chen et al. [69] introduced deep convolutional symmetric neural networks (DCSNNs), which use symmetric features to boost performance while staying efficient. Zhang et al. [70] proposed ResU-Net with attention to capturing details in both small and large tumors. These advancements showcase the dominance of deep learning, especially U-Net-inspired methods, in brain tumor analysis. While challenges like computational cost and data requirements persist, techniques like data augmentation and efficient architectures are paving the way for a stronger deep learning presence in medical image analysis. Fig. 2.7 draws data from reputable publishers such as Scopus, SCI, etc, and presents an overview of the growth in publications on deep learning-based brain tumor detection from 2019 to 2024.

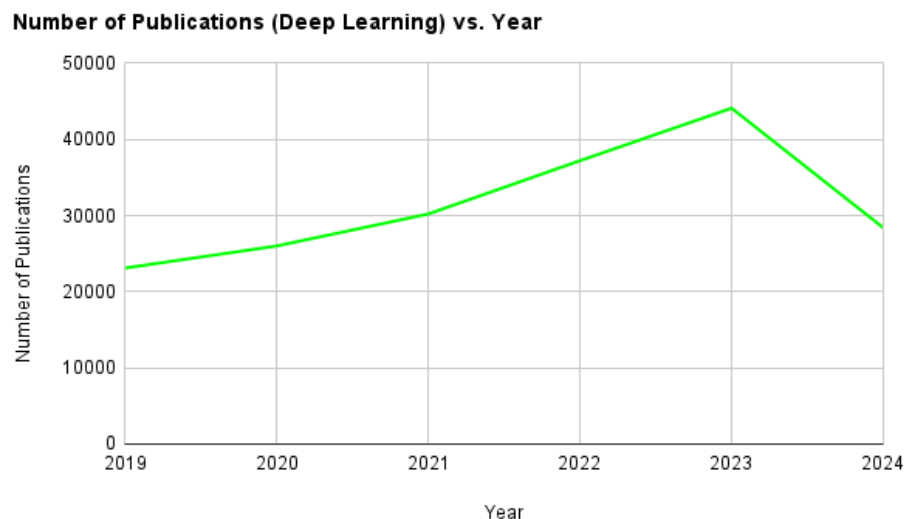


Fig.2.7. Last 5 years Publications on Deep Learning Based Brain Tumor Detection

Table 2.6 compares various techniques for enhancing and denoising brain MRI images. Each entry details a different study, including the publication (authors and year), datasets used, specific methods applied (e.g., filtering, transformations), and performance metrics (PSNR for noise reduction, SSIM for structural similarity). For instance, Awudong et al. [71] achieved a PSNR of 38.7532 and SSIM of 0.9781 using Variance-Stabilizing Transformation and Morphological Component Analysis, but suggest incorporating edge detection for further improvement. The table showcases the effectiveness of different techniques while acknowledging limitations like

computational cost and parameter tuning, emphasizing the ongoing advancements in MRI image processing.

Table 2.6 Image Enhancement and Denoising

Reference	Year	Dataset Used	Technique Used	Results	Advantage(s)	Limitation(s)
[71]	2024	Simulated brain database and two real clinical brain MRI datasets	Variance-stabilizing transformation (VST) with Morphological Component Analysis	PSNR: 38.7532 MSSIM: 0.9781	Preservation of complex texture details of MRIs, and provide effective denoising even when the noise distributions and intensities are unknown	Incorporating an edge detection technique or the multi-modal fusion technique into the proposed denoising process can yield high-quality clinical images with enhanced edges and textures
[72]	2023	The BrainWeb and OASIS datasets, simulated dataset	Regularized Neighborhood Pixel Similarity Wavelet Algorithm	PSNR: 46.80 SSIM: 0.9903	Features are also preserved, visually better improvement with the edges and textures being preserved	More computational complexity
[73]	2021	Both synthetic and clinical MRI datasets	Variational mode decomposition (VMD) with total variation (TV) smoothing	PSNR: 34.22 SSIM: 0.55	Preserves edge and texture details	The necessity of fine-tuning parameter estimation to achieve optimal results in T1-w and PD-w images
[74]	2021	Both synthetic and real MRI datasets	Conditional GANs	SSIM: 0.9489 Normalized MSE: 0.0167	Achieved high SSIM compared to other state of art techniques in high noise levels	Experiments are conducted with little data and lack comprehension.
[75]	2019	Synthetic dataset from Brain web and Private MRI dataset	Kolmogorov Smirnov Non-Local Mean (KSNLM)	MSE: Gaussian noise: 0.0034	Does not require the knowledge of its statistical description as algorithm is	More processing time

Continued on Page-42

Read Table 2.6 (continued)						
				Rician Noise: 0.0036 SSIM: Gaussian noise: 0.9309 Rician Noise: 0.9309	not limited to specific noise distribution	
[76]	2018	Three datasets from two public data sources of IXI dataset and Brainweb.	Multi-channel Denoising CNN (DnCNN)	PSNR: 46.32 SSIM: 0.99	Achieve visually good result, removed noise but also retain details of the image, also remove noise in the background sufficiently	Performed well for specific noise level only.
[77]	2017	Brain MRI image	Curvelet transform combined with Wiener filter	PSNR: 50.421 MSE :0.590 SSIM:0.996	Remove gaussian noise in low and high frequency sub bands	Remove only gaussian noise
[78]	2016	simulated MR images of brain	Median and Wiener filter in conjunction with the Non local means filtering technique	PSNR: 41.01 MSE: 5.18	Maintains the edges and fine structures of the image without causing significant blurring	Computation time increased
[79]	2015	Four images, two MRI images of the human brain, CT image of pectus of a hollowed chest, and CT image of the brain	Multistage median filtering (MMF)	PSNR: 36.0898 MSE: 15.9994 RMSE:3.9999	Improved denoising efficiency in comparison to wavelet transform	Blur the boundaries of the tissues

Table 2.7 dives into the world of brain MRI segmentation techniques, comparing their strengths and weaknesses. Each entry analyzes a different study, providing details on the publication (authors and year), the type of MRI data used (private dataset vs. public dataset like BRATS 2018), and the specific segmentation method employed (e.g., 3D CNN, cascade CNN). The table then highlights the performance using metrics like the Dice Similarity Coefficient (which quantifies overlapping between segmentation and ground truth), sensitivity (ability to detect true positives), and specificity (ability to detect true negatives).

For instance, Rajendran et al. [80] achieved a remarkable F-score of 99.40% for the whole tumor and 98.29% for the tumor core using a 3D CNN on their private dataset. However, their dependence on a limited dataset restricts the generalizability of their findings. Conversely, Ranjbarzadeh et al. [22] tackled overfitting by utilizing a cascade CNN approach on the public BRATS 2018 dataset. This method achieved Dice scores of 0.923 for the whole tumor and 0.9113 for the enhancing tumor, demonstrating its ability to extract various image features. However, it might struggle with segmenting very large tumors.

By showcasing these contrasting techniques, the table emphasizes the ongoing challenge of balancing computational complexity (e.g., 3D CNNs require more resources) with the need for substantial amounts of training data to ensure generalizability across different brain MRI scans.

Table 2.7 Image Segmentation

Reference	Year	Dataset Used	Technique Used	Results	Advantage(s)	Limitation(s)
[80]	2023	Private MRI dataset	3D CNN	F-Score: For the whole tumor: 99.40% For tumor core:98.29%, Sensitivity: For the whole tumor: 99.39% For tumorcore:98.25%	Produces much more precise and complete estimates	Small dataset

Continued on Page-44

Read Table 2.7 (continued)						
[81]	2023	Three publicly available dataset	Improved version of U-Net	Hausdorff distance (HD) metric 1.82, 1.31, and 2.30 for the three datasets	Reduction of network parameters and an improved ability to capture detailed information	Only capture 2D information
[82]	2022	BRATS 2015 dataset	Otsu's thresholding technique	TP: 68.7955% Accuracy: 95.56% for Level 4	Class 4 is ideal and appropriate for segmenting MRI scan	Computational cost could be high.
[83]	2021	BRATS 2015 dataset	Two-stage GAN	Dice score: 0.85	Requires less training data	Computationally complex
[22]	2021	BRATS 2018 dataset	Cascade CNN	Dice scores: 0.9203 (whole tumor), 0.9113 (enhancing tumor) and 0.8726 (tumor core)	Reduced computational time	Limitations when encountering large tumor volume
[84]	2021	BRATS 2018	3D residual neural network	Dice scores: 0.9121 (whole tumor), 0.818 (enhancing tumor) and 0.8662 (tumor core)	Better segmentation process as well as lowest computational complexity	High number of decoder parameters used. Decoder should be lightweight
[85]	2020	BRATS 2015 dataset and BRATS 2017 dataset	Encoder-decoder based network	Dice Similarity Coefficient (Whole): 94%	Requires less data for training	Redundancy in skip connection
[86]	2020	BRATS 2013	Superpixel Segmentation	Specificity: 0.957, Sensitivity: 0.840, PPV: 0.832, Jaccard: 0.751, Dice Score: 0.858	Quick computation time and strong resistance to rotational and scale-invariant changes	Sensitive to noise and distortions as well as extraction of redundant features

Continued on Page-45

Read Table 2.7 (continued)						
[87]	2019	BRATS 2013 dataset	CNN	Dice score: 0.86, Sensitivity: 0.86 and Specificity:0.91	Data preprocessing such as bias field correction and utilizes dropout regularize to deal with the over-fitting problem.	Not perform very well for LGG.
[88]	2018	BRATS-2017 dataset	cascaded anisotropic CNN	Dice score: For enhancing tumor: 0.7859, whole tumor: 0.9050, tumor core: 0.8378	Residual connections and multi-scale predictions are employed in these networks to boost the segmentation performance	Require long time for training and testing

Table 2.8 focuses on the effectiveness of various classification techniques used to analyze brain MRI images. It provides a breakdown of different studies, comparing their performance based on datasets, methodologies, and evaluation metrics. These metrics surrounds key measures such as overall accuracy, sensitivity (able to identify true positives), specificity (able to identify true negatives), and F1 Score (a balanced metric summarizing precision and recall).

A study by ZainEldin et al. [89] marks its exceptional performance. The reason for their exceptional mark was due to the combination of Convolutional Neural Network (CNN) with a specific optimization technique (Sine-Cosine Fitness Grey Wolf Optimization) on the BRaTS 2021 dataset. This strong approach led to the yield of near-perfect scores across all the metrics, making it to achieve 99.98% accuracy, sensitivity, specificity, and F1 score. However, the table adheres to some trade-offs related to the optimization steps significantly increasing the process time. Unlikely, some studies like Ayadi et al. [90] went for some simpler methods. Their technique employment includes Speeded Up Robust Features (DSURF) and Histogram of Oriented Gradients (HoG) in combination with the support vector machine (SVM) classifier. With a good accuracy of 90.27%, this process encompasses a very minimal

computation time. The table below encompasses the progress made in brain MRI accuracy with the adoption of deep learning and optimization techniques. However, it also emphasizes the ongoing challenges of balancing processing time and the generalizability of these models across various MRI datasets.

Table 2.8 Image classification

Reference	Year	Dataset Used	Technique Used	Results	Advantage(s)	Limitation(s)
[89]	2023	BRaTS 2021	CNN and Sine-Cosine Fitness Grey Wolf Optimization	Accuracy: 99.98% Sensitivity: 99.98% Specificity: 99.98% PPV: 99.98% NPV: 99.98% F1 score: 0.9998	Enhancement of the CNN's performance by the CNN optimization's hyperparameters to achieve higher accuracy	The extra optimization steps cause it to take long time, and challenges of generalization and data diversity emerged.
[91]	2023	Kaggle brain MRI dataset	Semi-Supervised Network	Accuracy: 94.3% Sensitivity: 92.5% Specificity: 96.5%	Improved brain MRI classification performance	Not validated on a bigger dataset, unable to reveal the location of the brain tumor
[90]	2022	Figshare	DSURF and HoG as feature extractors, SVM as classifier	Accuracy: 90.27%	Minimal computation time as well as can also assist physicians in improving their classification results for brain cancers.	Classification accuracy Need to be increased.
[92]	2022	BRATS 2018, 2019 and 2020	Kernel based SVM	Accuracy: 99.2% (BRATS 2018), 99.36% (BRATS 2019), 99.15% (BRATS 2020)	Higher classification accuracy for timely and proper diagnosis	Only two class classification is done

Continued on Page-47

Read Table 2.8 (continued)						
[93]	2021	4 publicly available datasets	CNN model for multi classification	Accuracy: 99.33% for two tumor classification, 92.66% for brain tumor type classification, 98.14% for brain tumor grade classification	CNN models can assist medical professionals in validating the initial multi class brain tumor assessment.	Accuracy reduced for brain type classification
[94]	2020	REMBR ANDT	CNN-based DL model	Accuracy: 100% recall: 100% Precision: 100% F-score: 100% and AUC: 1 ($p < 0.0001$) for two class classification	AI-based transfer learning outperforms traditional machine learning methods in brain tumor classification.	Incorporation of multiple MRI protocols needed
[95]	2020	BraTS 2018 dataset	3D CNN	Accuracy: 96.49%	Deep CNN approach is used for accurate glioma tumor classification, learning local as well as global contextual information with reduced weights	In order to achieve improved performance, capsule networks (CapsNet) were not utilized in this method for MRI brain tumor classification.
[96]	2020	BraTs2015, BraTs2017, BraTs2018	Transfer learning-based approach with ELM based feature selection	Accuracy (98.16%, 97.26%, and 93.40%, for the BraTs2015, BraTs2017, and BraTs2018 datasets	Extraction of features enhanced classification accuracy and decreased processing time	More classification accuracy is required.
[97]	2020	Private dataset containing 204 brain MR images	MLP with improved whale optimization	Accuracy: 96.5%	Improved version of whale optimization algorithm is used to design an optimized system	Small dataset

Continued on Page-48

Read Table 2.8 (continued)						
[98]	2019	Figshare	CNN technique for a three-class classification	Accuracy: 97.1% recall: 98.9% Precision: 98% and AUC: 0.99	Transfer learning approaches are extremely effective strategies in situations where medical images are inadequate.	Samples falling under the meningioma category were incorrectly categorized, overfitting occur due to less training data.

2.9 SUMMARY

This chapter dives deep into the wider segments of the brain MRI analysis, meticulously giving rich learning in segmentation and classification techniques. It gives highlight regarding the strengths and weaknesses of the various approaches in order to give good help to researchers in this critical and crucial field. The dissection of the prominent techniques given in the review like 3D Convolutional Neural Networks (CNN) that encompasses the capability to learn intricate spatial relationships within MRI data. It also gives a good dive into exploring the cascade CNNs which is a method that tackles overfitting by employing a multi-stage learning process. Moreover, the chapter sheds light upon optimization techniques like Sine-Cosine Fitness Grey Wolf Optimization, which can promote the boundaries of performance by fine-tuning the hyperparameters of deep learning models.

This chapter also encompasses the near-perfect scores in metrics like accuracy, sensitivity (correctly identifying true positives), and specificity (correctly identifying true negatives). These advancements give a promising aspect for the early and accurate detection of the brain tumor. This review also encompasses the hurdles that need to be addressed, reviewed, and henceforth need to be overcome such as the computational complexity in the 3D CNNs, that require good processing power. This chapter also gives highlight regarding the risks of overfitting when being used with the limited datasets to train the models. The need for substantial training data is also encompassed by this review in order to ensure generalizability across diverse brain MRI scans.

This chapter further paves the way for innovation in brain tumor detection using MRI analysis by emphasizing aspects such as striking a balance, navigation through the trade-off between computational resources needed for complex techniques, and the further ability to achieve generalizable results across different datasets.

2.10 RESEARCH GAPS

Brain tumors are considered a major global health concern, requirement for the development of efficient screening techniques and alternative therapies is essential to enhancing patient outcomes. Identification of any gaps or restrictions in the literature on DL methods for brain tumor detection is quite a critical aspect. addressing computational complexity, guaranteeing the generalizability of models across various populations, and requiring larger datasets to improve model training are the possible gaps that need to be covered. Therefore, the obstacles such as overfitting and the need for significant computational resources need to be overcome.

- Need for a technique that effectively integrates both local contextual information on disease appearance and the global anatomical context in medical imaging, particularly for small and diverse datasets.
- Removing noise from medical images and stripping the skull region are challenging tasks. So, an efficient noise removal method that enhances medical image quality without losing critical diagnostic details is needed.
- Need an approach that balance reduced computational time with improved discrimination accuracy, especially for complex medical image analysis.
- Need an approach that increases segmentation accuracy and effective visualization capability to enhance diagnostic precision.

As inspiration, we need to develop some novel strategies for the analysis of various brain abnormalities, especially brain tumors in light of the difficulties posed by the traditional approaches.

2.11 RESEARCH OBJECTIVES

Based on research gaps findings following are the research objectives: -

- 1) To develop an approach which includes both local contextual information on disease appearance and global contextual information on disease location for accurate classification.
- 2) To examine the enhancement of medical images and to find an efficient approach for noise removal.
- 3) To develop an approach that results in less computation time as well as better discrimination capability.
- 4) To develop an approach for improved diagnosis of diseases by amalgamation of segmentation and rendering techniques.

CHAPTER 3

DENOISING OF MAGNETIC RESONANCE IMAGES

3.1 PREAMBLE

The field of medical diagnosis and treatment has changed greatly with advanced imaging technologies. MRI is notable for being non-invasive and producing clear images of soft tissues. However, the noise in MRI scans can be a big problem. This noise can make images worse, lead to wrong diagnoses, and put patient care at risk. Accurate diagnosis and treatment planning are thus significantly affected by this inherent noise.

This chapter presents two related studies on denoising MRI images and essential post-processing steps for accurate diagnosis. The first study examines various denoising methods, focusing on Rician noise in MRI images. It offers a detailed comparison between a modern deep learning-based technique and traditional denoising methods such as median, Gaussian, average, bilateral, non-local means, and wavelet-based filters. The work exposes the shortcomings of conventional techniques in terms of retaining significant image details and shows how well the suggested deep learning model performs in terms of efficiently lowering noise without compromising image integrity. The image quality of MRI is enhanced as well as the accuracy of later diagnostic procedures by the deep learning-based denoising technique.

The second study focuses on the particular use of denoised MRI images in the diagnosis of brain tumors, building on the groundwork established by the first study. Whether benign or malignant, brain tumors need to be precisely imaged in order to be diagnosed, segmented, and classified. The study presents the Modified Triplet Cross Fusion Learning (MTCL-MRI) scheme for denoising MRI images. The MTCL-MRI scheme is then used, U-Net segmentation of brain tumors, and ResNet to classify

various categories of brain tumors, including glioblastoma, meningioma, and pituitary adenoma. The study evaluates the proposed MTCL-MRI algorithm using the BraTS 2020 dataset. The results illustrate that the MTCL-MRI scheme can provide comprehensive local and global features for improved brain tumor diagnosis. It gives improved diagnostic accuracy for brain tumor classification compared to other cutting-edge techniques. Achieve higher classification accuracy and precision when classifying the various categories of brain tumors in the dataset.

The study also discusses any limitations or challenges encountered, such as the potential impact of MRI image quality or the dataset size on the MTCL-MRI scheme's performance.

The collective knowledge gained from these investigations advances medical imaging technologies, improving patient care outcomes and diagnostic capacity. By making substantial contributions in the medical imaging field, this chapter hopes to direct future study and advancement in this crucial area of healthcare.

3.2 DENOISING MAGNETIC RESONANCE IMAGES WITH RICIAN NOISE BASED ON DEEP LEARNING

MRI images with noise are reduced while maintaining their key characteristics through the use of denoising techniques. Reducing noise in MRI scans effectively not only increases their visual quality but also improves the precision of subsequent diagnostic procedures. Numerous denoising algorithms have been created, from more modern DL-based techniques to more conventional linear and nonlinear filters. Conventional filters, like Gaussian, median, and average filters, frequently have trouble achieving a balance between noise reduction and the retention of important image details. Although nonlinear approaches such as wavelet-based techniques, bilateral filtering, and anisotropic diffusion offer improvements, they are not without limitations, especially when dealing with the complex nature of Rician noise. Although they show great potential, CNN-based denoising methods have not been thoroughly investigated in the context of MRI images, so substantial progress remains to be made.

This work objectives to close this gap by providing an in-depth analysis and comparison of DL-based and conventional denoising approaches for MRI images. Specifically, the work focuses on how well these techniques remove Rician noise while retaining the essential information required for a precise diagnosis. This work determines the best method for denoising MRI images by comparing a number of filters, such as wavelet-based filters, median, Gaussian, average, bilateral, non-local means, and a suggested deep learning model. Fig. 3.1. shows the architecture of the proposed methodology.

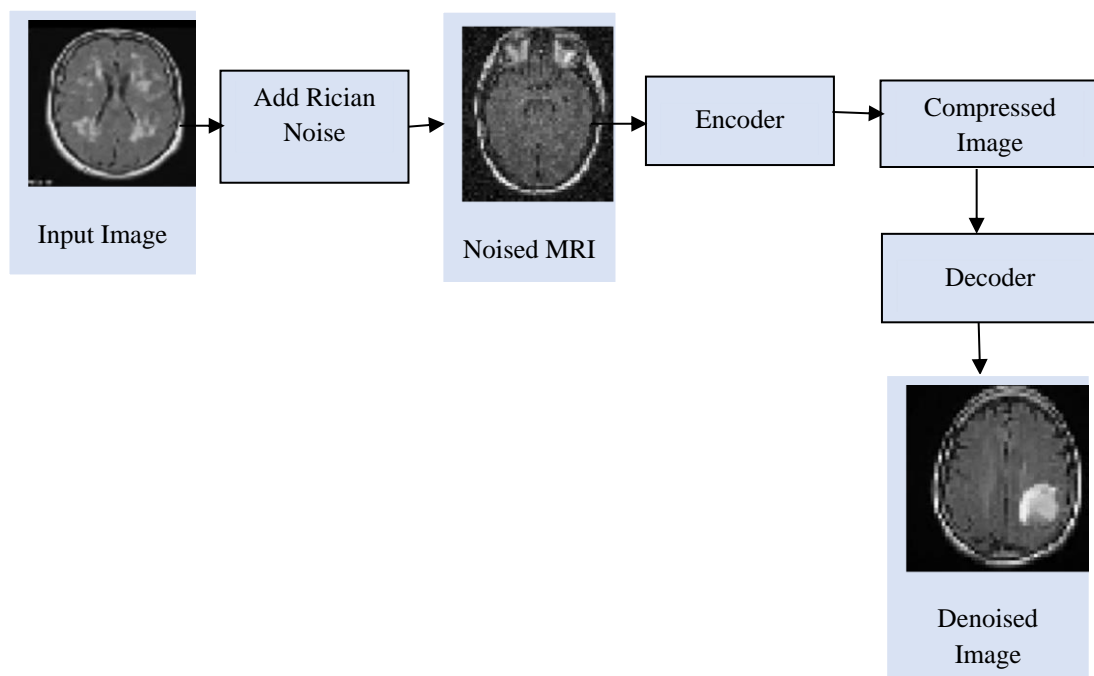


Fig.3.1. Architecture of Proposed Methodology

3.2.1 Methodology

A comparison between DL techniques and conventional denoising techniques is given in this section.

a. Median Filter

Median filters are the spatial filter that follows the sliding window technique in which the window's central value is substituted with its median value. Although the

median filter removes noise, it also removes precise details [99]. The median filter has a number of advantages compared to linear filters, including the ability to reduce the impact of exceptionally large noise levels and providing edge preservation with less blurring for different kinds of noises. In the presence of noise, though, it significantly blurs image edges.

b. Gaussian Filter

By smoothing/blurring the image with Gaussian filtration, random noise in the image can be decreased. Gaussian blurring is similar to convolving an image and with a Gaussian function [100]. The capability of Gaussian filters to preserve edge is superior than that of average filters, but they lose the fine details of the image.

c. Average Filter

Average filters are the spatial filter that follows the sliding window technique in which the window's central value is substituted with its mean/average value. The average filter is disadvantageous as its edge preservation capability is very bad and it results in the blurring of images [101].

d. Bilateral Filter

A bilateral filter is described by the weighted average of neighboring pixels, which is equivalent to its Gaussian convolution. The applicability of bilateral filters has quickly evolved. It is employed in various applications including for noise removal and enhancement. This filtering technique is used for smoothing of images while keeping their edges intact. The bilateral filter looks at the change in adjacent variables to preserve the edge during smoothing [102].

e. Nonlocal Means

By averaging the identical pixels in an image dependent on intensity distance for retrieving the value of each pixel, the Nonlocal Means (NLM) approach [103] uses redundancies in the image and eliminates noise from the images. The weighted mean of complete pixels in MRI scans is used to estimate every specific pixel value.

f. Wavelet-based Denoising

The Wavelet domain is beneficial since the discrete wavelet transform tries to concentrate the desired signal energy into a smaller number of coefficients; thus, the discrete wavelet transform of noisy image limits the number of high SNR coefficients (that are to be kept) and a substantial amount of low SNR coefficients (that are to be discarded) [15]. The image is then restored after the removal of noise components by utilizing the method of inverse discrete wavelet transform.

3.2.2 Performance Metrics

The effectiveness of various image denoising methods can be assessed in two forms i.e., objectively and subjectively. The subjective approach involves visually comparing the original image with the denoised image with the naked human eye. In the objective approach, different metrics are used that measure denoising effectiveness. Various performance measures, including Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), and Root Mean Square Error (RMSE), are often utilized to evaluate the effectiveness of different filters.

Distinct visual quality assessments for MRI scans using different evaluation metrics are addressed in detail [103]. MSE is determined by Eqn. 3.1 as

$$MSE = \frac{1}{L * W} \sum_{x=0}^{L-1} \sum_{y=0}^{W-1} (I_{xy} - \widehat{I}_{xy})^2 \quad (3.1)$$

where, L and W define the image's length and width, respectively, I_{xy} signifies the original image's pixel value, and \widehat{I}_{xy} signifies the denoised image's pixel value. RMSE is determined by Eqn. 3.2 as

$$RMSE = \sqrt{MSE} \quad (3.2)$$

PSNR is determined by the Eqn. 3.3 as

$$PSNR = 10 \log_{10} \left(\frac{(Max - 1)^2}{MSE} \right) = 20 \log_{10} \left(\frac{Max - 1}{RMSE} \right) \quad (3.3)$$

where Max signifies the maximum intensity level of an image which is possible. E.g., In 8-bit Image, Max= 28 i.e., 256 levels. Less values of MSE means, better PSNR.

3.2.3 Results

The proposed DL-based approach is compared to traditional denoising methods such as median filter, gaussian filter, average filter, nonlocal means filter, bilateral filter, and wavelet smoothed filters in a comparative study. The data collected for this research is composed of 646 T1 weighted MRI scans of the brain that were obtained from The Cancer Imaging Archive (TCIA) Publicly Accessible repository [41]. The images belonged to patients who had been diagnosed with glioma. The original input and the noisy MRI images with 5% contamination and 10% contamination are displayed in Fig. 3.2.

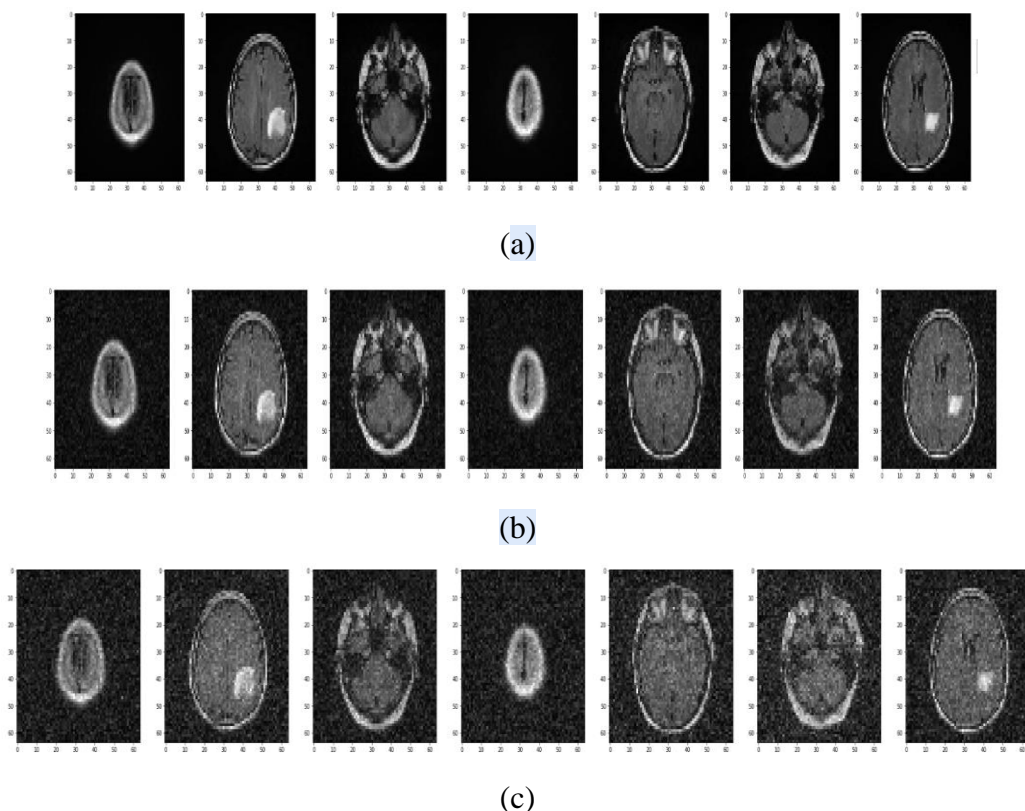


Fig.3.2. (a) Input Image (b) Noised Image (Noise Level=5%) (c) Noised Image (Noise Level=10%)

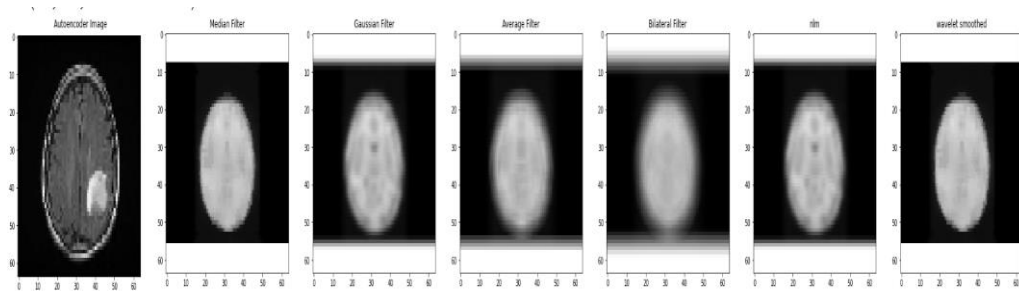


Fig.3.3. Images after denoising using different techniques

Fig.3.3 shows the output denoised images obtained using different techniques explained in this work. The images are contaminated by 5% and 10% Rician noise. However, noise is effectively reduced as we can see that the white spots that were intentionally added to the input images have vanished from the restored images. The evaluation of several denoising approaches on noisy images with 5% contamination for various metrics is summarized in Table 3.1. When compared with other denoising algorithms, the findings clearly show that the DL-based technique gives improved performance across all parameters.

Table 3.1 Comparative results of various denoising algorithms used in the paper for 5% noise level

Technique Used/ Performance Measures	MSE	RMSE	PSNR
Median Filter	0.2862	0.5350	53.5637
Gaussian Filter	0.2653	0.5151	53.8932
Average Filter	0.2553	0.5053	54.0606
Bilateral Filter	0.2195	0.4685	54.7166
NLM	0.2424	0.4923	54.2852
Wavelet-Based	0.1909	0.4369	55.3231
Proposed DNN-Based Technique	0.1019	0.3192	58.0491

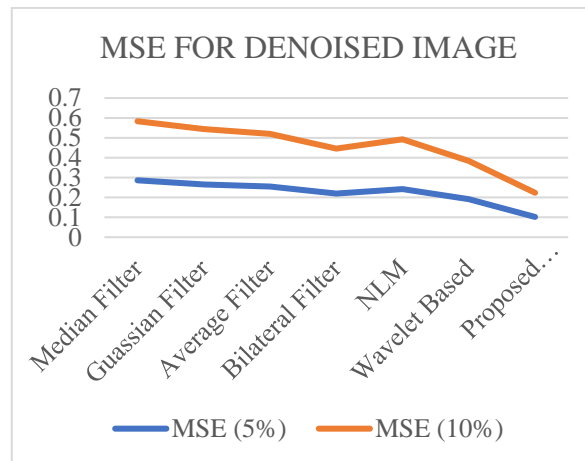
The evaluation of several denoising approaches on noisy images with 10% contamination for various metrics is summarized in Table 3.2. Table 3.2 also shows that the DL-based method gives better performance in terms of all parameters than the other traditional techniques. The acquired result indicated that visuals were blurred

and lost information when linear filtering methods were used. The NLM and bilateral filters benefit in terms of edge preservation. The results indicated that the DL-based technique is used to get better results by decreasing noise to the highest extent possible, as evaluated by PSNR. By considering the results of all performance measures, we found that our deep learning-based model outperforms other traditional denoising approaches.

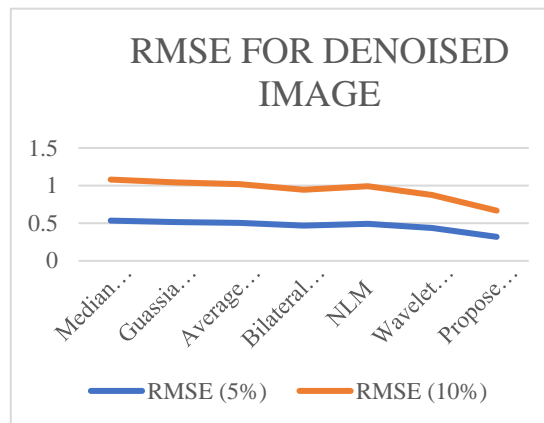
Table 3.2 Comparative results of various denoising algorithms used in the paper for 10% noise level

Technique Used/ Performance Measures	MSE	RMSE	PSNR
Median Filter	0.2971	0.5451	53.4013
Gaussian Filter	0.2789	0.5281	53.6765
Average Filter	0.2647	0.5145	53.9031
Bilateral Filter	0.2265	0.4759	54.5804
NLM	0.2498	0.4998	54.1549
Wavelet-Based	0.1927	0.4390	55.2815
Proposed DNN-Based Technique	0.1216	0.3487	57.2818

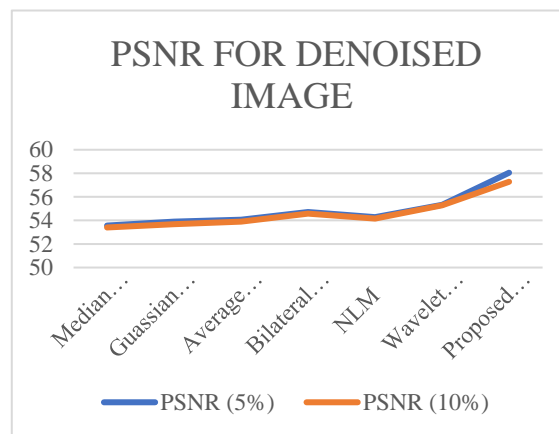
Fig. 3.4. graphically represents the comparison between denoised images using different techniques used in this paper in terms of MSE, RMSE, and PSNR parameters for 5% and 10% noise levels. These graphs also show that utilizing our proposed technique, the benefits are quite significant for both noise levels. Table 3.3 discusses the performance comparison of Rician noise-corrupted brain MRI denoising techniques.



