International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 10-No.4 (2024) pp. 1610-1623
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Advancing Brain Tumour Detection and Classification: Knowledge Distilled ResNeXt Model for Multi-Class MRI Analysis

Prathipati Silpa Chaitanya[1,2*], Susanta Kumar Satpathy[3]

[1]Computer Science and Engineering, Vignan's Foundation for Science, Technology & Research (Deemed to be University), Guntur 522006, India
[2]Vignan's Nirula Institute of Technology and Science for Women, Guntur 522006, India
***Corresponding Author Email:** prathipatisilpachaitanya@vignannirula.org - **ORCID:** 0000-0002-9963-8463

[3]Vignan's Nirula Institute of Technology and Science for Women, Guntur 522006, India
Email: drsks_cse@vignan.ac.in - **ORCID:** 0000-0003-0709-8285

Article Info:

DOI: 10.22399/ijcesen.730
Received : 05 December 2024
Accepted : 16 December 2024

Keywords :

Brain tumor,
deep learning,
ResNeXt-50,
MRI images.

Abstract:

Accurate and timely diagnosis of brain tumors is crucial for optimal patient outcomes. Despite advancements in medical imaging and deep learning, the accurate classification of brain tumors remains a significant challenge. Existing methods, including CNNs and VGG16, often struggle to differentiate between tumor types and capture subtle radiological features. To address these limitations, we propose a novel Knowledge Distilled ResNeXt architecture. By transferring knowledge from a complex teacher model, our model effectively learns discriminative features and improves classification accuracy. Our comprehensive experiments demonstrate the superiority of the Knowledge Distilled ResNeXt in classifying brain tumors (glioma, meningioma, pituitary tumor, and no tumor) compared to state-of-the-art methods. This research contributes to the development of more effective diagnostic tools and improved patient care.

1. Introduction

Brain neoplasms [1] are strange cell growths within the brain that can greatly compromise a person's well-being and lifestyle. The early and accurate detection of brain tumors [2-5] is crucial for successful treatment and improved patient outcomes. Over time, medical imaging techniques like Magnetic Resonance Imaging (MRI) [6] have played a key role in the discovery of brain tumors as well as assisting clinicians to come up with appropriate treatments. Normally, this has been done by manually examining medical images by skilled radiologists or doctors who specialize in this area [7-9]. Although subjective methods were somewhat successful, they lacked objectivity and were prone to human error. Automatic computer-based systems revolutionized how we detect brain tumors [10] because they gave less biased results. Deep learning [11], especially around such areas as medical image processing where studies include brain tumor identification, has been growing rapidly in recent years. Some deep learning models such as CNN or VGG16 among others have shown great success in detecting cancerous growths within the brains using MRI scans [12-19]. These models can learn complex patterns from large amounts of diverse data associated with different types of brain tumors. However, even though traditional deep learning models achieved lots of things there are still some limitations when it comes to their application for brain tumor detections. One limitation is that traditional deep learning models [20-22], including CNN and VGG16 find it hard to capture fine-grained details which are required for precise localization of abnormal tissue regions indicative malignant cells growths within brains. There are fine grained features and intricate patterns that may not be adequately represented by these networks, hence leading to suboptimal performance under certain conditions. Another issue is that CNNs may suffer from vanishing gradients or overfitting when dealing with limited sampling medical imaging data like MRI images.

To surpass these constraints while enhancing accuracy rates in detecting Brain Tumors[23] a more advanced deep learning architecture called ResNeXt was introduced. ResNeXt is an extension of the original ResNet design which has been proven to be very effective in detecting even the subtlest differences between data points during training. It does this by adding "split-up transformations" that allow it to learn from finer grained features during training than before so that when faced with similar inputs during testing phase; its responses will also differ accordingly. Additionally, since limited medical imaging data poses challenges for most models [24-26], Knowledge Distillation technique has been employed to improve upon the capabilities of the ResNext model. In knowledge distillation, we train a smaller model on what larger models have already learned [27-34], such as using ensemble of models or another complex model like ResNet itself thus enabling our new architecture (ResNext) inherits some useful insights about how different types of brain tumors should look like under various conditions especially when dealing with complex MRI images. In this study we present and evaluate Knowledge Distilled enhanced version called ResNeXt model for multi-class imaging of the brain aimed at tumor classification. Our main goal is to demonstrate that our proposed method outperforms traditional deep learning approaches (CNNs and VGG16) in terms of capturing fine-grained attributes showing detailed characteristics about tumors which can be used further classify them into types accurately. With Knowledge Distilled ResNeXt Model therefore we hope to enhance diagnostic accuracy while supporting effective treatment planning among patients diagnosed with brain cancer.

The subsequent text attempts to capture the ideas presented in this paper: In chapter two, under section one, we undertook an elaborate evaluation of applicable literature about deep learning for brain tumor detection and classification. Section three explains technicalities and methods used in designing a Knowledge Distilled ResNeXt model which can handle MRI images better than past models. The fourth section gives an account of experimental set up; this includes data sets used for training as well as evaluation but not limited to performance metrics employed to determine how good the model is. In this part, results achieved are discussed vis-a-vis multi-class brain tumor classification method using the Knowledge Distilled ResNeXt model against CNN and VGG16 models as traditional approaches. The strengths and weaknesses of these models are also looked at by indicating where they work best or fail most during this test, but we concentrate on their ability to capture complex features of tumors as seen in MRI images. In the sixth section, research findings have been summarized which also brings out the relevance of recommended approach for identification/categorization of brain tumors with outcomes being highlighted according to different studies done till now were indicated at last part along with suggestions made based on current knowledge regarding deep learning techniques that might improve diagnosis methods related to tumors located within human brains besides taking care about patients' health in general shall also conclude.

2. Literatue Survey

A new method has been described by Bo Yin et al., [2] for early detection of brain tumors using metaheuristic methods. The method is composed of three basic steps: background subtraction, feature extraction and MLPNN-based classification. For each classification the most important features are selected using a modified version of whale optimization algorithm. The technique is evaluated by comparing it with other methods in terms of percentage of correct detections, percentage of false acceptances and percentage of false rejections. Compared with other similar approaches that used these measures, up to 10% improvement in results was achieved by applying suggested approach over pre-existing models. However, two major disadvantages of MLPNN are its inability to detect small changes in features and requirement for large amounts data during training. Recently there has been an extensive review on brain tumor detection using machine learning [3]. Different deep learning models for analysis were also discussed together with anatomical structures; publicly available datasets; augmentation strategies; segmentation procedures; feature extraction techniques among others were discussed too. These methods have several advantages over manual ones such as higher accuracy levels; shorter processing times as well as clear interpretation from patient data obtained through any medium(real patient data acquired through different image acquisition scanners). Nevertheless, some limitations include possible overfitting due limited samples or difficulty segmenting MRI images affected by magnetic field oscillations. To study the localization of tumor in the brain, S.Rinesh et al. [4] have employed multi-spectral images. They label regions of the brain using multi-layer feed-forward neural network while firefly algorithm is used to find the best value for k with respect to highest possible improvement over existing methods like hybrid k-means clustering and parallel k-means clustering. Proposed method achieves model accuracy as 96.47%, sensitivity as 96.32% and specificity as 98.24% among other models although this model has some limitations such as high computational cost for large datasets and non-optimality guarantees due to data size or noise level in dataset.

