Original Article

A Review on Multimodal Medical Image Fusion

Sreelekshmi A N[1], Dr. N. Sujatha[2]

[1]Research Scholar, P.G & Research Department of Computer Science, Sri Meenakshi Govt. Arts College for Women (Autonomous), Madurai Kamaraj University, Madurai, Tamilnadu, India.

[2]Assistant Professor, P.G & Research Department of Computer Science, Sri Meenakshi Govt. Arts College for Women (Autonomous), Madurai Kamaraj University, Madurai, Tamilnadu, India.

Email: sree.an1989@gmail.com

DOI: 10.47750/pnr.2023.14.03.257

Abstract

Multimodal medical image fusion is a current technique applied in applications related to medical field to combine images from the same modality or different modalities to improve the visual content of the image to perform further operations like image segmentation. Biomedical research and medical image analysis highly demand medical image fusion to perform higher level of medical analysis. Multimodal medical fusion assists medical practitioners to visualize the internal organs and tissues. Multimodal medical fusion of brain image helps to medical practitioners to simultaneously visualize hard portion like skull and soft portions like tissue.

Keywords: Multimodal medical image fusion, Image segmentation, Computed tomography.

1. Introduction

Images are the largest source of data in healthcare and, at the same time, one of the most difficult sources to analyze. Clinicians today must rely largely on medical image analysis performed by overworked radiologists and sometimes analyzes can themselves. Computer vision software based on the latest deep learning algorithms is already enabling automated analysis to provide accurate results that are delivered immeasurably faster than the manual process can achieve. Multimodal medical imaging can provide us with separate yet complementary structure and function information of a patient study and hence has transformed the way we study living bodies. The motivation for multimodal imaging is to obtain a superior exquisite image that will provide accurate and reliable statistics than any single image while retaining the best functions for the snapshots software program for medically testing, diagnosing and curing diseases.

Diagnostic tools consist of Computed tomography (CT) and Magnetic resonance imaging (MRI) and thus these are the two modalities that we will consider for Image Fusion Process.

We aim to approach a three-step process:

- Image Registration
- Image Fusion
- Image Segmentation
- Image Registration

Image registration is the process of transforming images into a common coordinate system so corresponding pixels represent homologous biological points. Registration can be used to obtain an anatomically normalized reference frame in which brain regions from different patients can be compared.

Landmark-Based Registration

Image landmark registration is a simple process where a number of points (landmarks) are defined in the same locations in two volumes. The landmarks are then matched by an algorithm, and the volumes are thus registered. The CT scan image is taken as the reference (fixed) image and the MRI scan image is aligned as per the points selected by the user.

Transfer Learning

Transfer learning is an optimization that allows rapid progress or improved performance when modeling the second task. We aim to use the VGG-19 CNN architecture with its pre-trained parameters which would help us to achieve our target. Visual Geometry Group (VGG-19) is a convolutional neural network that is trained on more than a million images from the Image Net database. The network is 19 layers deep and can classify images into 1000 object categories.

We convert our images to YCbCr color format because it preserves detailed information of luminance component.