(a)



(b)



(c)

Fig.3.4. (a) MSE comparison between denoised images using different techniques (b) RMSE comparison between denoised images using different techniques (c) PSNR comparison between denoised images using different techniques

Table 3.3 Performance comparison of Rician noise corrupted brain MRI denoising techniques

Technique Used	Technique Used		MSE	RMSE	PSNR
[104]	BT-Autonet (Autoencoder based)		41.684	-	29.64
[105]	Bm3D		-	-	36.11
	NLM method with non-				37.17 (for
[106]	subsampled shearlet transform (NSST)		-	-	5% noise level)
[107]	Anisotropic Weighted Kolmogorov-Smirnov (KS)-NLM Filter		-	-	38.52(5% noise), 34.18(10% noise)
[108]	Spatially adaptive nonlocal means (SANLM) filter		3.2141	1.7928	43.0942
[72]	Regularized Neighborhood Pixel Similarity Wavelet algorithm (PixSimWave)		-	-	46.62 (for 5% noise level)
Proposed Technique	Proposed DNN-Based Technique	5%	0.1019	0.3192	58.0491
		10%	0.1216	0.3487	57.2818

The results shown that the proposed method demonstrates a remarkable enhancement in denoising performance, outperforming other state of art techniques across all evaluated metrics. Comparative results highlighted that the proposed model achieve significantly higher PSNR values, leading to markedly lower value of MSE and RMSE, thereby ensuring superior noise reduction and better preservation of image details.

3.2.4 Discussion

This work presents a comparative analysis demonstrating the notable progress made in DL-based denoising techniques for MRI images, especially those tainted with Rician noise. While median, Gaussian, and average filters—three classic denoising

techniques—have been invaluable in the early phases of image processing, they clearly have limits when it comes to striking a balance between minimizing noise and preserving important image details. For example, median filters tend to blur the edges and reduce fine details even though they are good at reducing noise. Even though they are better at maintaining edges, Gaussian filters nevertheless cause the loss of subtle image details. Because of their oversimplified methodology, average filters have poor edge preservation performance, producing excessively smooth and low-detail images.

On the other hand, more advanced methods of noise reduction are provided by nonlinear techniques like wavelet-based denoising, bilateral filtering, and non-local means (NLM). Bilateral filters smooth images while preserving edges because they take into account both spatial proximity and pixel value variations. Unlike simple averaging techniques, the NLM approach preserves fine structural details by utilizing the redundancy of similar patches within the image and averaging them to reduce noise. By utilizing the discrete wavelet transform, wavelet-based denoising effectively separates noise from the signal while maintaining important image features and offering significant improvements over linear filters. But even with these improvements, these techniques are still unable to effectively handle the complexity and variability that come with Rician noise, especially when dealing with the extremely complex context of MRI images.

The suggested deep learning-based method offers a significant improvement in denoising performance. This method surpasses the modern approaches in all assessed metrics, such as RMSE, MSE, and PSNR, credit to the usage of Convolutional Neural Networks (CNNs). Comparative outputs exhibit DL based models achieving higher PSNR values thereby leading to a significant reduction in MSE and RMSE, showcasing superior noise reduction and detailed preservation. Compared to other approaches, the DL model is able to learn and generalize features more efficiently because of its architecture, which is specifically designed for the unique properties of MRI data and Rician noise. As a result, it offers sharper, more accurate images, which are necessary for accurate medical diagnosis. These conclusions highlight deep learning's revolutionary potential in medical image

processing and open the door to more accurate and thorough diagnostic instruments in clinical settings.

3.3 MRI IMAGE DENOISING, SEGMENTATION, AND CLASSIFICATION PIPELINE: LEVERAGING ENHANCED MTCL, U-NET SEGMENTATION, AND RESNET CLASSIFICATION

The strategic framework of this study is illustrated in Fig. 3.5, detailing the workflow that begins with the acquisition of MRI images, known as the image acquisition phase. This phase employs datasets containing human brain MRI scans. Subsequently, the images go through a number of pre-processing and post-processing stages, each involving a range of specialized procedures.

The methodology is outlined as follows:

- i. Noise elimination from MRI images through deep learning techniques: To start, any unwanted noise in the MRI images is eliminated using the MTCL-MRI (modified triplet cross-fusion learning [109]) method for unpaired MRI image denoising. This process is designed to enhance the clarity and quality of the images by reducing undesired artifacts or distortions. It does this by utilizing a noisy image, a high-quality MRI image, and an additional auxiliary image derived from another MRI sequence or modality to produce a more precise denoised result. The combined features are then inputted into a denoising network, which extracts the denoised features used to reconstruct the final denoised MRI image.
- ii. Statistical feature extraction using Segmentation (Local and Global Feature Extraction): After noise removal, extraction of statistical features takes place from the images by utilizing local and global features. This involves analyzing the spatial relationships of pixel intensity values within the images to derive valuable statistical information.
- iii. Classification of tumors into benign and malignant categories using ResNet based on the extracted features: Upon acquiring the statistical features, the ResNet-50 algorithm [110] is used for classification. This method utilizes the

features to distinguish between benign and malignant brain tumors, aiding in medical diagnosis and treatment planning.

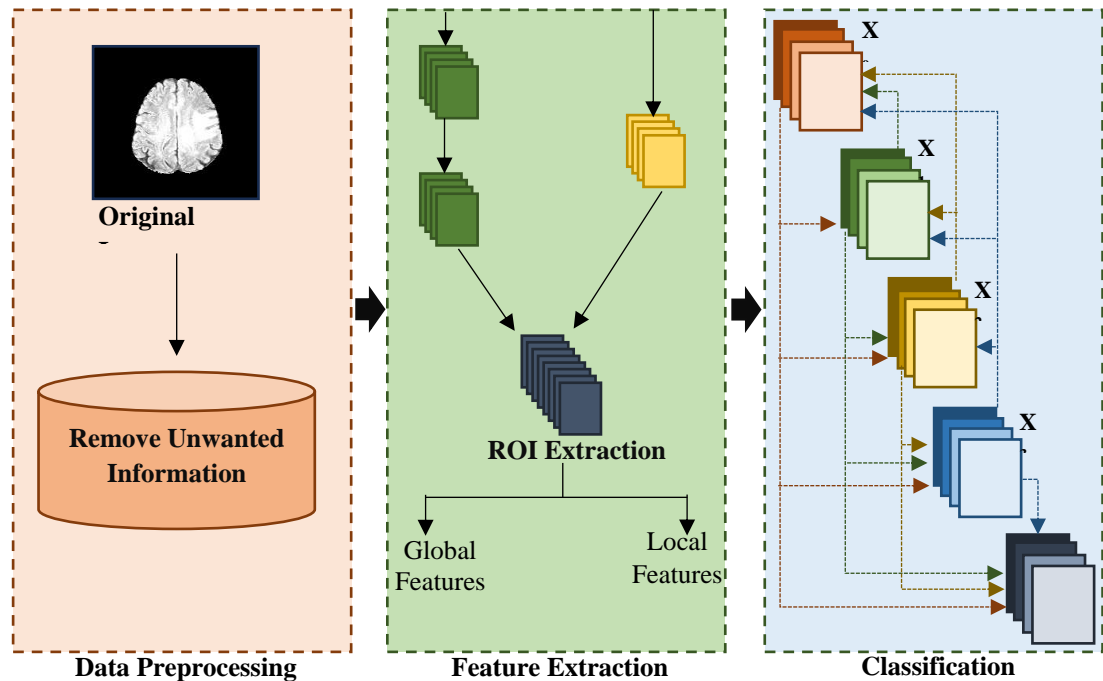


Fig.3.5. Outlined Strategy of The Proposed Work

3.3.1 Acquisition

The first step in the suggested approach entails obtaining an MRI image, which is used as the input for additional processing. In this method, we use brain tumor MRI images. This step is vital because it sets the groundwork for all subsequent analysis and processing. MRI images are chosen because they can provide detailed and high-resolution views of internal structures, making them well-suited for medical applications such as the detection and analysis of brain tumors.

3.3.2 Dataset

The research has utilized the BraTS 2020 dataset, which is specifically designed for brain MRI images and emphasizes the segmentation of brain tumors

across multiple MRI scan types. The BraTS dataset comprises images in MRI. NIfTI files format, each with dimensions of 240×240 pixels, encompassing samples from both male and female subjects. The BraTS dataset is widely recognized and utilized in the area of medical imaging, predominantly for tasks related to brain tumor analysis and segmentation. Its comprehensive collection of MRI images allows researchers to study and develop algorithms for accurately identifying and delineating brain tumors across various MRI modalities, providing insightful information for medical diagnosis and treatment planning [38].

3.3.3 Preprocessing

During the preprocessing stage of this study, the first step involves image enhancement. Specifically, a contrast stretching filter is applied to improve the quality of MRI images, particularly those in grayscale format, as it has shown relatively superior performance. Unlike histogram equalization, which can sometimes create an unnatural appearance by altering the relative intensities of grey levels, contrast stretching improves image contrast without distorting these relative intensities. Contrast stretching works by extending the image's range of contrast values to cover the full scale from 0 to 1. This process effectively enhances the differences between light and dark areas in the image, resulting in a clearer and more detailed representation. It removes uncertainties that may be present in certain regions of the image, leading to a more visually appealing and diagnostically valuable image [111].

Fig.3.6 in the study showcases the qualitative comparison between original, noisy, and denoised images, demonstrating how this technique enhances the overall quality of MRI images without introducing unnatural artifacts or altering the essential grey level information. This step is vital as it prepares the images for additional analysis and feature extraction, ensuring that subsequent processing steps can yield accurate and meaningful results in tasks such as tumor classification and medical diagnosis.

In image processing, a filtering operation was used to enhance sharpness, smoothness, and edge definition. The anisotropic filter [112] was chosen for this purpose over other filters like the Gaussian, median, or mean filters. This decision was made because while a mean filter is better at reducing noise, it doesn't preserve edges effectively.

The primary goal of image scanning is to eliminate digital noise, which can greatly affect the accuracy of the image parameters in Table 3.3. Noise can reduce image quality and make processing more difficult. To address this, various noise reduction techniques are used in image processing. In this study, the anisotropic filter was specifically used for denoising. The implementation uses the generalized anisotropic diffusion equation to describe the image diffusion procedure and enhance the quality of scanned images in noisy environments.

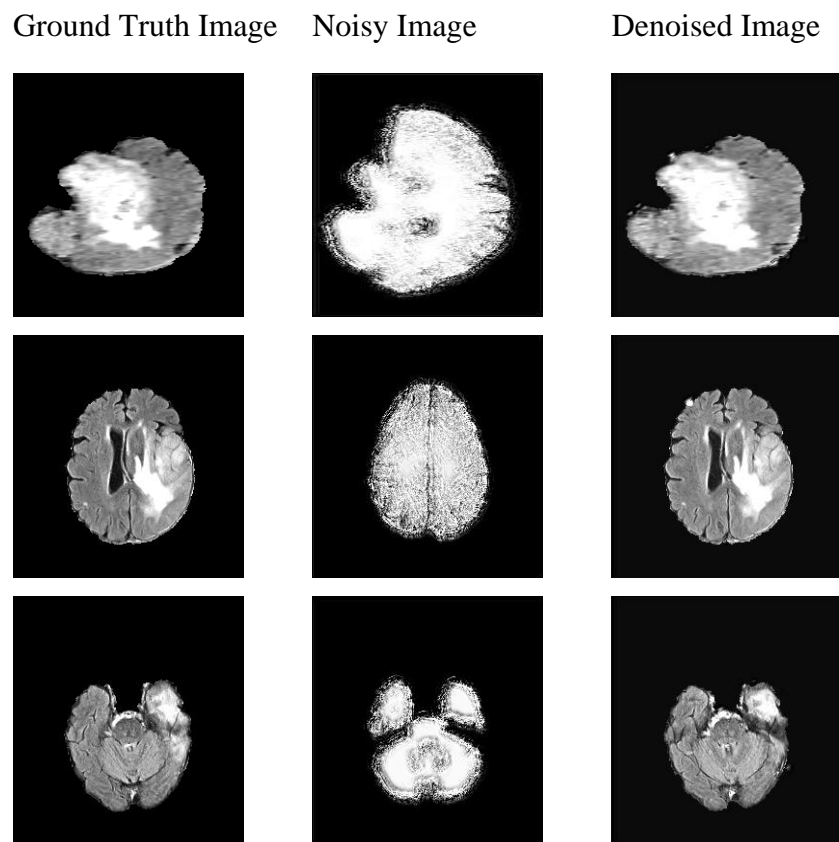


Fig.3.6. Denoising of MRI images using MTCL

The generalized anisotropic diffusion equation is employed to describe and model the image diffusion process. This equation is utilized in the following manner:

$$\frac{x_a}{x_b} = \text{div}(v(k, l, m)C_1) = C_v C_a + v(k, l, m)C_a \quad (3.4)$$

where, symbol C_a represents the gradients of an image, while $v(k, l, m)$ is utilized to denote the coefficient of diffusion.

Table 3.3 Qualitative evaluation of the proposed noise removal scheme

Noisy images				Denoised Images			
Slice No.	RMSE	PSNR	SSIM	Slice No.	RMSE	PSNR	SSIM
Slice 89	10.3544	16.1162	0.2339	Slice 89	10.4520	24.7876	0.3225
Slice 113	10.2743	16.1897	0.2294	Slice 113	10.4560	24.7836	0.3245
Slice 320	10.4627	16.0524	0.2351	Slice 320	10.4535	24.7861	0.3230

An assessment of the forward and backward changes, devoid of specific dimensions, is represented by the following terms:

$$y(l_{p,q}^m - l_{p,q}^m) = \frac{v_{p,q}^m + v_{p,q}^m}{2} \quad (3.5)$$

This statement describes the neighborhood of a pixel positioned at coordinates p, q in an image. Specifically, the neighborhood includes the four adjacent pixels surrounding the origin pixel, which are denoted as $N4 = \{(p - 1, q), (p + 1, q), (p, q - 1), (p, q + 1)\}$.

The notation suggests that there is a strong diffusion effect on noise pixels and a weaker diffusion effect on signal pixels. This approach aims to eliminate noise while preserving the signal within the image. Throughout each iteration and the optimization process of the image, various diffusion designs are applied to maintain a consistent phase scale. Segmentation is applied to divide a filtered image into distinct regions and to extract local and global features based on similar characteristics or attributes.

3.3.4 Feature Extraction

The Global and Local Feature Reconstruction Network proposed in the study includes a segmentation component. The global feature extraction module operates on each level feature map through three main steps:

1. It generates global descriptors for each class to represent the features [113].
2. The current level global descriptors are combined with the global descriptors from the higher level in order to make a cross-level global descriptor.
3. It predicts rebuilding the weight for every location, thereby utilizing the cross-level global descriptors to rebuild every pixel. This permits every pixel to incorporate features from its own level as well as permits them to derive the features from different levels.

Generates Global Descriptors. $\beta_{lev} \in \mathbb{R}^{ch_n \times HW}$ is a flattened feature map where lev highlights the level of this feature map. This feature map has ch_n channels, with H and W denoting its spatial dimensions. β_{lev} is fed to a convolutional layer to produce an attention map $AT_{lev} \in \mathbb{R}^{ch_k \times HW}$ with dimensions ch_k representing the number of global descriptors and an embedding feature $BT_{lev} \in \mathbb{R}^{ch_h \times HW}$ with dimensions ch_h representing the dimension of the embedding feature. In the proposed work, ch_h is set to 64 thereby reducing the size of weights and computational cost, considering that feature maps at different levels may have varying channel dimensions.

The process of generating global descriptors involves:

1. The flattened feature map $\beta_{lev} \in \mathbb{R}^{ch_n \times HW}$ when fed into a convolutional layer, it tends to produce an attention map $AT_{lev} \in \mathbb{R}^{ch_k \times HW}$ and an embedding feature $BT_{lev} \in \mathbb{R}^{ch_h \times HW}$.
2. Using the attention map $AT_{lev} \in \mathbb{R}^{ch_k \times HW}$ to compute global descriptors representing the spatial attention across the feature map.
3. Employing the embedding feature $BT_{lev} \in \mathbb{R}^{ch_h \times HW}$ to generate global descriptors that capture the essential features of the input feature map at that level.

This approach allows us to extract meaningful global descriptors g from the feature map at each level, facilitating the subsequent processing steps for feature aggregation and reconstruction.

$$g_{lev} = [g_{lev}^1, g_{lev}^2, \dots, g_{lev}^{ch_k}] = \varphi(AT_{lev})BT_{lev}^\top \in \mathbb{R}^{ch_k \times HW} \quad (3.6)$$

Here, φ refers to the process of applying SoftMax normalization across the spatial dimensions.

Global Descriptors Cross-Level. For the feature map at the lev^{th} level, utilizing the global descriptors g_{lev} obtained on this level for feature rebuilding, each pixel at each location tends to be interconnected through g_{lev} . Moreover, merging all the global descriptors from the higher levels with those from this level resulted in the creation of a cross-level global descriptor, which can be expressed in the form given below in Eqn. 3.7:

$$g'_{lev} = con(g_{lev}, g_{lev+1}, \dots, g_{lev+z}) \in \mathbb{R}^{(z+1)ch_k \times ch_h} \quad (3.7)$$

The global descriptors $g_{lev+1}, \dots, g_{lev+z}$ are derived from the $(lev+1)^{th}$ to $(lev+z)^{th}$ level feature maps. This process ensures that not only are pixels at the same level associated with each other through their corresponding global descriptors, but they also capture features from higher-level global descriptors. This integration of cross-level information enriches the representation of each pixel, contributing to a more comprehensive and context-aware feature reconstruction process.

Reconstruction of Global Features. The subsequent step involves utilizing the hybridized global descriptors g'_{lev} to rebuild the characteristics of every pixel location. 1×1 convolution operation is used to predict the global reconstruction weights $V_{lev} \in \mathbb{R}^{(z+1)ch_k \times HW}$ is based on the current feature map β_{lev} . Here, z denotes the total number of global descriptors connections produced by higher-level feature maps, as illustrated in Fig. 3.7. The SoftMax function is applied along the channel dimension for the normalization of the reconstruction weights, which improves the selection of global descriptors. The method of reconstructing the feature map β_{lev} can be expressed in Eqn. 3.8 as follows:

$$\tilde{\beta}_{lev} = g'_{lev} \text{softmax}(V_{lev}) \quad (3.8)$$

The reconstructed feature map is denoted as $\tilde{\beta}_{lev} \in \mathbb{R}^{ch_k \times HW}$.

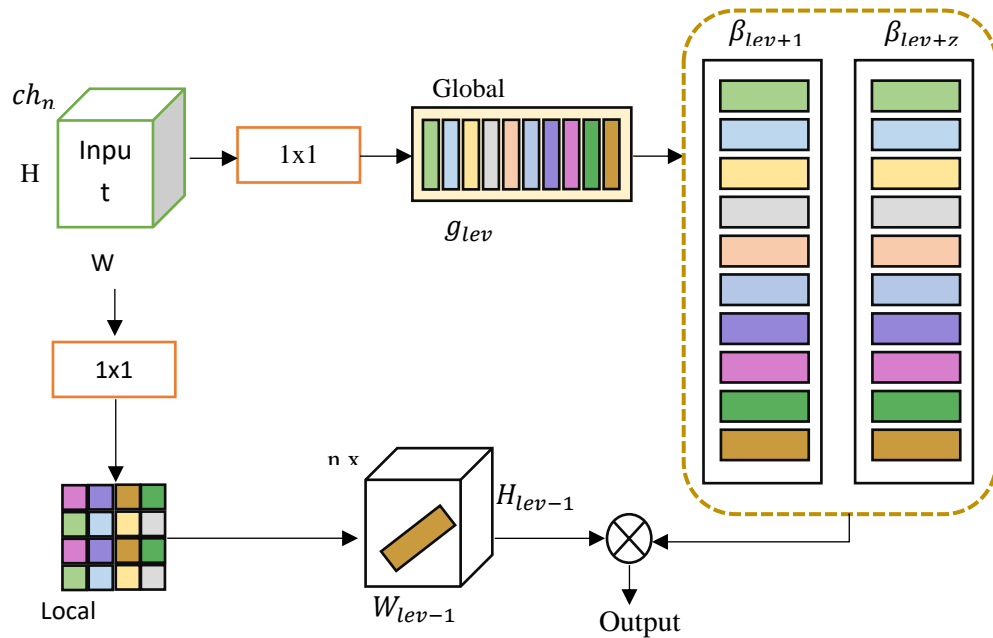


Fig.3.7. Proposed Local-Global Feature Extraction Model

Local Feature Reconstruction. The Local Features Reconstruction module is suggested to facilitate the transfer of spatial details from lower-level features to upper-level features. This module utilizes lower-level features to lead the reconstruction of local features in upper-level features, thereby enabling semantic alignment as well as feature-up sampling.

As depicted in Fig. 3.7, the Local Features Reconstruction module receives 2 adjoining feature maps, β_{lev-1} and β_{lev} , as in the feeds and produces a feature map β_{out} . as an outfeed This resultant feature map tends to combine both the spatial and semantic information from β_{lev-1} and β_{lev} . To maintain the relative spatial information while reconstructing and enhancing the quality of the up-sampled feature with semantic richness, the local feature rebuild method tends to linearly integrate $n \times n$ the collinear neighborhood surrounding every unit point.

3.3.5 Classification

In the classification process, a widely recognized deep residual network called ResNet-50 is employed as the initially trained network inside the deep convolutional neural network (DCNN). In order to solve the vanishing gradient problem that frequently arises in deep networks, this network is designed for the same. The output from ResNet-50 is then fed into a SoftMax classifier to continue the classification process. There are various types of layers within the DCNN [114] termed as Down Sampling Layers, Convolutional Layers, and Fully Connected Layers marked as the procedures for identifying and categorizing the features. The depth of a DL model, which refers to the number of layers, plays a crucial role in improving classification accuracy. Additional layers are often added to a CNN to increase its depth. However, it's observed that beyond a certain depth threshold, the accuracy of the network starts to decline gradually.

$$M_f = K_f + f \quad (3.9)$$

To mitigate this degradation problem, ResNet includes a mapping function in its architecture. The mapping function, as described in Eqn. 3.9, is designed to reduce the impact of degradation and improve the overall performance of the network.

The mapping function M_f is created using a feed-forward neural network in combination with a skip connection. Here, f signifies the input to the network. Typically, the skip connection is an identity mapping, achieved by directly bypassing certain layers. The residual mapping function is denoted as $K(f, N)$, which shows the variation between the input, f and the output of the network, N . This relationship is expressed in Eqn. 3.10.

$$R = K(f, N) + f \quad (3.10)$$

In the ResNet model's convolutional layers, 3x3 filters are used, and down sampling is performed with a stride of two. Classification outcomes are generated by

combining a SoftMax layer with global average pooling. The ResNet architecture is built after inserting shortcut connections.

3.3.6 Experiment and Results

The proposed scheme involved a segmentation component using U-Net [115] that is well suited to segment medical image by utilizing encoder-decoder architecture with skip connections for efficient capturing of spatial information and generation of precise segmentation masks. It provides the information on identification and delineation of tumor regions within MRI images. The model was trained to segment LGG and HGG [116] within the MRI images as depicted in Fig. 3.8 (a) & (b).

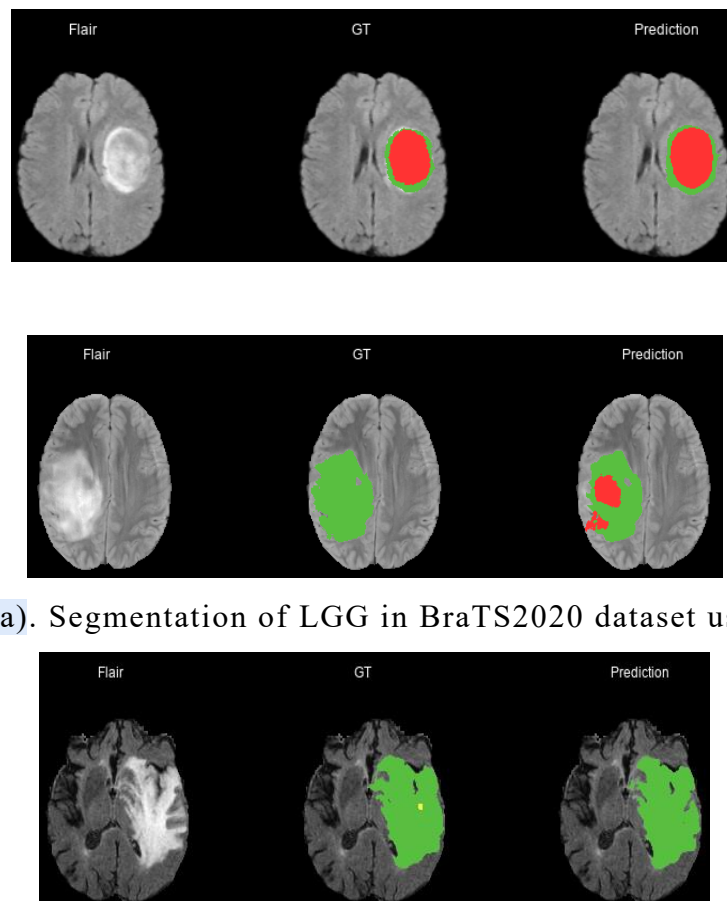


Fig. 3.8 (a). Segmentation of LGG in BraTS2020 dataset using U-Net

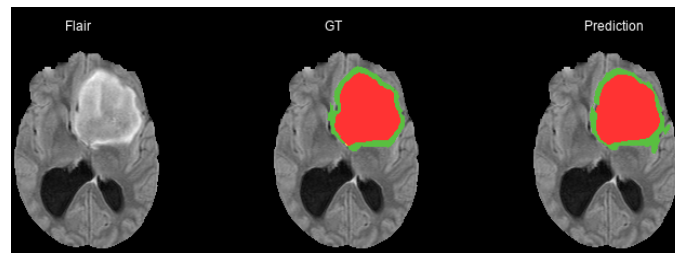


Fig. 3.8 (b). Segmentation of HGG in BraTS2020 dataset using U-Net

In this subsection, qualitative as well as quantitative analysis are conducted in order to assess the accuracy and loss in comparison to other existing techniques, and the results are juxtaposed with those of the proposed algorithm. Utilization of ResNet for classification leveraged its ability to diminish the degradation problem associated with deep networks, resulting in robust classification outcomes. The qualitative analysis of the proposed solution is given with accuracy and loss function derived while training and testing the model and that is depicted in Fig. 3.9.

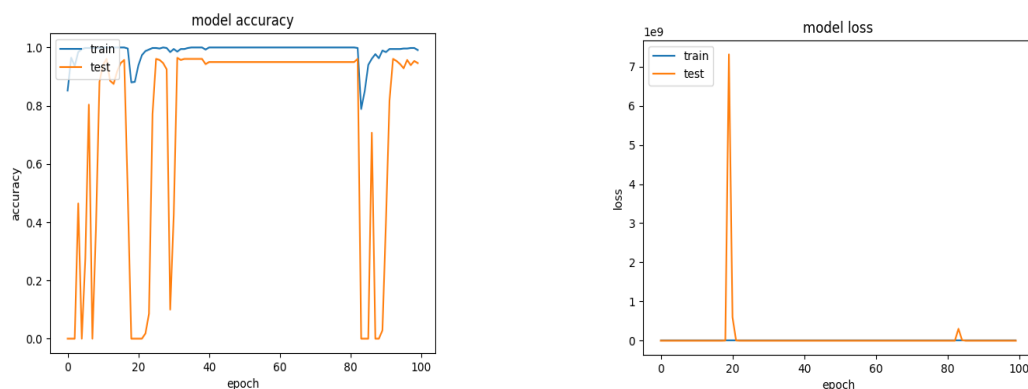


Fig. 3.9. Accuracy and loss function of the proposed algorithm in training and testing

Finally, the quantitative comparison of this study's results with the cutting-edge scheme indicates that the developed algorithm shows higher visual as well as quantitative performance. Details of this study's performance are shown in Table 3.4.

Table 3.4 Comparative analysis of the suggested model with other types of brain tumor detection methods in recent literature

Ref	Year	Technique	Purpose	Dataset	Acc	Precision
[117]	2022	MobileNetV2 and LSTM hybrid network	Improve The Edge Details	BraTS 2018 Figshare	97.47% 98.92%	92.9
[98]	2019	GoogLeNet	Feature Extraction	Figshare	98%	89.9
[118]	2022	DarkNet-53	Automatically classify, and localize tumor	T1W-CE MRI	98.54%	93.3
[96]	2020	CML-ELM	Feature Fusion	BraTs2015, BraTs2017, BraTs2018	98.16%, 97.26%, and 93.40%	91.4
Proposed	-	Leveraging Modified Triplet Cross Fusion Learning	To reduce unwanted artifacts	BraTS2020	99.56%	97.3

3.3.7 Discussion

This research introduces a thorough method to improve brain tumor detection accuracy in MRI images. The approach involves denoising, segmentation, and classification using advanced techniques customized to overcome the challenges of medical image processing, specifically in reference to brain tumors.

The first step involves using the Modified Triplet Cross Fusion Learning (MTCL-MRI) technique to remove noise. This deep learning method is specifically designed to effectively denoise unpaired MRI images. By using a set of three images including a noisy MRI image, a high-quality MRI image, and additional images from other MRI sequences, the MTCL-MRI method significantly minimizes noise while retaining vital structural details. This is crucial for enhancing the clarity and quality of the images, which is essential for accurate subsequent processing stages.

After removing noise from the images, the study uses segmentation to extract statistical features using both local and global methods. This combination ensures a comprehensive analysis of the spatial relationships of pixel intensity values within the images, yielding valuable statistical information for accurate tumor characterization. The segmentation is done with the help of a reliable architecture, U-Net which is also renowned for its effectiveness in medical image segmentation, thanks to its encoder-decoder structure with skip connections.

ResNet-based approach is utilized for the classification task. ResNet is known for its deep residual learning capabilities, that address the vanishing gradient problem often seen in deep networks, thus improving classification accuracy. The features extracted are then used to differentiate between benign and malignant tumors, aiding in precise medical diagnosis as well as treatment planning. Additionally, integration of the SoftMax layer with global average pooling in ResNet further enhances the classification outcomes.

The suggested methodology was thoroughly assessed using the BraTS 2020 dataset, which contains detailed MRI images specifically intended for brain tumor analysis. The results showcased superior performance in terms of accuracy and noise reduction in comparison to existing techniques. Both qualitative and quantitative analysis, including comparisons of accuracy and loss functions throughout the training and testing phases, emphasized the reliability of the suggested pipeline. The study reported a classification accuracy of 99.56%, which is noticeably more than other contemporary methods mentioned in recent literature.

The suggested method was compared with cutting-edge techniques, and the results confirmed its effectiveness. We considered methods like MobileNetV2 and LSTM hybrid networks, GoogLeNet, and DarkNet-53, and found that our method outperformed them when assessed for accuracy and noise handling. These findings demonstrate the potential of combining MTCL-MRI, U-Net segmentation, and ResNet-50 classification to advance medical imaging technologies and improve

diagnostic accuracy for brain tumors. this study introduces a robust and innovative pipeline for brain MRI image processing aimed at enhancing brain tumor detection.

By leveraging advanced methodologies in denoising, segmentation, and classification, the proposed approach addresses key challenges such as image noise, artifact reduction, and accurate tumor classification. The comprehensive evaluation and comparative analysis underscore the potential of this pipeline to significantly improve medical imaging and support more accurate and efficient diagnosis of brain tumor as well as treatment planning.

3.4 SUMMARY

This work incorporates two notable contributions to the area of medical images utilizing MRI. These works address the issue of Rician noise in MRI images. In the first work, we proposed an advanced denoising technique with DNN to improve image quality. The study was successful in preserving important image features while achieving significant noise reduction. The second work focuses on the detection of brain tumors making use of the MTCL scheme along with U-net segmentation and ResNet classification. MTCL helped in reducing Rician noise in MRI images. The integration of U-net segmentation and ResNet classification ensures robust outcomes, overcoming the degradation problem associated with deep networks. Together these work focuses on providing advanced algorithms that enhance the effectiveness of MRI by denoising and improving the accuracy.

CHAPTER 4

DETECTION OF BRAIN TUMOR USING TRANSFER LEARNING

4.1 PREAMBLE

Brain tumors are considered to be a threat to the patient's neurological functions as they offer challenges in medical diagnostics due to their critical location. With the help of MRIs, clinicians tend to visualize the internal brain structures with high resolutions and contrast. The conventional methods pose a challenge to the detection as they rely upon the skills of the radiologists which becomes time-consuming which further leads to variations in interpretations. Due to the introduction of CNN, the accuracy and efficiency of medical image analysis have been increased as it has automated feature extraction and classification tasks.

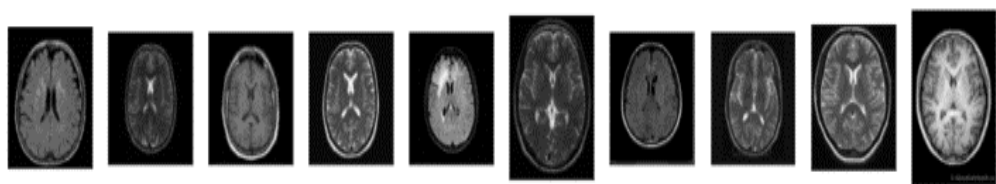
The chapter suggests the challenges posed by the limited annotated medical images, that offer challenges in training the deep learning models from scratch. In order to overcome these challenges, there is the utilization of transfer learning that further utilizes the pre-trained Convolutional Neural Network architectures that include VGG-16, VGG-19, InceptionV3, ResNet50, ResNet101, and EfficientNetB1. These originally trained models on large datasets such as the Image-net are calibrated for our specific tasks of detection of brain tumors by utilizing MRI scans. This research focuses upon the pre-trained models such as VGG16 and ResNet101 depicting its high accuracy in classifying the images of the brain as tumor and non-tumor images thereby leading to the provision of reliable and effective diagnostic of brain tumors.