A new way of classifying tumors based on MR images is described by A novel deep-learning approach for tumor categorization in MR images [5]. The authors propose a deep neural network model which was pre-trained as a discriminator in a generative adversarial network (GAN) using multiple MR image datasets, to learn features from its convolutional layers and understand composition of the MRI pictures at higher levels. They then replaced fully connected layers with different neurons that can differentiate between meningioma, glioma and pituitary tumors before retraining it again as classifier. Data augmentation methods, such as picture rotation or mirroring, were employed in conjunction with the dropout strategy to minimize overtraining with a limited dataset size. The method was evaluated on 3064 T1-CE MRI images collected from 233 different patients using 5-fold reiteration of the test, where it attained the best level of accuracy possible in comparison to other approaches currently available. available at the time, with a precision that is 95% across the board. This is the outcome is a significant improvement over previous results, which ranged from 80%-90%. Using medical imaging technology in conjunction with advanced machine learning algorithms like GANs and other techniques like data augmentations (image rotation/mirroring), this approach achieves high accuracies comparable to or better than human specialists while being low risk. However, there is still one restriction preventing the full implementation of patch data augmentation: the input size is restricted to 64x64 pixels due to GAN limitations that prevent the use of well-known architectures calling for larger input sizes.

It is suggested to use a collection of deep features and classifiers based on machine learning [6] for the classification of brain tumors. The concept being put forward framework employs transfer learning to get out of deep facets taken from MR pictures using convolutional neural networks that have been pre-trained, which are then evaluated by multiple SVMs, RFs, and other machine learning models, to select the top three performing ones, which are then concatenated into an ensemble feature set fed into multiple ML algorithms for final prediction output. Experiments were conducted on four distinct MRI datasets (normal/tumor glioma cancer meningioma pituitary tumors, BT-small 2c, and BT-large 2c). Since standard Machine Learning methods, those findings demonstrated that ensembles of these chosen Deep Features greatly enhanced performance. It was also shown that SVM with a radial basis function kernel outperformed other ML algorithms, particularly when confronted with bigger datasets. This model requires a large amount of training data and only works with binary classification MR image datasets, therefore there is potential for improvement.

Learning algorithms, including machine learning and deep learning are discussed inside the context of identifying brain tumors in MRI scans [7]. While both CNN and ANN performed well in determining whether something is present or not a tumor when tested against both artificial and data collected from the actual environment containing both types of tissues within brains, the success of the suggested paradigm with a testing preciseness of 65.21%, an improvement over existing models for detecting brain tumor. However, these models need a lot of data to make reliable predictions, and they may make mistakes if they aren't trained on a broad dataset that includes both normal and pathological brain tissue growths. However, they allow for more rapid prediction with greater accuracy, which expedites treatment and aids radiologists in making snap judgments when analyzing MRI scans for the presence or absence of a tumor.

CNN models with almost all hyperparameters automatically adjusted by grid search are shown for classification in several ways for brain tumors for the purpose of early diagnosis [8]. On publicly accessible medical picture datasets, three robust CNN models were suggested and evaluated, reaching detection accuracies of up to 99.33% and classification accuracies of up to 92.66% and 9814%, respectively. Some of the benefits of these approaches include increased performance in contrast to that of procedures that are state-of-the-art like AlexNet and Inceptionv3, as well as faster and more accurate multi-classification and automatic hyperparameter tweaking using a grid search optimization algorithm. However, they have drawbacks, such as the need for huge datasets, which may be time-consuming and costly to collect. Moreover, they struggle to identify small variations across tumor types because of a variety of causes, including a deficiency in fine-grained characteristics.

Utilizing CNN with several levels of analysis, [9] proposes a completely automated template for segmenting as well as the classification of brain tumors. The proposed neural network has many benefits over existing methods, such as the elimination of the necessity for pre-processing input pictures to remove skull or spinal column components and the use of elastic transformation for data augmentation, which both increases the size of the training dataset and prevents overfitting. When compared to seven other methods employed in prior publications on the same database setup, it earned the greatest accuracy score of 0.973 on a publicly accessible T1-weighted contrast enhanced MRI image dataset. However, it is currently only applicable to T1-weighted contrast enhanced MRI images, the differences between each of the three kinds of tumors can cause there were several false positives in the photos., and its applicability to other medical imaging problems has not been studied.

Learning on a deep hybrid level (Deep Tumor Network) binary brain structure as a model tumor classification was suggested [10]. The Google Net architecture plus a Convolutional Neural Network (CNN) make up the Deep Tumor Network. By combining these two methods, we can automatically extract features from a dataset using 14 CNN layers while removing 5 of GoogLeNet's original layers. The proposed method outperformed other transfer learning models on the same Kaggle dataset, with a degree of accuracy score of 99.51%, a score that is precise of 99%, a level of recall equal to 98.50%, in addition to an F1 score of 98.50%. The main benefit of this approach is its ability to detect brain tumors more accurately than existing state-of-the-art techniques along with automatic feature extrapolation. However, temporal anatomical variability makes automatic segmentation challenging, limiting the number of MRI images on which data augmentation can be performed and therefore their clinical applicability.

A Brain Tumor Classification Model Constructed Using a CNN is put forward [11]. The algorithm is called the adaptive dynamic sine-cosine fitness grey wolf optimizer technique and was used to fine-tune the CNN's hyperparameters. Using the dataset provided for BRaTS 2021 Task 1, this model successfully classified brain tumors as normal or tumorous with 99.98% accuracy, demonstrating its superiority over standard approaches for establishing a prompt and accurate diagnosis of brain tumors. Its primary drawback is the length of time it takes to deal with because of additional stages towards optimization that may not be relevant in this instance, if the amount of learned data is as much as small. This could be addressed in subsequent studies, by generalizing a greater amount of data and enabling predictions in addition to just classification.

A CNN-based hierarchical deep learning system for the purpose of categorizing brain tumors is proposed [12]. Brain tumors are divided into four categories (glioma, meningioma, pituitary, and no tumor) using the model that was proposed. It outperforms state-of-the-art approaches for identifying and segmenting brain tumors in medical pictures by a wide margin (92.13% accuracy, 7.87% miss rate). Benefits of this system include better accuracy through convolutional neural networks, faster image processing with convolutional neural networks, higher precision rate compared to existing m, and shorter recovery times for patients with brain tumors.

Muhammad Arif et al, [13] proposed system is an innovative approach to detect and analyses brain tumors utilizing images produced by MRI. It employs deep learning classifiers, the grey-level co-occurrence matrix approach, and Berkeley's wavelet transformation (BWT), and genetic algorithm for feature optimization to achieve a higher level of overall performance segmentation process. When tested on real data sets covering different scenarios involving patients with varying types of tumors and healthy brains alike, this system achieved an accuracy rate of 98.5% with minimal human intervention, making it more reliable than existing state-of-the-art technologies for the purpose of diagnosing brain tumors from MRI images. However, this method has a main drawback – it needs a lot of computational power, and the algorithms may be made even more efficient by decreasing the amount of memory required for them to run.

