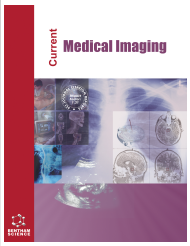
BENTHAM SCIENCE

CrossMark

Current Medical Imaging

Content list available at: https://benthamscience.com/journals/cmir



RESEARCH ARTICLE

Automated Brain Tumour Detection and Classification using Deep Features and Bayesian Optimised Classifiers

S.Arun Kumar[1,*] and S. Sasikala[1]

[1]*Department of Electronics and Communication Engineering, Kumaraguru College of Technology, Tamil Nadu 641049, India*

Abstract:

Purpose:

Brain tumour detection and classification require trained radiologists for efficient diagnosis. The proposed work aims to build a Computer Aided Diagnosis (CAD) tool to automate brain tumour detection using Machine Learning (ML) and Deep Learning (DL) techniques.

Materials and Methods:

Magnetic Resonance Image (MRI) collected from the publicly available Kaggle dataset is used for brain tumour detection and classification. Deep features extracted from the global pooling layer of Pretrained Resnet18 network are classified using 3 different ML Classifiers, such as Support vector Machine (SVM), K-Nearest Neighbour (KNN), and Decision Tree (DT). The above classifiers are further hyperparameter optimised using Bayesian Algorithm (BA) to enhance the performance. Fusion of features extracted from shallow and deep layers of the pretrained Resnet18 network followed by BA-optimised ML classifiers is further used to enhance the detection and classification performance. The confusion matrix derived from the classifier model is used to evaluate the system's performance. Evaluation metrics, such as accuracy, sensitivity, specificity, precision, F1 score, Balance Classification Rate (BCR), Mathews Correlation Coefficient (MCC) and Kappa Coefficient (Kp), are calculated.

Results:

Maximum accuracy, sensitivity, specificity, precision, F1 score, BCR, MCC, and Kp of 99.11%, 98.99%, 99.22%, 99.09%, 99.09%, 99.10%, 98.21%, 98.21%, respectively, were obtained for detection using fusion of shallow and deep features of Resnet18 pretrained network classified by BA optimized SVM classifier. Feature fusion performs better for classification task with accuracy, sensitivity, specificity, precision, F1 score, BCR, MCC and Kp of 97.31%, 97.30%, 98.65%, 97.37%, 97.34%, 97.97%, 95.99%, 93.95%, respectively.

Conclusion:

The proposed brain tumour detection and classification framework using deep feature extraction from Resnet 18 pretrained network in conjunction with feature fusion and optimised ML classifiers can improve the system performance. Henceforth, the proposed work can be used as an assistive tool to aid the radiologist in automated brain tumour analysis and treatment.

Keywords: Brain tumour, Resnet 18, Detection, Classification, Deep learning, Transfer learning, Hyper parameter tuning.

Article History	Received: November 02, 2022	Revised: January 06, 2023	Accepted: January 25, 2023

1. INTRODUCTION

Brain tumour is the uncontrolled growth of cells in the brain. Brain tumours are classified as benign/cancerous or malignant/non-cancerous. Tumours can originate in the brain (primary tumour) or spread from the affected parts of the body to the brain (secondary tumour). Brain tumours are characterized by headaches, changes in vision and hearing, seizures, confusion, difficulty in making decisions, and behaviour changes.

Brain tumours are classified based on the cell in which tumours develop, and the severity of the tumour determines the damage to the nervous system. Usually, brain tumours originate from glial cells and are termed gliomas. Gliomal tumours affect the brain, spinal cord, and peripheral nervous system. Glioblastoma, a type of glioma, is a malignant brain

* Address correspondence to this author at the Department of Electronics and Communication Engineering, Kumaraguru College of Technology, Tamil Nadu 641049, India;
E-mail:arunkumar.s.ece@kct.ac.in

tumour, accounting for 1 out of 5 cancer deaths. Meningioma tumour develops in the layers covering the brain and spinal cord (meninges), accounting for 20% of the tumours. Pituitary tumours originate in the pituitary gland located near the brain. Pituitary tumours are benign or cancerous (adenomas) [1].

As per GLOBOCAN 2020 statistics [2], 3,08,102 new brain cancer cases and 2,51,329 deaths were reported. Brain tumour is the major cause of cancer deaths in people below 40 years. Brain tumours are most common in older adults but can also occur at a young age. Brain tumour is the second leading cancer type in children, demanding novel research and early detection techniques for rapid diagnosis.

Manual classification of brain tumour images is a difficult task for radiologists, as there is a high resemblance in the structure of MR images. Accurate detection and classification of brain tumours demands radiologist expertise and is time-consuming. On the other hand, Computer Aided Diagnosis (CAD) systems can automate brain tumour diagnosis and aid the radiologist in making accurate decisions [3].

Brain tumour is the deadliest disease that is quite common among adults and young children. Timely diagnosis of brain tumours can improve treatment options and save the lives of affected patients. Early detection of brain tumours using CAD systems can accelerate treatment and reduce mortality [4]. In recent times, advancements in medical imaging techniques and the use of Artificial Intelligence (AI) techniques have improved the CAD-based diagnosis of brain tumours. MRI-based approaches have proved to be an efficient imaging modality for brain tumour analysis. Manual MRI analysis for disease detection is time-consuming, demands subject knowledge and is error-prone. The use of AI techniques such as ML and DL to automate disease detection proved to be an effective choice in brain tumour detection. ML techniques are feature driven. Deep learning has been used in the recent past to boost the diagnostic performance of classification and detection tasks in problem domains of biomedical engineering. DL techniques offer improved performance because of their ability to extract deep features for efficient detection and classification [5]. Therefore, the proposed CAD system uses ML and DL techniques for the identification and grading of the types of brain tumours from brain Magnetic Resonance Images (MRI).

ML algorithm uses handcrafted features to detect/classify the tumour. Furthermore, feature extraction is the key step in ML which demands expertise based on the problem domain [6]. However, deep learning algorithms learn feature representation from the images directly. Self-learning ability of DL makes it easier for non-experts to use it as a tool for aiding tumour diagnosis. Convolutional Neural Network (CNN) and other DL algorithms surpass conventional ML algorithms in terms of accuracy and other performance metrics in disease diagnosis [7].

Despite the improved performance, the training of deep CNN needs huge training data. To overcome the above limitation, rather than training DL networks from scratch, the use of Transfer Learning (TL) approach on a small, labelled dataset has resulted in significant performance improvement [8]. TL is a ML research paradigm that transfers knowledge learned from an earlier task to a new task. In transfer learning, features learned from pretrained models trained on a large dataset are transferred to a network trained on a target dataset. Thus, TL overcomes the challenges in conventional DL approaches. Pretrained deep CNN networks can be used as a feature extractor or as a base learner for TL. Deep CNN networks learn hierarchical feature representation from image data. Initial layers generate shallow features, and the final layers generate deep features. The use of deep features extracted from the global pooling layer and convolutional layers of a pretrained network followed by machine learning classifiers has been experimented to improve the performance of biomedical disease detection systems [9]. Thus, image classification using DL algorithms consumes more time and results in improved performance only for large dataset. In biomedical disease detection problems such as brain tumour, data is scarce, and TL can be employed to improve the detection capability of the classifier. The use of TL as a classifier consumes more training time and needs high computational resources. To leverage the advantage of TL, deep features extracted from pretrained networks are classified using ML classifiers. ML classifiers perform well in smaller datasets and consume less training time and reduced computational time. The use of deep features classified by ML classifiers results in better accuracy vs. training time trade-off [10].

Therefore, the proposed system experiments with three different approaches for enhancing brain tumour detection and classification. A baseline CNN, Pretrained Resnet18, Pretrained Resnet 18 features extracted from the pooling layer and fusion of shallow and deep layer features in Resnet 18 followed by ML classifiers and Bayesian optimized ML classifiers are proposed in this study. The work is evaluated on a publicly available brain tumour detection and brain tumour classification dataset.

The major contributions of the proposed work are:

- To propose a multilayer CNN and deep feature-based CAD framework for the detection and classification of brain tumour from 2D MRI images.
- Deep features extracted from the global pooling layer of Resnet-18 pretrained are classified by SVM, KNN, and DT classifier. Furthermore, the hyperparameters of the abovesaid classifiers are optimised using a Bayesian algorithm for improved detection and classification performance.
- Feature fusion of shallow and deep layers of Resnet-18 pretrained network is used for feature extraction followed by classification using aforesaid ML classifiers and Bayesian optimised ML classifiers for enhanced performance.
- The proposed research is among the first to employ feature fusion of shallow and deep layers of Resnet-18 pretrained network in conjunction with Bayesian optimised ML classifiers to improve the performance of the proposed CAD system in comparison with existing literature.

