International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Image Fusion of MRI and CT Scan for Brain Tumor Detection Using VGG- 19

Rashmi Ashtagi[1], Swati Jadhav[2], Abbas Madhavaswala[3], Ripul Handoo[4], Devanshu Rathi[5], Rishi Purohit[6]

Submitted: 26/10/2023 **Revised**: 16/12/2023 **Accepted**: 26/12/2023

Abstract: Brain tumor (BT) detection is crucial for patient outcomes, and bio-imaging techniques like Magnetic Resonance Image (MRI) and Computed Tomography (CT) scans play a vital role in clinical assessment. However, manual analysis of these images is time-consuming and requires expertise. To address this, we propose an image fusion model that combines MRI and CT images using Wavelet-based fusion and leverages the VGG-19 architecture for improved accuracy. Image fusion combines modalities, enhancing their strengths while mitigating weaknesses. Our method employs the Wavelet fusion technique, decomposing images into frequency bands. The low-frequency LL band holds key structural information. The VGG-19 network, with its convolutional and pooling layers, is used to merge LL bands, reconstructing fused images. We conduct evaluations on brain MRI and CT images, employing preprocessing, feature extraction, and fusion stages. Our approach not only reduces the doctor's workload and analysis time but also enhances tumor detection accuracy. Automation of image analysis and early, accurate tumor identification lead to better patient care.

***Keywords**: Brain tumor detection, MRI, CT scan, Wavelet-based fusion, VGG-19 architecture, image analysis..*

1. Introduction

A brain tumor (BT) is a form of brain abnormality that can arise due to various factors. Left unrecognized and untreated, brain tumors can significantly increase morbidity and mortality rates. Clinical assessment of brain tumors is commonly conducted using bio-imaging techniques, such as Computed Tomography (CT) scans and Magnetic Resonance Imaging (MRI). Early detection is crucial for improving patient outcomes. MRI utilizes magnetic fields and radio frequencies to generate detailed images, particularly useful for soft tissue imaging, including tumors and brain tissues. On the other hand, CT scans employ X-rays for detailed imaging, primarily capturing hard tissues like bones. Despite their utility, each modality possesses distinct strengths and weaknesses. The manual analysis of a large volume of MR and CT images by specialists for disease detection is labor-intensive and time-consuming. To address this challenge, this study proposes a model that fuses MR and CT images using Wavelet-based fusion. Additionally, the model employs the VGG19 architecture, enhancing accuracy and scalability. This innovative approach aims to streamline the analysis process and contribute to more efficient and accurate detection of brain abnormalities.

[1]*Computer Department, Vishwakarma Institute of Technology Pune.*
rashmi.ashtagi@vit.edu
[2]*Department of Computer engineering Vishwakarma Institute of Technology Pune.*
swati.jadhav@vit.edu
[3]*Computer Department, Vishwakarma Institute of Technology Pune.*
abbas.madhavaswala21@vit.edu
[4]*Computer Department, Vishwakarma Institute of Technology Pune.*
ripul.handoo21@vit.edu
[5]*Computer Department, Vishwakarma Institute of Technology Pune.*
devanshu.rathi21@vit.edu
[6]*Computer Department, Vishwakarma Institute of Technology Pune*
rishi.purohit21@vit.edu

Image fusion is a technique that combines images from different modalities and combines them into one single image that has the best features of both the images. There are various ways to fuse MRI and CT scan images such as intensity-based fusion (combines the intensity values of the modalities), feature-based fusion (fuses the extracted features from the modalities), Wavelet-based fusion wavelet transforms to decompose the MRI and CT images into different frequency bands. These frequency bands are then fused to create a single image). We have used Wavelet-fusion method to fuse the MRI and CT scan images. The proposed method generates LL band, LH band, LV band and LD band by applying fusion based on the VGG-19 network to four sets of pictures. These are not an anatomical structure that can be generated by the different modalities. These are mathematical constructs used in image fusion techniques.

Wavelet transforms are a technique that breaks down an image into a series of frequency bands. Among these bands, the LL band is recognized as the low-frequency band, encompassing crucial information about the overall structure of the image. Additionally, the LH band represents the horizontal band, capturing details about horizontal edges in the image, while the LV band pertains to the vertical band, preserving information about vertical edges. The LD band corresponds to the diagonal band, focusing on diagonal edges within the image. In the process of image fusion, LL bands take precedence, as they encapsulate a substantial amount of information and play a pivotal role in preserving the image's fundamental structure. During the fusion process, these bands are typically merged first. For image processing tasks, particularly identification and classification, deep convolutional neural networks such as the VGG-19 network are widely utilized. In the context of the proposed method, it involves the merging of four sets of images, each having undergone discrete wavelet transformation into four distinct frequency bands. This innovative approach aims to leverage the strengths of wavelet transformations and deep learning to enhance picture identification and classification tasks.