This study not only offers a diverse analysis of different deep learning models and their effectiveness but also draws special attention to the potential of transfer learning in medical imaging. The findings are quite indicative towards the

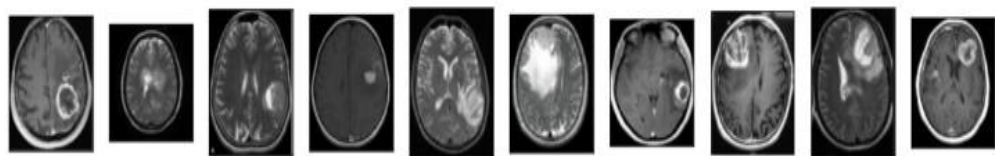
development of automated systems that offer assistance to radiologists in diagnosing tumors more effectively and efficiently.

4.2 DATASET

This study utilized a dataset comprising of 646 T1-weighted MRI brain images, which are accessible for free and are tagged as either tumor or non-tumor. These images have been acquired from TCIA's (The Cancer Imaging Archive) publicly accessible repository and were captured from 20 patients who were diagnosed with glioblastoma, which is quite a hostile form of brain cancer [41]. From the total images, 548 images are showcasing tumors while 98 showcased as non-tumors, thereby providing varying representations of brain scans for robust model training and its evaluations. Due to the availability of the images in the JPG/JPEG formats, it is compatible with various image processing tools and software. The images as represented in Fig 4.1 are from both the tumor and the non-tumor classes, thereby indicating the variations present within the data. These well-chosen datasets thereby help in efficiently training the deep learning models for the accurate diagnosis and detection of brain tumors.



(a) Images representing brain MRI with non-tumor



(b) Images representing brain MRI with tumor

Fig.4.1. Dataset described tumorous and non-tumorous images

4.3 DATA PRE-PROCESSING AND ENHANCEMENTS

Data Pre-processing offers raw data in formats that are more understandable and suitable for analysis and processing ahead. This transformation thereby upgrades the quality and usability of the images and therefore making them suitable for training and evaluating deep learning models. There are various steps involved in pre-processing pipelines, each giving an outline to address the challenges included in the raw medical imaging data. The steps usually include noise reduction to eliminate artifacts, intensity homogenization in order to standardize the pixel values across images, and also changing the size to ensure uniform dimensions. Moreover, the process of skull stripping is utilized to remove the non-brain tissues and enhance the contrast thereby improving the visibility of brain structures. The various stages involved in the data pre-processing are depicted in Fig. 4.2. It showcases the systematic approach initiated for the preparation of the MRI images for the detection of the brain tumor accurately and efficiently. Meticulous Pre-processing tends to refine the data to a state that tends to enhance the performance and reliability of the subsequent deep learning models for training and evaluation.



Fig.4.2. Steps involved in Data Pre-processing

A total of 646 images are consisted of this MRI dataset. These images are further bifurcated into 323 for training, 193 for validation, and 130 for testing. The images further go under the cropping process in order to isolate the brain region, which assists in focusing upon the relevant anatomical structures. The technique utilized by this cropping process as described in the reference [119], tends to be inclusive of assessing the extreme points within the contour lines in order to define the boundaries of the brain quite precisely. This helps in the removal of the extraneous areas from the images, thereby representing only the essential portions of the brain required for

accurate analysis. Fig.4.3 represents the resized image after the identification of the extreme points lying inside the outlines, further depicting the effectiveness of pre-processing steps in preparation of the images for further processing. For improved accuracy in the brain tumor, this refined approach makes sure that the deep learning models are trained on the data so as to be relevant and optimized for performance.

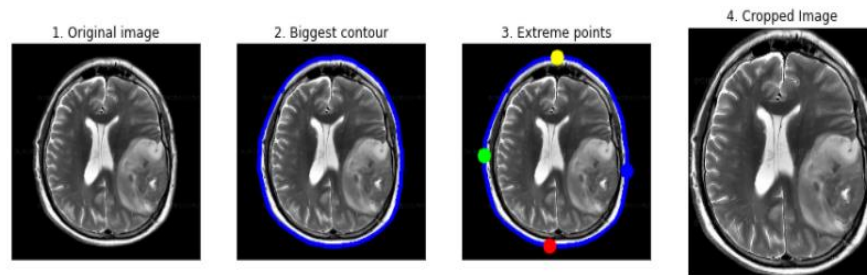


Fig.4.3. Cropped image after finding extreme points in contour

Acquiring the MRI image is quite costly and complex, leading to the limited dataset sizes for certain applications. The dataset used in this study is comparatively smaller even though the DL models perform better with the larger datasets. The data augmentation has been done in this study to address these limitations so that the size of the data increases artificially. This includes initiating variations to the original images like rotations, translations, and flips, in order to give rise to a diverse set of modified images. This process tends to enhance the model's learning ability as well as reduce the risk of overfitting [120] by making it more generalized to the unseen data. Fig. 4.4 represents the images post data augmentation, showcasing the various modified samples created from the original datasets. The training datasets are enhanced effectively by this approach therefore further enhancing the robustness and performance of DL models in the detection of brain tumors.

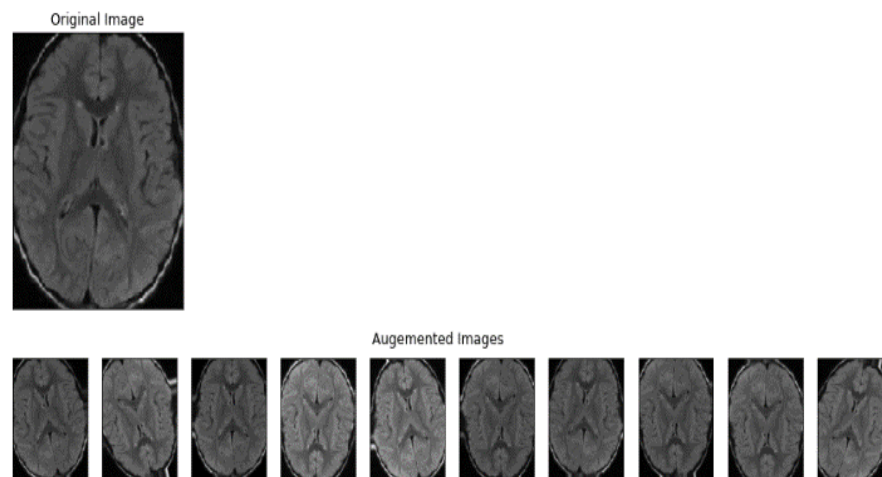


Fig.4.4. Image Samples after augmentation

4.4 PRE-TRAINED CNNs ARCHITECTURE

Deep Learning has now become increasingly popular and demanding over the last few decades due to its accessibility to large datasets and advancements of high-performance graphical processing units also referred to as GPUs which were part of the gaming industry initially. The multiple layers consisting of the Neural network are capable of learning and extracting complex features from input images thereby enabling the analysis of different types of images by utilizing the labeled as well as the unlabeled datasets. The capability of DL to process and learn from massive quantities of data has greatly improved image analysis applications, allowing for more accurate and efficient interpretation of complex medical images [121].

CNNs are commonly used in DL for analyzing medical images [122]. CNNs are designed to learn spatial hierarchies of features through a series of layers, that includes convolutional layers, non-linear activation layers, pooling layers, as well as fully connected layers. The convolutional layers extract features from the input images, while the pooling layers reduce the dimensionality of the data, preserving essential information. Fully connected layers then select the most efficient characteristics and tend to pass into the classification layer for final prediction. In this study, various pretrained CNN architectures have been used to take advantage of their feature extraction capabilities. Table 4.1 provides details on the boundary limits that

are used in various CNNs architectures, offering insights into their configurations and performance, offering insights into their configurations and performance.

In reality, it is unlikely that a person can train a full CNN model from scratch as datasets with adequate sample data are usually not feasible. Insistently, pre-training a CNN on extensive datasets e.g. ImageNet seems to have become a standard exercise. Transfer learning (TL) [123],[124] is being considered as the improved learning prospects in a novel problem as it tends to extract characteristics attained from an analogous problem that occurs. TL incorporates the usage of the characteristics obtained from one dataset across the other datasets as well.

Table 4.1. Convolutional Neural Network Architectures used Boundaries

Simulation Environment	Framework	Mini Batch	Data Augmentations Parameters	CNN Architecture	Input Size	Optimization Technique	Learning Rate	Loss Type
Python	Keras	16	Vertical Flip = True Horizontal Flip = True Random Brightness Range = [0.3 1] Shear Range = 0.1 Rotate Limit = 10	VGG16	240x 240	Ftrl	0.00 1	Binary cross entropy
				VGG19	240x 240	Ftrl	0.00 1	Binary cross entropy
				InceptionV3	240x 240	Ftrl	0.01	Binary cross entropy
				ResNet 50	240x 240	Ftrl	0.00 1	Binary cross entropy
				ResNet 101	240x 240	Ftrl	0.00 1	Binary cross entropy
				EfficientNetB1	240x 240	Ftrl	0.00 1	Binary cross entropy

VGG16 and VGG19 models [125]- A thorough study related to the effect of Convolutional neural network penetrating its performance on computational vision has been done by Simonyan K. and Zisserman A. in their work on the VGGNet network. They incrementally increased the network penetration from 11 to 19 weight layers by utilizing a comparatively little convolution filter. Among these variations, models with 16 to 19 weight layers, identified as VGG16 and VGG19 respectively, demonstrated superior performance. The findings indicated that as the depth of the network increased, the classification error decreased, eventually reaching a saturation point when the penetration exceeded Nineteen layers. The authors emphasized the importance related to network penetration in enhancing visual representations and confirmed its critical role in improving the accuracy and effectiveness of CNN models in computer vision tasks.

InceptionV3 model [126]- The Inception model, developed by the Google Brain Team, is a 48-layer Convolutional Neural Network (CNN) designed to improve the efficiency and performance of DL models. It has been trained on the extensive ImageNet database and is capable of categorizing objects into 1,000 distinct classes. One notable advantage of this model is its fast training speed when compared to other inception models and processes. This speed is achieved through a sophisticated architecture that optimizes computational resources, while still maintaining high accuracy. As a result, it is a valuable tool in various computer vision applications.

ResNet model [127]- The development of ResNet by Microsoft, led to the onset of an in-depth residual learning platform that simplifies training of deeper networks and reduces errors as network depth increases. ResNet introduced configurations including 18-layer, 34-layer, 50-layer, and 101-layer frameworks. Its innovative design allows for greater depth without common issues associated with deep networks, such as vanishing gradients. By using shortcut connections that skip one or more layers, ResNet effectively mitigates these issues, enabling the network to learn more efficiently. Compared to the VGG network, ResNet is less complex and significantly deeper, offering improved performance and accuracy in various computer vision tasks.

EfficientNetB1 model [128]- The Brain team of Google progressed with the EfficientNet, a Convolutional Neural Network that focuses on optimizing model scaling. Their research showed that if carefully balanced the depth, width, and resolution of the network can lead to superior performance. By scaling the Convolutional neural network to give rise to deeper understandable models, they achieved remarkably upgraded efficiency and accuracy as compared to formerly utilized CNNs. EfficientNet's innovative approach to model scaling allows for more effective utilization of computational resources, achieving state-of-the-art performance in several computer vision tasks even though maintaining a comparatively low computational cost.

4.5 PROPOSED METHODOLOGY

The diagnosis method proposed tends to focus upon improving the accuracy of the detection of brain MRI images by utilizing the DL model and TL methodology. Fig 4.5 represents the flow diagram of the suggested work performed.

Transfer Learning (TL) indicates the technique that utilizes pre-trained models to efficiently memorize new tasks by using features extracted from these models. In the context of medical imaging, TL is particularly advantageous for analyzing complex medical images as it allows for the utilization of low-level features learned from vast datasets encompassing diverse medical image modalities or various domains. This approach facilitates the rapid acquisition of new tasks by building on the pre-existing knowledge embedded within deep learning models. In this study, we first meticulously pre-processed MRI data to ensure its compatibility with our model framework, followed by extensive testing. We then utilized 6 pre-trained CNNs models—VGG16, VGG19, InceptionV3, ResNet50, ResNet101, and EfficientNetB1.

These models were trained using brain MRI images to effectively distinguish between tumorous and non-tumorous images. The application of TL not only enhanced the training efficiency but also improved the accuracy of the classification,

demonstrating the capabilities of pre-trained DL models in advancing medical image analysis.

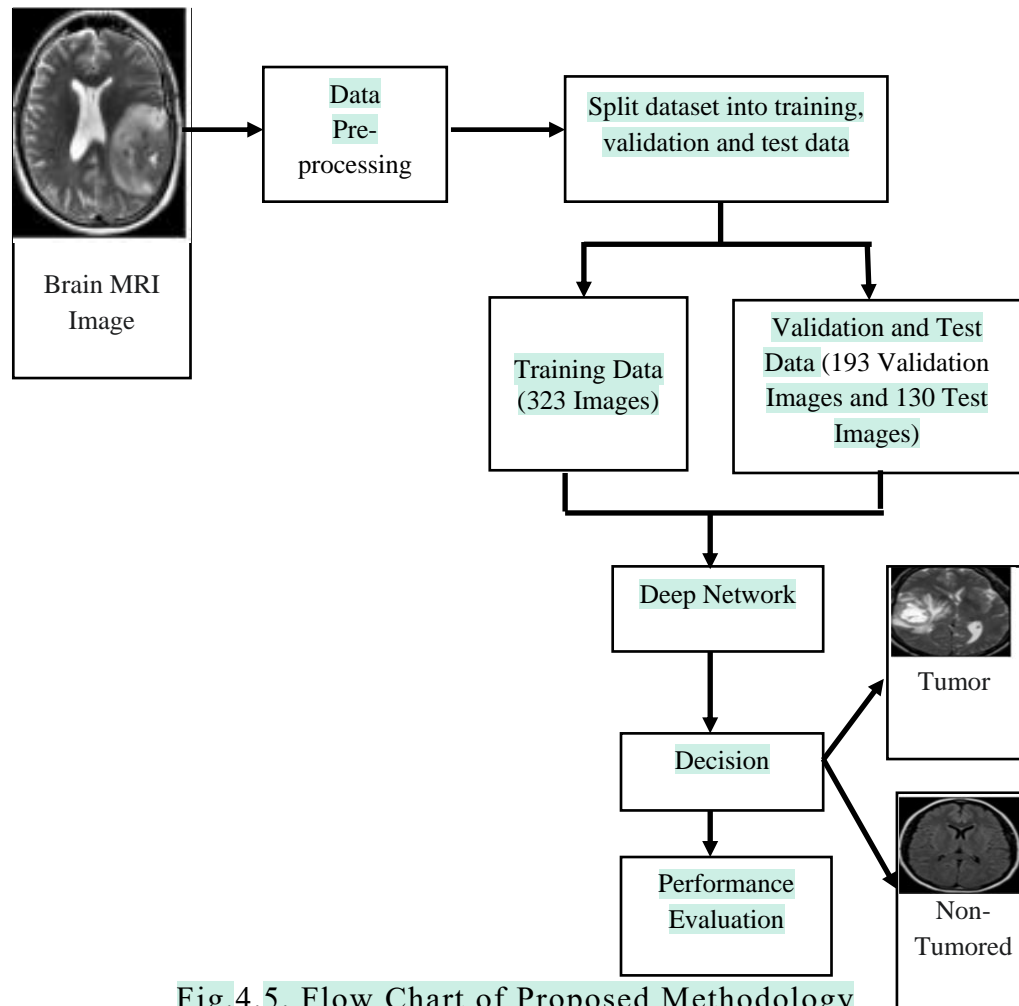


Fig.4.5. Flow Chart of Proposed Methodology

4.6 EXPERIMENTAL RESULTS

All the models were trained for 50 epochs. Fig. 4.6 and Fig. 4.7 were attained on the mark of training the models utilizing brain MRI image dataset for 50 epochs and thereby representing the accuracy and loss curve related to training and validation set for VGG16 as well as ResNet101 models.

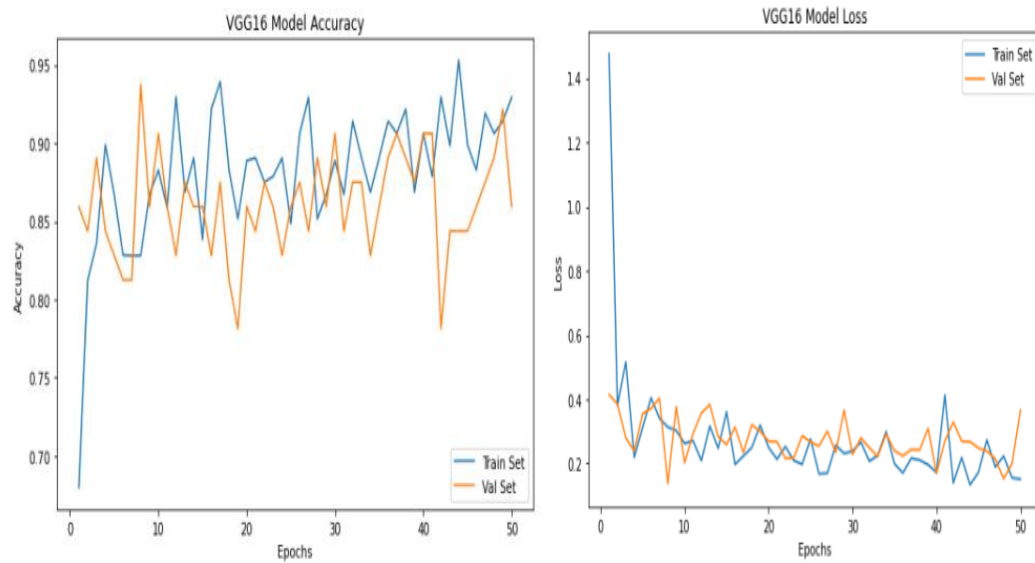


Fig. 4.6. Accuracy and loss curve for training and validation set for VGG16 model

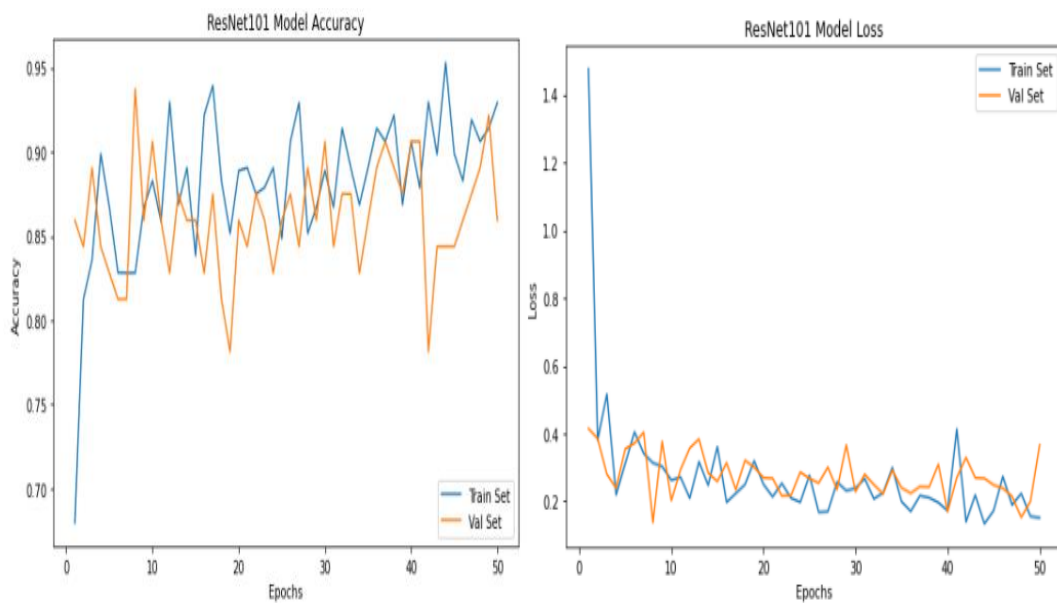


Fig. 4.7. Accuracy and loss curve for training and validation set for ResNet101 model

Table 4.2 displays the performance of various models utilized in our study. Table 4.3 shows the comparison of the proposed VGG16-based model with relative studies.

Table 4.2. Performance of different pretrained models

CNN Architecture	Acc	Precision	Recall	F1 score	Kappa	AUC	Confusion Matrix
VGG16	0.92	0.97	0.94	0.95	0.73	0.89	$\begin{bmatrix} 17 & 3 \\ 7 & 103 \end{bmatrix}$
VGG19	0.87	0.93	0.91	0.92	0.53	0.78	$\begin{bmatrix} 13 & 7 \\ 10 & 100 \end{bmatrix}$
InceptionV3	0.85	0.85	1.0	0.92	0.00	0.50	$\begin{bmatrix} 0 & 20 \\ 0 & 110 \end{bmatrix}$
ResNet50	0.91	0.94	0.95	0.95	0.63	0.80	$\begin{bmatrix} 13 & 7 \\ 5 & 105 \end{bmatrix}$
ResNet101	0.92	0.96	0.94	0.94	0.69	0.87	$\begin{bmatrix} 16 & 4 \\ 7 & 103 \end{bmatrix}$
EfficientNetB1	0.89	0.94	0.94	0.94	0.59	0.79	$\begin{bmatrix} 13 & 7 \\ 7 & 103 \end{bmatrix}$

Table 4.3. Comparison with relative studies.

Ref	Dataset	Technique	Acc	S_n	S_p	Precision	F1-Score	AUC
[129]	TCIA	K-Nearest Neighb or method	89.5%	-	-	-	-	-
[130]	TCIA	SVM	90%	90%	-	90%	90%	-
[131]	TCIA	GoogleNet	90.9%	-	-	-	-	0.939
[132]	TCIA	VGG16	92%	92%	92%	-	-	-
Proposed Technique	TCIA	Based on VGG16	92%	94%		97%	95%	0.89

Performance evaluation was carried out for the different techniques on the basis of various measures such as Accuracy, Precision, Recall, F1 score, Kappa, and AUC. VGG16 and ResNet101 report the highest test accuracy. ResNet101 marks the training accuracy of 94%. Comparatively, VGG16 has the high F1 score as the precision and recall are higher in this case. As per the analysis based on the performance measures, both VGG16 and ResNet101 have better performance when compared to the other pretrained networks.

4.7 DISCUSSION

153 In this study, we used six pre-trained CNN architectures for the classification of brain MRI as tumor or non-tumor. The models we used were VGG16, VGG19, InceptionV3, ResNet50, ResNet101, and EfficientNetB1. Our main goal was to accelerate the training process and achieve high classification accuracy by making use of transfer learning. This involved utilizing features learned from a large dataset of images that are obtained from numerous medical image modalities.

The first step involved preparing the MRI data to make sure it was ready for the models. This process included normalizing the data, resizing the images, and using augmentation techniques to surge the variety of the training dataset. This helped in the improvement of the model's robustness and ability to generalize.

113 Each of the pre-trained models (using the brain MR images dataset) has been fine-tuned. The proficiency of these models has been assessed on the basis of the accuracy, precision, recall, and F1-score metrics. Amongst the various models, VGG16 and ResNet101 possessed superior results by attaining the good accuracy and effectiveness. Both VGG16 and ResNet101 have attained 92% accuracy rates and were found to be quite effective in learning the intricate patterns and features that are linked with the tumor and non-tumor brain MRIs.

121 Transfer learning has now emerged as the most useful learning technique in medical image classification. VGG16 and ResNet 101 have showcased quite a superior performance and also depicted that sufficient depth and residual learning tend to have a good distinguishing skill in order to distinguish between the brain MR images with tumor and without tumor. By utilization of the pre-trained models, this has been identified that there is a decrease in the time as well as the computational resources required for the training, so attaining high accuracy thereby making them suitable for real-world medical applications.

Furthermore, the variances in proficiency between the numerous models highlight the significance of choosing a model architecture that suits the specific characteristics of the medical imaging data. Future studies could delve into additional fine-tuning and the possible integration of these models to improve accuracy and reliability. Moreover, expanding the dataset and including images from diverse sources could enhance the model's ability to generalize and be applicable to a wider range of clinical scenarios.

The transfer learning models such as VGG16 and ResNet101 have an effective approach to detecting brain tumors by utilizing Magnetic Resonance images and lays an important ground for further advancement in automated medical diagnostics.

4.8 SUMMARY

This research evaluates the performance of 6 pre-trained CNN models - VGG16, VGG19, InceptionV3, ResNet50, ResNet101, and EfficientNetB1 in the classification of brain MRI images as either tumorous or non-tumorous by utilizing Transfer learning. After systematic preprocessing of the MRI data, each of these models was fine-tuned and evaluated on the basis of different metrics such as accuracy, precision, recall, and F1-score. The VGG16 and ResNet101 were found to be performing with the highest level of accuracy. The findings further ensure that the TL improves the training efficiency and accuracy thereby showcasing that the DL models, specifically VGG16 and ResNet101, VGG16 and ResNet101 are quite accurate in the detection of brain tumors. This study therefore focuses upon the potential of these models in advancing medical image analysis and diagnostics thereby providing a strong background for future clinical research.

CHAPTER 5

DETECTION OF BRAIN TUMOR USING MULTI-LAYER CONVOLUTION NEURAL NETWORK

5.1 PREAMBLE

Diagnosing brain tumors accurately and promptly is a major challenge in medical image processing. Traditional human-assisted diagnostics struggle with the large volume of data and the complex, varied appearances of tumors, which can be mistaken for normal tissues. To tackle these issues, this chapter presents a new approach that utilizes advances in deep learning, specifically convolutional neural networks (CNNs).

The triplanar Multi-Layer Morphological Convolutional Neural Network introduced in this study is designed to improve the detection of brain tumors from 2D MRIs. The approach is based on the DeepLab framework but includes enhancements that decrease computational requirements and time complexity without sacrificing accuracy. By incorporating spatial pyramid pooling and Atrous convolution techniques, the system can take hold of both local and global contextual details, which is essential for precise tumor detection.

The aim of this research is to provide a contribution to the field of medical image analysis by providing a more efficient, accurate, and cost-effective solution for brain tumor diagnosis, potentially improving patient outcomes through earlier and more reliable detection.

5.2 PROPOSED METHODOLOGY

This work introduces an advanced technique for predicting brain tumors from MRI images. It combines triplanar isolated spatial pyramidal pooling with Atrous convolution to minimize computational costs and enhance system performance. This innovative approach captures both local and global contextual information more effectively and maintains the resolution of feature maps, ensuring that critical details are preserved without the excessive computational burden typically associated with 3D convolutional networks. The proposed methodology not only streamlines the computational process but also significantly improves the accuracy of tumor detection, ultimately facilitating earlier and more precise diagnoses. This advancement has the potential to improve patient outcomes by enabling timely and optimal treatment interventions, thereby enhancing survival rates for individuals diagnosed with brain tumors.

The DeepLab [133] network discovered a significant problem with the traditional method of downsampling the image size by a factor of 32. This led to a loss of important information, which became more problematic during the process of increasing the image size, requiring a lot of computing power. To solve this problem, DeepLab used Atrous convolution, a technique that maintains the spatial resolution of feature maps while capturing all the image details. They also used Atrous spatial pyramid pooling, a type of dilated pooling, to expand the area that each filter covers. This helped in better feature extraction. Building on this, our research introduces Spatial Isolated Pyramidal Pooling combined with Atrous convolution to accurately predict brain tumors from MRI images. This approach maintains high accuracy in tumor detection while significantly reducing the required number of computations. This lowers computational costs and improves the overall system performance. The proposed method's architecture, illustrated in Fig. 5.1, shows an innovative technique that effectively balances precision and efficiency, offering a robust solution for brain tumor prediction.

5.2.1 Morphological Neural Network

A spatial pyramidal pooling technique that aims to reduce the computational costs associated with local and global contextual feature learning, while improving tumor detection performance in MRI images. The technique enhances accuracy by classifying the local tumor context and preserving the global context, which is essential for predicting brain tumors accurately. A crucial aspect of this method is the use of morphological operations, particularly dilation, within the spatial pyramidal pooling framework. Dilation helps the model capture important features by expanding the receptive field without compromising spatial resolution. By effectively balancing the extraction of both local and global features, this approach significantly reduces computation costs and optimizes the performance of brain tumor detection systems, providing a more efficient and accurate diagnostic tool.

Dilation is a fundamental operator in the field of mathematical morphology, crucial for image-processing tasks. The dilation operator requires two inputs: the main input is the image to be dilated, in addition, the secondary input is a structural element, or kernel, which is usually a small set of coordinate points. This structural element determines how the dilation process modifies the input image. Grayscale dilation, using a disc-shaped structuring element, effectively brightens the image by increasing the size of bright areas surrounding dark regions and decreasing the size of dark areas around bright regions. The image may lose small dark regions and expand tiny bright spots as the surrounding intensity values adjust. This effect is most noticeable in areas of the image with rapidly changing intensity, while areas with uniform intensity mostly remain the same, except at their boundaries. This property of dilation makes it a powerful tool in morphological operations. It enhances the detection of critical features in medical images, such as brain tumors, by effectively highlighting relevant structures and smoothing out noise. The proposed network learns the features by adjusting weights and errors.

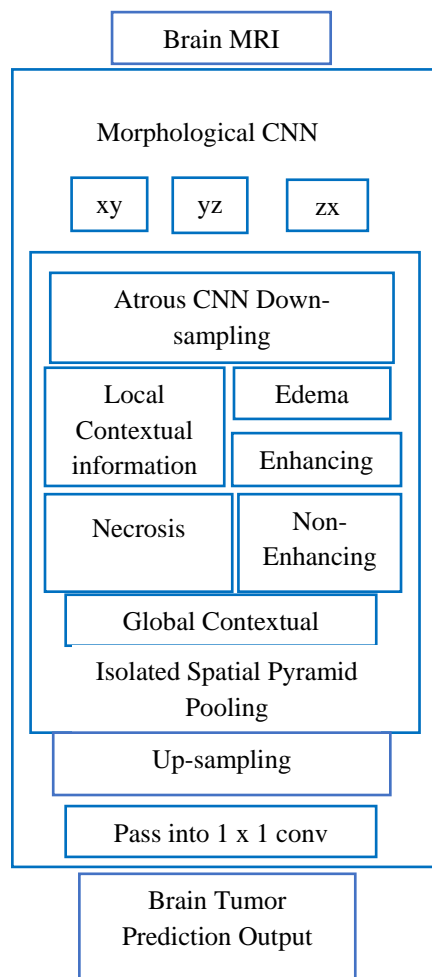


Fig.5.1. Architecture of the Proposed Work

5.2.2 Comparative mathematical analysis of Isolated spatial pyramid pooling with DeepLab model

Let the input image be of size 5×5 . For dilation, it uses a filter. Let the filter size be 3×3 .

DeepLab uses spatial pyramidal pooling with Atrous convolution neural network for feature learning of an input image of size 5×5 using filter 3×3 , which results in an output image of 5×5 with 81 feature learning computations. Fig. 5.2. shows the input image size and different filter sizes.

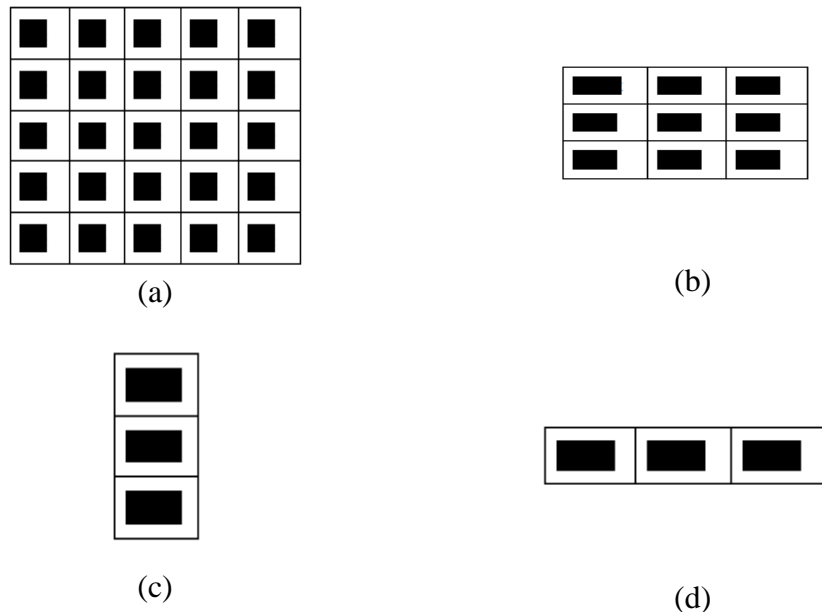


Fig.5.2. (a) 5*5 Input Image (b) 3*3 Filter (c) 3*1 Filter (d) 1*3 Filter

The steps done by DeepLab are as follows:

- (i) For a 3*3 matrix, column-wise it has taken 3 operations for the feature to learn the complete width of the input image.
- (ii) For a 3*3 matrix, row-wise it has taken 3 operations for the feature to learn the complete height of the input image.
- (iii) So, the entire number of operations for the positions= $3*3=9$. So, there are 9 positions.
- (iv) The total number of computations = $9 * \text{filter size} = 9 * 3 * 3 = 81$.

The proposed isolated spatial pyramidal pooling technique isolates the 3*3 filter into two parts of the filter.

- (i) Let the input image be the same, of size 5*5 and filter 3*3.
- (ii) Input image is passed through the 3*1 filter. An intermediate output image is generated with this. This intermediate image obtained at the output is then passed through the 1*3 filter. The image obtained is the final convoluted output image.
- (iii) Considering the 3*1 filter matrix, the column-wise proposed method took 5 operations to feature learn the entire width of an input image.

- (iv) Row-wise it took a total of 3 operations to feature learn the entire height of the input image when considering a 3×1 filter matrix.
- (v) The entire number of operations for the positions is $5 \times 3 = 15$. So, there are 15 positions.
- (vi) The total number of computations = $15 \times \text{filter size} = 15 \times 3 \times 1 = 45$. This is the number of computations to generate an intermediate image from filter 1.
- (vii) The intermediate image now has a size = 5×3 as shown in Fig. 5.3.

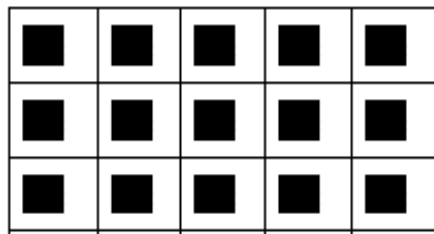


Fig.5.3. 5×3 intermediate image

- (viii) Now, on this intermediate image, for a 1×3 filter matrix, it took 3 operations column-wise to feature learn the entire width of the input image.
- (ix) For a 1×3 filter matrix, it took 3 operations row-wise to feature learn the entire height of the input image.
- (x) Entire number of operations for the positions = $3 \times 3 = 9$. So, there are 9 positions.
- (xi) The total number of computations = $9 \times \text{filter size} = 9 \times 1 \times 3 = 27$
- (xii) This is the number of computations for images convoluted from filter 2 i.e. 1×3 filter.
- (xiii) Therefore, the total number of convolutions from the spatial isolated pooling is $45 + 27 = 72$.