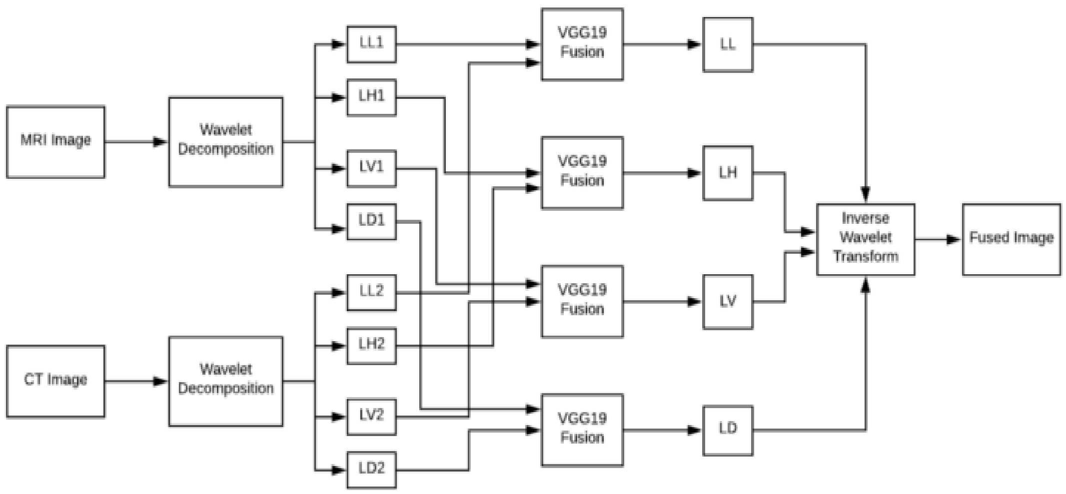
Discrete Wavelet Transform

Wavelet transform provides high-frequency resolution at low frequencies and high timer solution at high frequencies. A discrete wavelet transform (DWT) is a wavelet transform for which the wavelets are discretely sampled. It captures both frequency and location information (location in time).

Watershed Algorithm

Watershed segmentation is a region-based technique that utilizes image morphology. It requires selection of at least one marker ("seed" point) interior to each object of the image, including the background as a separate object. The markers are chosen by an operator or are provided by an automatic procedure that takes into account the application-specific knowledge of the objects. Once the objects are marked, they can be grown using a morphological watershed transformation.

Image Fusion Architecture:



2. Proposed System

Brain tumour is the most destructive disease which leads to a very short life expectancy in the highest grade. Manual brain tumour diagnosis with a single image is less accurate yet still a time-consuming procedure. So, the Multi-modal Fusion technique is adopted which merges two or more images to obtain a highly informative image with increased accuracy. In this paper, it is proposed to combine computed tomography (CT) and magnetic resonance imaging (MRI) images in wavelet domain. The proposed method uses deep learning to fuse the images through the feature maps and obtain its features for further classification. Experimental results show that, the tumour can be detected accurately which will be useful for diagnosis. Experimental results give fused multimodal medical image so high quality with better statistical assessments as compared to existing methods.

2.1. Pre-processing in image fusion

Image Fusion starts with capturing of image with the help of multi-sensors and then the pre-processing takes place. The work like adjustment of brightness and stretching of contrast is taken care of during pre-processing. This is done because two different images that have been taken at different angles may cause distortion.

2.1.1. Image Registration

The initial phase in the method involving the fusion of images is registering the input images. Thus, image registration can be characterized as the procedure for mapping the inputs with the assistance of referenced image. Image registration is purposefully done to adjust images regarding each other. This procedure requires two images as input: the first image is called the referenced (control) image, while the second image is called sensed image. The sensed image is aligned with the referenced image during image fusion. Henceforth, the procedure of image registration involves setting up of point-by-point agreement be twixt various images, portraying a similar section. Based on this, there are two sorts of image registrations utilized in the process of image fusion, uni-modal and multimodal techniques. Uni-modal strategies enlist images in a similar methodology procured by a similar device, while multimodality registration techniques enlist images gained by various devices (multimodal registration). Contingent upon the sorts of images, the registration can be delivered with either global transforms, for example, scaling, translation, non-rigid, rotation, and affine transforms of which it is fit for local warps.

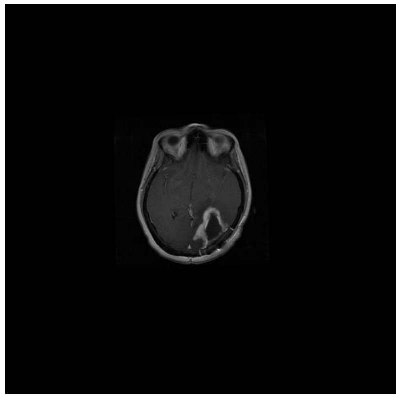


Figure 2.1. Registered MRI image

2.1.2. Multimodal image fusion process in medical domain

The process of image fusion particularly multimodal medical image fusion in tents to ameliorate the quality of image by reduction of redundant data to increment the pathological pertinence of the images in medical classifications and diagnosis [23]. The algorithms for fusion are reliant on inputs. Hence, building a fusion algorithm depends on 3 main factors: imaging modality, organ under study, and fusion algorithm employed (see Figure 2.2). Figure 2.3 represents the flow diagram of the fusion process for medical MRI and CT images.

