FMDB *Transactions on Sustainable Health Science Letters*

Detection of Brain Tumour in Sensor-Based Microwave Brain Imaging Using Dung Beetle Optimiser-Based Hybrid Transformer Enhanced CNN Model

B. Gunapriya[1,*], V. Revathi[2], Thirumalraj Karthikeyan[3], S. Gopikha[4], Sheila Agnes Vidot[5]

[1]Department of Electrical and Electronics Engineering, New Horizon College of Engineering, Bengaluru, Karnataka, India.
[2]Department of Research and Development, New Horizon College of Engineering, Bengaluru, Karnataka, India.
[3]Department of Artificial Intelligence, Trichy Research Labs, Quest Technologies, Tiruchirappalli, Tamil Nadu, India.
[4]Department of Information Technology, St. Joseph's College of Engineering, Chennai, Tamil Nadu, India.
[5]Department of Medical Consultant, IPS Health, Mahe, Victoria, Seychelles.
gunapriya1978@gmail.com[1], revshank153@gmail.com[2], thirumalraj.k@gmail.com[3], gopikha.re@gmail.com[4], vidotsheila@gmail.com[5]

Abstract: Deep learning and artificial neural networks can diagnose brain tumours as benign or malignant. Brain disease research requires image categorisation and automatic BT segmentation from reconstructed microwave (RMW) brain images. The tumour's architecture requires manual tumour identification, classification, and segmentation, which is highly time-consuming yet vital. RMB images were originally utilised to create a sensor-based microwave brain imaging (SMBI) image library. The 1320 photos include 300 "normal" photographs, 215 "malignant" images, 200 "double benign" and "double malignant" images, and 190 "single benign" and "single malignant" images. Images were resized and normalised. The dataset was expanded to 13,200 training photos per fold for 5-fold cross-validation. Though promising in BT classification, deep learning techniques cannot extract global features and preserve long-range associations. Vision Transformer (ViT) uses a self-attention mechanism to model long-range associations, which are essential for BT categorisation. For the BT organisation, the researcher employs a TECNN-based model with a CNN for extraction and an attention instrument in the transformer for global feature extraction. The Dung Beetle Optimisation Algorithm (DBO) chooses the best weight value here. The TECNN model achieved 89%-91% six-class categorisation accuracy after training on raw RMB images. The SMBI scheme's RMB photo classification can use the suggested model to diagnose tumours accurately.

Keywords: Brain Classification; Reconstructed Microwave; Vision Transformer; Double Malignant; Microwave Images; Dung Beetle Optimisation Algorithm; Global Pooling.

Received on: 22/11/2024, **Revised on:** 31/01/2025, **Accepted on:** 23/03/2025, **Published on:** 07/09/2025

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSHSL

DOI: https://doi.org/10.69888/FTSHSL.2025.000501

Cite as: B. Gunapriya, V. Revathi, T. Karthikeyan, S. Gopikha, and S. A. Vidot, "Detection of Brain Tumour in Sensor-Based Microwave Brain Imaging Using Dung Beetle Optimiser-Based Hybrid Transformer Enhanced CNN Model," *FMDB Transactions on Sustainable Health Science Letters*, vol. 3, no. 3, pp. 148–164, 2025.

1. Introduction

Brain tumours and other abnormalities of the brain are becoming a leading cause of mortality globally. A brain tumour develops from the growth of malignant cells inside the skull. It can eventually evolve into cancer, which is devastating to the brain's key

*Corresponding author.

components. Brain cancer is a serious health problem since it threatens human life, can be deadly, and has a major impact on patients' quality of life [1]. The risk of having brain cancer rises with each passing day as a result of the unchecked expansion of brain tumours. Radiologists and clinicians face significant challenges when analysing, classifying, and identifying brain tumours. The proper treatment of brain cancer necessitates a prompt and precise diagnosis [2]. Segmenting specific tumour areas from a head image is a common task in medical imaging applications, and brain tumour segmentation can be an important approach to this. In addition, clinical evaluation and treatment planning for brain cancer require accurate automated segmentation of brain tumours from clinical imaging. Brain cancer is the tenth largest cause of mortality in both adults and children, per the American Cancer Society [3]. Brain tumours can be effectively treated, but early diagnosis, categorisation, and research are crucial. Current imaging techniques are used to diagnose brain tumours in cutting-edge hospitals [4]. Medical professionals, such as radiologists, can use these imaging guidelines to detect better disorders, such as brain tumours [5]. Because of their large doses of radioactivity, poor sensitivity, strong ionising characteristics of brain tissues, high cost, and potential risk for pregnant women and elderly patients, these imaging modalities pose a significant danger of cancer [6].

Non-ionising, risk-free ionisation, cost-effectiveness, and a low profile all contribute to microwave imaging's (MWI) attractiveness to researchers for medical applications [7]; [8]. To address the limitations of conventional medical imaging techniques, scientists have recently turned to microwave imaging [9]; [10]. Microwave head imaging (MWHI) technology relies heavily on antennas, with single-antenna sensors used to reconstruct backscattered biological data [11]. After collecting and organising the necessary data, the image reconstruction algorithm is used to generate the final images. Brain tumours may be detected using a variety of microwave head-imaging methods and image reconstruction techniques [12]; [13]. However, the main drawbacks of the current MWHI modalities are (i) noise, blurriness, and low-resolution images generated by the system, (ii) difficulty in identifying tumour regions by automatic detection, and (iii) difficulty in identifying the tumour with a radiologist. Researchers have begun using methods to circumvent these restrictions [15]. The proposed TECNN consists of a CNN for extracting features and a transformer for transforming those features. By employing many convolutional and pooling layers, CNN, which is primarily concerned with pixel-wise information, is used to extract the local characteristics. The Transformer looks at data in patches rather than pixels. Images are first convolved before being fed into our suggested model. At this stage, the processed picture is split up and supplied into the feature extractor pipeline and the transformer pipeline, respectively [14].

There is a noticeable conceptual difference among CNN's feature maps due to their varying dimensionalities. To ensure that the advantages of both feature maps and patches are preserved during the conversion process, researchers employ CNN and transformer architectures, along with an intelligent module. The FFM and IMM draw on both CNN and transformer feature extraction methods. In the FFM, transformer patches are combined with the CNN's feature maps once the latter are translated into transformer space. Instead of just converting the input feature maps, researchers employ global pooling (GP) to select a useful channel. In this area, researchers focus on channels that cater to certain classes. This makes it much easier to switch between feature maps and patches. Spreading CNN's local news and information advice to the transformer stations is possible. The IMM performs average pooling (AG) after the Transformer input has been converted to a feature map. Subsequently, a channel-wise attention approach is applied to emphasise the class-sensitive channels of the pooled features. To further enhance the long-range dependence, CNN feature maps are adaptively fused with the transformer's input. The aggregated characteristics still contain both locally relevant data and globally significant traits. Our model effectively executes BT classification thanks to its hybrid structure. The following are some of our work's most notable contributions:

- The study recommends a hybrid approach for efficient BT categorisation because it (1) extracts local information while concurrently retaining global information.
- DBOA selects the best hyperparameter tuning, thereby increasing classification accuracy.
- The suggested approach uses the FFM and IMM representations: the former transforms CNN chin maps into PE, and the latter fuses them adaptively.
- Our model has been tested on publicly accessible datasets, and its experimental consequences for BT classification are superior to those of state-of-the-art approaches.

2. Related Works

Using a saliency map and deep learning feature optimisation, Khan et al. [16] offer an automated technique for detecting and classifying brain tumours. Step by step, the intended framework was put into action. An enhancement method is provided as the first step of the proposed framework. After that, an active-contour-based saliency-map-based tumour segmentation method is presented and applied to the original photos. After that, a CNN model, EfficientNetB0, which has already been partially trained, is fine-tuned and trained in two ways: on improved images and on tumour images. Features are taken from the average pooling layer, and both models are trained by deep transfer learning. Next, researchers apply the features. In the last stage, an enhanced dragonfly optimisation is used to categorise the most useful attributes. Using three publicly accessible datasets, the experimental approach improved accuracy to 95.14%, 94.89%, and 95.94%. The SVM classifier for brain tumour identification

has been solved using a unique parallel optimisation approach proposed by Qin et al. [17] that relies on a shared-memory environment. First, the features of MR images of brain tumours are the HOG procedure, compared to those obtained using the wavelet transform technique. As a method, researchers employ SVM with an L-norm loss function. Since it is sparse, it may be detected much more rapidly. At last, researchers propose and implement the SMP-SGD, Adagrad, and SMP-Adam algorithms on the classifier's solution. The experimental findings demonstrate that the HOG procedure is superior to the discrete wavelet transform approach for extracting features from brain MRI scans to detect brain tumours, as achieved by the suggested SMP-SGD algorithm and its modifications.