The VGG19 architecture is made up of a series of 18 convolutional layers and pooling layers. The convolutional layers extract features from the input image, and the pooling layers reduce the size of the feature maps. The VGG19 architecture has been shown to be very effective at learning to recognize objects in images. The VGG-19 network is used to combine the LL band, and the inverse discrete wavelet transform is used to reconstruct the fused picture. An evaluation of the suggested method is done using a collection of brain MRI and CT images. The proposed methodology consists of three key stages: preprocessing, feature extraction, and fusion. In the preprocessing stage, MRI and CT images are pre-processed to enhance image quality and address differences in imaging characteristics. The feature extraction stage involves employing the pre-trained VGG-19 model to extract deep features from both MRI and CT images, capturing high-level representations specific to each imaging modality. Subsequently, a fusion strategy is employed to combine the extracted features from MRI and CT images to create a fused representation. The fused feature representation aims to preserve salient information from both modalities while minimizing redundancy, leading to an enhanced and more informative tumor representation.

In Real Life, this approach of detection of brain tumors can reduce the workload of doctors as it is quite tedious to study the scan. saves the time of the doctors. With image fusion, the doctor can use a software tool to automatically segment the tumor, which can save them a significant amount of time. In addition to reducing the workload of doctors, image fusion can also improve the quality of care that patients receive. By automating the process of image analysis and improving the accuracy of tumor detection, image fusion can help doctors to identify tumors earlier and more accurately. This can lead to better outcomes for patients.

2. Literature Survey

R.Jain et al. [1] has discussed how Perfusion Computed Tomography (PCT) and Magnetic Resonance (MR) perfusion imaging have become crucial tools in the assessment of brain tumors. These imaging modalities offer unique insights beyond standard morphological imaging by providing physiologic and hemodynamic data. PCT, in particular, stands out for its ability to simultaneously measure blood volume and permeability, streamlining the assessment process. Additionally, PCT's availability of an arterial input function enhances its accuracy. Both PCT and MR perfusion imaging serve as valuable imaging biomarkers for glioma grading, angiogenesis assessment, treatment planning, and response evaluation, significantly improving the diagnosis and management of brain tumors.

Marco Alfonse and Abdel-Badeeh M.Salem et al. [2] have introduced a diagnostic system designed for distinguishing between benign and malignant brain tumors. This system relies on Magnetic Resonance Imaging (MRI) images in DICOM format and encompasses several key phases. The initial stages involve dataset acquisition, preprocessing, and segmentation. Segmentation is carried out utilizing the Expectation Maximization (EM) algorithm and adaptive thresholding. This process is instrumental in isolating the primary brain region while eliminating extraneous components. Notably, structural element techniques are also employed during segmentation. Following segmentation, the system proceeds with feature extraction utilizing the Fast Fourier Transform (FFT) technique. Feature selection is then executed through the Minimum Redundancy Maximum Relevance (MRMR) criterion, identifying the most pertinent features for further analysis. The final phase involves classification, where a Support Vector Machine (SVM) is employed. The overall performance of the proposed system is notable, achieving an impressive classification accuracy rate of 98.9%. The combination of segmentation techniques, feature extraction, selection, and classification contributes to the system's effectiveness in diagnosing brain tumors accurately.

Soobia Saeed, Afnizanfaizal Abdullah and NZ. Jhanjhi et al. [3] have presented research which delves into 3D image analysis to understand neuronal interactions, focusing on identifying tumors and their impact on brain tissue due to increased glucose metabolism. The study examines slowly developing brain and oral tumors, aiming to mitigate their effects through chemotherapy.The author employs MATLAB for 3D segmentation and statistical modeling, with the goal of detecting CSF leakage and distinguishing tumor regions from non-tumor areas. Ultimately, this work seeks to enhance cancer treatment outcomes and envision the future of this technology in the medical field.

Abhishek Anil, Aditya Raj, H Aravind Sarma, Naveen Chandran R, Deepa P L et al. [4] explores the detection of brain tumors utilizing deep learning networks, particularly employing a classification network that categorizes input MR images into two classes: those with tumors and those without. Transfer learning is the method of choice for retraining the classifier, yielding superior results compared to existing methods. The research involves three distinct networks, with VGG19 emerging as the most effective detector among them. The networks are trained for classification with a dataset featuring 1,000 classes, highlighting the potential of transfer learning in brain tumor detection.