2. RELATED STUDIES

A brief literature survey on the use of DL approaches for brain tumour detection and classification is presented in the following section.

A DL framework extracts automatic deep features and performs classification without the need for handcrafted features [11]. Furthermore, fine-tuning of CNN using VGG 19 TL architecture is experimented to boost the accuracy of the brain tumour classification system. An accuracy of 94.82% is achieved on CE- MRI dataset using 5-fold cross-validation. A deep CNN with Long Short Term memory (LSTM) is experimented for enhancing the performance of brain tumour classification system [12]. LSTM layer added before the convolutional layer resulted in an improved accuracy of 92% in Kaggle brain tumour classification dataset.

Feature fusion of 4 different DL architectures, CNN, LSTM, Gated Recurrent Unit (GRU), and Recurrent Neural Network (RNN) classified by fuzzy minmax meta-classifier resulted in an accuracy of 92.86% on a brain tumour detection dataset [13]. Deep features extracted from 315-layer Inceptionv3 TL model and score values obtained from the softmax layer were used to train a 6-layer light weight quantum network to classify the tumour types. Furthermore, semantic segmentation was done to segment the tumour volume. An average accuracy of 96% and 90.91 was obtained on Kaggle dataset and BRATS 2020 dataset, respectively [14].

MRI image is preprocessed using median filtering and contrast enhancement techniques. The preprocessed image is converted to a wireframe model using the cellular logic array processing (CLAP) model. 3D patterns of MRI images are determined, and tumour classification is achieved using the 3D Alexnet architecture. Improved accuracy is obtained using the proposed approach on three different brain tumour datasets [15]. Pretrained networks such as Alexnet, Googlenet, and Resnet18 are used as feature extractors. Feature extraction is followed by fusion of three networks and the fused feature is classified using SVM and KNN classifier. An accuracy of above 97% is achieved on 3 publicly available datasets [16].

A transfer learning framework using Alexnet CNN, and Resnet 18 CNN was experimented on Kaggle MRI dataset. Furthermore, the deep features extracted from the above pretrained networks were classified using SVM classifier. An accuracy of 95.1% was reported using Alexnet features classified by SVM [17]. A customised CNN with augmented Kaggle image dataset was used to improve the brain tumour CAD. The custom CNN outperforms the existing architecture in terms of number of parameters and accuracy. An accuracy of 88% was achieved on Kaggle dataset [18]. Image pre-processing was performed to enhance the image by noise removal. Pre-processing was followed by morphological segmentation. A custom CNN with 10 layers was used to classify the segmented image [19].

Various parameter optimisation techniques are applied to diagnose brain tumours using CNN model. Optimisation of learning rate, training size, number of epochs, in CNN are experimented with the objective to minimise training loss and maximize the detection accuracy [20]. Tumour classification using features extracted from a 17-layer deep CNN has been experimented on BRATS dataset. An accuracy of 96.19% is achieved from the features of fully connected layers classified using four different ML classifiers [21]. Three different DL CNN architectures with 13, 25, and 16 layers are used for detection, classification, and severity grading of brain tumours. Hyperparameters are optimised by grid search algorithm. Custom CNN exhibits superior performance over the TL models [22].

A lightweight 13-layer CNN is proposed to classify brain tumours. CNN model is trained using images from Figshare dataset and tested using images from Kaggle dataset. Maximum accuracy of 97.2% is reported for a 3-class tumour classification [23]. The use of pretrained deep features extracted from 9 different pretrained networks is classified using 9 different classifiers. Ensembles of the three best deep features are fused to improve classification accuracy. An average accuracy of 98.83% is obtained on the 2-class Kaggle dataset [24].

A Custom CNN model with 3 different number of layers (19, 22, 25) is used as the baseline system for tumour detection (2 class and 4 class). A 22-layer Custom CNN model is used as a source model for transfer learning to identify subclasses of tumours from a target dataset. A classification accuracy of 96.9% is reported on 4-class tumour classification [25]. A multiphase deep boosted framework using TL based pretrained CNN and customised CNN is analysed to improve the performance of breast cancer CAD [26]. Deep CNN frameworks are used to automate brain tumour diagnosis [27, 28]. Improved accuracy is observed in brain tumour detection and classification using DL networks when compared to ML approaches. The deep networks are experimented on publicly available MRI datasets.

From the literature studies, the following observations are made:

- ML approaches are time consuming, and feature driven.
- DL approaches improve the performance at the cost of more labelled data.
- TL approach using pretrained network improves the system performance with a reduced dataset but consumes more training time.
- Deep feature extraction from pretrained networks provides a trade-off between accuracy and computational complexity.

Henceforth, the proposed brain tumour detection and classification system experiments with the use of Resnet18 deep features for improved performance. Features extracted from the global pooling layer and fusion of features from shallow and deep layers of the pretrained network are analysed in the proposed work. The workflow is presented in Fig. (**1**).

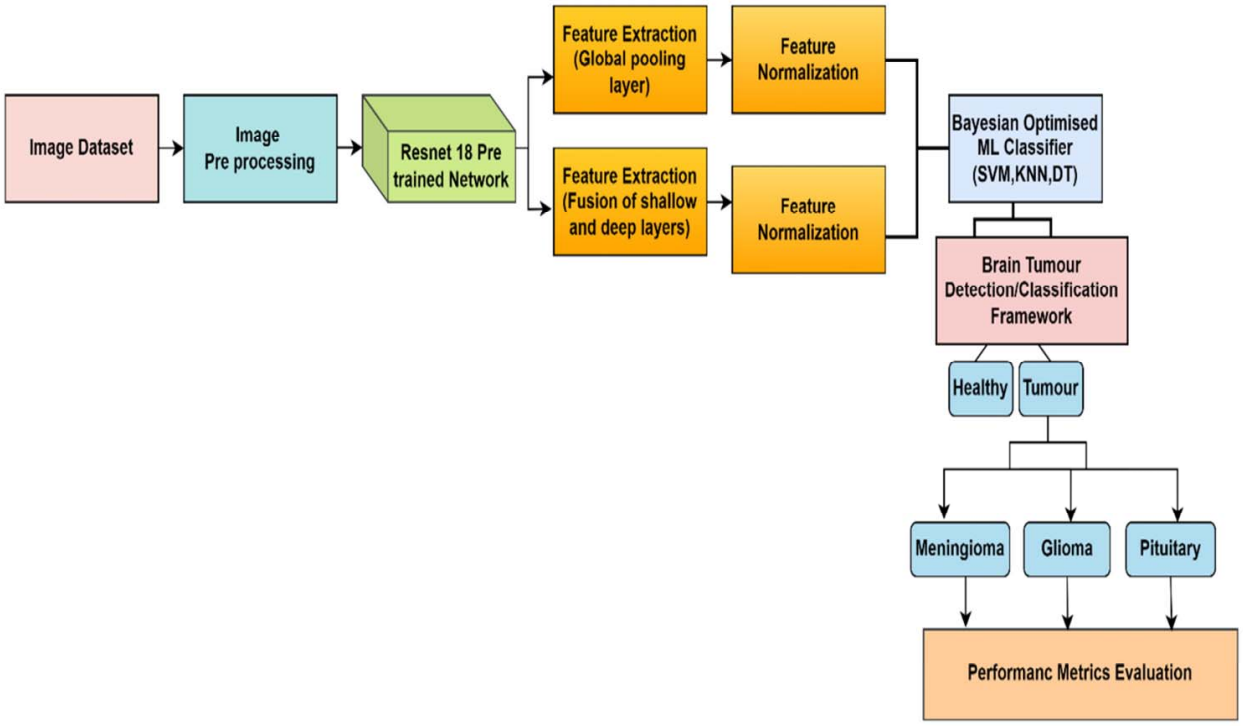


Fig. (1). Proposed brain tumour CAD system.

3. MATERIALS AND METHODS

3.1. Image Dataset

MRI images of brain tumour and normal individuals available in dataset [29, 30] were used for the proposed work. 2080 MR images of healthy people and 2426 MR images of tumour-affected persons available were used for detection [29]. Dataset [30] consisted of 1026 glioma tumour, 1053 meningioma tumour and 1006 pituitary tumour MRI images and was used for brain tumour classification. MRI brain image used for brain tumour detection and classification is presented in Fig. (**2**).

