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

A Survey on Brain Tumour Segmentation Techniques in Deep Learning

[1]Bharathi Panduri, [2]Dr. O. Srinivasa Rao

Submitted: 03/10/2023 **Revised**: 20/11/2023 **Accepted**: 01/12/2023

Abstract: Machine learning has recently been used in hospitals to speed up the diagnosis and analysis process. To speed up the beginning of the recovery process, doctors can now get help with diagnosis. The future of AI in health care may involve tasks ranging from simple to complex, including answering the phone, assessing medical records, primary care trending and analytics, therapeutic medicine and computer design, reading radiology images, constructing medical treatment and diagnosis plans, and even conversing with patients. Deep learning models can interpret medical images like X-rays, MRI scans, CT scans, etc. to establish a diagnosis. The algorithms can spot risks and detect inconsistencies in the medical images. Deep learning is often used for cancer detection. MRI scans are needed to properly segment brain tumours, which can help with clinical diagnosis and treatment planning. However, medical practice is particularly difficult because some testing methods are not included in MRI pictures. When comparing the quantitative and qualitative findings of medical image analysis as it is now practised, the suggested approach performs better. The Chest CT images are superior at precisely identifying malignant lung nodules in the case of lung cancer detection. For patients' prospects of survival, early identification of lung cancer is essential. To identify between malignant and benign nodules, develop a multi-view knowledge-based collaborative (MV-KBC) deep model using sparse chest computed tomography (CT) data from previous study work. But the MV-KBC model was more precise. The model, however, is only usable with supervised image data. In order to address the model's limitation, we provide a unique deep learning-based multi view model in our research. For semi-supervised medical picture applications, the proposed model's accuracy was greatly increased, and calculation and classification times were reduced.

Keywords*: classification, MV-KBC, inconsistencies, computed tomography (CT), algorithms*

1. Introduction

Medical image processing entails the use of 3D images of the human body, often obtained from a CT or MRI [1] scanner. The doctors are investigating with the medical report and finding the illnesses of the patient. So then they can diagnose the medical procedures or any surgery. Medical image analysis applications have experienced a significant increase in interest in deep learning methods over the past ten years. A particularly promising field is medical imaging. Implementing medical image analysis using deep learning is a novel innovation [2]. The use of deep learning to efficiently analyse medical picture data. Many techniques in deep learning can be used to analyse medical picture data. For researching the analysis of medical images, deep learning used classes such as categorization, segmentation, and detection.

In medical image analysis, picture reconstruction, image collection, and cascade activities are generally thought of as three discrete activities. A cascade of mistakes from one task to the next can be introduced by applying medical image analysis algorithms [3] to the original data in a serial fashion, especially when the non-organized data.Despite the fact that medical images in clinical decision-making is growing, the assurance of image quality is a step that is frequently skipped in automated image analysis pipelines. This is a crucial stage since good quality medical images are essential for downstream operations like segmentation to be accurate [4].

Digital photography is a must-have for everyday use. Medical imaging processing is the term used to describe the handling of pictures by computers [5]. The methods and processes used in this processing span a wide spectrum, including image capturing, preservation and display. A function that displays an assessment of a sight's properties, like illumination or colour. Digital images provide a number of benefits, including quick and affordable processing, straightforward transfer and storage, instant evaluations, the ability to make numerous copies without compromising quality, quick and affordable reproduction, and flexible transformation. The image processing need for huge speed and database to store. Sometimes the digital photographs storage is violating copyright while keeping quality [6].

Research involving a "3D Medical Multi-modal Segmentation Network" has notable limitations that should be acknowledged. Firstly, the effectiveness of these networks heavily relies on the availability and diversity of annotated medical imaging data. Limited access to diverse and high-quality datasets from various

[1]*Research Scholar,Department of Computer Science & Engineering, JNTUK, Kakinada.*
Email: bharathiraaz@gmail.com
[2]*Professor & Head, Department of Computer Science & Engineering, JNTUK, Kakinada.*
Email: osr_phd@yahoo.com

sources and patient populations can hinder the network's generalization capability. Moreover, the accuracy of segmentation heavily depends on the quality and consistency of manual annotations, introducing the potential for annotation variability and inaccuracies, particularly for complex structures. The inherent complexity of deep learning models can demand significant computational resources, potentially limiting their accessibility and applicability in resource-constrained environments. Interpretability remains a concern, as these models often lack transparency in decision-making, raising challenges in critical medical applications where understanding the basis of segmentation decisions is crucial. The model's ability to generalize to unseen data or new imaging domains and the risk of overfitting further underscore the complexity of achieving robust and reliable segmentation. Additionally, the translation of research findings into clinical practice requires thorough clinical validation and addressing regulatory and ethical considerations, making the real-world deployment of such models a multifaceted challenge. Finally, as the field of deep learning evolves rapidly, staying up-to-date with emerging architectures and techniques becomes essential to maintain research relevance and practicality.

Digital image editing on a computer is referred to as image processing. Numerous benefits exist for this approach, including adaptability, flexibility, data storage, and communication. Effective photo maintenance has been made feasible by the invention of several image scaling algorithms. With this approach, the photos must be subjected to multiple sets of criteria at once. Both 2D and 3D images can accommodate many dimensions [7]. In the 1960s, the first image processing methods were developed. These methods were applied in a variety of sectors, including space exploration, medicine, the arts, and TV image enhancement. With the advent of computers in the 1970s, image processing became more affordable and quick. In the 2000s, image processing improved in speed, cost, and simplicity [8].

Several deep learning techniques have shown significant promise in enhancing the efficiency and effectiveness of medical image categorization, segmentation, and detection. These techniques leverage advancements in neural network architectures, data augmentation, transfer learning, and attention mechanisms.Brain tumor segmentation using deep learning techniques involves the process of automatically identifying and delineating the boundaries of different regions within brain images, specifically focusing on areas affected by tumors. This is a critical task in medical image analysis, as accurate segmentation aids in diagnosis, treatment planning, and monitoring of brain tumors. Deep learning techniques have shown significant promise in this domain due to their ability to automatically learn relevant features from data without the need for explicit feature engineering. Training deep learning models to analyze diverse medical image modalities such as X-rays, MRI scans, and CT scans presents several unique challenges and considerations due to the differences in imaging characteristics, data acquisition, and clinical context.

Lung lesions are frequently identified and confirmed with multidetector CT (MDCT) [9]. It is regarded as the gold standard for finding these lesions and is a particularly sensitive approach for finding lung nodules. Between 85% and 95% of nodules measuring 5 to 11 mm can be detected by MRI [10]. Even though MDCT will show the size as 1 or 2 mm, depending on the lung cancer risk classification, a prompt is advised for lesions size more than 7- or 8-mm. Lesions less than 7 mm in diameter should be monitored to determine their development pattern [11]. According to Koyama et al. [16], thin-section MDCT cannot detect malignant nodules as effectively as non-contrast enhanced pulmonary MRI (Fig.1). The rates for malignant nodules ($p>0.05$) did not substantially differ from the total detection rate. It will be giving value lower than the MDCT which is (97.0%, p0.05).

