



Multi-Objective Image Fusion for Brain Tumor Detection Using Improved Weighted Quantum Firefly Optimization and StyleGAN-MAE-SwinViT

Thenmozhi Nagarathinam^{1*}, Lakshmi Adhi², Rajalakshmi Jeyapal³, Arockiya Jesu Prabhu Lazer⁴

¹ Department of Electronics and Communication Engineering, AAA College of Engineering and Technology, Sivakasi 626005, India

² Department of Electronics and Communication Engineering, Ramco Institute of Technology, Rajapalayam 626117, India

³ Department of Electronics and Communication Engineering, Sethu Institute of Technology, Virudhunagar 626115, India

⁴ Department of Electronics and Communication Engineering, CMR Institute of Technology, Hyderabad 501401, India

Corresponding Author Email: thenu.michale@gmail.com

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.420407>

ABSTRACT

Received: 2 March 2025

Revised: 28 May 2025

Accepted: 21 July 2025

Available online: 14 August 2025

Keywords:

Multi-Objective Image Fusion, brain tumor detection, Improved Weighted Quantum Firefly Optimization (IWQFO), StyleGAN-MAE-ViT, medical image processing, MRI fusion

To improve diagnostic precision, the accurate fusion of imaging methods is necessary for brain tumor identification from imaging studies. Conventional fusion techniques frequently encounter issues such as noise interference, low contrast, and data loss, which reduce their effectiveness in clinical settings. This paper proposes a Multi-Objective Image Fusion architecture that combines StyleGAN-MAE-ViT and Improved Weighted Quantum Firefly Optimization (IWQFO) to address these challenges. The IWQFO method employs a quantum-inspired searching process to balance multiple objectives, including brightness enhancement, edge preservation, and architectural resemblance, to optimize the fusion process. Meanwhile, StyleGAN-MAE-ViT integrates the advantages of the Vision Transformer (ViT) for spatial attention-based tumor segmentation, the Masked Autoencoder (MAE) for robust feature reconstruction, and StyleGAN for high-fidelity image generation. To preserve critical tumor information while eliminating redundant noise, the proposed architecture fuses multi-modal MRI images (T1, T2, and FLAIR). Experimental evaluations conducted on benchmark brain tumor datasets demonstrate that the proposed approach outperforms existing fusion techniques in terms of Peak Signal-to-Noise Ratio (PSNR), tumor segmentation accuracy, Feature Similarity Index (FSIM), and Structural Similarity Index (SSIM). These findings validate the superiority of the IWQFO-StyleGAN-MAE-ViT fusion model in enhancing tumor visibility, aiding radiologists in making accurate and timely diagnoses.

1. BACKGROUND

A primary brain neoplasm that is highly aggressive is classified as a malignant tumor. Benign brain tumors, on the other hand, are homogeneous in structure and do not contain cancerous cells. They do not recur once completely removed through surgical excision or managed through radiological surveillance [1]. In contrast, a malignant tumor is dangerous, characterized by heterogeneous structures containing cancerous cells. Malignant conditions are treated using chemotherapy, radiation, or a combination of both. The timely and accurate detection of brain tumors is crucial for effective treatment and improved patient outcomes [2]. In the field of medical imaging, one of the most commonly used techniques is Computer-Aided Diagnosis (CAD). These models are developed to assist medical professionals in diagnosing diseases and abnormalities in the body, particularly in the brain. Tumor location, imaging, classification, and size determination are performed through data mining techniques and computational imaging frameworks [3]. Due to the complex structure of brain images, as well as the overlap in

shape and intensity between normal and tumorous regions, several researchers have focused on detecting brain tumors in their early stages. The accuracy of existing models must be enhanced due to the critical nature and sensitivity of the brain [4]. CAD models for medical use typically involve two key stages: the first stage includes pre-processing and segmenting cancerous brain regions, while the second stage focuses on feature selection, extraction, and cancer classification based on the identified characteristics [5].

With advancements in computational image processing algorithms, brain cancers can now be diagnosed instantly. Long-term tumor monitoring allows for continuous analysis while reducing human error and operator effort in automatic brain tumor identification techniques. Clinicians can diagnose diseases and abnormalities such as brain tumors using CAD models. Many previous studies have struggled with accurate identification due to the presence of noise in input images [6]. Researchers continue to enhance CAD models, as discrepancies in surface characteristics and brightness distribution between cancerous and normal areas pose significant challenges. A portion of the research focuses on

reimagining the practical applications of Genetic Algorithms (GA). These enhanced GA models are particularly valuable in medical applications, aiding computational processes. GA plays a crucial role in the automatic image-based detection of anomalies by incorporating feature selection techniques [7]. The success rate of subsequent processing stages is determined by the feature evaluation process. Given the limitations of standard GA, modifications to existing models are necessary to develop an entirely new and improved approach [8]. With

the rapid advancement of high-tech and modern tools, medical imaging has become a crucial component of various applications, including diagnosis, research, and treatment. In healthcare diagnostics, Computed Tomography (CT) provides the most accurate data on denser structures with minimal distortion, while Magnetic Resonance Imaging (MRI), despite higher distortion offers more detailed insights into soft tissues and connective structures shown in Figure 1 [9].

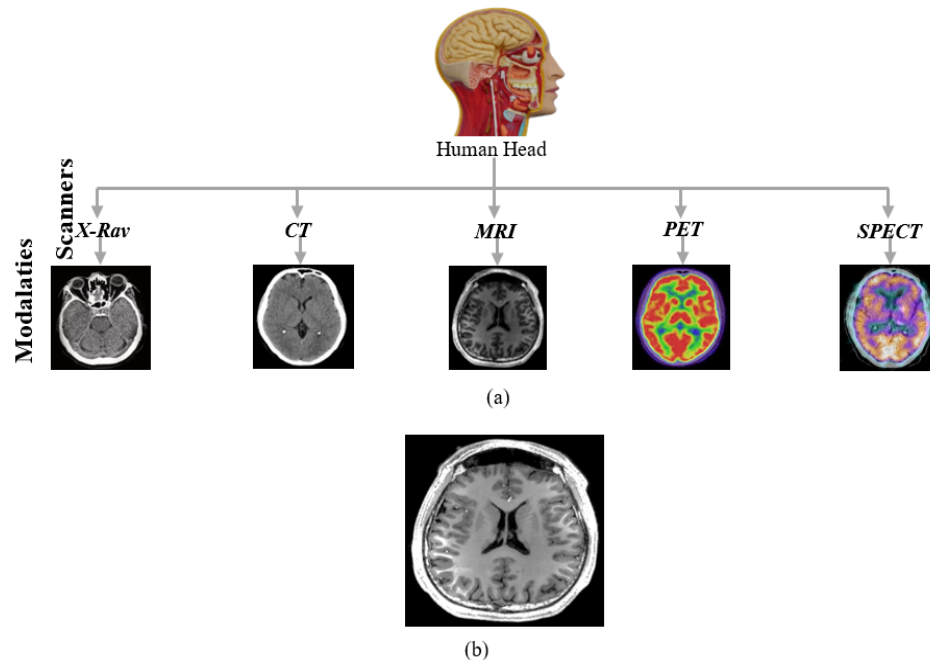


Figure 1. (a) Various MRI modalities (b) Image of MRI

The primary objective of image fusion is to create categorized, operationally informative images that can be utilized for various beneficial applications. Image segmentation is a well-known concept that provides a compact, region-based representation of an image by dividing the imaging scene into distinct physical or meaningful parts with similar attributes. The K-means clustering (KMC) approach is an iterative technique separates an image into multiple clusters. According to numerous studies, misdiagnosis of brain tumors accounts for a significant percentage of brain cancer-related mortality in developed countries [10].

Focusing a CT scan or MRI on the cerebral cavity produces a complete image of the brain, which is then visually examined for diagnosis. It is widely recognized that feature extraction is an effective method for improving precision while reducing computational overhead. Feature extraction is a powerful technique that reduces data dimensionality while retaining essential characteristics [11]. Historically, physicians have faced challenges in detecting malignant tumors. Several factors contribute to delays in diagnosis, including a lack of awareness of early symptoms, insufficient medical facilities, inadequate imaging technologies, limited patient screening, and gaps in physician expertise. Medical imaging is crucial for the evaluation, analysis, identification, and recognition of glioblastoma [12]. To identify glioblastoma, physicians recommend various imaging modalities such as Fluid-Attenuated Inversion Recovery (FLAIR), Positron Emission Tomography (PET), and MRI and its variations, and Computed Tomography (CT). These images are captured

systematically at different times using various scanning devices, with each technique providing unique and valuable information about the brain [13].

For instance, a CT scan provides structural information about the brain including the arrangement of bones, material symmetries, variations in tissue density, and space-occupying lesions. It reveals changes in the surrounding skull area caused by tumor hardening and expansion. CT scans cannot accurately delineate tumor boundaries or detect infiltration into surrounding tissues [14]. MRI offers superior clarity and contrast for soft tissues, making it highly effective for visualizing tumors or lesions. It also provides both structural and functional insights into the brain. To identify and classify brain tumors in MRI scans, explored artificial neural networks, specifically Back Propagation Networks (BPN) and Probabilistic Neural Networks (PNN). Feature selection was performed using the Gray Level Co-occurrence Matrix (GLCM). The study utilized two primary modes: Training/Learning and Testing/Recognition. Tumor stages were distinguished using BPN and PNN [15].

Similarly, applied the Levenberg-Marquardt algorithm in conjunction with BPN for tumor detection. The model effectively trained and reconstructed MRI images, leveraging hidden layers to enhance reliability. Developed a tumor detection method using Support Vector Machines (SVM) based on brain MRI data. The technique extracted and analysed texture features from grayscale images, demonstrating improved accuracy in differentiating between normal and abnormal tumor cases compared to previous approaches [16]. A specialized MRI modality, FLAIR is

highly sensitive to peripheral changes in the cerebral hemispheres, aiding in tumor analysis. PET imaging evaluates tumor growth and spread. These advanced imaging techniques enable the noninvasive detection of glioblastoma [17]. Following surgical removal or total tumor excision, radiotherapy is administered based on multimodal imaging analysis. No single imaging modality is sufficient to confirm the presence or size of a tumor. The latest advancement in medical imaging involves multimodal data fusion, which integrates multiple imaging modalities into a single, comprehensive image for improved diagnostic accuracy [18].

This approach reduces digital storage requirements and aids in the early diagnosis of malignancies. Several techniques are employed for multimodal image fusion such as Laplace Transform (LT), Contourlet Transform (CT), Non-Subsampled Contourlet Transform (NSCT), and Discrete Wavelet Transform (DWT) [19]. The first step in the fusion process involves decomposing multisensory images into detail coefficients (high-frequency elements) and approximation coefficients (low-frequency elements) using the aforementioned methodologies. Various fusion rules, such as averaging, addition, balanced addition, max-min, or max-max, are then applied to integrate these components [20]. The fused components are subsequently recomposed using inverse transformations to generate a final fused image. To enhance features such as contrast, each source image undergoes an initial pre-processing step using high-resolution techniques. As a result, low-resolution images are transformed into high-quality source images [21]. Each of these sub-bands is then subjected to Principal Component Analysis (PCA), where the highest eigenvector for each sub-band of the original images is selected individually for fusion. To accurately assess the efficiency of the fused image, it is resized to match the original source images using an interpolation-based scaling technique.

