



# Enhanced performance of Dark-Nets for brain tumor classification and segmentation using colormap-based superpixel techniques

Sakshi Ahuja<sup>\*</sup>, Bijaya Ketan Panigrahi, Tapan Kumar Gandhi

Electrical Engineering Department, Indian Institute of Technology Delhi, New Delhi, 110016, India

## ARTICLE INFO

### Keywords:

Brain tumor  
MRI  
Deep learning  
Superpixel segmentation

## ABSTRACT

The brain tumor is the deadliest disease in adults as it arises due to an abnormal mass of cells that grows rapidly and it alters the proper functioning of the organs. In clinical practice, radiographic images of different modalities are used to diagnose types of brain tumors, their size, and location. The proposed work aims to automatically classify, localize, and segment brain tumors from T1W-CE Magnetic Resonance Image (MRI) datasets. The T1W-CE MRI dataset is divided into 8:1:1, i.e., 80% training set, 10% of each validation, and testing set. To address the overfitting issues, the training data set is augmented using 2-levels wavelet decomposition and geometrical operations (scaling, rotation, translation). Performance of pre-trained DarkNet model (DarkNet-19 and DarkNet-53) is evaluated for the multi-class classification and localization of brain tumors. The best performing pre-trained DarkNet model achieved 99.60% of training accuracy and 98.81% of validation accuracy. The performance evaluation parameters confirm the superiority of the proposed methodology in comparison to the state-of-the-art on the T1W-CE MRI dataset. On 1070 T1W-CE testing images, the best-performing pre-trained DarkNet-53 model obtained a testing accuracy of 98.54% and Area Under Curve (AUC) of 0.99. The tumor is segmented using a 2-D superpixel segmentation technique with an average dice index of  $0.94 \pm 2.6\%$  on the 793 brain tumor testing data. To prove the superiority of the proposed technique, it is implemented on MRI images from the BraTS2018 dataset. The comparative analysis of performance evaluation parameters of the proposed methodology with the state-of-the-art technique proves its robustness and clinical significance.

## 1. Introduction

The clinical diagnosis of the tumors includes analysis of the nature of cancerous cells, i.e., cells are cancerous (malignant) or non-cancerous cells (benign), size, and location of malignant cells. Tumors are given a name based on (a) origin of tumor cells, (b) number ranging from 1–4 or grade in roman numerals I–IV (Community, 2013), for an e.g. grade I–II refers to Low-Grade Glioma (LGG), and grade III–IV refers to High-Grade Glioma (HGG). In a recent study, out of all types of Central Nervous System (CNS) tumors, there are more than 80% chances that it is a brain tumor (Observatory, 2018). As per World Health Organization (WHO) report, there are 120 types of brain tumors based on types of tumor cell, their location, and complexity (Patel, Fisher, Nichols, A.-Allah, & J. Abdela, 2019). The most deadly tumor types are Meningioma, Glioma, Pituitary, and lymphoma (Board, 2019).

Brain tumors are generally divided into two categories, i.e., metastatic and primary brain tumors. In primary brain tumor types, the cancerous cells originate within brain cells and the most common tumors are: (a) Meningioma (35%), (b) Glioblastoma (16%), and (c)

Pituitary tumor (14%) (Ostrom et al., 2015). Meningioma is a non-cancerous tumor that grows from the membranes surrounding the brain and spinal cord. The Pituitary tumors develop in the pituitary gland and come under the benign tumor category with abnormal growths (Ostrom et al., 2019; Patel et al., 2019). If the cancer cells spread into the brain from another infected area of the body parts then it is a metastatic tumor.

Medical imaging techniques use a variety of imaging techniques like X-ray, MRI, functional Magnetic Resonance Imaging (fMRI), Magnetic Resonance Spectroscopy (MRS), Computed Tomography (CT) scan, Positron Emission Tomography (PET) scan, etc. Radiologist prefers structural MRI scan for brain tumor diagnosis because it provides brain tumor information on cellular, functional, metabolic and vascular level. The multi-modality and multi-orientation MRI are useful for the detailed diagnosis of brain tumor tissue. The multi-modal MRI includes T1, T1-Gadolinium (T1-Gd), T1-Weighted (T1-W), T2, T2-W, Fluid Attenuated Inversion Recovery (FLAIR), etc. These modalities are used to examine the variety of tumor tissues, namely, edema, necrosis, enhancing, and non-enhancing tissues. T1Gd is useful to visualize active

<sup>\*</sup> Corresponding author.

E-mail addresses: [Sakshi.Ahuja@ee.iitd.ac.in](mailto:Sakshi.Ahuja@ee.iitd.ac.in) (S. Ahuja), [bkpanigrahi@ee.iitd.ac.in](mailto:bkpanigrahi@ee.iitd.ac.in) (B.K. Panigrahi), [tgandhi@ee.iitd.ac.in](mailto:tgandhi@ee.iitd.ac.in) (T.K. Gandhi).

tumors, tumor core is visible in T2, and complete tumor region is visible in T2 FLAIR modality (Villanueva-Meyer, Mabray, & Cha, 2017). The edema tissue that surrounds the tumor appears bright in T2-W MRI. Also, MRI is available in multi-orientation, i.e., axial, coronal, and sagittal.

Various medical units such as radiology, chemotherapy, and neurosurgery need to be consulted for the effective treatment of the tumor patient. Further, the choice of the treatment procedure, i.e., surgery or chemotherapy depends on tumor grades, location, size, and types. The major challenges of manual diagnosis of a brain tumor are complexity in the size of types of tumor, different shapes, types, and size of tumors, etc. In clinical setup, manual diagnosis of tumor tissues using a large set of MRI slices for each orientation and in different modalities is a complex task. Thus, automatic brain tumor classification and segmentation technique is required.

The proposed methodology aims to classify, localize and segment the tumor from the multi-view T1W-CE MRI dataset. The input brain MRI is classified into four categories, namely, normal, Meningioma, Glioma, and Pituitary tumor. For the tumor classification, the performance of pre-trained DarkNets models is evaluated on multi-orientation T1W-CE MRI, i.e., axial, coronal, and sagittal views. Initially, input brain MRI data is pre-processed (Pre-processing-I) to make input data compatible with the pre-trained DarkNets model. The pre-processing steps used are normalization, resize, and image data type converted to JPEG format. The pre-processed data is divided into 8:1:1, i.e., 80% of the training set, 10% of each validation set, and testing set. The training dataset is not sufficient for DarkNets model training and leads to an overfitting issue. To resolve the overfitting issue, training data is augmented using geometrical operations (scaling, translation, rotation) and 2-levels of the wavelet decomposition. DarkNets models are trained on augmented training data and numerous evaluation parameters are used to compare their performance. Further, feature maps of the finest performing DarkNets model are used for tumor localization. Then, MRI data having tumors are pre-processed (Pre-processing-II) for the tumor segmentation task. The pre-processing-II includes skull stripping, noise removal. Finally, the tumor is segmented using a color-map based on 2-D superpixel segmentation techniques. The segmented tumor is compared with the ground truth mask based on the following parameters: dice index, and error in the area of the tumor. Thus, the proposed framework addresses the following issues:

1. Address the data overfitting issue by training the pre-trained deep learning model on augmented data using geometrical and 2-levels of wavelet decomposition techniques.
2. A multi-class classification framework for tumor classification in all three views (axial, coronal, and sagittal views).
3. Superpixel techniques are used for the effective discrimination of tumor tissue regions.

The rest of the paper is organized as follows: Section 2 illustrates the techniques available in the literature for brain tumor classification and segmentation; Section 3 is dedicated to the materials and methods for tumor classification using DarkNets models, and segmentation using superpixel technique; Section 4 put forth the results and discussion. Conclusions are put forth in Section 5.

## 2. Literature review

In the state-of-the-art, the brain tumor diagnosis process includes tumor detection, classification, localization, and segmentation. Radiography images in association with artificial intelligence-based techniques are useful for the accurate diagnosis of the disease. The binary class and multi-class classification of brain tumors are done using handcrafted features, machine learning, and deep models. This section put forth the brief details of techniques available in the literature for brain tumor diagnosis.

### 2.1. Brain tumor classification

In Sriramakrishnan, Kalaiselvi, and Rajeswaran (2019), a Support Vector Machine (SVM) classifier is used to classify tumor slices. Fuzzy C Means (FCM) algorithm is implemented to obtain tumor regions. Further, the Probabilistic Local Ternary Patterns (PLTP) technique is used to obtain the tumor substructures. The model achieved a dice index of 76% for complete tumor region, 53% for core, and 58% for enhancing tumor on BraTS 2013. Whereas for BraTS 2015, the dice values are 81% for complete tumor, 49% for core, and 47% for enhancing tumor. The CNN model is implemented to classify three types of tumor, i.e., Meningioma, Pituitary, and Glioma tumor with an accuracy of 94.39% (Das, Aranya, & Labiba, 2019).

In Nayak, Dash, and Majhi (2016), features are extracted using Discrete Wavelet Transform (DWT), dimensionality reduction is done using Probabilistic Principal Component Analysis (PPCA). Further, classification is done using the AdaBoost algorithm and random forests classifier. Classification accuracy of 100% is achieved with Dataset-66, and Dataset-160. Whereas, an accuracy of 99.53% is achieved with Dataset-255. In Chaplot, Patnaik, and Jagannathan (2006), the wavelet-based technique is implemented on 52 axial T2-weighted images. The model obtained a classification accuracy of 94% with Self-Organizing Maps (SOMs) and 98% with SVM. Probabilistic Neural Network (PNN) and Discrete Wavelet Transform (DWT) are used to classify normal, and tumor-infected tissues. The model obtained a classification accuracy of 100% on 25 DICOM images dataset having 18 brain tumor infected tissues, and the rest normal tissues (Shree & Kumar, 2018).

Brain tumor grades and types are classified using feature extraction and classified using binary SVM (Zacharaki et al., 2009). The gradient and context-sensitive features are extracted to detect and classify HGG, and LGG (Zhao et al., 2019). The obtained accuracy, sensitivity, and specificity for classification of metastases and gliomas are 85%, 87%, and 79% respectively. Whereas HGG and LGG are classified with an accuracy of 88%, a sensitivity of 85%, and a specificity of 96%. The tumor is classified into binary classes from 3-D volumetric FLAIR images using an SVM classifier without PCA with an accuracy of 88% (Gupta, Gandhi, Gupta, & Panigrahi, 2017).