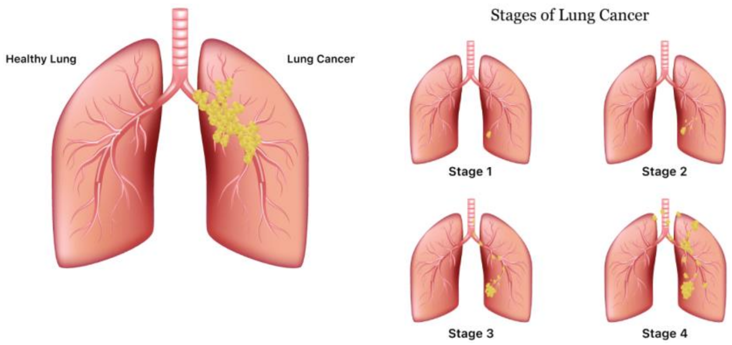

Fig 1. Normal and Affected patient's images

The main objective of the "MV-KBC deep model" is to accurately distinguish between malignant and benign nodules in chest CT scans. CT scans are three-dimensional images that can provide detailed insights into structures within the body, including the lungs. Identifying whether a nodule is malignant (cancerous) or benign (non-cancerous) is a critical task in diagnosing lung diseases, especially lung cancer.

The "MV-KBC deep model" likely combines various viewpoints or representations of CT scan data and incorporates existing medical knowledge to make more informed decisions about the nature of nodules. This could involve extracting features from different image slices, leveraging anatomical information, and integrating clinical guidelines to enhance the model's accuracy in nodule classification.Semi-supervised learning is a powerful approach that combines both labeled and unlabeled data to improve the accuracy and efficiency of medical image analysis. Leveraging unlabeled data can be particularly beneficial in scenarios were obtaining labeled data is challenging and costly.

To choose the right course of treatment, mediastinal lymph nodes must be accurately assessed. N1 illness is indicated by nodes present in the ipsilateral per bronchial area or hilum; treatment choices are unaffected by this. N2 illness is characterised by ipsilateral mediastinal or sub carinal lymphadenopathy, which may be respectable if only one station is affected. The categorization of the stages in between N1 to N3 showing the severity of the disease. N3 shows the crucial stage in the disease. N3 illness is characterised by hilar lymph nodes, scalene, or Virchow's nodesand is contraindicated for aggressive surgery [12].

Although there are scanners and images that are integrated MRI and PET, the technique is the same which are used in research tool for some time to come [13]. However, important developments have allowed for the exclusive use of Diagnostic imaging for physiological and structural imaging. Compared to dynamic MDCT, the operational information offered by MRI is far from an additional dimension. MRI is acknowledged comparable to delivering operational and molecular analysis, although not measuring glucose metabolism [14].

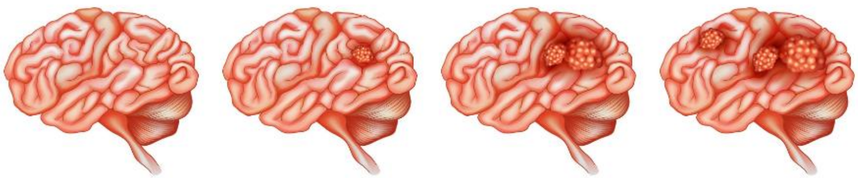

Fig.2 Different stages of Brain tumour

Fig 2 depicts the tumour improvement from stage by stage. There are 4 to 5 stages to determine the severity of the cancer. The machine learning techniques are used to find the tumours early stage which can helpful to the patients.A tumour develops more quickly the more aggressive it is. Typically, it can take months or even years for a brain tumour to grow. The most prevalent and dangerous type of brain cancer is glioblastoma. They are one of our main examples since the immune system cannot detect their growth. It is increasing tiredness and sluggishness and constant need for sleep, often spending the majority of the day in bed or relaxing. Reduction of muscular mass and weight loss. Having little to no appetite and having trouble swallowing food or liquids.

Artificial Intelligence (AI) has revolutionized the healthcare industry by driving advancements across various domains. In medical imaging, AI algorithms analyze complex images like X-rays, CT scans, and MRIs, aiding in the detection of diseases such as cancer and fractures. AI's capability to sift through vast patient data enables early disease diagnosis and personalized treatment planning. Moreover, AI accelerates drug discovery, predicts disease outbreaks, and streamlines administrative tasks through automation. Remote patient monitoring and telemedicine powered by AI facilitate virtual consultations and data collection. Natural Language Processing (NLP) enhances electronic health records and communication, while genomic analysis assists in understanding genetic contributions to health conditions. As AI continues to transform healthcare operations, its integration requires careful attention to ethical, privacy, and regulatory considerations, underscoring the collaborative nature between AI experts and medical professionals. Establishing guidelines and standards for the responsible and equitable deployment of AI in medical image analysis is crucial to ensure patient safety, ethical considerations, and equitable access to healthcare advancements.

Diffusion-weighted imaging (DWI) MRI signals are produced by the molecules of water that move in the internal molecules of the water, and they aid in the detection of neoplastic tumours [15]. Recent studies discovered that lung cancer may be detected by DWI and could be distinguished from post-obstructive lobar collapse results collected by DWI [17]. With accuracy comparable to PET-CT, whole-body MRI with DWI can be used to diagnose lung cancer patient's' M-stages, and quantitative DWI analysis allows for the distinction of lymph glands containing and without lesions [18].

2. Literature Survey

In the modern world, brain tumours are more harmful than cancer. Early discovery is crucial for clinical assessment and planning of the treatment. The best imaging method for brain tumours is an MRI scan, which delivers an excellent soft tissue contrast without using radiation. Traditional brain tumour segmentation techniques have a lot of success in terms of identification. The machine learning functions and methods are used for this research. However, segmenting brain tumours remains difficult, particularly when several modalities are absent. There are three different sorts of missing modalities. The first kind involves patient-to-patient variations in brain anatomy [19]. The second issue is the variation in glioma sizes, forms, and textures from one patient to the next. The ability of low contrast, low intensity MR imaging methods to vary.

The fusion technique plays a vital role in achieving an appropriate segmentation result for problems involving multi-modal clinical feature extraction. Brain tumour segmentation on MRI with missing nodes, H. Fu et al., [20]. With the help of latent representation learning and robust segmentation, this approach overcame the difficulty of segmentation on missing modalities.