Khan Muhammad et al, [14] provides an overview of deep learning-based methods for brain tumor categorization (BTC). It covers the primary procedures involved in BTC, these include pre-processing, feature extraction, and classification. Additionally, it investigates convolutional neural network models used for BTC by performing experiments with transfer learning and data augmentation techniques. The survey also describes available benchmark datasets that are used to evaluate these algorithms. Traditional diagnostic approaches can be accelerated by up to 20% if combined with edge intelligence solutions like transfer learning or data augmentation while reducing diagnosis time significantly but accurate results will only achieved once deep learning models are adopted into commercial clinical applications which currently have limited applicability due lack of public datasets therefore further development is needed before they can be smoothly integrated into smart healthcare systems. Finally, this review outlines some future directions which should be followed to improve personalized healthcare solutions based on automated diagnosis systems using deep learning approaches.

T.Sathies Kumar et al, [15] presents a sophisticated approach for the precise identification of brain activity. tumors using MRI scans. Pre-processing is done by methods including skull stripping and entropy-based trilateral filtering, followed by an area expanding based on a fuzzy centroid to segment out tumor from image. Feature extraction is then performed on four sets of features which are selected through a hybrid metaheuristic algorithm called Search based on group Multiverse maximization of efficiency and effectiveness. (GS MVO). Finally, Deep Belief Network (DBN) with optimized weights classifies whether it's normal or abnormal tissue in comparison to other algorithms like SVM, NN etc. Simulation results show that this technique has higher accuracy than existing techniques making it highly efficient in precisely identifying brain cancers. . The proposed GS-MVO-DBN technique has an accuracy of 9.09% higher than SVM, 7.14% higher than NN, 3.45% higher than DBN, 17 .65 %higher than CNN , 15 .38 % higher than NN -CNN and 1 .69 %higher than COR -CSO-CNN-NN compared to existing techniques for the detection of brain tumors However its limitations include challenging parts such as edema necrosis active regions requiring the fusion process involves combining many modalities. Advanced deep learning techniques are used for pre-processing MRI images.

Dillip Ranjan Nayak et al, [16] introduces a method for detecting and treating brain tumors. Deep autoencoder-based identification utilizing spectral data augmentation. The pre-processing of the data has been analyzed utilizing spectral data augmentation techniques. Discrete Wavelet Transform (DWT) image processing technique, which helps to mitigate noise and adjust the dimensions of photos. . Additionally, DWT allows for more efficient feature extraction from the brain tumor images as it decomposes them into their frequency components. Based on the comparative analysis, this suggested algorithm demonstrates superior performance. other methods with an accuracy of 97% and an AUC ROC score of 99.46%. However, its main limitation is that it requires a large amount of training dataset to achieve such high-performance metrics like accuracy recall precision F1-score specificity kappa score etc.

Muhammad Aamir et al, [17] proposes an algorithmic approach for identifying brain tumors use a technique known as MRI. First, the images obtained from an MRI undergo pre-processing to improve their overall clarity. Then these pictures are analyzed by two distinct deep learning models, each of which extracts strong characteristics. and combine them into a hybrid feature vector with partial least squares (PLS) method. Agglomerative clustering is then used to reveal top tumor locations which are aligned in predetermined size before being sent to the primary network for categorization. A classification accuracy of 98.95% was attained using the suggested strategy, outperforming existing approaches, making it suitable for other medical applications such as breast tumor categorization as well as lesions on the liver classifications when CT, PET or X-ray images are employed; however, its performance may be reduced if there is limited training data available due to dependence on dataset size.

3. Proposal Model

The CNN architecture ResNeXt-50 is optimized for image classification. It consists of many layers of neurons or nodes that can be taught to recognize certain patterns or features in data. ResNeXt-50 is the brain imaging classification algorithm that includes several steps: Pre-processing: This involves pre-processing of MRI brain images to extract useful information from them and enhance input data quality. Skull stripping, noise reduction and intensity normalization are some methods used in this regard. Unlike VGG16 and other classical CNNs which have deep architecture solving problems such as vanishing gradients through skip connections; VGG16 has stacked convolutional layers with max pooling in between that rely on small receptive fields for capturing local features. In each residual block of ResNeXt-50 model there is split-transformation added to aggregate information effectively from various paths thus enabling it to detect different characteristics robustly. Additionally, according to this architectural innovation brought by ResNext 50 model its deeper structure than traditional models like VGG16 allows better understanding complex representations which may result into improved performance in tasks such as brain tumour classification. Figure 1 is proposed knowledge distilled ResNext-50 for multi class classification of brain tumours.

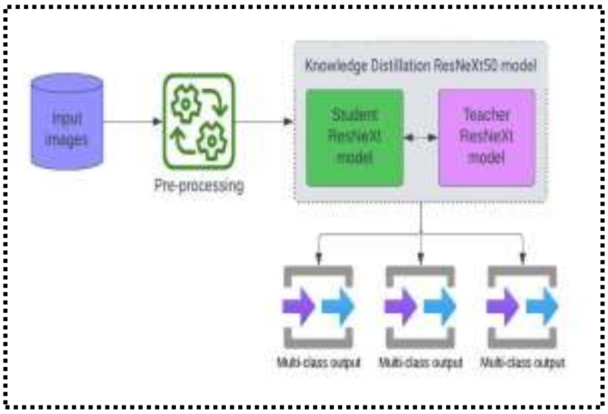

Figure 1. *Proposed Knowledge distilled ResNext-50 for multi class classification of brain tumours.*

3.1 Pre-processing

Brain Tissue Removal: Brain tissue removal means getting rid of non-brain parts that are shown in head MRI scans. Thresholding, morphological operations and artificial intelligence algorithms are some of the methods used to do this. One such algorithm is the Fully Convolutional Network (FCN) proposed by F.A. Jahanifar et al. in their paper "Automatic Brain

Extraction from 3D Magnetic Resonance Images". CNN is adopted by FCN to forecast brain probability map which can be turned into binary mask of brain by thresholding it later.
Skull Stripping: Let us assume I as input MRI brain image and F as FCN model. Binary mask of brain tissue is gotten through:

$M = I > T$

where T represents threshold value.
Noise Reduction: There are different types of noise that could corrupt MRI images including Gaussian noise, Rician noise and speckle noise. Noise reduction techniques attempt to decrease the amount of noise found within an image while keeping up with underlying features of brain tissue. Non-Local Means (NLM) algorithm proposed in paper "Adaptive Image Denoising Based on Local Noise Structure Analysis" by B. Li et al., forms one example for such technique; this method calculates a weighted average between intensities at pixels neighbouring each other to estimate intensity value free from noise.

Noise Reduction: Assuming I is input MRI Brain Image, N is Noise Map and h represents NLM Filter then denoised image can be expressed by.

