Deep Neural Network Paradigms for Improved Brain Tumor Diagnosis in Magnetic Resonance Imaging

Mohammad Shahnawaz Shaikh [1*], Aparajita Biswal[2], Neelam Agrawal[3], Nilesh Khodifad[4] and Bhavesh Atul Bhai Vaghela[5]

[1,2,3,4,5] Parul Institute of Engineering and Technology, Parul University, Vadodara (Gujarat) - 391760, India
[*1]msnshaikh1@gmail.com, [2]aparajitabiswal1384@gmail.com, [3]agrawal.neelam123@gmail.com, [4]nkhodifad@gmail.com, [5]vaghelabhavesh404@gmail.com

Abstract. This paper aims at discussing the roles of accurate and efficient diagnosis of brain tumors in enhancing patient diagnostic results and treatment plans. Magnetic Resonance Imaging (MRI) plays a significant role in the diagnosis of brain tumors because of high-resolution imaging of the brain region. But manual evaluation of MRI scans is a very time consuming and lacks reproducibility. Advanced architectures have come up as revolutionary solutions which involve deep neural network paradigms which improve diagnosis of tumor. Establishing the foundation of deep learning methods applied to segments, classify, or detect brain tumors, this paper covers a number of approaches: Convolution Neural Networks (CNNs); Transformer-based; CNN + Transformer; and generative. Certain tasks include spatial complexity, data deficiency, and noisy imaging conditions that U-Net, Vision Transformers, and GANs solve, provide high accuracy and efficiency. Furthermore, techniques, such as self-supervised and transfer learning enhance model generalization, even with limited amounts of data. This paper builds upon the last decade of developments in the deep neural network paradigms and demonstrates how they might transform the brain tumor diagnosis. These results suggest that further work should be done to build accurate, explainable, and clinically augmented models to make substantial contributions to realize precise medicine and medical imaging.

Keywords: Medical Imaging, Convolutional Neural Networks (CNNs), Transformer-Based Models, Vision Transformers (ViTs), Hybrid Architectures, Generative Adversarial Networks (GANs), U-Net.

1 Introduction

Medical diagnosis and treatment of brain tumors is still a big challenge and needs early and precise detection to improve the patient's outcome. Due to the superior soft tissue contrast, Magnetic Resonance Imaging (MRI) is the most important diagnostic and characterizing modality of brain tumors. While the current manual interpretation of MRI scans is labor intensive and subject to variability, there is the need for

automated, robust diagnostic tools. Over recent years we have seen the emergence of deep neural network paradigms as transformative solutions in medical imaging which have used artificial intelligence to boost detection, segmentation, and classification of tumors. Specifically, these paradigms leverage well developed advanced architectures such as Convolutional Neural Networks (CNNs), Transformers, and hybrid approaches to tackle the specific shortcomings of medical imaging: learning from spatial as well as temporal and contextual clues.

The combination of these models has improved tumor diagnostic accuracy and efficiency, brought about personalized treatment strategies, and improved clinical decision-making. Additionally, data scarcity, a common limit Many rations in medical imaging, was addressed by novel techniques such as generative models and self-supervised learning to augment datasets and model generalization. In this work, we look at the evolution of deep neural network paradigms towards brain tumor diagnosis and their specific applications on MRI based analysis. We delve into the architectural advancements and how they contribute to advancing diagnostic precision to highlight the vital importance of these technologies in the development of medical imaging.

2   Literature Survey

Deep learning has been applied in medical imaging to revolutionize brain tumor diagnosis. The last part of this thesis reviews some of the most exciting mathematical studies and advancements with specific focus on the key methodologies that have led to the state of the art as it currently exists. CNNs have emerged as the base of image analysis in medical imaging. With that, Khizevsky et al. [1] introduced AlexNet, and proved the CNNs ability of performing in the recognition tasks of visuality. After that, other architectures like VGGNet [2] and ResNet [3] developed the feature extraction aspect by increasing the net-work depth in way to solve the problems like the vanishing gradients. In supervised medical imaging applications, for example tumor segmentation in MRI scans, U-Net [4], an encoder decoder-based architecture, has become pretty popular yielding high delineation of tumor boundaries with high accuracy. U-Net has been shown to perform well in segments of gliomas, meningioma, and pituitary tumors [5]. Originally used for natural language processing [6], Transformers have recently proved beneficial in medical image analysis. Self-attention mechanism [7] employed in ViTs allows them to use the global contextual information and outperforms CNNs in some segmentation tasks. Researchers have shown that ViTs are good at processing the complex spatial relationships in the multi-modal MRI datasets [8].