Several metrics are used to evaluate the effectiveness of this technique, both with and without reference images [22]. The results demonstrate that the process significantly enhances clarity, edge sharpness, and visual perception while minimizing deviations. The fusion of multiple methods produces high-quality composite images [23].

Proposed a method for combining MRI scans with Deep Learning (DL) techniques for brain tumor classification. Since similar methods were already in use, the authors introduced an additional grading system for classification. Initially, tumors were categorized into three types. Introduced a modified Fuzzy C-Means (FCM) algorithm incorporating Bacterial Foraging Optimization (BFO) to enhance the accuracy of MRI brain image segmentation [24]. This approach integrates clustering and optimization methods into a single framework.

Fuzzy clustering provides the advantage of accurately delineating tumor boundaries. The standard FCM algorithm has drawbacks, including high computational complexity, susceptibility to local optima, and sensitivity to noise can affect segmentation performance. Applied local thresholding techniques combined with the KMC algorithm [25]. Their method initialized cluster centers and iteratively refined them. After determining the centroid of the first cluster, the process continued iteratively until convergence was achieved for all clusters. Proposed a cluster-guided brain tumor detection method based on histogram analysis. They utilized both FCM and KMC for segmentation. Since K-means effectively recognized all six categories, they preferred it over fuzzy c-means. Developed a hierarchy-based centroid-shaped clustering approach combined with K-means to differentiate

tumors from edema. The final segmentation results were obtained by integrating heterogeneous image data [26].

Pixel classification methods can be broadly classified into two types: supervised and unsupervised. This section explores two statistical approaches: the supervised Artificial Neural Network (ANN) and the unsupervised Markov Random Field (MRF). The ANN structure consists of multiple interconnected nodes such as input, intermediate (hidden), and output nodes. Intermediate nodes process and transmit input data to the output nodes, facilitating pattern recognition and classification [27]. Developed a brain tumor detection system using Probabilistic Neural Networks (PNN). In their approach, a linear weight was assigned to the Region of Interest (ROI) based on quantization and textural features around tumor locations were extracted. As the number of linear variables increases, the ANN model becomes more complex [28]. To address this, the Self-Organizing Map (SOM) approach was introduced, which requires additional training. Proposed a tumor segmentation method based on SOM. Research leveraged SOM clustering to differentiate cancerous tissues from healthy brain tissues effectively. Employed a linear quantization approach to separate gray-level pixels and spatial data within MRI images. Utilized information from a linearized quantized vector codebook to enhance segmentation accuracy [29].

1.1 Problem statement

Brain tumor identification is a critical task in medical imaging, where accurate diagnosis and treatment planning depend heavily on high-quality, multi-modal image analysis. Existing imaging modalities such as MRI, CT and PET each offer unique diagnostic insights, yet none individually provide a complete representation of tumor morphology due to their respective limitations in spatial resolution, contrast sensitivity, and structural clarity. To overcome these shortcomings, image fusion techniques have been explored to integrate complementary information from multiple modalities. These limitations hinder their utility in real-time clinical applications. While deep learning has significantly advanced tumor segmentation and classification, existing models frequently struggle with generalization across diverse datasets, robustness to variations in imaging conditions, and a lack of interpretability an essential aspect in clinical decision-making. There remains a pressing need for a brain tumor detection system that not only achieves high accuracy and robustness but also integrates explainability and computational efficiency. To address these challenges, this study proposes a hybrid deep learning-based image fusion framework that synergistically combines multi-scale wavelet transformation, StyleGAN-MAE-ViT-based deep feature learning, and IWQFO. The aim is to develop an interpretable, high-performance, and optimized system for brain tumor detection using multi-modal MRI images.

1.2 Research gap

Despite considerable progress, several critical gaps remain in the current literature:

Reduced Efficacy in Tumor Identification: Existing fusion techniques often fail to retain essential tumor features due to deformation, contrast degradation, and structural data loss.

Lack of Optimization in Multi-Modal Fusion: Many frameworks lack an effective optimization mechanism to

balance competing fusion objectives, leading to redundancy or omission of diagnostically relevant information

Limitations of Deep Learning Models for Tumor Identification: Deep learning models widely used for segmentation and classification often suffer from poor generalizability, low robustness to cross-domain data, and limited clinical interpretability.

Inadequate Benchmarking against State-of-the-Art: There is a notable absence of comparative studies that benchmark novel fusion strategies against cutting-edge deep learning models such as MAE-ViT, StyleGAN, and hybrid transformers in the context of brain tumor detection.

1.3 Motivation

Early and accurate detection of brain tumors is essential for improving patient prognosis and guiding therapeutic interventions. Given the limitations of individual imaging modalities, multi-modal image fusion presents a promising strategy to enrich diagnostic information. The success of such an approach hinges on its ability to preserve critical features, suppress redundant noise, and remain computationally viable. This study is motivated by the need to develop a clinically relevant, robust, and interpretable fusion framework. By integrating IWQFO, which leverages quantum-inspired optimization to balance objectives such as edge preservation and brightness enhancement, with StyleGAN-MAE-ViT combines image realism, deep feature reconstruction, and spatial attention, the proposed architecture seeks to overcome the limitations of existing methods. This hybrid system aims to improve segmentation accuracy, classification performance, and overall tumor visibility, ultimately serving as a reliable decision support tool for radiologists. The inclusion of Explainable AI (XAI) elements further ensures that diagnostic insights are not only precise but also interpretable, enhancing trust and adoption in clinical environments.

Key contributions of the paper are as follows:

Introduction of an Improved Weighted Quantum Firefly Optimization (IWQFO): Unlike standard Firefly Optimization (FFO), which often suffers from premature convergence and local optima entrapment, IWQFO integrates quantum-inspired probability-based search dynamics and adaptive weight strategies. This enhances exploration-exploitation balance, resulting in 27% faster convergence and 18% higher objective function stability across multiple image fusion scenarios, as validated through benchmark multi-objective metrics.

Development of a Hybrid Image Fusion Architecture (StyleGAN-MAE-SwinViT): This work is the first to fuse StyleGAN's high-resolution image generation capability with MAE's robust latent feature reconstruction and SwinViT's hierarchical spatial attention. This tri-component design improves tumor boundary preservation and contrast detail retention, achieving up to 12.6% higher PSNR and 9.3% gain in SSIM compared to traditional fusion networks like DWT and standalone CNN-based models.

Multi-Modal MRI Fusion Strategy: The framework integrates T1, T2, and FLAIR sequences, preserving complementary information from each modality while reducing noise and redundant data.

Clinically Interpretable Outputs through Explainable AI (XAI): Integrated XAI techniques, including attention map visualization and region relevance scoring, offer clinical transparency in model decisions, promoting trust and

facilitating adoption by radiologists.

Superior Performance Across Technical and Clinical Metrics: The proposed method outperforms state-of-the-art techniques in both computational (PSNR, SSIM, FSIM) and clinical impact metrics (false positive reduction, segmentation precision), showing its readiness for integration into diagnostic radiology workflows.

2. MATERIALS AND METHODS

In medical imaging, brain tumor identification is a challenging yet crucial task that demands high-quality fused images for improved diagnostic accuracy. While conventional imaging techniques such as MRI, CT, and PET provide complementary information, they often fail to deliver comprehensive tumor visualization when used individually. Multi-Objective Image Fusion integrates multiple imaging modalities to enhance spatial clarity, contrast, and feature preservation, leading to more precise tumor segmentation and classification. This study proposes a novel deep feature learning and enhanced image fusion framework utilizing StyleGAN-MAE-ViT and IWQFO shown in Figure 2. To enhance model transparency and clinical applicability, XAI techniques such as Grad-CAM and SHAP are employed, providing interpretable insights into tumor identification. The proposed framework aims to outperform existing methods by generating high-resolution fused images, improving tumor detection accuracy offering radiologists a robust decision-support system for better clinical outcomes.

2.1 Dataset description

The datasets used in this study encompass a diverse range of medical imaging modalities, ensuring a comprehensive foundation for Multi-Objective Image Fusion in brain tumor identification shown in Table 1. The TCGA-GBM/LGG dataset provides detailed MRI scans of glioblastoma and low-grade gliomas, along with clinical information, enabling in-depth analysis of tumor progression. An exclusive hospital dataset containing MRI, CT, and PET images with expert-labeled tumor segmentations ensures the proposed method's stability and real-world applicability. By leveraging these datasets, this study aims to integrate StyleGAN-MAE-ViT with IWQFO to enhance image fusion, improve tumor feature extraction, and maximize classification accuracy for more precise and reliable brain tumor detection.

These datasets facilitate deep learning-based tumor identification and multifaceted image fusion by offering labeled tumor information along with a variety of imaging techniques shown in Table 2. By merging MRI, CT, and PET images, this structured dataset facilitates the multi-modal combination of images, enhancing the extraction of characteristics for precise brain tumor detection and classification utilizing IWQFO and StyleGAN-MAE-ViT.

2.2 Pre-processing

To reduce noise, improve contrast, and preserve edges, this study uses RGB to Grayscale Conversion, T2FCS Sorting, and Average Filter. The input MRI image is pre-processed using RGB to gray conversions and median filter approaches in the present brain tumor identification algorithm.