X Ma, S Hu, S Liu, J Fang, S Xu et al. [5] **i**ntroduces a remote sensing image fusion method leveraging sparse representation (SR). The process involves learning an adaptive dictionary from source images, obtaining sparse coefficients through coding with this dictionary, and fusing these coefficients using an improved hyperbolic tangent function (tanh) and l0−max. The resulting initial fused image is further enhanced with spatial domain fusion (SF). Ultimately, the final fused image is reconstructed via guided filtering. Experimental results demonstrate the superiority of this approach over existing methods in both visual and quantitative assessments.

J Gao, J Li, M Jiang in et al. [6] addresses the trade-off between spectral and spatial resolution in hyperspectral imagery by introducing a self-supervised fusion method for hyperspectral and multispectral images.Unlike traditional techniques, this innovative approach does not rely on training datasets, making it versatile for scenarios where such data is limited or of low quality. The method leverages constraints derived from low-resolution hyperspectral images and synthetic high-resolution hyperspectral images. Extensive simulations and real-data experiments confirm the superiority of this approach, offering enhanced fusion performance compared to conventional methods, thus unlocking new possibilities for high-resolution hyperspectral imaging applications.

S Sharma, S Gupta, D Gupta, A Juneja, H Khatter, S Malik, ZK Bitsue in et al. [7] discussed that brain tumors, whether benign or malignant, pose severe health risks by increasing intracranial pressure and impairing cognitive functions. Current diagnostic methods like CT, MRI, PET scans, and blood tests are often time-consuming and may yield inaccurate results. To address this, a deep learning model based on pretrained VGG19 is proposed in this paper. The model, enhanced with normalization and data augmentation, achieves an impressive 98% accuracy and 94.73% sensitivity using a dataset of 257 images, 157 with brain tumors and 100 without. These results suggest the potential for clinically valuable BT detection solutions in CT images.

J Fu, W Li, A Ouyang, B He - Optik, 2021 – Elsevier et al. [8] have introduced a novel approach to enhance multimodal biomedical image fusion quality. By combining the rolling guidance filter and deep convolutional neural networks (CNNs), the method addresses issues like low luminance and blurred edges. It involves three key steps: base-detail extraction, perceptual image generation using the VGG network, and fusion with three distinct strategies. Experimental results show that this approach outperforms existing methods, offering improved image quality and detail enhancement in biomedical image fusion.

Hoo-Chang Shin; Holger R. Roth; Mingchen Gao; Le Lu; Ziyue Xu; Isabella Nogues et al. [9] addresses challenges in applying deep CNNs to medical image analysis, emphasizing architecture, dataset scale, and transfer learning. It investigates thoraco-abdominal lymph node detection and interstitial lung disease classification, achieving state-of-the-art results. Insights provide valuable guidance for high-performance computer-aided diagnosis systems in medical imaging across various tasks.

S Banerjee, DP Mukherjee, DD Majumdar in et al. [10] introduces a novel approach to biomedical image registration using point landmarks and leveraging geometric invariance properties. These landmarks are identified at the entrance and exit points of concave structures and inflection points of curves, obtained from the convex hull of the structures. The proposed registration technique operates in a standard reference frame and offers speed, semi-automation, and computational efficiency. The method is particularly suitable for registering functional images like Positron Emission Tomography and Single Photon Emission Computed Tomography with morphological images such as CT and MRI. In functional images, concavities manifest as valleys amid intensity variations. The method's potential extension to three-dimensional scenarios is also explored. Overall, the approach presents a promising avenue for improving image registration in biomedical contexts.

3. Methodology:

The procedural for the model is proposed in such a way where image registration of Magnetic Resonance Image (MRI) and Computed Tomography (CT) Scan images of the same patient is done using template made using Hypertext Markup Language (HTML) and Cascading Style Sheets (CSS). To fuse the image manual datapoints need to be registered for resizing the images. Then using transfer learning which is used to enhance the performance of multimodal image fusion models, particularly when there is a dearth of labeled data or when one modality suffers from severe noise or other drawbacks. If the images are of not the same dimensions, they are resized and then fused to give a fused image. After the images are fused using watershed algorithm image is being deployed to get the flood height of the images for better boundary separation. Watershed Flooding is applied to get the segmented image. A technique named Region Merging is applied to remove the false boundaries creating during the Watershed Flooding. The results of this model will give us the fused image of the MRI and CT scan and the segmented image of this fused image which helps to enhance the accuracy of detecting the brain tumor location.