Three types of brain tumors are classified using CNN based models such as AlexNet, GoogleNet, VGG-16, VGG-19, ResNet-18, ResNet-50, ResNet-101, ResNet-Inception-v2, and SENet. The model obtained an accuracy of 98.71% on the 3074 (60% training data) T1W-CE MRI dataset. In A-Ellah, Awad, Khalaf, and Hamed (2018), tumor localization is evaluated using CNN models (AlexNet, VGG16, and VGG19). An accuracy of 99.55% is obtained with CNN based AlexNet model on the RIDER Neuro MRI database. A dice similarity index of 0.87 is achieved on 804 3D MRI from BraTS 2013 database for the brain tumor segmentation.

The CNN-based model is proposed for the multi-class classification of a brain tumor (Zacharaki et al., 2009). The proposed work obtained following accuracies: (a) 99.3% for classification-1 (tumor/non-tumor), (b) 92.66% for classification-2 (normal brain, glioma, meningioma, pituitary and metastatic), and (c) 98.14% for glioma grades classification-3 task. The deep learning-based model is implemented with data augmentation for tumor classification with an accuracy of 94.74% on the figshare dataset (Ayadi, Elhamzi, Charfi, et al., 2021). In Sasank and Venkateswarlu (2021) brain tumor is classified from segmented brain tumor from MR images from BraTS 2014, BraTS 2015, BraTS 2018 dataset. The proposed methodology is implemented with a Soft plus extreme learning machine (KSELM) classifier.

### 2.2. Brain tumor segmentation

Superpixel-based segmentation is used to localize brain tumors with a dice index of 0.91 in 2-D MRI-based FLAIR modality from BraTS 2012 database (Rehman, Naqvi, Khan, Khan, & Bashir, 2019). Numerous automatic superpixel methods are as follows: graphical models, 2-D

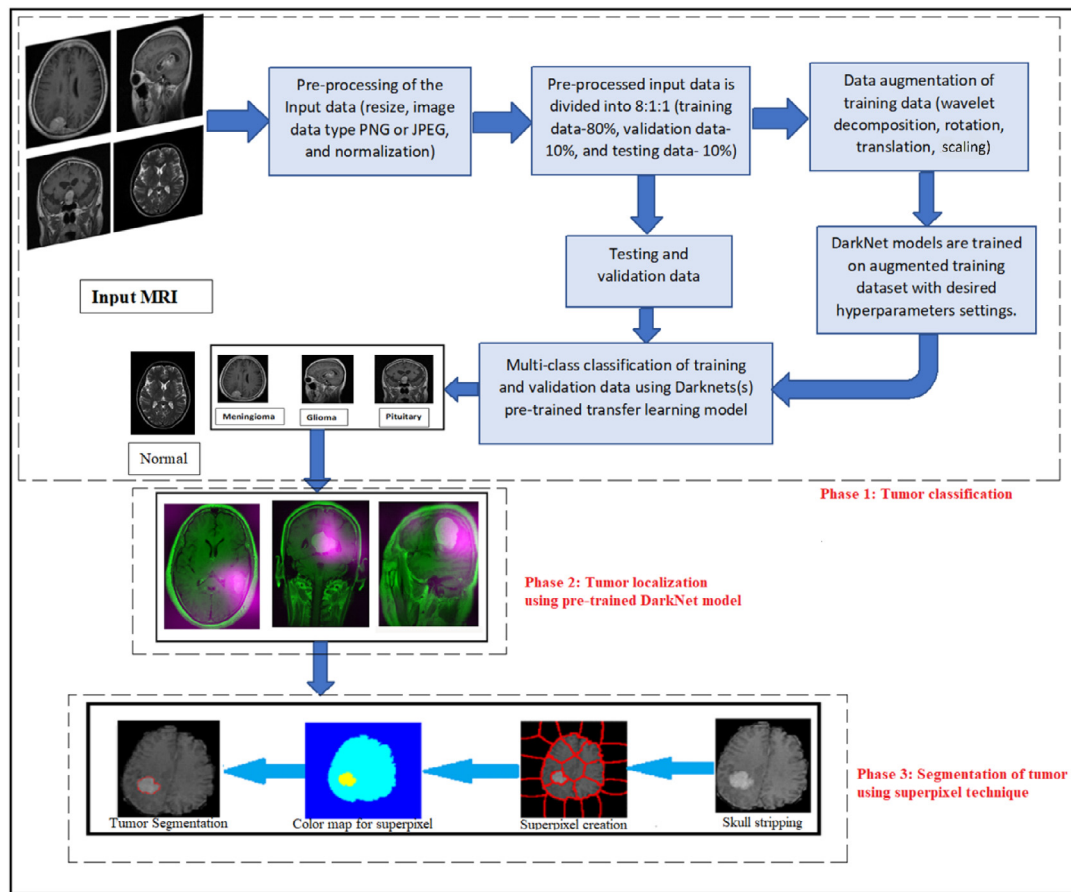


Fig. 1. Self-explanatory block diagram for the proposed methodology for brain tumor diagnosis model.

and 3-D superpixels, Simple Liner Iterative Clustering (SLIC), and Turbo pixels (A.-Dmour & A.-Ani, 2018; Achanta et al., 2012; Chen, Nguyen, Chui, & Ong, 2016). Superpixel-based spectral clustering is used to identify the tumor region labeled as Region Of Interest (ROI) with a dice similarity coefficient of 0.87 (tumor core), and 0.92 (edema) (Angulakshmi & Priya, 2018). Each superpixel is classified into normal and abnormal tissue via a Graphical model. The Graphical model implemented on BraTS 2013, and BraTS 2015 data set obtained a dice index of 91.5% (Zhao, Yang, Lin, Pang, & Wang, 2018).

On BraTS 2013 database, superpixel, and SVM-based techniques are implemented to segment tumors from FLAIR MRI with a dice index of 86.12% (Chen et al., 2017). The Superpixel along with Leave-One-Out Cross-Validation (LOOCV) obtained a dice index of 88% for whole tumor segmentation (Kong et al., 2019). The complete tumor and non-tumor areas are localized using low-level features of superpixels from multi-modal MRI scan from BraTS 2012 dataset (Bjoern et al., 2014). The superpixel segmentation with Bayesian inference is used to obtain whole-brain tissue-type maps of multi-modal glioma tumor types (Raschke et al., 2019). A graph-cut based superpixel technique obtained an average dice score of 81.1 on the multi-label BraTS 2013 dataset (Abiwinanda, Hanif, Hesaputra, Handayani, & Mengko, 2019). For superpixels based methods, the selection of an appropriate superpixel size is an important criterion to obtain a high similarity index (A.-Dmour & A.-Ani, 2018; Amiri, Mahjoub, & Rekik, 2018; Diniz et al., 2018).

The tumor is segmented from multi-orientation MRI slices using Fully Convolution Neural Networks (FCNNs) technique, and the Conditional Random Field (CRF) technique (Zhao, Wu et al., 2018). On BraTS 2017 dataset (Soltaninejad et al., 2018), a Fully Convolution Network (FCN) obtained a mean Dice index of 0.86 for WT, 0.78 for CT, and 0.66 for ET. In Amin, Sharif, Yasmin, and Fernandes (2018), lightweight

DNN is used on benchmark dataset ISLES and BRATS. The method obtained a dice index of 99.8% on Flair modality, 98% on T2, 100% on DWI, 95.4% on T1-C, and 97.4% on T1 modalities in 5.502 s. Also, brain tumors are segmented from FLAIR MRI using FCNNs (Lorenzo et al., 2019). An average dice index of 0.81 is obtained on 20 real HGG BraTS 2012 datasets (Chen et al., 2016). The spectral clustering-based superpixel technique obtained the dice index of 0.87 on tumor core and 0.92 in tumor edema in 20 s time duration (Angulakshmi & Priya, 2018).

A brain tumor is segmented using the Extremely Randomized Trees (ERT) technique based on appearance and context-based features (Soltaninejad et al., 2017). The ERT technique obtained a dice index of 0.85 for complete tumor, 0.79 for core, and 0.75 for enhancing tumor on the BraTS 2013 database, (Pinto, Pereira, Rasteiro, & Silva, 2018). Supervoxels and RF classifier-based techniques are used to segment the brain tumor sub-regions with a dice index of 0.80 for TC and 0.89 for WT (Mohammadreza et al., 2018).

Semantic segmentation is implemented using deep learning, i.e., labeling each pixel with a class of objects or non-objects is gaining researcher attention (Lateef & Ruichek, 2019). In Zhang et al. (2021), a four encoder-based ME-Net model is implemented on multimodal brain MRI. On BraTS 2020 dataset, the dice index obtained is 0.70249 for ET, 0.88267 for WT, and 0.73864 for TC. Fully automatic UNet based deep CNN is used to segment and detect brain tumors from BraTS 2015 dataset (Dong, Yang, Liu, Mo, & Guo, 2017). The average dice index obtained are WT-0.86, CT-0.86, and ET-0.65. The Cascade Convolutional Neural Network (C-CNN) obtained a dice index of 92.03% for WT, 91.13% for ET, and 87.26% for TC on the BraTS 2018 dataset (Ranjbarzadeh et al., 2021).

From the literature review, it can be inferred that MRI is the highly recommended medical imaging technique for brain tumor diagnosis by



researchers. MRI is based on 3-D data that provides useful information about contrast differences in soft tissues. Manual diagnosis of tumors from multiple modalities and multi-orientation MRI is a complex task. Thus, researchers are making efforts to develop fully automatic and semi-automatic techniques for brain tumor diagnosis. The automatic brain tumor diagnosis technique must possess the following properties, i.e., (a) enhanced classification accuracy, sensitivity, specificity, etc., (b) improved dice coefficient in less time, and c) reduced computational cost for the testing phase.

### 3. Materials and methods

The AI-based approach is the most reliable method for the faster and accurate diagnosis of brain tumors. In the proposed work, the tumor is classified, localized, and segmented from the T1W-CE MRI dataset. However, the limited availability of medical imaging modalities for the training of the deep learning model is a major challenge. To address the limited medical dataset issue, CNN-based transfer learning models along with data augmentation are preferred. The schematic hierarchy of the proposed methodology for brain tumor diagnosis on the considered dataset is shown in Fig. 1.

#### 3.1. Phase-1: Tumor classification

In phase-1, the input brain MRI dataset is classified into four classes. i.e., normal, Meningioma, Glioma, and Pituitary tumor. Further, the input MRI dataset is of different sizes and image formats. Thus pre-processing of input data is needed to make it compatible for training pre-trained DarkNets model. The pre-processed images set is then divided into 8:1:1, i.e., 80% training set, 10% of each validation, and testing set. The training set is not sufficient for the enhanced performance of the pre-trained DarkNets model and may cause an overfitting issue. Therefore, a data augmentation technique is used on the training dataset. The data augmentation process includes 2-levels of wavelet decomposition, scaling, rotation, translation. The DarkNet-19 and DarkNet-53 pre-trained models are trained on augmented training data with desired hyperparameters settings.