Hence, the Isolated spatial pyramidal pooling technique computes local and global contextual information at a lesser cost as compared to DeepLab spatial pyramidal pooling, hence reducing the feature learning cost of the overall system and improving performance.

5.2.3 Transposed Convolution Network

For up-sampling, generally, a 1×1 convolution is used. For a 3-D brain MRI image, the 3 dimensions are height, width, and depth. With the use of 1×1 convolution, the depth information is only reduced, while preserving height and width information. Therefore, to reduce the amount of information from all the 3 domains this work uses a transposed convolution.

Generally, transposed convolution is used to transform an input image in the decoding layer of an auto-encoder in the CNN. This helps in retaining global contextual information for the prediction of brain tumors effectively.

The properties of transposed convolution don't interact with the axial information of the input image. In a 1×1 convolution, a unit stride is used.

To obtain the convolution's output size of the convolved image take the stride to be 1 and no padding using the formula as given in Eqn.5.1:

$$\text{Output size} = i + k - 1 \quad 5.1$$

where, i = input image, k = kernel

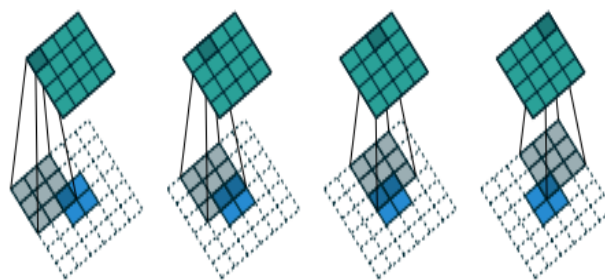


Fig.5.4. The transpose convolution [134]

So, the proposed methodology consists of two major improvements over the previously existing techniques:

- (a) Use of Spatial Pyramid Pool, which results in a lesser number of computations and hence, a better approach.

- (b) Use of a Transposed convolution network, instead of a simple 1×1 convolution, efficiently learns the global contextual information of an image for better brain tumor detection.

Brain MRI images are passed through a transposed convolution neural network for feature learning. The process of down-sampling the image to learn local contextual information and it reduces the excessive loss of global contextual information. Local contextual information includes the following tumor features:

1. Edema
2. Enhancing
3. Non-Enhancing
4. Necrosis

5.3 IMPLEMENTATIONS AND RESULTS

5.3.1 Dataset

For the implementation of the proposed work, tumorous images from the BraTS 2019 dataset and non-tumor images from the Masoud2021 dataset were used. BraTS 2019 makes use of MRI scans which are pre-operative, and it comprises 259 cases of high-grade glioma and 76 cases of low-grade glioma. The focus is on brain tumors that are heterogenic in nature and are also intrinsic. These are heterogenic in terms of shape, appearance, and histology. Manual annotation of all the images is done by radiologists. Fig 5.5 shows segmented tumor with Edema, Enhancing, Non-Enhancing, and Necrosis features.

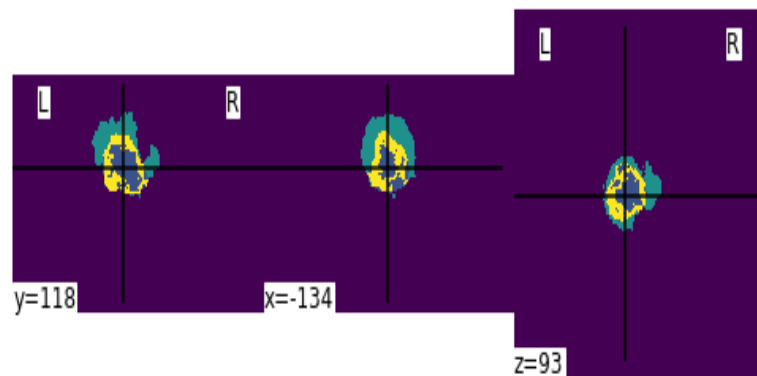


Fig.5.5. The segmented tumor with Edema, Enhancing, Non-Enhancing and Necrosis features

Images are normalized and divided into different folders, which are further divided into training and testing. Artificial augmentation of images is done so that the dataset increases and the model's capacity for learning increases, making it more broadly useful for data that hasn't been seen earlier. On the other hand, it becomes more resistant to overfitting.

5.3.2 Result and Analysis

The evaluation methods used in our proposed model are precision, recall, accuracy, mIOU, and dice score. The proposed method yields 97% accurate results in the prediction of local contextual information accomplished by proposed methodology is represented in Table 5.1.

Table 5.1. Metrics: Precision, Recall and Accuracy

Feature	Precision	Recall	Acc
Edema	0.93	0.90	0.97
Enhancing	0.94	0.83	0.97
Necrosis	0.93	0.83	0.95
Non-Enhancing	0.92	0.99	0.97
Non-Tumor	0.97	0.98	0.99

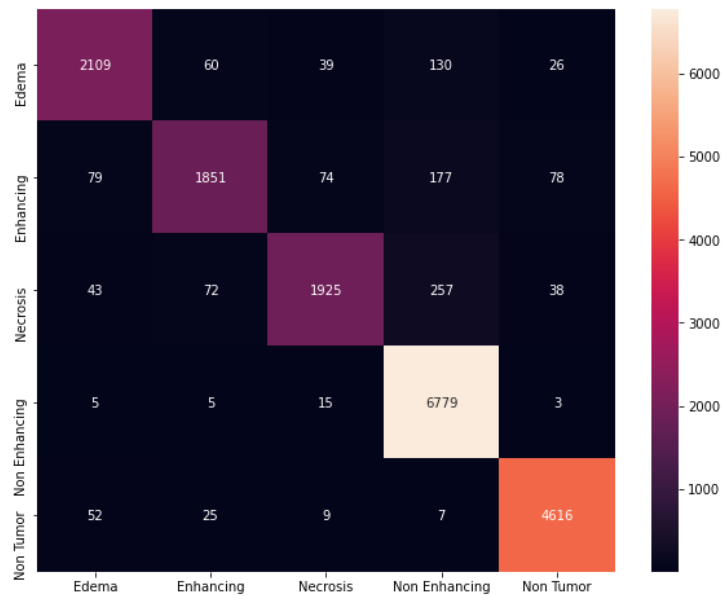


Fig.5.6. Local contextual information confusion matrix

The intermediate output of local contextual information is then fused in an isolated spatial pyramidal pooling stage with transposed convolution to retain global contextual information. The confusion matrix of the results computed on testing data is represented in Fig 5.6.

Fig 5.7 shows the brain tumor prediction confusion matrix using the Proposed Isolated Spatial Pyramidal Pooling CNN. Proposed isolated spatial pyramidal pooling with a transposed convolution network performed better when compared with the traditional methods.

The proposed methodology implementation results in identifying tumor and non-tumor at an accuracy of 91% on the testing dataset. Interpretation and evaluation of the testing dataset are represented using precision, recall, and F1-score in Table 5.2. Table 5.3 represents a comparative analysis of the proposed method to state-of-the-art methods in terms of mIoU Score.

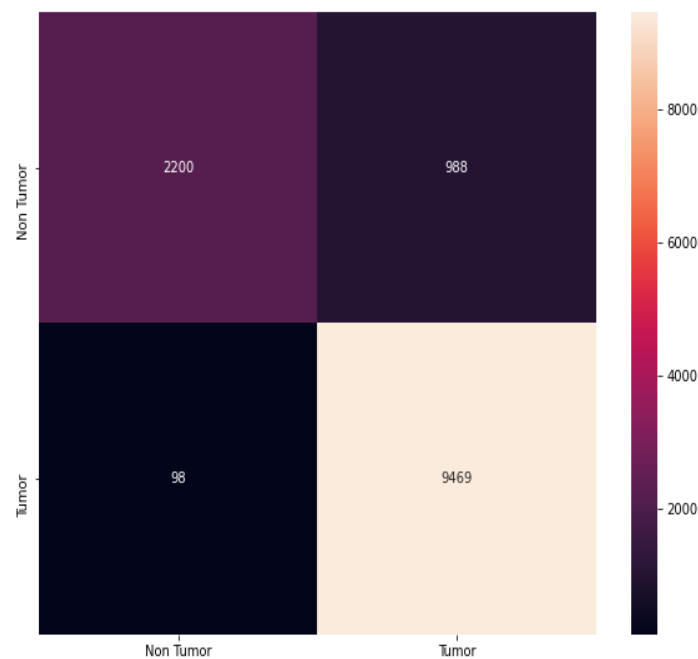


Fig.5.7. Brain tumor prediction confusion matrix using Proposed Isolated Spatial Pyramidal Pooling CNN

Table. 5.2 Performance Evaluation of the proposed model and comparison with other techniques

Ref	Dataset	Technique	Precision	Recall	F1- Score	Acc
[131]	Private Dataset (52 patients (grade II: 25, grade III: 27) with LGG and 61 patients with HGG)	Google Net	-	-	-	86.7%
[135]	TCIA dataset	VGG16	-	93.48%	-	86%
[136]	BRATS 2019	Fused Densenet201 and ResNet50-	85.25%-LGG and 88.5%-HGG	84.75%-LGG and 87%-HGG	-	87.8% and 84.6% for HGG and LGG
[137]	BRATS 2017	CNN	92%	91%	-	91%

Continued on Page-100

Read Table 5.2 (continued)						
[138]	BRATS 2019 and whole brain atlas data sets	CNN with hyperparameters adjustments	-	-	-	88.91%
Proposed Technique	BRATS 2019	Isolated Spatial Pyramidal Pooling CNN	99%	91%	95%	91%

The proposed methodology of isolated spatial pyramidal pooling with a transposed convolution neural network performed better with 85.4 mIoU score in comparison to the existing system with mIoU ranging between 64.3 to 79.7. Base reference model DeepLab resulted in 70.4 mIoU score.

Table. 5.3 Comparative analysis of the proposed method to state of art methods in terms of mIoU Score

Method	mIoU Score
Proposed Isolated Spatial Pyramidal Pooling with Transposed Convolution Network	85.4
[133]	79.7
[139]	77.8
[140]	76.4
[141]	67.8
[142]	67.6
[143]	64.3

Interpretation and evaluation of deep learning models for knowledge extraction are also done through dice scores, which is a competitive metric such as mIoU. Comparative analysis of existing brain tumor prediction model vs proposed isolated spatial pyramidal pooling with transposed convolution network is represented in Table 5.4.

Table. 5.4 Comparative analysis of Brain tumor prediction existing model's vs Isolated pyramidal pooling with transposed convolution neural network

Method	Dice Score
Proposed Isolated Spatial Pyramidal Pooling with Transposed Convolution Network	0.935
[144]	0.90
[145]	0.895
[146]	0.88
[147]	0.868
[148]	0.86

The proposed methodology successfully achieved a 0.935 dice score signifying great similarity between predicted tumor, non-tumor images, and the truth. The existing most significant models reviewed in the literature range between 0.86 to 0.88.

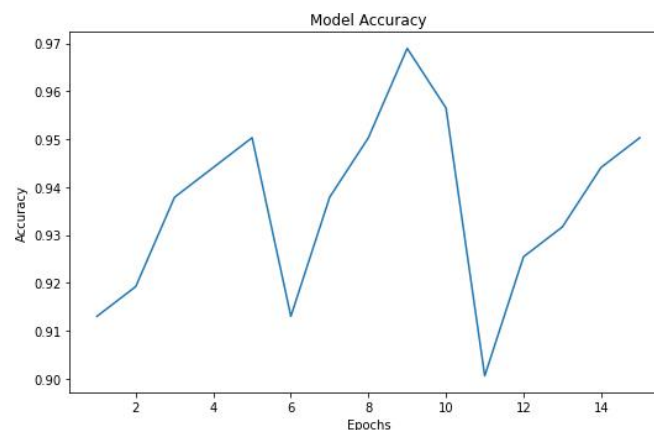


Fig.5.8. The plot of Model Accuracy vs Epochs

Fig 5.8. represents a plot of accuracy vs each iteration of feature learning with adjustment of weights based on error minimization in proposed Isolated spatial pooling with transposed convolution neural network in a percentage ranging 0-1 (0 is 0% and 1 is 100%). Fig 5.9 represents the model loss of each iteration. Loss is represented on a scale of 0-1 where the maximum loss of local and global contextual information in

the prediction of brain tumor is 0.8 and the minimum loss is 0.26. Epochs ranging from 0-15 represent the loss in each iteration of the feature learning process.

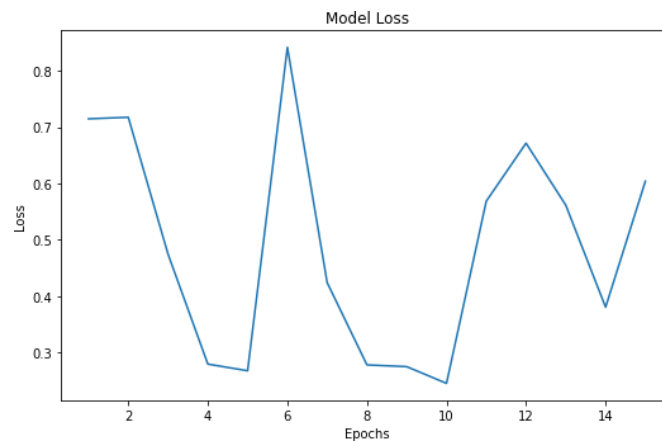


Fig.5.9 The plot of Model loss vs Epochs

5.4 DISCUSSION

The new approach, using an advanced triplanar isolated spatial pyramidal pooling technique combined with Atrous convolution, shows a significant improvement in predicting brain tumors from MRI images. This method tends to overcome the limitations of traditional convolutional techniques by reducing computational costs and thereby improving the accuracy of tumor detection. By utilizing morphological operations, especially dilation, the proposed system effectively preserves both local and global contextual details, which is essential for precise tumor identification.

Atrous convolution, as implemented in the DeepLab network, addresses the problem of information loss caused by down-sampling by preserving the input image size while capturing important features. The addition of isolated spatial pyramidal pooling further improves this capability by enabling efficient feature learning and reducing overall computation. As a result, this leads to a more accurate and cost-effective system for predicting brain tumors.

18 The morphological operation of dilation is important in this methodology. By using a structural component or kernel, dilation effectively brightens the image by increasing the size of bright areas around dark regions and decreasing the size of dark areas around bright regions. This operation is especially useful in highlighting areas with rapid intensity variations, which are important for accurate tumor detection. Dilation helps the image retain its essential features, allowing for better segmentation and classification of tumor regions.

8 The proposed method outperforms others, as shown by evaluation metrics such as mIoU and dice score. The isolated spatial pyramidal pooling with a transposed convolution network achieved a mIoU score of 85.4, which is significantly higher than DeepLab's score of 70.4 and other existing methods, whose scores ranged from 64.3 to 79.7. Additionally, the dice score of 0.935 indicates a high degree of resemblance between the predicted tumor regions and the ground truth, further confirming the effectiveness of the proposed approach.

A comparison with current methods highlights the progress made by this approach. The precision, recall, and accuracy measures for different tumor characteristics, including edema, enhancing, non-enhancing, necrosis, and non-tumor, consistently demonstrate improved performance. For example, the accuracy rates for identifying non-tumor regions can be as high as 99%, showcasing the strength of the proposed system in differentiating between tumor and non-tumor areas.

5.5 SUMMARY

This work describes a new way to use MRI images to predict brain tumors. It does this by combining an advanced triplanar isolated spatial pyramidal pooling technique with Atrous convolution. Addressing the limitations of traditional down-sampling methods, the proposed system reduces computational costs while enhancing detection accuracy by retaining crucial local and global contextual information. Adding dilation to the morphological operations effectively brings out important features, making it easier to separate and group tumor areas. Comparing this method

to the most up-to-date ones shows that it is better, as it gets higher mean intersection over union (mIoU) and dice scores and better precision, recall, and accuracy metrics across a number of tumor features. This methodology not only enables early and precise detection of tumors but also significantly contributes to improving patient survival rates.

CHAPTER 6

DETECTION OF BRAIN TUMOR BY AMALGAMATION OF SEGMENTATION AND RENDERING TECHNIQUE

6.1 PREAMBLE

Brain tumors due to their complex nature tend to pose a significant challenge in the field related to medical imaging and diagnostic as it has the potential to impair neurological functions. In order to improve patient outcomes by means of an effective treatment, it is necessary to have early as well as accurate detections. The traditional/conventional MRI scans are quite time-consuming and thereby also bring into account human errors. Therefore, this chapter introduces a new method where advanced segmentation and rendering techniques are combined by utilizing Deep Learning and Neural Architecture search.

This chapter showcased the advancements in medical image analysis by demonstrating the integration of the Meta Assistant Network (MAN) into the Hypernet framework together with the 3D transformer-based HyperSwinNet architecture. The suggested framework when compared to others, performs better by delivering the best results in the key metrics such as Dice Similarity Coefficient (DSC) and Hausdorff distance (HD) after an ample amount of testing on the BRATS dataset.

The goal of this chapter is to provide a solution to the challenges involved during the detection of brain tumors in the MRI. This chapter not only focuses upon the improvement in the field of medical imaging but also focuses upon improvement in the patient's care by earlier and more accurate detection of the tumor.

6.2 PROPOSED METHODOLOGY

In order to detect brain tumors efficiently, this study offers a comprehensive method for developing an automated system to segment brain tumors. The discussed methodology integrates advanced image processing techniques with machine learning models, thereby offering high segmentation accuracy and robustness across diverse MRI datasets. The suggested pipeline consists of image acquisition, image preprocessing, feature extraction, and segmentation, with every stage precisely designed in such a way that it can classify brain tumors and address the challenges within the MRI image analysis. The below given sub-sections offer a detailed description of each of the components in the pipeline.

The primary contributions of our work are outlined as follows:

- To address the large search space of segmentation networks, we provide a new HyperNet structure called Meta Assistant Network (MAN). This method feeds Neural Architecture Search (NAS) with relevant metadata. Interestingly, post-training MAN can be removed without adding computational expense for further searching and retraining procedures.
- We suggest HyperSwinNet, a ground-breaking 3D transformer-based architecture, for this reason. This architecture consists of an encoder used for pre-training that extracts feature representations at different resolutions from a HyperNet.
- By fine-tuning with the trained Swin Transformer on BraTS, the effectiveness of the proposed framework is improved and it achieves state-of-the-art performance on datasets mentioned.

See Fig. 6.1(a) and Fig. 6.1(b) for the overview. Fig. 6.1(a) illustrates the working of HyperNet and Fig. 6.1(b) shows the working of SwinNet.

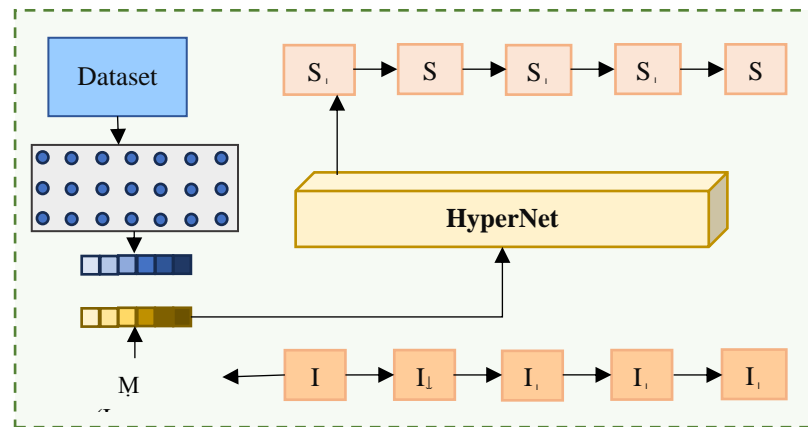


Fig.6.1(a) HyperNet Framework

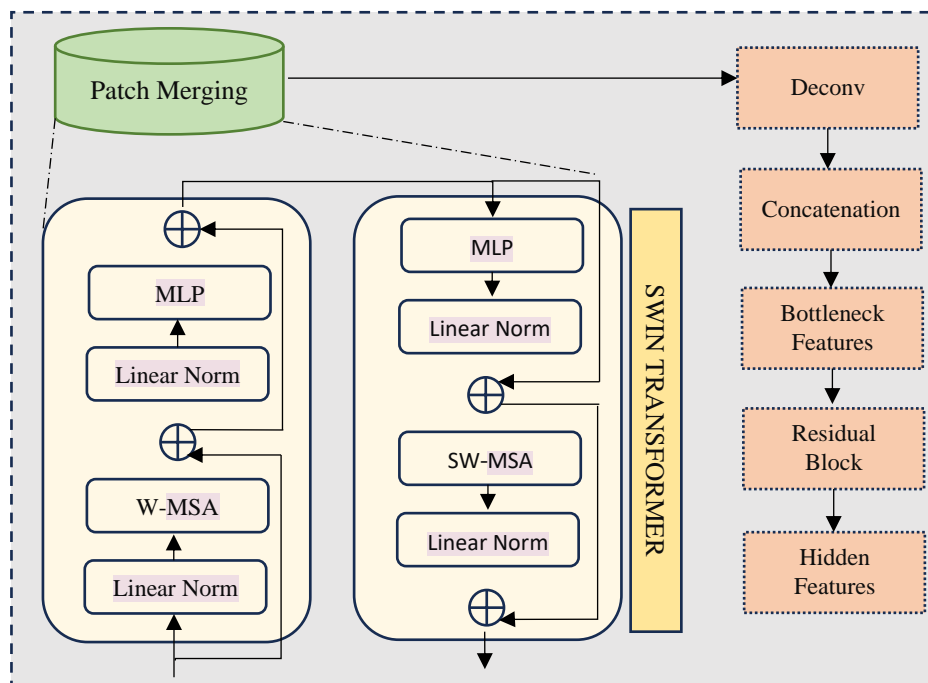


Fig.6.1(b) SwinNet Framework

6.2.1 HyperNet

Neural Architecture Search (NAS) seeks to identify the optimal sub-network s within search space P , given a computing constraint C .

$$\arg \min_s \mathcal{L}_{val}(s, w_s^*), \forall s \in P | Cost(s) < C, \quad (6.1)$$

where w_s^* represents s ideal weights following training. Given that obtaining w_s^* an exhaustive training process is often prohibitively expensive, the one-shot NAS provide

definition to a super-net with weights w that covers the whole search space. Following training, w is applied as a sub-set weight $w_s \in w$ performance estimator of s . With this method, shared weights are used to evaluate each sub-net, and the objective of the training is given in Eqn. 6.2 as:

$$\arg \min_w \frac{\sum \mathcal{L}_{train}(s, w_s)}{|A|}, \forall s \in A, \quad (6.2)$$

In order to improve the accuracy, the Meta Assistant Network (MAN) is being utilized which tends to remove the superfluous super-net features. MAN, not only improves subnetwork learning but also has the potential to introduce bias into evaluations because of the addition of a latest network component and comprehensive increase in representational capacity. There might be an impact on the relationship, particularly for smaller architectures, between the estimated performance of the subnetwork based on shared weights and its actual performance during deployment.

This incorporates a second stage that eliminates MAN for assessment, which in turn guarantees search fairness and production of the shared weights that tends to be independent of MAN. To be more precise, Annealing process is introduced in order to gradually replace $w_{H_\theta}^{s,l}$ with w_{fix} and a constant channel-wise scalar weight, $w_{fix} = 0.5$. One way to describe such a process is:

$$\arg \min_{w, \theta} \frac{\sum \mathcal{L}_{train}(s, (\lambda w_{fix} + (1 - \lambda) w_{H_\theta}^{s,l} * w_s))}{|A|} \quad (6.3)$$

where the annealing temperature, represented by λ , gradually increases to 1 from 0.

Even after the MAN is eliminated, it has been discovered that the linear annealing schedule still performs well and keeps w operative at a good level, leading to a comprehensive improvement in performance estimations solely depending just upon w .

6.2.2 Swin Transformer Encoder

A 3D token with a patch resolution of (H', W', D') has a dimension of H', W', D', S' , assuming that the encoder's input is a sub-volume $\chi \in \mathcal{R}^{H', W', D', S'}$. A series of 3D tokens with size $\frac{H}{H'} \times \frac{W}{W'} \times \frac{D}{D'}$ are produced by the patch partitioning layer and projected into a C-dimensional space by an embedding layer.

At layer ℓ , 3D token is divided evenly into $\left\lfloor \frac{H'}{M} \right\rfloor \times \left\lfloor \frac{W'}{M} \right\rfloor \times \left\lfloor \frac{D'}{M} \right\rfloor$ windows using a window of size $M \times M \times M$. The encoder blocks in layers ℓ and $\ell + 1$ have their outputs determined by the shifted windowing mechanism as in

$$o^\ell = WMSA \left(LN(o^{\ell-1}) \right) + o^{\ell-1} \quad (6.4)$$

$$o^\ell = MLP \left(LN(o^\ell) \right) + o^\ell \quad (6.5)$$

$$o^{\ell+1} = SWMSA \left(LN(o^\ell) \right) + o^\ell \quad (6.6)$$

$$o^{\ell+1} = MLP \left(LN(o^{\ell+1}) \right) + o^{\ell+1} \quad (6.7)$$

where, MLP stands for multi-layer perceptron, W-MSA stands for regular multi-head self-attention modules, and SW-MSA stands for window partitioning multi-head self-attention modules.

The outputs of W-MSA and SW-MSA are o^ℓ and $o^{\ell+1}$. We then use a 3D cyclic shifting to compute shifted windowing in batches more effectively. Additionally, we determine the self-attention, Att using Eqn. 6.8 as

$$Att(Q, K, V) = SM \left(\frac{QK^t}{\sqrt{d}} \right) V \quad (6.8)$$

where d represents the size of the query and key, and Q , K , and V stand for queries, keys, and values, respectively.

6.3 IMPLEMENTATIONS AND RESULTS

6.3.1 Dataset

One to four raters Brain Tumor Segmentation (BraTS) imaging dataset (manually segmented) using a standardized annotation protocol to ensure consistency and accuracy across the dataset. Experienced board-certified neuroradiologists meticulously examined and approved the segmentation process, ensuring the accuracy of the annotations. The dataset annotation includes the peritumoral edematous/invaded tissue (ED — label 2), necrotic tumor core (NCR — label 1), and the GD-enhancing tumor (ET — label 4) [149].

The dataset used in this study is a superset of the BraTS 2020 data, including 2,640 multiparametric MRI (mpMRI) scans. In the BraTS 2021 data, this was expanded to include 2,000 cases, totaling 8,000 mpMRI scans. This extensive dataset forms a strong basis for training and validating the proposed brain tumor prediction model.

The trained 3D reconstruction scheme's performance parameters are shown in Table 6.1. For each segmented mask, the Dice Similarity Coefficient (DSC) and Hausdorff Distance (HD) were found. The DSC measures the agreement between the predicted and ground truth segmentation, offering a quantitative evaluation of the segmentation accuracy. The Hausdorff distance (HD) evaluates the maximum distance between the margins of the predicted and actual segmentation, offering insights into the precision of the tumor boundary delineation.

The dataset is of high quality due to detailed segmentation and rigorous validation processes, which facilitate the development of a highly accurate and reliable brain tumor prediction model. By utilizing the extensive and well-annotated BraTS dataset, this study aims to advance medical image analysis, ultimately contributing to more effective and early diagnosis of brain tumors.

Table 6.1. Performance of the proposed schemes

Method	Metric	CSF	Gray Matter	White Matter	Average
HyperNet	DSC (%)	87.4 ± 4.2	92.3 ± 3.9	94.2 ± 2.2	91.3
	HD (mm)	10.6 ± 5.6	9.1 ± 5.0	12.4 ± 4.7	10.7
HyperSwinNet	DSC (%)	90.4 ± 4.5	94.7 ± 2.0	91.5 ± 3.0	92.2
	HD (mm)	8.9 ± 3.6	7.7 ± 3.2	10.3 ± 3.9	8.97

6.3.2 Result and Analysis

Patches with dimensions of $96 \times 96 \times 96$ are sampled in the training phase and then augmented by flipping and rotating. With instance normalization, six 3D convolutional layers make up the image encoder M and ReLU activation; $\ell images \mathcal{ER}^{256}$ is the output [149].

The HyperNet \mathcal{H} generates $w_{\mathcal{H}\theta}^{sl} \in \mathcal{R}^{ctot}$; which includes 5 fully connected layers that are separated by ReLU activations, culminating into a Sigmoid activation at the top. The inactivated portion of $w_{\mathcal{H}\theta}^{sl}$ is masked off from back-propagation for each sampled architecture s . Dice and cross-entropy loss are combined to create the training objective function \mathcal{L}_{train} . To improve efficiency throughout the search phase, foreground regions, up to a maximum size of $192 \times 192 \times 192 \text{ mm}^3$, are sampled as well as evaluated on a patch-wise basis. Several heuristics are used in the architecture sampling algorithm's implementation to weed out undesired configurations.

The proposed HyperSwinNet achieves state-of-the-art performance on the BraTS dataset, with an impressive average Dice score of 90.4%. In Fig. 6.2, illustrates the qualitative visualizations. The results demonstrate significant improvement with self-supervised pre-training. The experimental setup involves training models from scratch to address the inherent domain gap in MRI images, focusing primarily on brain tumor segmentation. Fig. 6.3 and Fig. 6.4 display image panels with annotated tumor sub-regions in various MRI scans and 3D reconstructions. These results emphasize the strength and effectiveness of HyperSwinNet in medical image analysis, highlighting its potential to enhance diagnostic and therapeutic results.

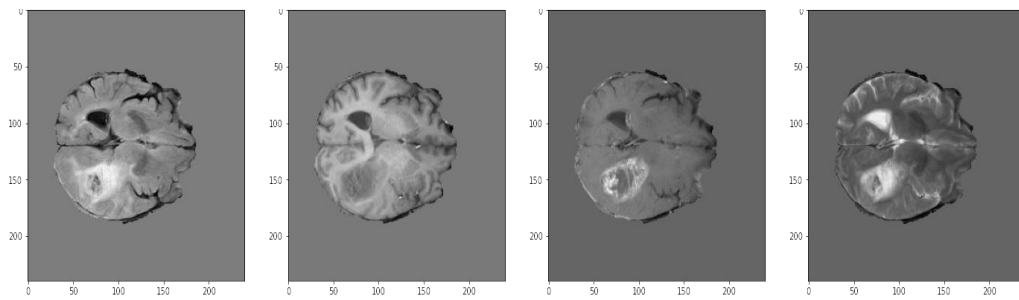


Fig.6.2. Qualitative results of representative BraTS data

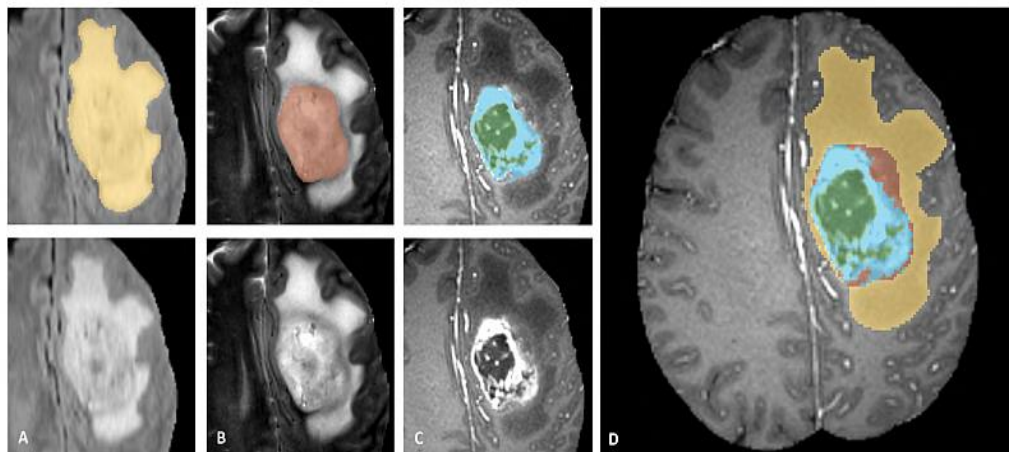


Fig.6.3. Image panels with the tumor sub-regions annotated in the different MRI scans

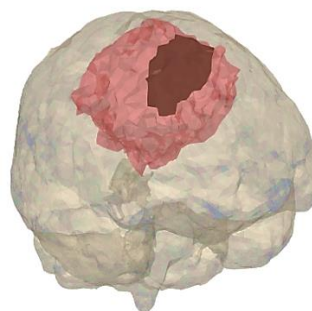


Fig.6.4. 3D reconstruction of tumor sub-region

The results of segmentation are promising, indicating significant progress in achieving reliable results. To further check the authenticity of the model, these segmentation results are tested using ResNet50. Table 6.2 shows the comparative analysis of the proposed work and the recent work done on the BraTS 2021 dataset.

Table 6.2. Comparative analysis of the proposed work and the recent work done on the BraTS 2021 dataset

Ref	Dataset	Technique	A _{cc}	Precision	S _n	S _p	F1 score
[150]	BraTS 2021	StyleGAN2-ADA	75.18%	-	-	-	-
[151]	BraTS 2021	CNN and RNN-based network	-	-	-	-	66.2
[152]	BraTS 2021	Enhanced ResNet10	66.5		0.783	0.533	-
[153]	BraTS 2021	Quantum Model	97.04	0.95	0.96	-	0.98
Proposed Method	BraTS 2021	HyperSwinNet	97.96%	0.97	0.98	0.96	0.98

The result showed that the proposed model demonstrates state-of-the-art performance on the BraTs dataset with 97.96% accuracy excelling in precision, sensitivity, and F1 score This highlights its potential for accurate and efficient brain tumor segmentation compared to alternative methods.

6.4 DISCUSSION

The Precision and efficiency of the brain tumor identification and its treatment planning has been significantly improved due to the integration of the Advanced Computer Aided diagnostic technologies. The combination of the segmentation and rendering technique presented in this study gives way to the self-supervised learning framework which is designed to address the challenges within the manual MRI analysis. Since the transformer-based encoder is utilized by the MAN (Meta Assistant Network) and the HyperSwinNet architectures to derive the features from the 3D Medical Images, for which this innovative approach thus tends to leverage this architecture.

There has been a substantial improvement in segmentation accuracy, essential for effective treatment strategies due to the proposal of this HyperSwinNet architecture. The given architecture consists of an encoder that tends to process and extract multi-resolution features by HyperNet, thereby leading to a major improvement in the segmentation's granularity and precision. Leveraging the hierarchical

architecture of SwinNet has given highly accurate segmented output as the shifted window self-attention mechanism helped in capturing multiscale features, providing context-aware information and exact localization of tumors.