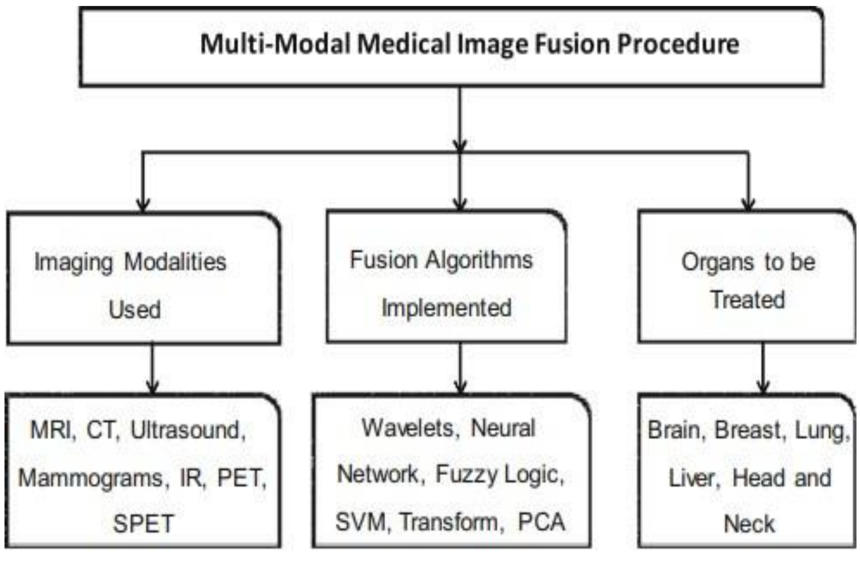


Figure 2.2. Multimodal image fusion in the medical domain

2.2. Image fusion techniques

Broadly there are two main categories of image fusion. The first one is spatial fusion and second is multi-resolution fusion. In the case of spatial fusion techniques, one can derive the highly focused and fused images with less effort and at the same time, also derives more spatial information about it. The issue, in this case, uses to be that the images in most of the cases are blur. In case of multi-resolution fusion methods, one can derive the enhanced spectral information along with higher signal-to-noise ratio (SNR). But the issue, in this case, is that this method is relatively complex and at the same spatial resolution received here is lower. In such cases here in both methods that have their own merits and demerits, one should adopt the technique as per their need. Figure 2.3 shows the different types of image fusion techniques that are used for fusion of images.

2.3. Wavelet transform fusion

The fusion techniques involve method as simple as pixel-based averaging or more complex techniques like ICA (independent component analysis), wavelet transformation fusion, and PCA. Furthermore, wavelet transformation–based image fusion is the most general and frequently used method that is utilized for image registration and fusion. Extraction of detailed and pertinent information from one image and injecting it into another image is the main concept behind wavelet-based image fusion (see Figure 2.4).