3.2. Image Pre-processing

The images in the database were resized to 224x224. Resizing was performed using bicubic interpolation, which was computationally efficient and produced an image of considerable quality. Bicubic interpolation considered 16 pixels at a time around an unknown pixel. Closer pixels to unknown pixels were given more weight to determine the interpolated pixel value. Bicubic interpolation retained fine details to produce a smoothened image. Resizing was followed by image normalization.

3.3. Baseline System

A baseline system was implemented using Convolutional Neural Network (CNN) and Resnet-18 Pretrained network. CNN can be used for classification and feature extractor. Convolution layer and pooling layer in CNN were used for feature extraction. Fully connected (FC) layer and Softmax (SM) layer enabled classification in CNN. Addition of convolution and pooling layers could be used to increase the depth of CNN. As CNN learned the spatial and temporal dependencies of an image, it was used to extract better features for classification. Despite the improved performance, CNN requires a careful selection of layer definitions and hyperparameters when developing from scratch. Two CNN architectures, CNN-3 and CNN-5, were experimented. CNN-3 and CNN-5 architectures used 3 and 5 stacked convolution layers respectively. Each stack consisted of a Convolution layer, Batch Normalization layer, Relu activation layer. Stacks 1 and 2 were followed by the max pooling layer. Stack 3 was followed by FC, SM and classification layer. The architecture of CNN-3 and CNN-5 network is presented in (Figs. 3 and 4) respectively. TL using Resnet18 pretrained network was also experimented.

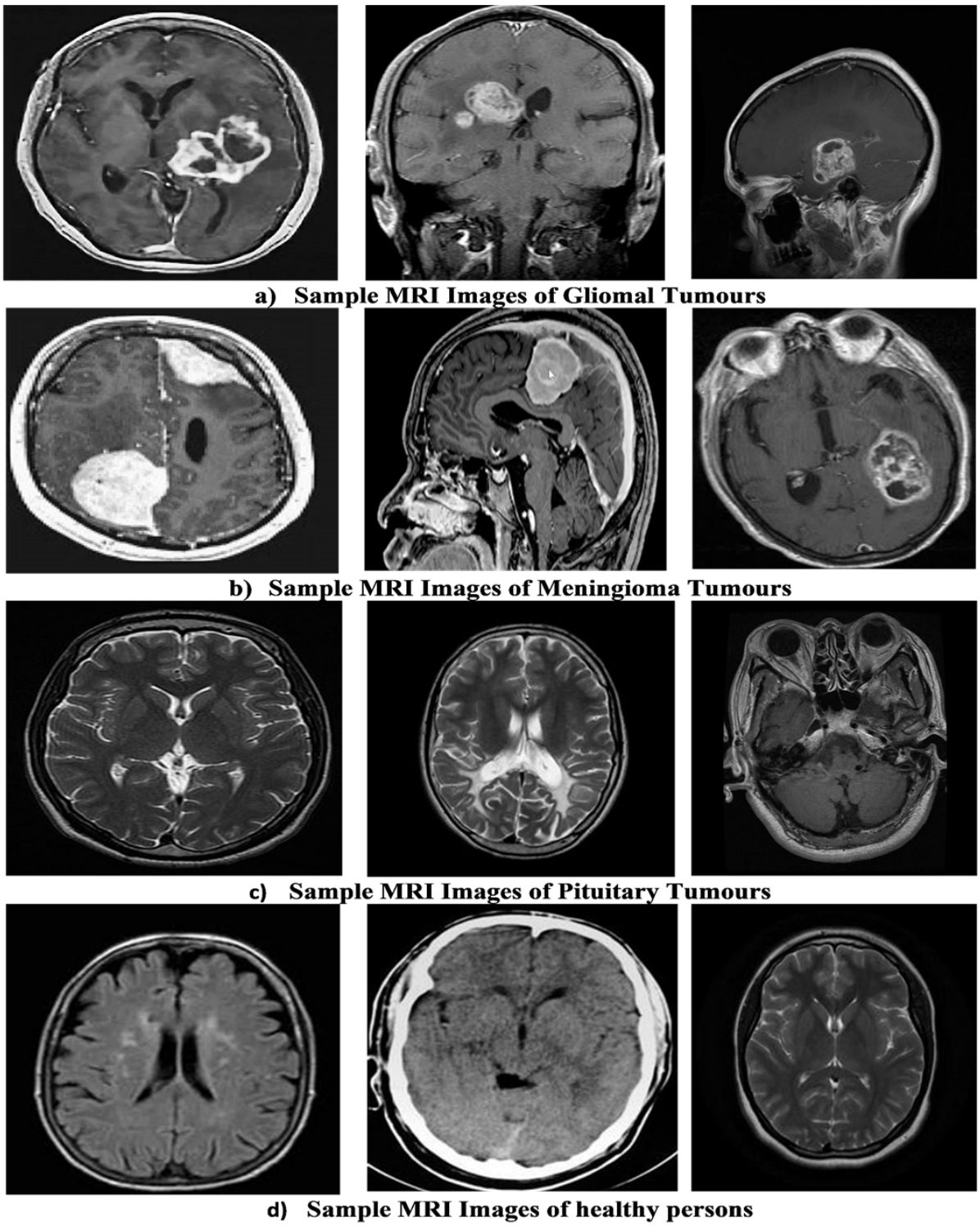


Fig. (2a-d). Brain MRI images used for detection and classification.

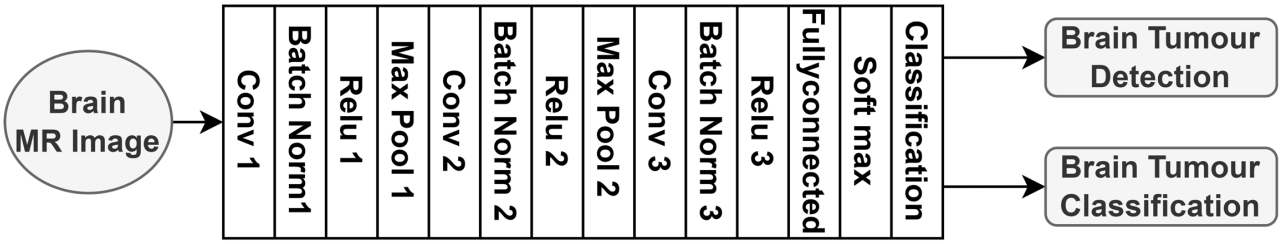


Fig. (3). CNN3- architecture.

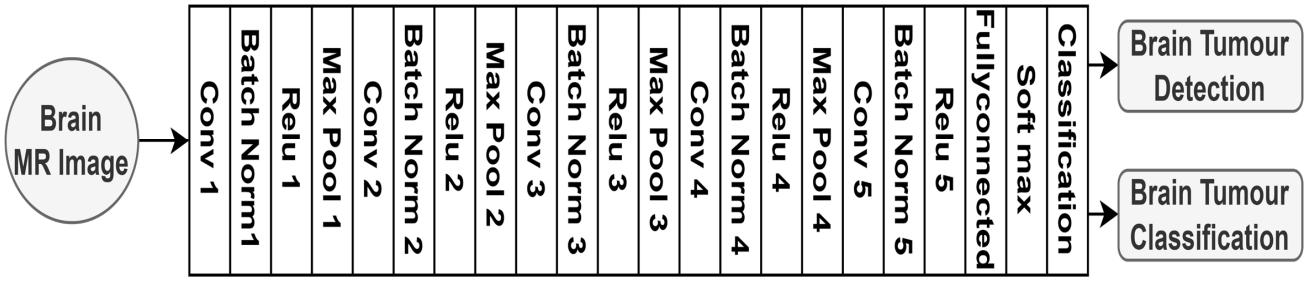


Fig. (4). CNN-5 architecture.

