Journal of
Information Systems Engineering
and Business Intelligence

Vol.8, No.2, October 2022
Available online at: http://e-journal.unair.ac.id/index.php/JISEBI

Hybrid Deep Learning Models for Multi-classification of Tumour from Brain MRI

Hafiza Akter Munira[1)], Md. Saiful Islam[2)*]

[1)2)]Department of Electronics and Telecommunication Engineering, Chittagong University of Engineering and Technology, Bangladesh

Kaptai, Highway Raozan Pahartali Rd, Chittagong 4349, Chittagong

[1)]hafizaakterete1601@gmail.com, [2)]saiful05eee@cuet.ac.bd

Abstract

Background: Brain tumour categorisation can be assisted with computer-aided diagnostic (CAD) for medical applications. Biopsies to classify brain tumours can be costly and time-consuming. Radiologists may also misclassify brain tumour types when handling large amounts of data with multiple classes. In this case, technological advancements and machine learning can help.
Objective: This study proposes hybrid deep learning approaches for classifying brain tumours using convolutional neural networks (CNN) and machine learning (ML) classifiers.
Methods: A new 23-layer CNN architecture is developed for brain deep feature extraction from magnetic resonance imaging (MRI). Random forest (RF) and support vector machine (SVM) classifiers are then used to evaluate the extracted in-depth features from the flattened layer of the CNN model. This study is unique because it employs CNN, CNN-RF, CNN-SVM, and tuned Inception V3 deep learning models on multi-class brain MRI datasets. The proposed hybrid method is run on two publicly available datasets.
Results: Among the four models, the CNN-RF model achieves 96.52% accuracy on the Fig share 3c dataset, while the CNN-SVM model achieves 95.41% accuracy on the large Kaggle 4c dataset with four classes (glioma, meningioma, normal, pituitary).
Conclusion: Experimental outcomes show that the hybrid techniques can significantly enhance the classification performance, especially on multi-class datasets (glioma, meningioma, normal, pituitary). This study also examines the various weight strategies for dealing with overfitting analytics.

Keywords: Brain Tumour, Convolutional Neural Network, Feature Extraction, Multi-Classification, Machine Learning Classifiers

Article history: Received 28 June 2022, first decision 16 August 2022, accepted 12 September 2022, available online 28 October 2022

I. INTRODUCTION

The brain is the most vital and intricate organ, regulating the entire nervous system, with 100 billion nerve cells [1]. Any irregularity within this organ can lead to threatening health problems. A brain tumour constitutes abnormal brain cells, which vary in size and type. According to the National Brain Tumour Society, approximately 18,020 people died from brain tumours in 2020 [2]. A brain tumour is an imbalanced proliferation of brain cells, which can be divided into major or minor. Major cases arise from within cells in the brain, whereas minor cases arise from cells other than brain cells [3]. The two types of tumours are glioma and meningioma, with glioma being the most common [4]. Meanwhile, the World Health Organization (WHO) [1] divides brain tumours into four types. Minor tumours, such as meningioma, are classified as grades 1 and 2. The remaining grades 3 and 4, such as glioma, are cancerous and lethal. The vast majority of pituitary tumours are benign. Large pituitary tumours are referred to as macroadenomas. In the clinical method, glioma, meningioma, and pituitary tumour prevalence rates are 45%, 15%, and 15%, respectively [5].

The treatments of brain tumours vary. Surgery is currently the most often used treatment for brain tumours since it has no adverse effect on the brain [6]. For imaging, magnetic resonance imaging (MRI) is often preferred over other techniques—such as computed tomography (CT), positron emission tomography (PET), and X-ray [7]. Manually analysing these images is time-consuming, labour-intensive, and prone to errors [8]. To overcome this limitation, an computer-aided diagnostic (CAD) system must be implemented to reduce the workload for brain MRI categorisation by physicians and specialists.

* Corresponding author

ISSN 2443-2555 (online) 2598-6333 (print)
doi: http://dx.doi.org/10.20473/jisebi.8.2.162-174

Methods for automatic brain tumour categorisation include inter- and intra-form, textural, and intensity alterations, but these remain ineffective. Traditional machine learning (ML) algorithms depend on handmade attributes, making them costly and non-sustainable. Deep learning methods can obtain valuable attributes and provide greater effectiveness but require massive labelled training data. A supervised learning model can outperform them but often produces an over-fitted model inappropriate for another extensive database. To address these challenges, the current study employs (1) a developed CNN model and tuned Inception V3 model for feature extrication for effective and discriminatory high-level features from brain MRI images and (2) multiple ML classifiers to distinguish normal and pathological brain MRI images.

We developed a feature descriptor method where deep features from the created CNN are computed using two ML classifiers. The best performing one was then chosen. The extensive evaluation is performed using the tuned pre-trained CNN model and the developed CNN model with two ML classifiers. The two datasets are (1) Fig share 3c dataset with three classes (glioma, meningioma, pituitary) and (2) Kaggle 4c dataset with four classes (glioma, meningioma, normal, pituitary). In conclusion, the followings are the contribution of this work:

1. The design and application of a completely automated hybrid method for brain tumour classification with (a) two types of CNN (the developed 23 layers CNN and tuned pre-trained Inception V3) model to extract profound features from brain MRI and (b) ML classifiers to efficiently distinguish the type of brain tumour.
2. A presentation of a three-step technique: (a) retrieving deep features by using two CNN models (the developed CNN and tuned Inception V3) for effective information retrieval and superior observation, and (b) grading of the optimum features employing tailored ML models, and (c) integrating them to develop a hybrid model for cutting-edge brain tumour classification using brain MRI.
3. The implementation of the method (CNN models and different ML classifiers) on two datasets (Fig share 3c and Kaggle 4c).
4. This method requires fewer pre-processing complexities.
5. The improvement of the overfitting issue of the model is analysed using the weight initialisation technique.

The rest of the paper is structured as follows: Section 2 lists the relevant paper works; Section 3 provides a brief explanation of the proposed methodology; Section 4 discusses the results of the proposed hybrid method; Section 5 presents the discussion; and Section 6 provides the conclusion.