Given the extensive research on binary classification, Hossain et al. [18] have considered multiclass classification of brain tumours. Researchers looked into the effectiveness of several deep learning (DL) architectures, such as Visual, for tumour detection to make the process quicker, more objective, and more reliable. Next, researchers present IVX16, a transfer learning (TL) based multiclass classification model that combines features from the three highest-performing TL models. Researchers utilise a dataset of 3264 photos as our test set. Extensive investigations show that VGG16, Xception, and IVX16 achieve peak accuracies of 95.11%, 93.88%, 94.19%, 93.88%, 93.58%, 94.5%, and 96.94%, respectively. Using MRI, Ali et al. [19] offer feature optimisation strategies for wrapper-based metaheuristic deep learning networks (WBM-DLNets). Here, 16 pre-trained deep learning networks are utilised to compute the features. Classification performance is measured using a support-based cost function across eight metaheuristic optimisation algorithms. These include the marine predator algorithm, whale optimisation algorithm (WOA), grey wolf optimisation algorithm (GWOA), bat algorithm (FA), and firefly algorithm (FFA). The optimal deep learning network is selected using a deep learning network selection method. At last, the finest deep learning networks' deep features are combined to build a support vector machine (SVM) model. The suggested WBM-DLNets method is verified using a publicly accessible dataset. Using the features chosen by WBM-DLNets greatly improves classification accuracy compared to results obtained using the complete set of deep features. - b0-ASOA achieves a classification accuracy of 95.7%.

A novel, accurate, and optimised approach for detecting brain tumours has been proposed by Ramtekkar et al. [20]. Preprocessing, segmentation, feature extraction, optimisation, and detection are all steps that the system takes. A compound filter consisting of segmentation is performed using threshold and histogram methods. The feature extraction process uses a grey-level co-occurrence matrix. Here, researchers employ a CNN for optimal feature selection. Using metrics such as accuracy, precision, and recall, this system evaluates its performance relative to another contemporary optimisation method and asserts its superiority. Python is the programming language used to create this system. This improved approach has been assessed to have an accuracy of 98.9% in detecting brain tumours. Classification utilising the Deep Belief Network Classifier (DBNQLBC) has enhanced the work of Kumar and Prince [21]. When categorising brain images, the suggested DBNQLBC method is used to improve accuracy, reduce false positives, and speed processing. The DBNQLBC method detects brain tumours by extracting features from input brain MRI scans. The DBNQLBC method comprises three types of layers. Features extracted from brain MRI scans are passed to a hidden layer. To further categorise the features, a quadratic logit boost classifier is used in the hidden layer. To detect brain tumours using the likelihood measure, the boosting classifier employs the classifier. Reduced logit loss is used to combine the findings of several sub-classifiers into a single reliable classifier. The Boosting classifier selects the most accurate categorisations. This allows for reliable classification of input MRI images into normal and pathological categories, with the results displayed at the output layer. This allows for more precise, faster, and easier identification of brain tumours.

Brain tumour detection accuracy and brain tumour detection per unit of MRI scans are some of the metrics used in the simulation evaluation. Using the Bagging K-Nearest Neighbour (BKNN) to improve the KNN's accuracy and quality rate is a unique approach to detecting brain tumours, as proposed by Archana and Komarasamy [22]. First, a U-Net architecture is used for image segmentation, and then a bagging-based k-NN classification method is used. U-Net is used to achieve more precise and consistent parameter distribution across all layers. Bagged decision trees work well because each tree has tiny variances and gives slightly diverse, skilful predictions, since each tree is a classification. Classification accuracy for brain MR images between normal and diseased tissue reached 97.7%, demonstrating the efficacy of the proposed technique over current methods. A unique and robust automated brain tumour fusion of characteristics has been proposed by Mostafa et al. [23]. Researchers found that pre-processing the images improved localisation results. Researchers used Figshare and Harvard as the training and test datasets, respectively. Accuracy, recall, precision, F1 score, and area under the curve (0.989) were all maximised by using the proposed strategy. Researchers compared our method to the standard DL, classical, and segmentation-based techniques already in use. Cross-validation was also performed on the Harvard dataset, achieving 99.3% accuracy. The results demonstrate that our strategy yields significantly better outcomes than the alternatives.

3. Materials for Stacked Antenna Imaging Scheme Expansion and Image Reconstruction Procedure

3.1. Design and Expansion Procedure of the Stacked Antenna

In this study, researchers present the design and implementation of a trial stacked-antenna sensor-based microwave scheme for reconstructing RMB images and evaluating system performance. It is important to note that the MBI system requires a wideband

antenna sensor with a frequency range [24]. An innovative 3D wideband stacking antenna sensor inspired by metamaterials (MTMs) has been printed on low-cost Rogers substrates. The antenna sensor consists of two (2) air spaces and three (3) substrate layers. To keep everything in place, researchers use double-sided foam tape. An RO4350B substrate is used for printing the bottom layer (BL), whereas an RT5880 substrate is used for printing the top and intermediate layers. The internal space is 2 mm in width. One MTM array component (14) is used in the OL and ML, whereas another (32) is used in the BL. By combining MTM array elements across many layers, antenna performance may be enhanced in terms of efficiency and realised gain in free space and in proximity to the head model.

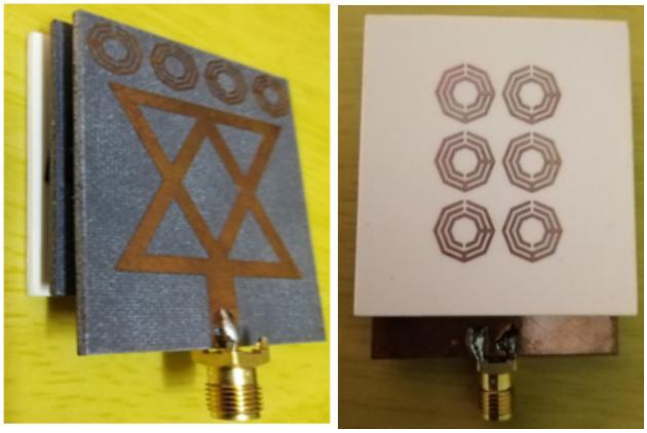

Figure 1: Fabricated antenna

An antenna sensor with dimensions of 50 by 40 by 8.66 mm³ is the optimal size. The antenna's presentation was validated by testing in both free space and in proximity to the head model. The antenna has 79.20% antenna efficiency (1.37-3.16 GHz), a radiation efficiency of 93%, a maximum fidelity of 98%, and a gain of 6.67 dBi, according to measurements. Figures 1 and 2 depict the manufactured antenna and the measured S-parameters (reflection coefficient). These findings confirm that the antenna can generate the required RMB images using the existing SMBI system. For this study, researchers used our recently developed MTM stacking antenna within the SMBI system architecture to produce RMB pictures.

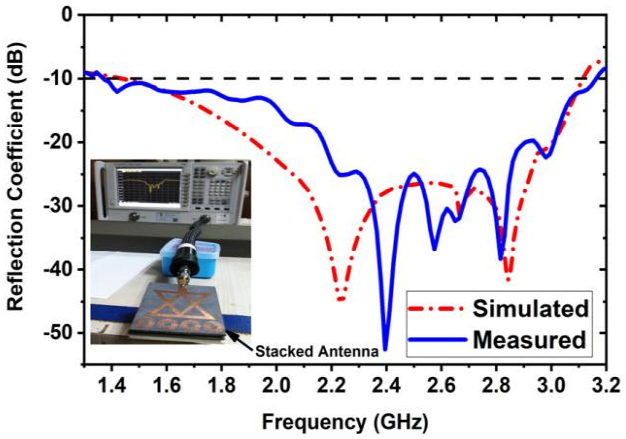

Figure 2: Reflection coefficient measurement

3.2. Phantom Perfect Fabrication Procedure and SMBI Scheme Implementation Procedure

To begin verifying the SMBI system, a six-layer tissue phantom containing tumour (benign and malignant) tissues was created. The electrical properties of malignant and benign tumours were considered to be specified amounts in Cheng and Fu [25]. The malignant tumour was fashioned in an atypical elliptical and triangular shape, whereas the benign tumour was designed to be generally spherical. The 3D head model was later layered with tissue phantoms and tumours. The model was modified to include both benign and malignant tumours at different locations, enabling measurements to be taken using brain imaging equipment. Both benign and malignant tumours were permanently attached to the skull model. S-parameters of the antenna sensor were calculated and evaluated using the developed head model, simulating tissue and tumours. The nine antennas installed within the helmet add even more functionality. An RF switch, a PNA E8358A transceiver, a microprocessor, and nine

arrays of antenna sensors were also included in the system. To complete a full revolution (360 degrees), the stepper motor-powered moveable platform tilts anticlockwise by 7.2 degrees with each revolution. The helmet has a secure grip on the engine shaft. The helmet has a diameter of 300 mm. The antenna sensor is fastened to the inside of the helmet using double-sided foam tape (DSFT). The optimal spacing between antennas for system-wide coverage is 40 degrees. To fine-tune the phantom head position, the sensor is placed 100 mm above the helmet's base.

