International Journal on Advanced Science Engineering Information Technology

Vol.12 (2022) No. 4
ISSN: 2088-5334

Two-Stage Approach of Hierarchical Deep Feature Representation and Fusion Frameworks for Brain Image Analysis

S. J. Prashantha [a,*], H. N. Prakash [b]

[a] Department of Computer Science and Engineering, AIT, Chikkamagaluru-577102, Visvesvaraya Technology University, Belagavi, India
[b] Department of Computer Science and Engineering, RIT, Hassan-573 201, Visvesvaraya Technology University, Belagavi, India
Corresponding author: *prasi.sjp@gmail.com

Abstract **— In recent decades, magnetic resonance (MR) brain images have initiated a wide range of image classification and segmentation methods. Feature representation is one of the essential aspects of medical image analysis. This paper proposes and investigates specific features that address the significance of high-level tasks with little annotation for medical images. Deep learning is a futuristic area of research in biomedical image analysis, in which the scope is moving us immediately to the goal of automating tasks in intelligent retrieval systems. This approach can incorporate many levels of feature representation to construct recognition of medical cells or images. We propose a novel approach based on the deep hierarchical features of two different convolutional neural network (CNNs) model choices to achieve competitive performance in the classification task. We explore feature representation through discriminative CNN models. The principal study of our proposed work is feature representations, feature-level fusion, and classification. Meanwhile, effective fusion frameworks were employed for brain MR image classification by using serial fusion and fusion operator strategies. The accuracy of the proposed technique is demonstrated using the Cancer Imaging Archive (TCIA) and Information eXtraction from Images (IXI) datasets. To the best of our knowledge, experiment results show that CNNs feature maps as input to the classifier and are superior to the original CNNs. The performance of the support vector machines (SVM) classifier has been used to evaluate in terms of training performance and classify subjects as either normal or abnormal.**

Keywords **—Medical image; feature representation; deep features; support vector machine; feature level fusion.**

Manuscript received 10 Aug. 2021; revised 28 Sep. 2021; accepted 16 Dec. 2021. Date of publication 31 Aug. 2022.

I. INTRODUCTION

Medical imaging and computer vision roles are growing day by day. Due to deep features learning, the current medical image analysis is experiencing a transition. This can be used for both learnings of features and classification. Meanwhile, objects that cells are characterized by pivotal or decisive features, detailed annotation of medical images, creating a group of discriminative specific features, etc., are often uncertain and tough tasks. Due to the nature of medical images, many tasks such as classification, segmentation, and detecting objects in clinical images are emerging. As a direct consequence of the recent success resulting from its application across a wide variety of scientific fields, including features design [1],[2] DNA analysis [3] etc., it has attracted significant interest from researchers in medical imaging techniques.

Recently, many deep architectures of feature learning algorithms have been proposed for segmentation [4],[5], classification [6], clustering [7], speech recognition [8], and so on. Deep architecture is a ladder or hierarchical formation of multiple layers, with every layer being a process by which individual parameters take the initiative to learn from the results of its preceding layer, which makes the feature learning concrete. Along with this, research of deep architecture can exponentially increase the knowledge and very successful frameworks for brain image classification. CNN's is one of the deep discriminative models and has recently become a popular medical image analysis choice. For instance, abnormal shapes or volumes of certain anatomical structures of the brain must be found in brain disorders like cancer or tumors.

The highest interesting challenges raised by deep architectures are learning the right features and how to compute its features representation, i.e., feature extraction. Deep CNNs are distinguished by utilizing numerous feature extraction stages to learn features automatically. Our work has addressed the research problem as image-level classification using brain MR images. Therefore, the analysis of these images is essential to detect disorders including Alzheimer's [9], Parkinson [10], and so on. However, manually classifying the brain is an expensive and time-consuming process on a larger scale. These drawbacks would have sparked considerable interest and enabled the systematic classification of brain MRI as soon as images are acquired. In medical imaging methods, significant clinical features characterize objects like cells of organs, including the brain, heart, kidneys, liver, and lungs. Therefore, features representation is essential for image classification [11]–[14]. We proposed a deep feature representations framework for extracting the prominent features of human brain MRI images through the proposed deep architectures.