II. LITERATURE REVIEW

This section reviews the deep and machine learning models for MRI brain tumour categorisation in past research. Table 1 depicts the strategies for computerised brain MRI classification associated with the enhanced traditional ML and deep learning methods.

The deep learning (DL) method has been widely employed for brain MRI categorisation in the last decade [9]. Because the attribute extrication and categorisation stages are built into conscience, the DL technique does not require handcrafted features. The DL technique involves a dataset, which may require further processing, and then prominent features are chosen [10]. A fundamental problem in MRI categorisation is minimising the gap between high-level sensory data detected by the human assessor and low-level sensory data acquired by the MRI machine. CNN can be utilised as an extractor to collect crucial attributes for the categorisation task to eliminate the semantic gap. In the preliminary and thicker layers, CNN models extricate primary and secondary elements, respectively. CNN extract profound features from brain MRI automatically. The primary convolutional block extracts feature from brain MRI such as edges and corners. These features are then combined in the next layer to produce high-level features. The extracted features from brain MRI are used to identify various image representation levels and capture relevant information such as image labels (class).

Researchers have employed CNNs to categorise brain MRI and evaluated their suggested methods on brain tumour datasets [11] [12][13]. Khawaldeh *et al.* [14] developed a CNN model for classifying the normalcy and abnormality of brain MRI to detect high-grade and low-grade glioma tumours. They developed their network upon the Alex-Net CNN architecture, achieving 91% accuracy. Afshar *et al.* [15] built a model for brain tumour categorisation by using capsule networks with 64 maps of MRI brain data, producing 86.56% accuracy. Charfi *et al.* [16] used histogram equalisation to segment images with PCA to reduce the dimension. Their accuracy was 90% based on the feed-forward neural network classifier. Despite the volume of work done throughout this area, a standard and robust solution to categorise brain MRI remains to be established.

Citak *et al.* [17] proposed three deep learning algorithms: multi-layer perceptron, logistic regression, and SVM. The SVM with linear kernel classification model acquired 93% accuracy on the MRI dataset. Past research has also used CNN models to classify brain lesions [18]. The architectures of these CNNs extract attributes from brain MRI through convolution and pooling processes. Pan *et al.* [19] suggested a CNN algorithm in their work to inspect MRI brain

tumour images, indicating that CNN sensitivity and specificity were 18% more than the ANN. Díaz *et al.* [20] developed a multi-pathway CNN framework for automatically segmenting glioma, meningioma, and pituitary brain tumours. They tested their suggested model against an openly accessible T1-weighted contrast-enhanced MRI dataset, gaining 97.3% accuracy. However, their learning environment was rather costly.

In conclusion, as seen in past research, the accuracy of categorising brain MRI using deep learning approaches is much higher than classical ML techniques. However, deep learning models require vast data for training to outperform classic ML techniques. Furthermore, the previously stated approaches have limitations that should be addressed before being applied to brain tumour classification: only evaluating binary classification MRI image datasets and overlooking multi-class datasets [21]. Kang *et al.* [5] developed a CNN-based model for classifying brain tumours using multi-class datasets. However, the accuracy is lower, and the computation time is not reasonable.

TABLE 1
WORK INVOLVED IN THE CLASSIFICATION OF BRAIN TUMOURS

Authors	Classification Type/Motive	Techniques Used	Dataset	Method of Classification	Method of Feature Extraction	Accuracy of the Model
[22]	Binary (normal and abnormal)	Classical Machine Learning-based Techniques	71 MRI	Feed-forward neural network	DWT	95.8%
[23]	Classification (Glioma, Meningioma, and Pituitary)	Deep learning-based techniques	3064 T1 weighted MRI	CNN	CNN	94.39%
[12]	Binary (normal and abnormal)		220 MRI	CNN	CNN	94.5%
[5]	Multi-classification		Three datasets (large dataset 3264 MRI)	ML classifiers	Pre-trained CNN networks	93.72% (large dataset)

III. METHODS

The framework of the proposed method is outlined in this section. Then, the details of the major components are explained in the subsequent sections. Fig. 1 depicts the workflow of the proposed methodology. The images are pre-processed (thresholding, cropping, resizing, and rescaling) before being fed into the CNN models. Two CNN models are used to extract deep features from brain MRI. The first is the newly created CNN model, and the second is the tuned Inception V3 model with the transfer learning approach. The ML classifiers then analyse the retrieved deep attributes.

This section includes the following major components: dataset collection, data pre-processing, CNN architecture, RF classifier, SVM, the tuned Inception V3 model, and training specifications of the proposed model.

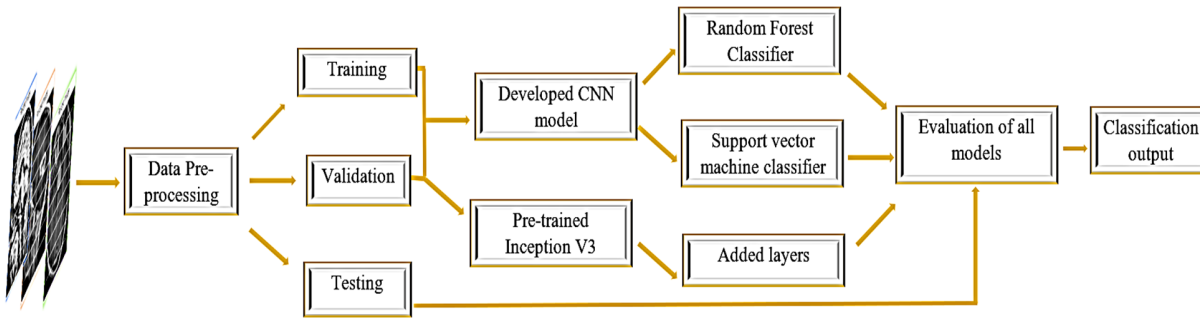

Fig. 1 The working flow of the proposed methodology

A. Dataset Collection

The image datasets are obtained from [21] and [24]. The first dataset, Fig share 3c, contains T1 weighted data from 233 patients, and the second dataset, Kaggle 4c, contains 3264 MRI images. There are 1426 gliomas, 708 meningiomas, and 930 pituitary class images in the Fig share 3c dataset. The number of gliomas, meningiomas,

normal, and pituitary class images in the Kaggle 4c dataset is 926, 937, 500, and 901, respectively. The image height and width in each dataset are 512 pixels. Table 2 displays the properties of the two image datasets.