##### 3.1.1. Dataset

The tumor-containing images are obtained from the figshare dataset. In medical image analysis, the pre-trained deep learning model requires data annotations from the expert. In the proposed work, the data annotation and image labels with tumor masks are provided by the neuro-radiologist expert (Cheng, 2017). The figshare dataset contains a total of 3064 T1W-CE brain MRI slices with three types of brain tumor, namely, (a) Meningioma, (b) Glioma, and (c) Pituitary. The T1W-CE modality contains a contrast agent in brain vessels and offers the following advantages: (a) necrotic and active tumor area is properly visible, and (b) tumor borders appear brighter. The input images are available in  $512 \times 512$  pixels in multi-orientation views, i.e., axial, coronal, and sagittal. The sample images of T1W-CE MRI images are shown in Fig. 2.

For the normal images, the IXI dataset having 600 MRIs without any lesion is used (Group). The MR images are available in NIFTI format in different acquisition protocols. The numerous modalities available are MRA images, T1, T2, PD-W, and DWI in 15 directions. The normal patient data is collected from three different hospitals in London, namely, (a) Institute of Psychiatry using a GE system (1.5 T), (b) Guy's Hospital using a Philips system (1.5 T), and (c) Hammersmith Hospital using a Philips system (3 T). The input images are pre-registered with the T1W-CE MRI dataset and bias field correction is performed using the Multiplicative Intrinsic Component Optimization (MICO) algorithm (Li, 2021).

#### 3.1.2. Pre-processing

The primary concern of pre-processing stage is to make the input MRI dataset compatible with the DarkNets pre-trained model. The pre-processing step includes (a) modification in image data type, (b) normalization, and (c) re-size input data.

The input MRI data is available in NIFTI file format. The proposed deep learning model is compatible with input images of JPEG or PNG image format. With the medical expert supervision, the considered MRI dataset is converted into .png format as it is lossless in-compared to the JPEG format for efficient diagnosis (Dodge & Karam, 2016). Further, resize the input images to  $256 \times 256 \times 3$  and make them compatible with DarkNets pre-trained models. After resizing, MR images are normalized. Using Eq. (1), the input image is normalized to the confined range of 0 to 1. Consider 'B' denotes the input brain MRI having a size ( $p \times q$ ), and 'Bnorm' is the normalized MRI.

$$B_{norm} = \frac{(B - \min(B))}{(\max(B) - \min(B))} \quad (1)$$

The pre-processed images are divided into 8:1:1, i.e., 80% training set, 10% of each validation, and testing set. However, pre-processed training sets are not sufficient for the efficient training of the pre-trained DarkNet model. Thus, the training set needs to be augmented using numerous parameters to reduce the overfitting issues.

#### 3.1.3. Training data augmentation

The pre-processed training set is augmented to enhance the performance of the pre-trained DarkNets model. The input images are augmented to address the data overfitting issue. The augmented data is created by following ways: (a) 2-level wavelet decomposition (b) scaling - [0.9 1.1], (c) translation in both axis within [-60 60] pixel range, and d) random rotation from [-90 90] range (Mohamed et al., 2021).

The wavelet decomposition is performed based on scale and orientation. Wavelets have the following properties: similarity, vanishing moments, regularity, and uniqueness property, etc. Thus, wavelet decomposed images are used to extract the relevant data from an image. In the level 'L' decomposition, the input image is down-sampled to two levels and gives approximation coefficient 'A' and detailed coefficient 'D' (Horizontal detail coefficient (H), Vertical detail coefficient (V), and Diagonal detail coefficient (D) (G.P. Nason, 1995). The mathematical expression for wavelet decomposition up-to level (L) 'WD\_L' is put forth in Eq. (2). In the proposed work, input images are decomposed to two levels using a 2-D discrete stationary wavelet. The dimensions of output, i.e., decomposed image depend on the size of the input image and level of decomposition (J., Krim, & Carfantan, 1996). Further, the wavelet decomposed images undergo geometrical augmentation.

$$WD_L = [A(:, :, L); H(:, :, 1 : L); V(:, :, 1 : L); D(:, :, 1 : L)] \quad (2)$$

#### 3.1.4. Transfer learning based pre-trained DarkNets model

The major challenge of the application of deep learning models in the medical field is the limited availability of labeled datasets for training purposes. Training the deep CNN model from scratch on such a small medical dataset leads to data overfitting and time-consuming tasks (Ho & Kim, 2021). To overcome these issues, transfer learning-based pre-trained models that are trained on the ImageNet dataset. However, there is a difference in dataset sizes, features, and classes between the non-medical image dataset and the MRI medical datasets classification. Thus, the transfer learning approach uses weights and features from the trained ImageNet model (Mehrotra, Ansari, Agrawal, & Anand, 2020). Also, the transfer learning-based model are having numerous layers such as convolution layers, batch normalization, max pooling, softmax, and fully connected layers, etc. For fine-tuning the pre-trained transfer learning model for the new classification task on medical images, initial layers are freezed and the deeper layer is updated with the required weights as per the MRI dataset. Also, the initial convolution layer provides the details of low-level features, i.e., edges,

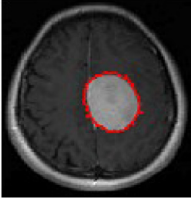

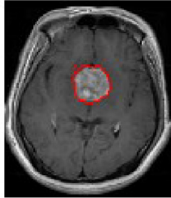
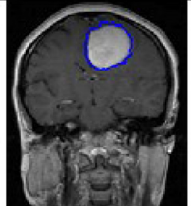
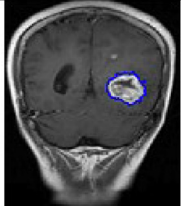
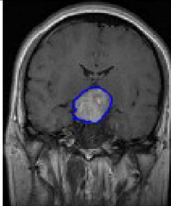

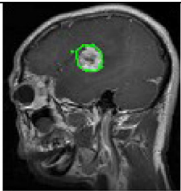
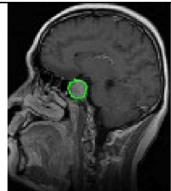
Tumor type View	Meningioma (708 slices)	Glioma (1426 slices)	Pituitary (930 slices)
<b>Axial</b>  (M=209; G=494; P=291 slices)			
<b>Corononal</b>  (M= 268; G= 437; P=319 slices)			
<b>Sagittal</b>  (M=231; G=495; P=320 slices)			

Fig. 2. Three types of tumors with three different views in T1W-CE MRI slices (Cheng, 2017).

color, etc., and the final layer depicts the high-level features of the tumor. Thus, due to this layer functionality, the desired features are extracted in fully connected layers for the new classification task, and weights are updated accordingly (Tandel et al., 2020). Due to the transfer learning-based pre-trained network, there is no need to train the proposed methodology on a large-scale MRI-based dataset.

The proposed method uses pre-trained transfer models, namely, DarkNet-19, and DarkNet-53. The DarkNet pre-trained models offer the following advantages: (a) identify small objects in an input image, (b) reduce error rate rapidly, (c) the pre-trained transfer learning models are useful to diagnose disease on limited datasets, and (d) fast in object detection. Thus DarkNet models are preferred in real-time object detection and classification with high accuracy (Vasavi, Priyadarshini, & Vardhan, 2021). The DarkNet-53 is a 53 layer deeper CNN and DarkNet-19 is 19 layers deeper network. These DarkNet models are trained on  $256 \times 256 \times 3$  image size in png format and classify data in 1000 object categories (Krizhevsky, Sutskever, & Hinton, 2012; Simonyan & Zisserman, 2014).

To classify the considered dataset, the layers of the DarkNets pre-trained model are updated according to the hyperparameters settings. The DarkNets model include convolution layer, i.e., conv(x), leaky(x), Rectified Linear Unit, i.e., leakyReLU(x) layer, batch normalization layer, i.e., batch norm(x), softmax, global average pooling (avg1), and addition layer res(y). Here, For DarkNet-53:  $x = [1:1:52]$ , and  $y = [1:1:23]$  and for DarkNet-19:  $x = [1:1:18]$ , and res (y) is not available. The pre-trained DarkNet-53 is 184 layers deeper network whereas DarkNet-19 is 64 layers deeper network. For fine-tuning, the final convolution layer is replaced with the 'new conv' layer.

The first convolution layer provides the details of the low-level features, and deeper convolution layers are used to obtain the tumor details. Batch normalization eliminates the need for a dropout layer. Also, hidden layers are not dependent on features, and normalize the output of the hidden layer. Thus it resolves the vanishing gradient issue and speeds up the training progress. In batch normalization, the normalized output is multiplied by geometrical parameters, namely, mean, and standard deviation. Therefore, two trainable parameters

are added to each layer. Consider, mean is symbolized by  $\beta$ , and the standard deviation is represented by  $\gamma$ . Now, over mini-batch, apply the batch normalization to activation 'A'. The outputs obtained for batch normalization are: mini-batch mean using Eq. (3), variance using Eq. (4), and batch normalization output using Eq. (5) (Ioffe & Szegedy, 2015).

$$\mu\beta \leftarrow \frac{1}{l} \sum_{i=1}^l A_i \quad (3)$$

$$\sigma^2\beta \leftarrow \frac{1}{l} \sum_{i=1}^l (A_i - \mu\beta)^2 \quad (4)$$

$$Y_i \leftarrow \gamma \frac{(A_i - \mu\beta)}{\sqrt{(\sigma^2\beta + \epsilon)}} + \beta \quad (5)$$

The ReLU threshold the matrix of activation to zero. However, if the learning rate is set high, some neurons will not be activated during the training and that leads to dead ReLU. To avoid this, DarkNet uses a leaky ReLU unit. The DarkNet-19 uses leaky(x):  $x$  varies from  $[1:1:18]$ , and DarkNet-53 uses leakyReLU (x):  $x$  varies from  $[1:1:52]$ . Global Average Pooling, i.e., 'avg1' generates a single feature map for each corresponding category of the classification task. In the proposed methodology, the global average pooling has the activation  $1 \times 1 \times 1024$ . Overfitting is not possible in this layer as no parameter is to be optimized. The output of 'avg1' is the input for the final convolution layer, i.e., new conv having activation  $1 \times 1 \times 4$ . The output of the final convolution layer is fed to the softmax layer with activation  $1 \times 1 \times 4$ . The final predicted output of testing images is obtained from the classification output layer, i.e., new class output. In the medical image analysis field, accuracy is not sufficient to compare the classification performance of pre-trained models. Thus, the detailed evaluation of the DarkNets pre-trained model is done using the following parameters: Critical Success Index (CSI), AUC, Sensitivity, Specificity, NPV, Precision, Mathews Correlation coefficient (MCC), and F1-score.