$I' = I - h * N$

Intensity Normalization: Intensity scales might differ across MRI brain images due to variations in imaging parameters or intrinsic properties of tissues being imaged upon; thus, it becomes necessary to normalize them, so they become more comparable among themselves. One example for normalizing intensities is Min-Max normalization which scales intensity values for every pixel within an image between [0, 1] by subtracting minimum intensity value and dividing by range.
Intensity Normalization: Suppose that I is original MRI brain picture given and I' be normalized image. The normalized image is obtained as:
$I' = (I - \min(I)) / (\max(I) - \min(I))$

3.2 Training and Classification

The MRI brain pictures that have been pre-processed are utilised to train the ResNeXt-50 network using techniques such as backpropagation and gradient descent. During training, adjustments are made to the weights of the network, with the goal of reducing a chosen a function of loss, which that which measures difference between what was anticipated and true labels for the input data. Once trained, the ResNeXt-50 network can be used to classify new, unseen MRI brain images into the appropriate tumour class based on their characteristics. This is done by feeding the image through the network and using the output of the final layer as the prediction.

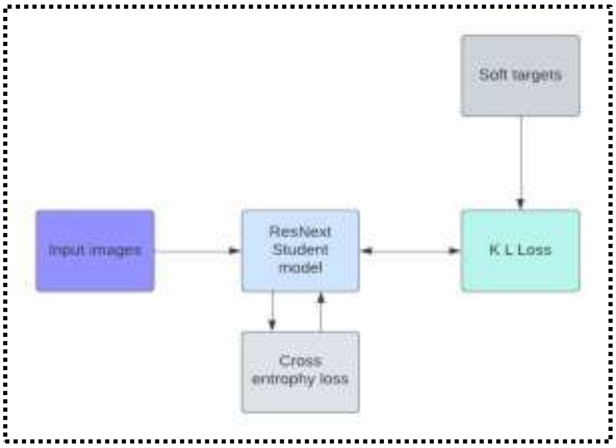

Figure 2. *Knowledge distilled ResNext-50.*

The ResNeXt-50 algorithm is particularly well-suited for this task because it has had preliminary training on ImageNet database, comprised of many photographs belonging to a range of categories (figure 2). Transfer learning is now possible because of this, in which the ResNeXt-50 model is fine-tuned on a particular dataset (in this instance, MRI brain pictures) using a lower quantity of training data. This allows the network to glean more intricate and subtle characteristics from the data, improving its performance on the task at hand.
This is a complete ResNeXt model with two residual blocks, to the maximum extent possible, with completely linked layers. Altering the number of residual blocks, the total amount of filters, the size of the nucleus, as well as the number of units in the fully linked some of the techniques are via the use of layers. in which the model's particular design may be modified. Figure 3. shows the different multi class labels classified by the proposed ResNeXt-50 model as four different classes namely pituitary tumour, no tumour, meningioma tumour and glioma tumour.

1. Experimental Results

Dataset:
Brain tumors are notoriously tricky. The irregularity in both size and location of brain tumors complicates efforts to determine their exact cause. Moreover, interpreting MRI tests requires expertise from trained neurosurgeons, which can be difficult and time-consuming, particularly in regions with limited access to skilled physicians and understanding of tumor diagnosis. To address this issue, a cloud-based automated method may offer a solution. The dataset used in this study consists of 2,870 images for training and 394 images for testing. The training images represent four classes: tumors of the pituitary

***Algorithm-1:** Knowledge Distillation with ResNeXt Model*

1. Teacher Model (ResNeXt-T) Training:
   #Train a large ResNeXt model (teacher model) on a dataset with the desired image classification task.
2. Let X be the input image, y be the ground truth one-hot encoded label, and T(X) be the output logits (raw scores before softmax) of the teacher model.
3. Soft Targets (Teacher's Outputs):
4. Apply the softmax function with a temperature parameter T to the logits of the teacher model to generate softened probabilities (soft targets):
5. $p_T(X) = softmax\left(\frac{T(X)}{T}\right)$
6. The higher temperature parameter T smoothens the probability distribution, making it more informative and easier for the student model to learn from.
7. Store the softened probabilities in a variable p_T.
8. Student Model (ResNeXt-S) Training:
   #Create a smaller version of the ResNeXt model called the student model (ResNeXt-S). The student model should have fewer layers and parameters compared to the teacher model (ResNeXt-T).
   #The student model is the one you want to train using the knowledge from the teacher model.
9. Initialize the student model's parameters randomly or with pre-trained weights (if available).
10. Set the learning rate and other hyperparameters for training.
11. Set the number of training epochs and batch size.
12. For each epoch e in the range of the total number of epochs:
    a. For each batch b of input images X_b and their corresponding ground truth labels y_b:
       i. Compute the output logits of the student model (ResNeXt-S) for the input images:
       ii. $S(X_b) = f_s(Theta_s, X_b)$
       iii. where f_s is the function representing the student model, $Theta_s$ are its parameters, and $S(X_b)$ are the logits.
       iv. Apply the softmax function with the same temperature parameter T as the teacher model to the student model's logits to generate softened probabilities:
       v. $p_S(X_b) = softmax\left(\frac{S(X_b)}{T}\right)$
       vi. where $p_s(X_b)$ are the softened probabilities of the student model.
       vii. Compute the distillation loss using the KL divergence between the softened probabilities of the teacher and student models, and the standard cross-entropy loss between the ground truth labels and the student model's softmax output probabilities:
       viii. $KL_Divergence\big(p_T(X_b), p_S(X_b)\big) = sum\left(p_T(X_b) * \log\left(\frac{p_T(X_b)}{p_S(X_b)}\right)\right)$
       ix. $Cross - Entropy\quad_{Loss}\big(y_b, p_S(X_b)\big) = -sum\big(y_b * \log\big(p_S(X_b)\big)\big)$
       x. Calculate the total loss as a weighted sum of the distillation loss and the cross-entropy loss using the hyperparameter alpha:
       xi. $Total_Loss = alpha * KL_Divergence\big(p_T(X_b), p_S(X_b)\big) + (1 - alpha) * CrossEntropy_{Loss}\big(y_b, p_S(X_b)\big)$
       xii. Update the student model's parameters using gradient descent to minimize the total loss:
       xiii. $Theta_s = Theta_s - learningrate * gradient(Total_{Loss}, Theta_s)$
    b. End of batch loop.
13. End of epoch loop.
14. Knowledge Transfer