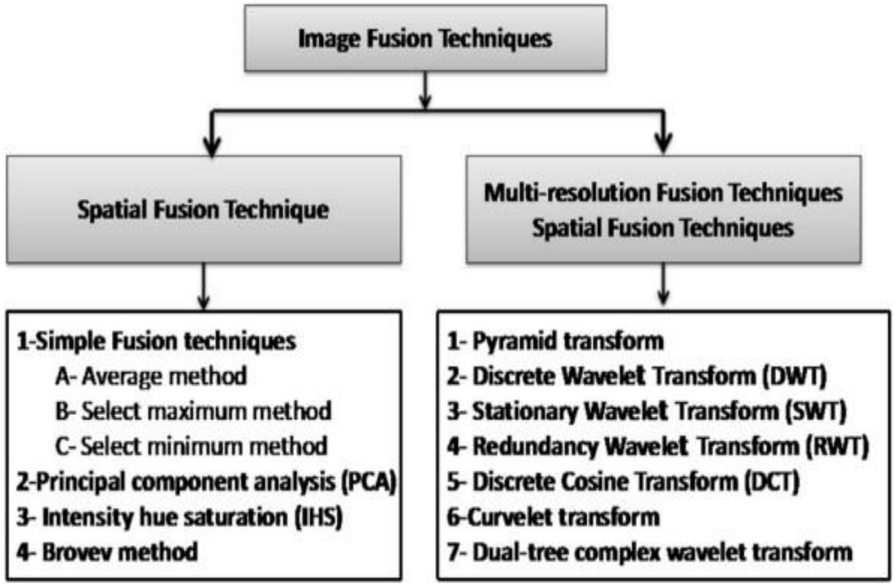


Figure 2.3. Various methods for image fusion

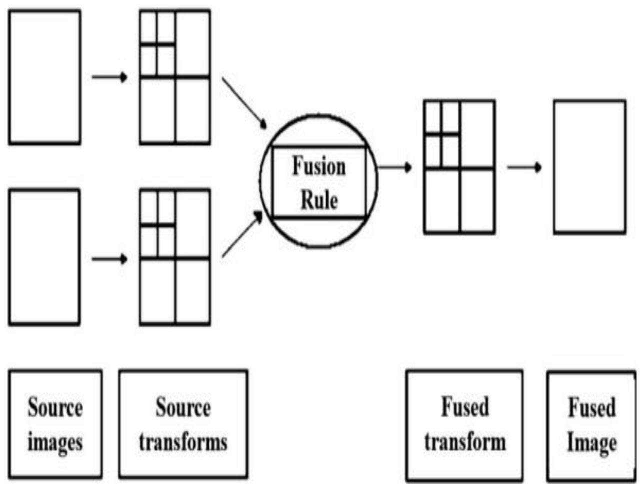


Figure 2.4. Fusion of input images based on wavelet transform (fusion process)

Wavelets can select and manipulate the frequencies in time and space both, therefore, extract the quality of detailed information that is stored in the high frequency in images. Wavelets have many known applications in medical image fusion such as image pseudo coloring, medical diagnosis, feature-level image fusion, medical segmentation, super-resolution, lifting scheme, 3D conformal radiation therapy (3D-CRT) treatment planning, and color visualization.

The neural network is the outstanding and most common approach of wavelet image fusion in which the wavelet plays the role of a fusion operator whereas feature processing is often done by the neural network. Such like the neural network, feature-level image fusions can be achieved by wavelets using kernel-based operators, e.g., SVM or support vector machines.

In the wavelet transform technique, the selected mother wavelet signals go through various decomposition, shifts (translations), and scaling (expansion and dilation) on application of wavelet transform. Other child window functions are derived from the primary wavelet functions, thus, the term mother wavelet. Discrete and continuous domains are the two domains in which wavelet transform is applied. Figure 2.5 shows concept of a discrete wavelet transform (DWT).