However, the approach solely examined the single-modal image when evaluating the attention mechanism. However, effective feature fusion is impossible to perform with multi-source segmentation modalities [21]. Identification of anatomical landmarks and reference points was a primary requirement for the medical picture analysis. Different techniques are used in the procedure to diagnose patients and organise their care. These techniques include image registration, segment method initialization, and clinical measurement computation. Although manually identifying anatomical positions may seem simple, it is frequently time-consuming and difficult. Manual identification can be replaced by quick and exact automatic landmark localization techniques, which may be especially useful when correct localization of numerous image positions is required. Many segmentation techniques have emerged recently that use latent extraction of feature to account for missing nodes. What sort of latent traits must be learnt, and how they should be learned, are the issues. The proposed HeMIS [22] computes the consecutive features across various modalities to predict the final segmentation after learning the visual features of each modality separately utilising the most recent cutting-edge framework from Havaei. However, learning the common latent representation cannot be accomplished by calculating the mean and variance of each separate representation.

X. Xu and P. Prasanna [23] created a basic representation which gives the input modality into identify the mean function for segmentation than latent representation. Chen et al. split the input modalities into content code and presentation code. Utilising feature disentanglement in [24]. The content code was then fused into a common representation for segmentation using a gating method. Despite the fact that the method requires two encoders for each of the multimodality, it takes more time and is more difficult.

Accurate and speedy anatomical structure detection is required for patient’s treatment in terms of medical analysis and diagnostics.To understand how the recorded structure looks, current systems for anatomy detection frequently rely on machine learning algorithms that take advantage of enormous annotated image databases [25]. The employment of technologically inefficient search-schemes for human biology detection and the use of subpar feature engineering approaches are two of these systems' most significant flaws. J. Liu[26] provided the method for identifying the behaviour of the disease with the help of artificial intelligent concepts.In this category, deep learning concepts are also helping to get the images as an input to predict the anatomy of the human body. It will create a model to characterize the multi scale image analysis.And, the huge volume of morphological object in the scanned image is trained to differentiate the target anatomical object in our body. Despite the fact that state-of-the-art solutions are greatly outperformed by the Multi-Scale Deep Reinforcement Learning approach [27]. However, improving accuracy performance is crucial.

We can choose the features from to combine features from various heterogeneous neural networksand the collected data will be classify using multi-view classifiers. Then the model will calculate their correlation to improve the performance for sentiment analysis. The experimented deep network performs categorization using the intermediate properties of convolutional neural network and recursive neural network. According to the experiment, the single view pattern gave thebetter effectiveness and performance to get the focuson the cancer disease.

The suggested model, based on deep learning, demonstrated tremendous promise for enhancing the accuracy of image detection and minimize the radiologist’spain to initial screening of COVID-19 pneumonia. It was trained using multi-view pictures of chest CT images.

Using chest computed tomography, it is essential to distinguish between favourableand harmful lung nodules to identify lung cancer early, when there is the greatest possibility that it will be cured. The most efficient way for identifying images right now is deep learning 28 however a lot of training examples are needed, and these are frequently not used in standard medical imaging systems. In this study, the author used the transferable multi-model ensemble (TMME) technique to segregate benign from malignant cancer nodules using the scant quantity of chest computed tomography data that is currently accessible. This methodology transmits the image data with ResNet-50 models, then the ImageNet database is applied to predict the lung nodes. The nodes are classified with an adaptive weighting scheme when applied into the error back propagation. The three models are used to characterise the measurement metrics such as shape, voxel values and appearance.

One of the most important steps in the detection of lung cancer in initial stage and a potent method to improve patient survival is the classification of harmfuland graciouslung nodules using scan images. The lung nodules and tissues around the wound in the cancer cells are very critical to classify. Hence it is very challenging to accurate to predict using machine learning concepts. To solve this issue the author, introduce a multi-model ensemble learning architecture lies on 3D convolutional neutral networks (MMEL-3DCNN).

Cancers are most prevalent disease which will end up the human life. In 2017, the cases were increased 26% in world wide. The disease cure is only 18%, despite major improvements in predict and cure in recent years. Prior to diagnosis, classifying lung nodules is crucial, notably since computerised classification can benefit medical professionals by offering a useful viewpoint. The classification of CT images using modern machine learning and computer vision techniques is incredibly speedy and precise.

More people die from lung cancer compared with other cancer and diseases worldwide. The risk of mortality in early-stage patients is believed to be reduced by low-dose computed tomography (LDCT) of the chest. An autonomous and intelligent system that matches or exceeds the analytical skills of human experts would improve the accuracy and efficacy of cancer screening. Methods: We created a framework for lung cancer screening based on the deep neural network (DNN) artificial intelligence (AI) technology [29]. To begin with, the photos used for training were labelled using a semi-automated annotation technique. Then, DNN-based models for the detection of lung nodules (LNs) and despitefulnessclassification were created in order to detect cancer specially lungs from LDCT pictures. The developed LN detection was given the name Deep Learning, and it was verified using a sizable dataset using deep learning concepts.

Each imaging modalities (X-ray, CT, MRI, ultrasound), and other imaging systems are having the rhyme sequence and some continuous noise mechanisms, which allow them to record the underneath anatomy in unique ways. Any two modalities or sequences have a significantly non-linear intensity conversion. For example, Reddy et al. [30] used a deep learning network to find the better precision in between the various images. Sometimes we use the synthesis data to create the database with the original images, such as the creation of CT images from MRI data. Radiation can be avoided, therefore this can be advantageous.

We put forth a system that uses a convolutional neural network to segregate brain tumours from 2D MRIs of the brain, followed by conventional classifiers and deep learning techniques. To properly train the model, we have collected a wide range of MRI pictures with variable tumour sizes, locations, forms, and image intensities. To double-check our work, we also used an SVM classifier and other activation algorithms (softmax, RMSProp, sigmoid, etc.). We use TensorFlow and Keras to build our suggested solution because Python is an effective programming language for quick work.

Deep Neural Systems (DNN) An ANN having n umber of hidden layers between the input and output layers. Based on the algorithm and system we can create the hidden layers that is referred to as a deep neural network (DNN). The compositional models produced by DNN architectures express the object as a layered composition of primitives. Effectively allowing for the modelling of complicated information with smaller units than a shallow network that performs similarly. The extra layers enable for the compilation of features from lower levels. DNNs are typically feed-forward networks in which information moves straight from layer to layer means input to output. Convolutional neural networks (B) (CNN) CNN is a feed-forward network variation with a multi-layered, intricately coupled architecture. A modified version of the back-propagation technique was used to train CNN.

According to a survey, brain tumours are the cause of the greatest death rate worldwide. Changes in hormones, blood clots, weakness, unsteady gait, slurred speech, mood fluctuations, eyesight loss, etc. are among the symptoms. The type of tumour is determined by its location, and an accurate diagnosis can prolong the patient's life [31]. Non-cancerous growths called benign tumours cannot infiltrate nearby tissue. They are entirely erasable and quite unlikely to return. Although benign brain tumours cannot spread to nearby tissue, they can nonetheless result in severe discomfort, irreversible brain damage, and even death. Brain tumours that are malignant have no clear boundaries. They spread throughout the brain or spinal cord and proliferate swiftly, increasing the pressure there.