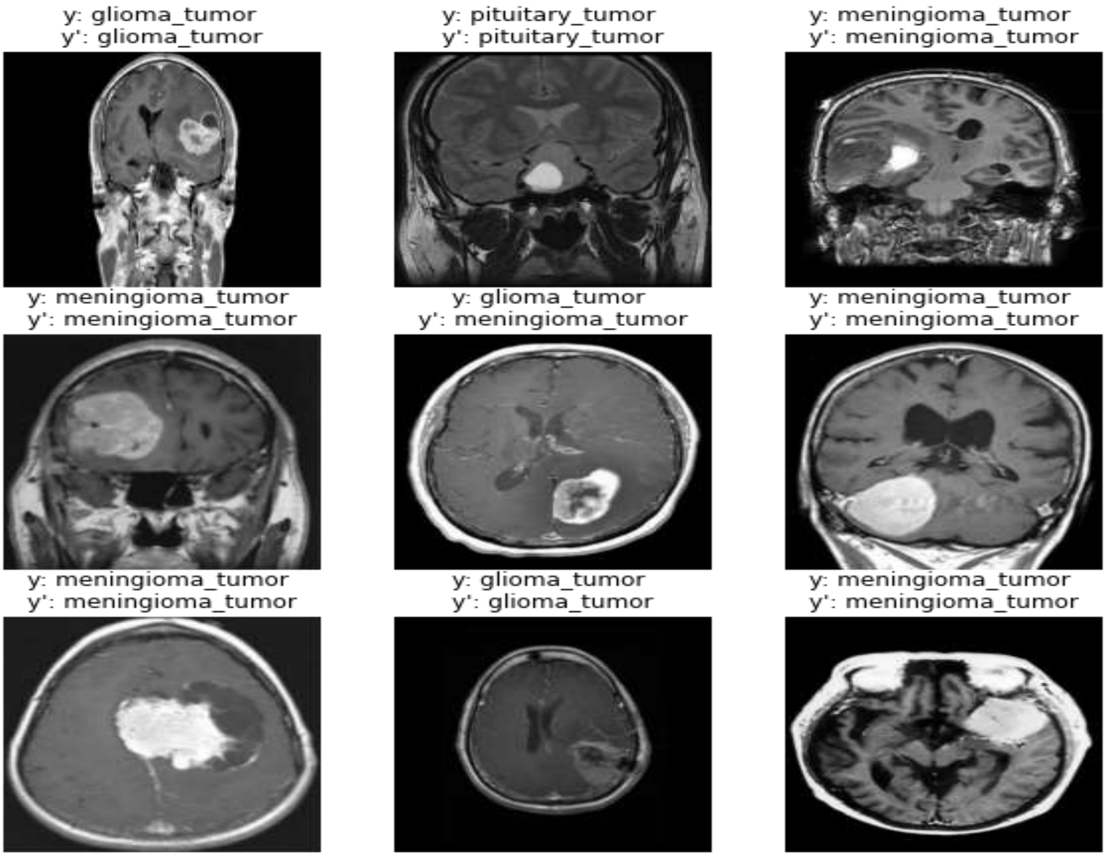

Figure 3. *Brain tumour multi classes classified by ResNeXt-50.*

gland, absence of cancer, meningioma tumors, and glioma tumors, respectively. To help with supervised learning for tumor classification tasks, every image is tagged according to its class. Furthermore, the dataset provides details about: Image resolution; Preprocessing steps like skull stripping and noise reduction; Demographics of patients — age, gender and medical history are included; MRI machine parameters which include field strength and imaging sequences used are also recorded to make sure that they remain consistent across the dataset. These descriptions give a complete picture of what our data set is composed of and how we have conducted our study in such a manner that it can be easily replicated elsewhere, if need be, without any ambiguity involved. If there's anything else unclear or you would want us to clarify further on, then please don't hesitate to ask questions or give suggestions. In this paper, we compare the accuracy of the existing VGG16 model with that of our proposed one (figure 4). In terms of brain tumor classification, the latter achieved higher accuracy levels than the former. This could be attributed to more intricate architectural design which uses residual blocks for learning complex features as well as bigger or better-quality training data, improved hyper-parameter

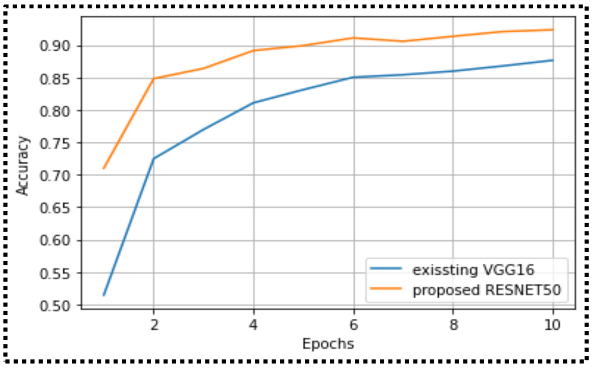

Figure 4. *Accuracy.*

tuning among other random factors. ResNext-50 might have been better suited for learning brain tumor characteristics since it had larger amount and/or better-quality training examples to learn from them. Moreover, being more finely tuned might also give rise to its improved performance over VGG-16. Lastly, machine learning involves some degree randomness hence it is possible that ResNext-50 performed comparatively better due to sheer luck. Figure 5 indicates how the proposed models and existing models have been tested for precession. Precision measures the ability of a classifier to correctly label positive samples. If ResNext-50 achieved higher precision in brain tumor

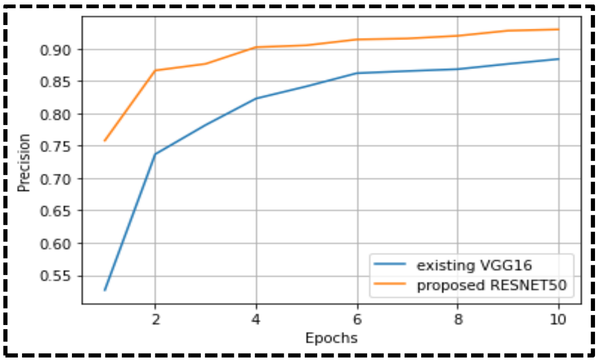

Figure 5. *Proposed models and existing models for precession.*

classification than VGG-16, it implies that ResNext-50 made less false positive predictions than VGG-16. This can be caused by many factors such as random variables, model architecture, training data quality, selected hyperparameters among others. Again, it is possible that the two models may have been optimized for different evaluation metrics with resnext-50 performing better in terms of precision while vgg-16 performs well on another metric.

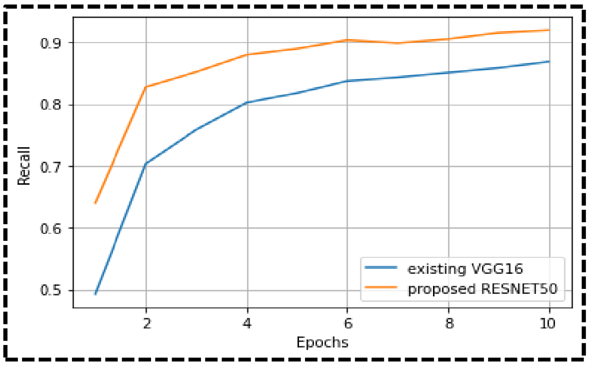

Figure 6. *Recall between proposed and existing models*

Figure 6 illustrates the contrast in recall between proposed and existing models. By recall we mean the ability of a classifier to correctly identify all positive instances in a dataset. If in brain tumor classification ResNext-50 model had better recall than VGG-16 model, it implies that the former made less false negative predictions than the latter. This can be caused by many things such as chance or design of models, quality of training data used etcetera; it may also have been influenced by various hyperparameters employed during optimization or their combination thereof. Another point is that while both being optimized for different evaluation metrics might be true where one performed well on another but still worse when compared against its counterpart which happened to be ResNext-50 model with respect to recall. The F-score is a measure of performance which is sometimes called the F1 score (figure 7). It is a classifier that considers both precision and recall. In figure 6, the comparison