The PNA communicates with the personal computer connected to the radio frequency (RF) transmitter, while Port B is connected to the RF receiver or switch. A six-layered 3D phantom model is built and attached to the centre of the helmet for testing the system's efficacy. The PNA collected reflected sensor signals after every 7.2-degree rotation. Next, the image creation method was applied to the collected sensor signals to generate RMB images. It's worth noting that in vivo imaging has recently been used to diagnose brain tumours in clinical practice. Because it is non-invasive, it may be used to see inside live organs (such as in brain tumour imaging) and pinpoint the precise position of objects of interest. To create a picture, however, a living item like an animal or human body is required, which is both difficult and problematic because of issues around medical approval. Due to ethical constraints imposed by ongoing clinical trials, researchers are unable to utilise a real human brain or body in our study to generate pictures of tumours [26]. This renders the collection of in vivo picture samples impossible. This is why researchers obtained all our data using a synthetic brain phantom model that closely mimics the features of a real brain. The data collected from the phantoms were then processed to generate images of simulated brain tumours for diagnostic purposes. However, in the future, researchers hope to obtain in vivo images for further analysis, which will aid medical professionals in easily categorising the tumour.

3.3. Illustration of RMB Image Trials

Trials of the tissue-like phantom, various tumour situations, and RMB pictures are shown in Figure 3. The phantom model's layout is shown in Figure 3 for clarity. As an additional detail, the tumours were placed in several different areas of the head model.

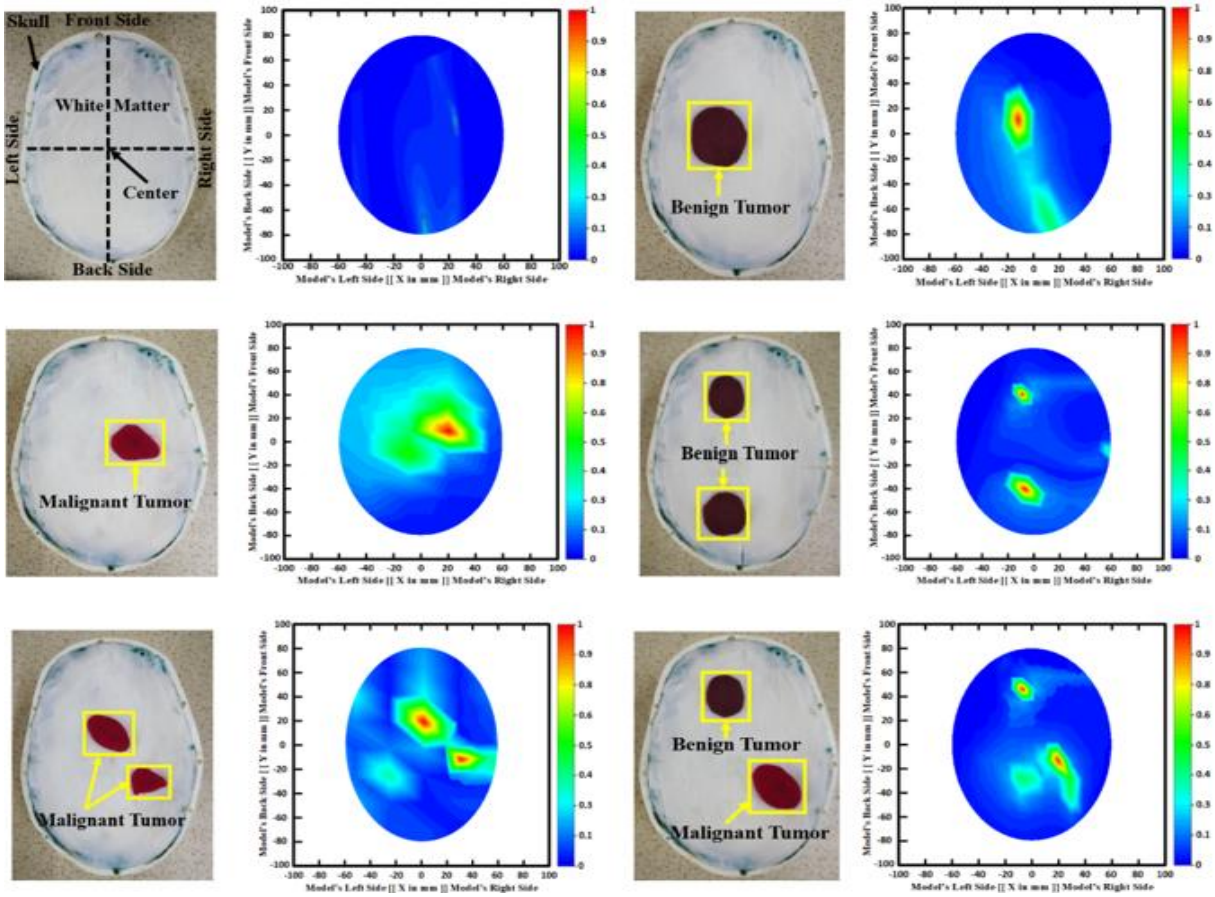

Figure 3: Phantom models of the skull that have been created to seem like real tissue, including RMB images: Noncancerous tissue (a), benign tumour (b), malignant tumour (c), benign and malignant tumour (BMT), and malignant tumour (MBT) images

The normal tissue (NT) picture, the benign tumour (BT) image, the malignant tumour (MT) image, the double benign and malignant tumour (BBT) image, and the single benign and malignant tumour (BMT) image are all shown in Figure 3. After that, 920 samples across all classes were gathered to form the foundation of the RMB image dataset. Later, the photos underwent several preprocessing steps to prepare them for training, validation, and testing the classification models. Authentic RMB photos were used to probe the suggested classifier. The obtained picture samples were subjected to image preprocessing and augmentation to generate a sufficiently sized training dataset. In this section, researchers detail our research strategy, dataset description, preprocessing technique, data enhancement steps, and analytical findings. Reconstructed microwave brain (RMB) pictures were used to supplement the dataset for classification models, with data coming from:

- This study built an experimental MBI system.
- Our prior study

Most studies have either:

- Examined healthy
- Examined diseased brain images (i.e., tumour-based images)

Images of a diseased brain are broken down into five groups:

- Those showing a single benign tumour (BT).
- Those showing a single malignant tumour (MT).
- Those showing two benign tumours (BBT).
- Those showing two malignant tumours (MMT).
- Those showing both single benign and single malignant tumours (BMT).

Classification performance across the six classes was examined, and a new network was investigated.

3.4. Preparation of Image Dataset

The images used in this study come from two different places: our currently operational MBI system and Hossain et al. [27]. Better categorisation results can be achieved by combining two datasets. Excluding duplicates, the dataset contains 1320 original RMB photos: 300 for the NT class, 215 for the BT class, 200 for the MT class, 200 for the BBT and MMT classes, and 190 for the BMT class.