### 3.2. Phase-2: Tumor localization

The feature map of the best pre-trained DarkNet model is used to localize tumors in three types of tumor detected in multi-orientation views. The first convolution layer (conv1) of the pre-trained network detects features like color, texture, and edges. Further, evaluate the testing set of brain tumor data using the best performing pre-trained DarkNet model. Examine the output activations and scale the activations from range 0 to 1 as activation may have any value. The activation parameters are (a) white pixel signifies strong positive activation, (b) black pixel signifies strong negative activation, and (c) gray signifies no activation. However, deeper layers need to be analyzed to efficiently localize multiple tumors in multi-orientation. Thus, the tumor is localized using an updated deeper layer and an updated layer of the best performing pre-trained DarkNet model.

### 3.3. Phase-3: Segmentation of tumor using superpixel technique

In phase-3 tumor is segmented using the superpixel technique from multi-orientation T1W-CE MR images. Initially, skull stripping is performed on the input testing MR images containing the tumor. The skull from the MR images is removed using the Brain Extraction Tool (BET) (Jenkinson, Pechaud, & Smith, 2005; Smith, 2002). This skull stripping step is required to eliminate the unnecessary superpixel creation on the skull portion. Then, the tumor is segmented using the superpixel segmentation technique by grouping pixels into informative regions. Using superpixel technique, tumor containing MRI is partitioned into equal-sized square shape patches. The superpixels technique divides an image into a set of structural regions with similar pixels. Further, the shape of the superpixel is confined in the range 0 to infinity using the compactness parameter (' $m$ ' denotes the compactness value). The distance ' $D$ ' between two pixels ' $r$ ' and ' $s$ ' of input MR image ' $M$ ' is calculated using Eqs. (6)–(8). Here  $D(I)$  is the intensity-based distance and  $D(s)$  is the spatial distance between two adjacent pixels ' $r$ ' and ' $s$ '.

$$D(I) = \sqrt{(Mr - Ms)^2} \quad (6)$$

$$D(S) = \sqrt{(Xr - Xs)^2 + (Yr - Ys)^2 + (Zr - Zs)^2} \quad (7)$$

$$D = \sqrt{(D(S)/S)^2 + (D(I)/m)^2} \quad (8)$$

In the proposed work, super-pixels are merged using spatial distance  $D(s)$  and intensity  $D(I)$  between adjacent pixels ' $p$ ' ( $X_p, Y_p$ ) and ' $q$ ' ( $X_q, Y_q$ ). Further, the K-means clustering technique partitions the testing data into ' $k$ ' mutually exclusive clusters. Then, K-means clustering is used to cluster the color feature of each superpixel and returns the index of the cluster. Pixels are assigned the required label using the nearest cluster center. In the proposed algorithm, for the tumor containing T1W-CE MR images, numerous clusters are obtained. The desired tumor region is obtained based on the minimum area criterion. The performance of the superpixel technique is evaluated based on the dice index, and error in area calculation of tumor.

## 4. Results and discussion

This section illustrates the brief details of performance evaluation of tumor classification and localization using the DarkNets model. Further, the superpixel segmentation technique is examined for multi-class tumor segmentation. The results are obtained on the Matlab 2020a software version installed on Intel core i7 processor 10th Generation. The system is having 8 GB of RAM, a 64-bit operating system with a CPU of 2.60 GHz clock speed, and a GPU of 8 GB capacity Geforce RTX 2070 supermax design.

**Table 1**

Hyper-parameters settings for pre-trained DarkNets model.

Sr. No.	Details	Quantitative value
<b>A. DarkNet pre-trained model parameters</b>		
1	Layers	184 (DarkNet-53)64 (DarkNet-19)
2	Size of the input image	$256 \times 256 \times 3$
3	Final convolution (new conv)	Weight- $1 \times 1 \times 1024 \times 4$ Bias- $1 \times 1 \times 4$
<b>B. Optimization parameters</b>		
4	Epochs	80
5	Weight learn rate factor	20 units
6	Bias learn rate factor	20 units
7	New class output	4 classes
8	Mini batch	32
9	Learning rate	0.0004
10	Optimizer	'sgdm'
11	Verbose	False
12	Validation frequency	20

### 4.1. Performance comparison of DarkNets

The proposed work aims to classify brain tumors into multiple classes in multi-orientation directions (axial, coronal, and sagittal). The DarkNets models have trained up to 80 epochs on augmented training data. The pre-specified training options are as follows: (a) the number of classes: 4 (Normal, Meningioma, Glioma, and Pituitary tumor), (b) the number of epochs, (c) initial learning rate, (d) size of mini-batch. The brief details of the hyper-parameters are put forth in Table 1.

The performance of the DarkNet model is evaluated using performance evaluation parameters. The performance evaluation parameters are False Positive (FP), True Positive (TP), True Negative (TN), and False Negative (FN), AUC, accuracy (Eq. (9)), sensitivity (Eq. (10)), specificity (Eq. (11)), precision (Eq. (12)), Negative Predictive Value (NPV) (Eq. (13)), F1-score (Eq. (14)) and Matthews Correlation Coefficient (MCC) (Eq. (15)).

$$Accuracy = \frac{(TP + TN)}{(P + N)} \quad (9)$$

$$Sensitivity = \frac{TP}{(TP + FN)} \quad (10)$$

$$Specificity = \frac{TN}{(TN + FP)} \quad (11)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (12)$$

$$NPV = \frac{TN}{(TN + FN)} \quad (13)$$

$$F1 - score = \frac{2 \cdot TP}{(2 \cdot TP + FP + FN)} \quad (14)$$

$$MCC = \frac{(TP \cdot TN) - (FP \cdot FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (15)$$

The Region Operating Characteristic (ROC) curve of DarkNet-19 and DarkNet-53 with and without augmentation is shown in Fig. 3. From the detailed analysis of obtained results, it is concluded that the DarkNet-53 pre-trained model with data augmentation is the best performing model. The convergence graph of accuracy and loss function w.r.t. epoch for the finest performing pre-trained DarkNet-53 model trained on augmented data is put forth in Fig. 4.

Table 2 presents the performance evaluation parameters of the pre-trained DarkNet models used on the considered T1W-CE MRI dataset. Table 3 put forth the comparative analysis of the proposed technique for tumor classification with the existing techniques available in the literature. Also, the time consumption in training the best performing pre-trained DarkNet-53 model on augmented data is 81.4 min on a single GPU. The testing is done in a time duration of 8.24 s. The comparative analysis of the proposed methodology on the T1-W CE MRI dataset with state-of-the-art proves the superiority of the technique

**Table 2**

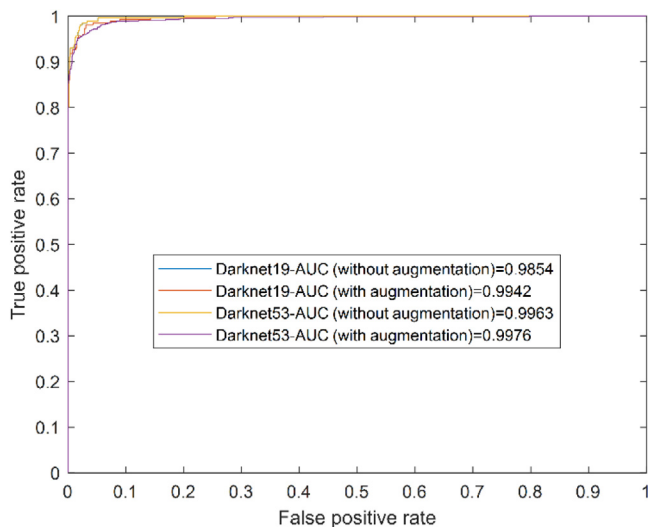
Performance evaluation of pre-trained DarkNets model on the testing set (N—Normal, M—Meningioma, G—Glioma, P—Pituitary)..

Classes	TP	TN	FP	FN	Accuracy (%)	Precision (%)	NPV (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	MCC (%)
<b>(A) DarkNet-19 with wavelet+geometrical augmentation</b>											
N	277	793	0	0	100	100	100	100	100	100	100
M	292	767	11	0	98.97	96.40	100	98.59	98.15	100	97.47
G	255	809	1	5	99.44	99.60	98.10	99.88	98.84	99.39	98.47
P	233	828	1	8	99.16	99.61	96.70	99.88	98.11	99.04	97.58
<b>(B) DarkNet-53 with wavelet+geometrical augmentation</b>											
N	277	793	0	0	100	100	100	100	100	100	100
M	296	765	7	2	99.16	97.69	99.33	99.09	98.50	99.74	97.92
G	254	813	2	1	99.72	99.22	99.61	99.75	99.41	99.88	99.23
P	234	830	0	6	99.44	100	97.50	100	98.73	99.28	98.39

**Table 3**

Performance comparison of the proposed methodology for tumor classification on the T1W-CE brain MRI dataset with the existing methods.

Technique(s)	Augmentation	Augmentation techniques	Accuracy (%)	Sensitivity (%)	Specificity (%)
VGG19 (Swati et al., 2019)	x	NA	94.82	94.25	94.69
ResNet50+ global average pooling (Kumar et al., 2021)	✓	Rotation by 0°, 90°, 180°, and 270°	97.08	97.20	–
GoogleNet (Deepak & Ameer, 2019)	x	NA	97.10	97.60	98.96
Gray Level co-occurrence Matrix (GLCM) and Bag-of-Words (BoW) model (Cheng et al., 2015)	✓	Tumor region augmentation with ring form partition	91.28	89.90	95.70
MANet (Shaik & Cherukuri, 2021)	x	NA	95.10	–	–
Transfer learning (VGG16,VGG19, ResNet50, DenseNet21) (Irmak, 2021)	x	NA	99.02(ResNet50)	–	–
Proposed	✓	2-levels of wavelet decomposition + geometrical operations	99.43	98.84	99.60

**Fig. 3.** ROC curve of DarkNets model with and without augmentation on testing data.

for tumor diagnosis. To prove the efficacy of the proposed technique, the state-of-the-art techniques are re-implemented for a four-class classification framework on the considered T1W-CE MRI dataset. A brief comparison with the proposed methodology is shown in Table 4.