TABLE 2
ATTRIBUTES OF THE DATASETS

Dataset Names	Number of Class	Training	Testing	Validation	Total
Fig share 3c	3	2343	460	261	3064
Kaggle 4c	4	2480	653	131	3264

B. Data Pre-processing

Minimal pre-processing is carried out to enhance the MRI images as input. Fig. 2 depicts the pre-processing steps, which include thresholding, cropping, resizing, and rescaling. An example is as follows:

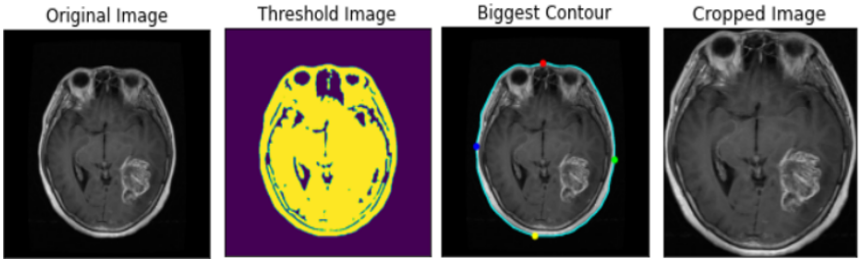

Fig. 2 Pre-processed images (a) Original MRI data; (b) Threshold image; (c) Contoured image; (d) Cropped image

1) Image Thresholding

To remove any noise in the original picture dataset, the threshold of the original image is first applied, followed by erosion and dilation.

2) Image Cropping

The cropping of the image is as follows:

- Step 1: The brain MRI images' threshold contours are mapped, and the largest contour is selected;
- Step 2: The contour comprising four edges' extreme points is plotted. The contour is essentially a NumPy array that contains the coordinates of the points on the map (x, y);
- Step 3: The extreme locations are used to crop the brain MRI.

3) Image Resizing

To reduce the computational cost of the original image size, the cropped image is resized to the preferred size of [224,224,3].

4) Image Rescaling

The images are rescaled for features ranging from 0 to 1, with a pixel having a maximum value of 255 and a minimum value of 0. Equation (1) is used for rescaling, with X denoting the MRI data.

$$X = \frac{X - X_{min}}{X_{max} - X_{min}} = \frac{X - 0}{255 - 0} = \frac{X}{255} \qquad (1)$$

C. CNN Model Architecture

Convolution using the resized MRI brain input image is achieved using a filter (3×3)that generates a feature map. This filter is used in conjunction with 'valid' padding and strides= (1,1). Fig. 3 depicts the suggested CNN design. The CNN model consists of five convolutional 2D layers, four max-pool 2D layers, one average pool 2D layer, five batch normalisation layers, four dropout layers, one flattened layer, and three dense layers. The total number of layers is 23.

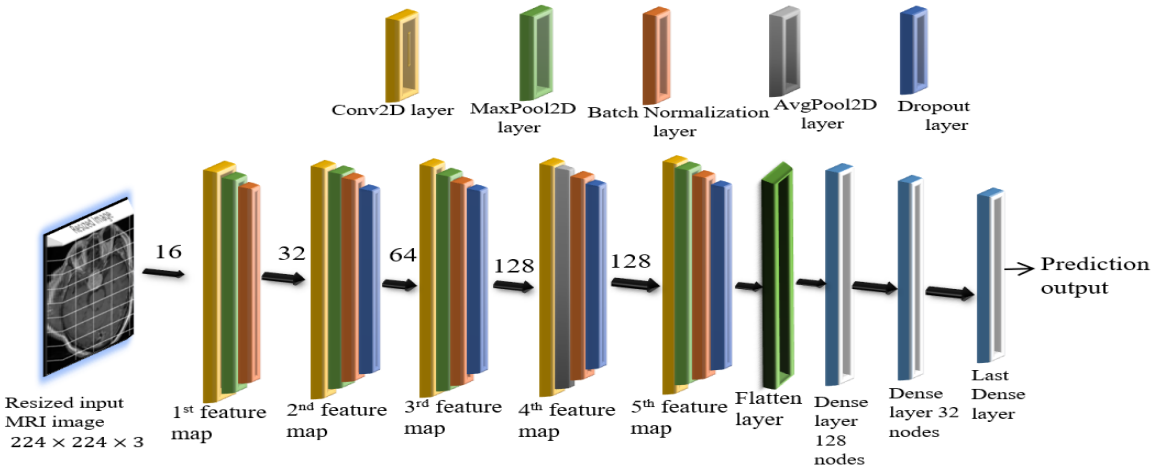

Fig. 3 Proposed CNN architecture model

TABLE 3
ALL THE LAYERS IN THE CNN MODEL WITH PARAMETERS AND ACTIVATION

Number of Layers	Names of layer	Activation shape	Parameters
1	Input Layer	224 × 224 × 3	0
2	1st Conv2D layer (3 × 3 filter)	222 × 222 × 16	448
3	MaxPool 2D (4 × 4 filter)	55 × 55 × 16	0
4	Batch Normalization	55 × 55 × 16	64
5	2nd Conv2D layer(3 × 3 filter)	53 × 53 × 32	4640
6	MaxPool 2D (2 × 2 filter)	26 × 26 × 32	0
7	Batch Normalization	26 × 26 × 32	128
8	Dropout (0.3)	26 × 26 × 32	0
9	3rd Conv2D layer(3 × 3 filter)	24 × 24 × 64	18,496
10	MaxPool 2D (2 × 2 filter)	12 × 12 × 64	0
11	Batch Normalization	12 × 12 × 64	256
12	Dropout (0.25)	12 × 12 × 64	0
13	4th Conv2D layer(3 × 3 filter)	10 × 10 × 128	73,856
14	AvgPool 2D (2 × 2 filter)	5 × 5 × 128	0
15	Batch Normalization	5 × 5 × 128	512
16	Dropout (0.25)	5 × 5 × 128	0
17	5th Conv2D layer(3 × 3 filter)	3 × 3 × 128	1,47,584
18	MaxPool 2D (2 × 2 filter)	1 × 1 × 128	0
19	Batch Normalization	1 × 1 × 128	512
20	Dropout (0.25)	1 × 1 × 128	0
21	Flatten layer	128	0
22	Dense1 (128) layer	128	16,512
23	Dense 2(32 neurons with regularise 0.0001)	32	4128
24	Dense 3 (class number× 1 neurons with regularise 0.0001)	4	132
TOTAL			2,67,268