When used in conjunction with intracranial imaging techniques like MRI or CT scan [32], the entire structure of a tumour can be seen. MRI scans use powerful magnetic fields and high radio frequencies to deliver precise information about soft tissues. Compounded Tomography scans are performed using X-ray beam transmission. The following steps are taken in the detection of brain tumours: an image's pre-processing, feature extraction, segmentation, and post-processing.

3. Technology

CNN

CNN is having two components like convolutional and pooling layers. These are two considering as components of convolutional neural network. It is very crucial to build a model with research and inculcate with the dataset. The created model will imply with many neurons to build the structure of CNN

Studying successful applications is a helpful strategy for discovering how to create efficient convolutional neural network architectures [33]. Due to the extensive research and deployment of CNNs for the ImageNet Large Scale Visual Recognition Challenge, or ILSVRC, from 2012 to 2016, this is very simple to accomplish. Due to this issue, the revolutionary for incredibly difficult computer vision applications has advanced quickly, and general breakthroughs in the design of convolutional neural network models have also been developed.

If we reduce overfitting time, then it reduces operational time in a Cascade Deep Learning model [34]. In the following step, a straightforward and effective Cascade Convolutional Neural Network (C-ConvNet/C-CNN) is suggested because we are only working with a smaller portion of each slice of the brain. This C-CNN model uses two alternative approaches to mine both local and global characteristics. Additionally, a novel Distance-Wise Attention (DWA) mechanism is introduced to increase the accuracy of brain tumours in comparison to with the different models. The central location of the brain inside the model, along with the tumor's [35] influence, are both taken into account by the DWA procedure.

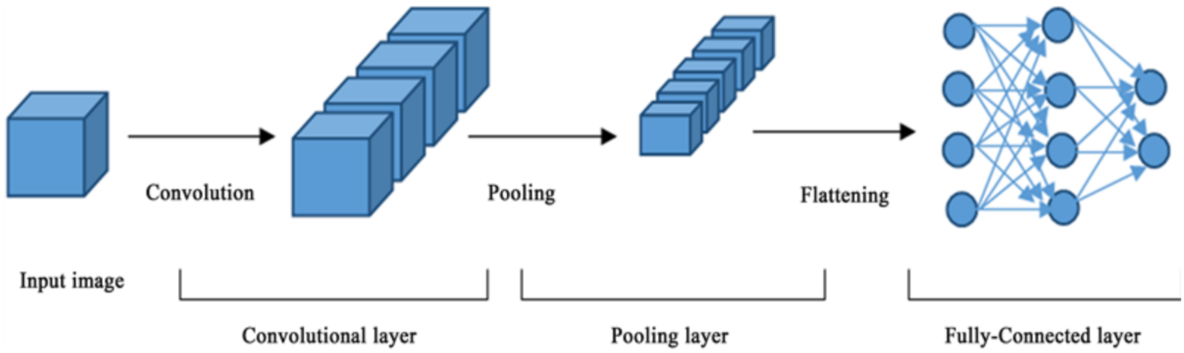

Fig 3. CNN architecture for image processing

Fig 3 explains about the CNN architecture for image processing. It contains the various layers which will segment and filter the image frames.

RNN

When the network layers are employing with sequential data or time series data is representing a recurrent neural network (RNN) [36]. There are different application like Siri, voice search, and Google Translate using deep learning concepts. Like we also use this powerful technique in medical field to provide value to human life. In this models, the researchers used to build a model with the required features. Then the model is trained with desired network design [37]. In RNN, feed forward concept is used to stand out of its memory which will affect the current input and result. The sequence of hidden layers used to refine the result which are collected from the previous layers. So, the latent value and correlated values are considered as coefficient values. That values are refined in the multiple hidden layers. Always, RNN outputs are mostly dependent on the previous layers in the network. But the prediction is accurate in image values because the multiple layer correlated values.

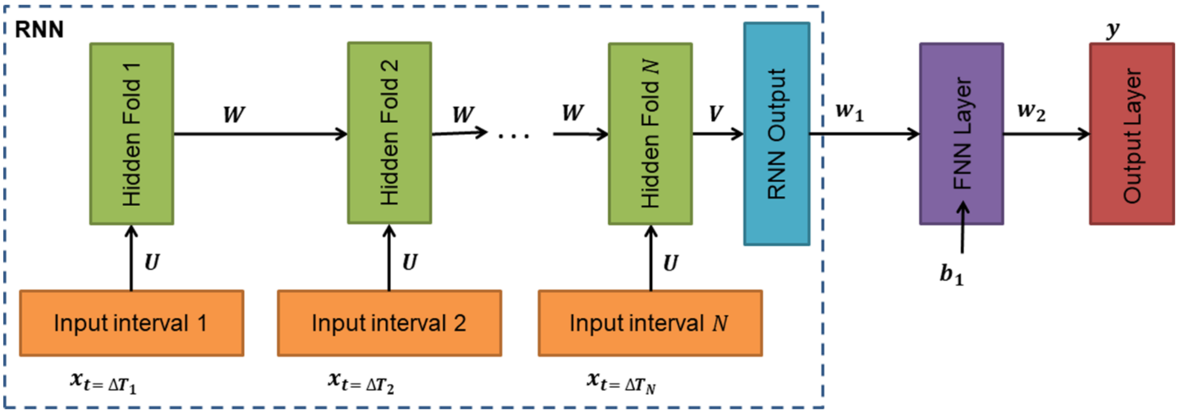

Fig 4. The RNN layers structure

Bidirectional recurrent neural networks (BRNN): Fig. 4 explains this alternative RNN network topology [38]. To improve the precision of predictions provided by unidirectional RNNs, which may only use previous data, bidirectional RNNs incorporate prediction performance. Going back to the earlier example of "feeling under the weather," the model would be better able to forecast "under" if it anticipated that "weather" would be the last word in the string.

Long short-term memory (LSTM): Sepp Hochreiter and Juergen Schmidhuber created this well-known RNN architecture as a solution to the diminishing gradient problem. In other words, the RNN model mightnot be predictedcorrectly forecast the current state if the state from earlier is having an impact with the current prediction.For illustration, suppose we wished to foretell the italicised phrase, someone has a nut allergy. She has a peanut butter allergy. We can determine whether a food contains nuts dependent on the circumstances of a nut allergy. However, if that information had been a few phrases previously, it would be difficult for the RNN to correlate the inputs.

The "cells" of LSTMs [39] in the deep levels of the neural network have three gates: 1) an input gate, 2) an output gate, and 3) a forget gate. These gates regulate the information flow that is necessary for the network to predict output. For instance, you may remove a gender pronoun like "she" from the cell state if it appeared more than once in earlier phrases.

Gated recurrent units (GRUs): This RNN version is comparable to LSTMs since it attempts to solve the temporary memory issue that plagues RNN models. It employs hidden states to govern information rather than "cell states," and includes two gates—1) a reset gate and 2) an update gate—instead of three. The reset and update gates regulate the amount of and specific information to preserve, similar to the gates in LSTMs.