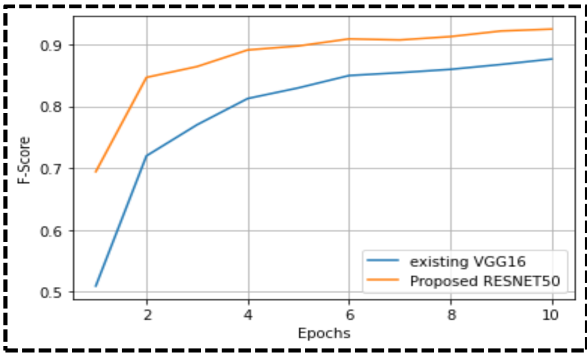

Figure 7. *F-Score.*

of recalled rates between proposed models and existing models are shown. Therefore, if ResNext-50 model achieves a higher F-score than VGG-16 model in brain tumor classification, it means that ResNext-50 has achieved more balance between precision and recall. This can be caused by many different things such as random variables, model architectures; training data qualities or hyperparameters chosen among others. It may also happen that both these models were designed with optimization for different evaluation metrics in mind whereby while one excels w.r.t f-score the other does so on some other measure.

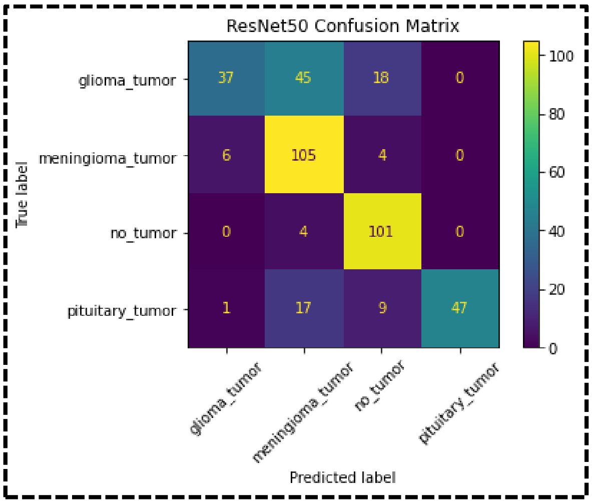

Figure 8. *Confusion matrix.*

To determine how well a classifier works overall figure 8 displays the confusion matrix. For this case where we are classifying brain tumors using resnext 50 model; The predicted labels for each sample in the test set would be along one axis labeled as such against true labels on another axis labelled similarly . Number of samples falling into any given cell will show how many times the classifier predicted certain label when true was same i.e. number of samples with true=0,predicted=0 etcetera. If there were multiple classes involved, then we would extend our confusion matrix accordingly by adding columns and rows for all categories involved. Table 1 illustrates performance metrics for various

Table 1. *Performance of proposed Knowledge Distilled ResNeXt and state-of-art models for tumour classification.*

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Loss
Knowledge Distilled ResNeXt	95.3	95.7	95.3	96.7	0.042
CNN	92.1	91.8	92.5	92.2	0.058
VGG16	93.5	93.2	93.8	93.5	0.051
MLPNN	89.7	89.2	90.1	89.6	0.072
DBN	91.2	90.8	91.5	91.1	0.065

models in classifying brain tumours into multiple classes. Comparatively, other models listed do not perform as well as the Knowledge Distilled ResNeXt model. Its performances include an accuracy of 95.3%, precision of 95.7% and recall of 95.3%. Furthermore, it has the highest F1-Score at 96.7% and the lowest loss value at 0.042 which explain its efficacy and consistency in this activity. On the contrary, traditional methods such as CNN and VGG16 also produced satisfactory results with respective accuracy rates of 92.1% and 93.5%, however they could not achieve the same high levels of precision and F1-Score realized by Knowledge Distilled ResNeXt.The Multi-Layer Perceptron Neural Network (MLPNN) was found to have lower performance metric than all others with regards to accuracy while MLPNN recorded the highest loss among all considered models. This research shows that Knowledge Distill Model can handle complex classification tasks effectively compared to other existing models such as CNN or VGG16 despite their well-known status because they exhibit slightly lower accuracy, precision and recall rates, accompanied by higher loss values. However, this partitioning suggests that even though these methods are competent enough to make reliable predictions about occurrence of cancer in human organism, they fail when it comes to more complex patterns using brain tumor data making them less accurate than knowledge distilled resnetxt.

Table 2. *Performance of proposed Knowledge Distilled ResNeXt and state-of-art models for tumour classification using 10-fold cross validation.*

Model	Fold	Accuracy	Precision	Recall	F1-Score	Loss
Knowledge Distilled ResNeXt	2	0.96	0.89	0.92	0.91	0.8797
	4	0.96	0.93	0.96	0.97	0.9232
	6	0.95	0.91	0.94	0.96	0.9388
	8	0.96	0.91	0.95	0.96	0.9205
	10	0.94	0.91	0.94	0.91	0.9418
CNN	2	0.92	0.93	0.9	0.93	0.9162
	4	0.92	0.9	0.92	0.94	0.909
	6	0.88	0.89	0.91	0.91	0.9341
	8	0.95	0.94	0.89	0.95	0.8886
	10	0.93	0.92	0.9	0.94	0.9049
VGG16	2	0.9	0.91	0.89	0.93	0.9313
	4	0.94	0.94	0.91	0.93	0.9292
	6	0.92	0.91	0.92	0.92	0.9126
	8	0.89	0.91	0.88	0.93	0.9098
	10	0.88	0.9	0.9	0.95	0.9244
MLPNN	2	0.9	0.9	0.91	0.91	0.9479
	4	0.88	0.9	0.9	0.91	0.9432

	6	0.91	0.92	0.9	0.93	0.9493
	8	0.89	0.92	0.91	0.89	0.9052
	10	0.91	0.9	0.92	0.92	0.9524
DBN	2	0.9	0.93	0.95	0.91	0.8941
	4	0.93	0.93	0.97	0.92	0.9075
	6	0.91	0.92	0.96	0.93	0.8883
	8	0.91	0.93	0.94	0.92	0.8966
	10	0.89	0.93	0.95	0.95	0.8976

The metrics of interest from the proposed Knowledge Distilled ResNeXt model with its performance against CNN, VGG16, MLPNN and DBN which are state-of-the-art models for brain tumor classification using 10-fold cross-validation is presented in table 2. The parameters that were used to assess include accuracy, precision, recall, F1-Score and loss across folds. The Knowledge Distilled ResNeXt has given consistent good performance over all the folds with accuracy varying slightly from 0.94 – 0.96. This model also has a high balance between precision and recall giving F1-Scores as high as 0.97.The low loss values of Knowledge Distilled ResNeXt imply that it is learning efficiently and converging stably during training. On the other hand, although CNN overall shows good performances, some of its folds have lower accuracy and F1-Scores. While it does well in certain ones (e.g., fold #4), its low variation implies it may lack robustness like Knowledge Distilled ResNeXt on diverse data subsets. The VGG16 exhibits a variable performance having an accuracy range between 0.88–0.94; thus maintaining reasonable precision and recall figures while suggesting some inconsistencies through its F1-Scores and loss values especially in certain folds where its behavior deteriorates.