The utilization of the Neural Architecture Search into this framework has led to further improvement in the performance due to the ensured of optimal sub-network identification. Therefore, this study tends to give an outlook regarding the importance of advanced machine learning techniques in medical imaging, thereby giving a future scope for improvements in the diagnostics and treatment of not only in brain tumors but also in the cardiac, hippocampal, and prostate imaging tasks.

6.5 SUMMARY

This research tends to provide an innovative approach to diagnosing brain tumors by utilizing the integration of segmentation and rendering techniques, therefore giving way to the self-supervised learning framework for the re-building of 3D medical images. This framework overcomes the limitations of conventional MRI analysis, which is time-consuming as well as prone to human errors, by introducing the Hypernet structure called as Meta Assistant Network and 3D transformer-based architecture called as HyperSwinNet. The combination of these components tends to give further way to the extensive search space of segmentation networks and improve feature extraction through a pre-trained encoder. This framework provides a cutting-edge result by showcasing a major advancement in brain tumor diagnostics and is validated using the BraTs datasets. The segmentation accuracy is optimized and enhanced by utilizing the NAS which identifies optimal sub-networks in Hypernet while HyperSwinNet employs a 3D Swin Transformer encoder to extract multi-resolution features from MRI scans. The implementation consists of instance normalization, 3D convolutional layers, and a combination of Dice and cross-entropy loss for training objectives. Results showcased substantial improvements, with the proposed method achieving a Dice Similarity Coefficient (DSC) of 90.4% and a Hausdorff distance (HD) of 8.9 mm, concluding that the HyperSwinNet framework

significantly improves the brain tumor segmentation accuracy and efficiency as well as holds potential for application in other medical imaging tasks.

CHAPTER 7

BRAIN TUMOR IMAGE PIXEL SEGMENTATION AND DETECTION USING AN AGGREGATION OF GAN MODELS WITH VISION TRANSFORMER

7.1 PREAMBLE

Detection and classification of brain tumors are done using image pixel segmentation along with a vision transformer in this chapter. A tumor in the brain denotes the emergence of abandoned brain cells, whereas a tumor in the body reflects the uncontrolled growth of cancerous cells. A brain tumor could be cancerous or benign. Although benign brain tumors are not cancerous, they share structural similarities with malignant (heterogeneous) tumors, which are composed of active (cell) cells. To accelerate the treatment process, plan, and increase patient survival, brain tumors must be classified according to their type. MRI is considered to be a vital tool for the diagnosis of tumors, as it allows radiologists to assess the tumor's anatomical position and size. Brain tumor identification and magnetic resonance imaging (MRI) segmentation are challenging but essential challenges for many medical analysis applications. Many automated methods based on magnetic resonance imaging have been developed to identify tumors and categorize various types of tumors. There are some issues associated with the presently employed methods. These are present because brain MRI consists of the non-uniform region and its segmentation takes an extended period which makes it a difficult process. Also, the segmentation and classification of tumors are challenging because tumors vary in size, location, and appearance and have complex features. Inspired by the challenges presented by the conventional methods, a novel strategy is employed to segment and categorize the brain tumor. The aim is to develop a model for tumor classification based on MRI. Since every type of brain scan has unique information about the characteristics of the individual tumor components, we first propose employing pixel segmentation along

with a normalization preprocessing technique to create an adaptable and efficient brain tumor classification network. When it comes to the creation of synthetic images, Generative Adversarial Networks (GAN) have an advantage in several sectors. Using multiple GANs can make the model very complex and confusing, which is also harder to train. While Standalone GAN only retrieves the locally specified feature. So, we used a standalone GAN and Vision Transformer (ViT) together to improve the model's ability to detect tumors in medical images, so that we could obtain both locally and globally specified features in a single model. It will also enhance the images' resemblance more significantly. The constraints of minimal discrimination ability, large computational time, and limited data can be effectively addressed by our model. In addition, it is capable of understanding the variation in information between different representations of original images and getting better accuracy and computation time. BraTS 2020 and Masoud2021 datasets, i.e. a combination of SARTAJ, FigShare, and BR35H were used to evaluate the proposal model. The results obtained show that the proposed model can produce high-quality images with an accuracy of 0.9765 and sensitivity values of 0.977 on BraTS 2020 as well as 0.9899 and 0.9683 on Masoud2021 data sets. A deep learning model, which has been suggested for the accurate detection of brain tumors via scanning methods, is shown in Fig. 7.1. Two modules, namely image enhancement as well as training and validation of classification, are used in the proposed approach to diagnose tumors accurately.

7.2 METHODOLOGY

From the literature, some issues are identified in the traditional methods used for brain tumor classification. One of the most challenging tasks to evaluate in medical image processing can be identifying brain tumors. Since brain tumors can take on various kinds of shapes and textures, the identification task is difficult because there is a lot of variation in the images. Also, using neuroimaging techniques to determine a tumor's specific phase can be difficult and time-consuming. This is because it may be challenging to appropriately differentiate the tumor's sub-regions. After all, their intensities can resemble those of the nearby healthy tissues. Properly identifying and

separating irregular parts of the brain takes time, which makes brain segmentation in MRI images challenging. Also, tumor positions in medical images can be accurately identified by deep neural networks without any pre-processing technique. However effective network training may be hampered by the scarcity of appropriately labeled datasets, which is a potential drawback of this strategy. Thus, a new approach is required to segment and categorize brain tumors. For the extraction of deep features, a novel segmentation technique in combination with the vision transformer has been employed in the present research. Fig.7.1 depicts the suggested deep learning model for accurately detecting brain cancers. The training and validation modules for classification and image enhancement are the two modules used in the suggested framework for appropriate cancer diagnosis.

In MR images, the intensity levels are known to vary extensively and are not predefined. In this work, normalization is utilized in MRI pre-processing. To get a similar set of data that can verify a uniform convergence, normalization is done. To achieve enhanced outcomes, min-max normalization is used for scaling intensity values during the MRI preprocessing phase of the process. Numerous techniques of data augmentation are used, including rotation, random flipping to the left or right, and brightness and contrast alterations. In addition, a super-pixel segmentation method is done on the pre-processed image and that is used to segment the tumor. Following this, features are obtained through the creation of sequences of image patches containing nearby location pixels with similar attributes such as color, brightness, and texture. These segments are then used to generate GAN features. Lastly, the classification process is completed using a vision transformer as shown in Fig.7.1.

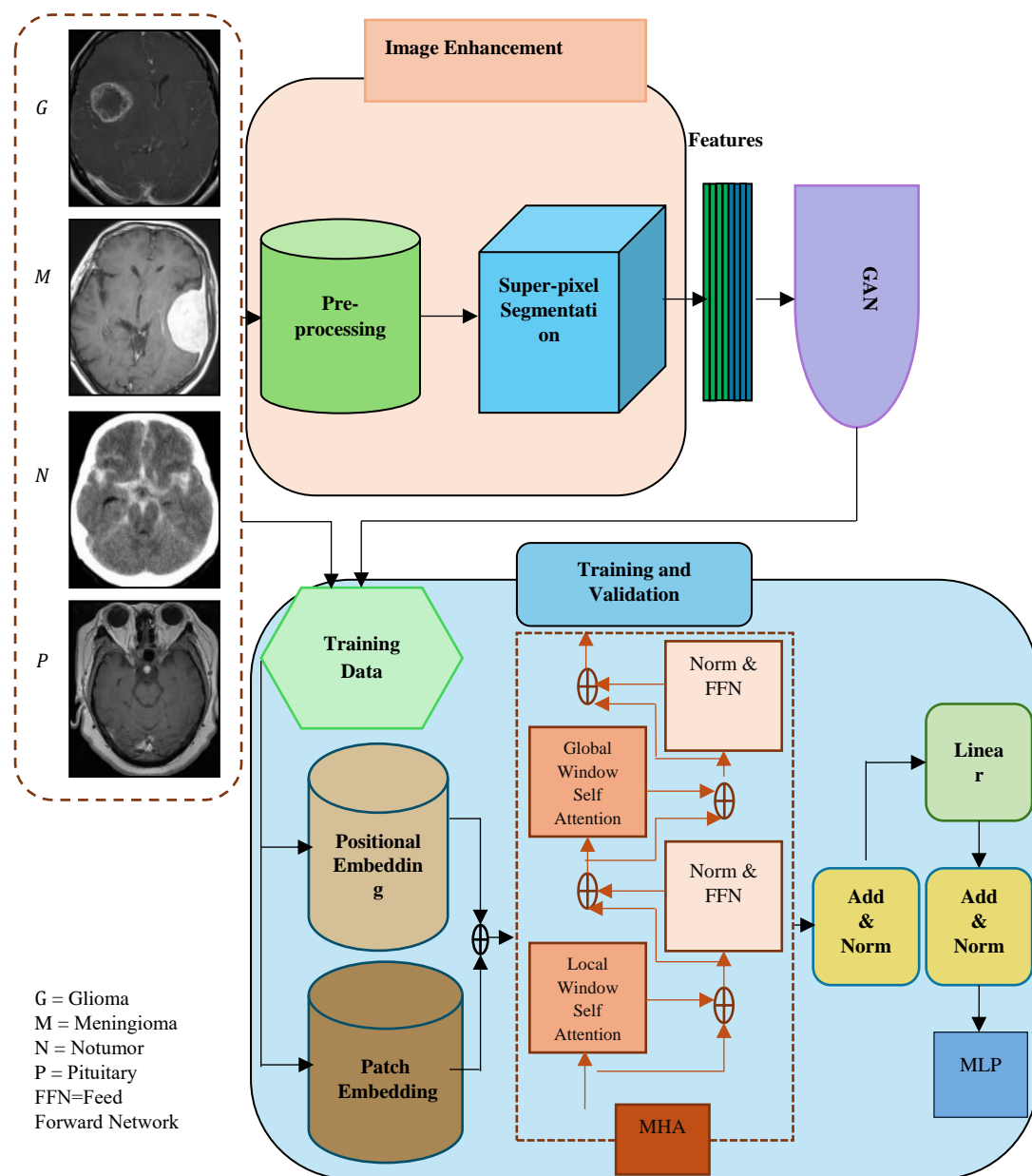


Fig.7.1. Training and testing framework of our suggested model

7.2.1 Step-1: Dataset Collection

There are a lot of publicly available datasets that can be used for study on different types of medical image analysis, however most of the time these data are too limited or inaccessible to the researcher because of privacy concerns. Within this field of study, augmenting the data is frequently employed to solve the issue of small datasets. Although, it is possible to create new images by employing basic techniques

for image processing. However, such approaches are unable to provide the generated images with sufficient detail to appear natural, they may not be especially useful in the discipline of medical diagnostics. The investigations are carried out on the Masoud2021 datasets [43] to validate the suggested model. We have also used the BraTS2020 dataset to verify the robustness of the developed model. Masoud2021 comprises 7023 MR images, that were created by combining three datasets i.e. Figshare [44], SARTAJ [154], and BR35H [155]. The MRI images in this dataset represent a particular category of healthy brains i.e. notumor images along with 3 classes of brain malignancies, encompassing gliomas, meningiomas, and pituitary tumors. 369 labeled training images and 125 non-labeled validation 3 D MR images, including T1, T1-W, post-contrast T1-W(T1gd), T2-W, and T2-flair images, constitute the BraTS2020 [38] [39] [40] dataset. In addition, trained neuroradiologists labeled the peritumoral edema (ED), enhancing tumor (ET), and necrotic and non-enhancing tumor core (NET) for each investigation. There are fifty-five brain slices in all, with each measuring 240×240 pixels within these three-dimensional MR scans.

7.2.2 Step-2: Pre-processing

Let us consider a set of m images representing a brain tumor, I .

$$I = \{I_1, I_2, \dots, I_k, \dots, I_m\} \quad (7.1)$$

where I_k denotes the k^{th} input image and m denotes the total number of images.

To prepare an MRI for training, preprocessing is a crucial step. The use of MRI preprocessing is vital since it reduces distortions and improves the quality of the image. It standardizes the photo for use with various subjects and scanners as well.

Motion artifacts, which are brought through a patient moving during the scan and occasionally resulting from magnetic field inhomogeneities, are among the many artifacts that can impact MRI images. This may result in an impairment in image quality and sometimes introduce errors during further analysis. Therefore, one of the main goals of preprocessing MRI data is to remove distortions and artifacts from the images.

Standardization and normalization techniques are used during preprocessing to ensure uniformity amongst samples and accelerate the processing. A consistent process is implied by pre-processing. The brain image I_k is used to feed the pre-processing in this case, and assessing that image is an important task. Overall, it is seen that the intensity values taken into account in MRIs are highly variable and lack any fixed value. The MRI intensity is highly dependent on the acquisition setting. The magnetic resonance imaging (MRI) intensity is impacted by the acquisition circumstances. Pre-processing refers to having the availability of a uniform procedure. We implemented a preprocessing framework by normalizing the input to attain a similar range of results. So, to prepare for the MRI, normalization of the intensity is essential. Therefore, normalizing the input MRI using min-max normalization allows for the scaling of intensity values in the interval $[0,1]$, which can be expressed in Eqn. 7.2 as,

$$I_n = \frac{(p_n - \min(p))}{(\max(p) - \min(p))} \quad (7.2)$$

where, $\min(p)$ signifies the lowest intensity value and $\max(p)$ signifies the greatest intensity value, and I_n represents the normalized intensity value of the position p_n .

Normalizing intensity makes it possible to achieve desired levels of intensity in a given range and facilitates effective learning. Furthermore, resizing the image is done to accelerate the process and for less memory accessibility. Thus, the obtained pre-processed image is expressed using I'_k .

7.2.3 Step-3: Segmentation

The technique of splitting a digital image into several segments, or groups of pixels (sometimes called super-pixels), is called image segmentation. Segmentation is to transform an image's representation into something which is simpler to comprehend and easy to examine. Pre-processed image, I'_k passed through the segmentation process.

Superpixel approaches use perceptual information to establish similarity indicators that are used to segment an image into groups. By applying the super-pixel

segmentation (SPS) technique, the input image created during the preprocessing phase is split up into unpredictable patches [156].

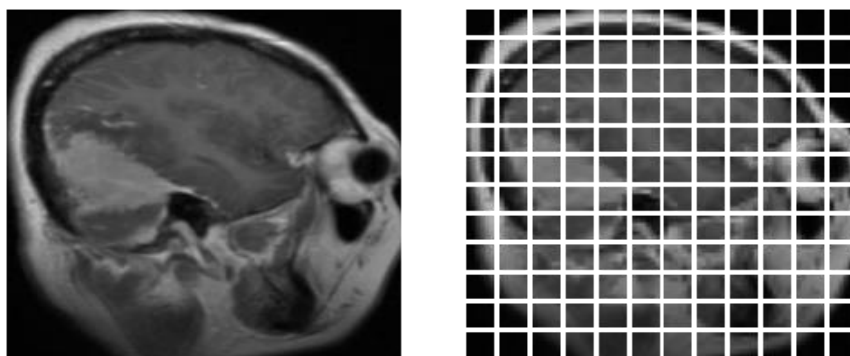
A reference image (R_r) for SPS is chosen based on the multi-exposure input images. Out of all the input images, the study's reference image has the fewest underexposed or overexposed pixels [157]. As illustrated in Fig. 7.2, the R_r reference image has been split into 'a' number of image patches, designated as $R_1, R_2, R_3, \dots, R_a$, by the use of a super-pixel segmentation technique aided by Simple linear iterative clustering (Slic). Through the clustering of pixels according to their colour similarity and closeness in the image plane, this technique creates superpixels.

$$SPS_{Slic}(R_r) = \{R_1, R_2, R_3, \dots, R_a\} \quad (7.3)$$

The following can be used to represent the image patches that were obtained through segmentation:

$$\{\{R_y(R_a)\}_{y=1\dots Y, a=1\dots m}\} \quad (7.4)$$

The a^{th} image patch of the y^{th} input image is indicated here by the notation $\{R_y(m_a)\}$. To ensure uniformity, every other input image is segmented using the same method as the reference image.



(a) Input Image

(b) Segmented Image

Fig.7.2. Super-pixel segmentation

7.2.4 Step-4: Creating Synthetic Images

Generative adversarial networks (GANs) have shown to be particularly effective for data preprocessing in a wide range of computer vision tasks since the influential research done by Goodfellow et al. [158].

GANs have demonstrated a dedication to generating remarkably realistic images even in the absence of a well-defined objective function. The GAN generator is also capable of learning from even minor differences in the data. Two models comprise the GAN deep learning framework: the generative model G_N and the discriminative model D_s . The generative model captures the distribution of the data. The discriminative model calculates the likelihood that the sample will be selected from the training set as opposed to the generative model; see Fig. 7.3. An adversarial process is used to simultaneously train the two models. The architecture is based on game theory and is equivalent to a two-player minimax game [158]. The generator aims to deceive the discriminator by making increasingly realistic samples during training, while the discriminator is trying to make sure that synthetic samples are accurately identified. The goal of G_N training process is to increase the likelihood that D_s will make a mistake. Encouraging the discriminator (D) to commit an error in differentiating between generated and real data samples is the main objective of the GAN training process.

A multilayer perceptron with characteristics β that shows a mapping to the data space is used to represent the differentiable function, $G_N(\epsilon, \beta)$ by the multilayer perceptron. A prior is defined on the independent random noise variable, ϵ to learn the generator's distribution G_d over the data space, x . The discriminator, $D_s(x, \beta)$ is a neural network that receives a sample from either the real training dataset or the synthetic dataset generated by G_N , and it outputs a single scalar value, indicating that the input data is from the real training dataset.

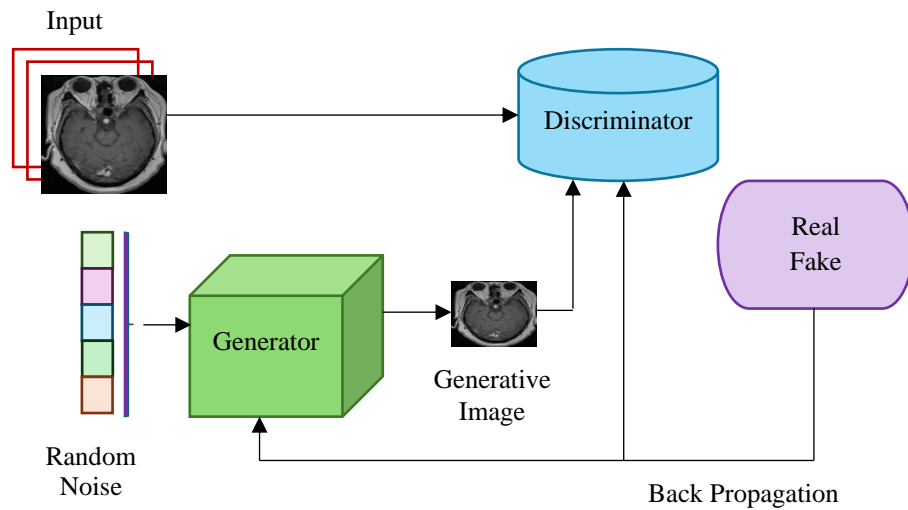


Fig.7.3. Block diagram of a generative adversarial network

The goal of the training process is to have the discriminator D_s maximize the likelihood of correctly labeling the training instances and generating samples from G_N . In order for D_s to be unable to tell them apart from real data, G_N is trained to produce data samples that are analogous to the real dataset. It is described as a two-player minimax game with the value function $v(G_N, D_s)$, which is defined in Eqn.7.5 as

$$\frac{\min_{G_N} \max_{D_s} v(G_N, D_s) = E_{x \sim G_{Ndata}(x)} [\log D_s(x)] + E_{\epsilon \sim G_{Ndata}(\epsilon)} [\log(1 - D_s(x))]} \quad (7.5)$$

where, the input random noise (ϵ) is the input, and x is the real data. Here, the likelihood of successfully categorizing actual data sample is maximized by the first term, while the likelihood of correctly classifying false samples is maximized by the second. Additionally, the goal is to motivate the generator to generate samples that are practically identical to real data because doing so reduces the likelihood that the discriminator will accurately identify them as being fake.

Real data distribution and input noise distribution are represented, respectively, by the variables data and $G_{N(\epsilon)}$. The alleviation of the Jensen-Shannon divergence between the distribution's data and a different distribution G_d , derived from ϵ and G_N , can be used to express this. While $G_N(\epsilon)$ denotes the mapping to synthesize the real data, $D_s(x)$ denotes the likelihood that x came from the real data. The generator, G appears more complex due to its deep neural network, more convolutional layers, and

non-linearities. Upsampling the noise vector, ε occurs as backpropagation is used by G_N to learn the weights. Ultimately, the discriminator labeled the data produced by the generator as real.

In GAN, both real and fake images are used during training to improve the generator's ability to create realistic images. However, during validation, it is important to only use real images to evaluate the generator's performance accurately. This is because using fake images during validation can lead to overestimating the generator's performance as it may have memorized the training data rather than learning to generate new images.

7.2.5 Step-5: Classification

In the end, we used the vision transformer on the GAN results. The input is initially converted into an array of image patches in order to be compatible with the transformer's input form. Instead of focusing on individual pixels, ViTs work with image patches. Fixed-size patches are extracted from the input image and then linearly embedded into high-dimensional vectors. The encoder in the Transformer receives these embeddings as input. To maintain positional information, position embedding has been integrated into patch embedding as well, producing a series of embedding vectors. Consequently, the following sequence of steps is used to transform the input:

$$y_0 = [\mathcal{X}_{class}; \mathcal{X}_p^1 \alpha; \mathcal{X}_p^2 \alpha; \dots \dots \mathcal{X}_p^N \alpha] + \alpha_{pos} \quad (7.6)$$

where, $\alpha_{pos} \in \mathbb{R}^{(N+1) \times D}$ denotes trainable positional embedding vector and $\alpha \in \mathbb{R}^{(p^2 \cdot C) \times D}$ denotes a trainable embedding vector. Further, y_0 provides as input to the encoder.

Position-wise feed-forward networks are placed after a multi-head self-attention mechanism in each layer of the Transformer encoder. A typical self-attention technique captures global dependencies by having each token pay attention to every other token. With longer input sequences, nevertheless, this method becomes computationally costly.

Local Window Self-Attention (*Att*) is incorporated by ViTs to solve this problem. In *Att*, each token only manages a fixed-size window-delimited local neighborhood within the image. Because of this local attention mechanism, ViTs may expand to greater input sizes more effectively and with less computing complexity. Tokens are capable of capturing spatial dependencies in an image while keeping their computing cost under control by focusing on adjacent patches.

Even while *Att* is good at capturing local information, it could have trouble capturing long-range dependencies in the image. ViTs additionally make use of Global Window Self-Attention (*Att*) to get around this restriction. In *Att*, each token manages a global context window that covers the entire image, in along with the surrounding patches. Tokens are capable of capturing long-range dependencies and contextual information that may cover the entire image because of this global attention approach. ViTs can efficiently represent both local and global image features because of the combination of Local Window Self-Attention (*Att*) and Global Window Self-Attention (*Att*). So, we combine both global window self-attention (*Att*) and local window self-attention (*Att*) in the proposed work. Local window self-attention improves the local information by considering all of the spatial positions. Conversely, global window self-attention maintains computational efficiency while capturing global contextual information [159]. Let's assume that the inputs to the transformer encoder are Query Q, Key K, and Values V. The multi-head attention module divides the Q, K, and V parameters in N different ways and then separately feeds every split via a distinct Head independently. Then, a final Attention score is generated by combining all of these related Attention calculations. Eqn. 7.7 to 7.10 are used to denote the multi-head attention module as follows:

$$[Q, K, V] = y \cup_{Q,K,V}, \cup_{Q,K,V} \in \mathbb{R}^{D \times D_h} \quad (7.7)$$

$$A = \text{softmax} \left(\frac{QK^T}{\sqrt{D_h}} \right) V, A \in \mathbb{R}^{N \times N} \quad (7.8)$$

$$\text{Att}(y) = AV \quad (7.9)$$

$$\text{MSA}(y) = [\text{Att}_1(y); \text{Att}_2(y); \dots \dots \text{Att}_K(y)] \cup_{msa} \cup_{Q,K,V} \in \mathbb{R}^{K \cdot D_h \times D} \quad (7.10)$$

Vision Transformer layers with multilayer perceptron (MLP) and Multi-head attention (MHA) approach can be written as shown in Eqn. 7.11 and Eqn.7.12.

$$V_T' = MHA(L_N(V_T^{n-1}) + V_T^{n-1}) \quad (7.11)$$

$$V_T = MLP(L_N(V_T'^{n-1}) + V_T'^{n-1}) \quad (7.12)$$

V_T' and V_T represents the corresponding output global and local features of the MHA and MLP modules in the used Vision Transformer block. The terms Layer Normalization and Multilayer Perceptron Module are referred to as L_N and MLP, respectively.

The multi-head attention module's output is normalized. Furthermore, this leverages modelling capabilities and also improves the accuracy of the proposed work.

$$\mathcal{U} = norm(MSA(y)) \quad (7.13)$$

Further transforming the features and classifying the output as detected Gaussian Error Linear Unit (GELU) activation function is used.

7.3 IMPLEMENTATIONS AND RESULTS

The experiments carried out based on the proposed work and the results obtained are discussed in this section.

7.3.1 Implementation Details

To identify and retrieve features from images, we train the model from end to end. The resolution of the input images has been changed to 124×124 , and the patch size has been adjusted to 9. There are twelve transformer layers in all. The whole algorithm has been trained via the Adam optimizer and the binary cross-entropy loss over datasets with an average rate of learning of 2×10^{-5} . To assess the effectiveness of the suggested model, the Receiver Operating Characteristic (ROC) and Area Under the ROC (AUC) curves are generated, and the accuracy of the proposed model is estimated.

7.3.2 Performance Evaluation

The performance is evaluated through other available methods in this subsection using both qualitative and quantitative analysis, with the outcomes being compared with the suggested algorithm. In this case, training takes up 70% of the dataset, with the remaining 30% going toward testing. Before being sent through the segmentation process, the images undergo initial preprocessing. The tumor zones need to be divided, however, occasionally they develop irregular overcrowded, and noisy patterns.

To accurately detect and classify the tumor, GAN is employed in conjunction with a vision transformer. A confusion matrix, training and validation accuracy, and training and validation loss are used for assessing this model. Fig. 7.4 (a) and Fig. 7.4 (b) illustrate the classification results for the Masoud 2021 dataset using the confusion matrix and ROC curve, respectively.

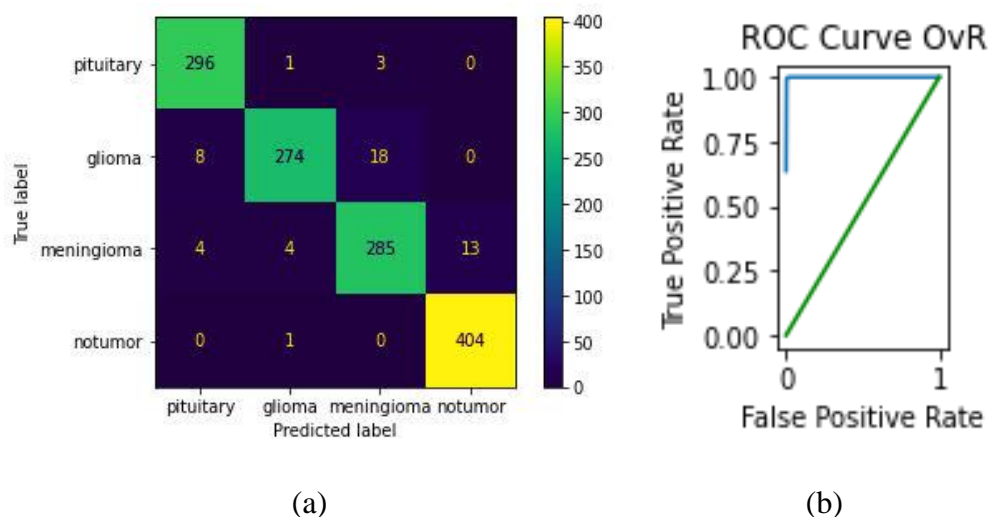


Fig.7.4. (a) Confusion matrix

(b) ROC curve

Additionally, the outcomes of the suggested work on the Masoud2021 dataset with the training-testing accuracy and loss for 100 epochs are shown in Fig. 7.5 (a) and Fig. 7.5 (b). With this dataset, the lowest train/validation loss and the highest total train/validation accuracy are 0.009/0.012 and 0.9899/0.9582, respectively.

We used the BraTS 2020 dataset to further examine the stability of the suggested model after achieving promising outcomes on the Masoud 2021 dataset. We were able to attain a minimum train/validation loss of 0.083/0.195 and a maximum overall train/validation accuracy of 0.9765/0.9498. The outcomes of the suggested approach on the BraTS 2020 dataset, including the training-testing accuracy and loss for 30 epochs, are shown in Fig. 7.6 (a) and Fig. 7.6 (b).

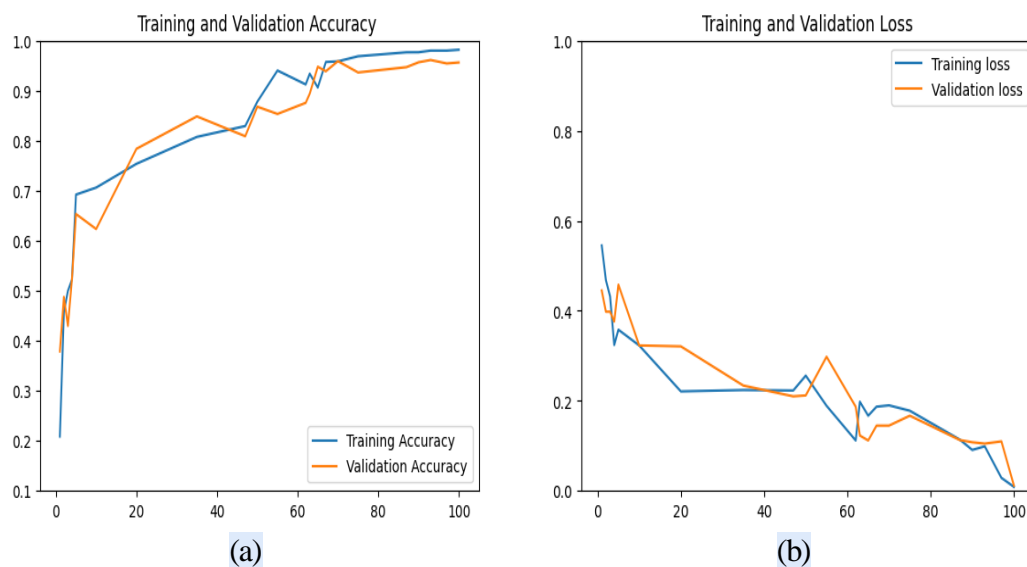


Fig.7.5. For Masoud2021 dataset: (a) Training-Testing accuracy curve, (b) Training-Testing loss curve for proposed work

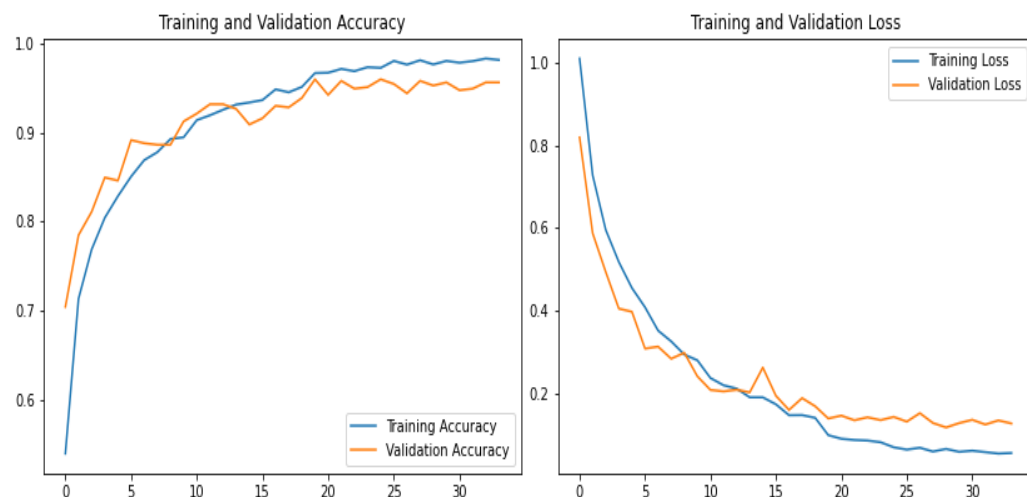


Fig.7.6. For BraTS2020 dataset: (a) Training-Testing accuracy curve, (b) Training-Testing loss curve for proposed work

Table 7.1 Summarizes the performance assessment

Fold	Classes	No. of Samples	Performance Measure				
			Acc	S_n	Precision	F1-Score	S_p
Fold I	glioma	1642	0.870	0.924	0.886	0.908	0.912
	meningioma	1650	0.909	0.860	0.940	0.87	0.925
	No-tumor	2000	0.994	0.980	0.991	0.96	0.993
	pituitary	1765	0.989	0.990	0.943	0.949	0.994
Fold II	glioma	1642	0.910	0.946	0.906	0.928	0.956
	meningioma	1650	0.899	0.896	0.950	0.90	0.906
	No-tumor	2000	0.974	0.977	0.989	0.93	0.986
	pituitary	1765	0.969	0.968	0.973	0.966	0.988
Fold III	glioma	1642	0.896	0.947	0.899	0.910	0.927
	meningioma	1650	0.917	0.892	0.922	0.893	0.919
	No-tumor	2000	0.988	0.990	0.997	0.97	0.991
	pituitary	1765	0.989	0.983	0.961	0.955	0.992
Performance measures (in term of mean \pm standard deviation (σ))	glioma	1642	0.892 ± 0.020	0.939 ± 0.013	0.897 ± 0.010	0.915 ± 0.011	0.932 ± 0.022
	meningioma	1650	0.908 ± 0.009	0.883 ± 0.020	0.937 ± 0.014	0.888 ± 0.016	0.917 ± 0.01
	No-tumor	2000	0.985 ± 0.010	0.982 ± 0.007	0.992 ± 0.004	0.953 ± 0.0208	0.99 ± 0.004
	pituitary	1765	0.982 ± 0.012	0.980 ± 0.011	0.959 ± 0.015	0.957 ± 0.009	0.991 ± 0.003

To test the suggested model, the dataset is shuffled three times, as shown by the folds in Table 7.1—fold I, fold II, and fold III. As indicated by Table 7.1, the no tumor class obtained the best results for the suggested approach using the Masoud2021 dataset, with values of 0.994, 0.990, 0.997, 0.97, and 0.993 in terms of accuracy, sensitivity, precision, F1-score, and specificity respectively. Based on the DL approaches utilized for brain tumor identification, Table 7.2 presents the comparison with results from related studies for Masoud dataset and Table 7.3 presents the comparison with results from related studies for BraTS2020 dataset.