The created CNN model comprises five convolutional blocks followed by classification layers. The classification layers consist of one flattened and three dense layers. Except for the first convolutional block, each contains one convolutional layer, one pooling layer, and one batch normalisation layer, followed by a dropout layer. There is no dropout layer in the first convolutional block. The initial Conv2D layer generates 16 feature maps of 222×222, which are subsequently fed into the max-pool 2D layer. Using the max-pool layer decreases the dimensionality to 55×55 while retaining the critical information. The batch normalisation layer's 16 feature maps are then sent into the convolutional layer of the second convolutional block, which yields 32 feature maps. Fig. 4 depicts the 32 feature maps produced by the second convolutional layer. This feature extraction procedure is carried on to the fifth convolutional block. More specific features are extracted as each block proceeds. The deep features are then flattened, followed by three dense layers. Table 3 displays the CNN model's layer parameters for each layer. At the last dense layer (Dense 3), the activation and parameters are shown for Kaggle 4c dataset.

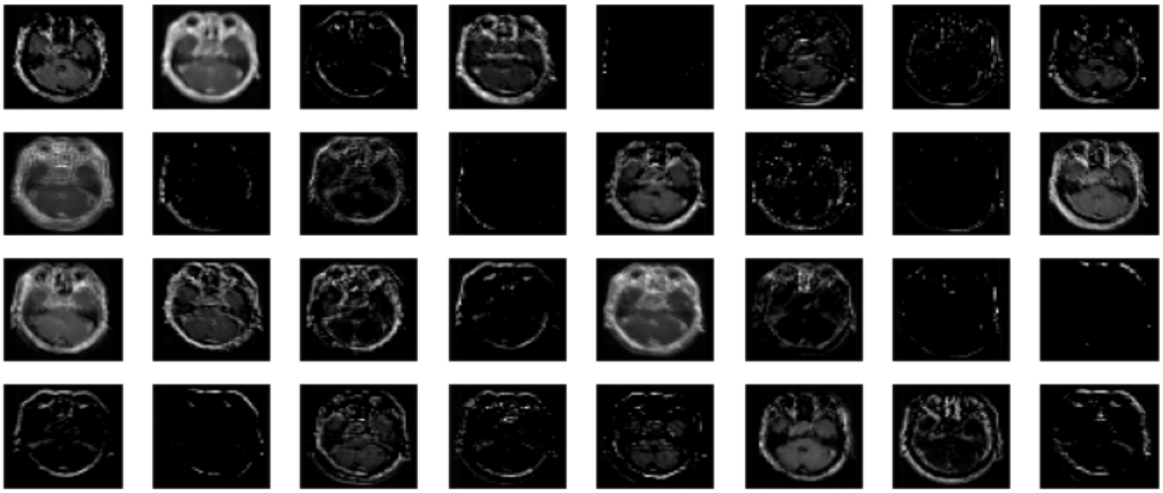

Fig. 4 Second feature maps of 32 filters

D. Random Forest Classifier

Random forest is a versatile supervised algorithm for classification. This method is composed of two random processes. The first is a random training set selection, and the second is induced by the tree-building process [25]. The following is the random forest classifier algorithm:

Step 1: The number of random estimators is selected.
Step 2: For every estimator, a decision tree is created.
Step 3: For each decision tree, a prediction result is created.
Step 4: Voting is carried out for each decision result.
Step 5: The forecasted outcome with the most votes is chosen as the final result.

In this established model, many estimators are set up, along with various random states. The final value (number of estimator=33, random state= 40) is chosen based on the highest accuracy.

E. Support Vector Machine Classifier

SVM is a learning system that uses regularisation to accomplish linear learning in non-linear domains [26]. The support vector machine technique aims to find a hyperplane in an N-dimensional space that recognises relationships between data points, where N is the number of features. The radial basis function achieves maximum accuracy among the various kernel functions. The SVM algorithm's two hyperparameters, degree (C) and gamma, are set to multiple values, and the final value (C=2, gamma=auto) is chosen based on the highest accuracy.

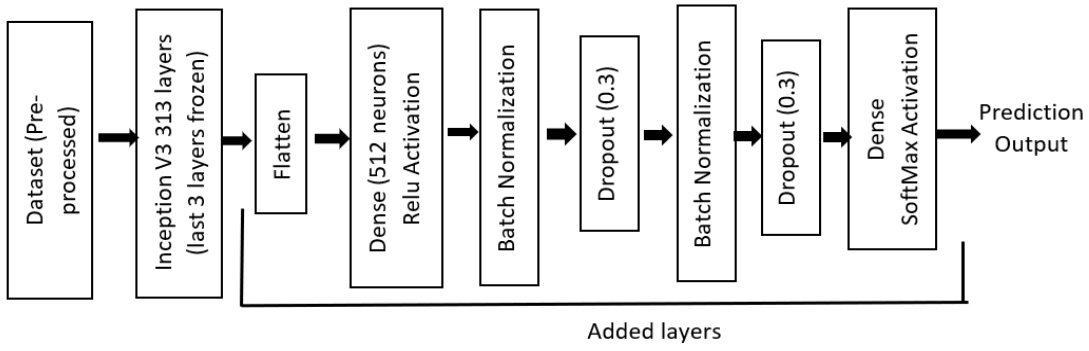

Fig. 5 Added layers in the inception V3 model

F. Tuned Inception V3

Transfer learning is utilised in this pre-trained Inception V3 model, where the last fully connected layer and the classification layer are frozen. Inception V3 is used among several pre-trained models because it improves performance with less computational weight than deep nets. The input image size is changed to a pre-processed input

image shape, i.e., [224,224,3]. Then seven layers are added, as shown in Fig. 5. Other combinations of extra layers are also used, chosen based on the model's performance.