In the past few decades, many proposed frameworks have highlighted the brain tumor analysis phases like classification, features detection and representation, features fusion, treatment plans, and outcomes predictions.

Chahal *et al.* [15] exhibit the current state-of-art concerning brain tumor detection approaches using MR images. The main goal is to aid researchers in medical image processing to derive the characteristics and identify different classification or segmentation models. Several deep learning approaches, namely auto-encoders, CNNs, deep neural networks (DNN) etc., are combined with other techniques to improve the recognition rate. Additionally, deep learning algorithms are self-learning feature representational capacity and enable good quantitative analysis.

Liu *et al.* [16] present a survey on deep learning applications to MRI images. Combining the manifold and different architectures can help the model enhance robustness and generalization on several categories of images by extracting many levels of semantic representations. More recently, applying deep learning algorithms to medical imaging has led to impressive performance enhancements in diagnosing and classifying complicated pathologies such as breast cancers, brain tumors, etc. For the most part, in medical image processing and computer vision, the exploitation of different experimental new ideas in CNN architectural design has changed the research direction. Ensemble feature learning is also one of the forthcoming research areas on CNNs.

Hasan *et al.* [17] proposed a method to enhance the recognition rate of magnetic resonance imaging brain scans classification from combined deep learning and modified gray level co-occurrence matrix features. The design of CNN architecture was selected by a trial-and-error method, which was used to estimate the optimal number of hyperparameters. As a result, the developed model performs as a novel feature extraction method. The experimental results showed that the dimension of the feature was set at 23, and the support vector machines (SVM) classifier was used to achieve the highest accuracy based on the collected brain MRI scans.

Mao *et al.* [18] proposed a unique feature representation approach for lung nodule image classification. It incorporates both global and local features. First, Superpixels divide a lung nodule image into local patches. Using an unsupervised deep autoencoder (DAE), these patches are reconstructed into a fixed length of local feature vectors. The global feature representation is the bag-of-visual-words model, with the visual vocabulary produced. The SoftMax algorithm is applied to lung nodule type classification. The experimental result shows the effectiveness of the feature representation approach based on different parameter settings.

Due to the complexity of medical image analysis, designing a set of definitive features is a natural routine for predicting the classification tasks. The detailed interpretation of medical images is generally a tedious and indefinite task. In the past few years, several challenges faced by deep architectures have been extracting the augmented feature set, ranging from minor to complicated abstractions, which can support learning complex problems. Interestingly, the domains of brain tumor analysis under detection or classification tasks are becoming challenging and still an open research problem due to the diversity of shapes, areas, and sizes of tumors.

From 2015 to 2020, the research studies in deep architectures are still moving towards those significant performance improvements. Deep architectures are composed of feature detector units stacked in layers. Lower layers recognize simple features and pass them on to higher-level layers, which detect more complex features, and there have been various techniques for deep network learning.

Liu *et al.* [19] proposed a fully connected CNN-based deep learning representation. It has strong global expression ability and contains high-level semantic information. It has been generated after layers of convolutions with the input image, which has a global receptive field. Zheng et al. [20] have contributed to the Convolutional representation study. It uses activations of convolutional layers, followed by a global-pooling method. This produces a compact image representation with dimensions equal to the number of feature mappings of the corresponding layer.

Yang *et al.* [21] address the learning of structured and non-redundant representations with deep neural networks using Structured Decorrelation Constraint (SDC). It is flexible and can be applied to different types of network layers, including 3D convolutional layers and 1D fully connected layers, which is beneficial for enhancing the network's performance. In order to identify how to design effective regularizes for multiple deep networks, we will do comprehensive research on the effects of various regularizes on different network topologies.

Athiwaratkun and Kang [22] introduced a feature component of Convolutional Neural Network (CNN) models and illustrated various applications of CNN features. They designed three CNN1, CNN2, and CNN3 architectures for the experiments. In addition, even if the CNNs are not optimum, because they are not fully trained, they can nevertheless extract useful features that provide comparable prediction accuracy when compared to more computationally expensive approaches such as CNN trained with Dropout or model averaging. In addition, it shows the practice of CNN features as input to other models (like SVM etc.,) compared to the original CNN. Le *et al.* [1] proposed a new method for automated learning of invariant features of tumor signatures. It is a two-layer network having nonlinear responses to discover the features from low-cost unlabeled data. The results of the experiments show that this method outperforms expert-designed representations.