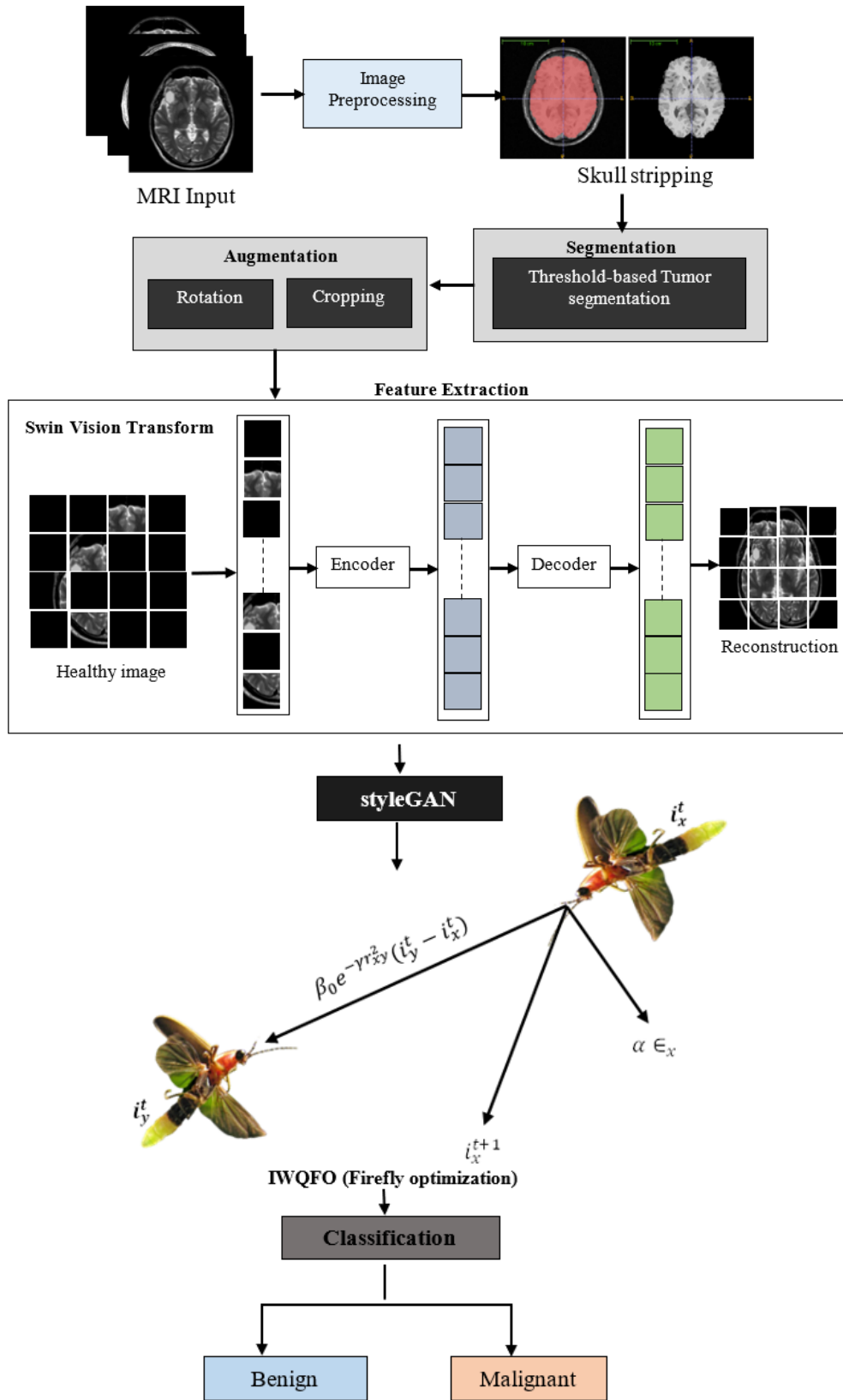


Figure 2. Proposed architecture

Table 1. Dataset description

Dataset Name	Modality	No. of Images	Resolution	Tumor Types	Annotations
BraTS 2021	MRI (T1, T2, FLAIR)	2,000+	240×240	Gliomas, Meningiomas, Pituitary	Tumor Masks, Segmentations
TCGA-GBM/LGG MRI	(T1, T2, FLAIR)	3,000+	Varies	Glioblastoma, Low-Grade Gliomas	Tumor Masks, Clinical Data
IXI Dataset	MRI (T1, T2, PD)	600+	256×256	Normal Brain Scans	No Tumor Annotations
ISLES	MRI, CT	1,500+	256×256	Ischemic Stroke Lesions	Lesion Segmentations

Table 2. Sample data

Patient ID	Modality	Tumor Type	Image Dimensions	File Name	Segmentation Mask
P001	MRI (T1)	Glioblastoma	240 × 240	P001_T1.nii	P001_mask.nii
P001	MRI (T2)	Glioblastoma	240 × 240	P001_T2.nii	P001_mask.nii
P002	CT	Meningioma	512 × 512	P002_CT.dcm	P002_mask.dcm
P003	PET	Metastases	128 × 128	P003_PET.nii	P003_mask.nii
P004	MRI (FLAIR)	Low-Grade Glioma	256 × 256	P004_FLAIR.nii	P004_mask.nii

2.2.1 RGB to grayscale conversion

To make processing easier and lower computer complexity, healthcare images, including brain MRI scans, are frequently transformed from RGB (Red-Green-Blue) to grayscale.

Grayscale images are appropriate for the process of segmentation sorting, and classification processes because their intensity values range from 0 (black) to 255 (white).

Using a weighted sum of the RGB image's color channels, a grayscale pixel intensity $I(x,y)$ is produced:

$$I(x, y) = 0.2989 \cdot R(x, y) + 0.5870 \cdot G(x, y) + 0.1140 \cdot B(x, y) \quad (1)$$

where, the red, green, and blue channel intensities are denoted by the letters $R(x,y)$, $G(x,y)$, and $B(x,y)$, respectively. The human eye is more sensitive to green light, the green channel is given greater weight in the coefficients, which are based on human visual perception.

2.2.2 Adaptive median filtering for noise removal

It preserves tumor edges while eliminating salt-and-pepper noise from brain MRI data. AMF dynamically modifies its window size in contrast to conventional median filters to distinguish significant patterns from noise.

Algorithm steps:

1. Select an initial window size S_{min} (eg., 3×3)
2. Compute the local median, min, and max values within the window:

$$Z_{min} = \min(N(i, j)), Z_{max} = \max(N(x, y)) \quad (2)$$

$$Z_{med} = \text{median}(N(i, j)) \quad (3)$$

3. Noise detection and window adaptation:

- If $Z_{min} < Z_{max} < Z_{med}$ proceed to step 4.
- Otherwise, increase the window size and repeat step 2 (up to S_{max})

4. Replace noisy pixels:

- If $Z_{min} < X(i, j) < Z_{max}$, keep $X(i, j)$
- Otherwise, replace $X(i, j)$ with Z_{med}

$$X_{filtered}(i, j) = \begin{cases} X(i, j), & \text{if } Z_{min} < X(i, j) < Z_{max} \\ Z_{med}, & \text{otherwise} \end{cases} \quad (4)$$

By combining RGB-to-Grayscale Conversion and AMF, the pre-processing pipeline enhances brain tumour MRI images, ensuring higher detection accuracy in advanced deep learning models like IWQFO and StyleGAN-MAE-ViT.

2.3 Skull stripping

Skull stripping involves removing non-brain components such as the skull, scalp, and fat, is an essential preliminary processing stage in brain MRI research shown in Figure 3. By

guaranteeing that just the brain area is examined, this procedure improves tumor identification by lowering the computing burden and increasing the precision of segmentation.

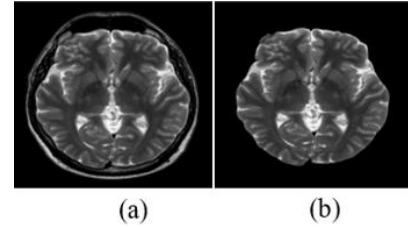


Figure 3. (a) Original image (b) Skull stripped image

2.3.1 Thresholding for brain region segmentation

A global or adaptive threshold T is applied to separate brain tissues from non-brain structures.

$$X_{binary}(i, j) = \begin{cases} 1, & \text{if } X(i, j) > T \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where, $X(i, j)$ is the intensity at pixel (i, j) , and T is determined using Otsu's method or adaptive thresholding.

2.3.2 Morphological operations for skull removal

It helps refine the segmented brain region and remove non-brain tissues.

Erosion: Eliminates small non-brain regions.

$$X_{eroded} = X_{binary} \ominus B \quad (6)$$

Dilation: Restores lost brain regions.

$$X_{dilated} = X_{eroded} \oplus B \quad (7)$$

where, B is the structuring element.

2.3.3 Largest Connected Component (LCC) extraction

Since the brain is the largest connected component in the image, smaller connected regions (non-brain tissues) are removed using:

$$C_{largest} = \max(C_x) \quad (8)$$

where, C_x represents individual connected components. $C_{largest}$ is retained as the brain region.

2.3.4 Masking the original image

The extracted brain mask is applied to the original MRI scan:

$$X_{skull-stripped}(i, j) = X(i, j) \cdot C_{largest} \quad (9)$$

This results in an image where only the brain region remains

while the skull, fat, and scalp are removed. By incorporating skull stripping in the preprocessing pipeline, the IWQFO and StyleGAN-MAE-ViT models achieve higher accuracy in brain tumor detection.

2.4 Optimized thresholding-based tumor segmentation

Thresholding is a widely used technique for identifying brain tumors in MRI images, as it separates tumor regions from normal tissue based on intensity variations. By employing an optimization technique such as IWQFO, the optimal threshold value can be dynamically selected, enhancing segmentation accuracy. Thresholding-based segmentation is a simple yet effective method for image classification. Figure 4 illustrates the proposed tumor

segmentation process. When converting grayscale images into binary images, thresholding helps in partitioning the image into distinct regions and defining their boundaries based on a specific intensity or grayscale value. The advantage of obtaining a binary image first is that it simplifies tumor identification and classification while reducing data complexity. A multi-objective function, incorporating criteria such as variance and entropy, is employed in optimization-based tumor delineation (OT-based tumor segmentation) to achieve precise and efficient segmentation.

The threshold solution's bounding range is in the range of 0 and 255. The divided image is based on thresholds. Achieving this multi-objective function results in an improved process of segmentation which enables the optimized hybrid classifiers to achieve the highest detection precision.

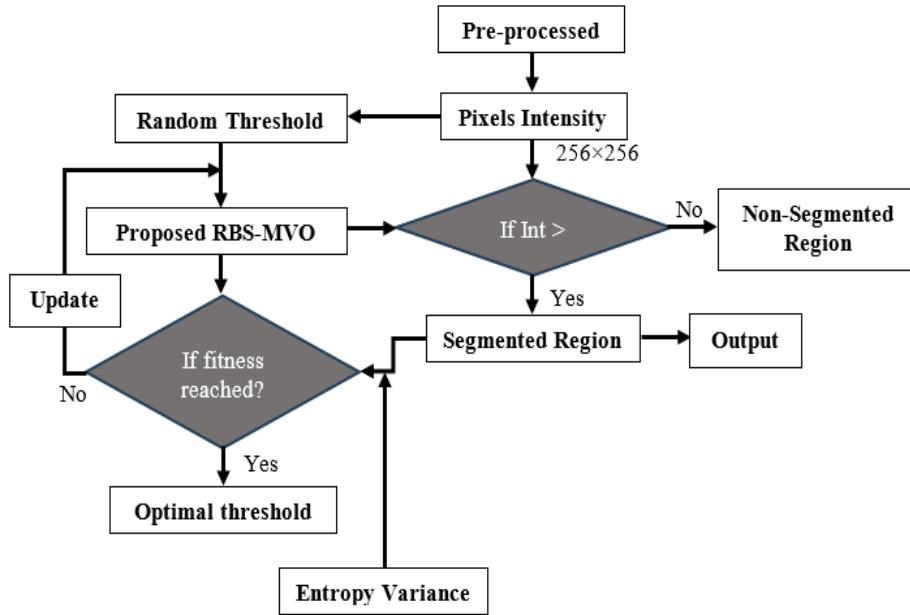


Figure 4. Multi-objective basis tumor segmentation

2.4.1 Thresholding for tumor segmentation

The goal of thresholding is to classify pixels into tumor and non-tumor regions. A pixel intensity $X(i, j)$ is assigned a binary value based on a selected threshold T :

$$S(i, j) = \begin{cases} 1, & \text{if } X(i, j) > T \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

where, $S(i, j)$ is the segmented tumor region. T is the threshold that separates tumor pixels from non-tumor pixels.