#### 4.2. Tumor localization using feature maps from DarkNet-53

The tumor localization task is performed on MR images classified as Meningioma, Glioma, and Pituitary tumor classes from testing data. The tumor is localized using feature maps of the best performing model, i.e., DarkNet-53. In the proposed work, edges and color details of tumors containing MR images are obtained from the feature maps extracted from the first convolution layer (conv1). The feature maps of the deeper layer are required to correctly localize three types of tumor in all three views, namely axial, coronal, and sagittal. Fig. 5 presents the tumor localized using deeper layer 'res23'. The neon pink color signifies the predicted tumor in an MR image and it is the strongest learned feature by the DarkNet-53 model. The neon green color signifies the background area in the MR brain image.

The tumor localized images helped the colormap-based superpixel technique to improve the tumor pixels localization and that leads to improvement in the dice index as well. The deeper layers examined for tumor localization are the final convolution layer new conv, conv52, leakyrelu52, and res23. The finest tumor localization results are obtained with 'res23'. The feature with the strongest activation 'res23' is then merged with the input image. Fig. 5 presents the tumor localized using deeper layer 'res23'. The neon pink color signifies the predicted tumor in an MR image and it is the strongest learned feature by the DarkNet-53 model. The neon green color signifies the background area in the MR brain image.



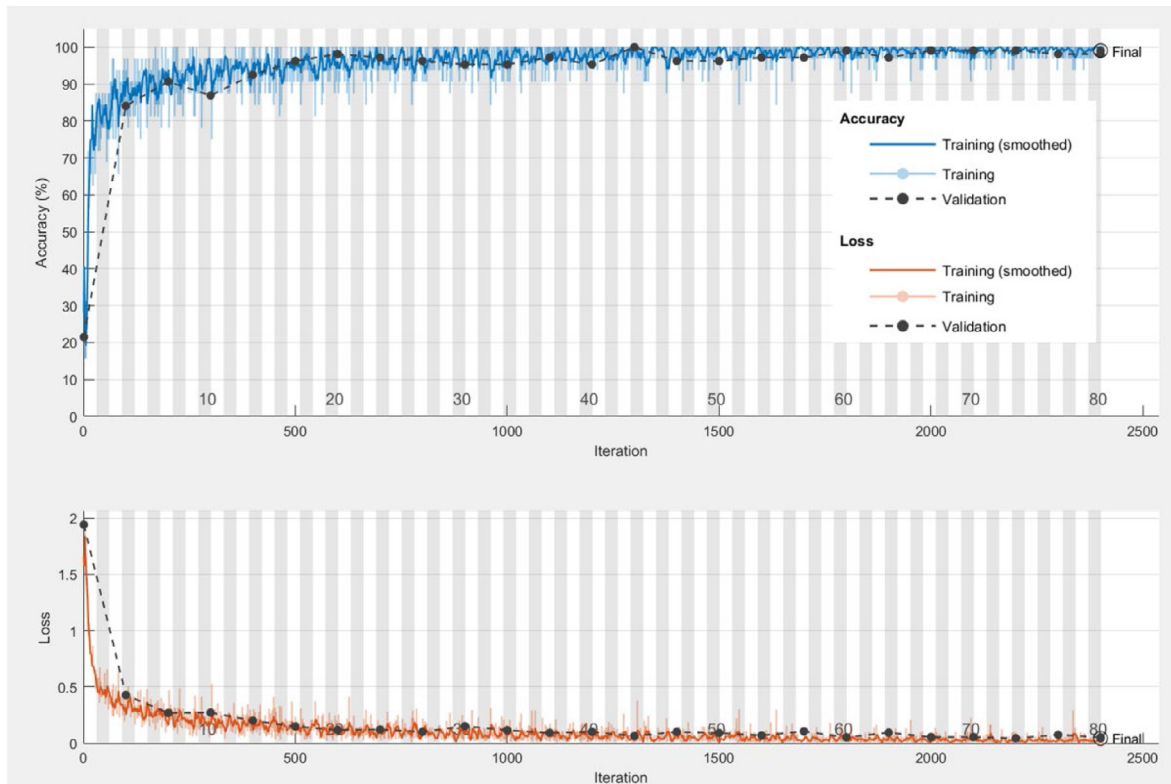


Fig. 4. Accuracy vs. epoch and loss function vs. epoch for DarkNet-53 trained on augmented data.

Table 4

Re-implementation of the existing techniques on proposed four-class classification framework on T1W-CE MRI dataset.

Technique(s)	Fine-tuning VGG19 (Swati et al., 2019)	ResNet50 (Kumar et al., 2021)	Two-channel DNN (Bodapati et al., 2021)	Proposed
Accuracy	95.68	97.28	98.12	99.43

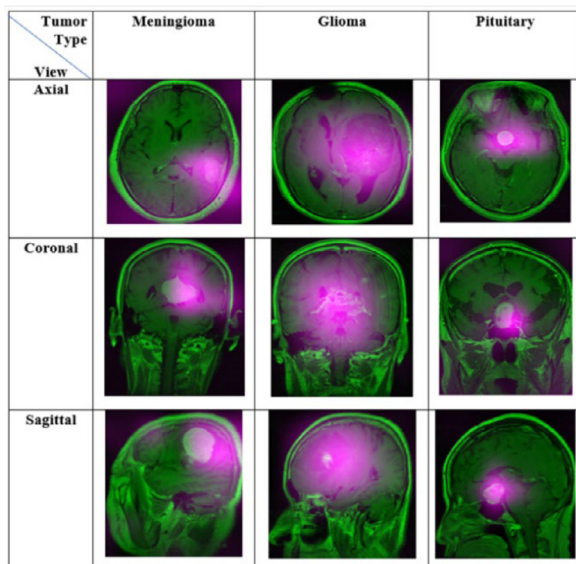


Fig. 5. Different tumor localization using DarkNet-53 pre-trained model from multi-orientation MR images.. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

#### 4.3. Tumor segmentation using colormap based superpixel technique

The T1W-CE MRI testing data is classified into four classes namely (a) Normal, (b) Meningioma, (c) Glioma, and (d) Pituitary. The colormap-based superpixel technique is used to segment tumors from the testing dataset containing three types of tumor and eliminate the normal patient slices. Out of 1070 testing MRI, the best performing pre-trained DarkNet-53 model classifies 793 images having three types of tumor (Meningioma, Glioma, and Pituitary tumor). Fig. 6 put forth the sample images of segmented tumors from tumors containing testing data in all three views (axial, coronal, and sagittal).

The performance of the superpixel segmentation technique is examined using the dice index, and error in tumor area calculation. For this, the segmented tumor is compared with the gold standard mask. The dice similarity coefficient or index varies from range 0 (no match) to 1 (100% match). The dice similarity coefficient between two segmented tumors (M seg) and gold standard mask (M ref) is calculated using Eq. (16). The error in the tumor area (Et) is calculated using Eq. (17), having 'Wp ref' represents white pixels of M ref, and 'Wp seg' represents white pixels of M seg.

$$Dice = \frac{2 * |M_{seg} \cap M_{ref}|}{|M_{seg}| + |M_{ref}|} \quad (16)$$

$$Et = \frac{\sum_{i=1}^n \sum_{j=1}^m Wp_{ref} - \sum_{i=1}^n \sum_{j=1}^m Wp_{seg}}{\sum_{i=1}^n \sum_{j=1}^m Wp_{ref}} * 100 \quad (17)$$

The proposed colormap-based superpixel technique implemented on the T1W-CE MRI dataset segmented the three types of tumor in



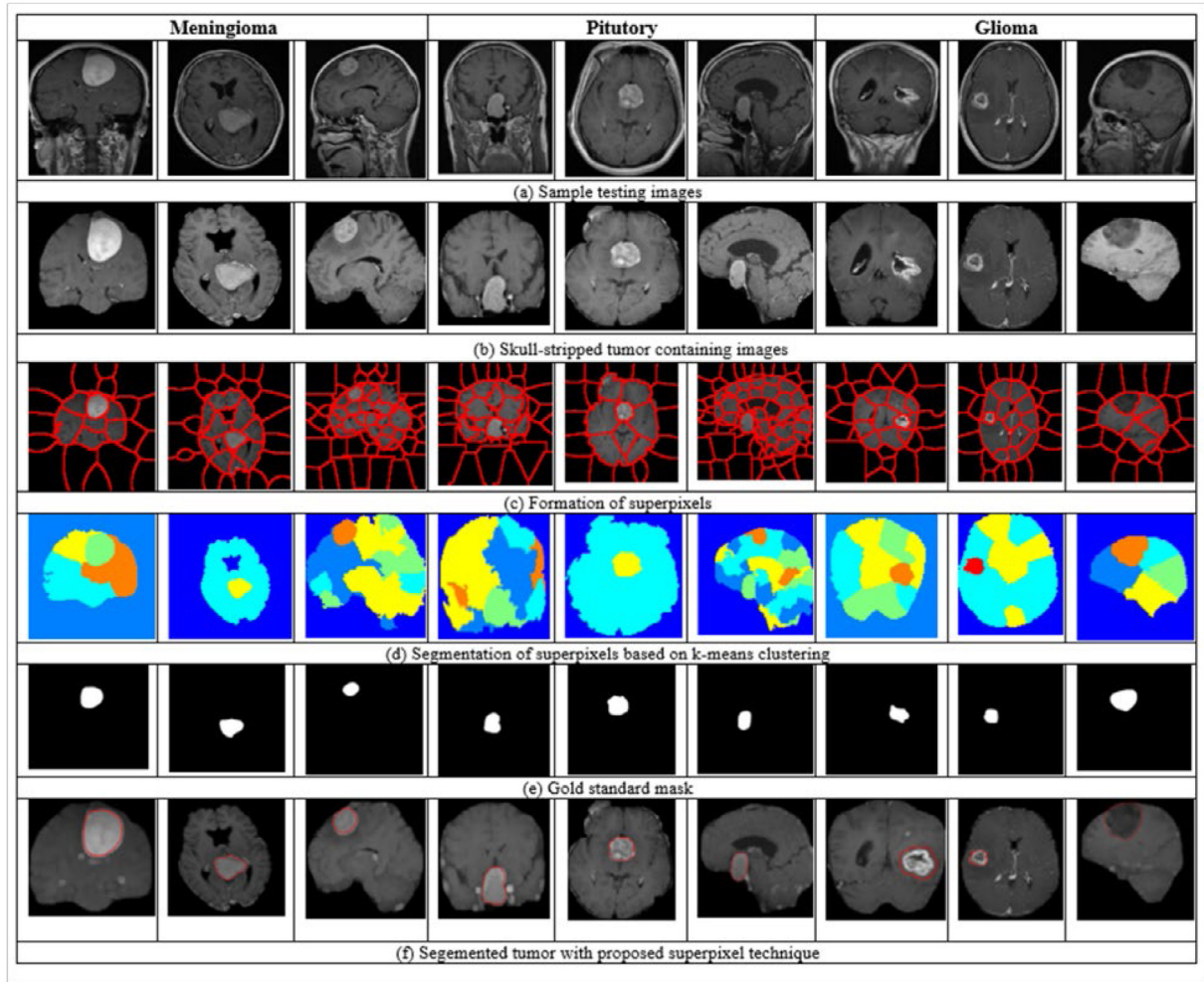


Fig. 6. Different types of tumor segmentation in all the three views using the colormap-based superpixel segmentation technique. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5

Comparative analysis of proposed tumor segmentation techniques with the techniques available in literature implemented on T1-CE MRI dataset (best-obtained values are represented in bold).