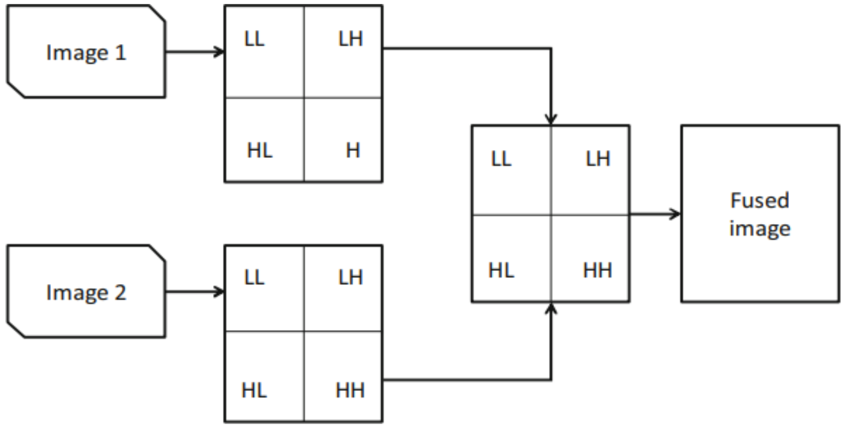


Figure 2.5. Schema of discrete wavelet transform

A continuous wavelet transform (CWT) is measured by altering the scale of the analysis window each time, shifting it or multiplying it by the signal. Wavelet Transform is represented by the connection between the signal and wavelet at different scales (frequency inverse) and the signal.

Smoothness and orthogonal as well as other desired properties are carried by the wavelet transform. Therefore, wavelet transform has wide applications in medical image fusion. Consistently, due to its good characteristic features and frequency responses, directionality, layered structures, and dimensionality coaxingly coincide with the human visuals; it is a critical method for image fusion.

In a paper, Singh et al. proposed pseudo coloring for medical images to observe the wavelet efficiency in separating an image in various sub-spaces preserving the features from the original images, using pseudo coloring for wavelet image fusion. The proposed method increased and improved the processing speed of the pseudo-coloring algorithm. Separation of appropriate tissues by the medical professionals is now possible by adopting this method for image fusion.

Medical image fusion technique of multimodal PET and MRI brain images based on the wavelet transform that has been studied by Yang et al. Following wavelet transform and gray-level fusion, the results of fusion are obtained by modifying the gray matter (GM) morphological and structural details and spectral statistics in a brain's white matter (WM) field normally gathering binary information from the neuroimaging of a single individual for diagnostic imaging studies. Calhoun and Adali, various researchers jointly proposed an effective and simple blind source separation technique that advantageously joined ICA (independent component analysis) and CCA ("canonical correlation analysis"), for the multitask fusion of data. It offers the proper links and strong precision in determining the two datasets in which the input can have characteristics or similar "properties" associations among datasets. Due to its minimal size, the task-related operation defined by a filtered database feature is essentially more easily manipulated than a four-dimension alone.

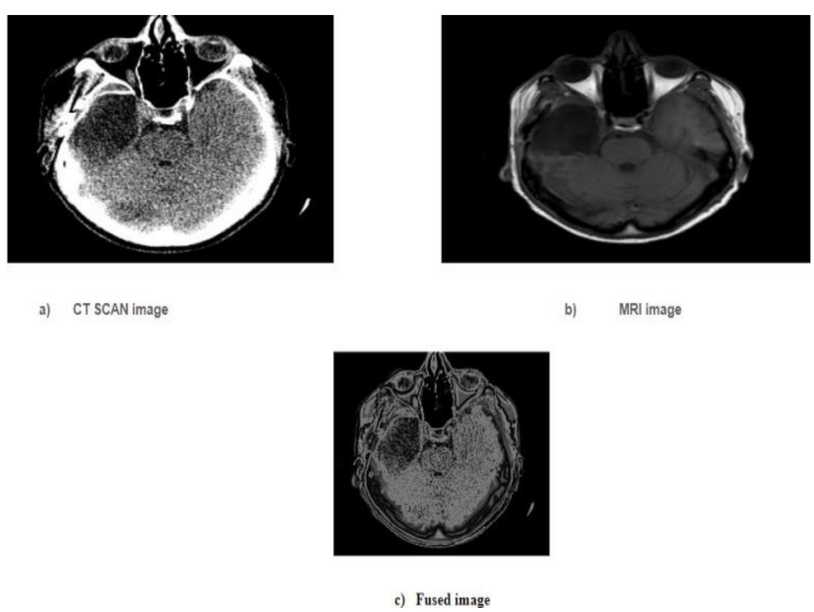


Figure 2.6. Multi-modal image fusion of brain CT and MRI images. a) CT image, b) MRI image, c) fused image