VGG 19:

The VGG19 Convolutional Neural Network (CNN) Network is an image classification convolutional neural network with a deep design. The VGG19 [22] model has one input layer, two output layers, and 18 hidden layers. You have the option to utilize a pre-trained variant of the network that has been trained on an extensive dataset containing over a million images sourced from the ImageNet database. This pre-trained network is capable of categorizing images into 1000 distinct object categories, encompassing items like keyboards, mice, pencils, as well as various animals. Consequently, this network has acquired comprehensive feature representations, making it suitable for analyzing a diverse array of images. The input layer accepts an image with dimensions of 224x224x3, and the output layers are a fully connected layer that creates a 4096-dimensional feature vector and a softmax layer that produces a probability distribution across the ImageNet dataset's 1000 classifications. VGG19 employs 3x3 filters with a stride of 1 pixel in its convolutional layers, allowing it to learn detailed spatial features. Its max pooling layers, using 2x2 filters with a stride of 2 pixels, reduce spatial dimensions while maintaining depth. VGG19 is versatile in computer vision, especially for transfer learning. It excels in object recognition due to its diverse feature learning. It's also useful in image segmentation, defining image regions, and medical image analysis, enhancing medical image tasks.

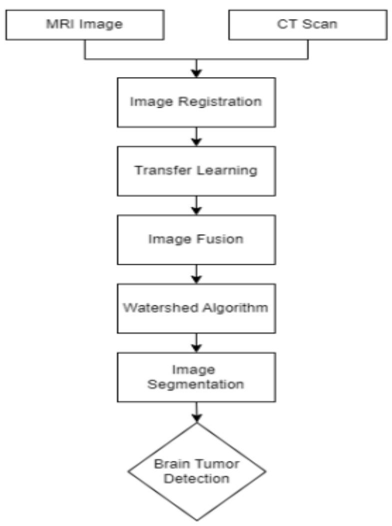

Fig.1.: Flow of Approach

(a) Image Registration:

This is the most crucial step in the image fusion process. This process involves setting up common coordinates for multimodal images such as CT scan, MRI images etc.in order to produce a complete fused image which provides more accurate presentation of the Brain tumor region.

Landmark Registration:

This technique is a form of landmark registration which includes finding and lining up shared landmarks between two images. This strategy involves creating distinctive geometric invariants for diverse sub-contours within both Computed Tomography (CT) and Magnetic Resonance (MR) images. When contours or sub-contours exhibit identical shapes and structures, they share the same set of invariants. While this doesn't guarantee exact matches, these invariants serve as hypotheses for potential matches, which can later be validated. To propose matches, absolute invariants are crucial, and employing local invariants ensures independence from overarching curve properties. This approach enables the detection of invariants even when segments are obscured or omitted due to segmentation issues. In summary, the key steps are as follows:

(1) Generating Geometric Invariants: Develop a collection of geometric invariants for distinct sub-contours within CT and MR images. These invariants encapsulate attributes that remain unaltered under specific transformations.[11]

(2) Identifying Potential Matches: Match sub-contour pairs by comparing their geometric invariants. This matching process can be optimized using hashing or simpler algorithms, considering the limited number of curves.[11]

(3) Validation in a Standardized Frame: After suggesting potential matches, pinpoint landmarks or inflection points within localized sub-contours. These points remain discernible even when subjected to transformations. Transform the sub-contours into a standardized frame using these points. Validate the proposed matches within this standardized frame[11].

(4) Sub-Contour Transformation and Image Alignment: The transformation of sub-contours within the standardized frame leads to image alignment. Aligning contours from different imaging modalities within this frame highlights disparities between the datasets. This information can provide insights into biological growth or the absence of specific structures within a region of interest. This approach primarily concentrates on localized sub-contours or regions of interest within the ventricles of the human brain. The implementation encompasses procedures to identify regions of interest, calculate geometric invariants, and shift these regions into a standardized frame. This method facilitates the analysis and comparison of images from distinct modalities, enabling the identification of variations and alterations in biological structures.[11]

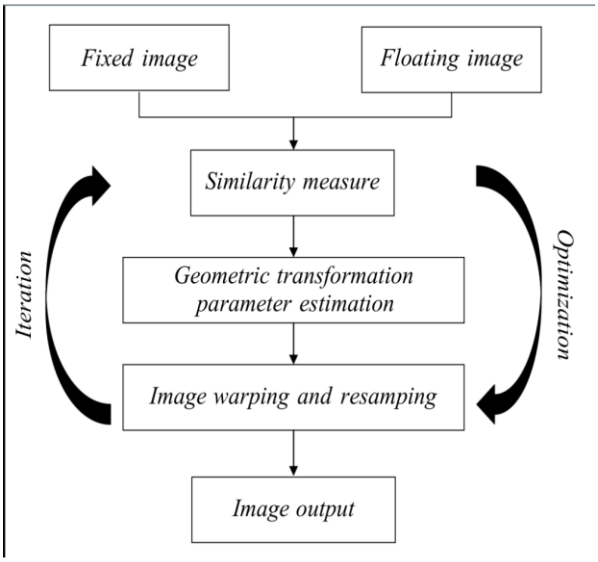

Fig.2. Landmark Registration

(b) Transfer Learning:

Transfer learning in multimodal image fusion involves adapting a pretrained deep neural network, originally trained on one modality, to enhance the performance of fusion tasks using another modality. For instance, a model designed for magnetic resonance imaging (MRI) can be fine-tuned for computed tomography (CT) images in the medical image fusion context.[12] Leveraging transfer learning, the model leverages learned features from one modality to improve feature representation and classification accuracy in another modality. This proves especially advantageous when dealing with data scarcity or noise challenges in one modality, as the learned features can compensate for these limitations. Transfer learning finds applications in various tasks such as face recognition, medical image analysis, and multispectral or hyperspectral remote sensing, yielding improved outcomes across diverse domains.[13][14][15][23].

Transfer learning is an effective method for enhancing the performance of multimodal image fusion models, particularly when there is a dearth of labeled data or when one modality suffers from severe noise or other drawbacks.

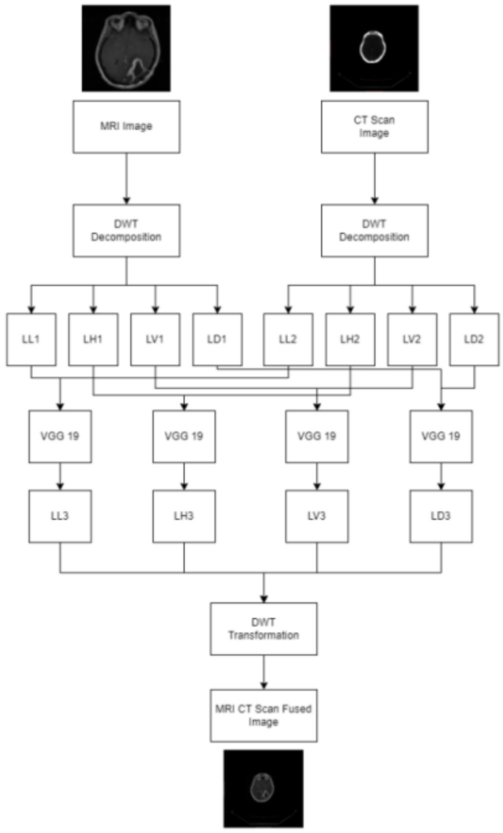

Fig.3. Transfere Learning using VGG19

Discrete Wavelet Transform:

Discrete Wavelet Transform (DWT) is a mathematical technique that decomposes a signal or image into different frequency components using wavelets. Unlike the traditional Fourier Transform, which provides information about frequency content across the entire signal, DWT breaks down the signal into both frequency and spatial information. It involves a series of filtering and down sampling operations, resulting in a multi-resolution representation. DWT is particularly useful for analyzing signals with both high and low-frequency components simultaneously, making it suitable for applications such as image compression, denoising, feature extraction, and pattern recognition. This transform has been widely applied in various fields, including signal processing, image analysis, and data compression.

Let $h(n)$ and $g(n)$ be the low-pass and high-pass filters, and let $x(n, m)$ be the input picture. These formulae may be used to calculate the DWT:

Approximation coefficients:

$$cA(j, n, m) = (h * h * x)(2^{-j}n, 2^{-j}m)$$

Horizontal detail coefficients:

$$cH(j, n, m) = (h * g * x)(2^{-j}n, 2^{-j}m)$$

Vertical detail coefficients:

$$cV(j,n,m) = (g * h * x)(2^{-j}n, 2^{-j}m)$$

Diagonal detail coefficients:

$$cD(j,n,m) = (g * g * x)(2^{-j}n, 2^{-j}m)$$

where j is the scale of the decomposition, n and m are the spatial coordinates of the picture, and * stands for convolution. The coefficients at each level j correspond to a particular frequency range and orientation of the picture.

(c)VGG 19:

The deep convolutional neural network (CNN) called VGG-19 was developed by the Visual Geometry Group at the University of Oxford [16]. With a total of 19 layers, it comprises 16 convolutional layers and three fully connected layers, showcasing remarkable performance in tasks related to image recognition. Within the context of image fusion, VGG-19 can be employed as a feature extractor to capture high-level features from input images. Subsequently, these features can guide the fusion process [16]. Typically, the network is pretrained using supervised learning techniques on substantial datasets such as ImageNet, and its weights are fine-tuned to minimize classification errors on the training set [16].

In VGG-19, each layer generates a set of feature maps that capture different levels of abstraction within the input images. Deeper layers encapsulate more complex entities such as objects and scenes, while shallower layers capture basic details like edges and textures. The structure of VGG-19 relies on convolutional processes, involving the application of a sequence of filters to the input image and the computation of dot products between filter coefficients and corresponding pixel values. To introduce non-linearity into the model's output after each convolutional layer, the rectified linear unit (ReLU) serves as a non-linear activation function. Suppose we denote the set of weights for a specific layer in VGG-19 as W, and the input image as X. The output of the layer can be computed as follows:

$$Y = f(W * X + b)$$

where b is the bias term, f is the activation function, and * stands for convolution. Gradient descent and backpropagation are used to learn the weights and bias terms during the training phase.