Tables 7.2 and 7.3 compare the outcomes of the proposed model with the relevant studies on deep learning techniques used in brain tumor detection. We summarize the relevant studies that used our dataset i.e. Masoud and BraTs dataset. It

includes metrics such as accuracy, sensitivity, specificity, precision, F1 score, and ROC to provide a comprehensive comparison of model effectiveness

Table 7.2 Comparison with relative studies for Masoud dataset (contain three datasets—Figshare, SARTAJ, and BR35H)

Ref	Dataset	Technique	Acc	S_n	S_p	Precision	F1-Score	M_{cc}^*	$C_A^\#$	ROC
[160]	Figshare datasets	17-layered CNN, Mobile NetV2 & M-SVM	98.92 %	98.82 %	99.02%	-	-	-	-	-
[161]	Figshare and Kaggle (Navone el) dataset	DenseNet-41 based Mask-RCNN and ResNet-50 based Mask-RCNN	96.3 %	95.3 %	-	-	-	-	-	-
[162]	Masoud dataset	CNN	97.84 %	97.85 %		97.85 %	97.90 %			
[163]	Masoud2 021 dataset	Various CNN models	Inception V3: 97.126 %	Inception V3: 96.59 %	Inception V3: 99.98%	Inception V3: 97.97 %	-	-	-	Inception V3: 99.84 %
Proposed Technique	Masoud2 021 (contain Three datasets — Figshare, SARTAJ, and BR35H)	GAN with ViT	Train: 98.99 % Test: 95.82 %	96.83 %	98.56%	94%	92.7 %	91.4 %	89.3 %	99.2 %

* M_{cc} – Mathew’s Correlation Coefficient

C_A – Cohen’s Kappa Coefficient

Table 7.3 Comparison with relative studies for BraTS dataset

Ref	Dataset	Technique	Acc	S_n	S_p	F1-Score
[160]	BraTS 2018	17-layered CNN, MobileNetV2 & M-SVM	97.47%	97.22%	97.94%	-
[164]	BraTs 2018, BraTs 2019, BraTs 2020 dataset	CNN model named EfficientNetB0	BraTs 2018: 94.89%, BraTs 2019: 95.14%, BraTs 2020: 95.94%	-	-	-
[165]	BraTs dataset	Semantic segmentation +CNN	95.7%	-	-	93%
[96]	BraTs dataset	CNN model named VGG16 and VGG19	BraTs 2015: 97.8%, BraTs 2017: 96.9%, BraTs 2018: 92.5%	-	-	-
[95]	BRATS 2018	3D Deep CNN	96.49%	-	-	-
[166]	BRATS 2015 [167]	CNN	98.49%	-	-	-
Proposed	BraTS 2020	GAN with ViT	Train: 97.65% Test: 94.98%	0.977	0.9606	-

7.4 DISCUSSION

The proposed methodology uses advanced machine learning and image processing techniques to improve the detection and classification of brain tumors from MRI images. This discussion explores the effectiveness of the techniques used, compares them with existing methods, and suggests potential areas for future research. The integration of Generative Adversarial Networks (GANs) and Vision Transformers (ViTs) in the proposed framework addresses several critical challenges in brain tumor detection. GANs are adept at generating high-quality, realistic images which, when combined with ViTs, improve the model's ability to capture both local and global features from medical images.

The combination of GANs and ViTs is particularly noteworthy when compared to methods that only rely on CNNs or traditional segmentation techniques. CNNs are usually constrained by their local receptive fields, which makes it difficult to capture the global context in images. On the other hand, ViTs employ self-attention mechanisms that enable a better understanding of global dependencies within the image. This ultimately leads to improved classification accuracy and robustness.

6 The results from the experiments conducted on the BraTS 2020 and Masoud2021 datasets show that the proposed model performs exceptionally well. The achieved accuracy and sensitivity values indicate a substantial enhancement over existing methods. This underscores the effectiveness of the combined GAN and ViT approach in accurately detecting and classifying brain tumors.

7.5 SUMMARY

122 Achieving accurate detection and classification of brain tumors using MRI screening is crucial for medical analysis, but it is also challenging. Our approach involves preprocessing with normalization and pixel segmentation to create an effective tumor segmentation system. Different brain imaging techniques provide unique information about tumor cells. To tackle issues such as limited data, long computation time, and limited discrimination, we employ a Vision Transformer (ViT) with a standalone GAN. This allows us to extract both global and local features and enhance image similarity. Unlike using multiple GANs, a standalone GAN simplifies training and improves model efficacy. This approach reduces computation time and increases accuracy. Our model produced high-quality images with accuracy and sensitivity values of 0.9899 and 0.9683 on the Masoud2021 dataset, and 0.9765 and 0.977 on the BraTS 2020 dataset.

CHAPTER 8

AN AUTONOMOUS AND INTELLIGENT HYBRID CNN-RNN-LSTM BASED APPROACH FOR THE CLASSIFICATION OF ABNORMALITIES IN BRAIN

8.1 PREAMBLE

This work represents the detection and classification of brain abnormalities using MRI testing. The MRI tests are quite advantageous but the only cons of the same is that it represents a lot of data in a single image thereby making the analysis time-consuming and difficult. To overcome this issue, it has been suggested to utilize the multi-mark classification method that takes account of several images exhibiting different abnormalities. Meningitis (MGT), encephalitis (ECP), granuloma (GNL), infract (IF), haemorrhage (HM), and ring-enhancing lesion (REL) are some examples of these abnormalities. In order to identify and categorize these abnormalities RNN and CNN are being utilized. In order to extract the important features from the input signal, the CNN utilizes a channel-wise model while the utilization of the dependency parameter to classify abnormalities is done by RNN. The RNN's LSTM protects against gradient failures.

In order to derive an accurate classification and also prevent the boundary conditions, performance metrics such as mean square error (MSE), and probability of occurrence (POC) are being used. Implementation of the individual CNN and RNN with the LSTM leads to poor performance and further needs laborious calculations. A Hybrid Intelligent controller is the suggested solution against the given problem as it combines the advantages of the CNN and RNN-based LSTM techniques. When compared with the CNN, RNN, RNN-based LSTM, and several other methodologies, the hybrid approach has significantly improved the performance parameters for the accurate classifications of the image related to brain abnormalities. Nearly 1000 image

data have been accumulated from the standard source in order to effectiveness of the hybrid approach.

Brain cancer is considered to be a global health issue and requires utmost attention for the diagnosis using MRIs. As conventional/manual tumor detection is quite intricate and prone to error, therefore automated systems are suggested and are essential. To have a systematic approach to the diagnosis of brain tumors, a model such as the BeatGAN [168] can be used. In order to have an accurate clinical diagnosis and further treatment, the segmentation and identification of the tumor border are required.

The literature identifies the gaps in the medical diagnosis with regard to the abnormalities of the brain. This work covers the diagnosis of specific brain abnormalities such as granulomas, meningitis, encephalitis, haemorrhage, infarct, and ring-enhancing lesions. It showcases the disadvantages of utilizing single-marking methods in MRI reports and therefore proposes a multi-mark classification approach as the solution.

80 The combination of the recurrent and convolutional neural network is referred to as the hybrid topology and is utilized for the multi-mark classification. To prevent gradient problems, RNN with LSTM handles the abnormality classifications while CNN extracts features. The performance metrics such as POC and MSE are being utilized to carry out the evaluation. The hybrid approach is known to be having a significant improvement in the performance parameters for accurate image classification while the CNN and RNN-based LSTM showcases effectiveness. The progress of the proposed model is tested on 1000 data samples, showcasing and validating the effectiveness of the hybrid approach. The literature offers a thorough and methodical approach to brain tumor diagnosis.

8.2 PROPOSED METHODOLOGY

Conventional methods can have multiple pieces of information from a single MRI scan. The resultant of these conventional methods is sometimes inappropriate,

uncertain, and inaccurate in classifications further leading to significant challenges in providing accurate diagnoses. Such inaccuracies pose serious outcomes related to patient care, as ineffective diagnosis or delayed diagnosis can affect the outcomes of the treatment.

A further review of the literature showcases that the single-mark approach scuffles with performance parameters such as probability of occurrence and mean square error.

The proposal of the multi-mark classification approach has been given for diagnosing the abnormalities of the brain. This approach offers multiple images with more than one mark to represent the abnormalities. Also, in order to improve the performance parameters, the hybrid intelligent approach integrating CNN and RNN is suggested.

The work includes:

- Identifying the limitations of the conventional single-mark approach in the analysis of MRIs.
- Proposal of a multi-mark classification approach for classifying brain abnormalities as it offers better performance in classifying the same.
- Development of the hybrid intelligent approach by integrating CNN and RNN with LSTM in order to enhance performance parameters.
- Testing of the suggested hybrid approach on a significant sample size and further comparing its effectiveness with the existing methods.

The goal of this study is to contribute to the field of medical imaging as it showcases the disadvantages of the current MRI analysis methods. This work improves the reliability of brain abnormality detection by proposing the multi-mark classification approach and hybrid intelligent system. The introduced methods have played a vital role in enhancing the diagnostic processes, thereby leading to good patient outcomes.

8.3 PROPOSED NETWORK

In order to detect and classify the abnormalities in the brain, firstly we have used CNN to extract the important features of the input signal based on a channel-wise model. The various stages of CNN which include kernel, padding, and pooling are shown in Fig.8.1. Then the size reduction of images is done as shown in Fig.8.2.

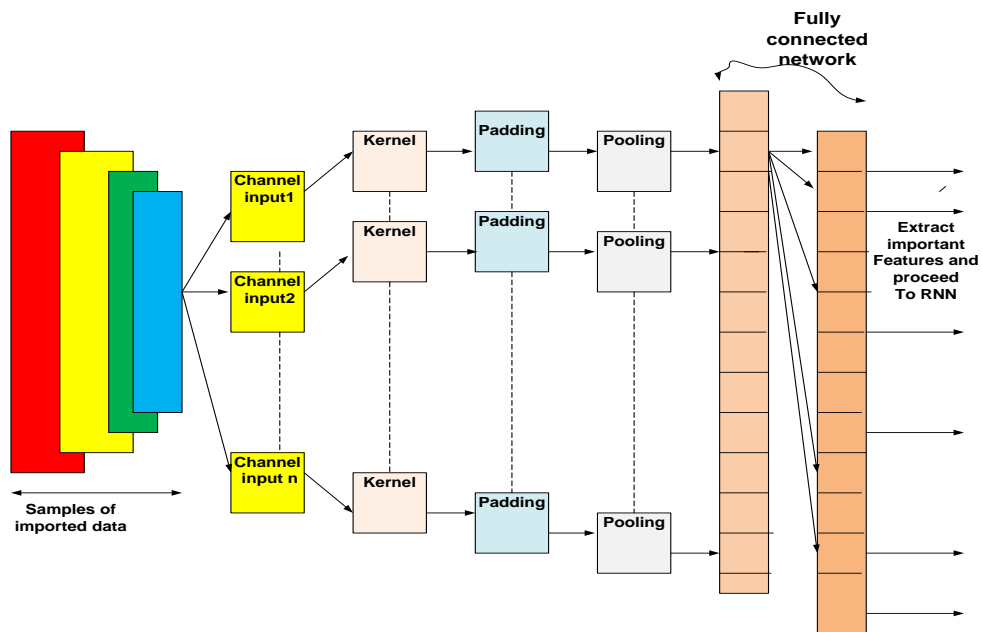


Fig.8.1 Structure of CNN considering the different stages of image classifications

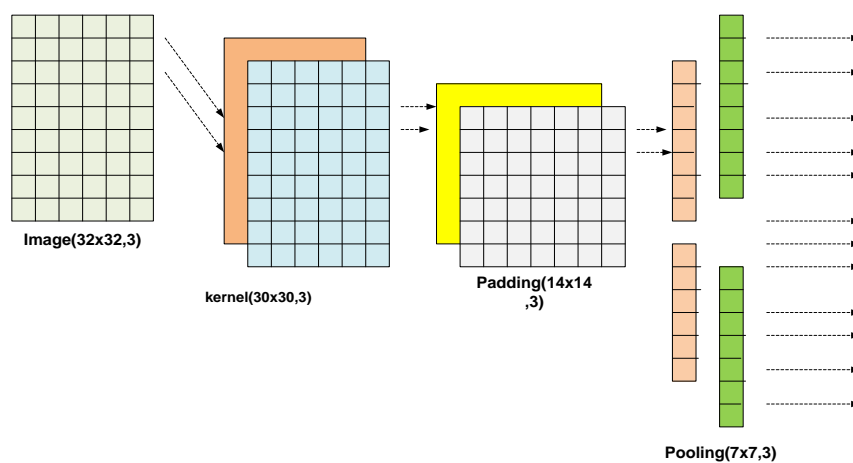


Fig.8.2 Image size reduction for extracting the important feature of determining the abnormalities

After extracting the main features from the image with reduced size, RNN is used to detect various abnormalities of the brain. The sub-section of the proposed network is classified into two parts: RNN and RNN-based LSTM [169][170].

8.3.1 Recurrent neural network (RNN)

One of the primary difficulties encountered with CNN is its inability to generate output that is connected to the time-sequential network. In this type of network, the relationship between sequential inputs and outputs is determined by their respective weights. There are four potential names for RNN structures: one-to-one, one-to-many, many-to-one, and many-to-many. For our work, we have opted for the many-to-many structure of RNN, as depicted in Fig.8.3.

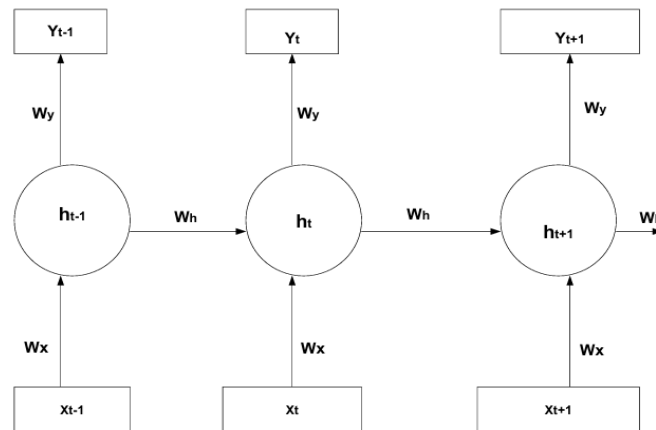


Fig 8.3. Architecture of RNN

The architecture of RNN is designed to resemble that of an Artificial Neural Network (ANN), where the activation function represents the actual output. Fig. 8.3 illustrates how the time sequence nature of three inputs is connected to three outputs using weights (W_x , W_y , W_h). This correlation considers three-time sequences ($t-1$, t , $t+1$). Let's focus on the output at the instant 't', as shown in Eqn. 8.1.

$$Z_h = (X_t W_x + h_{t-1} W_h) \quad (8.1)$$

Actual output is given in the form of the activation function as shown in Eqn. 8.2.

$$h_t = \text{act}(Z_h) \quad (8.2)$$

From the above figure, it can depict that

$$y_t = W_y h_t + b \quad (8.3)$$

The actual output is represented in Eqn.8.4 as

$$\hat{y}_t = \text{Act}(y_t) \quad (8.4)$$

The output at instant 't' will be compared with its preceded time which loses (L) as shown in Eqn. 8.5

$$L = \hat{y}_t - y_{t-1} \quad (8.5)$$

This loss will cause a change in weights as per in steepest descent algorithm as shown in Eqn 8.6.

$$\Delta W = \frac{\eta \partial L}{\partial W} \quad (8.6)$$

The weight adjustments are shown in Eqn. 8.6 has a problem. The gradient may disappear if the weight change is negligible, preventing weight updates. The idea of RNN-based LSTM is applied to solve this issue.

8.3.2 RNN based LSTM structure

The concept of Long Short-Term Memory (LSTM) is integrated with Recurrent Neural Networks (RNNs) to address the issue of gradient vanishing. LSTM effectively manages memory usage by discarding redundant information. This is achieved through the incorporation of various cell states, each assigned specific weights, to prevent gradient failure, as illustrated in Fig. 8.4.

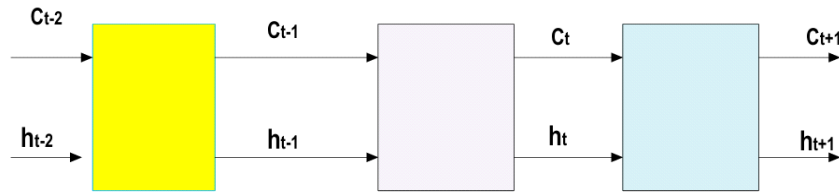


Fig 8.4. The concept of Long Short-Term Memory (LSTM) is integrated with Recurrent Neural Networks (RNNs)

Eqns 8.7 through 8.12 display the general expression of the RNN-based LSTM structure, which is mentioned in [170], [171].

$$f_t = \sigma(W_{fh}h_{t-1} + W_{fx}h_t + b_f) \quad (8.7)$$

$$i_t = \sigma(W_{ih}h_{t-1} + W_{ix}h_t + b_i) \quad (8.8)$$

$$g_t = \sigma(W_{gh}h_{t-1} + W_{gx}h_t + b_g) \quad (8.9)$$

$$o_t = \sigma(W_{oh}h_{t-1} + W_{ox}h_t + b_o) \quad (8.10)$$

$$c_t = f_t c_{t-1} + i_t g_t \quad (8.11)$$

$$h_t = o_t \tanh(c_t) \quad (8.12)$$

Four variables are involved in the basic equation of the RNN-based LSTM: weights (w), bias (b), hidden layer (h), and gate (g). Eqn. 8.13 is a matrix representation of these variables:

$$W = \begin{bmatrix} W_f \\ W_i \\ W_g \\ W_t \end{bmatrix}, b = \begin{bmatrix} b_f \\ b_i \\ b_g \\ b_t \end{bmatrix}, h = \begin{bmatrix} h_f \\ h_i \\ h_g \\ h_t \end{bmatrix}, g = \begin{bmatrix} g_f \\ g_i \\ g_g \\ g_t \end{bmatrix} \quad (8.13)$$

The LSTM's current time step is indicated by t , the input data is represented by x_t , the hidden state is represented by h_t , the cell state is represented by c_t , the input gate is represented by i_t , the forget gate is represented by f_t , and the output gate is represented by o_t . The weights are represented by W , the bias by b , the gate by g , the sigmoid function by σ , and the hyperbolic tangent function by \tanh . Fig. 8.5 shows the structure of the RNN-based LSTM image classification system.

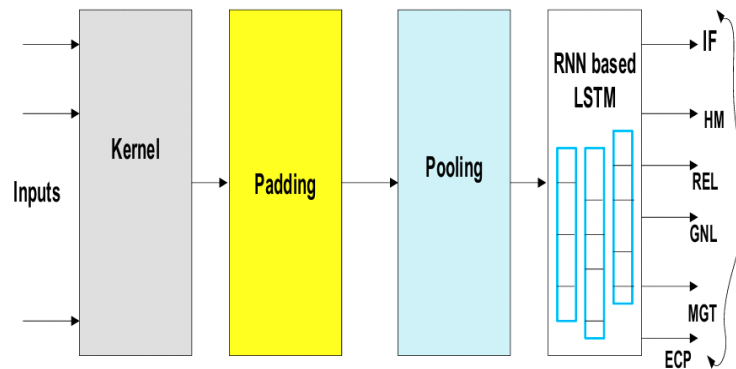


Fig.8.5. RNN –LSTM structure for image classification

8.3.3 Hybrid CNN-RNN based LSTM structure

The finest characteristics of CNN and RNN-based LSTM are combined to create a hybrid approach that is controlled by an intelligent technique (fuzzy logic controller) that improves abnormality classification with optimal performance parameters. Eqns 8.7 through 8.12 demonstrate that the result is a mixture of the weights and hidden layers multiplied together with an additional bias. Nevertheless, the controller design is complicated by the weights and hidden layer multiplication. Logic gates such as AND or OR can typically be used for this kind of multiplication; the AND gate is particularly useful when multiplying two variables. Weights and hidden layers input must be converted into digital form because the main problem with logic gates is their inability to handle multiple inputs in binary form.

Since two inputs demand four binary inputs, a contemporary controller must be used in place of the logic gate. An intelligent fuzzy logic controller is implemented between the two inputs in the proposed model.

Eqns 8.14 through 8.19 show the altered equations that the fuzzy logic controller (FLC) is controlling.

$$f_t = \sigma(W_{fh}(\text{FUZZY})h_{t-1} + W_{fx}(\text{FUZZY})h_t + b_f) \quad (8.14)$$

$$i_t = \sigma(W_{ih}(\text{FUZZY})h_{t-1} + W_{ix}(\text{FUZZY})h_t + b_i) \quad (8.15)$$

$$g_t = \sigma(W_{gh}(\text{FUZZY})h_{t-1} + W_{gx}(\text{FUZZY})h_t + b_g) \quad (8.16)$$

$$o_t = \sigma(W_{oh}(\text{FUZZY})h_{t-1} + W_{ox}(\text{FUZZY})h_t + b_o) \quad (8.17)$$

$$c_t = f_t(\text{FUZZY})c_{t-1} + i_t(\text{FUZZY})g_t \quad (8.18)$$

$$h_t = o_t(\text{FUZZY})\tanh(c_t) \quad (8.19)$$

The first step in implementing the fuzzy logic controller between the hidden layer and weights is to design a particular set of rules. It is observed that there is 0.8 uncertainty with the hidden layer and 0.9 uncertainty with weights. Consequently, five sets of rules will be applied to this uncertainty. Eqn 8.20 and Eqn. 8.21 divide the uncertainty into five parts for the two inputs, W and h , while Equation 8.22 divides the output uncertainty into five parts.

$$W = \frac{0.3}{W_1} + \frac{0.6}{W_2} + \frac{0.9}{W_3} + \frac{0.6}{W_4} + \frac{0.2}{W_5} \quad (8.20)$$

$$h = \frac{0.5}{h_1} + \frac{0.7}{h_2} + \frac{0.8}{h_3} + \frac{0.4}{h_4} + \frac{0.1}{h_5} \quad (8.21)$$

$$\text{Output} = W \cap h = \min(W, h)$$

$$O = \frac{0.3}{O_1} + \frac{0.6}{O_2} + \frac{0.8}{O_3} + \frac{0.4}{O_4} + \frac{0.1}{O_5} \quad (8.22)$$

Table 8.1 Fuzzy membership functions

W/h	NB (0.5)	NS (0.7)	ZS (0.8)	PB (0.4)	PS (0.1)
NB (0.3)	NS (0.6)	ZS (0.8)	ZS (0.8)	PS (0.1)	PB (0.4)
NS (0.6)	NB (0.3)	NS (0.6)	ZS (0.8)	NS (0.6)	PS (0.1)
ZS (0.9)	NS (0.6)	NB (0.3)	NB (0.3)	ZS (0.8)	PS (0.1)
PS (0.6)	PB (0.4)	ZS (0.8)	NB (0.3)	NS (0.6)	PS (0.1)
PB (0.2)	PS (0.1)	NB (0.3)	NS (0.6)	ZS (0.8)	PB (0.4)

As demonstrated above, it is observed that the FLC functions as a controller between the different inputs. The inputs are taken from CNN in their truncated form. Table 8.1 provides the membership functions for the two inputs.

Table 8.2 Updated values of weights and hidden layer

Parameter	1 st value	2 nd value	3 rd value	4 th value	5 th value	6 th value
$W_f(x, h)$	12.5	13.6	14.5	15.6	16.9	18.2
$W_i(x, h)$	13.6	15.6	18.6	19.8	21.3	22.9
$W_g(x, h)$	16.9	18.4	21.3	22.9	25.6	27.8
$W_t(x, h)$	18.7	19.8	25.6	27.8	29.9	30.2
$h_f(x, h)$	15.6	17.5	18.9	20.6	25.6	27.6
$h_i(x, h)$	34.6	37.8	38.9	41.2	45.6	49.5
$h_g(x, h)$	26.9	28.9	31.2	33.5	34.6	35.9
$h_t(x, h)$	28.9	29.8	34.6	38.9	39.5	44.8

A fuzzy logic controller is implemented as an output between the hidden layer (h) and the weights (W). As covered in the above section, the controller and weights are continuously trained. Six sets of weights and hidden layers are needed in order to identify six different kinds of abnormalities. The most recent six values for these weights and the hidden layers that are employed to detect abnormalities are shown in Table 8.2. Eqn 8.22 is used to determine the output (O_t) using the hidden layer and known weight values. Then, a comparison is made between these computed O_t values and the corresponding set values shown in Table 8.3. As further shown in Table 8.3, the type of abnormality is identified by the measured value that is closest to its set value.

The hybrid approach's surface view of the FLC (Fuzzy Logic Controller) is shown in Fig. 8.6. The FLC rules are displayed in Fig. 8.7. Next, membership rules between the two inputs are defined; these rules will hold true in every situation.

Table 8.3 Detection of abnormalities on the basis of abnormality detection

Type of abnormalities	Set value of O_t	Measured value of O_t	Types of abnormalities detected
CST	15.6	23.4	GNL
IF	17.8	28.9	ECP
HM	19.6	15.9	CST
REL	21.6	17.5	IF
GNL	23.5	19.2	HM
MGT	25.6	25.2	MGT
ECP	27.8	21.9	REL

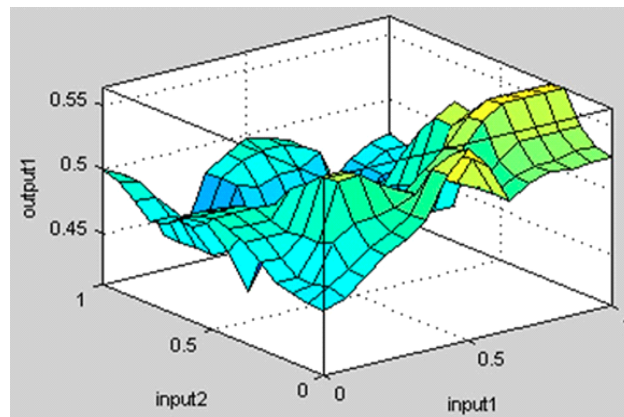


Fig.8.6 Surface view of FLC for hybrid approach

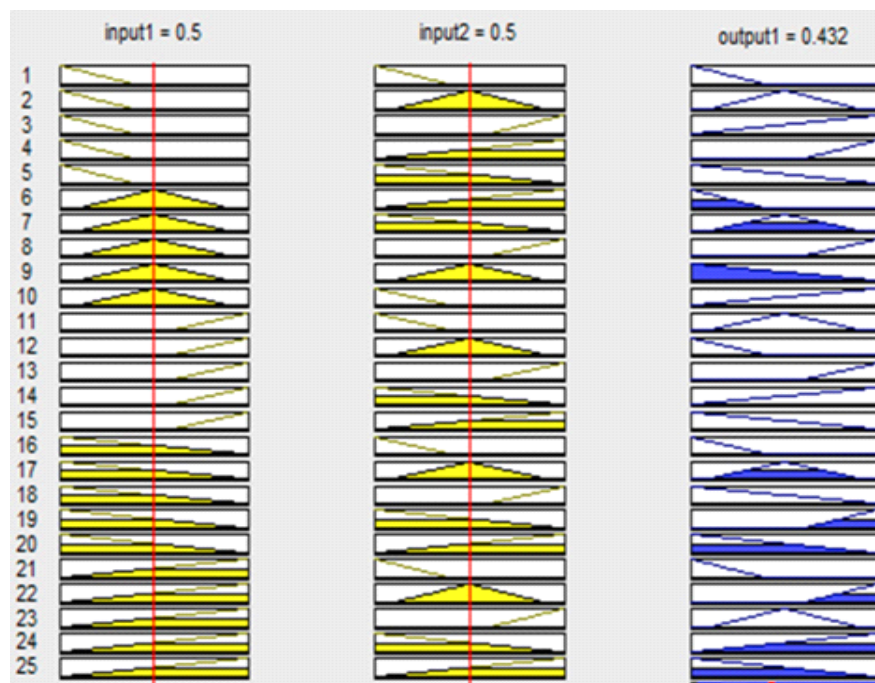


Fig.8.7 Rules structure of hybrid approach

8.4 EXPERIMENTAL RESULTS

Performance analysis is the primary focus of this section, which will be covered in detail. It has been mentioned in the previous sections that about 150 types of common abnormalities are found inside the brain. For performance analysis, a sample of 6 abnormalities has been used for the detection and classification of diseases. For assessing the purposes, an experiment has been done on 1000 image data for six

types of abnormalities which are referred from [42]. The two methods are used to analyze the 1000 collected data which are named as Chi-square test and Co-occurrence test.

8.4.1 Chi-square test (CST)

The sampling distribution of the statistical test is known as the chi-squared distribution when we assume that the null hypothesis is correct. If there is a significant difference between the anticipated frequencies as well as the observed frequencies for one or even more groups or categories, the chi-squared test can help to identify it. The likelihood of independent variables is shown. The general representation of the chi-square test is shown in Eqn. 8.23.

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (8.23)$$

where, O_i is the observed value and E_i is the expected value of the i^{th} variable.

1000 test data sets are distributed among the 6 abnormalities on the basis of CST as mentioned in Table 8.4.

Table 8.4 Test data allocation among six abnormalities

CST	IF	HM	REL	GNL	MGT	ECP	Total
LS	67	73	63	61	62	74	400
MS	50	50	40	50	45	65	300
HS	42	49	51	45	52	61	300
	149	172	154	156	159	200	1000

The various abnormalities have been allocated to the various samples on the basis of severity levels such as low severity (LS), medium severity (MS), and high severity (HS). The expected value of the sample can be estimated as given in Eqn. 8.23.

$$E_i = \frac{n_s * N_{ns(L,M,H)}}{N} \quad (8.23)$$

where, n_s = no. of samples for particular severity level, $N_{ns(L,M,H)}$ = No. of samples corresponding to abnormality for a particular severity level, N = Total no. of samples (1000). The expected value of various abnormalities corresponding to different severity levels is shown in Table 8.5.

Table 8.5 Expected value of abnormalities for different severity levels

CST	IF	HM	REL	GNL	MGT	ECP
LS	26.8	29.2	25.2	24.4	24.8	29.6
MS	15	15	12	15	13.5	19.5
HS	12.6	14.7	15.3	13.5	15.6	18.3

From Table 8.5, it can be observed that the chances of occurrence of ECP in terms of data set are highest in comparison to HM, IF, and other abnormalities. The second highest chances of occurrence with abnormalities HM for all three types of severities. Followed by IF, REL, GNL, and MGT. This is how the various abnormalities are identified and segregated on the basis of chance of occurrence. CST is able to distinguish the various abnormalities of the brain with the given tested data. Another way to identify the various abnormalities of the brain is by estimating the expected values. The higher the value of expectation means, the more chance of occurring of those particular diseases. From Table 8.5 sequences of occurring abnormalities are ECP, HM, IF, REL, GNL, and MGT.

CST also shows the data contribution of each abnormality with respect to given abnormalities. The data contribution of each abnormality for low severity level is shown in Fig. 8.8 while for medium severe level is shown in Fig. 8.9 along with high severe level is shown in Fig. 8.10 which concludes the complete dominance of ECP over others. Similar kinds of abnormalities can be expected for higher severity levels.

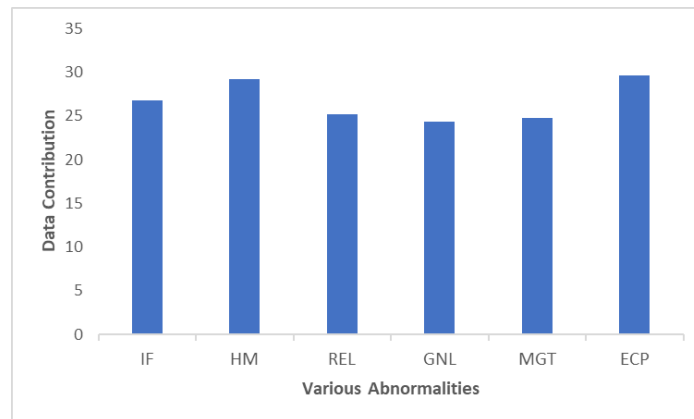


Fig.8.8 Percentage data contribution in various abnormalities for low severity abnormalities

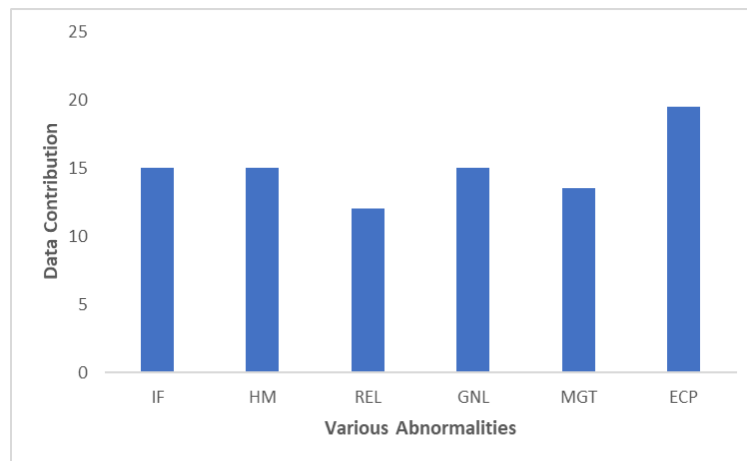


Fig.8.9 Percentage data contribution in various abnormalities for medium severity abnormalities

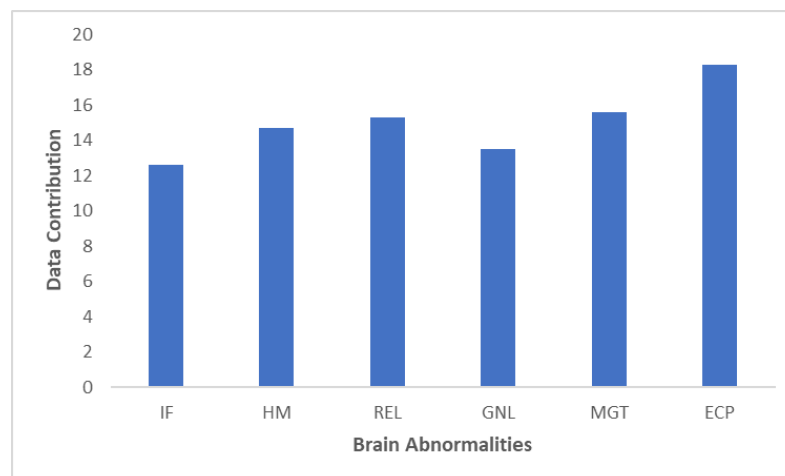


Fig.8.10 Percentage data contribution in various abnormalities for high severity abnormalities

8.4.2 Co-occurrence test

The co-occurrence test is used for the assessment of qualitative analysis of the system between two variables. The measurement of co-occurrence is given by the relationship [169] and its mathematical measurement is given in Eqn. 8.24.

$$R(1, m) = \frac{\sum_{n_s} conc(1, m)}{\sum_{n_s} I(l)} \quad (8.24)$$

l, m are the variables while n_s is the no. of the samples. $conc(l, m)$ and $I(l)$ are the perfect indicators for identifying the relationship which are given in Eqn.8.25 and Eqn.8.26 respectively.

$$conc(1, m) = \begin{cases} 1, Arr(1) \geq 0.5 \wedge Arr(m) \geq 0.5 \\ 0, else \end{cases} \quad (8.25)$$

$$I(1) = \begin{cases} 1, Arr(1) \geq 0.5 \\ 0, else \end{cases} \quad (8.26)$$

whereas $Arr(1)$ represents the strength of the abnormalities (1), \wedge shows the conjunction symbol.