3.4. Proposed Method

The use of deep features in pretrained networks leverages a large training time associated with transfer learning. The steps involved in the brain tumour detection/classification framework are outlined in the following section:

3.4.1. Resnet 18 Deep Feature Extraction

Resnet 18 is a 72-layer architecture with 18-layer depth [31]. Resnet architectures exhibit improved performance over Densenet and VGG architectures and are frequently employed in signal and image classification tasks [32]. Densenet architecture uses concatenation and preserves features from earlier layers. Concatenation in densenet and use of smaller convolution operations in the network consumes high computational and memory resources. Visual Geometry Group (VGG) architecture uses 2D convolutions and suffers from performance degradation as the number of layers increases. The performance degradation is caused by the vanishing gradient during the weight update. Identity/skip connections in Resnet architecture improves the generalization capability. Identity connections learn previous layer activations and allow the network to learn faster by avoiding vanishing gradient problems. The local gradient is always set to 1 in Resnet by means of identity connections. Resnet has multiple variants such as Resnet 18, Resnet 50, Resnet 101. Resnet 18 provides a simple and compact feature representation of the brain MRI images [33]. Thus, Resnet 18 architecture is better in terms of feature representation, performance and computation speed and is used in the proposed work.

Resnet 18 trained on Image net dataset (source) can be used to extract deep features from a small MRI brain tumour detection and classification dataset (target). Resnet offers impressive results in medical detection problems, hence the deep features extracted from Resnet18 pretrained architecture are used as a feature for brain tumour detection and classification. The reason for using pretrained networks as feature extractor, is to reduce the training time and to avoid training deep CNN from scratch. Fixed weights of deep CNN models pretrained on imagenet dataset is used for extracting deep features.

The resized images are fed to activation to extract deep features. Earlier layers correspond to low level features and the last layer in the pretrained network learns the hierarchical representation of images. Resnet18 pre-trained network is trained on colour images. To use Resnet 18 for MRI imaging modality, grayscale MRI images are converted to colour images. Gray scale value is repeated along R, G, B channels. A new dimension is created for colour image and the same grayscale image appears over 3 channels, with the performance similar to a colour image. The size of the feature map in layers of Resnet 18 is presented in Table **1**. M is the number of images.

Table 1. Feature map dimensions in Resnet 18 layers.

S.No.	Layer Name	Size of Feature Map	Activations
1.	Pool5	512 x M	1x1x512
2.	Res 2a layer	64 x M	56x56x64
3.	Res 2b layer	64 x M	56x56x64
4.	Res 3a layer	128 x M	28x28x128
5.	Res 3b layer	128 x M	28x28x128
6.	Res 4a layer	256 x M	14x14x256
7.	Res 4b layer	256 x M	14x14x256
8.	Res 5a layer	512 x M	7x7x512

Activations on the global pooling layer (pool5) result in 512 features, obtained through pooling the feature maps over the spatial locations. Features extracted in earlier shallow layers (res2a, 2b, 3a, 3b, 4a, 4b, 5a) have high spatial resolution and large number of activations. In Res 2a layer, the spatial size of activation is 56x56 and outputs 64 features for a

single image. In order to reshape the features of the form MxF (F-Feature dimension), the activations in the earlier shallow layers are averaged over all spatial locations.

Two feature sets extracted from Resnet 18 are analysed to improve the system performance. Activations extracted from the global pooling layer and fusion of shallow and deep layers generate 512 and 1920 features for a single image, respectively.

Feature Set1(FS1) - Global pooling layer (pool 5)

Feature Set2 (FS2) - Fusion of 8 layers (res2a, res2b, res3a, res3b, res4a, res4b, res5a, pool5)

The layer names used in FS1 and FS2 correspond to the layer definitions of Resnet 18 architecture as presented in Fig. (**5**). The structure of the residual blocks is further detailed in Fig. (**6**). Residual deep features from 7 residual blocks (res2a, res2b, res3a, res3b, res4a, res4b, res5a) and global pooling layer (pool 5) features are fused to form FS2.

Fig. (5). Resnet 18 architecture.

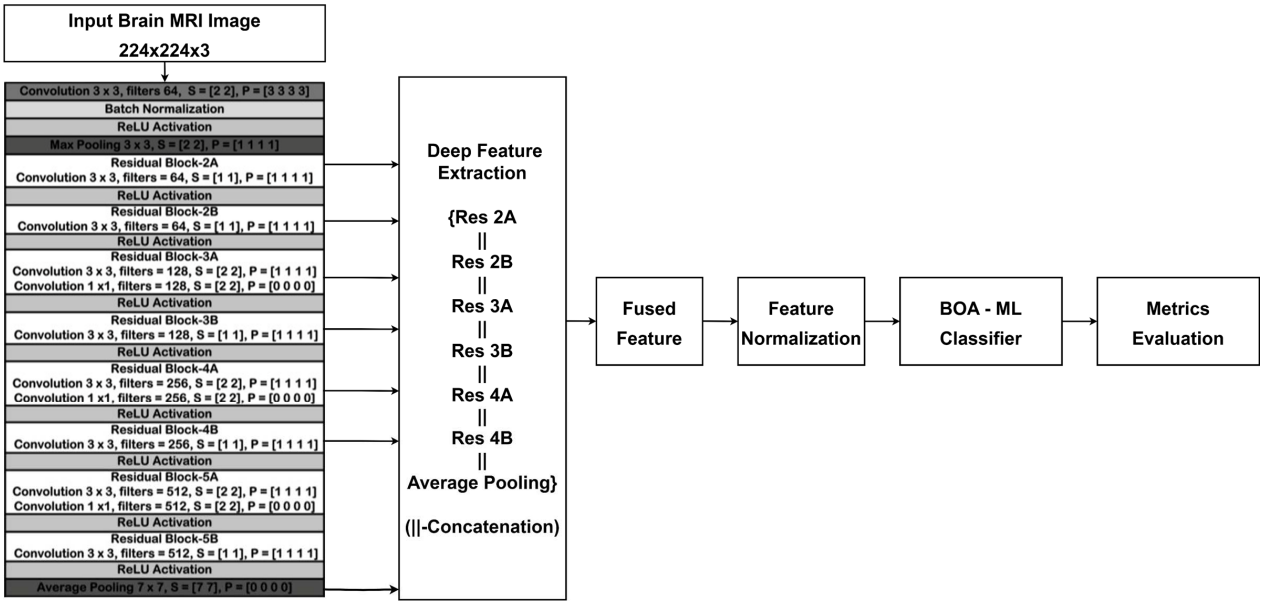


Fig. (6). Proposed feature fusion framework from resnet 18 network.

3.4.2. Feature Normalization

Feature normalization is done after deep feature fusion, as deep features contain features fused at different scales [34]. Feature normalization scales the feature vectors to [0,1] and enables easy classification and mapping of features to respective class labels. Feature normalization improves the classification ability and feature mapping of different classes by normalizing feature vectors to a similar range. Furthermore, feature normalization also gives equal importance to all feature vectors and prevents the bias caused by a feature with a high value. Extracted Resnet18 deep features are normalized using Z score. In Z score normalization, each data point in the feature vector is subtracted from the mean value (μ) and divided by the Standard Deviation (SD). Feature matrix is centered around μ of 0 and SD of 1.

$$Z - Score = \frac{x-\mu}{SD} \quad \textbf{(1)}$$

3.4.3. Brain Tumour Detection and Classification Framework

Deep features extracted from the Resnet-18 pretrained architecture were classified using the 3 ML classifiers listed below. Tumour detection (Normal/Healthy) and tumour classification (3 types of tumours) are experimented in the proposed work.

(i) Support Vector Machine (SVM): SVM determines the best hyperplane that maximises the margin between classes. Linear SVM is employed for linearly separable data. SVM applies the kernel trick to transform nonlinear data into linear data. Gaussian, Quadratic, Cubic and Intersection are the nonlinear kernel functions used in SVM.

(ii) K-Nearest Neighbour (KNN): KNN selects the K nearest data points (neighbours) and calculates the distance between the new data points and the neighbours and filters out the one with the shortest distance. The label of the new data point is assigned based on the majority labels of K nearest samples.

(iii) Decision Tree (DT): DT is a tree-structured algorithm used for classification. The nodes of the DT represent features /predictors of the data, branches represent the criteria, and the leaf node specifies the outcome. To predict a class from the feature set, an attribute selection criterion is used to choose the best attribute starting from the root node until further classification is not possible.

3.4.4. Bayesian Optimization

The hyperparameters of the above classifiers are optimised using Bayesian optimization. Steps involved in Bayesian optimisation are discussed in this section.

Hyperparameters are the parameters on which an ML algorithm operates. Selection of optimal hyperparameters in ML algorithm results in better classification performance.