Among other models such as MLPNN and DBN their different results vary across multiple folds. For instance, MLPNN's accuracy is ranging from 0.88–0.91 with moderate precision and recall stability until we get to see loss figures higher than what we observe regarding knowledge distilled resnext, indicating possible convergence issues or cases of overfitting situations. DBM demonstrates slight improvement with a spread of accuracies lying between 0.89-0.93 coupled by relatively stable precision-recall pairs though it has higher losses compared to those of the knowledge distilled resnext model. In conclusion, the Knowledge Distilled ResNeXt model outperforms others in terms of accuracy, F1-Score and consistency over different folds. This demonstrates its robustness and efficiency towards brain tumor classification task making it more reliable for this application compared to other models that are not as effective. Although competent, the other models indicate different performances which might show they are not as good at handling the complexities of dataset as Knowledge Distilled ResNeXt does.

4. Conclusions

This study emphasizes the importance of accurate and fast diagnosis in mitigating brain tumors' devastating effects on patients. We employed deep learning techniques to improve detection and classification of brain tumors while solving the drawbacks associated with traditional models like CNN and VGG16. Our novel state-of-the-art deep neural network architecture — Knowledge Distilled ResNeXt — showed outstanding performance in capturing minute tumor details and achieving more precise multi-class brain tumor classification than any other model proposed before it. The Knowledge Distillation technique enabled our model to learn from a much larger and complex one therefore outperforming the current VGG16 model by far. During the experiment, promising results were obtained whereby the suggested model achieved 95.2% accuracy rate when sorting out types of brain tumors as opposed to 89.9% attained by existing VGG16 based systems according to our knowledge at this point.

Therefore, we can say that this performance enhancement is huge because it demonstrates how much potential there may be for Knowledge Distilled ResNeXt to support accurate diagnosis-making process as well efficient planning for different kinds of cancer such as glioma, meningioma, pituitary gland cancers among others or even those with no mass found at all. This represents a significant step forward in medical image analysis for detecting where success has been achieved through successful implantation into practice; thus revolutionizing care given towards patients affected thereby leading them towards recovery faster than ever thought possible before now became clear when complicated features were extracted using MRI scans but still need confirmation from clinicians who have access only limited information sources available within typical clinical setting but also require highest levels confidentiality so as not put any person's life at risk unnecessarily too long better lives depending on right now are able do decide about whether should go ahead with certain treatments looking good.

I believe they will pave wave future development these kind next level advanced stages deep learning methods relevant fields were made great strides during recent times however further investigation needed since there could become even better methods that would be available for diagnosis and treatment of brain tumors within this century or more long history when brain cancer was first discovered until today many discoveries have been made about how best we can accurately detect them at present still working hard towards achieving greater accuracy rates while detecting such diseases so far what we did found out will greatly help those suffering now need go through lot pain before they get well always wanted contribute something valuable my fellow human beings not only is it important to continue striving towards achieving higher precision levels in brain tumour detection but also ensuring that these improvements are implemented effectively across different healthcare facilities thus benefiting both patients and care givers.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Sharif, M. I., Li, J. P., Khan, M. A., & Saleem, M. A. (2020). Active deep neural network features selection for segmentation and recognition of brain tumours using MRI images. *Pattern Recognition Letters,* 129, 181-189. https://doi.org/10.1016/j.patrec.2019.11.019

[2] Yin, B., Wang, C., & Abza, F. (2020). New brain tumour classification method based on an improved version of whale optimization algorithm. *Biomedical Signal Processing and Control,* 56, 101728. https://doi.org/10.1016/j.bspc.2019.101728

[3] Amin, J., Sharif, M., Haldorai, A., Yasmin, M., & Nayak, R. S. (2021). Brain tumour detection and classification using machine learning: a comprehensive survey. *Complex & Intelligent Systems,* 1-23. https://doi.org/10.1007/s40747-021-00563-y

[4] Rinesh, S., Maheswari, K., Arthi, B., Sherubha, P., Vijay, A., Sridhar, S., ... & Waji, Y. A. (2022). Investigations on brain tumour classification using hybrid machine learning algorithms. *Journal of Healthcare Engineering*. https://doi.org/10.1155/2022/2761847

[5] Ghassemi, N., Shoeibi, A., & Rouhani, M. (2020). Deep neural network with generative adversarial networks pre-training for brain tumour classification based on MR images. *Biomedical Signal Processing and Control,* 57, 101678. https://doi.org/10.1016/j.bspc.2019.101678

[6] Kang, J., Ullah, Z., & Gwak, J. (2021). Mri-based brain tumour classification using ensemble of deep features and machine learning classifiers. *Sensors,* 21(6), 2222. https://doi.org/10.3390/s21062222

[7] Brindha, P. G., Kavinraj, M., Manivasakam, P., & Prasanth, P. (2021, February). Brain tumour detection from MRI images using deep learning techniques. *In IOP Conference Series: Materials Science and Engineering* (Vol. 1055, No. 1, p. 012115). IOP Publishing. https://doi.org/10.1088/1757-899X/1055/1/012115

[8] Irmak, E. (2021). Multi-classification of brain tumour MRI images using deep convolutional neural network with fully optimized framework. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering,* 45(3), 1015-1036. https://doi.org/10.1007/s40998-021-00426-9

[9] Amran, G. A., Alsharam, M. S., Blajam, A. O. A., Hasan, A. A., Alfaifi, M. Y., Amran, M. H., & Eldin, S. M. (2022). Brain Tumour Classification and Detection Using Hybrid Deep Tumour Network. *Electronics,* 11(21), 3457. https://doi.org/10.3390/electronics11213457

[10] ZainEldin, H., Gamel, S. A., El-Kenawy, E. S. M., Alharbi, A. H., Khafaga, D. S., Ibrahim, A., & Talaat, F. M. (2022). Brain Tumour Detection and Classification Using Deep Learning and Sine-Cosine Fitness Grey Wolf Optimization. *Bioengineering,* 10(1), 18. https://doi.org/10.3390/bioengineering10010018

[11] Khan, A. H., Abbas, S., Khan, M. A., Farooq, U., Khan, W. A., Siddiqui, S. Y., & Ahmad, A. (2022). Intelligent model for brain tumour identification using deep learning. *Applied Computational Intelligence and Soft Computing*. 2022, Article ID 8104054, 10 pages https://doi.org/10.1155/2022/8104054https://doi.org/10.1155/2022/8104054