GNN

A fascinating recent development in machine learning is a technique called generative adversarial networks (GANs) [40]. Generative models, or GANs, produce fresh data instances that mimic your training data. GANs, for instance, have the ability to produce images that resemble snapshots of real humans even though those faces don't actually belong to any real people. In image synthesis, generative approaches are frequently employed. The most recent generation models include GAN, variational autoencoders, and autoregressive models [41]. Sharp image production is a GAN's strength, but its training process is unstable. To synthesise nodular images, the three most well-known GAN approaches were put to the test. First, the deep convolutional model of the genuinewGAN-GP (26) and pix-to-pix (27) was used. However, the quality of these generated images was subpar and they were hazy. We created a pgGAN-inspired growing more gradually wGAN using sliced Wasserstein distance loss, international segment, and pixel normalisation. It produced high-quality photos and stabilised the training process.

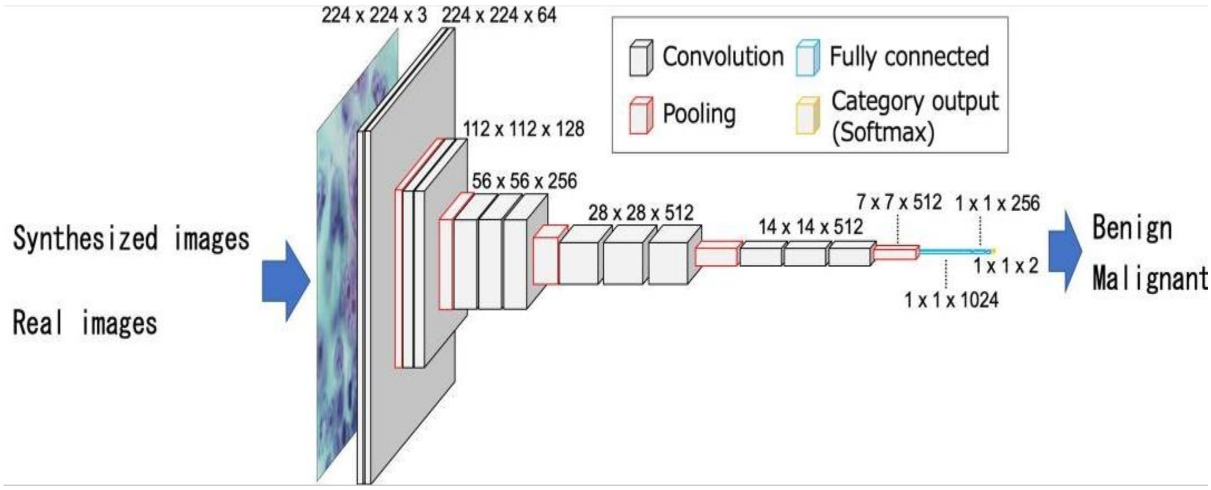

Fig 5. Image pooling in GAN

There were nine convolution layers in the discriminator as well. It began with a 64pixel picture, proceeded through four blocks made up of two convolution layers, and ended up with a 44 pixel feature map [42] that are shown in fig.5. This process was mirror imaged using the generator. To increase training stability, mini-batch discrimination was incorporated into the final layer in convolution network. To obtain the final true or false target, this feature map was processed with two compact, seamlessly connected layers.

4. Comparison Analysis

In the comparative analysis, the proposed multi-view model was pitted against existing single-view and traditional supervised methods. The same dataset was used for consistency, and the performance of each model was evaluated using the aforementioned metrics. Additionally, the models were subjected to different levels of data augmentation and noise simulation to test their robustness.

The 3D Medical Multi-modal Segmentation Network represents an advanced computational framework tailored for the precise segmentation of medical images acquired from multiple imaging modalities. This network is designed to handle three-dimensional medical image data, such as CT scans or MRI volumes, and integrates information from different imaging modalities to enhance the accuracy of segmentation tasks.

The network's architecture incorporates deep learning techniques, often based on convolutional neural networks (CNNs) or their variants, to automatically learn intricate features and patterns present within the volumetric medical images. By exploiting multi-modal data, which can include various imaging perspectives like T1-weighted, T2-weighted, and FLAIR MRI scans, the network can capture complementary information that improves the accuracy and reliability of segmentation results.

This kind of network is crucial in medical image analysis tasks, particularly in delineating structures of interest such as organs, tumours, or lesions. The fusion of information from different modalities helps the network to overcome challenges posed by variations in image appearance and texture, leading to more robust and precise segmentation outcomes.

The architecture of a 3D Medical Multi-modal Segmentation Network may involve encoder-decoder structures, often with skip connections or attention mechanisms to capture both local and global contextual information. Transfer learning and pre-trained models might also be utilized to enhance network performance, especially when training data is limited. The network's performance is typically assessed using metrics such as Dice coefficient, Jaccard index, or Hausdorff distance, which measure the overlap between the predicted and ground truth segmentations.

In the medical field, the 3D Medical Multi-modal Segmentation Network contributes significantly to clinical workflows, treatment planning, and disease assessment. Its ability to accurately segment structures from multiple modalities aids clinicians in making informed decisions, monitoring disease progression, and planning interventions. However, successful implementation requires rigorous training on diverse datasets and careful tuning of hyperparameters to achieve optimal results for specific medical imaging tasks.

5. Proposed System

Latent Correlation Representation Learning for Brain Tumour Segmentation with Missing MRI Modalities

The assessment of brain tumours is frequently done using magnetic resonance imaging (MRI). Clinical diagnosis and treatment planning depend on the precise segmentation of brain tumours using MR images. Multi-modal MR imaging [43] can also offer supplementary data for precise brain tumour segmentation. In clinical practice, it's typical to overlook several imaging modalities. We describe a brand-new brain tumour segmentation technique with missing modalities in this study. A correlation model is suggested to specifically depict the latent multi-source correlation because there is a significant link between multiple modalities. The segmentation becomes more reliable in the event that a modality is lacking as a result of the acquired correlation representation. The modality independent parameter is first estimated using the unique representation each encoder produces. The correlation model then converts each individual representation into a latent representation of a multi-source correlation.

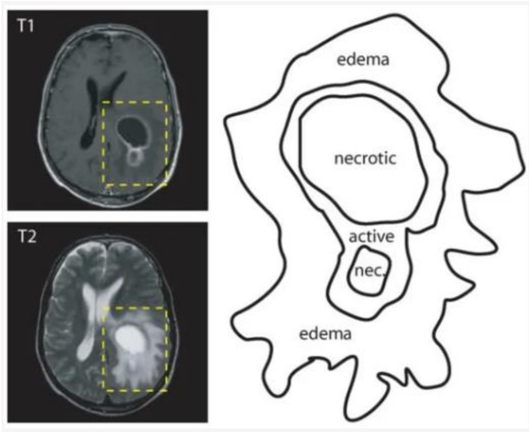

Fig.6 Brain segmentation

Fig 6 explains the brain tumours can be quantitatively analysed to learn more about their features and improve treatment strategy. Multiple imaging modalities with various contrasts are required for the precise segmentation of lesions. Manual segmentation, undoubtedly the most precise segmentation technique, would therefore be impractical for larger investigations. Deep learning has just lately become a viable option for quantitative analysis as a result of its record-breaking effectiveness. However, there are particular difficulties in medical image analysis.