G. Training Specifications of Proposed Model

The presented CNN model is trained using the 'Adam' optimizer as it combines the RMSprop and AdaGrad algorithms to maintain the sparse gradients. 'He-uniform' kernel weight initialisation is used in the proposed CNN model, which results in a non-overfitted model rather than a pre-trained 'ImageNet' InceptionV3 model. In 'he-uniform' weight initialisation, a uniform distribution is used with the limit $\sqrt{\frac{6}{number\ of\ input\ units\ in\ weight}}$. Early stopping criteria are employed with patience six and maximum mode. Table 4 depicts the training parameter configuration for the proposed approach.

TABLE 4
TRAINING PARAMETER CONFIGURATION

Parameter	Value/Types
Input size	$224 \times 224 \times 3$
Epochs	60
Optimiser	'Adam'
Learning rate	0.001
Loss function	Sparse Categorical Cross entropy
Weight initialiser	He-uniform
Kernel regularise	0.0001

IV. RESULTS

The experimental results are acquired from the two datasets (Fig share 3c and Kaggle 4c) to categorise brain tumours. This study uses a pre-trained Inception V3 model in addition to the custom-developed CNN model. On the ImageNet [27] dataset, the weight of the bottleneck layers of the pre-trained Inception V3 model is frozen. Two machine learning classifiers are utilised, i.e., random forest and support vector machine (RBF kernel). The proposed approach is implemented in python, and the GPU is used for training.

A. Performance Evaluation

Various matrices are used to evaluate the performance. The confusion matrix displays the correct and incorrect estimation of classes in a tabular style. For the Fig share 3c dataset, Table 5 illustrates the confusion matrix of the CNN-RF model. Table 6 depicts the confusion matrix of the CNN-SVM model used in this experiment on the Kaggle 4c dataset.

TABLE 5
CONFUSION MATRIX OF CNN-RF MODEL ON FIG SHARE 3C DATASET

		Predicted		
		Glioma	Meningioma	Pituitary
Actual	Glioma	204	9	1
	Meningioma	2	102	2
	Pituitary	1	1	138

TABLE 6
CONFUSION MATRIX OF CNN-SVM ON KAGGLE 4C DATASET

	Predicted			
Actual	Glioma	Meningioma	Normal	Pituitary
Glioma	176	9	0	0
Meningioma	3	176	2	7
Normal	3	1	94	2
Pituitary	0	3	0	177

Other performance metrics can be constructed from the confusion matrix to describe the model's performance. Precision, recall, specificity, and F1 score are essential metrics that can be determined using the equations below:

$$Precision = \frac{TP}{T\,P + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$Specificity = \frac{TN}{T\,N + FP} \quad (4)$$

$$F1 - score = \frac{2\times Precision \times Recall}{Precision + Recall} \quad (5)$$

$$Youden\ index = Recall + specificty - 1 \quad (6)$$

where, TP, FP, TN, and FN represent the number of classified cases of true positives, false positives, true negatives, and false negatives, respectively.

TABLE 7
EVALUATION OF THE MODELS ON THE FIG SHARE 3C DATASET

Performance	Classes	Tuned Inception V3 model	CNN model	CNN-RF model	CNN-SVM model
Precision	Glioma	0.88	0.99	0.99	0.97
	Meningioma	0.86	0.85	0.91	0.91
	Pituitary	0.87	0.99	0.98	0.97
Recall	Glioma	0.93	0.93	0.95	0.97
	Meningioma	0.63	0.99	0.96	0.91
	Pituitary	0.98	0.97	0.99	0.99
F1-score	Glioma	0.90	0.96	0.97	0.97
	Meningioma	0.73	0.91	0.94	0.91
	Pituitary	0.92	0.98	0.98	0.98
Accuracy (%)		87%	95.43%	**96.52%**	95.8%

TABLE 8
EVALUATION OF THE MODELS ON THE KAGGLE 4C DATASET

Performance	Classes	Tuned Inception V3 model	CNN model	CNN-RF model	CNN-SVM model
Precision	Glioma	0.87	0.95	0.96	0.97
	Meningioma	0.87	0.94	0.90	0.93
	Normal	0.92	0.98	0.98	0.98
	Pituitary	0.89	0.87	0.92	0.95
Recall	Glioma	0.86	0.95	0.92	0.95
	Meningioma	0.84	0.86	0.94	0.94
	Normal	0.87	0.89	0.86	0.94
	Pituitary	0.96	0.99	0.97	0.98
F1-score	Glioma	0.87	0.95	0.94	0.96
	Meningioma	0.85	0.89	0.92	0.93
	Normal	0.89	0.93	0.91	0.96
	Pituitary	0.92	0.93	0.95	0.97
Accuracy (%)		88.21%	93%	93.11%	**95.41%**

Tables 7 and 8 illustrate the class-wise performance matrices for Fig share 3c dataset and Kaggle 4c datasets, respectively. The bold text in the table represents the highest accuracy obtained among all the models. Based on the experimental results of all the models, the highest accuracy model is chosen from both datasets. The weighted performance measure of the best model on each dataset is presented in Table 9.

TABLE 9
PERFORMANCE OF PROPOSED METHOD ON BOTH DATASETS

Dataset	Proposed method	Weighted F1 score (%)	Weighted Specificity (%)	Accuracy (%)	Youden Index
Fig share 3c	**CNN-RF**	**96.6**	98.1	96.52	94.2
Kaggle 4c	**CNN-SVM**	95.4	97.8	95.41	93.2

Table 9 depicts that CNN-RF outstands other models in Fig share 3c dataset, and CNN-SVM outstands other models in the Kaggle 4c dataset. The following observations are drawn from the results:

1) Observation 1:

Table 5 shows that the meningioma class has the highest number of misclassified images in the Fig share 3c dataset, resulting in a lower recall, precision, and F1 score for the meningioma class in Table 7.