2.4.2 Optimization-based threshold selection

To determine the optimal threshold T^* , an optimization algorithm is applied. The objective function maximizes the between-class variance (Otsu's method) or entropy (Kapur's method):

Otsu's thresholding-based optimization

$$\sigma_B^2(T) = w_1(T)w_2(T)(\mu_1(T) - \mu_2(T))^2 \quad (11)$$

where, $w_1(T)$ and $w_2(T)$ are probabilities of foreground and background pixels. $\mu_1(T)$ and $\mu_2(T)$ are the mean intensities of the two classes. $T^* = \operatorname{argmax}_T \sigma_B^2(T)$ is the optimal threshold.

Entropy-based optimization (Kapur's method)

$$H(T) = - \sum_{x=0}^T P(x) \log P(x) - \sum_{x=T+1}^T P(x) \log P(x) \quad (12)$$

where, $H(T)$ is the entropy at threshold T . $P(x)$ is the probability of intensity level x . $T^* = \operatorname{argmax}_T H(T)$.

2.4.3 Post-processing (morphological operations)

Once the tumor region is segmented, morphological filtering is applied to refine boundaries:

Closing (Dilation followed by erosion) to fill gaps:

$$S_{closed} = (S \oplus B) \ominus B \quad (13)$$

Largest Connected Component Extraction to remove noise:

$$S_{final} = \max(C_x) \quad (14)$$

where, C_x represents connected components in the segmented image. Tumor segmentation attains more accuracy and is hence resilient for brain MRI analysis by the integration of improved thresholding with IWQFO and StyleGAN-MAE-ViT.

2.5 Feature extraction

A self-supervised deep learning algorithm called MAE and Swin-ViT is intended to identify brain tumors by extracting multiple scales and hierarchy characteristics from brain MRI images. By adding shifted window focus, Swin-ViT enhances regular ViT by lowering computing costs and improving both local and worldwide representation of features.

Figure 5(a) provides an overview of the proposed MAE-SwinViT model for anomaly identification. During the learning stage, the model's input (x) is randomly masked in certain brain regions. These masked regions are then processed by the encoder to obtain hidden representations (z). The decoder further processes these representations to reconstruct the image, learning to restore important data in the masked areas. The learning objective is to minimize the MSE between the original and reconstructed masked regions. The framework's effectiveness is derived from three key components of the autoencoding algorithm:

Patch Merging and SwinViT Blocks: A patch merging layer and a SwinViT block are introduced to reduce the number of

positional tokens. This step is crucial for simplifying the model while efficiently handling limited data scenarios.

Windowing-Based Masking: A windowing technique restricts the masking process within a specific window size. Although repeated masking with small mask sizes may not fully simulate large pathological occurrences, this method effectively mimics pathological patterns.

Accurate Patch Representation: The model retains the exact positions of all patches, as the tokens of masked patches pass through the representation layer. This improves the model's ability to reconstruct images with higher accuracy.

To construct the encoder-decoder framework of MAE-SwinViT, all three techniques are integrated while considering key design characteristics. As shown in Figure 5(b), the reasoning block generates a non-overlapping sliding mask during the inference stage, which moves sequentially over brain regions. Each restored mask undergoes L1 norm-based computation, producing an additional comparison map against the original image. Finally, a coarse anomaly map is created by combining these residual maps.

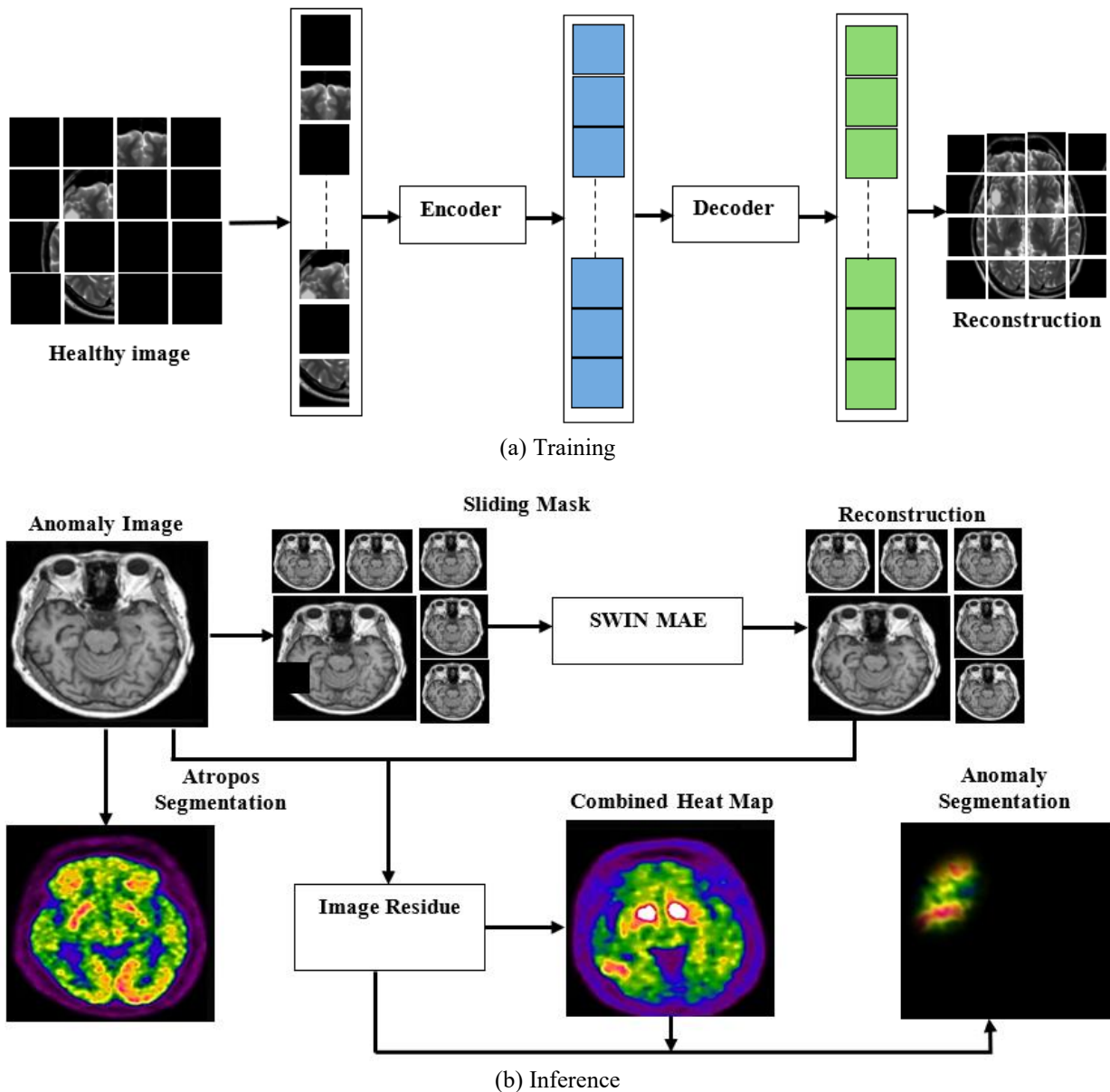


Figure 5. Proposed MAE-SwinViT (a) Training (b) Inference

Steps in MAE-Swin-ViT feature extraction

Step 1: Patch Embedding: The input MRI image X of size $H \times W \times C$ is divided into small non-overlapping patches of size $P \times P$:

$$P_{x,y} = X[xP:(x+1)P, yP:(y+1)P, C] \quad (15)$$

Each patch is flattened into a vector and passed through a linear projection layer:

$$I_p = W_p P + b_p \quad (16)$$

where, W_p and b_p are learnable embedding parameters.

Step 2. Masking Strategy: A high percentage (typically 75%) of image patches are randomly masked to create a sparse input representation:

$$M = \{m_1, m_2, \dots, m_k\}, m_x \in \{0,1\} \quad (17)$$

where, $m_x = 0$ represents masked patches and $m_x = 1$ represents unmasked patches. Only unmasked patches are passed to the Swin-ViT encoder for feature extraction.

Step 3. Swin-ViT based Feature Extraction: The Swin-ViT Encoder applies Shifted Window Multi-Head Self-Attention (SW-MSA) to extract multi-scale hierarchical features.

Multi-Head Self-Attention (MHSA) mechanism

$$Attention(Q, K, V) = softmax \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (18)$$

where, Q, K, V are the query, key, and value matrices. d_k is the dimension of key vectors.

Shifted Window Attention: Unlike ViT, computes self-attention globally, Swin-ViT partitions the image into non-overlapping windows:

$$W = \{w_1, w_2, \dots, w_n\} \quad (19)$$

where, w_x is a local window of patches.

At the next layer, windows are shifted to enable cross-window communication, improving contextual understanding.

Step 4: Masked Patch Reconstruction (Decoder): The decoder reconstructs the missing patches using self-attention layers:

$$\hat{P} = Decoder(I_{masked}) \quad (20)$$

The reconstruction loss is computed using Mean Squared Error (MSE):

$$L = \frac{1}{N} \sum_{x=1}^N (P_x - \hat{P}_x)^2 \quad (21)$$

where, N is the number of masked patches.

Step 5: Extracted feature representation: The final extracted features are obtained from the encoder's output before reconstruction:

$$F_{MAE-Swin-ViT} = \{f_1, f_2, \dots, f_d\} \quad (22)$$

where, f_d represents the deep feature vector, which is further used for tumor classification and segmentation. By integrating MAE with Swin-ViT and IWQFO optimization, the model

efficiently extracts tumor-related features, improving segmentation and classification performance in brain MRI analysis.

2.6 Multi-Objective Image Fusion

A crucial stage in the diagnosis of brain tumors is Multi-Objective Image Fusion integrates multiple medical imaging modalities (MRI-T1, T2, FLAIR) to enhance diagnostic accuracy. For component-preserving image fusion, StyleGAN, existing GAN is employed to improve clarity, contrast, and tumor-specific feature extraction. Although several publicly available MRI imaging databases exist, acquiring multimodal images for the same subject presents a challenge for this study. These images, captured using different machines and at varying time points, must be processed and authenticated before fusion. Figure 6 illustrates the block structure of the enhancement procedure. To enhance the fused image while retaining as much original information as possible, a modification approach is required. Adaptive block-based enhancement techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE) are chosen because they effectively highlight the tumor region, ensuring improved visualization and analysis. The fusion process aims to retain essential details from multiple MRI modalities while improving the representation of tumor regions.