Equation (**2**) represents the hyperparameter optimization

$$\dot{x} = arg \min_{x \in X} f(x) \quad \textbf{(2)}$$

f(x) represents the objective function to be minimised. ẋ denotes the hyperparameters that result in the minimising of the objective function and x takes any value in X . The objective function considered in this work is the classification error. The optimal value of hyperparameters that minimises the classification error is determined by the optimization algorithm.

Hyperparameters of ML algorithms are optimised using Bayesian Optimisation (BO). BO aims to create a probability model of the objective function and use it to identify the optimal hyperparameters to test the true objective function. BO uses Bayes theorem to find the minimum classification error.

BO uses the results of previous evaluations to construct a probabilistic model that maps hyperparameters to the score probability of the desired objective function.:

$$P(Score|Hyperparameters) \quad \textbf{(3)}$$

The model in (3) is called a surrogate for the objective function and is denoted as P(y|x). It is easier to optimise the surrogate function than the original objective function. BO chooses the next set of hyperparameters to test the actual objective function by determining the hyperparameter that performs well on the surrogate.

The steps involved in BO are:

1. Create a surrogate probability model for the chosen objective function.

2. Identify the hyperparameters performing well on the surrogate

3. Hyperparameters determined in Step 2 are applied to the true objective function

4. Finetune Surrogate model based on the results

5. Repeat steps 2 to 4 until the maximum iteration is reached

Bayesian optimization aims to achieve less error with more data; this is accomplished by continuously updating the surrogate model after each evaluation of the chosen objective function. The number of iterations for the optimization of the classification algorithm is set to 30 through empirical analysis.

3.4.5. Performance Metrics Evaluation

Confusion matrix obtained from the classifier output is used to determine True Negative (TRN), True Positive (TRP), False Negative (FAN), and False Positive (FAP). 10-fold cross-validation is used to avoid overfitting. In 10-fold cross-validation, the total feature set is divided into 10 folds with equal sizes. For the training set, 9 folds are used, and 1-fold is used for validation. Total accuracy is the average accuracy of 10 folds.

4. RESULTS

Metrics listed in Table **2** were used to evaluate the brain tumour detection and classification system performance. Multiple metrics are used to evaluate the system in an efficient way. All experiments are performed using MATLAB 2020a

software.

A baseline system using CNN is experimented and the results are tabulated in Table **3**. An accuracy of 96.67% and 91.36% is achieved for brain tumour detection and classification framework with 5 convolutional layers. More the number of layers in CNN, better is the accuracy and other metrics. Transfer Learning using Resnet18 pretrained network yields an accuracy of 98.96% for detection and 94.60% for classification. The improved accuracy using Resnet18 network results in a training time of 2.46 Hours and 1.78 hours for detection and classification respectively.

Table 2. Evaluation metrics.

S.No.	Metric	Equation
1.	Accuracy	$\mathrm{ACC} = \frac{\mathrm{TRP}+\mathrm{TRN}}{\mathrm{FAP}+\mathrm{FAN}+\mathrm{TRP}+\mathrm{TRN}}$ **(4)**
2.	Sensitivity	$\mathrm{SEN} = \frac{\mathrm{TRP}}{\mathrm{FAN}+\mathrm{TRP}}$ **(5)**
3.	Specificity	$\mathrm{SPC} = \frac{\mathrm{TRN}}{\mathrm{FAP}+\mathrm{TRN}}$ **(6)**
4.	Precision	$\mathrm{PRE} = \frac{\mathrm{TRP}}{\mathrm{TRP}+\mathrm{FAP}}$ **(7)**
5.	F1score	$\mathrm{F}_1 = 2 \times \frac{\mathrm{PRE} \times \mathrm{SEN}}{\mathrm{PRE} + \mathrm{SEN}}$ **(8)**
6.	Mathews Correlation Coefficient	$\mathrm{MCC} = \frac{\mathrm{TRP}.\mathrm{TRN} - \mathrm{FAP}.\mathrm{FAN}}{\sqrt{(\mathrm{TRP}+\mathrm{FAP})(\mathrm{TRP}+\mathrm{FAN})(\mathrm{TRN}+\mathrm{FAP})(\mathrm{TRN}+\mathrm{FAN})}}$ **(9)**
7.	Balance Classification rate	$\mathrm{BCR} = \sqrt{\mathrm{SEN}.\mathrm{SPC}}$ **(10)**
8.	Kappa Coefficient	$\mathrm{Kp} = \frac{\mathrm{p}-\mathrm{e}(\kappa)}{1-\mathrm{e}(\kappa)}$ **(11)**

Note:*p- overall percent agreement; e(κ)- Probability of chance agreement.

Table 3. Performance metrics of baseline system.

S.No.	System	Deep N/w	ACC	SEN	SPC	PRE	F1	BCR	MCC	Kp
1.	Detection	CNN -3	93.64	95.67	91.90	91.01	93.28	93.77	87.35	87.25
		CNN- 5	96.67	95.03	98.08	97.69	96.34	96.54	93.29	93.29
		Resnet -18	98.96	98.26	99.58	99.52	98.89	98.92	97.93	97.92
2.	Classification	CNN -3	75.16	75.21	87.55	75.42	75.31	81.14	62.93	44.11
		CNN- 5	91.36	91.38	95.67	91.48	91.43	93.50	87.11	80.56
		Resnet-18	94.60	94.66	97.34	94.67	94.66	95.99	91.98	87.85

Table 4. Performance metrics of brain tumour detection and classification system.

S.No.	System	Classifier	ACC	SEN	SPC	PRE	F1	BCR	MCC	Kp
1.	Detection FS1	SVM	92.99	92.79	93.16	92.08	92.43	92.97	85.90	85.90
		KNN	**95.38**	**94.57**	**96.08**	**95.39**	**94.98**	**95.32**	**90.71**	**90.71**
		DT	86.62	86.01	87.14	85.15	85.58	86.57	73.10	73.10
2.	Detection FS2	SVM	**97.40**	**97.50**	**97.32**	**96.89**	**97.20**	**97.41**	**94.78**	**94.78**
		KNN	95.41	94.42	96.25	95.57	94.99	95.33	90.76	90.75
		DT	92.08	91.92	92.21	91.00	91.46	92.07	84.08	84.07
3.	Classification FS1	SVM	**85.06**	**85.11**	**92.50**	**85.30**	**85.21**	**88.73**	**77.70**	**66.38**
		KNN	82.59	82.74	91.30	82.56	82.65	86.91	73.98	60.83
		DT	72.67	72.70	86.30	73.02	72.86	79.21	59.15	38.52
4.	Classification FS2	SVM	**94.94**	**97.46**	**95.08**	**95.08**	**95.01**	**96.19**	**92.47**	**88.62**
		KNN	91.28	91.33	95.64	91.28	91.31	93.46	86.95	80.38
		DT	80.49	80.52	90.22	80.62	80.57	85.24	70.80	56.09

To further enhance the accuracy and reduce training time, deep features extracted from Resnet18 pretrained architecture are proposed. The results obtained using FS1 and FS2 for detection and classification are tabulated in Table **4**. FS2 exhibits improved metrics than FS1. Fusion of features from the shallow and deep layers of resnet18 network has improved the detection and classification performance. SVM with linear kernel, KNN (K=10) with Euclidean distance metric and Decision Tree with 100 splits and Gini index as split criteria are used as classifier hyperparameters. An improved accuracy of 97.40% and 94.94% is achieved for detection and classification respectively, using FS2 classified by SVM.

Tuning of hyperparameters in an ML classifier was further experimented to improve the system performance. Hyperparameters of ML classifiers optimised by Bayesian algorithm are listed in Table **5**. The tuned hyperparameters were used in the classifier algorithm to enhance the metrics. Apart from accuracy, other metrics listed in Table **2** are determined to further quantify the performance of the system. Accuracy measures the ratio of correct prediction to total prediction. Sensitivity and specificity measured the true positive and true negative rates of a classifier. A tumour detection/classification system should have high sensitivity and specificity to avoid false positive and false negative rates. Precision is the classifier's ability not to classify a sample as a cancer if it is healthy. F1 score and BCR were used to measure a classifier performance under an imbalanced class setting. High value of F1 score and BCR (close to 1) indicated better classifier performance. The kappa score quantified the degree level of agreement between the obtained and actual values. MCC yields a high value (close to 1) only if all classes are predicted correctly. From Table **6**, it is observed that parameter tuning resulted in an accuracy of 99.11% and 97.31% for detection and classification, respectively, using FS2 classified by SVM with optimal hyperparameters. Furthermore, the high value of other metrics for FS2 classified by optimal SVM indicated that feature fusion enhanced the system performance. FS2 feature set classified by SVM consumed training time of 101.35 seconds for detection and 83.19 seconds for classification. Furthermore, FS2 feature classified by optimized SVM required 22 minutes and 17 minutes for detection and classification, respectively. Thus, the use of fusion of Resnet 18 deep features classified by ML classifiers consume less training time and improved performance when compared to Pretrained Resnet 18 architecture.