This research looks into how deep learning (CNNs) models may be used to represent features as inputs to other classification models and investigates the effectiveness of high-level tasks with automated learning of features.

The contribution of this paper is summarized as follows:

- Design and develop a two-stream of deep architectures for extracting features of images.
- Create salient fused feature vectors retrieved by two deep convolutional networks of the first fully connected layers.
- To obtain the compact fusion features, principal component analysis (PCA) is adopted to reduce the dimension.
- Study the impact of compact fusion features by the classifier.

This paper is organized as follows. Section II presents the proposed framework to study the effectiveness of the feature representation. Section III demonstrates experimental setup results. Then, the conclusion is drawn in section IV

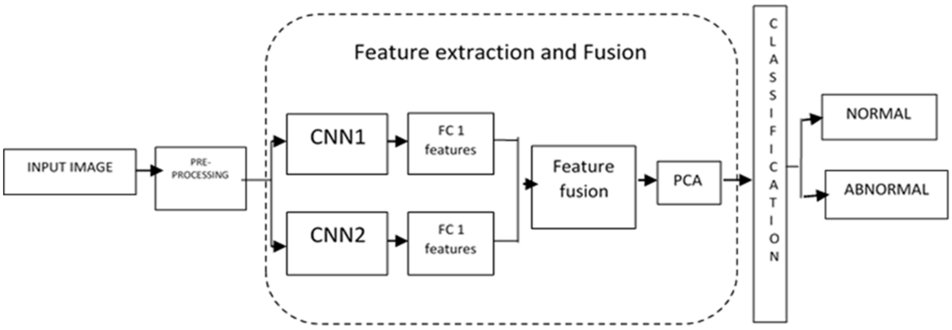


Fig. 1 Framework of the proposed Methodology

II. Material and Method

Feature representation is a vital building block of image classification. We formulate the problem as image-level classification. The proposed strategy is shown in Fig1. The proposed model comprises five stages. The first stage is pre-processing of the MR images using a unit-distribution transformation approach. The second stage explores the designed two-stream deep architecture (CNNs) frameworks for extracting the high-level features. The third stage is to construct fused features using serial strategy and fusion operators. The fourth stage is to apply the principal component analysis (PCA) to reduce the dimension of the deep fused features and yield a compact fusion feature vector. Finally, to obtain the image-level classification results by the classifier to predict whether an image is normal or abnormal.

A. Pre- Processing

The pre-processing process is generally designed to ensure the achievement of classification tasks. We applied intensity normalization to the input data using a unit-distribution transformation approach to assess brain image processing.

B. Deep Architecture

The Convolutional neural network (CNN) models are a redesign of traditional neural network (NN) layers that include convolution (C), pooling (P), and fully connected (FC) layers. As with ANN, it is primarily a hierarchical network structure, but the configuration and operation of the layer have been modified. The network model is subdivided into two parts: feature extraction and classification. The feature extraction part includes the convolution and pooling layers, whereas the fully connected layer was involved in the classification part. The Convolution layers are derived from mathematical operations consisting of filters/kernels, i.e., a 2D matrix of numbers. The filter is then convolved with the input to extract the output. The pooling layer plays a part in squeezing the spatial size of the feature representation. The fully connected (FC) layer gives the capability to map the features representation between the input and the output.

This research aims to provide an effective approach for automatically representing features based on hyperparameters and building CNN structures for a given specific classification task. The CNN architecture was encoded in this study as a number of hyperparameters, which are described in Table I, Where CNN represents the deep architecture, H_{conv} represents the convolutional layer parameters set, H_{pool} represents the parameters of pooling layer, H_{act} represents the activation function, and H_{fc} is the set of fully connected layer parameters.