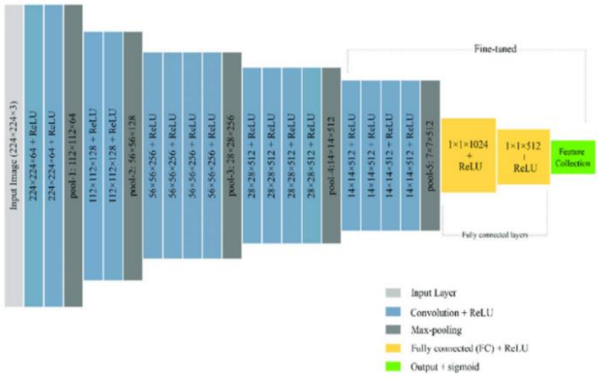

Fig 4. VGG 19 Architecture

(d) Image Segmentation:

In Medical Field image segmentation ids used to distinguish between different boundaries of brain tumor from healthy surrounding area after image fusion of multimodal images for brain tumor detection.[17] The post-multimodal image fusion result undergoes processing through various image segmentation techniques, aimed at distinguishing the tumor location from the adjacent healthy tissues. These methods draw upon diverse theories, including edge detection, region growing, intensity thresholding, and machine learning-based algorithms. One specific technique employed for image segmentation is the Watershed Transform. In summary, image segmentation subsequent to multimodal image fusion plays a pivotal role in precisely characterizing brain tumors and guiding clinical decision-making. Achieving accurate and reliable outcomes necessitates a deep understanding of medical imaging, signal processing, and machine learning.

Watershed Transform: In Image processing, watershed is transformation defined by a grayscale image. A mathematical method for segmenting images called the watershed algorithm is based on the idea of flooding a topographical surface. Each pixel in the image is given a height value depending on its intensity by the algorithm, which then divides the image into areas by flooding the surface from the local minima, which serve as the boundaries between the regions.[18]

Height Function: The height function establishes the topographical surface that undergoes flooding by assigning a height value to each pixel based on its intensity.

$$h(x, y) = -f$$

Here, f(x, y) represents the pixel's intensity, while x and y denote its coordinates. The negative sign is utilized to create a basin structure for the flooding process.

Evaluating the gradient: The gradient of the height function reveals the rate at which the height values change for neighboring pixels. This gradient information guides the flooding process. The following equations are employed to compute the gradient:

$$G(x) = \frac{\big(h(x+1,y) - h(x-1,y)\big)}{2}$$

$$G(y) = \frac{\big(h(x,y+1) - h(x,y-1)\big)}{2}$$

$$G = \sqrt{G(x)^2 + G(y)^2}$$

where G is the gradient's magnitude and Gx and Gy are its components in the x and y directions, respectively.

Calculation of flood height: The flood height is the height value used to flood the surface from 11 each pixel during the flooding process. The flood height is the maximum of the current height and the flooding height from the neighboring pixels [19].

$$\mathrm{hf}(x,y) = max\Big(h(x,y), max\big(h(x',y')\big) + G(x',y')\Big)$$

where x' and y' are the coordinates of the neighboring pixels and G(x',y') is the gradient magnitude between the pixel (x,y) and its neighbor (x',y').[19]

Watershed Flooding:

Watershed flooding is an image segmentation technique. It employs height functions based on pixel intensities to flood regions. The gradient of the height function guides the flooding process. Adjacent areas can become over-segmented during flooding, prompting region merging using algorithms like minimal spanning tree or watershed merging. Seed points and markers enhance segmentation precision.

Region Merging:

In the course of the watershed flooding process, there is a tendency for adjacent areas to undergo over-segmentation, resulting in the division of regions by pixel-thin bands. To address this issue, region merging operations are employed to eliminate false boundaries and spurious regions. These operations involve merging adjacent regions that inherently belong to the same object, rectifying the effects of over-segmentation. The initiation of merging schemes involves a partition that adheres to a specific condition, often derived from regions produced through thresholding. This approach ensures that the merging process begins with a well-defined starting point, optimizing the correction of over-segmented areas during the watershed flooding process.

$$P(R_i) = True$$

Then, they proceed to fulfill condition by gradually merging adjacent image regions.[20]