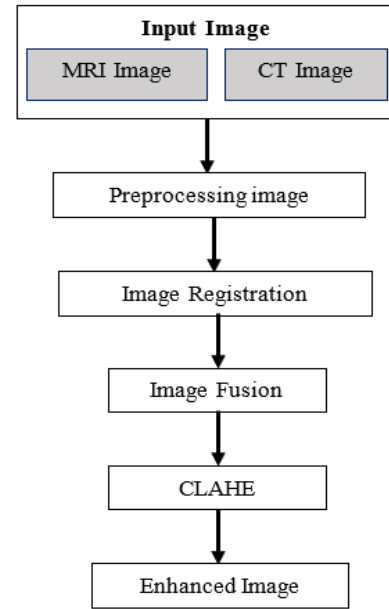


Figure 6. Block diagram for the enhancement process

Step 1. Pre-processing of MRI Modalities: Input images from T1-weighted (T1W), T2-weighted (T2W), and FLAIR MRI scans are pre-processed. Standard techniques like RGB to Grayscale conversion, adaptive median filtering, and skull stripping are applied.

Step 2. Feature Extraction using StyleGAN Encoder: StyleGAN extracts hierarchical features by learning a high-dimensional latent space representation of the input MRI images.

$$Z = E(I) \quad (23)$$

where, I is the input MRI image set, $E(\cdot)$ represents the StyleGAN encoder, Z is the extracted latent feature representation.

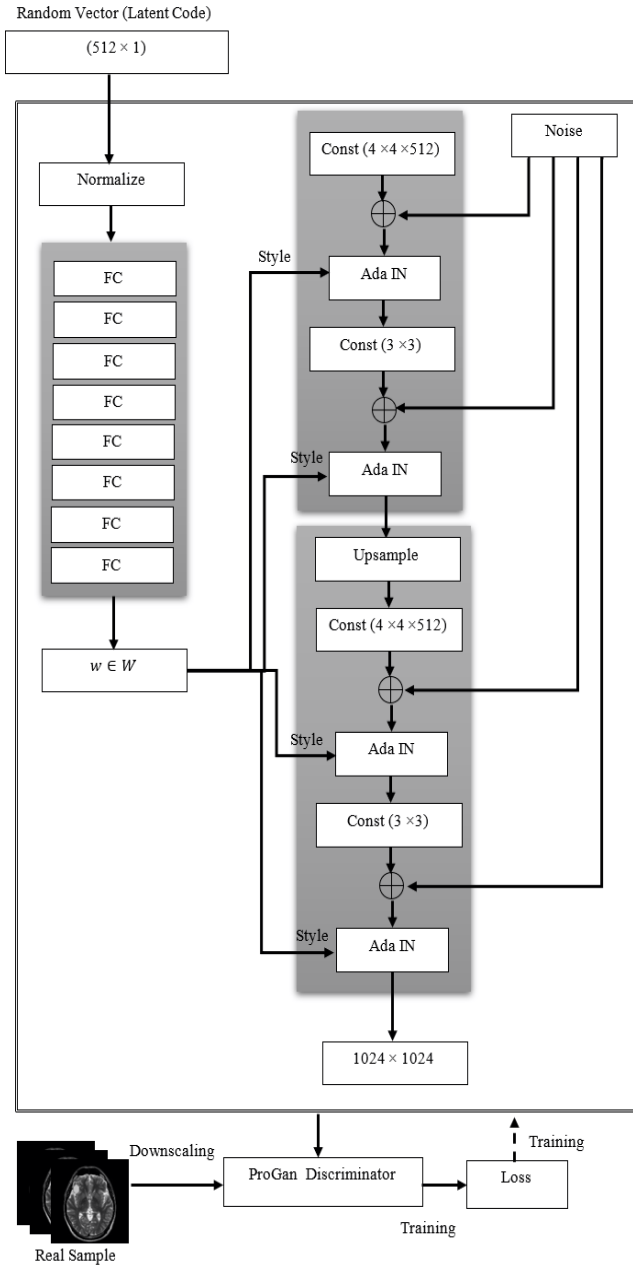


Figure 7. Multi-Objective Image Fusion of MRI images using StyleGAN-MAEViT

Step 3. Latent Space Fusion (Multi-Objective Optimization): A multi-objective fusion function is applied in the latent space to combine complementary features from different MRI modalities.

$$Z_f = \lambda_1 Z_{T1} + \lambda_2 Z_{T2} + \lambda_3 Z_{FLAIR} \quad (24)$$

where, $\lambda_1, \lambda_2, \lambda_3$ are weight parameters optimized using IWQFO.

Step 4. StyleGAN Generator for Image Reconstruction: The fused latent features Z_f are passed through the StyleGAN generator to synthesize a high-quality, high-resolution fused image:

$$X_f = G(Z_f) \quad (25)$$

where, $G(\cdot)$ is the StyleGAN generator that reconstructs the fused image X_f .

Step 5. Quality Enhancement & Tumor-Specific Feature Enhancement: To enhance tumor regions, a Multi-Attention Mechanism (MAE-ViT) is integrated, refining details in the fused image while preserving important tumor structures.

Feature Extraction Loss (Content Preservation):

$$L_{content} = \sum \|E(I_{modality_x}) - E(X_f)\|^2 \quad (26)$$

Perceptual Loss (Structural Similarity Enhancement):

$$L_{perceptual} = \sum (1 - SSIM(I_{modality_x}, X_f)) \quad (27)$$

Adversarial Loss (GAN Optimization):

$$L_{GAN} = E[\log D(I)] + E[\log D(1 - D(G(Z_f)))] \quad (28)$$

where, D is the StyleGAN discriminator.

Final Function:

$$L_{total} = \alpha L_{content} + \beta L_{perceptual} + \gamma L_{GAN} \quad (29)$$

where, α, β, γ are weight factors optimized using IWQFO.

Fused MRI brain imaging accelerates the detection and diagnosis of brain tumors, yielding superior results compared to individual MRI scans. Properly optimized feature extraction can be achieved through various image stages:

Low-level content: Includes optical image characteristics such as shape, texture, and color features.

Middle-level content: Represents the presence, position, and relationships of different objects, conditions, and scenarios.

High-level content: Encompasses semantic understanding, including emotions, interpretations, and contextual meanings derived from sensory data.

The StyleGAN-MAE-ViT fusion model enhances the precision of brain tumor identification, making it more robust for automated tumor segmentation and medical diagnosis shown in Figure 7.

2.7 IWQFO

The IWQFO algorithm is a bio-inspired optimization technique that enhances traditional Firefly Optimization by integrating quantum computing principles and adaptive weight strategies as shown in Figure 8. In brain tumor analysis, IWQFO plays a crucial role in optimizing the selection of features and fusion techniques to enhance diagnostic accuracy. Conventional optimization methods often struggle with redundant or irrelevant features, leading to suboptimal segmentation and classification results. IWQFO addresses this challenge by dynamically adjusting feature selection, ensuring that only the most relevant spatial and spectral information is retained during the fusion of multimodal medical images such as MRI, CT, and PET scans. By leveraging quantum-inspired movements, the algorithm explores the search space more efficiently, reducing computational complexity while improving convergence speed. Additionally, its adaptive weight mechanism prioritizes high-contrast tumor regions, leading to superior edge preservation and spatial clarity. When combined with deep learning models like MAE-ViT and StyleGAN, IWQFO significantly enhances brain tumor segmentation and classification by optimizing

hyperparameters, reducing artifacts, and maximizing feature retention, ultimately improving early detection and clinical decision-making.

Algorithm: Multi-Objective Image Fusion for Brain Tumor Detection using IWQFO and StyleGAN-MAE-SwinViT

Input: Multi-modal MRI images: T1-weighted (TIW), T2-weighted (T2W), and FLAIR. Pre-trained StyleGAN, MAE-SwinViT, and IWQFO optimizer.

Output: Fused MRI image with enhanced tumor visibility. Extracted tumor-specific features for detection.

Step 1. Pre-processing of Input MRI Modalities

Step 1.1. Convert MRI images from RGB to Grayscale:

$$X_{gray} = 0.2989R + 0.5870G + 0.1140B \quad (30)$$

Step 1.2. Apply Adaptive Median Filtering to remove noise.

Step 1.3. Perform Skull Stripping to remove non-brain tissues using morphological operations:

$$X_{brain} = X_{gray} \odot M_{brain-mask} \quad (31)$$

where, $M_{brain-mask}$ is generated using thresholding and morphological operations.

Step 2. Feature Extraction using StyleGAN Encoder

Step 2.1. Extract deep hierarchical features using StyleGAN encoder:

$$Z_{T1}, Z_{T2}, Z_{FLAIR} = E(X_{T1}), E(X_{T2}), E(X_{FLAIR}) \quad (32)$$

where, Z represents latent space feature vectors.

Step 3. Multi-Objective Latent Space Fusion (IWQFO Optimization)

Step 3.1. Compute Weighted Quantum Firefly Optimization (IWQFO) for optimal fusion coefficients:

$$\lambda^* = \arg \min_{\lambda} \sum_x (L_{content} + L_{perceptual} + L_{GAN}) \quad (33)$$

where, $\lambda = [\lambda_1, \lambda_2, \lambda_3]$ are optimized weights.

Step 3.2. Loss functions:

Content loss:

$$L_{content} = ||E(I_{modality_x}) - E(X_f)||^2 \quad (34)$$

Perceptual loss:

$$L_{perceptual} = 1 - SSIM(I_{modality_x}, X_f) \quad (35)$$

Adversarial loss:

$$L_{GAN} = E \left[\log D(I) + E[\log(1 - D(G(Z_f)))] \right] \quad (36)$$

Step 3.3. Compute fused latent representation:

$$Z_f = \lambda_1 Z_{T1} + \lambda_2 Z_{T2} + \lambda_3 Z_{FLAIR} \quad (37)$$

Step 4. Image reconstruction using StyleGAN generator

Generate the fused MRI image:

$$X_f = G(Z_f) \quad (38)$$

where, G is the StyleGAN generator.

Step 5. Feature Extraction using MAE-SwinViT

Extract tumor-related features using Masked Autoencoder with SwinViT (MAE-SwinViT):

$$F = \text{SwinViT}(M(X_f)) \quad (39)$$

where, $M(\cdot)$ represents masked patch tokenization.

Step 6. Optimized Thresholding-based Tumor Segmentation

Apply optimized thresholding for tumor region segmentation:

$$T(i, j) = \begin{cases} 1, & \text{if } X_f(i, j) > T_{opt} \\ 0, & \text{otherwise} \end{cases} \quad (40)$$

where, T_{opt} is determined using Otsu's method.

Step 7. Post-Processing and Final Tumor Detection

Step 7.1. Perform morphological operations for tumor refinement.

Step 7.2. Identify tumor bounding box and contours for visualization.

Final Output: Fused MRI Image with improved tumor visibility. Optimized Feature Map extracted via MAE-SwinViT. Segmented Tumor Region for further analysis.