2.4. Image Segmentation

Image segmentation is a commonly used technique in digital image processing and analysis to partition an image into multiple parts or regions, often based on the characteristics of the pixels in the image. Image segmentation could involve separating foreground from background, or clustering regions of pixels based on similarities in color or shape. Image segmentation is a method in which a digital image is broken down into various subgroups called Image segments which helps in reducing the complexity of the image to make further processing or analysis of the image simpler. Segmentation in easy words is assigning labels to pixels. All picture elements or pixels belonging to the same category have a common label assigned to them. For example: Let’s take a problem where the picture has to be provided as input for object detection. Rather than processing the whole image, the detector can be inputted with a region selected by a segmentation algorithm. This will prevent the detector from processing the whole image thereby reducing inference time. A common application of image segmentation in medical imaging is to detect and label pixels in animageorvoxelsofa3Dvolume that represent at patient’s brain or other organs.

Several algorithms and techniques for image segmentation have been developed over the years using domain-specific knowledge to effectively solve segmentation problems in that specific application area. These applications include medical imaging, automated driving, video surveillance, and machine vision.

Medical Imaging

During medical diagnosis for cancer, pathologists stain body tissue with hematoxylin and eosin (HE) to distinguish between tissue types. They then use an image segmentation technique called clustering to identify those tissue types in their images. Clustering is a method to separate groups of objects in a scene. The K-means clustering algorithm finds separations such that objects within each Cluster areas close to each other as possible, and as far from other objects in other clusters as possible.

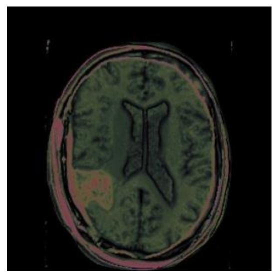


Figure 2.7. Segmented image

3. WATER SHED ALGORITHM

Watershed algorithm is based on extracting sure background and foreground and then using markers will make watershed run and detect the exact boundaries. This algorithm generally helps in detecting touching and overlapping objects in image. Watershed segmentation is a region-based technique that utilizes image morphology. It requires selection of at least one marker (“seed” point) interior to each object of the image, including the background as a separate object. The markers are chosen by an operator or are provided by an automatic procedure that takes into account the application-specific knowledge of the objects. Once the objects are marked, they can be grown using a morphological watershed transformation. To understand the watershed, one can think of an image as a surface where the bright pixels represent mountain tops and the dark pixels valleys. The surface is punctured in some of the valleys, and then slowly submerged into a water bath. The water will pour in each puncture and start to fill the valleys. However, the water from different punctures is not allowed to mix, and therefore the dams need to be built at the points of first contact. These dams are the boundaries of the water basins, and also the boundaries of image objects.