Table 5. Bayesian optimised hyperparameters used in ML classifier.

S.No.	System	Classifier	Parameter 1	Parameter 2	Parameter 3	Parameter 4
1.	Detection FS1	Optimized SVM	Kernel Function: **Quadratic**	Kernel Scale: **1**	Box Constraint Level: **990.4171**	Standardization: **True**
		Optimized KNN	Number of neighbors: **1**	Distance Metric: **City block**	Distance Weight: **Squared Inverse**	Standardization: **True**
		Optimized DT	Max no of splits : **230**	Split criterion: **Maximum Deviance reduction**	-	-
2.	Detection FS2	Optimized SVM	Kernel Function: **Gaussian**	Kernel Scale: **286.43**	Box Constraint Level: **362.97**	Standardization: **True**
		Optimized KNN	Number of neighbors: **1**	Distance Metric: **Spearman**	Distance Weight: **Equal**	Standardization: **True**
		Optimized DT	Max no of splits : **230**	Split criterion: **Maximum Deviance reduction**	-	-
3.	Classification FS1	Optimized SVM	Kernel Function: **Quadratic**	Kernel Scale: **1**	Box Constraint Level: **0.0010216**	Standardization: **True**
		Optimized KNN	Number of neighbors: **1**	Distance Metric: **Spearman**	Distance Weight: **Squared Inverse**	Standardization: **True**
		Optimized DT	Max no of splits : **185**	Split criterion: **Twoing rule**	-	-
4.	Classification FS2	Optimized SVM	Kernel Function: **Quadratic**	Kernel Scale: **1**	Box Constraint Level: **0.0010046**	Standardization: **True**
		Optimized KNN	Number of neighbors: **1**	Distance Metric: **Correlation**	Distance Weight: **Inverse**	Standardization: **True**
		Optimized DT	Max no of splits : 64	Split criterion: **Twoing rule**	-	-

Table 6. Performance metrics of brain tumour detection and classification system after bayesian optimisation.

S.No.	System	Classifier	ACC	SEN	SPC	PRE	F1	BCR	MCC	Kp
1.	Detection FS1	Optimized SVM	98.69	98.12	99.18	99.03	98.58	98.65	97.37	97.36
		Optimized KNN	**98.80**	**98.41**	**99.13**	**98.98**	**98.70**	**98.77**	**97.59**	**97.59**
		Optimized DT	88.17	87.79	88.50	86.75	87.26	88.14	76.23	76.22

(Table 6) contd.....

S.No.	System	Classifier	ACC	SEN	SPC	PRE	F1	BCR	MCC	Kp
2.	Detection FS2	Optimized SVM	**99.11**	**98.99**	**99.22**	**99.09**	**99.09**	**99.10**	**98.21**	**98.21**
		Optimized KNN	98.96	98.80	99.09	98.94	98.87	98.95	97.90	97.90
		Optimized DT	92.31	92.31	92.58	91.43	91.87	92.44	84.83	84.83
3.	Classification FS1	Optimized SVM	93.32	93.35	96.65	93.36	93.35	94.99	90.01	84.98
		Optimized KNN	**93.97**	**94.01**	**96.98**	**93.99**	**94.00**	**95.48**	**90.98**	**86.43**
		Optimized DT	73.84	73.87	86.89	74.09	73.98	80.12	60.86	41.14
4.	Classification FS2	Optimized SVM	**97.31**	**97.30**	**98.65**	**97.37**	**97.34**	**97.97**	**95.99**	**93.95**
		Optimized KNN	94.13	94.17	97.06	94.14	94.15	95.61	91.22	86.80
		Optimized DT	81.88	81.87	90.91	82.29	82.08	86.27	73.00	59.23

Table 7. Comparison with existing works in the literature.

S.No.	Reference No.	Classification	Accuracy (%)
1.	Swati *et al.* [11]	Glioma, Meningioma, Pituitary tumour	94.82
2.	Das *et al.* [13]	Tumour, non-tumour	97.62
3.	Shwetha *et al.* [19]	Tumour, non-tumour	88.1
4.	Irmak *et al.* [22]	Glioma, Meningioma, Pituitary tumour	92.66
5.	Kang *et al.* [24]	Tumour, non-tumour	97.85
6.	Das S *et al.* [28]	Glioma, Meningioma, Pituitary tumour	84.19
7.	Proposed work	Tumour, non-tumour	99.11
8.	Proposed work	Glioma, Meningioma, Pituitary tumour	97.31

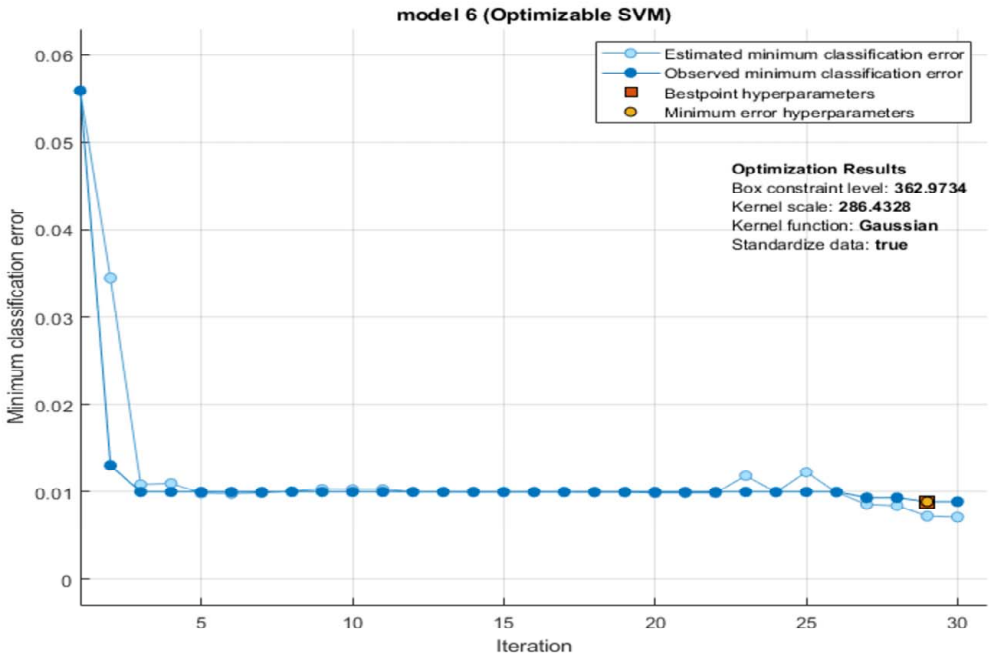


Fig. (7). MCE plot of Optimizable SVM for brain tumour detection.

5. DISCUSSION

The use of feature fusion of shallow and deep layers (FS2) in a Resnet 18 pretrained network improved the accuracy. The earlier layers in a pretrained network learned basic features and had better spatial resolution, while the deeper layers constructed using earlier layers contained meaningful features for efficient classification.

The choice of optimal hyperparameters using Bayesian probabilistic algorithm, improved the system performance by controlling overfitting and underfitting. From Table **6**, it is

observed that FS1 classified by KNN classifier with K=1 yields a maximum accuracy of 98.80% and 93.97% for detection and classification, respectively. City block distance metric for detection and Spearman metric for classification yielded optimised results. Furthermore, FS2 exhibited better detection and classification performance using nonlinear Gaussian and quadratic kernels, respectively. FS2 achieved accuracy, sensitivity, specificity, precision, F1 score, BCR, MCC, Kp of 99.11%, 98.99%, 99.22%, 99.09%, 99.09%, 99.10%, 98.21%, 98.21% respectively for detection. FS2 also performed better for classification task with accuracy, sensitivity, specificity, precision, F1 score, BCR, MCC and Kp of 97.31%, 97.30%, 98.65%, 97.37%, 97.34%, 97.97%, 95.99%, 93.95% respectively.