Analysis: The meningioma class has the highest number of misclassified photos. This is due to the fewer samples of this class in the dataset and the lack of class-specific data augmentation used to balance the dataset.

2) Observation 2:

SVM with RBF kernel outperforms other classifiers in larger datasets (Kaggle 4c).

Analysis: Table 8 shows that the CNN-SVM model with RBF kernel performs better on the Kaggle 4c dataset, which is larger than the Fig share 3c dataset. This is because the number of characteristics in each data point can easily cope with the number of training samples.

3) Observation 3:

On both datasets, the deep features derived from the CNN model outperform the tuned Inception V3 model, which is based on transfer learning.

Analysis: This is because the deep features extracted from the CNN model can assure smooth decision boundaries on the testing samples.

4) Observation 4:

From Table 9, it is observed that the weighted specificity value is high in each dataset.

Analysis: High specificity indicates that samples with no brain tumour are more accurately identified.

B. Analysis of CNN model

Analysis of the developed CNN model is executed to learn which layers of the CNN model contribute the most to overall model performance. Table 10 summarises the results of the trial (Kaggle 4c dataset). The greatest acquired accuracy model's time complexity is also noted. The following can be deduced from Table 10:

TABLE 10
ANALYSIS OF CNN MODEL LAYERS (KAGGLE 4C DATASET)

CNN Layers	CNN	CNN-RF	CNN-SVM	Time Complexity(sec)
Without 4th convo block	91.12	86.37	92.64*	304*
Without 6,10,19,20 no layer	87.44	86.06	91.42*	426*
Without 3rd convo block	92.0	88.51	**93.26***	**290***
Without 5thconvo block	89.58	87.44	93.72*	296*

1) Observation 1:

When some layers (such as max-pooling) from the second and fifth convolutional blocks are changed, all models underperform. This indicates that these layers are more influential.

Analysis: The max-pooling layer has no learnable parameters, but it decreases the feature dimension while preserving important information. As a result, the temporal complexity rises with the loss of accuracy.

2) Observation 2:

The model performs better than other criteria without the fifth convolutional block. As a result, it may be concluded that this block of layers is less influential.

Analysis: The model performance does not degrade much because the block is changed rather than just a few layers. The time complexity is also lower than in the first observation.

C. Analysis of Random Forest Classifier

The random forest classifier is used to solve the SoftMax activation function's failure to correctly detect the different viewpoints of multi-class images. This classifier can process both linear and non-linear data. Table 11 depicts a study of the RF classifier depending on the number of trees (Fig share 3c dataset). Table 11 shows that the model's performance degrades when the number of trees increases beyond 33. The number of trees chosen is significantly dependent on the number of characteristics. The maximum number of features is chosen as the square root of this experiment's total number of characteristics. Based on this dataset (Fig share 3c), the maximum number of characteristics can be accommodated with 33 trees. As a result, as the number of trees increases, the model's performance degrades.

TABLE 11
ANALYSIS OF RF CLASSIFIER (FIG SHARE 3C)

Number of Trees	Number of random States	Test Accuracy	Number of Incorrectly predicted classes
32	42	0.9609	18
33	**40**	**0.9652**	**16**
45	40	0.9522	22
50	42	0.9478	24

D. Effects of Different Weight Techniques on Overfitting

The developed models are trained using various kernel weight initialisation strategies, i.e., the he-normal, glorot-uniform, and he-uniform. Fig. 6 shows the effects of the weight techniques. Overfitting occurs when the data fits perfectly in the training set but performs poorly in the testing set. Fig. 6 shows that he-uniform minimises test loss and drives accuracy higher than the other two techniques, reducing the risk of overfitting. The weights are uniformly initialised in he-uniform. The CNN model achieves 95.43% accuracy, increasing to 96.52% when the RF classifier is employed.

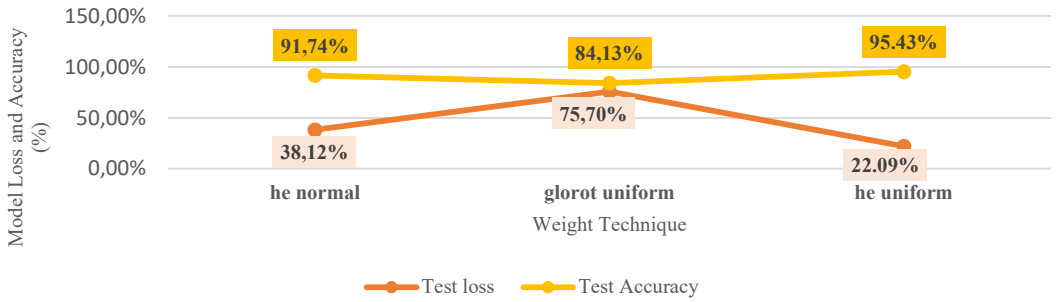

Fig. 6 Impact of various weight techniques on the CNN-RF model (fig share 3C)

V. DISCUSSION

The novelty of this method lies in the enhancement of the CNN model and the correct hyperparameter selection. This study is unique because it employs CNN, CNN-RF, CNN-SVM, and tuned Inception V3 deep learning models on multi-class brain MRI datasets. Two large multi-class datasets are well suited for the developed CNN architecture, indicating the model's effectiveness. The enhanced models are evaluated and compared based on the accuracy and weighted F1 score (for uneven distribution of classes in the datasets). In the Fig share 3c dataset, the CNN-RF model shows the highest F1 score in Table 13 compared to the existing methods. From the majority voting in the prediction of each decision tree, the CNN-RF model is the most accurate (Fig share 3c). Using the he-uniform weight initialisation technique, l1 regularisation, and hidden neuron dropout in the CNN architecture solves the overfitting problem. Among the four deep learning models, the CNN-RF model is the best for the Fig share 3c dataset and the CNN-SVM model for the Kaggle 4c dataset. Table 12 compares the computational complexity of the proposed methods to the existing ones. The pre-processing operations are less complex than manual segmentation [30]. Compared to the method in [5], the number of parameters is low, but the accuracy is high.