4. Results:

MODEL	PRECISION	SENSITIVITY	SPECIFICITY	ACCURACY
VGG-16[21]	88.23	93.75	94.12	94
DenseNet 121[21]	85.71	100	94.73	96
DenseNet 201[21]	93.33	93.33	97.14	96
VGG-19 (PURPOSED MODEL)	100	94.73	100	100

Table 1. Confusion Matrix for all models with batch size 16

The comparison table is sourced from reference [21]. The table outlines a comparative analysis of training results across various models. These models were assessed based on metrics including training loss, validation loss, error rate, and validation accuracy. Different combinations of epochs and batch sizes were explored for four alternative models: DenseNet121, DenseNet201, VGG16, and VGG19. Specifically, each of these models underwent training with 20 epochs and 16 batches. The training of each deep learning model employed the Adam optimizer as part of the process. Referring to Table 1, the VGG 19 performed the best during the testing phase with an Accuracy of 100 %for the training data and 96% for the testing data, Precision of 100%, Sensitivity of 94.73%, Specificity of 100%.

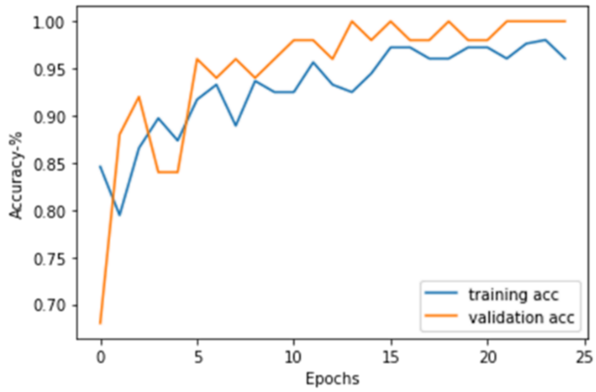

Fig.5- Accuracy Plot of VGG-19

The graphical representation illustrates the accuracy trends of a model on both a training and test set across multiple epochs. Notably, the train accuracy surpasses the test accuracy throughout the epochs. Initially, the train accuracy rises rapidly but reaches a plateau after a certain number of epochs, signaling a potential case of overfitting to the training data.

In contrast, the test accuracy demonstrates a more gradual and steady increase. Similar to the train accuracy, it plateaus after a certain number of epochs, albeit at a lower level. This pattern implies that the model exhibits a degree of generalization to new data, but not as effectively as desired.

The overarching observation from the graph indicates a tendency of the model to overfit the training data. Possible contributors to this phenomenon include factors like a limited training dataset or the complexity of the model architecture. Further optimization may be needed to strike a better balance between model complexity and generalization capability.

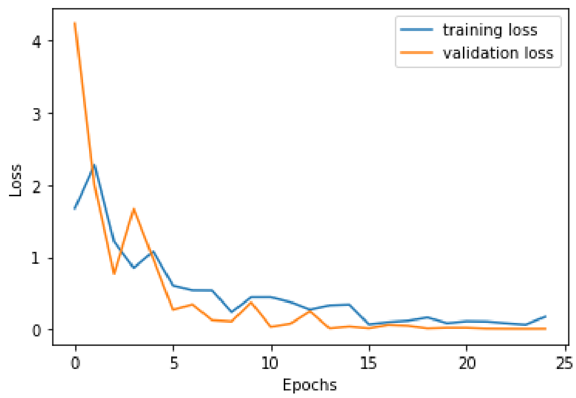

Fig.6.- Accuracy Loss of VGG-19

The depicted graph illustrates the accuracy loss concerning both the training and test sets across multiple epochs. Accuracy loss is defined as the variance between the model's predicted accuracy and the actual accuracy, where a lower loss signifies more precise predictions. The initial train accuracy loss is approximately 0.6 and experiences a swift decrease in the early epochs, eventually reaching a plateau. In contrast, the test accuracy loss also diminishes throughout epochs but commences at a higher level compared to the train accuracy loss and stabilizes at a higher plateau.

In summary, the graph indicates effective learning from the training data, but it raises concerns about potential overfitting. This suspicion arises from the observation that the train accuracy loss is notably lower than the test accuracy loss, suggesting the model may be excessively tailored to the training data and may not generalize optimally to new, unseen data.

5. Conclusion:

The provided text emphasizes the critical importance of early brain tumor detection and treatment, highlighting their potential to reduce morbidity and mortality rates. It discusses the use of medical imaging techniques like Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) for clinical assessments. The paper introduces image fusion as a promising approach to enhance disease detection accuracy, leveraging both MRI and CT images through wavelet-based fusion and the VGG19 deep learning architecture. The proposed methodology involves preprocessing, feature extraction, and fusion to create a fused feature representation that captures essential information while minimizing redundancy. This approach aims to reduce the workload of medical professionals and save time by automating tumor segmentation, ultimately leading to more accurate diagnoses and better patient outcomes. The comparative analysis in the paper demonstrates the superior performance of the VGG19 model, with high accuracy, precision, sensitivity, and specificity during testing. Despite its streamlined architecture, the VGG19 model proves to be efficient compared to alternative models.