The rank (r) is chosen as a variable for deciding the priority of selecting the abnormalities which is given as Eqn. 8.27.

$$r = \frac{R(1, m)}{R(m, 1)} \quad (8.27)$$

The rank decides the probability of occurrence of abnormalities for low and medium severity levels. It is observed that MGT occurs frequently and is larger in comparison to other abnormalities. The rank of abnormalities for low and medium level severity levels is shown in Tables 8.6 and 8.7.

Table 8.6 Rank of abnormalities of low severity level

Rank	IF	HM	REL	GNL	MGT	ECP
IF	1	0.89	0.85	0.82	0.79	0.91
HM	0.81	1	0.80	0.74	0.71	0.87
REL	0.79	0.78	1	0.71	0.69	0.83
GNL	0.77	0.71	0.79	1	0.65	0.82
MGT	0.71	0.73	0.72	0.69	1	0.79
ECP	0.68	0.77	0.69	0.67	0.62	1

Table 8.7 Rank of abnormalities of medium severity level

Rank	IF	HM	REL	GNL	MGT	ECP
IF	1	0.90	0.86	0.84	0.80	0.93
HM	0.83	1	0.82	0.75	0.73	0.89
REL	0.80	0.79	1	0.73	0.70	0.84
GNL	0.78	0.72	0.80	1	0.67	0.84
MGT	0.72	0.74	0.73	0.70	1	0.80
ECP	0.69	0.78	0.71	0.68	0.64	1

8.5 EXPERIMENTAL VALIDATION AND PERFORMANCE RESULT ANALYSIS

Experimental analysis has been carried out to estimate the various locations of abnormalities in the brain as shown in Fig. 8.11. In Fig. 8.11, various samples of MRI (A, B, C, D, E, F, G, H) have been taken to detect and classify the abnormalities based on a multi-mark approach. A total of six abnormalities are disbursed from the eight samples of MRI reports. The number of repetitions of ECP in the various samples of MRI reports is quite high in comparison to other abnormalities which shows the high possibility of occurrence of ECP. The next sequence of occurring of the abnormalities is HM which is followed by IF, REL, GNL, and MGT. This is how the various abnormalities can be identified by using a hybrid approach, CNN, RNN-based LSTM approach. The performance parameters like POC, and MSE are used for detecting the different abnormalities of the brain.

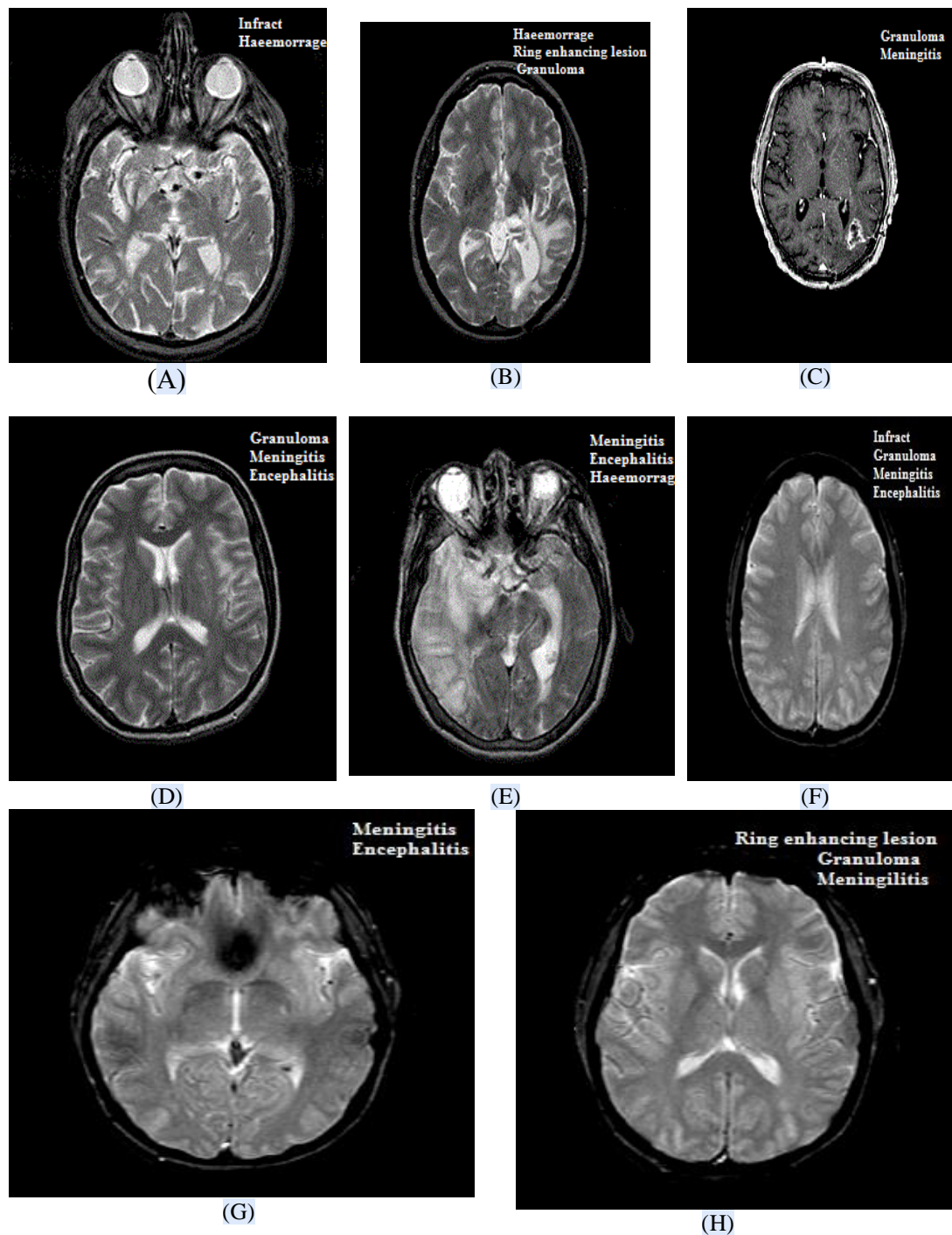


Fig.8.11 Experimental results of various abnormalities detected through the various samples of the MRI

It is seen that three different severity levels are used to identify the six types of abnormalities. The performance pattern is found to be the same for all levels. That's why from now onwards, critical or high severity levels have been considered for

estimating the performance parameters. Table 8.8 shows the POC of six abnormalities with the various methods.

Table 8.8 POC estimation with different methods of the various abnormalities

POC	IF	HM	REL	GNL	MGT	ECP
Hybrid approach	63.2	67	70.1	64.9	62.6	69.9
RNN-LSTM	57.2	61	64.1	58.9	56.6	63.9
CNN	52.1	55.9	59	53.8	51.5	58.8
[172]	36.5	40.3	43.4	38.2	35.9	43.2
[173]	34.2	38	41.1	35.9	33.6	40.9
[117]	33.9	37.5	40.6	35.1	33.1	40.1
[174]	33.5	36.9	39.8	34.8	32.9	39.7

Table 8.9 MSE estimation with different methods of the various abnormalities

MSE	IF	HM	REL	GNL	MGT	ECP
Hybrid approach	3.6	5.1	4.6	4.9	3.9	3.9
RNN-LSTM	5.9	7.4	6.9	7.2	6.2	6.2
CNN	6.4	7.9	7.4	7.7	6.7	6.7
[172]	9.8	11.3	10.8	11.1	10.1	10.1
[173]	10.7	12.2	11.7	12	11	11
[117]	10.9	12.9	11.9	12.5	11.5	11.6
[174]	11.1	13.6	12.1	13.1	11.9	12.1

The comparative analysis of performance parameters like the probability of occurrence (POC) and mean square error (MSE) of various abnormalities with various methods like hybrid approach, RNN-LSTM, CNN, Ref [172], Ref [173], Ref [117], Ref [174] are shown in Table 8.8 and Table 8.9 respectively. The probability of occurrence of various abnormalities is highest with the hybrid approach in comparison to other methods as shown in Table 8.8 while its graphical analysis is shown in Fig. 8.12.

In the same way, mean square error (MSE) is found to be optimally least with the hybrid approach in comparison to other methods as shown in Table 8.9 while its graphical analysis is shown in Fig. 8.13. The experimental validation and estimation of performance parameters like POC and MSE were found to be more fruitful with the hybrid approach for the classification and detection of various abnormalities.

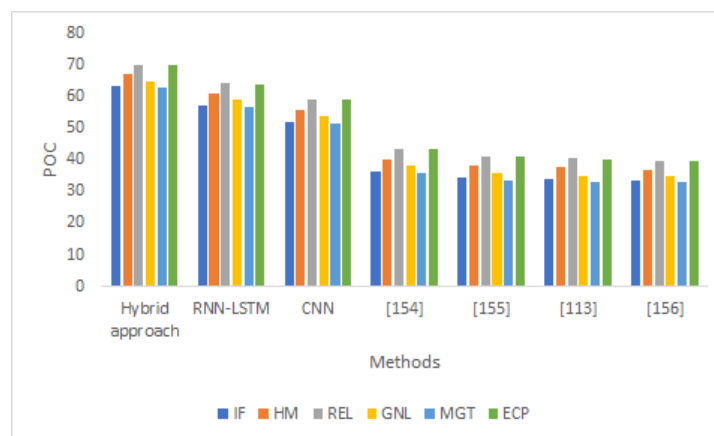


Fig.8.12. Graphical comparison of POC (%) with various methods

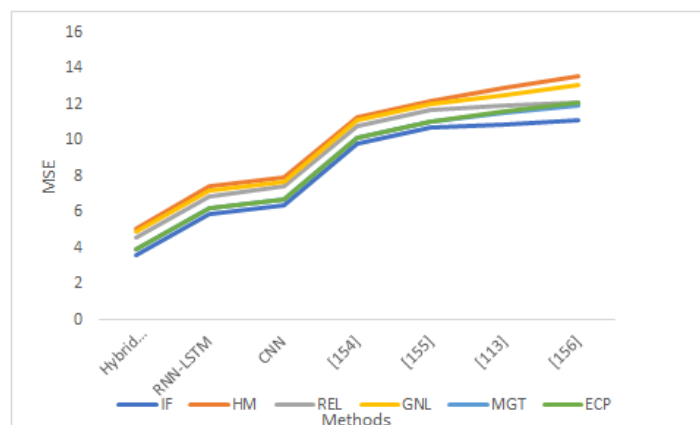


Fig.8.13 Graphical comparison of MSE (%) with various methods

8.5.1 Contribution of brain abnormalities in both Local and global contextual Levels

There are a total of six samples considered out of a total of 150 abnormalities. The contribution of each abnormality with respect to given samples (6) for the various severity levels (low severity, medium severity, high severity) has been shown in Table 8.10. However, the contribution of each sample abnormality with respect to total abnormalities (150) for the various severity levels (low severity, medium severity, high severity) has been mentioned in Table 8.11.

Table 8.10 Local share of different abnormalities in the brain

Local share (%)	IF	HM	REL	GNL	MGT	ECP
LS	16.75	18.25	15.75	15.25	15.5	18.5
MS	16.67	16.67	13.33	16.67	15.00	21.67
HS	14	16.33	17.00	15.00	17.33	20.33

Table 8.11 Global share of different abnormalities in the brain

Global share (%)	IF	HM	REL	GNL	MGT	ECP
LS	2.79	3.04	2.63	2.54	2.58	3.08
MS	2.78	2.78	2.22	2.78	2.50	3.61
HS	2.33	2.72	2.83	2.50	2.89	3.39

From Table 8.10, the approximate contribution of each sample's abnormalities is about ($= 15.33\%$) with respect to the total sample abnormalities (six). The graphical analysis of the local share of the sample of abnormalities is shown in Fig. 8.14. The global sharing of each sample's abnormalities is about ($= 2.5\%$) with respect to the total number of abnormalities (150) or global share which is shown in Table 8.11 and its graphical comparison is shown in Fig. 8.15.

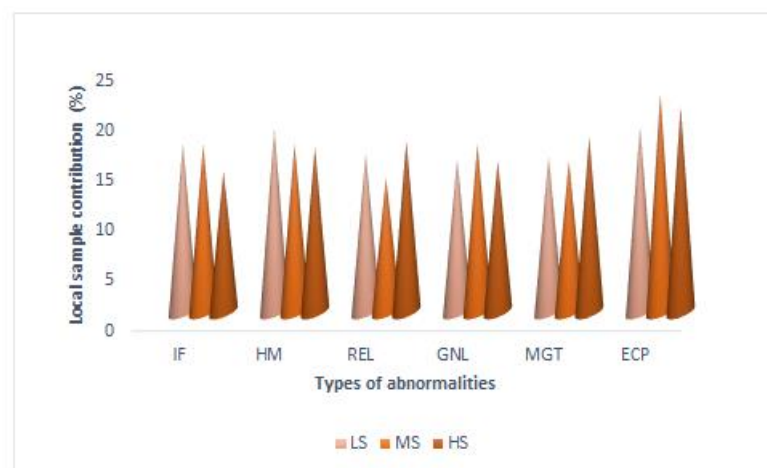


Fig.8.14. Graphical comparison of local contribution of samples of each abnormality with respect to different severity levels

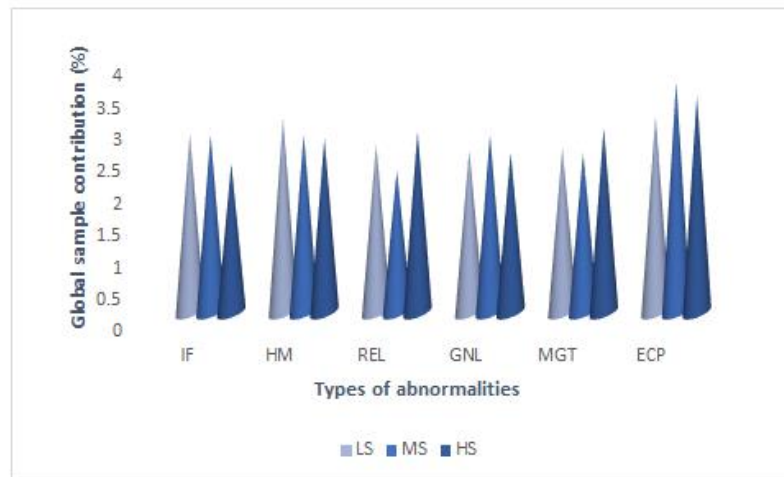


Fig.8.15 Graphical comparison of the global contribution of each sample's abnormalities with respect to different severity levels

8.5.2 Approach for noise reduction in the detection of abnormalities

The process of detection and classification of abnormalities in the brain produces a lot of noise. It is required to reduce noise levels with appropriate methods. In order to reduce the noise level, a mean square input-sharing algorithm is implemented. The component sharing of inputs and output is given in Eqn. 8.28.

$$c = \frac{X_i}{X_i^{max}}, d = \frac{y_i}{y_i^{max}} \quad (8.28)$$

c & d are called non-linear controller inputs and output-sharing ratios. Now the discrete time-varying mean square non-linear controllers ($c[n]$ & $d[n]$) are designed so that the problem of noise can be avoided. Thereafter, the actual measured value will be measured in terms of the modified controller as shown in Eqn. 8.29 and Eqn.8.30.

$$X_{measured}^{actual} = c[n]X_{measured} \quad (8.29)$$

$$y_{measured}^{actual} = d[n]y_{measured} \quad (8.30)$$

$$\text{where, } c[n] = \frac{\sum_{i=1}^6 \left(\frac{X_i(t)}{X_i^{max}} - \frac{X_{i+1}(t)}{X_{i+1}^{max}} \right)}{n^2 + 1} \text{ and } d[n] = \frac{\sum_{i=1}^6 \left(\frac{y_i(t)}{y_i^{max}} - \frac{y_{i+1}(t)}{y_{i+1}^{max}} \right)}{n^2 + 1}$$

where 'i' is the number of abnormalities.

Another important feature has been considered for the attenuation of the noise which is named as convergence factor. The convergence factor is represented in Eqn. 8.31 and Eqn.8.32.

$$\lim_{n \rightarrow \infty} (X_{measured}^{actual} - X_{measured}) \left| \frac{X_i(t)}{X_i^{max}} \frac{X_{i+1}(t)}{X_{i+1}^{max}} \right|^2 \rightarrow 0 \quad (8.31)$$

$$\lim_{n \rightarrow \infty} (y_{measured}^{actual} - y_{measured}) \left| \frac{y_i(t)}{y_i^{max}} \frac{y_{i+1}(t)}{y_{i+1}^{max}} \right|^2 \rightarrow 0 \quad (8.32)$$

8.6 DISCUSSION

In order to obtain a precise image of the brain structure, the MRI is considered to be quite essential for the diagnosis of brain anomalies, however, due to its complexity and the amount of data, the interpretation of these images is quite difficult. The single-mark approach or method is not that effective and tends to misclassify as it is not able to capture completely the complex nature of brain abnormalities. This impreciseness tends to initiate patient anxiety and thereby postpones the necessary treatment. Therefore, it is necessary to develop more advanced classification techniques in order to overcome the issue possessed by the single-mark approach/method in MRI.

The multi-mark classification approach thereby tends to overcome the disadvantages of the single-mark classifications. Utilization of this approach tends to give various MRI scan sections with multiple labels which further tends to denote different abnormality. A more thorough analysis of the MRI scans can be conducted as this approach tends to employ multiple marks, that enhance the POC and reduce the MSE during the process of classifying. Therefore, this approach is a huge improvement in the diagnosis of brain abnormalities using medical imaging.

A hybrid intelligent system combining CNN and RNN shows significant improvement in the performance of multi-mark technique. The CNNs are proficient in pattern recognitions, MRI image extractions, and distinguish between the normal and aberrant tissue structure while the RNN specifically the LSTM networks are efficient in handling the temporal dependencies and sequential data which plays an important role in classifying the abnormalities across the no. of MRI slices. Considering the CNNs and RNNs separately, the hybrid system tends to have a better classification performance. This combination improves the effectiveness of the classification system and tends to fix the gradient vanishing problem in RNNs.

The efficiency of the suggested hybrid approach is examined by utilizing the datasets of 1000 MRI images that showcase the range of brain abnormalities. The performance of the system is monitored by utilizing the performance metrics such as POC and MSE. The results showcase that the hybrid system performs better as compared to the conventional approaches/techniques thereby giving an accurate and more dependable classification. Therefore, the hybrid systems outcomes tend to deliver better diagnostic results due to which the creation of effective treatment plans can be carried out with accurate diagnosis therefore giving better patient outcomes.

8.7 SUMMARY

This chapter focuses upon the critical problem of correctly categorizing brain abnormalities from MRIs. Due to the complexity involved in brain imaging, the conventional single-mark classification tends to fail as the results provided by it is not reliable. This chapter showcases the hybrid intelligent system that integrates the RNN, more importantly, LSTM and CNN in order to initiate future improvements in this technique. The hybrid systems integrate the sequential data handling strengths of the RNN with the feature extraction of the CNN to achieve an improved classification. In order to check the effectiveness of this approach, a dataset of 1000 MRI images is being utilized thoroughly and also the performance metrics like mean square error, and probability of occurrence showcases the improvements over the conventional methods. Therefore, this technology has the potential to transform MRI-based

diagnostic procedures, thereby promoting an accurate as well as early diagnostic of brain abnormalities and thereby serving the patient for better outcomes.

CHAPTER 9

CONCLUSION, FUTURE WORK AND SOCIAL IMPACT

9.1 CONCLUSION

Using a variety of cutting-edge deep learning techniques, we have explored the complex field of brain tumor detection using MRI images in this extensive study. Every chapter has made a distinct contribution to our comprehension and progress in this important area, highlighting both successes and ongoing difficulties.

Chapter 2: A comprehensive review of different segmentation and classification methods for brain MRI analysis was given in this chapter. It emphasized the use of 3D Convolutional Neural Networks (CNNs) in MRI data to capture complex spatial relationships and investigated the use of cascade CNNs in multi-stage learning to reduce overfitting. Furthermore, the potential for improving model performance was investigated using optimization techniques such as Sine-Cosine Fitness Grey Wolf Optimization. The review emphasized how promising these techniques are for detecting brain tumors with high sensitivity, specificity, and accuracy. It also emphasized the need for substantial training data to ensure generalizability, pointing out the computational complexity and risk of overfitting when using small datasets.

Chapter 3: This chapter first introduced DNN based model that helped in improving the effectiveness of MRI scans that were adversely affected by Rician noise. The proposed model comprises of several convolution layers, parallelly gathering various image features while filtering out noise. The performed comparative evaluation showed that the proposed approach outperformed various traditional denoising techniques, and the median filter performed the least efficiently. While improving the effectiveness of MRI for subsequent processing, the proposed approach

also highlighted the potential of various filters that have been used to reduce noise in MRI scans.

Building on this, we also presented the Modified Triplet Cross Fusion Learning (MTCL-MRI) denoising technique, which led to the introduction of a robust pipeline for brain tumor detection in MRI images. The technique showed notable gains in noise reduction while maintaining structural details. Using ResNet-based techniques for classification and U-Net for efficient feature extraction, the segmentation phase produced a high classification accuracy of 99.56% on the BraTS 2020 dataset. This methodology demonstrated its potential to improve medical imaging and diagnostic accuracy by outperforming several cutting-edge techniques.

Chapter 4: This chapter evaluated six pre-trained CNN models (VGG16, VGG19, InceptionV3, ResNet50, ResNet101, EfficientNetB1) for brain MRI image classification using transfer learning. ResNet101 and VGG16 were the top performers, obtaining the highest classification accuracy, according to the results. According to the study, transfer learning greatly improves training effectiveness and precision, which makes these models very useful for automated brain tumor detection.

Chapter 5: In order to decrease computational costs and increase detection accuracy, the chapter presented a novel method that combines Atrous convolution with triplanar isolated spatial pyramidal pooling. In terms of mean intersection over union (mIoU), dice scores, accuracy, precision, recall, and F1 score, this method performed better.

Chapter 6: The HyperSwinNet framework was introduced in this chapter and it was able to obtain an average Dice score of 90.4% on the BraTS dataset. The framework effectively navigated the large search space of segmentation networks by combining a transformer-based architecture with self-supervised pre-training. Using a variety of MRI tasks, its robustness was shown, highlighting its clinical relevance and potential to enhance therapeutic and diagnostic results.

Chapter 7: In order to address the problems of limited data and lengthy computation times, this study used a Vision Transformer (ViT) with a standalone GAN for preprocessing, normalization, and pixel segmentation. The model's ability to extract global and local features and improve image similarity was demonstrated by its high accuracy and sensitivity values.

Chapter 8: The last chapter presented a hybrid intelligent system that combines CNNs and RNNs, specifically LSTM networks, for multi-mark classification. This method allowed for the accurate and early detection of brain abnormalities, demonstrating significant improvements in classification performance. The method has a significant impact on patient outcomes, as evidenced by its potential to revolutionize MRI-based diagnostic procedures.

Overall Conclusion: This work has significantly advanced the field of MRI-based brain tumor detection. Through the application of various cutting-edge deep learning techniques and creative approaches, we have tackled important problems like precise classification, accurate segmentation, and noise reduction. The results highlight how these methods can lead to better patient outcomes, lower computational costs, and increased diagnostic accuracy. Subsequent research endeavors will center on enhancing these techniques, broadening their range of applications, and surmounting outstanding obstacles to propel the domain of medical imaging and brain tumor identification forward.

9.2 FUTURE WORK

Even though our study has several advantages, we have found a few shortcomings that should be addressed in further research.

- In the future scope, the survival prediction of the patient can also be made possible.
- The current study could potentially be used for complex imaging techniques like PET, fMRI, and others in order to differentiate between brain images in three-dimensional image datasets.

- The semantic gap is an issue that mostly affects deep learning features. This issue was visible in the retrieval studies as well as in the visualized brain MRI features that deep networks had learned. The semantic gap between machine learning features and features developed by clinicians may be minimized by deep neural networks that use advanced feature-aware learning. Future studies in this area will make a substantial contribution to dealing with this issue.
- Our work is mainly focused on Brain MRI images and focused on abnormalities related to the brain. But in the future, we can focus on applying various algorithms to multiple organs.
- It is becoming more and more crucial to develop efficient multimodal fusion techniques as one modality may not be reliable or adequate for capturing the variety of complicated diseases in order to improve personalized medicine and customize medical care.

9.3 SOCIAL IMPACT

In the area of medical diagnosis and treatment, the proposed work has profound social implications as it gives advancements in MRI image denoising, segmentation, feature extraction, and classification. This research contributes to the precise and early detection of brain tumors by enhancing the quality of MRI scans. Early detection is an important parameter in treatment planning and improving patient outcomes.

In addition to it, the development of adequate models for MRI analysis also provides the way for their application in other complex imaging techniques and expands its scope of improved diagnostics other than brain tumors. This has the prospective of transforming personalized medicine as well as making treatments based on more accurate diagnosis. By overcoming this limitation, the overall quality of healthcare can be improved, as the advanced diagnostic tool will lead to timely and accurate identification of disease. This will ultimately improve patient survival rate and will provide the quality of life to patients suffering from serious medical conditions.

REFERENCES

- [1] Y. LeCun, Y. Bengio, and G. Hinton, 'Deep learning', *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [2] J. Tang, C. Deng, G.-B. Huang, and B. Zhao, 'Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine', *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1174–1185, Mar. 2015, doi: 10.1109/TGRS.2014.2335751.
- [3] M. I. Razzak, S. Naz, and A. Zaib, 'Deep Learning for Medical Image Processing: Overview, Challenges and the Future', 2018, pp. 323–350. doi: 10.1007/978-3-319-65981-7_12.
- [4] G. Litjens *et al.*, 'A survey on deep learning in medical image analysis', *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [5] D. Shen, G. Wu, D. Zhang, K. Suzuki, F. Wang, and P. Yan, 'Machine learning in medical imaging', *Comput. Med. Imaging Graph.*, vol. 41, pp. 1–2, Apr. 2015, doi: 10.1016/j.compmedimag.2015.02.001.
- [6] D. Shen, G. Wu, and H.-I. Suk, 'Deep Learning in Medical Image Analysis', *Annu. Rev. Biomed. Eng.*, vol. 19, no. 1, pp. 221–248, Jun. 2017, doi: 10.1146/annurev-bioeng-071516-044442.
- [7] K. Suzuki, 'Survey of deep learning applications to medical image analysis', *Med Imaging Technol.*, vol. 35, no. 4, pp. 212–226, 2017.
- [8] J. Ker, L. Wang, J. Rao, and T. Lim, 'Deep Learning Applications in Medical Image Analysis', *IEEE Access*, vol. 6, pp. 9375–9389, 2018, doi: 10.1109/ACCESS.2017.2788044.
- [9] A. Maier, C. Syben, T. Lasser, and C. Riess, 'A Gentle Introduction to Deep Learning in Medical Image Processing', Oct. 2018, doi: <https://doi.org/10.48550/arXiv.1810.05401>.
- [10] F. Movahedi, J. L. Coyle, and E. Sejdic, 'Deep Belief Networks for Electroencephalography: A Review of Recent Contributions and Future Outlooks', *IEEE J. Biomed. Heal. Informatics*, vol. 22, no. 3, pp. 642–652, May 2018, doi: 10.1109/JBHI.2017.2727218.
- [11] S. Liu *et al.*, 'Deep Learning in Medical Ultrasound Analysis: A Review', *Engineering*, vol. 5, no. 2, pp. 261–275, Apr. 2019, doi: 10.1016/j.eng.2018.11.020.
- [12] A. S. Lundervold and A. Lundervold, 'An overview of deep learning in medical imaging focusing on MRI', *Z. Med. Phys.*, vol. 29, no. 2, pp. 102–127, May 2019, doi: 10.1016/j.zemedi.2018.11.002.
- [13] J. Liu *et al.*, 'Applications of deep learning to MRI images: A survey', *Big Data Min. Anal.*, vol. 1, no. 1, pp. 1–18, Mar. 2018, doi: 10.26599/BDMA.2018.9020001.
- [14] A. D. Piersson and P. N. Gorleku, 'A national survey of MRI safety practices in Ghana', *Heliyon*, vol. 3, no. 12, p. e00480, Dec. 2017, doi: 10.1016/j.heliyon.2017.e00480.
- [15] R. D. Nowak, 'Wavelet-based Rician noise removal for magnetic resonance imaging', *IEEE Trans. Image Process.*, vol. 8, no. 10, pp. 1408–1419, 1999, doi: 10.1109/83.791966.
- [16] S. Hussain, S. M. Anwar, and M. Majid, 'Segmentation of glioma tumors in brain using deep convolutional neural network', *Neurocomputing*, vol. 282, pp. 248–261, Mar. 2018, doi: 10.1016/j.neucom.2017.12.032.
- [17] J. Liu *et al.*, 'IOUC-3DSFCNN: Segmentation of Brain Tumors via IOU Constraint 3D

- Symmetric Full Convolution Network with Multimodal Auto-context', *Sci. Rep.*, vol. 10, no. 1, p. 6256, Apr. 2020, doi: 10.1038/s41598-020-63242-x.
- [18] D. Nie, H. Zhang, E. Adeli, L. Liu, and D. Shen, '3D Deep Learning for Multi-modal Imaging-Guided Survival Time Prediction of Brain Tumor Patients', 2016, pp. 212–220. doi: 10.1007/978-3-319-46723-8_25.
 - [19] S. Lu, Z. Lu, and Y.-D. Zhang, 'Pathological brain detection based on AlexNet and transfer learning', *J. Comput. Sci.*, vol. 30, pp. 41–47, Jan. 2019, doi: 10.1016/j.jocs.2018.11.008.
 - [20] G. S. Tandel *et al.*, 'A Review on a Deep Learning Perspective in Brain Cancer Classification', *Cancers (Basel)*, vol. 11, no. 1, p. 111, Jan. 2019, doi: 10.3390/cancers11010111.
 - [21] J. Amin, M. Sharif, M. Yasmin, and S. L. Fernandes, 'Big data analysis for brain tumor detection: Deep convolutional neural networks', *Futur. Gener. Comput. Syst.*, vol. 87, pp. 290–297, Oct. 2018, doi: 10.1016/j.future.2018.04.065.
 - [22] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghouschi, S. Anari, M. Naseri, and M. Bendeche, 'Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images', *Sci. Rep.*, vol. 11, no. 1, p. 10930, May 2021, doi: 10.1038/s41598-021-90428-8.
 - [23] P. Tiwari *et al.*, 'CNN Based Multiclass Brain Tumor Detection Using Medical Imaging', *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–8, Jun. 2022, doi: 10.1155/2022/1830010.
 - [24] 'Statistics, Cancer-Types. Brain Tumor': Accessed: Jan. 31, 2023. [Online]. Available: <https://www.cancer.net/cancer-types/brain-tumor/statistics>
 - [25] Á. Györfi, L. Szilágyi, and L. Kovács, 'A Fully Automatic Procedure for Brain Tumor Segmentation from Multi-Spectral MRI Records Using Ensemble Learning and Atlas-Based Data Enhancement', *Appl. Sci.*, vol. 11, no. 2, p. 564, Jan. 2021, doi: 10.3390/app11020564.
 - [26] S. A. Y. Al-Galal, I. F. T. Alshaikhli, and M. M. Abdulrazzaq, 'MRI brain tumor medical images analysis using deep learning techniques: a systematic review', *Health Technol. (Berl.)*, vol. 11, no. 2, pp. 267–282, Mar. 2021, doi: 10.1007/s12553-020-00514-6.
 - [27] A. Tiwari, S. Srivastava, and M. Pant, 'Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019', *Pattern Recognit. Lett.*, vol. 131, pp. 244–260, Mar. 2020, doi: 10.1016/j.patrec.2019.11.020.
 - [28] S. Saman and S. Jamjala Narayanan, 'Survey on brain tumor segmentation and feature extraction of MR images', *Int. J. Multimed. Inf. Retr.*, vol. 8, no. 2, pp. 79–99, Jun. 2019, doi: 10.1007/s13735-018-0162-2.
 - [29] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, 'A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning', *Circuits, Syst. Signal Process.*, vol. 39, no. 2, pp. 757–775, Feb. 2020, doi: 10.1007/s00034-019-01246-3.
 - [30] N. P. Gupta, P. K. Malik, and B. S. Ram, 'A Review on Methods and Systems for Early Breast Cancer Detection', in *2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM)*, IEEE, Jan. 2020, pp. 42–46. doi: 10.1109/ICCAKM46823.2020.9051554.
 - [31] M. K. Abd-Ellah, A. I. Awad, A. A. M. Khalaf, and H. F. A. Hamed, 'A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned', *Magn. Reson. Imaging*, vol. 61, pp. 300–318, Sep. 2019, doi: 10.1016/j.mri.2019.05.028.
 - [32] M. W. Nadeem *et al.*, 'Brain Tumor Analysis Empowered with Deep Learning: A