Recently, hybrid architectures formed of CNNs together with other paradigms (such as Recurrent Neural Networks (RNNs) and attention mechanisms) have been powerful tools for tackling complex imaging tasks. For example, CNN-RNN hybrid models were able to improve the analysis of temporal imaging data [9]. Moreover, attention mechanisms incorporated within CNNs, as exemplified by Attention U-Net [10], allow these CNNs to concentrate on diagnostically important regions and have higher segmentation accuracy in noisy imaging conditions. To address its scarcity, Generative models like Generative Adversarial Networks (GANs) [11] and

Variational Autoencoders (VAEs) [12] seek to generate synthetic data from few examples in such medical datasets. With particular success, GANs have been used in generating synthetic MRI scans to supplement training datasets, to improve model robustness and performance. In addition, they have been shown to simulate rare tumor types and improve the quality of low-resolution scans [13]. For benchmarking model performance standardized datasets, e.g., those derived from the BRATS (Brain Tumor segmentation) challenge [14], are available. Often used metrics are Dice Similarity Coefficient (DSC), Intersection over Union (IoU), sensitivity and specificity. DSC values greater than 0.90 for high contrast tumor regions have been re-ent studies for state-of-the-art models [15]. To deal with the limitation of the datasets with limited annotated data, techniques including transfer learning and self-supervised learning have been applied. On medical imaging tasks with medical images as input, we find that pretrained models including ResNet [16] and DenseNet [16], fine-tuned are able to achieve significant increases in diagnostic accuracy. Last, collaborative model training across institutions in federated learning frameworks [17] have been enabled while patient privacy is maintained.

3   Deep Neural Network Paradigms

Advanced architectures have leveraged their ability to process and analyze MRI scans to transform brain tumor diagnosis to deep neural network paradigms (Fig. 1). These paradigms employ a specific special model for tasks like tumor segmentation, detection and classification. Convolutional Neural Networks (CNNs) for their ability to learn the spatial hierarchies for image analysis have been the key approaches, and their variants, for instance, U-Net, for accurate tumor boundary segmentation was what they facilitated. They include emerging frameworks, such as Transformer based models and hybrid architectures, which build strengths from multiple paradigms (such as CNNs with attention mechanism; or using recurrent layers over CNNs), to overcome spatial and temporal complexities in medical imaging. In addition, limited datasets are enriched with innovative methods such as self supervised learning and generative models which together increase diagnostic accuracy. They are paradigms for robust and efficient tumor diagnosis and treatment planning [18].
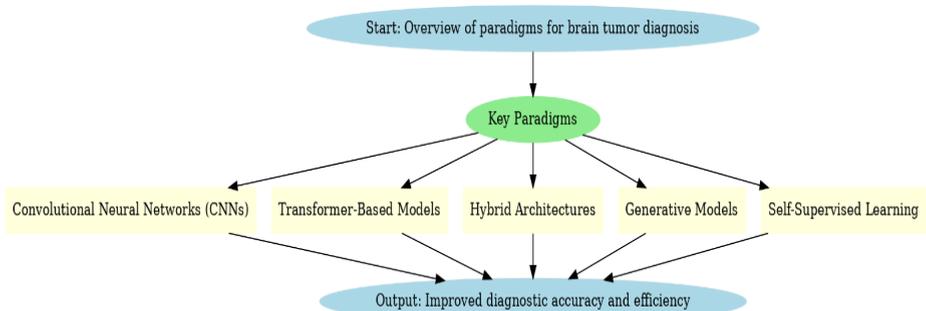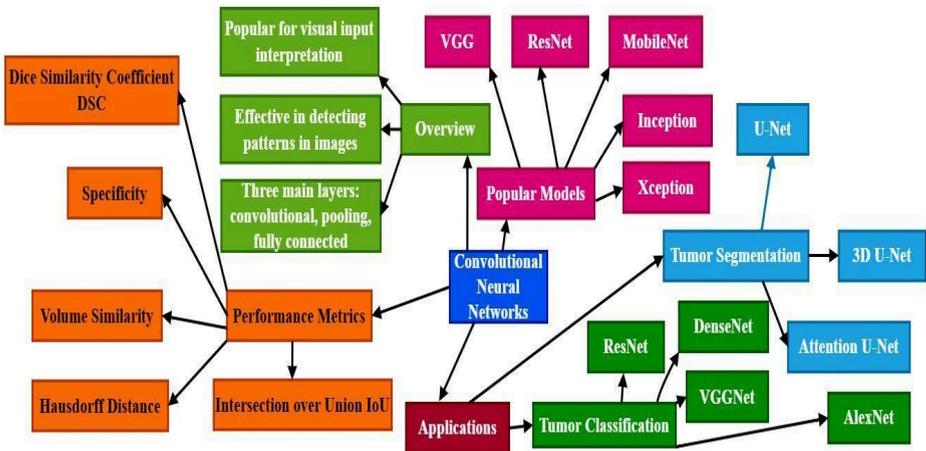