6. Fusion Strategy

In the context of hand gesture identification, we look into a number of deep multi-modal fusion techniques in this study. In other words, our objective is to recognise the executed hand motion while successfully combining the data from several streams given many video inputs (such as thickness and colour data). We concentrate especially on establishing a common representation at hidden layer, which was disregarded in earlier studies. Previously, individually trained networks for each paradigm were combined via late fusion. We use the C3D [44] backbone architecture for this, which has shown promising results for multi-modal gesture detection, and we investigate the conventional late fusion technique. We also assess merging at intermediate network levels and provide an easy way for tying the streams together earlier using 111 convolutions, which we analyse at various network levels. Finally, using the cross-stitch units, We propose a unique architecture called C3Dstitch that can simultaneously learn to combine the signals of both neurons at different layers.

Strategy for Late Fusion [45] called as initial multi-modal method combines the results of the two or more networks throughout their last fully linked layer using late fusion, a method frequently used in gesture detection. Three different training methods for the model are investigated: 1) independent training with networks with various nodes and two independent losses, 2) end-to-end joint operations between a numbers of networks with a single loss calculated after the aggregate average, and 3) a multi-step approach that combines pre-processing and fine tuning. With the exception of the network's fine-tuning phase, which was trained in many phases, the learning parameters are the same as those used models for the backbone that were independently trained for each modality.

This work primarily focuses on methods where information exchange occurs at the intermediate network layer's feature map level so that relevant early feature correlations [46] are considered. Our initial instinct is to employ distinct streams at the base levels before merging them into a single model later. A simplistic fusion technique is provided by simple convolutions followed by combining of the multiple output extracted features. The form of each of the two inputs to the fusion modules should match that of the entry for an unique shared network of the layer below (after the fusion). As a result, we reduced by half the quantity of output filters in each of the 111 convolution layers (i.e., we subtract the number of streams from the number of filters). We can apply convolutions to reduce the filter space's dimensionality ratio. The conclusion is that the final design is composed of three components: 1) two early-stage networks for each unique modality and 2) a shared networks for the final phase that uses the shared input representation.

Fusion on multiple Levels via Cross-stitch Units

We had to manually select the model phase where the streams would be mixed up until this point. In this paragraph, our goal is to create a model that allows for simultaneous information sharing on various layers without placing restrictions on where individual or group learning occurs. We present an innovative multi-stream approach with distinct C3D networks for each modality that communicate with one another at both the pooling layer and the fully linked layer. Through a learnt weighted average known as cross-stitch units, the output of each of these layers is mixed in this architecture [47]. In other words, all networks pairwise interact at each level, and the extent to which alien modalities interact is discovered throughout.

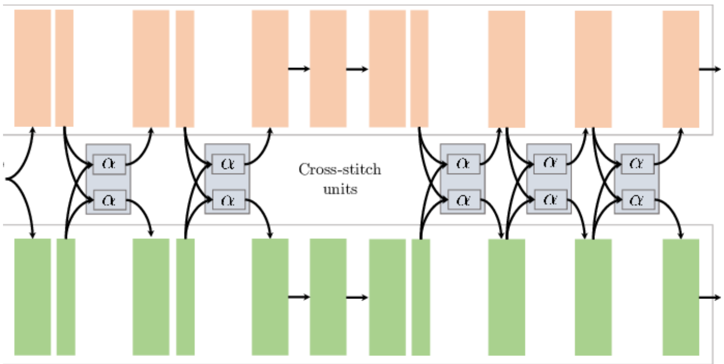

Fig 7. Cross stitch unit in network layers

3D Medical Multi-modal Segmentation Network Guided by Multi-source Correlation Constraint

The association between various modalities may be taken into consideration in the field of multifunctional segmentation [48] to enhance the segmentation outcomes. We suggest a sub classification system with a correlation restriction in this study. Our network has correlation constraint, feature fusion, decoding, and N model-independent encoding routes with N image sources. The N modalities' modality-specific features can be recorded using the model independent encoding route. We initially insist a linear correlation block to calculate the correlation between the different modalities because there is a substantial link between them. Then, using the linear correlation block as a guide, we apply a loss function to direct the network to learn the linked features.

This block drives the network to pick up the more useful latent correlated characteristics for segmentation. We suggest using a dual concentration proposed fusion block to modify the features along the modalities and geographical routes since not all of the features gathered from the encoders are appropriate for segmentation. This can help to mask less useful information and highlight relevant ones. The decoder [49] eventually projects the fused feature representation to produce the segmentation outcome.

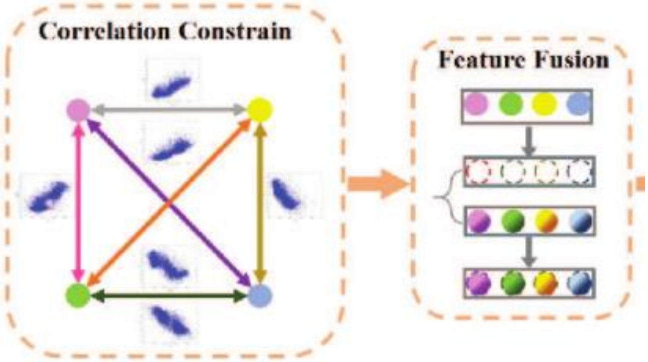

Fig 8. Correlation constraints

7. Conclusion

Early detection of brain tumours is essential for the diagnosis of cancer since an accurate diagnosis can improve survival rates. Given the difficulties of tumour biopsies, deep learning-based brain tumour analysis frequently makes use of three-dimensional (3D) magnetic resonance imaging (MRI). The N modalities' modality-specific features can be recorded using the model independent encoding route. We initially suggest a simple correlation block to study the connection between the modalities because there is a strong correlation among them. The network is then instructed to understand the associated attributes based on correlation values using the loss function. Thecorrelation block drives the network to pick up the more useful latent correlated characteristics for segmentation. Not all of the features that were retrieved from the encoders can be used to characterize the segmentation, hence we suggest using a dual attention based fusion block used to reconfigure the attributes and along modalities and temporal routes. This can help to mask less useful information and highlight relevant ones. In future, the following are the gaps identified and to be fulfilled in our research. The brain tumour MRI images are applied into a novel deep learning-based brain tumour segmentation for detecting diseases. We create a deep learning based framework for improving detection accuracy when using semi-supervised medical images and reducing computation and classification time. Future medical research can benefit from the method of combining all MRI sequences at once with strong generalisation capability, which can help radiologists diagnose tumours successfully.