TABLE I
HYPER PARAMETERS

Parameters	Description
CNN	{Hconv, Hpool, Hact ,Hfc }
Hconv	{C0, ..., Cn }
Ci	{kcount, ksize }*
Hpool	{Pi }
Hact	{Ai }

* COUNT: Number of filters SIZE: THE Size of filters

However, with CNNs, having more convolutional layers decreases the depth of feature maps at the network's end, resulting in a coarse classification. To address this issue, we expand the architecture of Fig.1 in the proposed framework by incorporating the CNN1 and CNN2 designed as features extractors described in Tables II and III. There are eight layers, ordered as I, C1, P1, C2, P2, C3, P3, and F1 in the sequence CNN1 model. Where I, C, P, F are denoted as input layers, convolutional layers, pooling layers, and fully connected layers. Similarly, eleven layers, I, C1, C2, P1, C3, C4, P2, C5, C7, P3, and F1, are sequences of the CNN2 model. The architecture design of CNN1 and CNN2 were optimized using a trial-and-error approach.

The effectiveness of CNN depends primarily on the nature of the designed network, and how the layers are stacked, and kernels/filters are set. Gradient back-propagation represents the main method for learning all types of neural networks. To design a CNN1& CNN2 for a specific task, it is necessary to understand the prerequisite to be met and how the input image slice is set up to the network. The size of every convolutional layer or a given input MRI slice can be calculated using equations (1) and (2), respectively.

$$\text{Convw}= ((\text{MRslicew}- \text{cfw}+ (2*p))/\text{sw}+1 \quad (1)$$

$$\text{Convh}= ((\text{MRsliceh}- \text{cfh}+ (2*p))/\text{sh}+1 \quad (2)$$

Where cf represents the convolution filter, p is the number of zero padding and s denotes number of strides.

TABLE II
ARCHITECTURE OF CNN1

Layer Name	Type of layer	Kernel size	Feature Map/shape
I	Input	-	240x240x1
C1	Conv1+ ReLU	x5 , 32 filters	240x240x32
P1	Max- Pooling	x2, stride 1	120x120x32
C2	Conv2+ ReLU	x5 , 48 filters	120x120x48
P2	Max- Pooling	2x2, stride 1	60x60x32
C3	Conv3+ ReLU	x5 , 64 filters	60x60x64
P3	Max- Pooling	2x2, stride 1	30x30x64
F1	Fully Connected (FC1)	1x384	1x384

TABLE III
ARCHITECTURE OF CNN2

Layer Name	Type of layer	Kernel size	Feature Map/shape
I	Input	-	240x240x1
C1	Conv1+ ReLU	5x5 , 32 filters	240x240x32
C2	Conv2+ ReLU	5x5 , 32 filters	240x240x32
P1	Max- Pooling	2x2, stride 1	120x120x32
C3	Conv3+ ReLU	5x5 , 48 filters	120x120x48
C4	Conv4+ ReLU	5x5 , 48 filters	120x120x48
P2	Max- Pooling	2x2, stride 1	60x60x48
C5	Conv5+ ReLU	5x5 , 64 filters	60x60x64
C7	Conv3+ ReLU	5x5 , 64 filters	60x60x64
P3	Max- Pooling	2x2, stride 1	30x30x64
F1	Fully Connected (FC1)	1x384	1x384

C. Features Fusion Modules

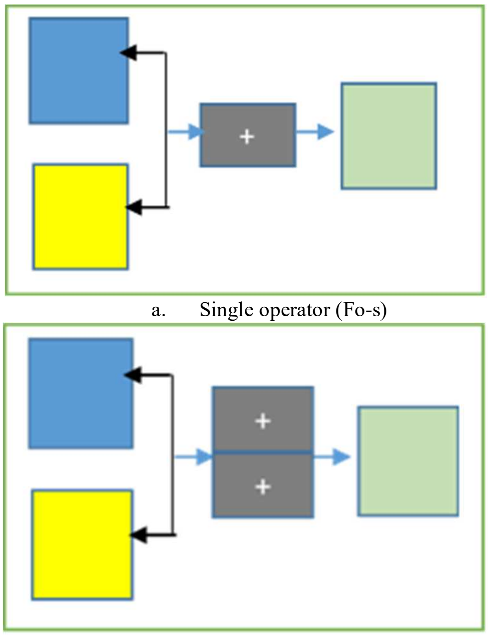
Feature fusion builds to learn image features completely and enhance their rich internal information. We use the proposed CNN1 and CNN2 models to extract features from the specified layer. The fused feature vectors, which include a rich information collection, can contribute to the classification process. As a result, fusing two distinct sets of characteristics is a critical challenge. We investigate the features fusion modules in four different ways. First, we employ a serial feature fusion technique [23], which concatenates the two feature sets. The dimension of the fused features is the sum of the two feature sets' dimensions. The output of the fusion feature can be written as:

$$\text{Xconcatfusion}= [\text{CNN1\_FC1, CNN2\_FC1}] \quad (3)$$

Second, summation of two sets of features, dimension of the fused features is equal to the source of the dimensions of the two sets of features

$$\text{Xsumfusion}= [\text{CNN1\_FC1} + \text{CNN2\_FC1}] \quad (4)$$

The final vectors $X_{concatfusion}$ and $X_{sumfusion}$ are input to the PCA to reduce the features, then these features were used as a classifier to produce the output. Third and fourth, we introduce fusion operators, namely single and multi-operator [24]. Fig.2a and Fig.2b show the detailed structures of the feature fusion operators. In particular, an input image Z is translated into 2 feature spaces using the CNN1-FC1 and the CNN2-FC1 feature extractors, represented as Ecnn1 and Ecnn2 with the features Ecnn1(z), Ecnn2(z) € ZFo Then, ZFo embeds Ecnn1(x) and Ecnn2(x) in a fusion feature space.



a. Single operator (Fo-s)

b. Multi operator (Fo-m)

Fig. 2 Fusion operators.

The single operator (Fo-s)) Using a learned weight scalar, compute the weighted sum between the CNN1-FC1 and CNN2-FC1 feature maps.

$$\text{Fo-s} = \beta \text{Ecnn1 (Z)} + (1\text{-}\beta) \text{ Ecnn2 (Z)} \quad (5)$$

CNN1-FC1 and CNN2-FC1 features are scaled from β and (1-β) respectively and then merged elementwise, as illustrated in figure2a. The multi-operator (Fo-m) computes the weighted sum of CNN1-FC1 and CNN2-FC1 feature maps using a learned weight vector β.

$$\text{Fo-m} = \beta \text{ Ecnn1 (Z)} + (1\text{-}\beta) \text{ Ecnn2 (Z)} \quad (6)$$

The weight vector β is first broadcasted to the feature shape and then multiplied by the features elementwise, as shown in figure 2b.

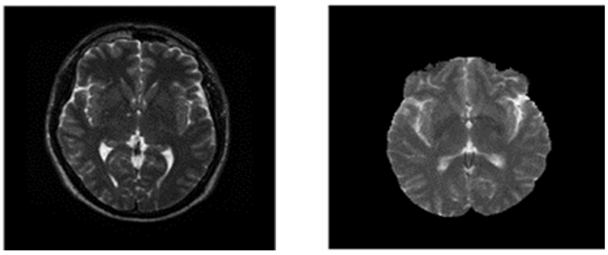
D. Dimensionality Reduction and Classification

The principal component analysis (PCA) is a technique and useful for classification tasks. After feature fusion, there exist higher-dimensional features and increased computation complexity. Consequently, to minimize dimensionality, we utilized principal component analysis (PCA). We adopted the support vector machines (SVM) classifier for classification in this work. The fusion features, whose dimension is reduced using PCA, are utilized for training a binary classification classifier. In addition, after comparing RBF and polynomial kernel functions, we choose the RBF kernel function with the best classification results. We extract the score feature from each image using a trained SVM model. Based on the score of the model has predicted the outputs.

III. Results and Discussion

A. Dataset

The proposed method applied on both the IXI and TCIA dataset of human brain T2-weighted MRI images. The IXI dataset images were collected from computational analysis of brain development website (https://brain-development.org), and TCIA data were gathered from the cancer imaging archive website (https://cancerimagingarchive.net). The IXI dataset involves MR images from normal, healthy subjects, whereas the TCIA MR images from abnormal, unhealthy subjects. All images were selected from 20 subjects. The MR slices are used in this work were acquired from Philips 3T / 1.5 T system. Each of the volumes studied is made up of 5 or 6 separate slices. The 200 image slices (100 abnormal slices and 100 normal slices) of subject volumes were considered in this work. Fig 3a and 3b are shown as a normal and an abnormal T2 –weighted MR brain image, respectively.



a. b.