- Review, Taxonomy, and Future Challenges’, *Brain Sci.*, vol. 10, no. 2, p. 118, Feb. 2020, doi: 10.3390/brainsci10020118.
- [33] R. Kaifi, ‘A Review of Recent Advances in Brain Tumor Diagnosis Based on AI-Based Classification’, *Diagnostics*, vol. 13, no. 18, p. 3007, Sep. 2023, doi: 10.3390/diagnostics13183007.
- [34] A. Younis, Q. Li, M. Khalid, B. Clemence, and M. J. Adamu, ‘Deep Learning Techniques for the Classification of Brain Tumor: A Comprehensive Survey’, *IEEE Access*, vol. 11, pp. 113050–113063, 2023, doi: 10.1109/ACCESS.2023.3317796.
- [35] C. Han *et al.*, ‘Infinite Brain MR Images: PGGAN-based Data Augmentation for Tumor Detection’, Mar. 2019.
- [36] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, ‘Classification using deep learning neural networks for brain tumors’, *Futur. Comput. Informatics J.*, vol. 3, no. 1, pp. 68–71, Jun. 2018, doi: 10.1016/j.fcij.2017.12.001.
- [37] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, ‘Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges’, *J. Digit. Imaging*, vol. 32, no. 4, pp. 582–596, Aug. 2019, doi: 10.1007/s10278-019-00227-x.
- [38] B. H. Menze *et al.*, ‘The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)’, *IEEE Trans. Med. Imaging*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: 10.1109/TMI.2014.2377694.
- [39] S. Bakas *et al.*, ‘Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features’, *Sci. Data*, vol. 4, no. 1, p. 170117, Sep. 2017, doi: 10.1038/sdata.2017.117.
- [40] S. Bakas *et al.*, ‘Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge’, Nov. 2018.
- [41] K. Clark *et al.*, ‘The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository’, *J. Digit. Imaging*, vol. 26, no. 6, pp. 1045–1057, Dec. 2013, doi: 10.1007/s10278-013-9622-7.
- [42] D. Summers, ‘Harvard Whole Brain Atlas: www.med.harvard.edu/AANLIB/home.html’, *J. Neurol. Neurosurg. Psychiatry*, vol. 74, no. 3, pp. 288–288, Mar. 2003, doi: 10.1136/jnnp.74.3.288.
- [43] “Brain Tumor MRI Dataset | Kaggle.” <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset> (accessed.)
- [44] “Figshare Brain Tumor Dataset | Kaggle.” <https://www.kaggle.com/datasets/ashkhagan/figshare-brain-tumor-dataset>.
- [45] V. Cheplygina, M. de Bruijne, and J. P. W. Pluim, ‘Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis’, *Med. Image Anal.*, vol. 54, pp. 280–296, May 2019, doi: 10.1016/j.media.2019.03.009.
- [46] A. Fatima, A. R. Shahid, B. Raza, T. M. Madni, and U. I. Janjua, ‘State-of-the-Art Traditional to the Machine- and Deep-Learning-Based Skull Stripping Techniques, Models, and Algorithms’, *J. Digit. Imaging*, vol. 33, no. 6, pp. 1443–1464, Dec. 2020, doi: 10.1007/s10278-020-00367-5.
- [47] H. Z. U. Rehman, H. Hwang, and S. Lee, ‘Conventional and Deep Learning Methods for Skull Stripping in Brain MRI’, *Appl. Sci.*, vol. 10, no. 5, p. 1773, Mar. 2020, doi: 10.3390/app10051773.
- [48] T. Magadza and S. Viriri, ‘Deep Learning for Brain Tumor Segmentation: A Survey of State-of-the-Art’, *J. Imaging*, vol. 7, no. 2, p. 19, Jan. 2021, doi: 10.3390/jimaging7020019.
- [49] A. Wadhwa, A. Bhardwaj, and V. Singh Verma, ‘A review on brain tumor segmentation of MRI images’, *Magn. Reson. Imaging*, vol. 61, pp. 247–259, Sep. 2019,

- doi: 10.1016/j.mri.2019.05.043.
- [50] L. Shapiro and G. Stockman, *Computer vision*. Upper Saddle River: Prentice-Hall, 2001.
 - [51] A. Helwan and R. Abiyev, 'Shape and texture features for the identification of breast cancer', *Lect. Notes Eng. Comput. Sci.*, vol. 2226, pp. 542–547, 2016.
 - [52] V. R. Pandit and R. J. Bhiwani, 'Image Fusion in Remote Sensing Applications: A Review', *Int. J. Comput. Appl.*, vol. 120, no. 10, pp. 22–32, 2015, doi: 10.5120/21263-3846.
 - [53] K. Rautela, D. Kumar, and V. Kumar, 'A Systematic Review on Breast Cancer Detection Using Deep Learning Techniques', *Arch. Comput. Methods Eng.*, vol. 29, no. 7, pp. 4599–4629, Nov. 2022, doi: 10.1007/s11831-022-09744-5.
 - [54] Y. Yang, W. Wan, S. Huang, F. Yuan, S. Yang, and Y. Que, 'Remote Sensing Image Fusion Based on Adaptive IHS and Multiscale Guided Filter', *IEEE Access*, vol. 4, pp. 4573–4582, 2016, doi: 10.1109/ACCESS.2016.2599403.
 - [55] V. Naidu and J. Raol, 'Pixel-level Image Fusion using Wavelets and Principal Component Analysis', *Def. Sci. J.*, vol. 58, no. 3, pp. 338–352, May 2008, doi: 10.14429/dsj.58.1653.
 - [56] G. Murtaza *et al.*, 'Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges', *Artif. Intell. Rev.*, vol. 53, no. 3, pp. 1655–1720, Mar. 2020, doi: 10.1007/s10462-019-09716-5.
 - [57] S. Agarwal, 'Data Mining: Data Mining Concepts and Techniques', in *2013 International Conference on Machine Intelligence and Research Advancement*, IEEE, Dec. 2013, pp. 203–207. doi: 10.1109/ICMIRA.2013.45.
 - [58] A. Srinivasa Reddy and P. Chenna Reddy, 'MRI brain tumor segmentation and prediction using modified region growing and adaptive SVM', *Soft Comput.*, vol. 25, no. 5, pp. 4135–4148, Mar. 2021, doi: 10.1007/s00500-020-05493-4.
 - [59] M. Thayumanavan and A. Ramasamy, 'An efficient approach for brain tumor detection and segmentation in MR brain images using random forest classifier', *Concurr. Eng.*, vol. 29, no. 3, pp. 266–274, Sep. 2021, doi: 10.1177/1063293X211010542.
 - [60] M. S. Alam *et al.*, 'Automatic Human Brain Tumor Detection in MRI Image Using Template-Based K Means and Improved Fuzzy C Means Clustering Algorithm', *Big Data Cogn. Comput.*, vol. 3, no. 2, p. 27, May 2019, doi: 10.3390/bdcc3020027.
 - [61] M. Angulakshmi and G. G. Lakshmi Priya, 'Walsh Hadamard Transform for Simple Linear Iterative Clustering (SLIC) Superpixel Based Spectral Clustering of Multimodal MRI Brain Tumor Segmentation', *IRBM*, vol. 40, no. 5, pp. 253–262, Oct. 2019, doi: 10.1016/j.irbm.2019.04.005.
 - [62] B. Devanathan and K. Venkatachalapathy, 'An Optimal Multilevel Thresholding based Segmentation and Classification Model for Brain Tumor Diagnosis', in *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, IEEE, Nov. 2020, pp. 1133–1138. doi: 10.1109/ICECA49313.2020.9297571.
 - [63] O. Ronneberger, P. Fischer, and T. Brox, 'U-Net: Convolutional Networks for Biomedical Image Segmentation', 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
 - [64] F. Lin, Q. Wu, J. Liu, D. Wang, and X. Kong, 'Path aggregation U-Net model for brain tumor segmentation', *Multimed. Tools Appl.*, vol. 80, no. 15, pp. 22951–22964, Jun. 2021, doi: 10.1007/s11042-020-08795-9.
 - [65] N. M. Aboelenein, P. Songhao, A. Koubaa, A. Noor, and A. Afifi, 'HTTU-Net: Hybrid Two Track U-Net for Automatic Brain Tumor Segmentation', *IEEE Access*, vol. 8, pp. 101406–101415, 2020, doi: 10.1109/ACCESS.2020.2998601.
 - [66] J. Zhang, J. Zeng, P. Qin, and L. Zhao, 'Brain tumor segmentation of multi-modality

- MR images via triple intersecting U-Nets', *Neurocomputing*, vol. 421, pp. 195–209, Jan. 2021, doi: 10.1016/j.neucom.2020.09.016.
- [67] M. I. Sharif, J. P. Li, M. A. Khan, and M. A. Saleem, 'Active deep neural network features selection for segmentation and recognition of brain tumors using MRI images', *Pattern Recognit. Lett.*, vol. 129, pp. 181–189, Jan. 2020, doi: 10.1016/j.patrec.2019.11.019.
 - [68] J. Sun, Y. Peng, Y. Guo, and D. Li, 'Segmentation of the multimodal brain tumor image used the multi-pathway architecture method based on 3D FCN', *Neurocomputing*, vol. 423, pp. 34–45, Jan. 2021, doi: 10.1016/j.neucom.2020.10.031.
 - [69] H. Chen, Z. Qin, Y. Ding, L. Tian, and Z. Qin, 'Brain tumor segmentation with deep convolutional symmetric neural network', *Neurocomputing*, vol. 392, pp. 305–313, Jun. 2020, doi: 10.1016/j.neucom.2019.01.111.
 - [70] J. Zhang, Z. Jiang, J. Dong, Y. Hou, and B. Liu, 'Attention Gate ResU-Net for Automatic MRI Brain Tumor Segmentation', *IEEE Access*, vol. 8, pp. 58533–58545, 2020, doi: 10.1109/ACCESS.2020.2983075.
 - [71] B. Awudong, P. Yakupu, J. Yan, and Q. Li, 'Research and Implementation of Denoising Algorithm for Brain MRIs via Morphological Component Analysis and Adaptive Threshold Estimation', *Mathematics*, vol. 12, no. 5, p. 748, Mar. 2024, doi: 10.3390/math12050748.
 - [72] R. G. Akindele, M. Yu, P. S. Kanda, E. O. Owoola, and I. Aribilola, 'Denoising of Nifti (MRI) Images with a Regularized Neighborhood Pixel Similarity Wavelet Algorithm', *Sensors*, vol. 23, no. 18, p. 7780, Sep. 2023, doi: 10.3390/s23187780.
 - [73] D. Pankaj, G. D., and N. K.A., 'A novel method for removing Rician noise from MRI based on variational mode decomposition', *Biomed. Signal Process. Control*, vol. 69, p. 102737, Aug. 2021, doi: 10.1016/j.bspc.2021.102737.
 - [74] M. Tian and K. Song, 'Boosting Magnetic Resonance Image Denoising With Generative Adversarial Networks', *IEEE Access*, vol. 9, pp. 62266–62275, 2021, doi: 10.1109/ACCESS.2021.3073944.
 - [75] F. Baselice, G. Ferraioli, V. Pascasio, and A. Sorriso, 'Denoising of MR images using Kolmogorov-Smirnov distance in a Non Local framework', *Magn. Reson. Imaging*, vol. 57, pp. 176–193, Apr. 2019, doi: 10.1016/j.mri.2018.11.022.
 - [76] D. Jiang, W. Dou, L. Vosters, X. Xu, Y. Sun, and T. Tan, 'Denoising of 3D magnetic resonance images with multi-channel residual learning of convolutional neural network', *Jpn. J. Radiol.*, vol. 36, no. 9, pp. 566–574, Sep. 2018, doi: 10.1007/s11604-018-0758-8.
 - [77] R. Biswas, D. Purkayastha, and S. Roy, 'Denoising of MRI Images Using Curvelet Transform', 2018, pp. 575–583. doi: 10.1007/978-981-10-4762-6_55.
 - [78] N. Joshi, S. Jain, and A. Agarwal, 'An improved approach for denoising MRI using non local means filter', in *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*, IEEE, Oct. 2016, pp. 650–653. doi: 10.1109/NGCT.2016.7877492.
 - [79] M. R. Maneesha Mohan, C. H. Sulochana, and T. Latha, 'Medical image denoising using multistage directional median filter', in *2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015]*, IEEE, Mar. 2015, pp. 1–6. doi: 10.1109/ICCPCT.2015.7159261.
 - [80] S. Rajendran *et al.*, 'Automated Segmentation of Brain Tumor MRI Images Using Deep Learning', *IEEE Access*, vol. 11, pp. 64758–64768, 2023, doi: 10.1109/ACCESS.2023.3288017.
 - [81] D. Cheng *et al.*, 'Brain tumor feature extraction and edge enhancement algorithm based on U-Net network', *Heliyon*, vol. 9, no. 11, p. e22536, Nov. 2023, doi: 10.1016/j.heliyon.2023.e22536.

- [82] M. T. Nyo, F. Mebarek-Oudina, S. S. Hlaing, and N. A. Khan, 'Otsu's thresholding technique for MRI image brain tumor segmentation', *Multimed. Tools Appl.*, vol. 81, no. 30, pp. 43837–43849, Dec. 2022, doi: 10.1007/s11042-022-13215-1.
- [83] Y. Ding *et al.*, 'ToStaGAN: An end-to-end two-stage generative adversarial network for brain tumor segmentation', *Neurocomputing*, vol. 462, pp. 141–153, Oct. 2021, doi: 10.1016/j.neucom.2021.07.066.
- [84] X. Zhou, X. Li, K. Hu, Y. Zhang, Z. Chen, and X. Gao, 'ERV-Net: An efficient 3D residual neural network for brain tumor segmentation', *Expert Syst. Appl.*, vol. 170, p. 114566, May 2021, doi: 10.1016/j.eswa.2021.114566.
- [85] S. Nema, A. Dudhane, S. Murala, and S. Naidu, 'RescueNet: An unpaired GAN for brain tumor segmentation', *Biomed. Signal Process. Control*, vol. 55, p. 101641, Jan. 2020, doi: 10.1016/j.bspc.2019.101641.
- [86] T. Imtiaz, S. Rifat, S. A. Fattah, and K. A. Wahid, 'Automated Brain Tumor Segmentation Based on Multi-Planar Superpixel Level Features Extracted From 3D MR Images', *IEEE Access*, vol. 8, pp. 25335–25349, 2020, doi: 10.1109/ACCESS.2019.2961630.
- [87] S. Sajid, S. Hussain, and A. Sarwar, 'Brain Tumor Detection and Segmentation in MR Images Using Deep Learning', *Arab. J. Sci. Eng.*, vol. 44, no. 11, pp. 9249–9261, Nov. 2019, doi: 10.1007/s13369-019-03967-8.
- [88] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, 'Automatic Brain Tumor Segmentation Using Cascaded Anisotropic Convolutional Neural Networks', 2018, pp. 178–190. doi: 10.1007/978-3-319-75238-9_16.
- [89] H. ZainEldin *et al.*, 'Brain Tumor Detection and Classification Using Deep Learning and Sine-Cosine Fitness Grey Wolf Optimization', *Bioengineering*, vol. 10, no. 1, p. 18, Dec. 2022, doi: 10.3390/bioengineering10010018.
- [90] W. Ayadi, I. Charfi, W. Elhamzi, and M. Atri, 'Brain tumor classification based on hybrid approach', *Vis. Comput.*, vol. 38, no. 1, pp. 107–117, Jan. 2022, doi: 10.1007/s00371-020-02005-1.
- [91] Z. Atha and J. Chaki, 'SSBTCNet: Semi-Supervised Brain Tumor Classification Network', *IEEE Access*, vol. 11, pp. 141485–141499, 2023, doi: 10.1109/ACCESS.2023.3343126.
- [92] C. S. Rao and K. Karunakara, 'Efficient Detection and Classification of Brain Tumor using Kernel based SVM for MRI', *Multimed. Tools Appl.*, vol. 81, no. 5, pp. 7393–7417, Feb. 2022, doi: 10.1007/s11042-021-11821-z.
- [93] E. Irmak, 'Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework', *Iran. J. Sci. Technol. Trans. Electr. Eng.*, vol. 45, no. 3, pp. 1015–1036, Sep. 2021, doi: 10.1007/s40998-021-00426-9.
- [94] G. S. Tandel, A. Balestrieri, T. Jujaray, N. N. Khanna, L. Saba, and J. S. Suri, 'Multiclass magnetic resonance imaging brain tumor classification using artificial intelligence paradigm', *Comput. Biol. Med.*, vol. 122, p. 103804, Jul. 2020, doi: 10.1016/j.compbiomed.2020.103804.
- [95] H. Mzoughi *et al.*, 'Deep Multi-Scale 3D Convolutional Neural Network (CNN) for MRI Gliomas Brain Tumor Classification', *J. Digit. Imaging*, vol. 33, no. 4, pp. 903–915, Aug. 2020, doi: 10.1007/s10278-020-00347-9.
- [96] M. A. Khan *et al.*, 'Multimodal Brain Tumor Classification Using Deep Learning and Robust Feature Selection: A Machine Learning Application for Radiologists', *Diagnostics*, vol. 10, no. 8, p. 565, Aug. 2020, doi: 10.3390/diagnostics10080565.
- [97] B. Yin, C. Wang, and F. Abza, 'New brain tumor classification method based on an improved version of whale optimization algorithm', *Biomed. Signal Process. Control*, vol. 56, p. 101728, Feb. 2020, doi: 10.1016/j.bspc.2019.101728.

- [98] S. Deepak and P. M. Ameer, 'Brain tumor classification using deep CNN features via transfer learning', *Comput. Biol. Med.*, vol. 111, p. 103345, Aug. 2019, doi: 10.1016/j.combiomed.2019.103345.
- [99] J. M. Waghmare and B. D. Patil, 'Removal of Noises In Medical Images By Improved Median Filter', *Int. J. Eng. Sci.*, vol. 2, no. 7, pp. 49–53, 2013.
- [100] E. S. Gedraite and M. Hadad, 'Investigation on the effect of a Gaussian Blur in image filtering and segmentation', *Proc. Elmar - Int. Symp. Electron. Mar.*, pp. 393–396, 2011.
- [101] . S. . ., 'Performance Analysis of Image Filtering Algorithms for Mri Images', *Int. J. Res. Eng. Technol.*, vol. 03, no. 05, pp. 438–440, 2014, doi: 10.15623/ijret.2014.0305080.
- [102] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, 'Bilateral filtering: Theory and applications', *Found. Trends Comput. Graph. Vis.*, vol. 4, no. 1, pp. 1–73, 2009, doi: 10.1561/06000000020.
- [103] A. Buades, B. Coll, and J.-M. Morel, 'A Non-Local Algorithm for Image Denoising', in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, IEEE, pp. 60–65. doi: 10.1109/CVPR.2005.38.
- [104] M. Juneja *et al.*, 'Denoising of magnetic resonance images of brain tumor using BT-Autonet', *Biomed. Signal Process. Control*, vol. 87, p. 105477, Jan. 2024, doi: 10.1016/j.bspc.2023.105477.
- [105] I. Mahmoud El-Henawy, M. Elbaz, Z. H. Ali, and N. Sakr, 'Novel Framework of Segmentation 3D MRI of Brain Tumors', *Comput. Mater. Contin.*, vol. 74, no. 2, pp. 3489–3502, 2023, doi: 10.32604/cmc.2023.033356.
- [106] A. Sharma and V. Chaurasia, 'MRI denoising using advanced NLM filtering with non-subsampled shearlet transform', *Signal, Image Video Process.*, vol. 15, no. 6, pp. 1331–1339, Sep. 2021, doi: 10.1007/s11760-021-01864-y.
- [107] B. Kanoun, M. Ambrosanio, F. Baselice, G. Ferraioli, V. Pascasio, and L. Gomez, 'Anisotropic Weighted KS-NLM Filter for Noise Reduction in MRI', *IEEE Access*, vol. 8, pp. 184866–184884, 2020, doi: 10.1109/ACCESS.2020.3029297.
- [108] S. Saladi and N. Amutha Prabha, 'Analysis of denoising filters on <scp>MRI</scp> brain images', *Int. J. Imaging Syst. Technol.*, vol. 27, no. 3, pp. 201–208, Sep. 2017, doi: 10.1002/ima.22225.
- [109] M. Geng *et al.*, 'Triplet Cross-Fusion Learning for Unpaired Image Denoising in Optical Coherence Tomography', *IEEE Trans. Med. Imaging*, vol. 41, no. 11, pp. 3357–3372, Nov. 2022, doi: 10.1109/TMI.2022.3184529.
- [110] D. Sarwinda, R. H. Paradisa, A. Bustamam, and P. Anggia, 'Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer', *Procedia Comput. Sci.*, vol. 179, pp. 423–431, 2021, doi: 10.1016/j.procs.2021.01.025.
- [111] R. Karthickmanoj and N. Manoharan, 'Median Filter density', vol. 7, no. October, pp. 61–63, 2014.
- [112] J. George and S. P. Indu, 'Fast Adaptive Anisotropic Filtering for Medical Image Enhancement', in *2008 IEEE International Symposium on Signal Processing and Information Technology*, IEEE, Dec. 2008, pp. 227–232. doi: 10.1109/ISSPIT.2008.4775677.
- [113] J. Song *et al.*, 'Global and Local Feature Reconstruction for Medical Image Segmentation', *IEEE Trans. Med. Imaging*, vol. 41, no. 9, pp. 2273–2284, Sep. 2022, doi: 10.1109/TMI.2022.3162111.
- [114] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, 'Medical Image Analysis using Convolutional Neural Networks: A Review', *J. Med. Syst.*, vol. 42, no. 11, p. 226, Nov. 2018, doi: 10.1007/s10916-018-1088-1.

- [115] N. Siddique, P. Sidike, C. Elkin, and V. Devabhaktuni, 'U-Net and its variants for medical image segmentation: theory and applications', Nov. 2020, doi: 10.1109/ACCESS.2021.3086020.
- [116] H. Ullah, Y. Zhao, L. Wu, A. Noor, and L. Zhao, 'Multi-modal Medical Image Fusion Technique to Improve Glioma Classification Accuracy', in *2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP)*, IEEE, Oct. 2021, pp. 321–325. doi: 10.1109/ICSIP52628.2021.9689018.
- [117] S. Maqsood, R. Damaševičius, and R. Maskeliūnas, 'Multi-Modal Brain Tumor Detection Using Deep Neural Network and Multiclass SVM', *Medicina (B. Aires)*, vol. 58, no. 8, p. 1090, Aug. 2022, doi: 10.3390/medicina58081090.
- [118] S. Ahuja, B. K. Panigrahi, and T. K. Gandhi, 'Enhanced performance of Dark-Nets for brain tumor classification and segmentation using colormap-based superpixel techniques', *Mach. Learn. with Appl.*, vol. 7, p. 100212, Mar. 2022, doi: 10.1016/j.mlwa.2021.100212.
- [119] 'Rosebrock A.: Finding Extreme Points in Contours with OpenCV.PyImageSearch, 11', [Online]. Available: <https://pyimagesearch.com/2016/04/11/finding-extreme-points-in-contours-with-opencv/>
- [120] M. Toğaçar, B. Ergen, and Z. Cömert, 'BrainMRNet: Brain tumor detection using magnetic resonance images with a novel convolutional neural network model', *Med. Hypotheses*, vol. 134, p. 109531, Jan. 2020, doi: 10.1016/j.mehy.2019.109531.
- [121] J. Schmidhuber, 'Deep Learning in Neural Networks: An Overview', Apr. 2014, doi: 10.1016/j.neunet.2014.09.003.
- [122] Y. LeCun, K. Kavukcuoglu, and C. Farabet, 'Convolutional networks and applications in vision', in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, IEEE, May 2010, pp. 253–256. doi: 10.1109/ISCAS.2010.5537907.
- [123] S. J. Pan and Q. Yang, 'A Survey on Transfer Learning', *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.
- [124] K. Weiss, T. M. Khoshgoftaar, and D. Wang, 'A survey of transfer learning', *J. Big Data*, vol. 3, no. 1, p. 9, Dec. 2016, doi: 10.1186/s40537-016-0043-6.
- [125] K. Simonyan and A. Zisserman, 'Very Deep Convolutional Networks for Large-Scale Image Recognition', Sep. 2014.
- [126] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, 'Rethinking the Inception Architecture for Computer Vision', Dec. 2015.
- [127] K. He, X. Zhang, S. Ren, and J. Sun, 'Deep Residual Learning for Image Recognition', in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- [128] M. Tan and Q. V. Le, 'EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks', May 2019.
- [129] R. H. Ramdlon, E. Martiana Kusumaningtyas, and T. Karlita, 'Brain Tumor Classification Using MRI Images with K-Nearest Neighbor Method', in *2019 International Electronics Symposium (IES)*, IEEE, Sep. 2019, pp. 660–667. doi: 10.1109/ELECSYM.2019.8901560.
- [130] G. Cinarer and B. G. Emiroglu, 'Classification of Brain Tumors by Machine Learning Algorithms', in *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, IEEE, Oct. 2019, pp. 1–4. doi: 10.1109/ISMSIT.2019.8932878.
- [131] Y. Yang *et al.*, 'Glioma Grading on Conventional MR Images: A Deep Learning Study With Transfer Learning', *Front. Neurosci.*, vol. 12, Nov. 2018, doi: 10.3389/fnins.2018.00804.
- [132] M. A. Naser and M. J. Deen, 'Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images', *Comput. Biol. Med.*, vol. 121, p. 103758,

- Jun. 2020, doi: 10.1016/j.compbimed.2020.103758.
- [133] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, 'DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs', *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018, doi: 10.1109/TPAMI.2017.2699184.
 - [134] V. Dumoulin and F. Visin, 'A guide to convolution arithmetic for deep learning', Mar. 2016.
 - [135] V. Singh, S. Sharma, S. Goel, S. Lamba, and N. Garg, 'Brain Tumor Prediction by Binary Classification Using VGG-16', in *Smart and Sustainable Intelligent Systems*, Wiley, 2021, pp. 127–138. doi: 10.1002/9781119752134.ch9.
 - [136] A. Aziz *et al.*, 'An Ensemble of Optimal Deep Learning Features for Brain Tumor Classification', *Comput. Mater. Contin.*, vol. 69, no. 2, pp. 2653–2670, 2021, doi: 10.32604/cmc.2021.018606.
 - [137] I. Aboussaleh, J. Riffi, A. M. Mahraz, and H. Tairi, 'Brain Tumor Segmentation Based on Deep Learning's Feature Representation', *J. Imaging*, vol. 7, no. 12, p. 269, Dec. 2021, doi: 10.3390/jimaging7120269.
 - [138] T. Kalaiselvi, T. Padmapriya, P. Sriramakrishnan, and V. Priyadharshini, 'Development of automatic glioma brain tumor detection system using deep convolutional neural networks', *Int. J. Imaging Syst. Technol.*, vol. 30, no. 4, pp. 926–938, Dec. 2020, doi: 10.1002/ima.22433.
 - [139] G. Lin, C. Shen, A. van den Hengel, and I. Reid, 'Efficient piecewise training of deep structured models for semantic segmentation', Apr. 2015.
 - [140] M. Oršić and S. Šegvić, 'Efficient semantic segmentation with pyramidal fusion', *Pattern Recognit.*, vol. 110, p. 107611, Feb. 2021, doi: 10.1016/j.patcog.2020.107611.
 - [141] Z. Liu, X. Li, P. Luo, C.-C. Loy, and X. Tang, 'Semantic Image Segmentation via Deep Parsing Network', in *2015 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Dec. 2015, pp. 1377–1385. doi: 10.1109/ICCV.2015.162.
 - [142] F. Yu and V. Koltun, 'Multi-Scale Context Aggregation by Dilated Convolutions', Nov. 2015.
 - [143] J. Uhrig, M. Cordts, U. Franke, and T. Brox, 'Pixel-Level Encoding and Depth Layering for Instance-Level Semantic Labeling', 2016, pp. 14–25. doi: 10.1007/978-3-319-45886-1_2.
 - [144] T. M. Ali *et al.*, 'A Sequential Machine Learning-cum-Attention Mechanism for Effective Segmentation of Brain Tumor', *Front. Oncol.*, vol. 12, Jun. 2022, doi: 10.3389/fonc.2022.873268.
 - [145] X. Guan, Y. Zhao, C. O. Nyatega, and Q. Li, 'Brain Tumor Segmentation Network with Multi-View Ensemble Discrimination and Kernel-Sharing Dilated Convolution', *Brain Sci.*, vol. 13, no. 4, p. 650, Apr. 2023, doi: 10.3390/brainsci13040650.
 - [146] M. Soltaninejad *et al.*, 'Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI', *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 2, pp. 183–203, Feb. 2017, doi: 10.1007/s11548-016-1483-3.
 - [147] A. Kermi, I. Mahmoudi, and M. T. Khadir, 'Deep Convolutional Neural Networks Using U-Net for Automatic Brain Tumor Segmentation in Multimodal MRI Volumes', 2019, pp. 37–48. doi: 10.1007/978-3-030-11726-9_4.
 - [148] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, 'Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks', 2017, pp. 506–517. doi: 10.1007/978-3-319-60964-5_44.
 - [149] U. Baid *et al.*, 'The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification', Jul. 2021.
 - [150] M. M. E. Yurtsever, Y. Atay, B. Arslan, and S. Sagiroglu, 'Development of brain tumor

- radiogenomic classification using GAN-based augmentation of MRI slices in the newly released gazi brains dataset', *BMC Med. Inform. Decis. Mak.*, vol. 24, no. 1, p. 285, Oct. 2024, doi: 10.1186/s12911-024-02699-6.
- [151] D. Kollias, K. Vendal, P. Gadhavi, and S. Russom, 'BTDNet: A Multi-Modal Approach for Brain Tumor Radiogenomic Classification', *Appl. Sci.*, vol. 13, no. 21, p. 11984, Nov. 2023, doi: 10.3390/app132111984.
 - [152] S. Das, 'Optimizing Prediction of MGMT Promoter Methylation from MRI Scans using Adversarial Learning', in *2022 IEEE 34th International Conference on Tools with Artificial Intelligence (ICTAI)*, IEEE, Oct. 2022, pp. 1047–1054. doi: 10.1109/ICTAI56018.2022.00160.
 - [153] J. Amin, M. A. Anjum, N. Gul, and M. Sharif, 'A secure two-qubit quantum model for segmentation and classification of brain tumor using MRI images based on blockchain', *Neural Comput. Appl.*, vol. 34, no. 20, pp. 17315–17328, Oct. 2022, doi: 10.1007/s00521-022-07388-x.
 - [154] 'Sartaj Bhuvaji, Ankita Kadam, Prajakta Bhumkar, and Sameer Dedge, "Brain Tumor Classification (MRI) | Kaggle," The folder contains MRI data. The images are already split into Training and Testing folders. Each folder has more four subfolders. , 2020. http'.
 - [155] 'AHMED HAMADA, "Br35H:: Brain Tumor Detection 2020," kaggle, 2020. <https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection?select=no>'.
 - [156] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, 'SLIC Superpixels Compared to State-of-the-Art Superpixel Methods', *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012, doi: 10.1109/TPAMI.2012.120.
 - [157] S. Wang and Y. Zhao, 'A Novel Patch-Based Multi-Exposure Image Fusion Using Super-Pixel Segmentation', *IEEE Access*, vol. 8, pp. 39034–39045, 2020, doi: 10.1109/ACCESS.2020.2975896.
 - [158] I. Goodfellow *et al.*, 'Generative adversarial networks', *Commun. ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2014, doi: 10.1145/3422622.
 - [159] M. Ding, B. Xiao, N. Codella, P. Luo, J. Wang, and L. Yuan, 'DaViT: Dual Attention Vision Transformers', Apr. 2022.
 - [160] S. Maqsood, R. Damaševičius, and R. Maskeliūnas, 'Multi-Modal Brain Tumor Detection Using Deep Neural Network and Multiclass SVM', *Medicina (B. Aires)*, vol. 58, no. 8, p. 1090, Aug. 2022, doi: 10.3390/medicina58081090.
 - [161] M. Masood *et al.*, 'A Novel Deep Learning Method for Recognition and Classification of Brain Tumors from MRI Images', *Diagnostics*, vol. 11, no. 5, p. 744, Apr. 2021, doi: 10.3390/diagnostics11050744.
 - [162] Z. Rasheed *et al.*, 'Brain Tumor Classification from MRI Using Image Enhancement and Convolutional Neural Network Techniques', *Brain Sci.*, vol. 13, no. 9, p. 1320, Sep. 2023, doi: 10.3390/brainsci13091320.
 - [163] M. A. Gómez-Guzmán *et al.*, 'Classifying Brain Tumors on Magnetic Resonance Imaging by Using Convolutional Neural Networks', *Electronics*, vol. 12, no. 4, p. 955, Feb. 2023, doi: 10.3390/electronics12040955.
 - [164] M. A. Khan *et al.*, 'Multimodal brain tumor detection and classification using deep saliency map and improved dragonfly optimization algorithm', *Int. J. Imaging Syst. Technol.*, vol. 33, no. 2, pp. 572–587, Mar. 2023, doi: 10.1002/ima.22831.
 - [165] G. Karayegen and M. F. Aksahin, 'Brain tumor prediction on MR images with semantic segmentation by using deep learning network and 3D imaging of tumor region', *Biomed. Signal Process. Control*, vol. 66, p. 102458, Apr. 2021, doi: 10.1016/j.bspc.2021.102458.
 - [166] M. O. Khairandish, M. Sharma, V. Jain, J. M. Chatterjee, and N. Z. Jhanjhi, 'A Hybrid CNN-SVM Threshold Segmentation Approach for Tumor Detection and Classification

- of MRI Brain Images', *IRBM*, vol. 43, no. 4, pp. 290–299, Aug. 2022, doi: 10.1016/J.IRBM.2021.06.003.
- [167] 'BRATS - SICAS Medical Image Repository'.
 - [168] S. Liu *et al.*, 'Time Series Anomaly Detection With Adversarial Reconstruction Networks', *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 4, pp. 4293–4306, Apr. 2023, doi: 10.1109/TKDE.2021.3140058.
 - [169] B. Zhao, X. Li, X. Lu, and Z. Wang, 'A CNN–RNN architecture for multi-label weather recognition', *Neurocomputing*, vol. 322, pp. 47–57, Dec. 2018, doi: 10.1016/j.neucom.2018.09.048.
 - [170] X. Zhou, Y. Li, and W. Liang, 'CNN-RNN Based Intelligent Recommendation for Online Medical Pre-Diagnosis Support', *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 18, no. 3, pp. 912–921, May 2021, doi: 10.1109/TCBB.2020.2994780.
 - [171] S. Montaha, S. Azam, A. K. M. R. H. Rafid, M. Z. Hasan, A. Karim, and A. Islam, 'TimeDistributed-CNN-LSTM: A Hybrid Approach Combining CNN and LSTM to Classify Brain Tumor on 3D MRI Scans Performing Ablation Study', *IEEE Access*, vol. 10, pp. 60039–60059, 2022, doi: 10.1109/ACCESS.2022.3179577.
 - [172] B. Deepa, M. Murugappan, M. G. Sumithra, M. Mahmud, and M. S. Al-Rakhami, 'Pattern Descriptors Orientation and MAP Firefly Algorithm Based Brain Pathology Classification Using Hybridized Machine Learning Algorithm', *IEEE Access*, vol. 10, pp. 3848–3863, 2022, doi: 10.1109/ACCESS.2021.3100549.
 - [173] M. Zhang, W. Song, and J. Zhang, 'A Secure Clinical Diagnosis With Privacy-Preserving Multiclass Support Vector Machine in Clouds', *IEEE Syst. J.*, vol. 16, no. 1, pp. 67–78, Mar. 2022, doi: 10.1109/JSYST.2020.3027758.
 - [174] E. U. Haq, H. Jianjun, X. Huarong, K. Li, and L. Weng, 'A Hybrid Approach Based on Deep CNN and Machine Learning Classifiers for the Tumor Segmentation and Classification in Brain MRI', *Comput. Math. Methods Med.*, vol. 2022, pp. 1–18, Aug. 2022, doi: 10.1155/2022/6446680.