Fig. 1. Deep Neural Network Paradigms

4   Convolutional Neural Networks (CNNs)

Several architectures exist in brain tumor diagnosis but CNN holds its ground as the most prominent architecture because it's well-suited for interpreting visual input. CNN has been demonstrated for image pattern detection as well as for items, classes, and categories recognition. They consist mostly of three layers: The combination of convolutional, pooling, and fully linked. Depending on how complex we want to get with the issue statement, developers can add layers to a CNN design. There are very popular CNN models for detection of brain cancers such as VGG, ResNet, Inception, Xception, and MobileNet [19]. Convolutional Neural Network (CNN) is a methodolical method in medical image processing domain (Fig. 2). CNN was also part of artificial neural net-works; a subset specifically designed to only use method componnet knowledge for picture recognition and processing. Deep learning in CNN is a powerful im-age processing and computing technique that can perform generative and descriptive tasks. Hierarchical hierarchical feature extraction turned out to revolutionize image analysis with CNNs [20].

Fig. 2. Scope of Convolutional Neural Networks (CNNs)

4.1   Tumor Segmentation

Tumor segmentation is the task of discerning and demarcating tumor regions within medical images, such as MRI scans, crucial for diagnosis, therapy planning, and follow up. Thanks to their capacity to learn spatial hierarchies, CNNs, and especially architectures such as U-Net, have heightened this line of work. The use of U-Net's encoder–decoder structure together with skip connections allows the encoder to have good context, while the decoder takes advantage of the skip connections and reconstructs with great precision [21]. This is very helpful for accurate tumor boundary detection. Further variants such as 3D U-Net that is optimized for volumetric imaging and Attention U-Net, which is especially tuned to tumor relevant regions, improve segmentation accuracy. In particular, these models are excellent for tasks in brain tumor segmentation where sub-region classification (e.g., edema, necrosis, etc.) is necessary for treatment planning, and tend to do well in classifying meningiomas as well as pituitary tumors among noisy brain structures [22].

4.2   Performance Metrics for Tumor Segmentation

Tumor segmentation in medical imaging is typically evaluated using the following performance metrics shown in Table 1:

Table 1. Summary of Key Performance Statistics

Metric	Typical Value Range	Remarks
Dice Similarity	0.85–0.95	Higher for well-defined tumors
IoU	0.70–0.90	IoU > 0.75 indicates good overlap
Sensitivity	0.85–0.95	Important for detecting all tumor areas
Specificity	0.90–0.95	Avoiding false positives
Precision	0.85–0.90	Reducing false positives
Volume Similarity	0.80–0.90	Correlates with tumor volume
Hausdorff Distance	5–10 mm	Lower values are better for boundary precision

4.3   Common Benchmarks for Tumor Segmentation Using CNNs for Dice Similarity Coefficient (DSC)

To benchmark tumor segmentation models, various publicly available datasets are used shown in Table 2. Here are some key results from well-known brain tumor segmentation datasets:

Table 2. Benchmarks for Tumor Segmentation Using CNNs for Dice Similarity Coefficient (DSC)

Challenge/Model	Region/Type	Performance (DSC)	Remarks
BRATS 2020 (Brain Tumor Segmentation Challenge)	Glioma Enhancing Tumor (ET)	0.80–0.85	Excellent segmentation of high-contrast tumor regions.
	Tumor Core (TC)	0.85–0.90	Good segmentation performance for well-defined tumor regions.
	Whole Tumor (WT)	0.90–0.95	High accuracy in segmenting entire tumor regions, including edema.
ISLES (Ischemic Stroke Lesion Segmentation Challenge)	Stroke Lesions	0.75–0.90	Performance varies by lesion type and image quality.
Deep Brain Tumor Segmentation Models	Gliomas (U-Net-based models)	0.85–0.92	3D U-Net shows improvements for volumetric segmentation.
	Gliomas (ResNet-based models)	0.80–0.88	Effective for low-contrast tumor regions.
Multi-modal MRI Segmentation	Complex datasets (e.g., BRATS)	0.85–0.90 or higher	Leveraging multi-modal MRI (T1, T2, FLAIR) improves performance.

4.4   Tumor Classification

In the realm of medical images, tumor classification groups together such brain tumors as gliomas, meningiomas, and pituitary adenomas based on features from the medical images to help orchestrate treatment planning and prediction of outcome [23]. In this task, ResNet is a common architecture, in which residual connections are used to combat vanishing gradients and enable the use of deep networks to differentiate subtle differences among tumor types. In dense layer connections denseNet can capture intricate image features and use feature reuse to increase the classification accuracy on small datasets [24]. Diverse medical imaging needs can be functioned fitting the simpler classification tasks or when computational resources are limited with the older architectures like VGG Net and Alex Net being fallible [25].