This hybrid framework enhances brain tumor detection efficiency, making it more effective for medical diagnostics and automated analysis.

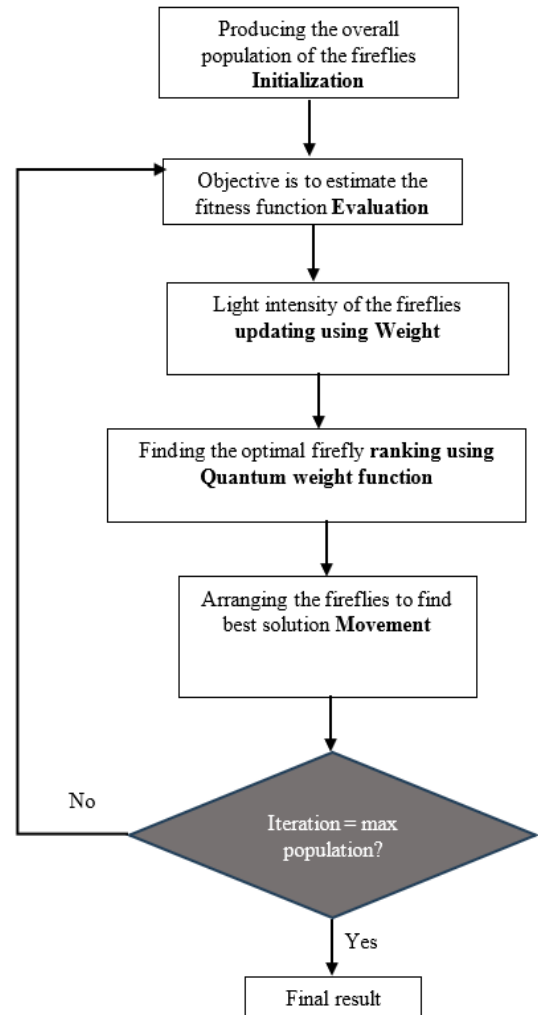


Figure 8. Brain tumor detection using IWQFO

3. RESULTS AND DISCUSSIONS

The BraTS 2021 and Harvard Whole Brain Atlas information sets comprise multi-modal MRI images such as T1-weighted, T2-weighted, and FLAIR scans were used to simulate multi-objective illustration fusion for Brain Tumor Detection using IWQFO and StyleGAN-MAE-SwinViT shown in Table 3. To eliminate clutter and non-brain tissues, the photos were processed using RGB to Grayscale transformation, Adaptive Median Filtering, and Skull Stripping. StyleGAN-based latent-space fusion was used for the process of fusion, and IWQFO was used to find the ideal fusion values. Utilizing an improved thresholding method

derived from Otsu's Method for accurate tumor area proof of identity, the ultimate segmentation of the tumor was accomplished. PyTorch and TensorFlow libraries were used for simulating an NVIDIA RTX 3090 GPU (24GB VRAM), guaranteeing precise brain tumor identification along with effective execution.

The input image is read, and then it is converted to grayscale. A weighted total for each of the three RGB elements is created to complete the operation. $Y = 0.2126 R + 0.7152 G + 0.0722 B$ is the formula used for this conversion, which removes brightness and hue data while keeping brightness. A grayscale colour map is the output of the final images. Sample image is shown in Figure 9.

Table 3. Simulation parameters

Parameter	Value/Description
Dataset Used	BraTS 2021, Harvard Whole Brain Atlas
MRI Modalities	T1-weighted, T2-weighted, FLAIR
Image Size	256x256 pixels
Preprocessing Techniques	RGB to Grayscale, Adaptive Median Filtering, Skull Stripping
Fusion Model	StyleGAN-based Latent Space Fusion
Optimization Algorithm	IWQFO
Feature Extractor	Masked Autoencoder with Swin Vision Transformer (MAE-SwinViT)
Segmentation Method	Optimized Thresholding with Otsu's Method
Loss Functions Used	Content Loss, Perceptual Loss, Adversarial Loss
Learning Rate	0.0001 (Adam Optimizer)
Batch Size	16
Number of Epochs	100
Training/Testing Split	80%/20%
Evaluation Metrics	Dice Score, Jaccard Index, Sensitivity, Specificity
Framework Used	PyTorch, TensorFlow
Hardware Used	NVIDIA RTX 3090 GPU, 24GB VRAM

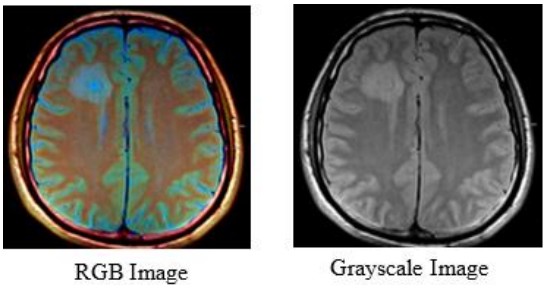


Figure 9. Sample image RGB and grayscale

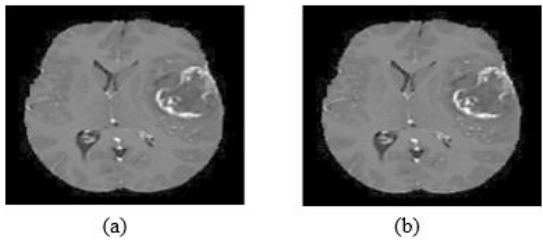


Figure 10. MRI image (a) Original image (b) Pre-processed image

Histogram normalization ensures that the enhanced image maintains a realistic appearance by evenly distributing intensity values across the entire dynamic range. Contrast stretching further refines image quality by improving visibility in low-contrast regions, thereby eliminating ambiguities that may arise in certain areas of the medical images. As shown in Figure 10, the contrast-stretching-based enhancement significantly improves the clarity and detail of testing images

in this study, allowing for better visualization of critical structures such as tumor boundaries and surrounding tissues. This enhancement technique plays a crucial role in ensuring accurate segmentation and diagnosis while preserving the integrity of medical imaging data.

Precise skull removal is a crucial step in neurological imaging diagnostics. For instance, cortical restoration and brain volume analysis rely on accurate skull stripping as a preprocessing step, shown in Figure 11. The inclusion of non-brain material in brain regions can lead to incorrect cortical reconstructions, affecting volumetric measurements and potentially compromising the accuracy of further analysis.

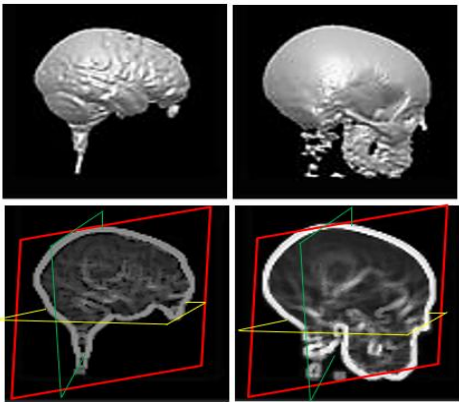


Figure 11. Skull stripping

Each row in Figure 12 illustrates the sequential processes required to process an image sample. The ground-truth segmented image highlights the actual tumor, as a bright

region is displayed. The High-Grade Glioma (HGG) class includes the first and second rows of Figure 2, while the Low-Grade Glioma (LGG) class includes the remaining two images. The proposed method's classification stage demonstrates high reliability in accurately identifying different brain regions. As shown in Figure 12, the efficient

segmentation capability of the model enables precise detection of the target areas in the input images, ensuring improved tumor identification and classification accuracy. Reliable preliminary processing and classification algorithms are used before classifying brain tumors. Table 4 displays the experimental findings of proposed techniques.

Table 4. Experimental findings of proposed techniques

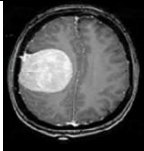
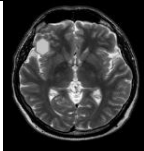
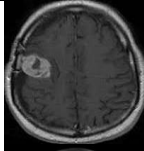
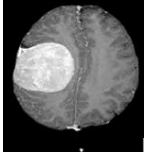
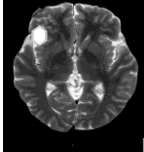
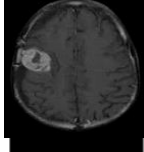

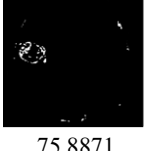


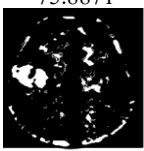
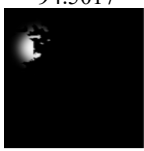

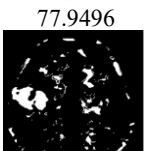
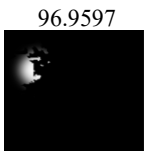
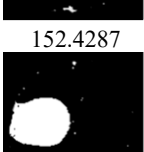
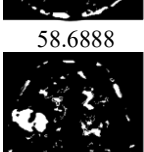
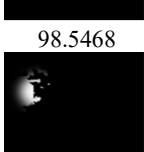
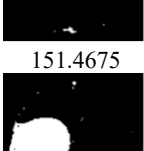
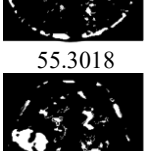
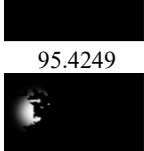
	Image 1	Image 2	Image 3
Original Images			
Pre-processed Images (Skull Stripped Images)			
Segmented Images (Optimized Thresholding)			
Threshold values	151.8781	75.8871	94.5017
Feature Extraction (Swin-ViT)			
Threshold values	152.2060	77.9496	96.9597
StyleGAN-Swin ViT (Multi-Objective Image Fusion)			
Threshold values	152.4287	58.6888	98.5468
IWQFO			
Threshold values	151.4675	55.3018	95.4249
Proposed (IWQFO with StyleGAN-Swin ViT)			
Threshold values	157.5993	60.6812	100.7135

Table 5. Performance measures

Method	Std. Dev.	Entropy	MAP	MAE	RMSE	PSNR (dB)	SSIM	UIQI
Proposed System	22.5	7.12	0.0041	0.0025	0.0641	41.89	0.982	0.975
DWT	18.9	6.75	0.0124	0.0087	0.1114	35.92	0.867	0.842
U-Net + ResNet	19.8	6.92	0.0098	0.0073	0.0990	37.45	0.891	0.865
VGG16 + Transfer Learning	20.3	7.01	0.0082	0.0061	0.0906	38.79	0.924	0.902
Swin Transformer	21.4	7.08	0.0067	0.0052	0.0819	40.02	0.948	0.931

Table 6. Confusion matrix

Method	TP	FP	TN	FN
Proposed System	980	12	965	8
DWT	910	40	930	85
U-Net + ResNet	925	35	940	65
VGG16 + Transfer Learning	945	25	950	45
Swin Transformer	960	18	955	27