TABLE 12
COMPARISON OF THE PROPOSED MODELS' COMPUTATIONAL COMPLEXITY WITH EXISTING MODELS

References	Pre-processing	Layer	Feature Extraction	Kernel	Parameter	Classifier	Accuracy
[30]	Manual segmentation	-	2D DWT, 2D Gabor filter	-	-	BPNN	91.9%
[5]	Thresholding+ Cropping+ Resizing rescaling+ augmentation	(82 + 50+ 7 convo block)	(DenseNet-169 + Shufflenet + MnasNet)	-	20M +299M +3.9M	SVM	93.72%
[15]	-	6	Caps Net	32, 64	-	FC	86.56%
Proposed method	Thresholding+ Cropping+resizing+ rescaling	23	CNN	16, 32, 64 128, 128	**2,67,268**	RF SVM	**96.52%** **95.43%**

TABLE 13
COMPARISON USING CLASS-SPECIFIC MATRICES (FIG SHARE 3C) WHERE G, M, P MEAN GLIOMA, MENINGIOMA, AND PITUITARY, RESPECTIVELY

References	Class	Precision	Recall	Specificity	Average F1 score
[28]	G	-	96.4	96.3	-
	M	-	86.0	95.5	
	P	-	87.3	95.3	
[29]	G	91.0	97.5	-	0.93
	M	94.5	76.8	-	
	P	98.3	100	-	
[30]	G	-	95.1	96.3	-
	M	-	86.9	96.0	
	P	-	91.2	95.7	
Proposed CNN-RF	G	99.0	95.0	98.8	**0.96**
	M	91.0	96.0	97.2	
	P	98.0	99.0	99.0	

TABLE 14
COMPARISON USING THE SAME FIG SHARE 3C DATASET

References	Method	Accuracy (%)
[13]	CNN	84.19
[15]	Caps Net	86.56
[28]	BoW-SVM	91.28
[9]	CNN	91.43
[23]	CNN	94.39
Proposed Method	Tuned Inception V3	87.0
Proposed Method	CNN	95.43
Proposed Method	CNN-SVM	95.8
Proposed Method	**CNN-RF**	**96.52**

TABLE 15
COMPARISON USING THE SAME KAGGLE 4C DATASET

References	Method	Data division	Accuracy (%)
[5]	Pre-trained features ensemble-ML	80% in training and 20% in testing	93.72
Proposed Method	Tuned Inception V3		88.21
Proposed Method	CNN		93
Proposed Method	CNN-RF	76% in training, 4% in validation,	93.11
Proposed Method	**CNN-SVM**	and 20% in testing	**95.41**

Regarding the specific three-class and four-class classification problems of brain tumours, the performance of the proposed method is compared to others. The suggested CNN-RF approach outperforms state-of-the-art methods across the board. A more detailed comparison is provided in Table 13. Based on the specificity matrix, the suggested method outperforms those in previous studies [28] [30]. The F1 score is also higher than the method in [29]. Table 14 shows that this strategy outperforms all current methods. Table 15 shows that the application on the Kaggle 4c dataset performs better than the method by [5] despite using an equal amount of data in the testing set.

Nonetheless, this research is limited in terms of applicability to other datasets. Compared to the CNN architecture's uniqueness and simplicity, this method may not perform well on different multi-class brain MRI datasets.

VI. CONCLUSIONS

This work proposes an efficient hybrid deep learning-based approach to classify the multi-class brain MRI. The three steps to construct the model are pre-processing, attribute extrication from CNN architecture, and classification of the recovered features using the RF and SVM classifiers. Image thresholding, cropping, resizing, and rescaling are performed in the pre-processing. To extricate deep attributes from brain MRI, this study utilised two types of deep CNN, a new CNN architecture and a tuned pre-trained Inception V3 model with the transfer learning approach. The classifiers then analyse the extricated deep attributes. This work performs detailed evaluations on two datasets (Fig sharing 3c and Kaggle 4c) to categorise brain MRI using two deep CNNs and two different ML (RF, SVM) classifiers. The findings show that the constructed CNN model outperforms the pre-trained Inception V3 model. The CNN model has a smooth boundary deception on testing samples. The CNN network is also simpler and faster to execute than pre-existing pre-trained networks. Deep networks, such as Inception V3, require dedicated hardware for real-time performance. The CNN-RF model (96.52%) outperforms the other four models in the Fig share 3c dataset. Meanwhile, the CNN-SVM model (95.41%) outperforms the large Kaggle 4c dataset with four classes (glioma, meningioma, normal, and pituitary). Despite the efforts described in this work, some enhancements are still possible. For example,

regarding the misclassification of meningioma classes in Fig share 3c dataset, further research can be carried out for class-specific data augmentation.

Author Contributions: *Hafiza Akter Munira*: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing. *Md. Saiful Islam*: Software, Investigation, Data Curation, Writing - Review & Editing, Supervision.

Funding: This research received no specific grant from any funding agency.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

[1] D. N. Louis *et al.*, "The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary," *Acta Neuropathologica*, vol. 131, no. 6. 2016. doi: 10.1007/s00401-016-1545-1.

[2] C. Rouse, H. Gittleman, Q. T. Ostrom, C. Kruchko, and J. S. Barnholtz-Sloan, "Years of potential life lost for brain and CNS tumors relative to other cancers in adults in the United States, 2010," *Neuro. Oncol.*, vol. 18, no. 1, 2016, doi: 10.1093/neuonc/nov249.

[3] G. S. Tandel *et al.*, "A review on a deep learning perspective in brain cancer classification," *Cancers*, vol. 11, no. 1. 2019. doi: 10.3390/cancers11010111.

[4] J. Liu *et al.*, "Applications of deep learning to MRI Images: A survey," *Big Data Mining and Analytics*, vol. 1, no. 1. 2018. doi: 10.26599/BDMA.2018.9020001.

[5] J. Kang, Z. Ullah, and J. Gwak, "Mri-based brain tumor classification using ensemble of deep features and machine learning classifiers," *Sensors*, vol. 21, no. 6, 2021, doi: 10.3390/s21062222.