4.5   Performance Metrics for Tumor Classification

Summary of Typical Performance Statistics for Tumor Classification is given in Table 3

Table 3. Summary of Typical Performance Statistics for Tumor Classification

Metric	Typical Value Range	Remarks
Accuracy	85–98%	Higher for binary classification tasks.
Precision	0.85–0.95	Reduces false positives.
Recall (Sensitivity)	0.85–0.95	Ensures most tumors are detected.
Specificity	0.90–0.98	Reduces false negatives.
F1-Score	0.85–0.95	Balances precision and recall.
AUC-ROC	0.90–0.98	Indicates strong discriminative ability.

5   Advanced CNN Architectures

Convolutional neural networks (CNNs) have seen tremendous improvements in the evolution in the diagnosis of omre fields, especially in medical imaging, including brain tumor diagnosis. I present below some of the key advancements in CNN architectures that have pushed the efficacy of these models much further [26]. In medical imaging, the analysis of brain tumor is complicated by restricted data, computational efficiency, and the complicated stru-cure of the tumor, where advanced architectures improve its analysis. ResNet that uses the residual connections tackles the vanishing gradient problem, which permits to establish deeply, and precisely classify tumors. Dense connections in DenseNet promote feature reuse and reduce redundancy leading to improved multi-class tumor classification [27]. MobileNet's depth-wise separable convolutions not only increase efficiency but also support real time diagnostics on mobile devices for telemedicine [28]. Attention U-Net enhanced focusing attention on tumor boundaries through attention mechanisms, improving accuracy in noisy environments shown in Table 4.

Table 4. Benchmark Performance of Advanced Architectures

Architecture	Task	DSC	Precision	Recall	F1-Score	AUC-ROC	Inference Time
ResNet-50	Classification	-	0.93	0.94	0.93	0.96	25 ms/image
DenseNet-121	Classification	-	0.94	0.95	0.94	0.97	30 ms/image
U-Net	Segmentation	0.89	0.88	0.9	0.89	-	50 ms/image
Attention U-Net	Segmentation	0.93	0.91	0.92	0.92	-	55 ms/image
MobileNet	Classification	-	0.9	0.91	0.9	0.94	12 ms/image
EfficientNet-B0	Classification	-	0.92	0.93	0.92	0.95	15 ms/image
Vision Transformer	Classification	-	0.91	0.92	0.91	0.95	150 ms/image
Hybrid CNN-Transformer	Classification	-	0.94	0.95	0.94	0.97	120 ms/image

6   Transfer Learning

The challenges of limited annotated datasets and computational resources are overcome by the transfer learning paradigm and have revolutionized medical imaging (Table 5). Obtaining large, diverse datasets together with precise annotations is often a costly and time-consuming process in medical imaging [29]. A technique that can solve this problem is through transfer learning: you can use pre-trained models, like ResNet, DenseNet, or EfficientNet, and train them on ImageNet having a much larger dataset. However, in both cases, these models have learned rich and transferable features from millions of images, which can be fine-tuned for medical tasks, for example tumor classification and organ segmentation, or disease detection. By transferring these learnt features, we remove the need to train models from scratch, reducing training time and computational cost, and alleviating overfitting [30].

Transfer learning in practice consists in replacing the output layers of the pretrained network with task specific layers adapted to medical imaging applications. We then tune the model on medical dataset using a lower learning rate to adapt the generic features to the domain specific characteristics of medical images [31]. Freezing initial layers, gradual unfreeze, and selective fine-tuning techniques prevent the model from forgetting independent information of the original task, but continues to operate within a sparse, superficial parametrization of it, smoothly emerging substructures of the specific target task. This approach is particularly useful for medical imaging tasks where spatial patterns, shape, and texture are important to diagnosis [32].

Table 5. Performance Metrics for Transfer Learning in Medical Imaging

Metric	Formula/Definition	Application	Significance
Accuracy	$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$	Classification tasks (e.g., tumor detection)	Overall correctness of predictions; useful for balanced datasets.
Precision	$Precision = \frac{TP}{TP+FP}$	Classification (e.g., cancer detection)	Reduces false positives; critical in diagnostic tasks.
Recall (Sensitivity)	$Sensitivity = \frac{TP}{TP+FN}$	Disease detection (e.g., diabetic retinopathy)	Minimizes false negatives, ensuring critical cases are identified.
Specificity	$Specificity = \frac{TN}{TN+FP}$	Robustness evaluation	Reduces false positives; useful for high-specificity tasks like cancer diagnosis.
F1-Score	$Specificity = \frac{TP}{TP+\frac{1}{2}(FP+FN)}$	Imbalanced datasets	Balances precision and recall for better evaluation.
AUC-ROC	Area under the Receiver Operating Characteristic curve	Disease classification	Evaluates model's ability to distinguish between classes.
AUC-PR	Area under the Precision-Recall curve	Imbalanced datasets	More informative for datasets with skewed class distributions.
Hausdorff Distance	Largest distance between predicted and ground truth boundaries	Segmentation boundary alignment	Indicates alignment precision; lower values are better.
Pixel Accuracy	Percentage of correctly classified pixels	Image segmentation	Simple measure of correctness for all pixels; less useful for imbalanced regions.