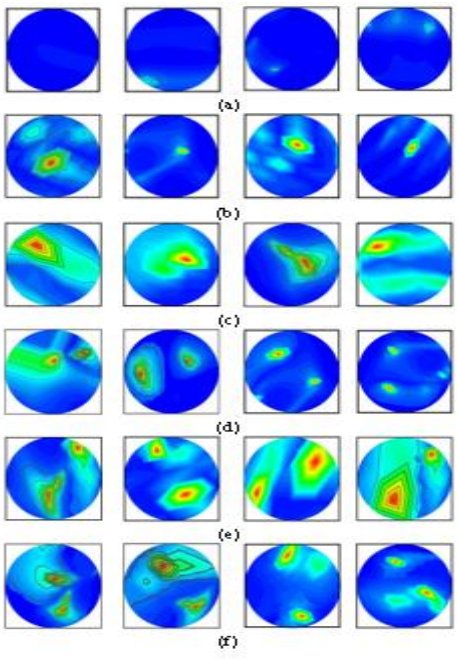

Figure 4: Randomly designated RMB image samples from the unique dataset: (a) NT, (b) single BT, (c) single MT, (d) BBT, (e) MMT, (f) BMT

Figure 4 shows examples of the unique RMB photos across all classes. On 20 November 2022, you may see the original RMB picture dataset and the source code at -Brain-picture. Due to the limited size of the training dataset, this research uses image augmentation to expand it. Image augmentation might boost model accuracy without the need for collecting more data or samples. It may generate a rich dataset from a small sample picture dataset, thereby improving network performance and considerably expanding the data available for training models. Rotation, scaling, and translation are examples of image augmentation techniques that can expand the training dataset. Image or data modern state-of-the-art image synthesis procedures, in particular, realistic augmentation, can also be explored. Image synthesis and data scaling were used as enhancement methods, and a diffusion probabilistic model was applied to both. Authors have assumed that a pixel value in an image dataset is an integer between 0 and 255, which linearly scales to (1, 1) when the data dimensionality is reduced. In addition, the images generated by our microwave imaging method are two-dimensional, nearly noise-free, and provide excellent perceptibility of the target object (i.e., tumour). Therefore, the procedures of anatomically directed distortion and adversarial diffusion augmentation cannot be applied to RMB images.

That's why this study employed four distinct picture augmentation strategies (rotation, scaling, translation, and inversion) to generate the test data. Both clockwise and anticlockwise rotations of 10-40° are applied to the photos. As a result, the tumour items are moved around to new positions in the pictures. Images may be scaled in either direction, making them smaller or larger. Here, magnifications of 10% to 12% are used on the images. The tumour objects are relocated in the photos using the image translation method, which involves a 10%- 15% vertical and horizontal translation. The vertical flipping technique is also used as an enhancement strategy. To achieve the objectives of training, validation, and testing, this study employed a method. For five-fold cross-validation, 80% of the photos were used for training, and 20% for testing. 80% of the dataset is used for training, and 20% for validation to avoid overfitting. Following the enhancement, 13,200 training photos, 264 testing images, and 231 validation images were generated per fold. Table 1 provides a comprehensive description of the data collection.

Table 1: Description of the datasets

Dataset	Sum of original Images	Image Lessons	Training Dataset					
			Sum of Copy Per class			Increased Train Image Per Fold	Testing Imageries Per Fold	Validation Copy Per Fold
			This Work	Ref [27]	Total			
Raw RMB Image Sample	1320	Non-Tumor(NT)	200	100	300	3000	60	48
		Single Benign (BT)	140	70	215	2151	43	35
		Single Malignant Tumours (MT)	140	70	215	2151	43	35
		Two Benign Tumours (BBT)	150	50	200	200	40	32
		Two Malignant Tumours (BBT)	150	50	200	200	40	32
		Solitary Benign and Solitary Malignant Tumour (BMT)	140	50	190	1900	38	31
		Total	920	400	1320	13,20	246	213

4. Methodology

4.1. Image Segmentation

For the watershed transform, a gradient magnitude picture is extracted from the edge data using the Sobel operators. When compared to noise, 3D Sobel operators produce more accurate segmentation results. The operator approximates the derivatives by convolving the original image with two separate 3×3 kernels, one for horizontal changes and one for vertical changes. The computations are as follows, Gu et al. [28], if A_i is the i-th slice of the source picture, GX_i and GY_i are two images that include the horizontal and vertical derivative approximations at each location, and G_i supplies the gradient magnitude:

$$GX_i = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \times A_{i-1} \begin{bmatrix} -2 & 0 & +2 \\ -4 & 0 & +4 \\ -2 & 0 & +2 \end{bmatrix} \times A_i + \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \times A_{i+1} \quad (1)$$

$$GY_i = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} \times A_{i-1} \begin{bmatrix} -2 & -4 & -2 \\ 0 & 0 & 0 \\ +2 & +4 & +2 \end{bmatrix} \times A_i + \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} \times A_{i+1} \quad (2)$$

$$G_i = \sqrt{Gx_i^2 + Gy_i^2} \quad (3)$$

Instead of performing the time-consuming 3D segmentation, the watershed is applied to the 2D image slices after 3D pre-processing is complete. As a result, Sobel operators are used on the X and Y axes of the images.

4.2. Classification Process

In some cases, it might be useful to classify tumours in an image purely based on local or global features. This research supports the use of TECNN to improve classification accuracy. The TECNN hybrid model combines the benefits of models. Figure 5 illustrates the system's general design.

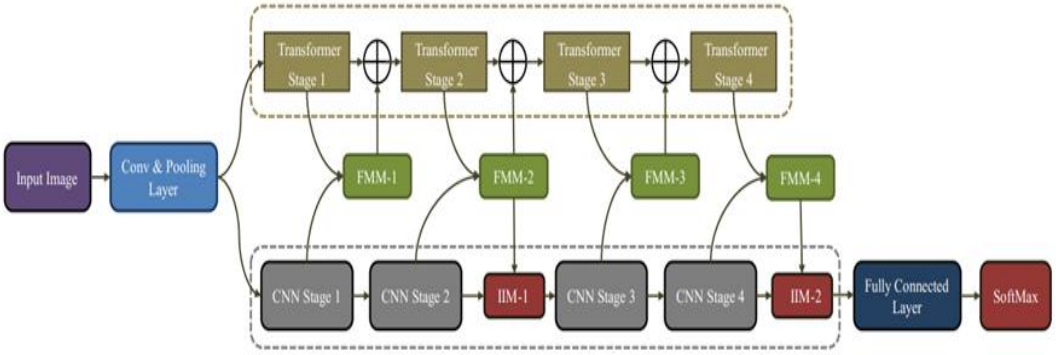

Figure 5: Overall structure of the projected model

The suggested model consists primarily of a CNN feature extractor and a transformer. To extract local information, CNNs typically use a combination of a convolutional layer and a max pooling layer. The input picture is max-pooled to preserve the most important local attributes and highlight features. A second transformer pathway is then used to choose the best features from the feature maps and feed them into the CNN. The CNN feature extractor uses a cascade of convolutional layers to extract spatial features. In the transformer route, self-attention modules are used for global feature extraction. In this case, both the Transformer and the CNN are active at once. Each CNN stage sends a feature map emphasising local characteristics to the transformer pathway to compensate for the latter's inherent lack of spatial inductive bias. In this case, the FFM merges the features of two sources, encouraging a seamless transition from feature maps to PE. The transformer then acts as if it were dependent on something far away by focusing on itself. After passing through the transformer stages, the PE are sent back to the FFM to be combined with feature maps in the IMM using CNN feature extraction. IMM can automatically combine local features from the CNN and global features from the transformer. The global features extracted by the transformer are then used to enhance the CNN's feature extraction. The IMM-2 result is linked to a classifier that generates our model's final prediction.

4.2.1. CNN Feature Extractor

The DenseNet is a cutting-edge convolutional neural network model with exceptional representation skills. Throughout the training phase, convolution techniques are utilised in these models to store local information hierarchically [29]. It might be difficult to train a model from scratch, especially with limited data. To save time and improve performance, researchers leverage a pre-trained model [30]. The weight parameters are set based on the DenseNet-121 model that has already been trained. Each layer in the DenseNet-121 is fed the sum of the outputs from the levels above it. Because of this connectivity, gradients from the loss function are instantly available to each layer, thereby accelerating the spread of features across the network. DenseNet-121 is a multi-stage architecture that employs stacked 3x3 and 1x1 convolution layers. The 1×1 convolution is inserted as a bottleneck layer before the 3×3 convolution to improve computational efficiency. Notably, a transition layer comes after every 33 convolutions. Here, researchers have a batch normalisation layer, an 11 convolutional layer, and an average pooling layer of size 22.

4.2.2. Transformer Pathway

The transformer route is meant to provide additional global direction to the CNN progressive features extractor. Before the first transformer step, the tensor is projected into a high-dimensional space with 768 channels via convolutional operations.