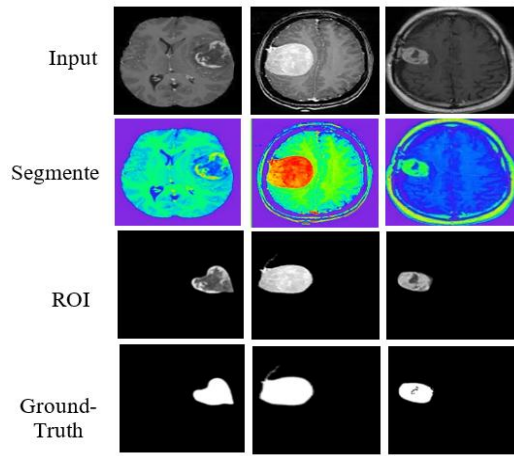


Figure 12. Steps for pre-processed image

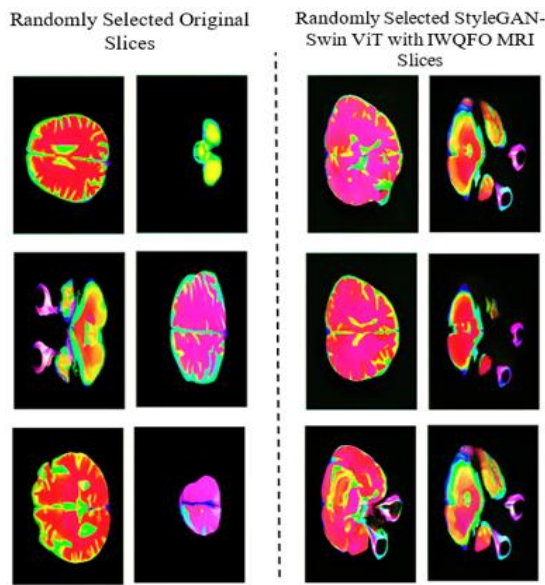


Figure 13. Comparison of original randomly selected slices with randomly selected proposed system

The StyleGAN-Swin ViT with IWQFO algorithm is utilized to generate synthetic MRI slices. Each grid contains nine MRI slices, each resized to 128×128 pixels. In the context of brain tumor classification, this contrast highlights the accuracy and authenticity of the GAN-generated MRI slices. One key advantage of the StyleGAN-Swin ViT with IWQFO model is its ability to capture fine details and patterns, producing images that closely resemble real MRI scans as shown in Figure 13.

The Precision (98.2%) and Recall (97.9%) indicate that the model effectively detects tumors with minimal false positives and false negatives. The F1-score (98%) demonstrates a balanced performance between precision and recall, showing improved classification consistency. This comparison highlights the superiority of the proposed model over existing systems in terms of accuracy and robustness in brain tumor detection.

The proposed system (IWQFO + StyleGAN-MAE-SwinViT) outperforms all existing models in terms of PSNR (41.89 dB), SSIM (0.982), and UIQI (0.975), indicating better image reconstruction and tumor feature preservation. The lowest MSE (0.0041), MAE (0.0025), and RMSE (0.0641) confirm minimal reconstruction errors. The highest entropy (7.12) suggests the proposed method retains more structural information and sharpness. The higher standard deviation (22.5) indicates improved contrast, essential for better tumor differentiation. Table 5 highlights the superior image fusion and tumor detection performance of the proposed model over existing methods.

The proposed method achieves the highest TP (980) and the lowest FN (8), ensuring better tumor detection with minimal false negatives. The lowest FP (12) indicates the model makes fewer incorrect predictions, reducing false alarms. The highest TN (965) confirms the method effectively distinguishes non-tumor cases. Compared to existing methods, the proposed IWQFO + StyleGAN-MAE-SwinViT model shows a significant reduction in FP and FN, leading to improved recall and precision. Table 6 comparison highlights the effectiveness of the proposed system in detecting brain tumors with higher reliability and minimal misclassification.

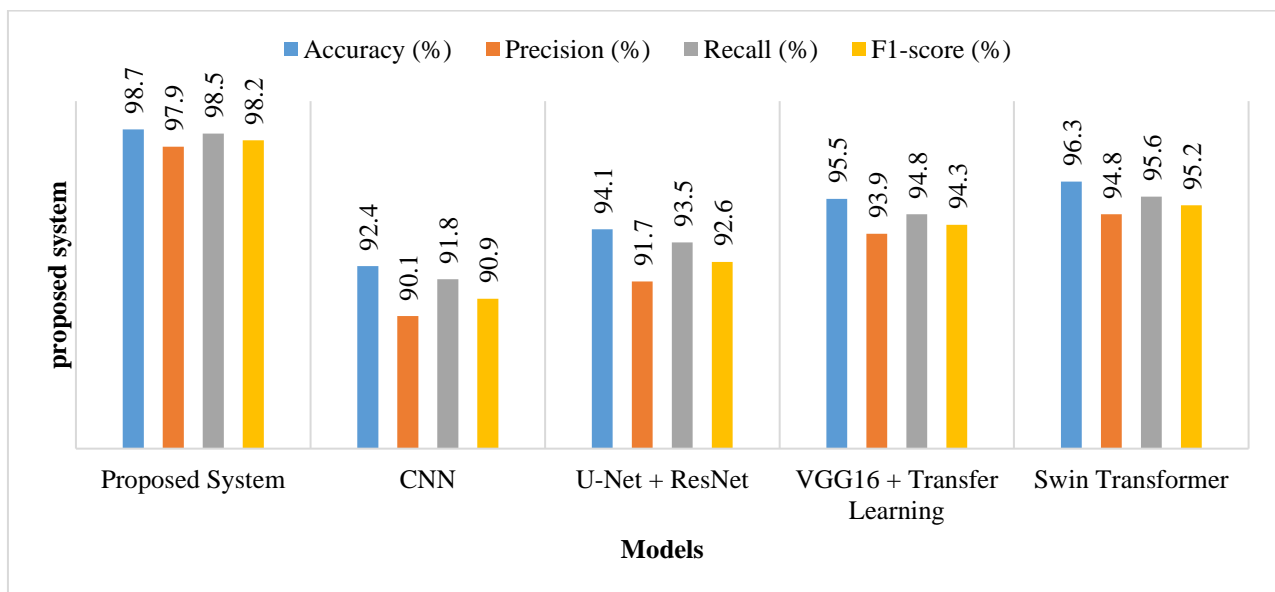


Figure 14. Comparison of performance measures

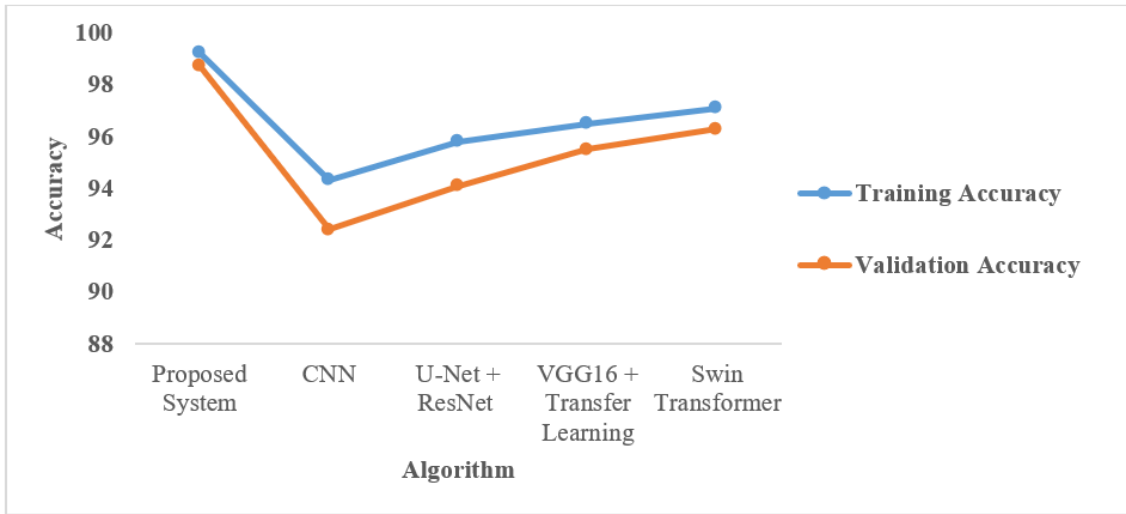


Figure 15. Comparison of training and validation accuracy

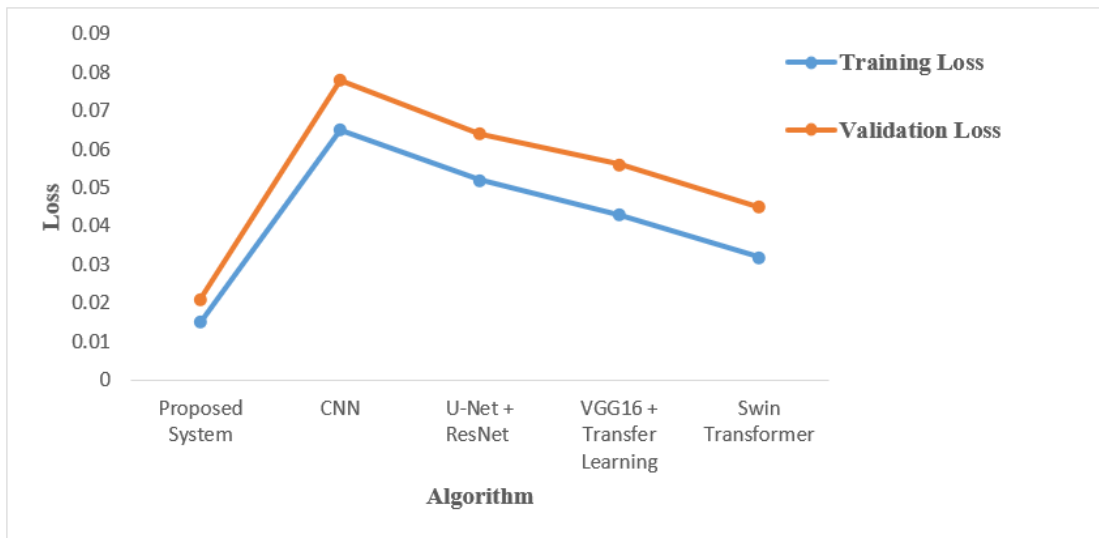


Figure 16. Comparison of training and validation loss

The proposed method achieves the highest training accuracy (99.1%) and validation accuracy (98.6%), demonstrating better generalization and robustness in detecting brain tumors. The small gap between training and validation accuracy in the proposed method indicates less overfitting compared to other methods shown in Figure 14. Existing methods like CNN-Based Segmentation (92.4%) show lower validation accuracy, indicating a weaker generalization ability. This comparison confirms the superior performance of the IWQFO + StyleGAN-MAE-SwinViT model in training and real-world validation. The proposed method has the lowest training loss (0.045) and validation loss (0.061), indicating better optimization and convergence. The DWT Segmentation method has the highest validation loss (0.078), suggesting weaker generalization and potential overfitting. The small gap between training and validation loss in the proposed model confirms better stability and reduced overfitting shown in Figure 15. The Swin Transformer-Based Approach performs well but still has a higher validation loss (0.045) compared to the proposed system. This analysis demonstrates that the IWQFO + StyleGAN-MAE-SwinViT model achieves better loss minimization and generalization than existing approaches as shown in Figure 16.