Technique	SegNet (Rehman, Naz, Naseem, Razzak, & Hameed, 2019)	MAG-Net (Gupta, Pun, Sonbhadra, & Agarwal, 2021)	Chan-Vese (Gunasekara, Kaldera, & Dissanayake, 2021)	Proposed
Dice index	0.93	0.74	0.92	<b>0.94 ± 2.6%</b>

all three views. An average error in tumor area calculation is 4.29%. Recently, attention-based mechanisms are gaining popularity due to the optimum accuracy and sensitivity in medical imaging modalities (Yang et al., 2020; Yongkai et al., 2020). The ResNet50 with feature pyramid attention model is implemented on the considered testing dataset (Liu et al., 2019). The model obtained a dice index of 0.91 0.8% in 10.12 s time duration. The attention 2-D UNet obtained a dice index of 0.87 1.6% in 15.34 duration (Noori, Bahri, & Mohammadi, 2019). The time complexity is resolved with the attention mechanism but the dice index is also dropped. The proposed segmentation technique is compared with the techniques available for segmentation as shown in Table 5. The colormap-based superpixel technique improves the similarity coefficients in comparison to the state-of-the-art in optimum time duration.

#### 4.4. Ablation study

In this section, we performed ablation studies to evaluate the impact of each component in our proposed methodology. The proposed methodology is implemented using pre-trained DarkNet53 trained on augmentation of the input T1-CE MRI dataset using two-level decomposition+ geometric parameters. For segmentation, the colormap-based 2-D superpixel technique is evaluated. The proposed work is further investigated for the following research queries: (1) How would different pre-trained CNN models affect the performance on the considered T1-CE MRI dataset? (2) How does the optimizer affect the performance of the best performing pre-trained model? (3) How does variation in the cross-validation scheme impact the classification performance of DarkNet53? (4) Impact of levels of wavelet decomposition on classification performance, and (5) How well the proposed methodology is effective on different multi-modal MRIs using the BraTS 2019 dataset?

**Table 6**  
Performance comparison of different pre-trained networks on T1W-CE MRI dataset.

Network	Input image size	Parameters (in millions)	Depth	Updated layers	Training accuracy (%)	Time (in minutes)
SqueezeNet	$227 \times 227 \times 3$	2	18	68	93.5	42.4
DenseNet201	$224 \times 224 \times 3$	20	201	708	92.5	163.2
ResNet18	$224 \times 224 \times 3$	12	18	71	91.9	54.6
Inception-resnetv2	$299 \times 299 \times 3$	56	164	824	95.3	154.4
Inceptionv3	$299 \times 299 \times 3$	24	48	315	96.2	82.5
MobileNetv2	$224 \times 224 \times 3$	4	53	154	95.6	63.2
DarkNet19	$256 \times 256 \times 3$	21	19	64	97.9	44.8
DarkNet53	$256 \times 256 \times 3$	42	53	184	99.6	81.4

**Table 7**  
Comparative analysis of different optimizer functions performance with DarkNet53 pre-trained network on T1W-CE MRI dataset.

Optimizer	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	NPV (%)	MCC (%)
Sgdm	99.58	99.38	99.40	99.80	99.38	99.80	99.19
Adam	99.43	98.95	98.84	99.60	98.89	99.62	98.52
RMSprop	99.29	98.72	98.58	99.52	98.62	99.54	98.17

**Table 8**  
Comparison of accuracy obtained in state-of-the-art with proposed methodology using different cross-validation schemes for T1W-CE MRI dataset.

Technique	Data augmentation	Cross-validation	Accuracy
ResNet50 (Kumar et al., 2021)	x	5-fold	97.48
ResNet50 (Kumar et al., 2021)	✓	5-fold	97.08
GLCM+SVM (Cheng et al., 2015)	✓	5-fold	89.72
Two channel DNN (Bodapati et al., 2021)	x	5-fold	97.23
Proposed	✓	5-fold	99.10
Proposed	✓	Hold-out (8:1:1)	99.43

**Table 9**  
Comparative analysis of the performance of data augmentation technique on the proposed methodology.

Augmentation	Training accuracy (%)	Testing accuracy (%)	Validation accuracy (%)
No augmentation	90.15	88.10	91.23
Geometric	93.90	92.61	92.55
Geometric+Level-1 wavelet decomposition	97.15	96.40	94.42
Geometric+Level-2 wavelet decomposition	99.60	98.54	98.81
Geometric+Level-3 wavelet decomposition	97.80	96.82	97.52

#### (6) Impact of pruning technique on classification and segmentation of tumor.

Firstly, the performance of different pre-trained CNN models is evaluated on the considered T1W-CE MRI dataset. The hyperparameters settings are the same as mentioned in Table 1 with the early stopping criterion. Table 6 presents brief details of the performance comparison parameters and depicts that DarkNet53 outperformed in comparison to other pre-trained networks. The performance of the best performing pre-trained model is evaluated for three different optimizer functions namely (a) stochastic gradient descent with momentum (sgdm), (b) RMSprop and (c) Adam. Table 7 shows that the DarkNet53 pre-trained network with an 'sgdm' optimizer achieved the best performance on the testing dataset. To briefly explain the impact of cross-validation schemes, the proposed methodology is evaluated with two types of cross-validation schemes: (a) hold out cross-validation scheme and (b) 5-fold cross-validation scheme. Table 8 put forth the detailed comparative analysis of the proposed cross-validation techniques with state-of-the-art techniques available for the considered dataset.

The wavelet augmentation technique is an important rationale of the proposed methodology. Thus, the T1W-CE MRI dataset is evaluated on different augmentation conditions as shown in Table 9. The obtained results from the best performing DarkNet53 model with the proposed data augmentation approach (Geometric+Level-2 wavelet decomposition) ensure the reduction of overfitting and provide the optimum performance with limited medical data. Fig. 7 put forth the Glyph-Plot to evaluate the overall performance and correlation of numerous performance parameters (accuracy, and AUC) for DarkNet models.

To prove the robustness of the proposed technique it is investigated on a multi-modality brain MRI dataset obtained from BraTS 2018 dataset (Florian et al., 2020; Menze et al., 2015). The BraTS 2018 dataset contains four types of modalities namely (a) T1 (b) post-contrast T1-weighted (T1Gd), (c) T2-weighted (T2-w) and (d) T2 Fluid Attenuated Inversion Recovery (T2-FLAIR). The multi-modality BraTS dataset is pre-processed of size  $224 \times 224 \times 155$  (co-registered), skull stripped, and isotropic resolution of  $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$ . The training dataset contains 75 Low-Grade Glioma (LGG) patients and 210 High-Grade Glioma (HGG) patients. In the proposed methodology, for each patient 20 slices containing tumor pixels are considered for each modality. Further, N4ITK bias field correction is used for finer segmentation results (Tustison et al., 2010).

Table 10 presents the performance evaluation parameters obtained using the proposed methodology on multi-modality MRI for binary classification. The proposed methodology obtained an average classification accuracy of 98.15% on multi-modal MRI. On BraTS 2018 dataset, the existing techniques obtained classification accuracy of 93.69% with two-channel DNN (Bodapati et al., 2021), 92.50% with the Inception V3 pre-trained CNN model (Sharif, Li, Khan, & Saleem, 2020). Thus, the proposed technique is robust for multi-modality MRI in comparison to the state-of-the-art techniques.

In recent times, pruning techniques are suggested to reduce computation time and memory space consumption of deep learning models (Pietron & Wielgosz, 2020). To evaluate the impact of pruning on the proposed methodology, the filter size of all convolution layers is reduced. The updated model is retrained with 20 epochs and the accuracy dropped by  $-0.75\%$ . Further, processing speed is increased

**Table 10**

Performance evaluation parameters obtained for multi-modality MRI using the proposed methodology for binary classification of tumor grades.

Modality	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	NPV (%)	MCC (%)
T1	96.17	94.20	97.30	97.30	95.53	94.20	91.44
T2	97.90	97.31	98.34	98.34	97.78	97.31	95.64
T1-CE	98.95	98.65	99.20	99.20	98.89	98.65	97.81
Flair	99.60	99.45	99.70	99.70	99.59	99.45	99.15
Mean	98.15	97.40	98.63	98.63	97.94	97.40	96.01

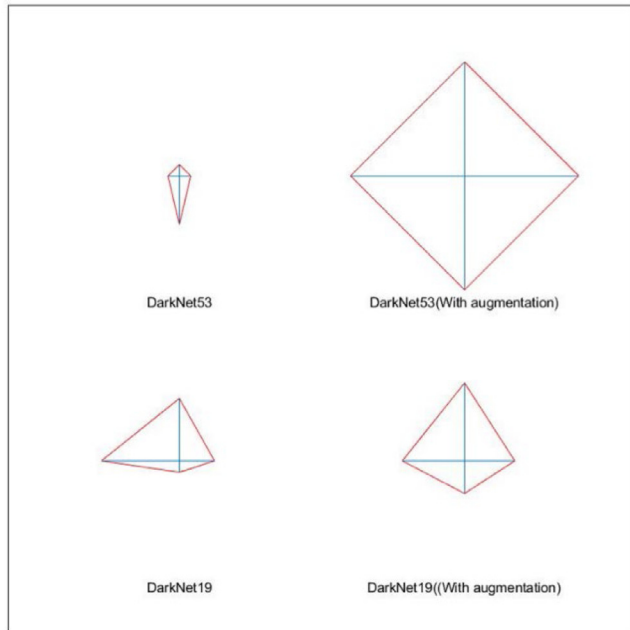


Fig. 7. Glyph plot to compare the overall performance of DarkNets with and without augmentation.

and time consumption is reduced by 26% on a single GPU. However, the size of the pre-trained DarkNet53 model is 155 MB so the training speed is not a major concern with the proposed technique. Thus, the pruning technique is not helpful for the proposed methodology.

#### 4.5. Limitations and future scope

The main limitations of the proposed work are as follows: (a) implemented for 2-D radiographic images, (b) T1 modality is classified with least accuracy. To implement the proposed transfer learning-based model on the MRI dataset, the input images are converted into .png format. The image conversion does not affect the classification and tumor localization performance due to the transfer learning-based fine-tuning approach implemented using a pre-trained model. However, due to the reduction in image pixel quality, the dice index is slightly impacted for small size tumors (for initial slices containing small tumor pixels).