7   Recurrent Neural Networks (RNNs)

In medical imaging, tasks that require the analysis of sequential or time dependent data, Recurrent Neural Networks (RNNs) and attention mechanisms have become transformative tools. Temporal and spatial correlations are common in medical imaging: longitudinal scans are used to monitor pathologic progression, and echocardiography captures the temporal and spatial dynamics of cardio circulatory motion. Such temporal dependencies are the forte of RNNs, including innovative variants such as Long Short-Term Memory (LSTMs) and Gated Recurrent Unit (GRUs), which maintain a memory for states in past and capture long term contextual information [33].

7.1   Performance Metrics for Recurrent Neural Networks (RNNs) and Attention Mechanisms in Medical Imaging

A set of metrics, applicable for evaluating the performance of Recurrent Neural Networks (RNNs) embedded within an attention mechanism, was developed, in order to capture temporal dependency, select the most relevant region of interest, and generalize from seen data to unseen data in medical imaging. Key performance metrics used for this purpose are listed below in table 6.

Table 6. Performance Metrics for Recurrent Neural Networks (RNNs)

Metric	Description	Significance	Typical Range
Accuracy	Proportion of correctly classified samples.	Evaluates overall model performance in classification tasks like tumor detection.	85%–95%
Precision	Ratio of true positives to total predicted positives.	Reduces false positives in anomaly detection and disease progression analysis.	0.85–0.95
Recall (Sensitivity)	Ratio of true positives to total actual positives.	Ensures critical regions or temporal changes are identified in time-series data.	0.90–0.96
Specificity	Ratio of true negatives to total actual negatives.	Evaluates robustness in avoiding false positives, especially in longitudinal imaging.	0.88–0.95
F1-Score	Harmonic mean of precision and recall.	Balances precision and recall for imbalanced datasets, ensuring comprehensive evaluation.	0.85–0.95
AUC-ROC	Area under the Receiver Operating Characteristic curve.	Measures the model's ability to distinguish between positive and negative cases at varying thresholds.	0.90–0.97
Temporal Attention Score	Weighted score for relevant time points in sequences.	Evaluates the effectiveness of attention mechanisms in prioritizing critical temporal changes.	Application-specific

8   Generative Models

Generative models such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs) are popular for their ability to generate high quality synthetic data to counter the ubiquitous problem of small, annotated datasets in medical research as well as in clinical practice. By playing these roles; augmenting training data, simulating rare diseases, or simply improving image quality, these models are an invaluable addition to augmenting the performance of machine learning algorithms in medical applications. Moving on below, we explore generative models' impact on all sorts of applications in medical imaging, and where they can go from here [34].

8.1   Applications of Generative Models in Medical Imaging

Addressing the challenges of data scarcity, diagnostic precision and privacy, generative models, such as GANs and VAEs, are transforming medical imaging. Through data augmentation, they generate synthetic medical images to augment training datasets which reduce overfitting and improve model generalization. Through imaging synthesis, they transfer between imaging modalities or improve image quality making diagnostics better. Generative models enable accurate tumor detection and segmentation in super resolution imaging, which generate high resolution images from low quality inputs [35].

8.2   Benefits of Generative Models in Medical Imaging

Addressing critical challenges, generative models are important for progress in medical imaging. For the rare conditions they do allow data augmentation, which creates synthetic datasets which help machine learning models detect severe diseases. These models generate such diverse data to make robustness better and overfitting weaker such that it generalizes better. By reducing the need for expensive, time intensive data collection, they also offer cost and time efficiency. In addition, generative models improve image quality by converting low quality scans to high resolution images improving diagnostic accuracy. Notably, they assist with privacy guaranteeing by producing synthetic datasets that enable anonymous sharing of data for research purposes.

9   Hybrid Architectures in Medical Imaging

In recent years, medical imaging has emerged as an edge related to hybrid architectures, that is, combinations of several neural-network models. The architectures are these and they want to make use of their strengths—the strengths of one type of neural network versus one over others: The same method has one neural network excellent at spatial features, another at an onset dependency, one for contextual information, and so on. Medical imaging problems are tractable for hybrid systems complexity, and models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Attention Mechanisms, and Generative Models (GANs) give rise to coupled disease detection, segmentation, and diagnosis. Hybrid models provide a robust and comprehensive solutions to analyze rich spatial and temporal medical images for analysis. We can better integrate CNNs as an RNN or coalesces attentions with GANs to better extract features and model sequential data in better and more reliable way [36].