Therefore, the transformer leverages the convolutional layer's local knowledge. This squares with the discovery that improving feature representation by adding locality to the initial layers of transformers. At each stage of the transformer, the CNN's feature maps are squashed down to patches. Therefore, each pixel is treated as a separate PE. After the spatial feature maps are normalised, the patch sequence is supplied to the transformer. These updates now include a trainable class token. In a standard vision transformer, the class token is utilised to make a final prediction after training on class-specific data [31]. The suggested method utilises this learnable token in conjunction with feature maps in the FFM to provide international characteristics. There are four transformer phases in the TECNN, and each one employs both local and global pathways. As shown in Figure 6, the global route evaluates the entire image before making a decision, whereas the local pathway gathers knowledge from a small region of the image. To boost performance, researchers combine data from many locations in our study. Each transformer stage uses a pair of identical transformers, building blocks. Each MLP and MSA layer in a transformer consists of multiple processing layers. Layer norm is applied before each MSA and MLP in this scenario. In MSA, the input of size $B \times H \times W \times C$ is converted to a query: $Q \in B \times H \times W \times D_q$, key $K \in B \times H \times W \times D_q$, and value $V \in B \times H \times W \times D_v$ Matrices. Here, C, D_q, and D_V are the input, follows:

$$\text{Attention}(Q, K, V) = \sigma\left(\frac{QK^T}{\sqrt{Dv}}\right)V \tag{4}$$

The SoftMax function is denoted by (). The MSA layer is formed by repeating the process of self-attention head (h) times. After a GELU activation function, the MLP block consists of two contiguous layers that each project PE to either 3072 or 768 dimensions, respectively.

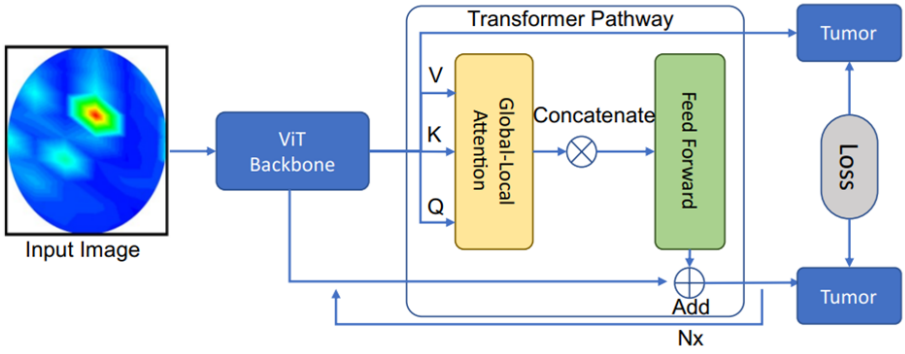

Figure 6: Characterises the transformer pathway for features

The subsequent can be used to accelerate the full process in a transformer block:

$$T_{output}^{MSA} = MSA\left(\tau\left(T_{input}\right)\right) + T_{input} \tag{5}$$

$$T_{output} = MLP\left(\tau\left(T_{output}^{MSA}\right)\right) + T_{output}^{MSA} \tag{6}$$

where T_{output}^{MSA} indicates MSL layer output, T_{output} is output, T_{input} is the transformer phase input, and $\tau(\cdot)$ is the layer normalisation purpose.

4.2.3. Feature Merge Module

The FFM receives two inputs, F_{input}^{c} from CNN and F_{input}^{T} from the transformer route. The F_{input}^{c} has $B \times C \times H \times W$ size, while F_{input}^{T} has a size of $B \times (1 + H \times W) \times C$ and 1 for the class token. Before being global pooling. This GP is viewed as a gate that selects educational channels and enhances the input's class sensitivity. Next, average pooling with a stride of 2 is used to down-sample the combined data. Average pooling helps transmit information globally by averaging the values of adjacent pixels. This brings the CNN input nearer to the transformer characteristics. After averaging all of the patches, the feature maps are converted to PE F_patchC. Input to a Transformer (F_{input}^{T}) and output (F_{patch}^{C}) are fused to create F_{output}^{T}. The projected model has four FFMs, all of which have F_{output}^{T}. As portrayed in Figure 7. The transformer stage is added to each F_{output}^{T} of the FFM, and the combined value is sent into the next phase. During the fusion procedure, all PE of F_inputT is discarded, save the class token. F_patchC is associated with the class token to obtain F_outputT, and the abundant local information in

F_patchC's characteristics aids CNN transformation. Once F_outputT is fed to the next transformer stage, the analysis global features based on local feature info to enhance the representation capabilities.

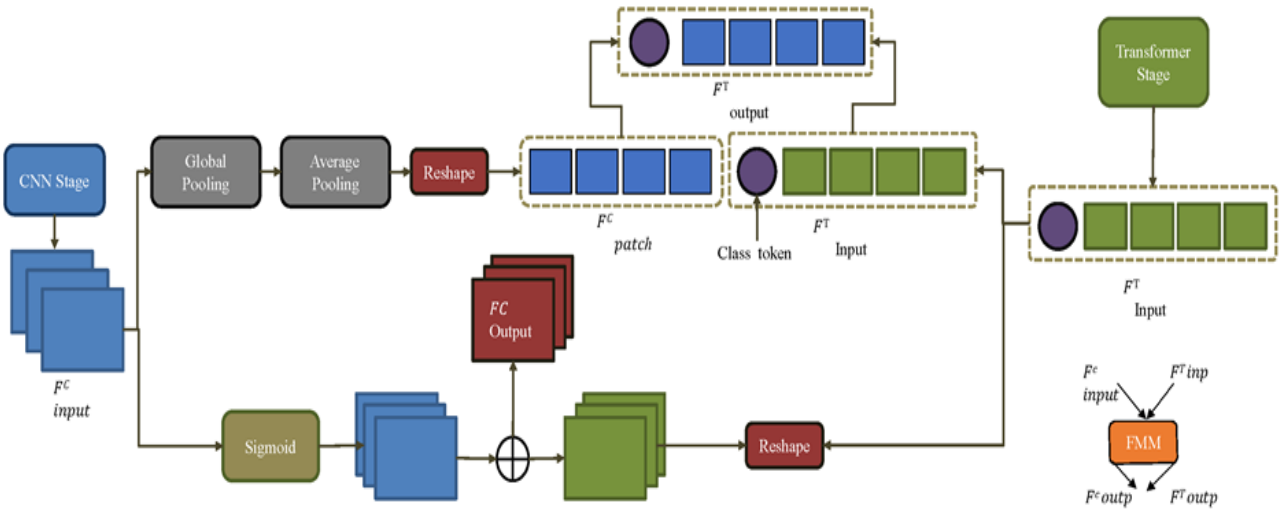

Figure 7: The projected FFM structure

Only the FFM-2 and FFM-4 comprise extractor, or F^{c}_{output}, as seen in Figure 7. Here, IMM-1 and IMM-2 receive F^{c}_{output} from FFM-2 and FFM-4. Transformer input (F^{T}_{input}) in the FFM space to produce F^{T}_{input}. The equation for F^{c}_{output} is as follows:

$$F^{c}_{output} = \theta(F^{c}_{input}) \otimes F^{T}_{input} \quad (7)$$

4.2.4. Intelligent Merge Module

In the IMM, it is given jointly with the CNN-based extractor. The IMM makes an exertion fold and chooses the most useful data for the user. I^{c}_{input} (B × C × H × W) and I^{F}_{input} (B × C × H × W), as demonstrated in Figure 8. I^{c}_{input} created from the CNN stage, whereas I^{F}_{output} comes from topographies in I^{F}_{input}, which is identical to F^{c}_{output} of the FFM. Here, the study of concept dependencies among the channels of I^{F}_{input} to recognise the most useful global aggregation. Figure 8 descriptor D_3 of size, B×C×1×1 reproduces these needs. C characterises the sum of channels of I^{F}_{input}. Each channel of D_3 each pixel designates the number of I^{F}_{input}. To reduce the number of educational channels, I^{F}_{input} is multiplied by D_3.

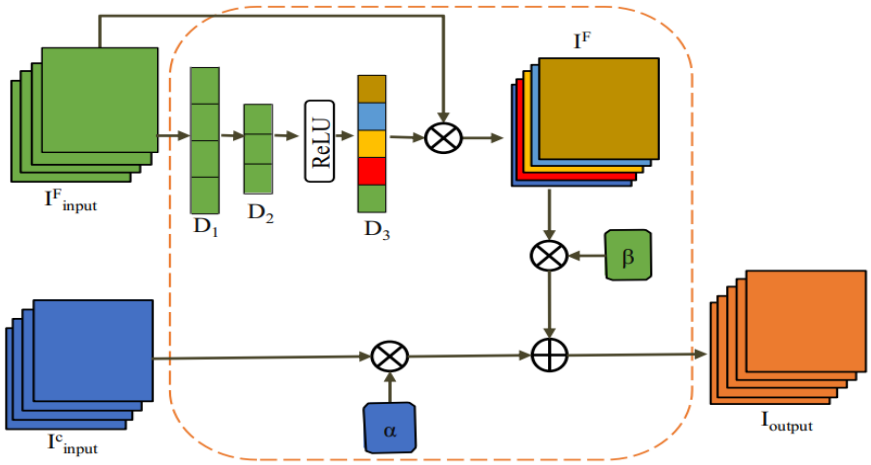

Figure 8: Representation of the IMM structure

Prior to gaining D_3, global syndicate the spatial info of I^{F}_{input} and crop descript tor D_1:

$$D_1^k = \frac{\left[\sum_{i=1}^{H}\sum_{j=1}^{W}\left(I_{input}^{F}\right)^{k}(i,j)\right]}{\frac{H}{W}} \quad (8)$$

where D_1 and D_3 both have a similar size. Here, $\left(I_{input}^{F}\right)^{k}(i,j)$ stipulates the kth channel of I_{input}^{F}, while D_1^k element of D_1. A form is obtained, and the feature images are averaged. Following, the researcher condenses D_1 using a totally connected layer L_1:

$$D_2 = L_1(D_1) \quad (9)$$

where $D_2 \in B \times (C/r) \times 1 \times 1, r$ is the After the data in D_2 has been processed by a ReLU layer, the channel number is restored, and the information can be accessed via the compact ratio (set to 16 for this research):

$$D_3 = \sigma(L_2(D_2)) \quad (10)$$

Where stands for function. During the training phase of the compact-restore process, D_3 undergoes dynamic changes in the values of its items, allowing it to learn about the interdependencies among I_{inputF}'s channels. In this case, IF is produced by a scaling of I_{inputF} by D_3:

$$I^F = I_{input}^{F}.D_3 \quad (11)$$

Descriptor D_3 is responsible for bringing greater attention to channels that deliver more relevant data and improving the feature representation of IF. Input C, which contains data unique to each class, is gathered by the CNN feature extractor with the help of the previously-trained model. CNN's core structure, on the other hand, requires that the input have a higher proportion of local features to global feature representations. Since PE already has global characteristics, researchers aggregate them and add them to the input C. As a result, the suggested TECCN gains in efficacy and has a greater capacity for representation. IMM's output is a weighted sum of I and C., and I^F:

$$(I_{output})^i = a^i.\left(I_{input}^{C}\right)^i + \beta^i.(I^F)^i \quad (12)$$

where $(I_{output})^i$ represents the output of I_output, (I_inputC) characterises the ith channel of input C, and (IF) represents the ith channel of input F. The length of the real arrangements in this scenario is the channels in the input signal (I_inputC). The and symbols both start with a value of 1. The parameters, like the rest of the model's parameters, are updated at each epoch during training. The features from I_inputC and IF receive different contributions depending on the values of a and b. These evolvable patterns impart intelligence onto the whole. The statistical characteristics of I_inputC and IF are dissimilar. There may be large differences between channels, for instance, in pixel count deviation. Therefore, regional or global features may be lost when I_inputC and IF are included. Each feature map can have its global and local trait proportions calculated independently, thereby preserving features at both scales. As a result, the CNN's regional characteristics may benefit from the global Transformer, and the TECNN's representational power can grow to meet expectations. Algorithm 1 also provides a detailed step-by-step explanation of the suggested model.

Algorithm 1: Training and Testing Stepladders of Our Perfect on Condition that

Input: Dataset of BT images.

Dataset division: Training, Validation, and **Testing Output**: Class labels.

1. Training:

- Model parameters
- 1. Image size: (224,224,3)
- P: 9
- Mini − batch size: 32
- N; the number of samples
- Learning rate: 0.0001
- Optimizer: DBO

2. Set the number of mini − batches as: $N_b = \frac{N}{b}$.

3. For iteration $= 1$: number of epochs

3.1. For batch $= 1$ number of mini − batches:

- Online image augmentation during training
- The obtained training set is fed to the TECCN Convolution and Transformer class
- branch
- The augmented images batch is fed to the TECCN encoder of the classification branch.
- The classification token is fed to the token classifier
- Calculated the loss function
- Loss backpropagation
- Updating the model parameters

Model testing:

1. Feed the input images to the model

2. Calculate the prediction label using output label Y

4.2.5. Hyper-Parameter Tuning Using Dung Beetle Optimiser (DBO)

DBO is an optimiser that draws on the thieving and reproductive habits of dung beetles to determine the optimal learning rate for the proposed model. Ball-, little dung-, and thief beetles are the DBO's four categories of search agents. In particular, different search agents use different update criteria. Please take note that the following description will focus on specifics.

4.2.5.1. Ball-Rolling Dung Beetle

Dung beetles use astronomical landmarks as guides to keep the dung ball moving in a straight course as they roll it. As a result, the dung beetle's current location may be stated as:

$$x_i(t+1) = x_i(t) + a \times k \times x_i(t-1) + b \times \Delta x \quad (13)$$

$$\Delta x = |x_i(t) - X^w| \quad (14)$$

Where t is the current iteration sum, $x_i(t)$ is the iteration, k, on a global scale; Dx models intensity. The dung beetle will dance around obstructions that might otherwise prevent it from continuing its journey. To change the ball's rolling direction, a tangent function is used to mimic the dance-like behaviour. As a result, researchers have revised the dung beetle's location to reflect the following:

$$x_i(t+1) = x_i(t) + \tan(\theta)|x_i(t) - x_i(t-1)| \quad (15)$$

where $\theta \in [0, \pi]$ is the ricochet angle.

4.2.5.2. Brood Ball

Dung beetles have to make sure their young are protected by choosing a good place to lay eggs. Therefore, in DBO, researchers suggest the selection approach to model the region where female eggs:

$$Lb^* = \max(X^* \times (1 - R), Lb) \quad (16)$$

$$Ub^* = \max(X^* \times (1 - R), Ub) \quad (17)$$

Where $X*$ represents the position; Lb^* and Ub^* characterise area; $R = 1 - t/Tmax$; Tmax denotes the supreme sum of repetitions; Lb and Ub are the space, correspondingly. Each female dung beetle is only expected to lay a single egg every cycle.

Since the spawning area's bounds move around dynamically as values change, the site of the egg ball moves around with each iteration as well, as shown below:

$$B_i(t+1) = X^* + b_1 \times (B_i(t) - Lb^*) + b_2 \times (B_i(t) - Ub^*) \quad (18)$$

where $B_i(t)$ is the site of the ith compass at the tth repetition; b1 and b2 are two freelancers random courses with dimension.

4.2.5.3. Small Dung Beetle

The little dung beetle is the common name for the kind of adult dung beetle that emerges from its burrow to forage for food. For tiny dung beetles, the following are the limits of their prime feeding territory:

$$Lb^b = \max\left(X^b \times (1 - R), Lb\right) \quad (19)$$

$$Ub^b = \min\left(X^b \times (1 - R), Ub\right) \quad (20)$$

where X^b represents the position; Lb^b and Ub^b are the limits on the suitable foraging region. So, here's an updated map showing where the little dung beetle is:

$$x_i(t+1) = x_i(t) + C_1 \times \left(x_i(t) - Lb^b\right) + C_2 \times \left(x_i(t) - Ub^b\right) \quad (21)$$

where $x_i(t)$ represents the iteration; C_1 is the random sum subject to the usual distribution; C_2 is the accidental vector within the variety of (0,1).

4.2.5.4. Thief

Thieving dung beetles are notorious for taking the dung balls of their fellow bugs. The best food source, as shown by Equation (5), is X^b. Let's pretend that the area around. X^b is where all the best competitive foods can be found. The following is how the thief's position is updated during the repetition procedure:

$$x_i(t+1) = X^b + S \times g \times \left(|x_i(t) - X^*| + \left|x_i(t) - X^b\right|\right) \quad (22)$$

where $x_i(t)$ characterises the ith thief's current position at the tth repetition; g is a normally distributed random vector of size 1D; S is a continuous value.

5. Results and Discussion

Nine different classification representations using the projected MBINet model are built and evaluated in this study on the Anaconda distribution platform. All of our tests run on a 64-bit Windows 10 installation with 128 GB of RAM and a CPU. An 11 GB NVIDIA GeForce is also used to boost the network training performance. Metrics for five-fold classification performance were computed.

5.1. Performance Metrics

Sensitivity and total measures were utilised to assess the performance of the various hybrid architectures during the picture classification validation process. Equations (23) and (25) allow these criteria to be calculated:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (23)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (24)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+Fp+TN} \quad (25)$$

Where $TP = \text{true positive}, FN = \text{false negative}, TN = \text{true negative}, \text{and } FP = \text{false positive}.$

Figures 9 and 10 show the loss graph and dice score for the training and test data.