Declaration:

Ethics Approval and Consent to Participate:

No participation of humans takes place in this implementation process

Human and Animal Rights:

No violation of Human and Animal Rights is involved.

Funding: No funding is involved in this work.

Conflict of Interest: Conflict of Interest is not applicable in this work.

Authorship contributions:

Authors Contributed Equally

Acknowledgment:

No acknowledge

Data Availability:

Based on the request authors can provide.

Reference

[1] Razzak, M.I., Naz, S. and Zaib, A., 2018. Deep learning for medical image processing: Overview, challenges and the future. Classification in BioApps, pp.323-350.

[2] Shen, Y. and Gao, M., 2019, June. Brain tumor segmentation on MRI with missing modalities. In International Conference on Information Processing in Medical Imaging (pp. 417-428). Springer, Cham.

[3] Zhang, D., Huang, G., Zhang, Q., Han, J., Han, J. and Yu, Y., 2021. Cross-modality deep feature learning for brain tumor segmentation. Pattern Recognition, 110, p.107562.

[4] Chattopadhyay, A., & Maitra, M. (2022). MRI-based Brain Tumor Image Detection Using CNN based Deep Learning Method. Neuroscience Informatics, 100060.

[5] Li, L., Zimmer, V. A., Schnabel, J. A., & Zhuang, X. (2022). Medical image analysis on left atrial LGE MRI for atrial fibrillation studies: A review. Medical Image Analysis, 102360.

[6] Byra, M., Han, A., Boehringer, A. S., Zhang, Y. N., O'Brien Jr, W. D., Erdman Jr, J. W., ... & Andre, M. (2022). Liver fat assessment in multiview sonography using transfer learning with convolutional neural networks. Journal of Ultrasound in Medicine, 41(1), 175-184.

[7] Chen, X., Wang, X., Zhang, K., Fung, K. M., Thai, T. C., Moore, K., ... &Qiu, Y. (2022). Recent advances and clinical applications of deep learning in medical image analysis. Medical Image Analysis, 102444.

[8] Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., ... & Xu, D. (2022). Unetr: Transformers for 3d medical image segmentation. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 574-584).

[9] Jin, Y., Zhao, Q., & Cheng, J. (2022, February). Oral Cone Beam Computed Tomography Images Segmentation Based On Multi-view Fusion. In 2022 7th International Conference on Intelligent Information Technology (pp. 59-64).

[10] Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H. R., & Xu, D. (2022). Swinunetr: Swin transformers for semantic segmentation of brain tumors in mri images. In International MICCAI Brainlesion Workshop (pp. 272-284). Springer, Cham.

[11] Soomro, T. A., Zheng, L., Afifi, A. J., Ali, A., Soomro, S., Yin, M., & Gao, J. (2022). Image Segmentation for MR Brain Tumor Detection Using Machine Learning: A Review. IEEE Reviews in Biomedical Engineering.

[12] Vasudevan, A., & Preethi, N. (2022). Analysis of MRI Images to Discover Brain Tumor Detection Using CNN and VGG-16. In International Conference on Ubiquitous Computing and Intelligent Information Systems (pp. 415-422). Springer, Singapore.

[13] Li, X., Ma, S., Tang, J., & Guo, F. (2022). TranSiam: Fusing Multimodal Visual Features Using Transformer for Medical Image Segmentation. arXiv preprint arXiv:2204.12185.

[14] Musallam, A. S., Sherif, A. S., & Hussein, M. K. (2022). A New Convolutional Neural Network Architecture for Automatic Detection of Brain Tumors in Magnetic Resonance Imaging Images. IEEE Access, 10, 2775-2782.

[15] Shah, H. A., Saeed, F., Yun, S., Park, J. H., Paul, A., & Kang, J. M. (2022). A Robust Approach for Brain Tumor Detection in Magnetic Resonance Images Using Finetuned EfficientNet. IEEE Access, 10, 65426-65438.

[16] Anaya-Isaza, A., &Mera-Jiménez, L. (2022). Data Augmentation and Transfer Learning for Brain Tumor Detection in Magnetic Resonance Imaging. IEEE Access, 10, 23217-23233.

[17] Song, Y., & Ma, R. (2022). Multiple Omics Analysis of the Rac3 Roles in Different Types of Human Cancer. IEEE Access, 10, 92633-92650.

[18] Zhu, J., Teolis, S., Biassou, N., Tabb, A., Jabin, P. E., &Lavi, O. (2020). Tracking the adaptation and compensation processes of patients' brain arterial network to an evolving glioblastoma. IEEE transactions on pattern analysis and machine intelligence, 44(1), 488-501.

[19] L. M. Oliveira et al., "Spectroscopic detection of pigments in tissues: correlation with tissue aging and cancer development," 2022 International Conference Laser Optics (ICLO), 2022, pp. 1-1, doi: 10.1109/ICLO54117.2022.9839939.

[20] H. Fu et al., "HMRNet: High and Multi-Resolution Network With Bidirectional Feature Calibration for Brain Structure Segmentation in Radiotherapy," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 9, pp. 4519-4529, Sept. 2022, doi: 10.1109/JBHI.2022.3181462.

[21] Kumar, E. Boopathi, and V. Thiagarasu. "Comparison and Evaluation of Edge Detection using Fuzzy Membership Functions." International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE), ISSN (2017): 2454-4248.

[22] I. Ezhov et al., "Geometry-Aware Neural Solver for Fast Bayesian Calibration of Brain Tumor Models," in IEEE Transactions on Medical Imaging, vol. 41, no. 5, pp. 1269-1278, May 2022, doi: 10.1109/TMI.2021.3136582.

[23] X. Xu and P. Prasanna, "Brain Cancer Survival Prediction on Treatment-Naïve MRI using Deep Anchor Attention Learning with Vision Transformer," 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), 2022, pp. 1-5, doi: 10.1109/ISBI52829.2022.9761515.

[24] Sundari, M. S., & Nayak, R. K. (2021, November). Efficient Tracing and Detection of Activity Deviation in Event Log Using ProM in Health Care Industry. In 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 1238-1245). IEEE.

[25] M. A. Hafeez, C. B. Kayasandik and M. Y. Dogan, "Brain Tumor Classification Using MRI Images and Convolutional Neural Networks," 2022 30th Signal Processing and Communications Applications Conference (SIU), 2022, pp. 1-4, doi: 10.1109/SIU55565.2022.9864962.

[26] J. Liu, "VGG, MobileNet and AlexNet on Recognizing Skin Cancer Symptoms," 2022 3rd International Conference on Electronic Communication and Artificial Intelligence (IWECAI), 2022, pp. 525-528, doi: 10.1109/IWECAI55315.2022.00107.

[27] M. Liu, C. Chen, Q. Yuan, L. Zhuang and P. Wang, "Olfactory Bulb Biosensor Monitoring the Effect of Treatment Drugs for AD in vitro," 2022 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), 2022, pp. 1-3, doi: 10.1109/ISOEN54820.2022.9789665.