For future study, the proposed methodology needs to be investigated on 3-D superpixel segmentation techniques for volumetric analysis. The performance evaluation is to be obtained for several other larger-scale multi-sequence MRI data such as Diffusion Tensor Imaging (DTI), Magnetic Resonance Spectroscopy (MRS), Diffusion-Weighted Imaging (DWI), etc. Also, the computer-aided technology needs to be implemented with the proposed framework on the real-time clinical dataset.

## 5. Conclusions

Radiography images (MRI, CT, PET, etc.) in association with artificial intelligence-based techniques are useful for the accurate diagnosis

of the disease. The multi-class classification of brain tumors is done using pre-trained DarkNet models. The comparative analysis of the pre-trained DarkNet model on the T1W-CE MRI dataset with augmentation (wavelet decomposition and geometrical parameter) reduces the data overfitting issue. The proposed methodology achieved the highest training accuracy of 99.60% with an sgdm optimizer for tumor classification using the pre-trained DarkNet-53 model. On 1070 testing images, an AUC of 0.99, and testing accuracy of 98.54% is achieved. From testing data, the best performing DarkNet-53 model classifies 277 images into normal and 793 images into tumor classes (Meningioma, Glioma, and Pituitary tumor). Further, the superpixel technique is used to segment three types of tumors from 793 testing images having a tumor. The superpixel technique segmented tumor in all three views (axial, coronal, and sagittal) with an average dice index of  $0.94 \pm 2.6\%$  and an average error in tumor area calculation is 4.29%. The dice index and classification performance are improved with the proposed method as compared to the state-of-the-art method.

## CRedit authorship contribution statement

**Sakshi Ahuja:** Conceptualization, Methodology, Software, Writing – original draft. **Bijaya Ketan Panigrahi:** Project administration, Supervision, Investigation, Visualization. **Tapan Kumar Gandhi:** Supervision, Formal analysis, Investigation, Validation.

## References

- A.-Dmour, H., & A.-Ani, A. (2018). A clustering fusion technique for mr brain tissue segmentation. *Neurocomputing*, 275, 546–559. <http://dx.doi.org/10.1016/j.neucom.2017.08.051>.
- A.-Ellah, M. K., Awad, A., Khalaf, A. A. M., & Hamed, H. F. A. (2018). Two-phase multi-model automatic brain tumour diagnosis system from magnetic resonance images using convolutional neural networks. *EURASIP Journal on Image and Video Processing Volume*, (1–10), <http://dx.doi.org/10.1186/s13640-018-0332-4>.
- Abiwinanda, N., Hanif, M., Hesaputra, S. T., Handayani, A., & Mengko, T. (2019). Brain tumor classification using convolutional neural network. In *World congress on medical physics and biomedical engineering*, vol. 68 (pp. 183–189). Singapore: Springer, [http://dx.doi.org/10.1007/978-981-10-9035-6\\_33](http://dx.doi.org/10.1007/978-981-10-9035-6_33).
- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34, 2274–2282. <http://dx.doi.org/10.1109/TPAMI.2012.120>.
- Amin, J., Sharif, M., Yasmin, M., & Fernandes, S. L. (2018). Big data analysis for brain tumor detection: Deep convolutional neural networks, future generation computer systems. <http://dx.doi.org/10.1016/j.future.2018.04.065>.
- Amiri, S., Mahjoub, M. A., & Rekik, I. (2018). Tree-based ensemble classifier learning for automatic brain glioma segmentation. *Neurocomputing*, 313, 135–142. <http://dx.doi.org/10.1016/j.neucom.2018.05.112>.
- Angulakshmi, M., & Priya, G. L. (2018). Brain tumour segmentation from mri using superpixels based spectral clustering. *Journal of King Saud University-Computer and Information Sciences*, <http://dx.doi.org/10.1016/j.jksuci.2018.01.009>.
- Ayadi, W., Elhamzi, W., Charfi, I., et al. (2021). Deep cnn for brain tumor classification. *Neural Process Letter*, 53, 671–700. <http://dx.doi.org/10.1007/s11063-020-10398-2>.
- Bjoern, M., Andras, J., Stefan, B., Jayashree, K.-C., Farahani, et al. (2014). The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 33. <http://dx.doi.org/10.1109/TMI.2014.2377694>, URL: <https://hal.inria.fr/hal-00935640>.
- Board, C. E. (2019). Brain tumor: statistics. URL: <https://www.cancer.net/cancer-types/brain-tumor/statistics>.
- Bodapati, J., Shaik, N. S., Naralasetti, V., et al. (2021). Joint training of two-channel deep neural network for brain tumor classification. *Signal, Image and Video Processing Volume*, 15, 753–760. <http://dx.doi.org/10.1007/s11760-020-01793-2>.