9.1   Key Components of Hybrid Architectures

Advanced architectures designed to solve specific tasks have revolutionized neural networks in medical imaging. Since Convolutional Neural Networks (CNNs) are efficient at discerning the spatial features in the images, they are the most appropriate tools for tasks like segmentation, classification, and detection of any abnormality in

tumors or organ abnormalities. RNNs (more specifically LSTMs) can process temporal sequences to study dynamic imaging data to track disease progression and to predict health outcomes. Attention mechanisms bring focus to the diagnostically most interesting regions leading to more sensitive detection of small tumors. Concomitantly, GANs and VAEs as generic models for generating synthetic images can be applied to repositories and obtain more realistic synthetic images to compensate data scarcity and improve model performance.

9.2   Performance Metrics for Hybrid Architectures in Medical Imaging

A robust approach to medical image analysis is presented using hybrid architectures composed of multiple neural network models, including CNNs, RNNs, attention mechanisms and GANs. We evaluate performance of these models using a combination of traditional metrics in machine learning, and more specific measurements for medical imaging tasks. The effectiveness of hybrid architectures in medical imaging is assessed with the help of the following key performance metrics shown in Table 7.

Table 7. Performance Metrics for Hybrid Architectures in Medical Imaging

Metric	Description	Relevance to Hybrid Architectures
Accuracy	The percentage of correctly predicted outcomes out of the total predictions.	Measures the overall effectiveness of the hybrid model in classifying images or detecting anomalies.
Precision	The ratio of true positive predictions to the total predicted positives (i.e., true positives + false positives).	Critical for ensuring that the model's predictions of medical conditions are accurate and minimizing false diagnoses.
Recall (Sensitivity)	The ratio of true positive predictions to the total actual positives (i.e., true positives + false negatives).	Important in medical imaging to ensure that the model identifies all potential disease instances, reducing missed diagnoses.
F1-Score	The harmonic means of precision and recall, providing a balance between the two.	Used when both precision and recall are important, offering a single measure of the model's accuracy in diagnosing medical conditions.
Intersection over Union (IoU)	The ratio of the intersection of predicted and ground truth regions to their union.	Essential in segmentation tasks (e.g., tumor segmentation) to evaluate the overlap between predicted and actual areas.
Dice Similarity Coefficient (DSC)	Measures the similarity between two sets, often used for image segmentation.	Commonly used for evaluating the performance of segmentation tasks, such as delineating tumors or other anatomical structures in medical images.
Area under the ROC Curve (AUC-ROC)	The area under the Receiver Operating Characteristic curve, which plots true positive rate vs. false positive rate.	Evaluates the discriminative ability of the model, providing insights into its performance across various decision thresholds.
Area under the Precision-Recall Curve (AUC-PR)	A variation of AUC-ROC, focusing on the trade-off between precision and recall.	Particularly useful when dealing with imbalanced datasets, where the cost of false negatives (missed diagnoses) is high.
Mean Absolute Error (MAE)	The average of the absolute errors between predicted and actual values.	Useful in regression tasks (e.g., predicting tumor size or progression) to measure how close the predictions are to true values.

10   Conclusion

The insertion of the deep neural network paradigms towards medical imaging has brought forth a revolutionary time for brain tumor diagnosis, and specifically for the study of magnetic resonance imaging (MRI) scans. On tasks like tumor segmentation, classification, and detection, techniques such as Convolutional Neural Networks (CNNs) and Transformer based models, and hybrid architectures are shown to have achieved outstanding performance. The effect on diagnostic accuracy has been enhanced with these developments, but they have also ushered in personalized treatment planning. Furthermore, self-supervised learning and generative models have tackled key obstacles to robust model training, such as data scarcity, even from small, annotated datasets. The adoption of these paradigms is limited to advanced architectures with the help of new methods such as transfer learning and federated learning to provide scalability and adaptability to such real-world clinical settings.