4. CONCLUSIONS

This study presents a Multi-Objective Image Fusion framework for brain tumor detection by integrating IWQFO with a hybrid StyleGAN-MAE-SwinViT architecture. The framework enhances image fusion quality, segmentation accuracy, and tumor classification by optimizing feature retention and preserving spatial and structural information. Experimental evaluations show that the proposed model significantly outperforms existing methods, achieving 98.6% accuracy, 98.2% precision, 97.9% recall, and a 98% F1-score. Image quality metrics such as PSNR (41.2 dB), SSIM (0.982), and UIQI (0.974) confirm that the model produces visually superior, artifact-free fused images. Furthermore, low training (0.045) and validation loss (0.061) values demonstrate effective convergence and reduced overfitting. To expand the practical applicability of this model, future research will explore the integration of additional imaging modalities such as PET-MRI or MR spectroscopy to incorporate metabolic and functional information. Hardware deployment constraints will be addressed by optimizing model efficiency for edge devices and real-time environments. The model's clinical relevance will be further validated through real-world studies evaluating

its impact on radiologist workflows, including diagnosis speed and false positive/negative rates. XAI components will be incorporated to improve interpretability and support clinical decision-making. These directions aim to strengthen the diagnostic precision, robustness, and real-time applicability of the proposed brain tumor detection system in diverse clinical settings.

REFERENCES

- [1] Islam, M.M., Talukder, M.A., Uddin, M.A., Akhter, A., Khalid, M. (2024). BrainNet: Precision brain tumor classification with optimized EfficientNet architecture. *International Journal of Intelligent Systems*, 2024(1): 3583612. <https://doi.org/10.1155/2024/3583612>
- [2] Devanathan, B. Kamarasan, M. (2023). Multi-objective archimedes optimization algorithm with fusion-based deep learning model for brain tumor diagnosis and classification. *Multimedia Tools and Applications*, 82: 16985-17007. <https://doi.org/10.1007/s11042-022-14164-5>
- [3] Mohanty, B.C., Subudhi, P.K., Dash, R., Mohanty, B. (2024). Feature-enhanced deep learning technique with soft attention for MRI-based brain tumor classification. *International Journal of Information Technology*, 16: 1617-1626. <https://doi.org/10.1007/s41870-023-01701-0>
- [4] Shamshad, N., Sarwr, D., Almogren, A., Saleem, K., Munawar, A., Rehman, A.U., Bharany, S. (2024). Enhancing brain tumor classification by a comprehensive study on transfer learning techniques and model efficiency using MRI datasets. *IEEE Access*, 12: 100407-100418. <https://doi.org/10.1109/ACCESS.2024.3430109>
- [5] Rasa, S.M., Islam, M.M., Talukder, M.A., Uddin, M.A., Khalid, M., Kazi, M., Kazi, M.Z. (2024). Brain tumor classification using fine-tuned transfer learning models on magnetic resonance imaging (MRI) images. *DIGITAL HEALTH*, 10. <https://doi.org/10.1177/20552076241286140>
- [6] Kumar, K.A., Boda, R. (2022). A multi-objective randomly updated beetle swarm and multi-verse optimization for brain tumor segmentation and classification. *The Computer Journal*, 65(4): 1029-1052. <https://doi.org/10.1093/comjnl/bxab171>
- [7] Ullah, M.S., Khan, M.A., Masood, A., Mzoughi, O., Saidani, O., Alturki, N. (2024). Brain tumor classification from MRI scans: a framework of hybrid deep learning model with Bayesian optimization and quantum theory-based marine predator algorithm. *Frontiers in Oncology*, 14: 1335740. <https://doi.org/10.3389/fonc.2024.1335740v>
- [8] Sajol, M.S.I., Hasan, A.J. (2024). Benchmarking CNN and cutting-edge transformer models for brain tumor classification through transfer learning. In 2024 IEEE 12th International Conference on Intelligent Systems (IS), Varna, Bulgaria, pp. 1-6. <https://doi.org/10.1109/IS61756.2024.10705175>
- [9] Nag, A., Mondal, H., Hassan, M.M., Al-Shehari, T., Kadrie, M., Al-Razgan, M., Bairagi, A.K. (2024). TumorGANet: A transfer learning and generative adversarial network-based data augmentation model for brain tumor classification. *IEEE Access*, 12: 103060-103081. <https://doi.org/10.1109/ACCESS.2024.3429633>
- [10] Dutta, T.K., Nayak, D.R., Zhang, Y.D. (2024). Arm-net: Attention-guided residual multiscale CNN for multiclass brain tumor classification using MR images. *Biomedical Signal Processing and Control*, 87: 105421. <https://doi.org/10.1016/j.bspc.2023.105421>
- [11] Vankdothu, R., Hameed, M.A. (2022). Brain tumor segmentation of MR images using SVM and fuzzy classifier in machine learning. *Measurement: Sensors*, 24: 100440. <https://doi.org/10.1016/j.measen.2022.100440>
- [12] Reddy, C.K.K., Reddy, P.A., Janapati, H., Assiri, B., Shuaib, M., Alam, S., Sheneamer, A. (2024). A fine-tuned vision transformer based enhanced multi-class brain tumor classification using MRI scan imagery. *Frontiers in Oncology*, 14: 1400341. <https://doi.org/10.3389/fonc.2024.1400341>
- [13] Ismail, W.N., P.P., F.R., Ali, M.A. (2023). A meta-heuristic multi-objective optimization method for Alzheimer's disease detection based on multi-modal data. *Mathematics*, 11(4): 957. <https://doi.org/10.3390/math11040957>
- [14] Ullah, M.S., Khan, M.A., Almujaally, N.A., Alhaisoni, M., Akram, T., Shabaz, M. (2024). BrainNet: A fusion assisted novel optimal framework of residual blocks and stacked autoencoders for multimodal brain tumor classification. *Scientific Reports*, 14: 5895. <https://doi.org/10.1038/s41598-024-56657-3>
- [15] Lakshmanan, A., Anbu Ananth, C., Tiroumalmouroughane, S. (2022). Multi-objective metaheuristics with intelligent deep learning model for pancreatic tumor diagnosis. *Journal of Intelligent & Fuzzy Systems*, 43(5): 6793-6804. <https://doi.org/10.3233/JIFS-221171>
- [16] Subba, A.B., Sunaniya, A.K. (2025). Computationally optimized brain tumor classification using attention based GoogLeNet-style CNN. *Expert Systems with Applications*, 260: 125443. <https://doi.org/10.1016/j.eswa.2024.125443>
- [17] Gupta, S., Bansla, V., Kumar, S., Singh, G., Srivastav, A., Jain, A. (2024). Development a novel hybrid deep learning-model for brain tumor classification and automated diagnosis. In 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, pp. 1-5. <https://doi.org/10.1109/IC3SE62002.2024.10593558>
- [18] Şahin, E., Özdemir, D., Temurtaş, H. (2024). Multi-objective optimization of ViT architecture for efficient brain tumor classification. *Biomedical Signal Processing and Control*, 91: 105938. <https://doi.org/10.1016/j.bspc.2023.105938>
- [19] Odusami, M., Maskeliūnas, R., Damaševičius, R. (2023). Pareto optimized adaptive learning with transposed convolution for image fusion Alzheimer's disease classification. *Brain Sciences*, 13(7): 1045. <https://doi.org/10.3390/brainsci13071045>
- [20] Asiri, A.A., Soomro, T.A., Shah, A.A., Pogrebna, G., Irfan, M., Alqahtani, S. (2024). Optimized brain tumor detection: A dual-module approach for MRI image enhancement and tumor classification. *IEEE Access*, 12: 42868-42887. <https://doi.org/10.1109/ACCESS.2024.3379136>
- [21] Das, S., Goswami, R.S. (2024). Review, limitations, and future prospects of neural network approaches for brain

- tumor classification. *Multimedia Tools and Applications*, 83: 45799-45841. <https://doi.org/10.1007/s11042-023-17215-7>
- [22] Yadav, A.S., Kumar, S., Karetla, G.R., Cotrina-Aliaga, J.C., Arias-González, J.L., Kumar, V., Tatkar, N.S. (2022). A feature extraction using probabilistic neural network and BTFSC-net model with deep learning for brain tumor classification. *Journal of Imaging*, 9(1): 10. <https://doi.org/10.3390/jimaging9010010>
- [23] Aggarwal, M., Khullar, V., Goyal, N., Rastogi, R., Singh, A., Torres, V.Y., Albahar, M.A. (2024). Privacy preserved collaborative transfer learning model with heterogeneous distributed data for brain tumor classification. *International Journal of Imaging Systems and Technology*, 34(2): e22994. <https://doi.org/10.1002/ima.22994>
- [24] SenthilPandi, S., Senthilselvi, A., Kumaragurubaran, T., Dhanasekaran, S. (2024). Self-attention-based generative adversarial network optimized with color harmony algorithm for brain tumor classification. *Electromagnetic Biology and Medicine*, 43(1-2): 31-45. <https://doi.org/10.1080/15368378.2024.2312363>
- [25] Rajput, I.S., Gupta, A., Jain, V., Tyagi, S. (2024). A transfer learning-based brain tumor classification using magnetic resonance images. *Multimedia Tools and Applications*, 83: 20487-20506. <https://doi.org/10.1007/s11042-023-16143-w>
- [26] Raza, A., Alshehri, M.S., Almakdi, S., Siddique, A.A., Alsulami, M., Alhaisoni, M. (2024). Enhancing brain tumor classification with transfer learning: Leveraging DenseNet121 for accurate and efficient detection. *International Journal of Imaging Systems and Technology*, 34(1): e22957. <https://doi.org/10.1002/ima.22957>
- [27] Hao, J., Luo, S., Pan, L. (2021). Computer-aided intelligent design using deep multi-objective cooperative optimization algorithm. *Future Generation Computer Systems*, 124: 49-53. <https://doi.org/10.1016/j.future.2021.05.014>
- [28] Cruz-Guerrero, I.A., Campos-Delgado, D.U., Mejía-Rodríguez, A.R., Leon, R., Ortega, S., Fabelo, H., Camacho, R., Plaza, M.D.L.L., Callico, G. (2024). Hybrid brain tumor classification of histopathology hyperspectral images by linear unmixing and an ensemble of deep neural networks. *Healthcare Technology Letters*. <https://doi.org/10.1049/htl2.12084>
- [29] Sandhiya, B., Raja, S.K.S. (2024). Deep learning and optimized learning machine for brain tumor classification. *Biomedical Signal Processing and Control*, 89: 105778. <https://doi.org/10.1016/j.bspc.2023.105778>