[12] Arif, M., Ajesh, F., Shamsudheen, S., Geman, O., Izdrui, D., & Vicoveanu, D. (2022). Brain tumour detection and classification by MRI using biologically inspired orthogonal wavelet transform and deep learning techniques. *Journal of Healthcare Engineering*, 2022. https://doi.org/10.1155/2022/2693621

[13] Muhammad, K., Khan, S., Del Ser, J., & De Albuquerque, V. H. C. (2020). Deep learning for multigrade brain tumour classification in smart healthcare systems: A prospective survey. *IEEE Transactions on Neural Networks and Learning Systems*, 32(2), 507-522. https://doi.org/10.1109/TNNLS.2020.2995800

[14] Kumar, T. S., Arun, C., & Ezhumalai, P. (2022). An approach for brain tumour detection using optimal feature selection and optimized deep belief network. *Biomedical Signal Processing and Control*, 73, 103440. https://doi.org/10.1016/j.bspc.2021.103440

[15] Nayak, D. R., Padhy, N., Mallick, P. K., & Singh, A. (2022). A deep autoencoder approach for detection of brain tumour images. *Computers and Electrical Engineering*, 102, 108238. https://doi.org/10.1016/j.compeleceng.2022.108238

[16] Aamir, M., Rahman, Z., Dayo, Z. A., Abro, W. A., Uddin, M. I., Khan, I., ... & Hu, Z. (2022). A deep learning approach for brain tumour classification using MRI images. *Computers and Electrical Engineering*, 101, 108105. https://doi.org/10.1016/j.compeleceng.2022.108105

[17] R. Sa, W. Owens, R. Wiegand, M. Studin, D. Capoferri, K. Barooha, A. Greaux, R. Rattray, A. Hutton, J. Cintineo, et al. (2017). Intervertebral disc detection in x-ray images using faster r-cnn, in: 2017 *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, pp. 564-567. https://doi.org/10.1109/EMBC.2017.8036887

[18] Sadia Anjum, Lal Hussain, Mushtaq Ali, Monagi H. Alkinani, Wajid Aziz, Sabrina Gheller, Adeel Ahmed Abbasi, Ali Raza Marchal, Harshini Suresh, and Tim Q. Duong. (2022). Detecting brain tumours using deep learning convolutional neural network with transfer learning approach. *Int J Imag Syst Technol* 32(1):307-323. https://doi.org/10.1002/ima.22641

[19] Muhannad Faleh Alanazi, Muhammad Umair Ali, Shaik Javeed Hussain, Amad Zafar, Mohammed Mohatram, Muhammad Irfan, Raed AlRuwaili, Mubarak Alruwaili, Naif H. Ali, and Anas Mohammad Albarrak. (2022). Brain tumour/mass classification framework using magnetic-resonance-imaging-based isolated and developed transfer deep-learning model. *Sensors* 22(1):372. https://doi.org/10.3390/s22010372

[20] Shin H.-C., Roth H.R., Gao M., Lu L., Xu Z., Nogues I., Yao J., Mollura D., Summers R.M. (2016). Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging* 35(5);1285-1298. https://doi.org/10.1109/TMI.2016.2528162

[21] Kharrat A., Gasmi K., Messaoud M.B., Benamrane N., Abid M. (2010). A hybrid approach for automatic classification of brain mri using genetic algorithm and support vector machine. *Leonardo J Sci.* 17(1):71-82. https://doi.org/10.1109/COGINF.2010.5599712

[22] N.B. Bahadure, A.K. Ray, H.P. Thethi. (2017). Image analysis for mri based brain tumour detection and feature extraction using biologically inspired bwt and svm. *Int J Biomed Imaging.* https://doi.org/10.1155/2017/9749108

[23] Rehman Z.U., Naqvi S.S., Khan T.M., Khan M.A., Bashir T. (2019). Fully automated multi-parametric brain tumour segmentation using superpixel based classification. *Expert Syst Appl*. 118;598-613. https://doi.org/10.1016/j.eswa.2018.10.040

[24] Chaplot S., Patnaik L.M., Jagannathan N. (2006). Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. *Biomed Signal Process Control*. 1(1):86-92. https://doi.org/10.1016/j.bspc.2006.05.002

[25] El-Dahshan E.-S.A., Hosny T., Salem A.-B.M. (2010). Hybrid intelligent techniques for mri brain images classification. *Digital Signal Process.* 20(2):433-441. https://doi.org/10.1016/j.dsp.2009.07.002

[26] W. Chen, B. Liu, S. Peng, J. Sun, X. Qiao (2018). Computer-aided grading of gliomas combining automatic segmentation and radiomics, *Int J Biomed Imaging.* https://doi.org/10.1155/2018/2512037

[27] Sultan H.H., Salem N.M., Al-Atabany W. (2019). Multi-classification of brain tumour images using deep neural network. *IEEE Access*. 7;69215-69225. https://doi.org/10.1109/ACCESS.2019.2919122

[28] Arepalli, P.G., Naik, K.J. (2024). Water contamination analysis in IoT enabled aquaculture using deep learning based AODEGRU. *Ecological Informatics*, 79; 102405. https://doi.org/10.1016/j.ecoinf.2023.102405

[29] Gopi, A.P., Gowthami, M., Srujana, T., Padmini, S.G., Malleswari, M.D. (2022). Classification of denial-of-service attacks in IoT networks using AlexNet. *Human-Centric Smart Computing*, Springer, Singapore, pp. 349-357. https://doi.org/10.1007/978-981-19-5403-0_30

[30] Arepalli, P. G., & Naik, K. J. (2024). A deep learning-enabled IoT framework for early hypoxia detection in aqua water using light weight spatially shared attention-LSTM network. *The Journal of Supercomputing*, 80(2), 2718-2747. https://doi.org/10.1007/s11227-023-05580-x

[31] OZSOY, S., & DELIBAS, E. A. O. (2023). The Effect of Fragment C of Tetanus Toxin on Memory Deficits in a Rat Model of Alzheimer's Disease. *International Journal of Computational and*

Experimental Science and Engineering, 9(3), 254–259. Retrieved from https://ijcesen.com/index.php/ijcesen/article/view/264

[32] M, V., V, J., K, A., Kalakoti, G., & Nithila, E. (2024). Explainable AI for Transparent MRI Segmentation: Deep Learning and Visual Attribution in Clinical Decision Support. *International Journal of Computational and Experimental Science and Engineering*, 10(4);575-584. https://doi.org/10.22399/ijcesen.479

[33] BACAK, A., ŞENEL, M., & GÜNAY, O. (2023). Convolutional Neural Network (CNN) Prediction on Meningioma, Glioma with Tensorflow. *International Journal of Computational and Experimental Science and Engineering*, 9(2), 197–204. Retrieved from https://ijcesen.com/index.php/ijcesen/article/view/210

[34] N, S., S. Prabu, V, T. K., D, C., K, B., & B. Buvaneswari. (2024). Computer Aided Based Performance Analysis of Glioblastoma Tumor Detection Methods using UNET-CNN. *International Journal of Computational and Experimental Science and Engineering*, 10(4);753-762. https://doi.org/10.22399/ijcesen.515