[6] R. Mehrotra, M. A. Ansari, R. Agrawal, and R. S. Anand, "A Transfer Learning approach for AI-based classification of brain tumors," *Mach. Learn. with Appl.*, vol. 2, 2020, doi: 10.1016/j.mlwa.2020.100003.

[7] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, 2016, doi: 10.1109/TMI.2016.2538465.

[8] K. Popuri, D. Cobzas, A. Murtha, and M. Jägersand, "3D variational brain tumor segmentation using Dirichlet priors on a clustered feature set," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 7, no. 4, 2012, doi: 10.1007/s11548-011-0649-2.

[9] J. S. Paul, A. J. Plassard, B. A. Landman, and D. Fabbri, "Deep learning for brain tumor classification," in *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 2017, vol. 10137. doi: 10.1117/12.2254195.

[10] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, and T. R. Mengko, "Brain tumor classification using convolutional neural network," in *IFMBE Proceedings*, 2019, vol. 68, no. 1. doi: 10.1007/978-981-10-9035-6_33.

[11] N. M. Balasooriya and R. D. Nawarathna, "A sophisticated convolutional neural network model for brain tumor classification," in *2017 IEEE International Conference on Industrial and Information Systems, ICIIS 2017 - Proceedings*, 2018, vol. 2018-Janua. doi: 10.1109/ICIINFS.2017.8300364.

[12] D. Jude Hemanth, J. Anitha, A. Naaji, O. Geman, D. E. Popescu, and L. Hoang Son, "A Modified Deep Convolutional Neural Network for Abnormal Brain Image Classification," *IEEE Access*, vol. 7, 2019, doi: 10.1109/ACCESS.2018.2885639.

[13] J. Seetha and S. S. Raja, "Brain tumor classification using Convolutional Neural Networks," *Biomed. Pharmacol. J.*, vol. 11, no. 3, 2018, doi: 10.13005/bpj/1511.

[14] S. Khawaldeh, U. Pervaiz, A. Rafiq, and R. S. Alkhawaldeh, "Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks," *Appl. Sci.*, vol. 8, no. 1, 2017, doi: 10.3390/app8010027.

[15] P. Afshar, A. Mohammadi, and K. N. Plataniotis, "Brain Tumor Type Classification via Capsule Networks," *Proc. - Int. Conf. Image Process. ICIP*, pp. 3129–3133, 2018, doi: 10.1109/ICIP.2018.8451379.

[16] S. Charfi, R. Lahmyed, and L. Rangarajan, "Impact Factor(JCC): 1.5548-A NOVEL APPROACH FOR BRAIN TUMOR DETECTION USING NEURAL NETWORK," *Citeseer*, vol. 2, 2014.

[17] F. Citak-Er, Z. Firat, I. Kovanlikaya, U. Ture, and E. Ozturk-Isik, "Machine-learning in grading of gliomas based on multi-parametric magnetic resonance imaging at 3T," *Comput. Biol. Med.*, vol. 99, 2018, doi: 10.1016/j.compbiomed.2018.06.009.

[18] X. Yang and Y. Fan, "Feature extraction using convolutional neural networks for multi-atlas based image segmentation," 2018. doi: 10.1117/12.2293876.

[19] Y. Pan *et al.*, "Brain tumor grading based on Neural Networks and Convolutional Neural Networks," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2015, vol. 2015-November. doi: 10.1109/EMBC.2015.7318458.

[20] F. J. Díaz-Pernas, M. Martínez-Zarzuela, D. González-Ortega, and M. Antón-Rodríguez, "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network," *Healthc.*, vol. 9, no. 2, 2021, doi: 10.3390/healthcare9020153.

[21] S. Bhuvaji, S., Ankita Kadam, A., Bhumkar, P., Dedge, S., Swati Kanchan. S., & kanchan, "Brain Tumor Classification (MRI)," 2020.

[22] Z. Ullah, M. U. Farooq, S. H. Lee, and D. An, "A hybrid image enhancement based brain MRI images classification technique," *Med. Hypotheses*, vol. 143, 2020, doi: 10.1016/j.mehy.2020.109922.

[23] S. Das, O. F. M. R. R. Aranya, and N. N. Labiba, "Brain Tumor Classification Using Convolutional Neural Network," *1st Int. Conf. Adv. Sci. Eng. Robot. Technol. 2019, ICASERT 2019*, no. May 2019, pp. 1–6, 2019, doi: 10.1109/ICASERT.2019.8934603.

[24] J. Cheng, "brain tumor dataset. figshare.," 2017. https://doi.org/10.6084/m9.figshare.1512427.v5 (accessed Jun. 21, 2021).

[25] F. Rotaru, S.-I. Bejinariu, M. Luca, R. Luca, and C. D. Niță, "Brain Tumor Segmentation Based on Random Forest," pp. 1–11, 2016, [Online]. Available: http://mss.academiaromana-is.ro/mem_sc_st_2016/9_Rotaru.pdf

[26] H. F. Chen, "In silico log p prediction for a large data set with support vector machines, radial basis neural networks and multiple linear regression," *Chem. Biol. Drug Des.*, vol. 74, no. 2, pp. 142–147, 2009, doi: 10.1111/j.1747-0285.2009.00840.x.

[27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, 2017, doi: 10.1145/3065386.

[28] J. Cheng *et al.*, "Erratum: Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition (PLoS ONE 10:12 (e0144479))," *PLoS ONE*, vol. 10, no. 12. 2015. doi: 10.1371/journal.pone.0144479.

[29] A. Pashaei, H. Sajedi, and N. Jazayeri, "Brain tumor classification via convolutional neural network and extreme learning machines," *2018 8th Int. Conf. Comput. Knowl. Eng. ICCKE 2018*, no. Iccke, pp. 314–319, 2018, doi: 10.1109/ICCKE.2018.8566571.

[30] M. R. Ismael and I. Abdel-Qader, "Brain Tumor Classification via Statistical Features and Back-Propagation Neural Network," *IEEE Int. Conf. Electro Inf. Technol.*, vol. 2018-May, pp. 252–257, 2018, doi: 10.1109/EIT.2018.8500308.