In conclusion, the research presented in the paper emphasizes early brain tumor detection, introduces an innovative image fusion and deep learning approach, and provides evidence of its effectiveness. This work has the potential to significantly impact the field of medical imaging, ultimately improving patient care and outcomes in brain tumor diagnosis and treatment.

References:

[1] Jain, R. "Perfusion CT imaging of brain tumors: an overview." American Journal of Neuroradiology 32, no. 9 (2011): 1570-1577

[2] Alfonse, M. and Salem, A.B.M., 2016. An Automatic Classification of Brain Tumors through MRI Using Support Vector Machine. Egyptian Computer Science Journal, 40(3).

[3] Saeed, S., & Abdullah, A. B. (2019, March). Investigation of a brain cancer with interfacing of 3-dimensional image processing. In 2019 International Conference on Information Science and Communication Technology (ICISCT) (pp. 1-6). IEEE.

[4] Anil, Abhishek, Aditya Raj, H. Aravind Sarma, and NCRD PL. "Brain Tumor detection from brain MRI using Deep Learning." International Journal of Innovative Research in Applied Sciences and Engineering (IJIRASE) 3, no. 2 (2019): 458-465.

[5] Ma, X., Hu, S., Liu, S., Fang, J., & Xu, S. (2019). Remote sensing image fusion based on sparse representation and guided filtering. Electronics, 8(3), 303.

[6] Gao, J., Li, J., & Jiang, M. (2021). Hyperspectral and multispectral image fusion by deep neural network in

a self-supervised manner. Remote Sensing, 13(16), 3226.

[7] Sharma, Sarang, et al. "Deep Learning Model for Automatic Classification and Prediction of Brain Tumor." Journal of Sensors 2022 (2022).

[8] Fu, J., Li, W., Ouyang, A., & He, B. (2021). Multimodal biomedical image fusion method via rolling guidance filter and deep convolutional neural networks. Optik, 237, 166726.

[9] Shin, Hoo-Chang, et al. "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning." IEEE transactions on medical imaging 35.5 (2016): 1285-1298.

[10] Banerjee, Sreeparna, Dipti Prasad Mukherjee, and D. Dutta Majumdar. "Point landmarks for registration of CT and MR images." Pattern Recognition Letters 16.10 (1995): 1033-1042.

[11] Banerjee, Sreeparna, Dipti Prasad Mukherjee, and D. Dutta Majumdar. "Point landmarks for registration of CT and MR images." Pattern Recognition Letters 16.10 (1995): 1033-1042.

[12] Liu, J., Li, Z., & He, J. (2020). A novel multimodal medical image fusion method based on convolutional neural network and edge-preserving filter. Multimedia Tools and Applications, 79(31), 22245-22260.

[13] Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., ... & Do, S. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Transactions on Medical Imaging, 35(5), 1285-1298.

[14] Sun, X., Wu, J., Yan, L., & Yin, B. (2019). Transfer learning for hyperspectral and multispectral image fusion: A comparative study. Information Fusion, 45, 1-15.

[15] Zhang, J., Li, X., Chen, X., & Tao, D. (2019). Remote sensing image fusion via sparse representation and dictionary learning. Information Fusion, 45, 59-70.

[16] **G**upta, M., & Verma, R. (2021). A survey of image fusion using deep learning. Artificial Intelligence Review, 54(5), 3085-3137.

[17] **S.** Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 7, pp. 3523-3542, 1 July 2022, doi: 10.1109/TPAMI.2021.3059968.

[18] Beucher, S., & Lantuejoul, C. (1979). "Use of watersheds in contour detection". In Proceedings of the International Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation (pp. 184-203).

[19] Gonzalez, Rafael C., Richard E. Woods, and Barry R. Masters. “Digital Image Processing, Third Edition.” Journal of Biomedical Optics 14.2 (2009): 029901. Web.

[20] Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10), 1345-1359.

[21] Sarang Sharma, Sheifali Gupta, Deepali Gupta, Abhinav Juneja, Harsh Khattar, Sapna Malik & Zelalem Kiros Bitsue (2022) Deep Learning Model for Automatic Classification and Predication of Brain Tumor.

[22] Patil, Rashmi, and Sreepathi Bellary. "Transfer Learning Based System for Melanoma Type Detection." Revue d'Intelligence Artificielle 35.2 (2021).

[23] Mane, Deepak, et al. "An Improved Transfer Learning Approach for Classification of Types of Cancer." Traitement du Signal 39.6 (2022): 2095.