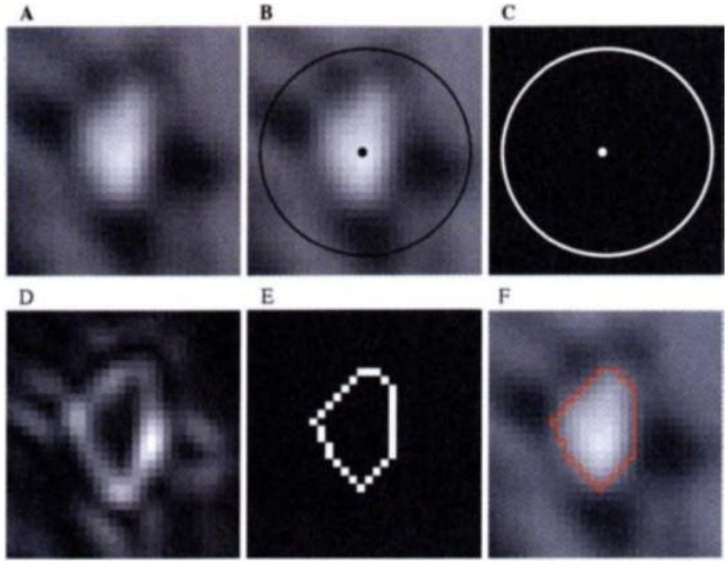


Figure 3.1. System Architecture

An application of watershed segmentation to extract lymph nodes on CT images is shown in Figure 3.1. In this implementation a 3×3 Sobel edge operator is used in place of the morphological gradient to extract edge strength. The originally lymph node image is shown in Figure 3.1A. In the first step, the operator positions a cursor inside the node (Figure 3.1B). All pixels within a radius of two pixels of the mark are used as seed points for the lymph node. To mark the exterior of lymph node, the operator drags the cursor outside of the node to define a circular region, which completely encloses the node (Figure 3.1C). All pixels outside this circle mark the background.

In the next step, an edge image is created using the Sobel edge operator (Figure 3.1D). The edge image has high values for the pixels with strong edges. With the seed point marking the node interior, the circle marking the background (Figure 3.1C), and the edge image generated by the Sobel operator (Figure 3.1D), the segmentation proceeds directly with the watershed operation (Figure 3.1E). The watershed operation operates on an edge image to separate the lymph node from the surrounding tissue. By using a technique called simulated immersion, the watershed considers whether a drop of water at each point in the edge image would flow to the interior seed point or the exterior marker. Points that drain into the interior belong to the lymph node, whereas points that drain to the exterior belong to the surrounding tissue.

Watershed analysis has proven to be a powerful tool for many 2-D image-segmentation applications). Higgins and Ojard applied a 3D extension of the Watershed algorithm to cardiac volumetric images.

3.1. VGG19

VGG-19 is a convolutional neural network that is 19 layers deep. You can load a pre-trained version of the network trained on more than a million images from the Image Net database. The pre-trained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by-224.

The main purpose for which the VGG net was designed was to win the ILSVRC but it has been used in many other ways.

VGG19 is used just as a good classification architecture for many other data sets and as the authors made the models available to the public they can be used as is or with modification for other similar tasks also. Transfer learning can be used for facial recognition tasks also. Weights are easily available with other frameworks like keras so they can be tinkered with and used for as one wants.

4. RESULT ANALYSIS

There are various steps that are used to analyses the result. The first steps of results is reading the input data's and perform the training and the testing of data.

4.1. Reading Data

Here we are going to open the data set with pandas, check distribution of labels, and over sample to reduce imbalance.

Figure 4.1. Reading data

4.2. Augmentation

Having sufficient training data is the key for training a neural network successfully; unfortunately, this requirement is seldom satisfied in most neural network applications. For medical imaging applications, the lack of data is more significant because of the cost of the annotations, and because of the imbalance in the occurrence between diseases. To mitigate shortages in data and fully utilize the data that are available, certain data augmentation techniques must be carried out in our experiment, we used the Augmentor software package. Specifically, we augmented our data through the following means:

- Flip the image horizontally
- Flip the image vertically
- Randomly rotate the image in the range of [25,25] degrees
- Randomly zoom in or out in the range of [0.85,1.15]
- Randomly distort the image

All of these methods were combined for augmenting each image, and a probability of 0.5 was used to determine whether or not to perform each of them.

4.3. Training

The process of modeling means training a machine learning algorithm to predict the labels from the features and validating it on holdout data. We can now train our model by using transfer learning. We load convolution neural networks from Fast AI. We use ‘fit one cycle function to train our model 14 epochs. The model works and interprets our model by using Classification Interpretation.

4.4. Validation

Validation is referred to as the process where a trained model is evaluated with a testing data set. The testing data set is a separate portion of the same data set from which the training set is derived. Model validation is carried out after model training. After validation, model will check the classes and ensure the trained data and its path is correct.