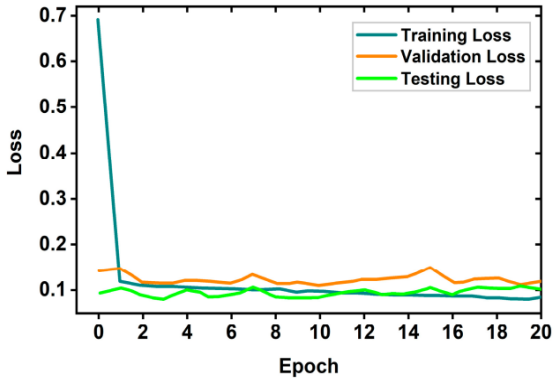

Figure 9: Loss value

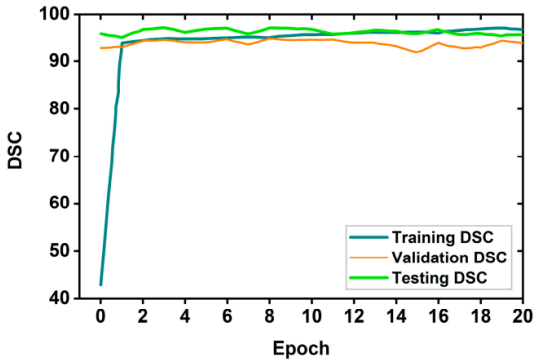

Figure 10: Dice score value

5.2. Performance Analysis of the Proposed Model

Table 2 shows the architectures for BT classification. The existing techniques from (16-22) are applied to our dataset, and the results are then averaged.

Table 2: Analysis of the projected model with existing procedures

Parameters	Training Dataset			Validation Dataset			
	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Run Time (s)
TECNN-DBO	97.25	94.43	96.89	96.63	94.09	96.13	832
ResNet50	96.08	92.51	95.53	95.14	91.88	94.16	986
VGG16	94.12	91.08	93.76	93.21	90.69	92.43	1056
InceptionV3	93.86	90.14	93.06	92.75	89.26	92.12	1008
WBM-DLNets	93.24	89.73	92.58	92.16	88.49	91.29	1161
DBNQLBC	92.86	88.52	91.89	91.64	87.29	90.83	1351
BKNN	92.07	87.39	90.96	91.09	86.73	90.08	1269
CNN	92.18	87.19	90.72	90.86	86.49	89.79	1412
ELM	91.52	86.72	90.81	90.19	86.08	89.45	1096
SVM	91.69	86.53	90.67	90.26	85.98	89.30	1196

In Table 2 above, it is characterised that the Analysis of projected faultless with existing events. In the analysis of the training dataset, the TECNN-DBO model reached a sensitivity of 97.25, a specificity of 94.43, and an accuracy of 96.89. In contrast, the ResNet50 model reached a sensitivity of 96.08, a specificity of 92.51, and an accuracy of 95.53. After the VGG16 model reached sensitivities of 94.12, specificities of 91.08, and accuracies of 93.76. Afterwards, the InceptionV3 model reached a sensitivity of 93.86, a specificity of 90.14, and an accuracy of 93.06, whereas the WBM-DLNets model reached a sensitivity of 93.24, a specificity of 89.73, and an accuracy of 92.58. After the DBNQLBC model reached sensitivities of 92.86, specificities of 88.52, and accuracies of 91.89. Afterwards, the BKNN model achieved a sensitivity of 92.07. Specificity: 87.39; accuracy: 90.96. The CNN model achieved a sensitivity of 92.18%, specificity of 87.19%, and an accuracy of 90.72%. The ELM model achieved a sensitivity of 91.52%, a specificity of 86.72%, and an accuracy of 90.81%. The SVM model achieved a sensitivity of 91.69%, a specificity of 86.53%, and an accuracy of 90.67%. Another, in the validation dataset, the TECNN-DBO model reached sensitivities of 96.63, 96.13, and 94.09, specificities of 94.09 and 96.13, and an accuracy of 832.

The ResNet50 model achieved a sensitivity of 95.14, a specificity of 91.88, a run time of 94.16, and an accuracy of 986. After the VGG16 model reached a sensitivity of 93.21, a specificity of 90.69, and a run time of 92.43, it also achieved an accuracy of 1056. Afterwards, the InceptionV3 model reached a sensitivity of 92.75, a specificity of 89.26, an accuracy of 92.12, and a runtime of 1008. After the WBM-DLNets model reached a sensitivity of 92.16, a specificity of 88.49, an accuracy of 91.29, and a run time of 1161, respectively, the DBNQLBC model reached the sensitivity of 91.64, 87.29, 90.83, and a run time of 1351, respectively. Afterwards, the BKNN model reached a sensitivity of 91.09, a specificity of 86.73, an accuracy of 90.08, and a runtime of 1269. Afterwards, the CNN model reached a sensitivity of 90.86, a specificity of 86.49, a runtime of 89.79,

and an accuracy of 1412. Afterwards, the ELM model reached a sensitivity of 90.19, a specificity of 86.08, an accuracy of 89.45, and a runtime of 1096. Afterwards, the SVM model achieved a sensitivity of 90.26, a specificity of 85.98, an accuracy of 89.30, and a runtime of 1196.

Table 3: Comparison of the procedures rendering to MSE

Techniques	Training Dataset	Validation Dataset
BKNN	0.68395	1.22368
ResNet50	0.00386	0.05856
VGG16	0.13201	0.29632
InceptionV3	0.29872	0.49632
WBM-DLNets	0.52361	0.97526
DBNQLBC	0.59368	1.10395
ELM	0.76589	1.38596
SVM	0.89652	1.42698

In Table 3, the Comparison of the algorithms according to MSE. In the analysis of the TECNN-DBO model, the training MSE was 0.00019, and the validation MSE was 0.00093. The ResNet50 model attained an MSE of 0.00386 during training and 0.05856 during validation. The VGG16 model achieved an MSE of 0.13201 on the training set and 0.29632 on the validation set. Then the InceptionV3 model attained an MSE of 0.29872 during training and 0.49632 during validation. The WBM-DLNets model achieved an MSE of 0.52361 on the training set and 0.97526 on the validation set. The DBNQLBC model attained an MSE of 0.59368 on the training set and 1.10395 on the validation set. The BKNN model attained an MSE of 0.68395 on the training set and 1.22368 on the validation set. The CNN model attained an MSE of 0.70196 on the training set and 1.31856 on the validation set. The ELM model attained an MSE of 0.76589 on the training set and 1.38596 on the validation set. The SVM model attained an MSE of 0.89652 on the training set and 1.42698 on the validation set.

Table 4: Performance of the projected model for various training and testing ratios

Model	Accuracy	Loss	Precision	Recall	F1-Score	AUC
70-30%	0.9102	0.2814	0.9213	0.8980	0.9090	0.9907
60-40%	0.8986	0.1970	0.9131	0.8981	0.9054	0.9944
80-20%	0.9689	0.1846	0.9609	0.8958	0.9271	0.9963

In Table 4, the Presentation of the projected model for various training and testing ratios. For a 70-30% model ratio, the accuracy is 0.9102, the loss is 0.2814, the precision is 0.9213, the recall is 0.8980, and the F1-score is 0.9090; the AUC is 0.9907. Then, with the 60-40% model ratio, the accuracy is 0.8986, the loss is 0.1970, the precision is 0.9131, the recall is 0.8981, the F1-score is 0.9054, and the AUC is 0.9944. Then, with the 80-20% model ratio, the accuracy is 0.9689, the loss is 0.1846, and the precision is 0.9609, along with a recall of 0.8958 and an F1-score of 0.9271; the AUC is 0.9963.

6. Conclusion

This research presents a lightweight Convolutional Neural Network (CNN) engineered for efficient brain cancer classification utilising Reconstructed Microwave Brain (RMB) pictures. The suggested TECNN framework serves as an independent operational neural network that can learn complex patterns with minimal human intervention. The SMBI system created the image data. It has a small, three-dimensional, stacked wideband sensor comprising nine antennas. This setup made it easy to obtain RMB photos, yielding a dataset of 920 samples. To improve the quality and variety of the training data, another RMB dataset from an earlier study was also added. The TECNN-DBO classifier then put the raw RMB images into six groups. This classifier combines the transformer topologies' ability to understand global context with the convolutional neural networks' ability to extract significant spatial features. The model increases network diversity, reduces processing requirements, and improves overall classification performance by introducing nonlinear processes. The proposed approach was shown to work in experiments. The TECNN-DBO model achieved 96.89% accuracy, 96.09% precision, and 92.71% F1-score when trained on 80% of the dataset and tested on the remaining 20%. These numbers show that the model is quite good at detecting tumours using microwave imaging and can identify both small and large features in RMB images. Future research will focus on exploring a more lightweight CNN design to achieve an optimal balance between computational efficiency and classification accuracy. Fine-grained tumour categorisation will be a major focus to increase diagnostic accuracy while still being useful in real-time medical and clinical settings with limited resources.