SVM classifier with nonlinear kernel outperformed KNN and DT classifiers. DT classifier exhibited reduced metric scores compared to other classifiers for FS1 and FS2. FS2 contained fused features from the earlier and deep layers of the Resnet18 network, thereby improving the classification performance. Minimum Classification error (MCE) plot and confusion matrix for the brain tumour detection task are presented in (Figs. **7** and **8**), respectively.

Figs. (**9** and 10) show the MCE plot and confusion matrix for the classification task. MCE plot gives the best point and minimum error hyperparameters for building an efficient classifier.

A comparative study of the proposed work with existing works in the literature on 3 class and 2 class brain tumour dataset is presented in Table **7**. The use of Resnet18 deep features classified by the optimised SVM classifier has outperformed the state-of-the-art results in the literature with increased accuracy and other relevant performance metrics. The dataset chosen for the proposed 3 class classification uses equal number of images in 3 classes to avoid class imbalance problem. Datasets in other literature use imbalanced datasets.

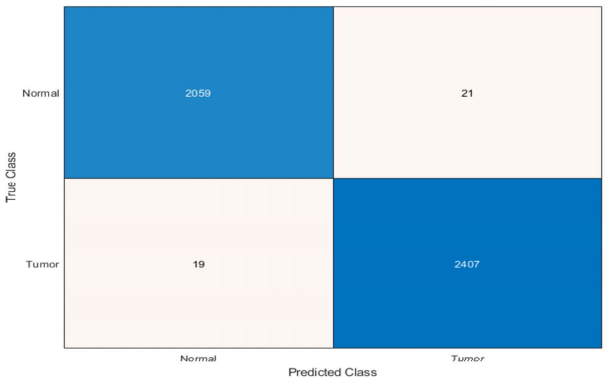


Fig. (8). Confusion Matrix of Optimizable SVM for brain tumour detection.

A paired sample t-test compares FS2 feature vectors for brain tumour detection and classification before and after hyperparameter tuning. A p-value of 0.000085267 for detection and 0.0023 for classification indicates that FS2 statistically outperforms Bayesian optimisation. As a result, it is concluded that the fusion of feature vectors from different layers of the Resnet18 architecture and classification by a hyperparameter-tuned classifier improved the system performance.

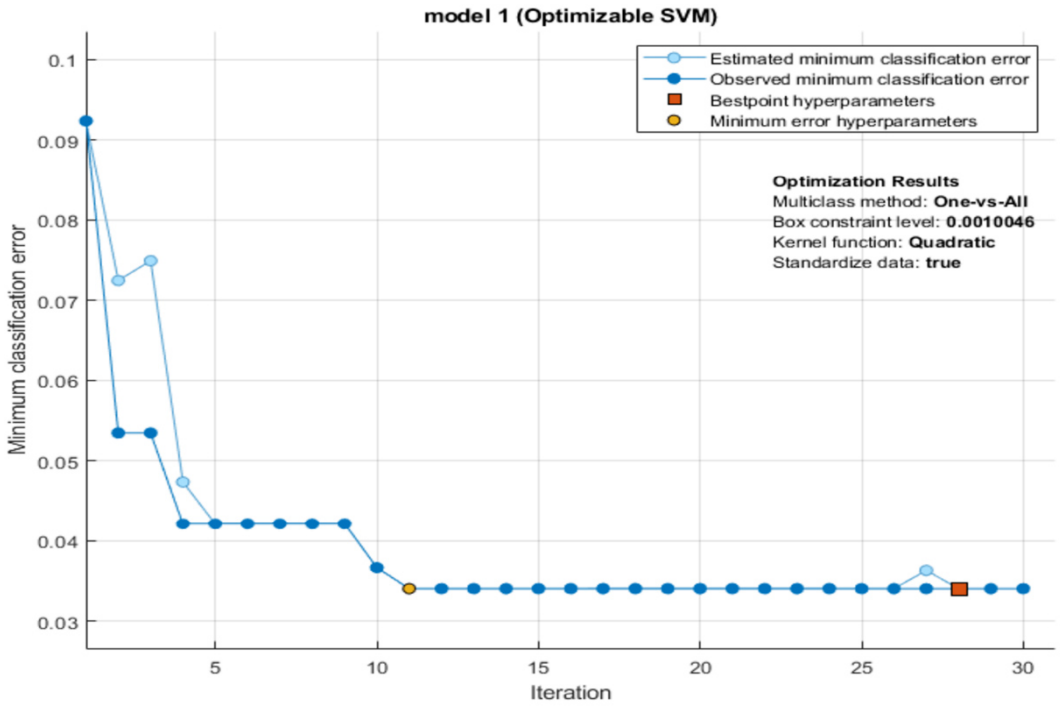


Fig. (9). MCE plot of Optimizable SVM for brain tumour classification.

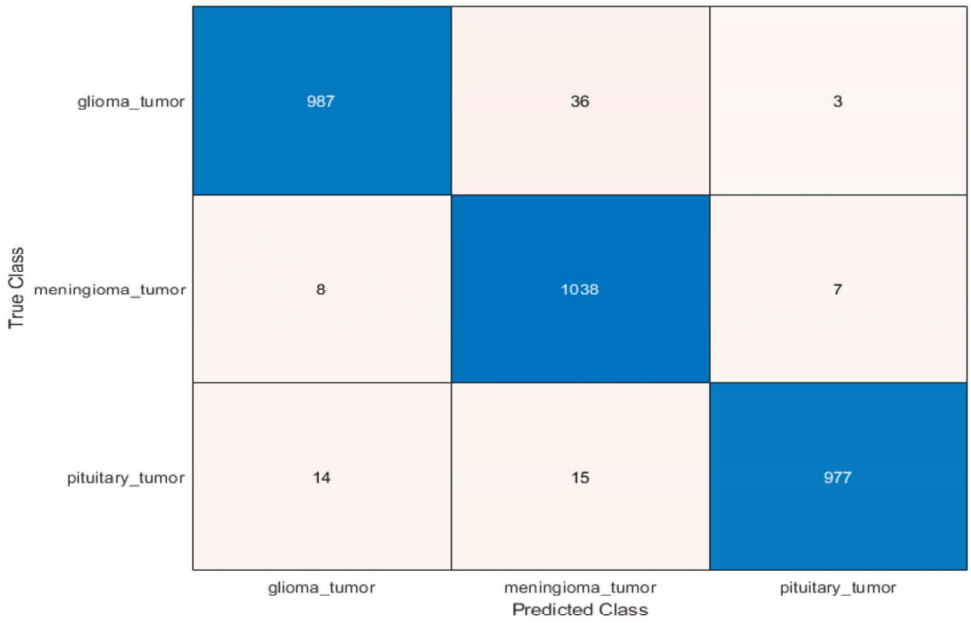


Fig. (10). Confusion Matrix of Optimizable SVM for brain tumour classification.

CONCLUSION

Manual analysis of brain tumours from MRI scan images requires a high level of radiologist expertise. A brain tumour detection and classification system using DL and ML is experimented in the proposed work to automate diagnosis. Feature extraction involves deep features extracted from the global pooling layer (FS1) and fusion of the deep and shallow layers (FS2) of Resnet 18 pretrained network. Further, to improve the performance, hyperparameter tuning of classifier parameters using Bayesian algorithm is performed. A maximum accuracy of 99.11% for detection and 97.31% for classification is achieved for FS2 classified by BA optimised SVM classifier. The proposed work is compared with existing works in the literature to claim improvement. Henceforth, the proposed work can be used as a tool to aid the radiologist in automating brain tumour detection and classification.

Techniques such as feature optimisation using metaheuristics and swarm-based techniques, feature transformation, and feature selection can be explored in the future to further improve the accuracy. A custom CNN with hyperparameters tuned using swarm-based optimisation can also be experimented to investigate the system performance.

LIST OF ABBREVIATIONS

CAD	=	Computer Aided Diagnosis
ML	=	Machine Learning
DL	=	Deep Learning
MRI	=	Magnetic Resonance Image
SVM	=	Support Vector Machine
KNN	=	K-Nearest Neighbour
DT	=	Decision Tree
BA	=	Bayesian Algorithm
BCR	=	Balance Classification Rate
MCC	=	Mathews Correlation Coefficient

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this study are available within the manuscript.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] King GG. Brain tumour classification: A comprehensive systematic review on various constraints. Comput Methods Biomech Biomed Eng Imaging Vis 2022; 1-3.