4.5. Testing

Testing is the process of evaluating a system or its component(s) with the intent to find whether it satisfies the specified requirements or not. In simple words, testing is executing a system in order to identify any gaps, errors, or missing requirements in contrary to the actual requirements.

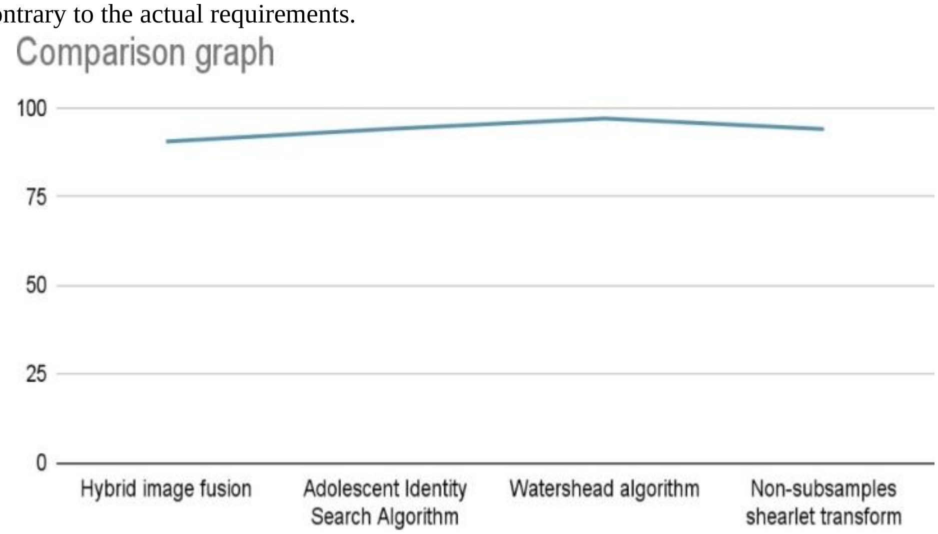


Figure 4.2. Graph of learning

4.6. Accuracy

Accuracy is one metric for evaluating classification models. Informally, accuracy is the fraction of predictions our model got right. Formally, accuracy has the following definition: Accuracy = Number of correct predictions by Total number of predictions. On testing different data in the given watershead algorithm we obtain a result of the image.

5. Conclusion

In the past few years, image fusion has been immensely used for several image-processing applications in various fields, particularly medical applications like retinopathy and brain tumor segmentation. The work presented a novel approach in performing medical image fusion using generative adversarial networks. The proposed approach efficiently captured the functional information from CT images and anatomical structure from MRI images and transformed them to the resulting fused image. The proposed approach generates fused images with less distortion and better structural information when compared to the existing approaches. The advantages of the proposed approach over the other existing approaches are that it can retain the textural information from the MRI image and the metabolic information from the PET image without losing pixel intensity.

References

[1] Multimodal Medical Image Fusion of Positron Emission Tomography and Magnetic Resonance Imaging Using Generative Adversarial Networks-R.Nandhini Abirami, P.M.Durai Raj Vincent, Kathiravan Srinivasan, K.Suresh Manic, and Chuan-Yu Chang 14 Apr 2022

[2] Detection of Brain Tumour by Image Fusion based on Convolutional Neural Network-Aishwarya N, Praveena NG, Rajalakshmi B, Reshma R, Rizwanabegam B, Sowmya D Nov 06, 2020

[3] Understanding Image Fusion-Yun Zhang

[4] Medical Image Segmentation: A Review- Dinesh D. Patil, Sonal G. Deore, 2013.

[5] Brain tumors–Lisa M. De Angelis Jan11, 2001.

[6] Fusion of CT and MRI for the detection of brain tumor by SWT and Probabilistic Neural Network- D.Arulkumar, V.Aswathi, R.Jenifer, R.S.Bhavani Mar 2020.

[7] Automated Multimodal image fusion for brain tumor detection-Harpreet Kaur, Deepika Koundal, Virendar Kadyan, Navneet Kaur, and Kemal Polat 2021.