[28] S. M. Kurian and S. Juliet, "Savitzky–Golay Pre-processed Kulczynski Similarity Feature Extractive Convolutional Deep Belief Network for Brain Tumor Detection," 2022 6th International Conference on Devices, Circuits and Systems (ICDCS), 2022, pp. 398-402, doi: 10.1109/ICDCS54290.2022.9780698.

[29] I. D. Romanishkin et al., "Investigation of a combined absorption-fluorescence-Raman spectroscopy approach to brain tumor tissue

differentiation ex vivo," 2022 International Conference Laser Optics (ICLO), 2022, pp. 1-1, doi: 10.1109/ICLO54117.2022.9840141.

[30] Reddy, Ch Subba, T. L. Yookesh, and E. Boopathi Kumar. "A Study on Convergence Analysis Of Runge-Kutta Fehlberg Method To Solve Fuzzy Delay Differential Equations." Journal of Algebraic Statistics 13.2 (2022): 2832-2838.

[31] S. Kumar Mishra, D. Kumar, G. Kumar and S. Kumar, "Multi-Classification of Brain MRI Using EFFICIENTNET," 2022 International Conference for Advancement in Technology (ICONAT), 2022, pp. 1-6, doi: 10.1109/ICONAT53423.2022.9725927.

[32] S. S and S. Aurelia, "A Machine Learning Entrenched Brain Tumor Recognition Framework," 2022 International Conference on Electronics and Renewable Systems (ICEARS), 2022, pp. 1372-1376, doi: 10.1109/ICEARS53579.2022.9752195.

[33] M. Jagadeesh, P. Chitra, K. Srilatha, M. Sumathi and I. RexilineSheeba, "Brain Tumour Classification using CNN Algorithm," 2022 International Conference on Electronics and Renewable Systems (ICEARS), 2022, pp. 1813-1816, doi: 10.1109/ICEARS53579.2022.9752096.

[34] K. Padmavathi, O. S. R. R. Y. Thalla, S. S. Reddy, P. Yadlapalli, T. Roshan and T. Charan, "Transfer Learning Using EfficientNet for Brain Tumor Classification from MRI Images," 2022 International Conference on Computer Communication and Informatics (ICCCI), 2022, pp. 1-4, doi: 10.1109/ICCCI54379.2022.9740923.

[35] L. M. Sao Khue and E. Pavlovskiy, "Improving Brain Tumor Multiclass Classification With Semantic Features," 2022 Ural-Siberian Conference on Computational Technologies in Cognitive Science, Genomics and Biomedicine (CSGB), 2022, pp. 150-154, doi: 10.1109/CSGB56354.2022.9865366.

[36] Kumar, E. Boopathi, and M. Sundaresan. "Edge detection using trapezoidal membership function based on fuzzy's mamdani inference system." 2014 International Conference on Computing for Sustainable Global Development (INDIACom). IEEE, 2014.

[37] H. Li et al., "View-Disentangled Transformer for Brain Lesion Detection," 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), 2022, pp. 1-5, doi: 10.1109/ISBI52829.2022.9761542.

[38] N. Çınar, B. Kaya and M. Kaya, "Comparison of deep learning models for brain tumor classification using MRI images," 2022 International Conference on Decision Aid Sciences and Applications (DASA), 2022, pp. 1382-1385, doi: 10.1109/DASA54658.2022.9765250.

[39] Chandra, J. V., &Pasupuleti, S. K. (2022, March). Machine Learning Methodologies for predicting Neurological disease using Behavioral Activity Mining in Health Care. In 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1035-1039). IEEE.

[40] Y. Zhang, B. Dai, M. Dong, H. Chen and M. Zhou, "A Lung Cancer Detection and Recognition Method Combining Convolutional Neural Network and Morphological Features," 2022 IEEE 5th International Conference on Computer and Communication Engineering Technology (CCET), 2022, pp. 145-149, doi: 10.1109/CCET55412.2022.9906329.

[41] Yookesh, T. L., et al. "Efficiency of iterative filtering method for solving Volterra fuzzy integral equations with a delay and material investigation." *Materials today: Proceedings* 47 (2021): 6101-6104.

[42] M. N. de Barros, E. D. S. Souza and J. Lisboa-Filho, "Using spatial analysis to calculate the incidence of lung cancer in the region of Muriaé-MG," 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), 2022, pp. 1-6, doi: 10.23919/CISTI54924.2022.9820194.

[43] C. Tejaswini, P. Nagabushanam, P. Rajasegaran, P. R. Johnson and S. Radha, "CNN Architecture for Lung Cancer Detection," 2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT), 2022, pp. 346-350, doi: 10.1109/CSNT54456.2022.9787650.

[44] S. A. Ali, N. Vallapureddy, S. Mannem, Y. Gudla and V. Malathy, "Detection of Cancer in Lung CT Image Using 3D CNN," 2022 2nd International Conference on Intelligent Technologies (CONIT), 2022, pp. 1-4, doi: 10.1109/CONIT55038.2022.9848092.

[45] D. Rawat, Meenakshi, L. Pawar, G. Bathla and R. Kant, "Optimized Deep Learning Model for Lung Cancer Prediction Using ANN Algorithm," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC),

2022, pp. 889-894, doi: 10.1109/ICESC54411.2022.9885607.

[46] G. Sruthi, C. L. Ram, M. K. Sai, B. P. Singh, N. Majhotra and N. Sharma, "Cancer Prediction using Machine Learning," 2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM), 2022, pp. 217-221, doi: 10.1109/ICIPTM54933.2022.9754059.

[47] S. Tomassini et al., "Cloud-YLung for Non-Small Cell Lung Cancer Histology Classification from 3D Computed Tomography Whole-Lung Scans," 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022, pp. 1556-1560, doi: 10.1109/EMBC48229.2022.9871378.

[48] M. Marathe and M. Bhalekar, "Detection of Lung Cancer using CT Scans with Deep Learning Approach," 2022 7th International Conference on Communication and Electronics Systems (ICCES), 2022, pp. 1026-1031, doi: 10.1109/ICCES54183.2022.9835901.

[49] K. K. Kumar and K. V. Kanimozhi, "Detection of Lung Disease using Novel Genetic Algorithm in comparison with Particle Swarm Optimization to improve Accuracy," 2022 International Conference on Business Analytics for Technology and Security (ICBATS), 2022, pp. 1-4, doi: 10.1109/ICBATS54253.2022.9759044.

[50] Keerthi, R.S., Dhabliya, D., Elangovan, P., Borodin, K., Parmar, J., Patel, S.K. Tunable high-gain and multiband microstrip antenna based on liquid/copper split-ring resonator superstrates for C/X band communication (2021) Physica B: Condensed Matter, 618, art. no. 413203

[51] Dhabliya, D. Delay-Tolerant Sensor Network (DTN) Implementation in Cloud Computing (2021) Journal of Physics: Conference Series, 1979 (1), art. no. 012031,