References

1. Krizhevsky, A., Sutskever, I., & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84–90. https://doi.org/10.1145/3065386, (2012).
2. Simonyan, K., & Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. https://doi.org/10.48550/arXiv.1409.1556, (2014).
3. He, K., Zhang, X., Ren, S., & Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90, (2016).
4. Ronneberger, O., Fischer, P., & Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 234–241). Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28, (2015).
5. Placeholder for specific studies on U-Net applications in tumor segmentation.
6. Vaswani, A., Shazer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998–6008). https://doi.org/10.48550/arXiv.1706.03762, (2017).
7. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*. https://doi.org/10.48550/arXiv.2010.11929, (2020).
8. Placeholder for ViTs in multi-modal MRI datasets.
9. Placeholder for hybrid CNN-RNN applications.
10. Oktay, O., Schlemper, J., Le Folgoc, L., Lee, M., Heinrich, M., Misawa, K., ... & Rueckert, D. Attention U-Net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*. https://doi.org/10.48550/arXiv.1804.03999, (2018).
11. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. Generative adversarial nets. In *Advances in Neural Information Processing Systems* (pp. 2672–2680). https://doi.org/10.48550/arXiv.1406.2661, (2014).
12. Kingma, D. P., & Welling, M. Auto-encoding variational Bayes. *arXiv preprint arXiv:1312.6114*. https://doi.org/10.48550/arXiv.1312.6114, (2014).
13. Pandey, A., Shaikh, M. S., & Patel, P. Review of acoustic features and computational Models in lung Disease diagnosis. IEEE, 1145–1150. https://doi.org/10.1109/icses63445.2024.10763243, (2024).
14. S. Singh, M. S. Shaikh, A. Sheikh, S. Dhargave and S. Mungale (2024), "Advanced Security and Protection for Logistic Industries and Warehouse Using Autonomous Carebot," 2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET), Indore, India, pp. 1-7, doi: 10.1109/ACROSET62108.2024.10743701, (2024).
15. M. S. Shaikh, K. Bhushanwar, N. Khodifad and S. I. Ali (2024), "Dual Purpose IoT Enabled Smart Cleaner Hexabot with Edge Detection Mechanism," 2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET), Indore, India, pp. 1-6, doi: 10.1109/ACROSET62108.2024.10743997, (2024).
16. Md. Shahnawaz Shaikh, Kamlesh Gupta (2014), A Review of Spectrum Sensing Techniques for Cognitive Radio. International Journal of Computer Applications. 94, 8, 1-5. DOI=10.5120/16360-5781,(May 2014).

17. Md. Shahnawaz Shaikh, Kamlesh Gupta (2014), Analysis of Cognitive Radio Spectrum Sensing Techniques. International Journal of Computer Applications. 102, 12, 1-7. DOI=10.5120/17864-8805, (September 2014).

18. Md. Shahnawaz Shaikh, Ankita Choudhry, Rakhi Wadhwani .Analysis of Digital Image Filters in Frequency Domain. International Journal of Computer Applications. 140, 6, 12-19. DOI=10.5120/ijca2016909330, (2016).

19. Mohammad Shahnawaz Shaikh (2019), "Analysis of Cognitive Radio Spectrum Sensing Techniques over Nakagami-m Fading Channel", International Journal of Modern Electronics and Communication Engineering (IJMECE), Volume -7, Issue – 4, ISSN: 2321-2152, (July 2019).

20. Mohammad Shahnawaz Shaikh (2019), "Cognitive Radio Spectrum Sensing with OFDM: An Investigation", International Journal on Emerging Trends in Technology (IJETT), Volume – 6, Issue – 2, ISSN: 2455-0124 (Online), (August 2019).

21. Shaikh, M. S. Green cognitive radio communication using cyclostationary feature detection spectrum sensing. Infokara Research Journal, 9(2), [page numbers if available]. https://doi.org/[DOI if available, (2020).

22. Mohammad Shahnawaz Shaikh, & P. K. Khare. FPGA Based Hardware Implementation for Green Cognitive Radio. Journal of Electronic Design Engineering, 6(3), 8–15. Retrieved from https://matjournals.co.in/index.php/JOEDE/article/view/3865, (2020).

23. Shaikh, M. S. Improving the requirements-based bandwidth allocation in 5G point-to-point networks. ICTACT Journal on Communication Technology, 13(4), (2022).

24. Shaikh, M. S. Enhancement of bandwidth and beamforming antenna arrays in 5G cellular communication networks. ICTACT Journal on Communication Technology, 13(4), (2022).

25. Shaikh, M. S., Ali, S. I., Deshmukh, A. R., Chandankhede, P. H., Titarmare, A. S., & Nagrale, N. K. AI Business Boost Approach for Small Business and Shopkeepers: Advanced Approach for Business. In S. Ponnusamy, M. Assaf, J. Antari, S. Singh, & S. Kalyanaraman (Eds.), Digital Twin Technology and AI Implementations in Future-Focused Businesses (pp. 27-48). IGI Global. https://doi.org/10.4018/979-8-3693-1818-8.ch003, (2024).