Acknowledgement: The authors extend their heartfelt appreciation to New Horizon College of Engineering, Quest Technologies, St. Joseph's College of Engineering, and IPS Health for their continuous support throughout this study. Their assistance, infrastructure, and collaborative efforts greatly contributed to the successful completion of this research.

Data Availability Statement: The data supporting this study are available from the corresponding authors upon reasonable request, subject to privacy and ethical restrictions.

Funding Statement: This research received no financial support or external funding.

Conflicts of Interest Statement: The authors declare no conflicts of interest related to this study.

Ethics and Consent Statement: This study was conducted by ethical standards and approved by the relevant institutional review board. Informed consent was obtained from all participants before their involvement in the study.

References

1. S. Gull, S. Akbar, S. A. Hassan, A. Rehman, and T. Sadad, "Automated brain tumor segmentation and classification through MRI images," *Proc. Int. Conf. Emerging Technology Trends in Internet of Things and Computing*, Erbil, Iraq, 2021.
2. A. K. Mandle, S. P. Sahu, and G. Gupta, "Brain tumor segmentation and classification in MRI using clustering and kernel-based SVM," *Biomedical and Pharmacology Journal*, vol. 15, no. 2, pp. 699–716, 2022.
3. S. Raghavendra, A. Harshavardhan, S. Neelakandan, R. Partheepan, R. Walia, and V. C. S. Rao, "Multilayer stacked probabilistic belief network-based brain tumor segmentation and classification," *International Journal of Foundations of Computer Science*, vol. 33, no. 5, pp. 559–582, 2022.
4. E. U. Haq, H. Jianjun, X. Huarong, K. Li, and L. Weng, "A hybrid approach based on deep CNN and machine learning classifiers for tumor segmentation and classification in brain MRI," *Computational and Mathematical Methods in Medicine*, vol. 2022, no. 1, p. 6446680, 2023.
5. S. Gupta, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, "MAG-Net: Multi-task attention guided network for brain tumor segmentation and classification," *Proc. Big Data Analytics*, Allahabad, India, 2021.
6. T. A. Jemimma and Y. J. Vetharaj, "Fractional probabilistic fuzzy clustering and optimization-based brain tumor segmentation and classification," *Multimedia Tools and Applications*, vol. 81, no. 13, pp. 17889–17918, 2022.
7. S. Krishnakumar and K. Manivannan, "Effective segmentation and classification of brain tumor using rough K-means algorithm and multi-kernel SVM in MR images," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 6751–6760, 2021.
8. A. R. Khan, S. Khan, M. Harouni, R. Abbasi, S. Iqbal, and Z. Mehmood, "Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification," *Microscopy Research and Technique*, vol. 84, no. 7, pp. 1389–1399, 2021.
9. T. Tazeen and M. Sarvagya, "Brain tumor segmentation and classification using multiple feature extraction and convolutional neural networks," *International Journal of Engineering and Advanced Technology*, vol. 10, no. 6, pp. 23–27, 2021.
10. F. J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, and D. González-Ortega, "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network," *Healthcare*, vol. 9, no. 2, p. 153, 2021.
11. R. Ranjbarzadeh, A. Caputo, E. B. Tirkolaee, S. J. Ghoushchi, and M. Bendechache, "Brain tumor segmentation of MRI images: A comprehensive review on artificial intelligence tools," *Computers in Biology and Medicine*, vol. 152, no. 1, p.105405, 2023.
12. K. A. Kumar and R. Boda, "A multi-objective randomly updated beetle swarm and multi-verse optimization for brain tumor segmentation and classification," *The Computer Journal*, vol. 65, no. 4, pp. 1029–1052, 2022.
13. E. S. Biratu, F. Schwenker, Y. M. Ayano, and T. G. Debelee, "A survey of brain tumor segmentation and classification algorithms," *Journal of Imaging*, vol. 7, no. 9, p. 179, 2021.
14. P. Sharma and A. P. Shukla, "A review on brain tumor segmentation and classification for MRI images," *Proc. Int. Conf. Advance Computing and Innovative Technologies in Engineering*, Greater Noida, India, 2021.
15. S. Vadhnani and N. Singh, "Brain tumor segmentation and classification in MRI using SVM and its variants: A survey," *Multimedia Tools and Applications*, vol. 81, no. 22, pp. 31631–31656, 2022.
16. M. A. Khan, A. Khan, M. Alhaisoni, A. Alqahtani, S. Alsubai, M. Alharbi, N. A. Malik, and R. Damaševičius, "Multimodal brain tumor detection and classification using deep saliency map and improved dragonfly optimization algorithm," *International Journal of Imaging Systems and Technology*, vol. 33, no. 2, pp. 572–587, 2023.
17. C. Qin, B. Li, and B. Han, "Fast brain tumor detection using adaptive stochastic gradient descent on shared-memory parallel environment," *Engineering Applications of Artificial Intelligence*, vol. 120, no. 4, p. 105816, 2023.

18. S. Hossain, A. Chakrabarty, T. R. Gadekallu, M. Alazab, and M. J. Piran, “Vision transformers, ensemble model, and transfer learning leveraging explainable AI for brain tumor detection and classification,” *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 3, pp. 1261 – 1272, 2023.
19. M. U. Ali, S. J. Hussain, A. Zafar, M. R. Bhutta, and S. W. Lee, “WBM-DLNets: Wrapper-based metaheuristic deep learning networks feature optimization for enhancing brain tumor detection,” *Bioengineering*, vol. 10, no. 4, p. 475, 2023.
20. P. K. Ramtekkar, A. Pandey, and M. K. Pawar, “Innovative brain tumor detection using optimized deep learning techniques,” *International Journal of System Assurance Engineering and Management*, vol. 14, no. 1, pp. 459–473, 2023.
21. V. V. Kumar and P. G. K. Prince, “Deep belief network assisted quadratic logit boost classifier for brain tumor detection using MR images,” *Biomedical Signal Processing and Control*, vol. 81, no. 3, p. 104415, 2023.
22. K. V. Archana and G. Komarasamy, “A novel deep learning-based brain tumor detection using the bagging ensemble with K-nearest neighbor,” *Journal of Intelligent Systems*, vol. 32, no. 1, pp. 1–13, 2023.
23. A. M. Mostafa, M. A. El-Meligy, M. A. Alkhayyal, A. Alnuaim, and M Sharaf, “A framework for brain tumor detection based on segmentation and feature fusion using MRI images,” *Brain Research*, vol. 1806, no. 5, p. 148300, 2023.
24. A. Hossain, M. T. Islam, M. E. Chowdhury, and M. Samsuzzaman, “A grounded coplanar waveguide-based slotted inverted delta-shaped wideband antenna for microwave head imaging,” *IEEE Access*, vol. 8, no. 10, pp. 185698–185724, 2020.
25. Y. Cheng and M. Fu, “Dielectric properties for non-invasive detection of normal, benign, and malignant breast tissues using microwave theories,” *Thoracic Cancer*, vol. 9, no. 4, pp. 459–465, 2018.
26. M. T. Islam, M. Samsuzzaman, S. Kibria, N. Misran, and M. T. Islam, “Metasurface loaded high gain antenna based microwave imaging using iteratively corrected delay multiply and sum algorithm,” *Scientific Reports*, vol. 9, no. 11, p. 17317, 2019.
27. A. Hossain, M. T. Islam, and A. F. Almutairi, “A deep learning model to classify and detect brain abnormalities in portable microwave-based imaging system,” *Scientific Reports*, vol. 12, no. 4, p. 6319, 2022.
28. P. Gu, W. M. Lee, M. A. Roubidoux, J. Yuan, X. Wang, and P. L. Carson, “Automated 3D ultrasound image segmentation to aid breast cancer image interpretation,” *Ultrasonics*, vol. 65, no. 10, pp. 51–58, 2016.
29. Z. Peng, W. Huang, S. Gu, L. Xie, Y. Wang, and J. Jiao, “Conformer: Local features coupling global representations for Visual Recognition,” *Proc. IEEE/CVF Int. Conf. Computer Vision*, Montreal, Quebec, Canada, 2021.
30. X. Lu, H. Sun, and X. Zheng, “A feature aggregation convolutional neural network for remote sensing scene classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 7894–7906, 2019.
31. H. Yar, T. Hussain, Z. A. Khan, M. Y. Lee, and S. W. Baik, “Fire detection via effective vision transformers,” *Journal of Korean Institute of Next Generation Computing*, vol. 17, no. 5, pp. 21–30, 2021.