- Chaplot, S., Patnaik, L., & Jagannathan, N. (2006). Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. *Biomedical Signal Processing and Control*, 1, 86–92. <http://dx.doi.org/10.1016/j.bspc.2006.05.002>.
- Chen, X., Nguyen, B. P., Chui, C.-K., & Ong, S.-H. (2016). Automated brain tumor segmentation using kernel dictionary learning and superpixel-level features. In *2016 IEEE international conference on systems, man, and cybernetics* (pp. 002547–002552). <http://dx.doi.org/10.1109/SMC.2016.7844622>.
- Chen, W., Qiao, X., Liu, B., Qi, X., Wang, R., & Wang, X. (2017). Automatic brain tumor segmentation based on features of separated local square. In *2017 Chinese automation congress* (pp. 6489–6493). <http://dx.doi.org/10.1109/CAC.2017.8243946>.
- Cheng, J. (2017). Brain tumor dataset. <http://dx.doi.org/10.6084/m9.figshare.1512427.v5>, URL: [https://figshare.com/articles/brain\\_tumor\\_dataset/1512427/5](https://figshare.com/articles/brain_tumor_dataset/1512427/5).
- Cheng, J., Huang, W., Cao, S., Yang, R., Yang, W., Yun, Z., et al. (2015). Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PLoS One*, <http://dx.doi.org/10.1371/journal.pone.0140381>.
- Community, C. S. (2013). Frankly speaking about brain tumors.
- Das, S., Aranya, OFMRR, & Labiba, N. N. (2019). Brain tumor classification using convolutional neural network. In *1st International Conference on Advances in Science, Engineering and Robotics Technology* (pp. 1–5). <http://dx.doi.org/10.1109/ICASERT.2019.8934603>.
- Deepak, S., & Ameer, P. (2019). Brain tumor classification using deep cnn features via transfer learning. *Computers in Biology and Medicine*, 111, 1–7. <http://dx.doi.org/10.1016/j.combiomed.2019.103345>.
- Diniz, P. H. B., Valente, T. L. A., Diniz, J. O. B., Silva, A. C., Gattass, M., Ventura, N., et al. (2018). Detection of white matter lesion regions in mri using slico and convolutional neural network. *Computer Methods and Programs in Biomedicine*, 167, 49–63. <http://dx.doi.org/10.1016/j.cmpb.2018.04.011>.
- Dodge, S., & Karam, L. (2016). Understanding how image quality affects deep neural networks. *arXiv:1604.04004*.
- Dong, H., Yang, G., Liu, F., Mo, Y., & Guo, Y. (2017). Automatic brain tumor detection and segmentation using u-net based fully convolutional networks. *arXiv:1705.03820*.
- Florian, K., Christoph, B., Diana, W., Jana, L., Ivan, E., Giles, T., et al. (2020). Brats toolkit: Translating brats brain tumor segmentation algorithms into clinical and scientific practice. *Frontiers in Neuroscience*, 14, <http://dx.doi.org/10.3389/fnins.2020.00125>.
- G.P. Nason, B. S. (1995). *The Stationary Wavelet Transform and Some Statistical Applications Vol. 103*. Springer, Group, B. I. A. Ixi dataset. URL: <http://brain-development.org/ixi-dataset/>.
- Gunasekara, S. R., Kaldera, H. N. T. K., & Dissanayake, M. B. (2021). A systematic approach for mri brain tumor localization and segmentation using deep learning and active contouring. *Journal of Healthcare Engineering*, <http://dx.doi.org/10.1155/2021/6695108>.
- Gupta, T., Gandhi, T. K., Gupta, R., & Panigrahi, B. (2017). Classification of patients with tumor using mr flair images. *Pattern Recognition Letters*, <http://dx.doi.org/10.1016/j.patrec.2017.10.037>.
- Gupta, S., Pun, N. S., Sonbhadra, S. K., & Agarwal, S. (2021). Mag-net: Multi-task attention-guided network for brain tumor segmentation and classification. *arXiv:2107.12321*.
- Ho, N., & Kim, Y. (2021). Evaluation of transfer learning in deep convolutional neural network models for cardiac short axis slice classification. *Scientific Reports*, 1839, 1–11. <http://dx.doi.org/10.1038/s41598-021-81525-9.22>.
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv:1502.03167*.
- Irmak, E. (2021). Multi-classification of brain tumor mri images using deep convolutional neural network with fully optimized framework. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, 45, 1015–1036. <http://dx.doi.org/10.1007/s40998-021-00426-9>.
- J., Pesquet, Krim, H., & Carfantan, H. (1996). Time-invariant orthonormal wavelet representations. *IEEE Transactions on Signal Processing*, 44, 1964–1970. <http://dx.doi.org/10.1109/78.533717>.
- Jenkinson, M., Pechaud, M., & Smith, S. (2005). Bet2: Mr-based estimation of brain, skull and scalp surfaces.
- Kong, Y., Wu, J., Yang, G., Zuo, Y., Chen, Y., Shu, H., et al. (2019). Iterative spatial fuzzy clustering for 3d brain magnetic resonance image supervoxel segmentation. *Journal of Neuroscience Methods*, 311, 17–27. <http://dx.doi.org/10.1016/j.jneumeth.2018.10.007>.
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). Imagenet classification with deep convolutional neural networks. *Neural Information Processing Systems*, 25, <http://dx.doi.org/10.1145/3065386>.
- Kumar, R., Kakarla, J., Isunuri, B. V., et al. (2021). Multi-class brain tumor classification using residual network and global average pooling. *Multimedia Tools and Applications*, 80, 13429–13438. <http://dx.doi.org/10.1007/s11042-020-10335-4>.
- Lateef, F., & Ruichek, Y. (2019). Survey on semantic segmentation using deep learning techniques. *Neurocomputing*, 338, 321–348. <http://dx.doi.org/10.1016/j.neucom.2019.02.003>.
- Li, C. (2021). Mri segmentation and bias field correction. url: <https://in.mathworks.com/matlabcentral/fileexchange/59752-mri-segmentation-and-bias-field-correction>.
- Liu, Y., et al. (2019). Automatic prostate zonal segmentation using fully convolutional network with feature pyramid attention. *IEEE Access*, 7, 163626–163632. <http://dx.doi.org/10.1109/ACCESS.2019.2952534>.
- Lorenzo, P. R., Nalepa, J., B.-Billewicz, B., Wawrzyniak, P., Mrukwa, G., Kawulok, M., et al. (2019). Segmenting brain tumors from flair mri using fully convolutional neural networks. *Computer Methods and Programs in Biomedicine*, 176, 135–148. <http://dx.doi.org/10.1016/j.cmpb.2019.05.006>.
- Mehrotra, R., Ansari, M., Agrawal, R., & Anand, R. (2020). A transfer learning approach for ai-based classification of brain tumors. *Machine Learning with Applications*, 2, 1–12. <http://dx.doi.org/10.1016/j.mlwa.2020.100003>.
- Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., et al. (2015). The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging*, 34, 1993–2024. <http://dx.doi.org/10.1109/TMI.2014.2377694>.
- Mohamed, E., Umer, N. M., Qunfeng, T., David, S., John-Paul, G., Catherine, B., et al. (2021). The effectiveness of image augmentation in deep learning networks for detecting covid-19: A geometric transformation perspective. *Frontiers in Medicine*, 8(153), <http://dx.doi.org/10.3389/fmed.2021.629134>.
- Mohammadreza, S., et al. (2018). Supervised learning based multimodal mri brain tumour segmentation using texture features from supervoxels. *Computer Methods and Programs in Biomedicine*, 157, 69–84. <http://dx.doi.org/10.1016/j.cmpb.2018.01.003>.
- Nayak, D. R., Dash, R., & Majhi, B. (2016). Brain mr image classification using two-dimensional discrete wavelet transform and adaboost with random forests. *Neurocomputing*, 18, 8–197. <http://dx.doi.org/10.1016/j.neucom.2015.11.034>.
- Noori, M., Bahri, A., & Mohammadi, K. (2019). Attention-guided version of 2d unet for automatic brain tumor segmentation. In *9th International conference on computer and knowledge engineering* (pp. 269–275). <http://dx.doi.org/10.1109/ICCKE48569.2019.8964956>.
- Observatory, T. G. C. (2018). Globocan 2018: India factsheet. URL: <http://cancerindia.org.in/globocan-2018-india-factsheet/>.
- Ostrom, Q. T., Cioffi, G., Gittleman, H., Patil, N., Waite, K., Kruchko, C., et al. (2019). Cbtrus statistical report: primary brain and other central nervous system tumors diagnosed in the united states in 2012–2016. *Neuro-Oncology*, 21, v1–v100.
- Ostrom, Q. T., Gittleman, H., Fulop, J., Liu, M., Blanda, R., Kromer, C., et al. (2015). Cbtrus statistical report: primary brain and central nervous system tumors diagnosed in the united states in 2008–2012. *Neuro-Oncology*, 17, iv1–iv62.
- Patel, A. P., Fisher, J. L., Nichols, E., A.-Allah, F., & J. Abdela, e. a. (2019). Global, regional, and national burden of brain and other cns cancer, 1990–2016: a systematic analysis for the global burden of disease study 2016. *The Lancet Neurology*, 18, 376–393. [http://dx.doi.org/10.1016/S1474-4422\(18\)30468-X](http://dx.doi.org/10.1016/S1474-4422(18)30468-X).
- Pietron, M., & Wielgosz, M. (2020). Retrain or not retrain? - efficient pruning methods of deep cnn networks. In *Lecture notes in computer science*, vol. 12139. Cham: Springer, [http://dx.doi.org/10.1007/978-3-030-50420-5\\_34](http://dx.doi.org/10.1007/978-3-030-50420-5_34).
- Pinto, A., Pereira, S., Rasteiro, D., & Silva, C. A. (2018). Hierarchical brain tumour segmentation using extremely randomized trees. *Pattern Recognition*, 82, 105–117. <http://dx.doi.org/10.1016/j.patcog.2018.05.006>.
- Ranjbarzadeh, R., Kasgari, A. B., Ghouschi, S. J., Anari, S., et al. (2021). Brain tumor segmentation based on deep learning and an attention mechanism using mri multi-modalities brain images. *Scientific Reports*, 10930. <http://dx.doi.org/10.1038/s41598-021-90428-8>.
- Raschke, F., Barrick, T. R., Jones, T. L., Yang, G., Ye, X., & Howe, F. A. (2019). Tissue-type mapping of gliomas. *NeuroImage: Clinical*, (21), <http://dx.doi.org/10.1016/j.nicl.2018.101648>.
- 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 Systems with Applications*, 118, 598–613. <http://dx.doi.org/10.1016/j.eswa.2018.10.040>.
- Rehman, A., Naz, S., Naseem, U., Razzak, I., & Hameed, I. A. (2019). Deep autoencoder-decoder framework for semantic segmentation of brain tumor. *Australian Journal of Intelligent Information Processing Systems*, 15, 53–60.
- Sasank, V., & Venkateswarlu, S. (2021). Brain tumor classification using modified kernel based softplus extreme learning machine. *Multimedia Tools and Applications*, 80, 13513–13534. <http://dx.doi.org/10.1007/s11042-020-10423-5>.
- Shaik, N. S., & Cherukuri, T. (2021). Multi-level attention network: application to brain tumor classification. *Signal, Image and Video Processing*, <http://dx.doi.org/10.1007/s11760-021-02022-0>.
- 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 tumors using mri images. *Pattern Recognition Letters*, 129, 181–189. <http://dx.doi.org/10.1016/j.patrec.2019.11.019>.
- Shree, N. V., & Kumar, T. N. R. (2018). Identification and classification of brain tumor mri images with feature extraction using dwt and probabilistic neural network. *Brain Informatics*, (23–30), <http://dx.doi.org/10.1007/s40708-017-0075-5>.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17, 143–155. <http://dx.doi.org/10.1002/hbm.10062>.



- Soltaninejad, M., Yang, G., Lambrou, T., Allinson, N., Jones, T. L., & Barrick, T. R. (2017). Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in flair mri. *International Journal of Computer Assisted Radiology and Surgery*, 12, 183–203. <http://dx.doi.org/10.1007/s11548-016-1483-3>.
- Soltaninejad, M., Zhang, L., Lambrou, T., Yang, G., Allinson, N., & Ye, X. (2018). Mri brain tumor segmentation and patient survival prediction using random forests and fully convolutional networks. In A. Crimi, S. Bakas, H. Kuijff, B. Menze, & M. Reyes (Eds.), *Lecture notes in computer science: vol. 10670, Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries. brainles 2017*. Springer, <http://dx.doi.org/10.1007/978-3-319-75238-918>.
- Sriramakrishnan, P., Kalaiselvi, T., & Rajeswaran, R. (2019). Modified local ternary patterns technique for brain tumour segmentation and volume estimation from mri multi-sequence scans with gpu cuda machine. *Biocybernetics and Biomedical Engineering*, 39, 470–487. <http://dx.doi.org/10.1016/j.bbe.2019.02.002>.
- Swati, Z. N. K., Zhao, Q., Muhammad Kabir, F. A., Ali, Z., Ahmed, S., & Lu, J. (2019). Brain tumor classification for mr images using transfer learning and fine-tuning. *Computerized Medical Imaging and Graphics*, (34–46), <http://dx.doi.org/10.1016/j.compmedimag.2019.05.001>.
- Tandel, G. S., Balestrieri, A., Jujaray, T., Khanna, N. N., Saba, L., & Suri, J. S. (2020). Multiclassmagnetic resonance imaging brain tumor classification using artificial intelligence paradigm. *Computers in Biology and Medicine*, 122, <http://dx.doi.org/10.1016/j.compbiomed.2020.103804>.
- Tustison, N. J., et al. (2010). N4itk: Improved n3 bias correction. *IEEE Transactions on Medical Imaging*, 29, 1310–1320. <http://dx.doi.org/10.1109/TMI.2010.2046908>.
- Vasavi, S., Priyadarshini, N. K., & Vardhan, K. H. (2021). Invariant feature based darknet architecture for moving object classification. *IEEE Sensor Journal*, 21, 11417–11426. <http://dx.doi.org/10.1109/JSEN.2020.3007883>.
- Villanueva-Meyer, J., Mabray, M., & Cha, S. (2017). Current clinical brain tumor imaging. *Neurosurgery*, 81, 397–415. <http://dx.doi.org/10.1093/neuros/nyx103>.
- Yang, G., et al. (2020). Simultaneous left atrium anatomy and scar segmentations via deep learning in multiview information with attention. *Future Generation Computer Systems*, 107, 215–228. <http://dx.doi.org/10.1016/j.future.2020.02.005>.
- Yongkai, L., et al. (2020). Exploring uncertainty measures in Bayesian deep attentive neural networks for prostate zonal segmentation. *IEEE Access*, 8, Article 151817-151828. <http://dx.doi.org/10.1109/ACCESS.2020.3017168>.
- Zacharakis, E., Wang, S., Chawla, S., Yoo, D. S., Wolf, R., et al. (2009). Classification of brain tumor type and grade using mri texture and shape in a machine learning scheme. *Magnetic Resonance in Medicine*, 62, 1609–1618. <http://dx.doi.org/10.1002/mrm.22147>.
- Zhang, W., Yang, G., Huang, H., Yang, W., Xu, X., et al. (2021). Me-net: Multi-encoder net framework for brain tumor segmentation. *International Journal of Imaging Systems and Technology*, (1–15), <http://dx.doi.org/10.1002/ima.22571>.
- Zhao, J., Meng, Z., Wei, L., Sun, C., Zou, Q., & Su, R. (2019). Supervised brain tumor segmentation based on gradient and context-sensitive features. *Frontiers in Neuroscience*, 13(144), <http://dx.doi.org/10.3389/fnins.2019.00144>.
- Zhao, X., Wu, Y., Song, G., Li, Z., Zhang, Y., & Fan, Y. (2018). A deep learning model integrating fcnn and crfs for brain tumor segmentation. *Medical Image Analysis*, 43, 98–111. <http://dx.doi.org/10.1016/j.media.2017.10.002>.
- Zhao, Z., Yang, G., Lin, Y., Pang, H., & Wang, M. (2018). Automated glioma detection and segmentation using graphical models. *PLoS One*, <http://dx.doi.org/10.1371/journal.pone.0200745>.