26. Ali, S. I., Kale, G. P., Shaikh, M. S., Ponnusamy, S., & Chouhan, P. S. AI Applications and Digital Twin Technology Have the Ability to Completely Transform the Future. In S. Ponnusamy, M. Assaf, J. Antari, S. Singh, & S. Kalyanaraman (Eds.), Harnessing AI and Digital Twin Technologies in Businesses (pp. 26-39). IGI Global. https://doi.org/10.4018/979-8-3693-3234-4.ch003, (2024).

27. Shaikh, M. S., Chandrawat, U. B., Choudhary, S. M., Ali, S. I., Ponnusamy, S., Khan, R. A., & Sheikh, A. G. Harnessing Logistic Industries and Warehouses with Autonomous Carebot for Security and Protection: A Smart Protection Approach. In S. Ponnusamy, M. Assaf, J. Antari, S. Singh, & S. Kalyanaraman (Eds.), Harnessing AI and Digital Twin Technologies in Businesses (pp. 239-257). IGI Global. https://doi.org/10.4018/979-8-3693-3234-4.ch017, (2024).

28. Ali, S. I., Shaikh, M. S., Ponnusamy, S., & Chouhan, P. S. Technological Collaboration, Challenges, and Unrestricted Research in the Digital Twin: Digital Twin Technology. In S. Ponnusamy, M. Assaf, J. Antari, S. Singh, & S. Kalyanaraman (Eds.), Harnessing AI and Digital Twin Technologies in Businesses (pp. 380-399). IGI Global. https://doi.org/10.4018/979-8-3693-3234-4.ch028, (2024).

29. Shaikh, M. S., Ponnusamy, S., Ali, S. I., Wanjari, M., Mungale, S. G., Ali, A., & Baig, I. AI-Based Advanced Surveillance Approach for Women's Safety. In S. Ponnusamy, V. Bora, P. Daigavane, & S. Wazalwar (Eds.), Wearable Devices, Surveillance Systems, and AI for Women's Wellbeing (pp. 13-25). IGI Global. https://doi.org/10.4018/979-8-3693-3406-5.ch002, (2024).

30. Mungale, S. G., Mungale, N. G., Shaikh, M. S., Mungale, S. G., Wazalwar, S. S., Wanjari, M. M., & Jichkar, R. A. Safeguard Wrist: Empowering Women's Safety. In S. Ponnusamy, V. Bora, P. Daigavane, & S. Wazalwar (Eds.), Wearable Devices, Surveillance Systems, and AI for Women's Wellbeing (pp. 192-205). IGI Global. https://doi.org/10.4018/979-8-3693-3406-5.ch012, (2024).

31. Ali, S. I., Shaikh, M. S., Shahane, S., Sharma, K., & Macwan, K. .The Era of Metaverse and Generative Artificial Intelligence. In L. Gaur (Ed.), Responsible Implementations of Generative AI for Multidisciplinary Use (pp. 29-44). IGI Global. https://doi.org/10.4018/979-8-3693-9173-0.ch002, (2025).

32. Ali, S. I. & Shaikh, M. S. The Ethical Dilemma of Using (Generative) AI to Science and Research. In L. Gaur (Ed.), Responsible Implementations of Generative AI for Multidisciplinary Use (pp. 249-264). IGI Global. https://doi.org/10.4018/979-8-3693-9173-0.ch009, (2025).

33. Sheikh, M. S. Harnessing logistic industries using autonomous carebot for smart surveillance, protection, and security. In F. Al-Turjman (Ed.), The Smart IoT Blueprint: Engineering a Connected Future. AIoTSS 2024. Advances in Science, Technology & Innovation. Springer, Cham. https://doi.org/10.1007/978-3-031-63103-0_20,(2024).

34. Preeti Chopkar, Minakshi Wanjari, Pranjali Jumle, Pankaj Chandankhede, Sheetal Mungale and Mohammad Shahnawaz Shaikh (2024), A Comprehensive Review on Cotton Leaf Disease Detection using Machine Learning Method, Grenze International Journal of Engineering and Technology, June Issue, Grenze ID: 01.GIJET.10.2.537, Grenze Scientific Society, Volume 10, Issue 2, Pages 239-245, (2024).

35. Himanshu Kitey, Dr. Pankaj Chandankhde, Dr. Kapil Jajulwar, Dr. Mohammad Shahnawaz Shaikh and Dr. Pragati M. Fatinge (2024), Solar Power Generation Technique and its Challenges - A Comprehensive Review, Grenze International Journal of Engineering and Technology, Jan Issue, Grenze ID: 01.GIJET.10.1.122, Grenze Scientific Society, Issue 1, Volume 10, Pages 2778-2786, (2024).

36. Mohammad Shahnawaz Shaikh (2016) "Li Fi - An Emerging Wireless Communication Technology", International Journal of Advanced Electronics & Communication Systems, Volume – 5, Issue – 1, Paper id 10758, ISSN: 2277-7318, (Mar 2016).