[2] Sharma R, Aashima , Nanda M, *et al.* Mapping cancer in africa: A comprehensive and comparable characterization of 34 cancer types using estimates from GLOBOCAN 2020. Front Public Health 2022; 10: 839835. [http://dx.doi.org/10.3389/fpubh.2022.839835] [PMID: 35548083]

[3] Arabahmadi M, Farahbakhsh R, Rezazadeh J. Deep learning for smart healthcare-A survey on brain tumour detection from medical imaging. Sensors 2022; 22(5): 1960. [http://dx.doi.org/10.3390/s22051960] [PMID: 35271115]

[4] Mzoughi H, Njeh I, Slima MB, Hamida AB. Review of computer aided-diagnosis (CAD) systems for MRI gliomas brain tumours explorations based on machine learning and deep learning. In2022 6th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP). 1-6.

[5] Nazir M, Shakil S, Khurshid K. Role of deep learning in brain tumour detection and classification (2015 to 2020): A review. Comput Med Imaging Graph 2021; 91: 101940. [http://dx.doi.org/10.1016/j.compmedimag.2021.101940] [PMID: 34293621]

[6] Devi PR, Vemuru S. Brain tumour detection with feature extraction and tumour cell classification model using machine learning–A survey. 2022 International Conference on Electronics and Renewable Systems (ICEARS). 1250-9. [http://dx.doi.org/10.1109/ICEARS53579.2022.9752080]

[7] Bharathi M, Kumar SA, Sasikala S. Deep learning architectures for improving effectiveness of Covid detection–A pilot study. 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA). 1-5. [http://dx.doi.org/10.1109/ICAECA52838.2021.9675714]

[8] Zhuang F, Qi Z, Duan K, *et al.* A comprehensive survey on transfer learning. Proc IEEE 2021; 109(1): 43-76. [http://dx.doi.org/10.1109/JPROC.2020.3004555]

[9] Sethy PK, Behera SK. A data constrained approach for brain tumour detection using fused deep features and SVM. Multimedia Tools Appl 2021; 80(19): 28745-60. [http://dx.doi.org/10.1007/s11042-021-11098-2]

[10] Ragab DA, Attallah O, Sharkas M, Ren J, Marshall S. A framework for breast cancer classification using Multi-DCNNs. Comput Biol Med 2021; 131: 104245. [http://dx.doi.org/10.1016/j.compbiomed.2021.104245] [PMID: 33556893]

[11] Swati ZNK, Zhao Q, Kabir M, *et al.* Brain tumour classification for MR images using transfer learning and fine-tuning. Comput Med Imaging Graph 2019; 75: 34-46. [http://dx.doi.org/10.1016/j.compmedimag.2019.05.001] [PMID: 31150950]

[12] Vankdothu R, Hameed MA, Fatima H. A brain tumour identification and classification using deep learning based on CNN-LSTM method. Comput Electr Eng 2022; 101: 107960. [http://dx.doi.org/10.1016/j.compeleceng.2022.107960]

[13] Das A, Mohapatra SK, Mohanty MN. Design of deep ensemble classifier with fuzzy decision method for biomedical image classification. Appl Soft Comput 2022; 115: 108178. [http://dx.doi.org/10.1016/j.asoc.2021.108178]

[14] Amin J, Anjum MA, Sharif M, Jabeen S, Kadry S, Moreno Ger P. A new model for brain tumour detection using ensemble transfer learning and quantum variational classifier. Comput Intell Neurosci 2022; 2022: 1-13. [http://dx.doi.org/10.1155/2022/3236305] [PMID: 35463245]

[15] Rani S, Ghai D, Kumar S, Kantipudi MVVP, Alharbi AH, Ullah MA. Efficient 3D alexnet architecture for object recognition using syntactic patterns from medical images. Comput Intell Neurosci 2022; 2022: 1-19. [http://dx.doi.org/10.1155/2022/7882924] [PMID: 35634047]

[16] Kibriya H, Amin R, Alshehri AH, Masood M, Alshamrani SS, Alshehri A. A novel and effective brain tumour classification model using deep feature fusion and famous machine learning classifiers. Comput Intell Neurosci 2022; 2022: 1-15. [http://dx.doi.org/10.1155/2022/7897669] [PMID: 35378808]

[17] Senan EM, Jadhav ME, Rassem TH, Aljaloud AS, Mohammed BA, Al-Mekhlafi ZG. Early diagnosis of brain tumour mri images using hybrid techniques between deep and machine learning. Comput Math Methods Med 2022; 2022: 1-17. [http://dx.doi.org/10.1155/2022/8330833] [PMID: 35633922]

[18] Alsaif H, Guesmi R, Alshammari BM, *et al.* A novel data augmentation-based brain tumour detection using convolutional neural network. Appl Sci 2022; 12(8): 3773. [http://dx.doi.org/10.3390/app12083773]

[19] Shwetha V, Madhavi CR, Nagendra Kumar M. Classification of brain tumours using hybridized convolutional neural network in brain MRI images. International Journal of Circuits, Systems and Signal Processing 2022; 16: 561-70.

[20] Nayak DR, Padhy N, Mallick PK, Bagal DK, Kumar S. Brain tumour classification using noble deep learning approach with parametric optimization through metaheuristics approaches. Computers 2022; 11(1): 10. [http://dx.doi.org/10.3390/computers11010010]

[21] Latif G, Ben Brahim G, Iskandar DNFA, Bashar A, Alghazo J. Glioma tumours' classification using deep-neural-network-based features with SVM classifier. Diagnostics 2022; 12(4): 1018. [http://dx.doi.org/10.3390/diagnostics12041018] [PMID: 35454066]

[22] Irmak E. Multi-classification of brain tumour MRI images using deep convolutional neural network with fully optimized framework. Iran J Sci Technol Trans Electr Eng 2021; 45(3): 1015-36. [http://dx.doi.org/10.1007/s40998-021-00426-9]

[23] Kibriya H, Masood M, Nawaz M, Nazir T. Multiclass classification of brain tumours using a novel CNN architecture. Multimedia Tools Appl 2022; 81(21): 29847-63. [http://dx.doi.org/10.1007/s11042-022-12977-y]

[24] Kang J, Ullah Z, Gwak J. Mri-based brain tumour classification using ensemble of deep features and machine learning classifiers. Sensors 2021; 21(6): 2222. [http://dx.doi.org/10.3390/s21062222] [PMID: 33810176]

[25] Alanazi MF, Ali MU, Hussain SJ, *et al.* Brain tumour/mass classification framework using magnetic-resonance-imaging-based isolated and developed transfer deep-learning model. Sensors 2022; 22(1): 372. [http://dx.doi.org/10.3390/s22010372] [PMID: 35009911]

[26] Zahoor MM, Qureshi SA, Bibi S, *et al.* A new deep hybrid boosted and ensemble learning-based brain tumour analysis using MRI. Sensors 2022; 22(7): 2726. [http://dx.doi.org/10.3390/s22072726] [PMID: 35408340]

[27] Seetha J, Raja SS. Brain tumour classification using convolutional neural networks. Biomed Pharmacol J 2018; 11(3): 1457-61. [http://dx.doi.org/10.13005/bpj/1511]

[28] Das S, Aranya OR, Labiba NN. Brain tumour classification using convolutional neural network. 2019 1st international conference on advances in science, engineering and robotics technology (ICASERT). 1-5. [http://dx.doi.org/10.1109/ICASERT.2019.8934603]

[29] Vashishtha P. Brain tumour classification. 2021. Available From: https://www.kaggle.com/code/purvitsharma/brain-tumour-classification-98-4-accuracy/data

[30] Bhuvaji S, Kadam A, Bhumkar P, Dedge S, Kanchan S. Brain tumour classification (MRI). 2020. [http://dx.doi.org/10.34740/KAGGLE/DSV/1183165]

[31] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. 770-8.

[32] Polat Ö, Güngen C. Classification of brain tumours from MR images using deep transfer learning. J Supercomput 2021; 77(7): 7236-52. [http://dx.doi.org/10.1007/s11227-020-03572-9]

[33] Ayyachamy S, Alex V, Khened M, Krishnamurthi G. Medical image retrieval using Resnet-18.Medical imaging Imaging informatics for healthcare, research, and applications. Int Soc Opt Photo 2019; Vol. 10954: p. 1095410.

[34] Sannasi Chakravarthy SR, Rajaguru H. Automatic detection and classification of mammograms using improved extreme learning machine with deep learning. IRBM 2022; 43(1): 49-61. [http://dx.doi.org/10.1016/j.irbm.2020.12.004]