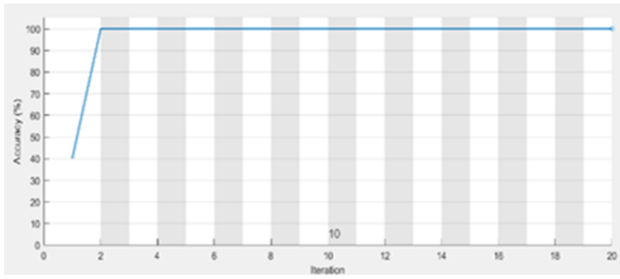
Fig. 3 Axial T2- weighted MR brain images: (a) Normal brain; (https://brain-development.org) (b) Abnormal brain (https://cancerimagingarchive.net).

B. Experimental setup

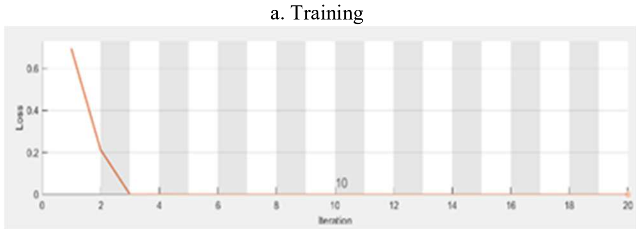
In this study, we attempted two- deep network experimental setups to evaluate the efficiency of the proposed model. These networks were optimized based on a trial-and-error method. In both networks, the fully connected (fc1) layer features are denoted as in-depth features. The evaluation of the proposed method was carried out in three different sets of experiments. First, we report all the experiments based on each network. Then, experiments with the fused features. Finally, we performed experiments on PCA. Moreover, experimental results reported the impact of the fusion pattern on classification strategy.

C. Results

1) Feature Extraction: For our experiments, we use two CNN architectures, namely CNN1 and CNN2, as feature extractors. The detailed description of these networks is shown in section II. The configuration of the network's architecture is a difficult and problem-specific task, which can significantly impact the performance of the model. We involved two distinct deep CNN architectures in this research. After developing CNN models, each model complies with the Adaptive Moment Estimation optimizer, convergence is faster, and results of the model training properties are shown in the Table II and III. To fit the training data with the model, the batch size of 128 and 20 epochs is used. Figs 4a and 4b illustrate the training set's loss and accuracy. These two architectures have one fully connected layer (FC1), composed of 384 deep CNN features, followed by a final classification, which outputs whether an image is normal or abnormal. In this research, the last fully connected layer is a feature representation of images and is considered a feature fusion operation. This is not a conventional approach in the literature. We are interested to see if the higher-level layer features fuses with fusion domains are more suitable for classification. Fig 5 shows the illustrations of the CNN1and CNN2 of the first 30 features representation maps.



a. Training



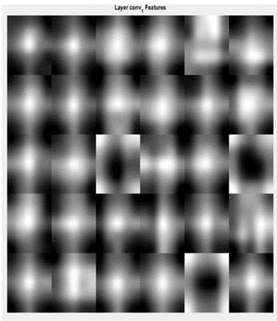
b. loss

Fig.4 Illustrates the (a) accuracy and (b) loss of the training set.

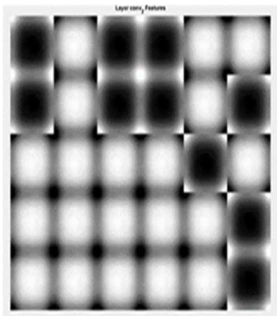
2) The CNN Features for classification: we trained the proposed network configuration with a pre-processed data sample (N=200), yielding different trained features, resulting in two differently trained networks. Unless stated otherwise, we use the first fully connected layer of the CNN1 and CNN2 networks as our feature vectors for all experiments. We have two settings: First, the 384-dimensional feature vectors combined with fusion operators with a SVM to solve the classification task. Second, we further add dimensionality reduction by PCA and report the individual results. To further study this, we trained an RBF and polynomial SVM for all classes using the output of each features fusion module.

TABLE IV
SVM Classifier Performance

Classifier	Features Fusion Modules	Accuracy (%) Dimension of features with PCA						
		50	100	150	200	250	300	350
RBF SVM	Concat	64.0	61.0	59.5	58.0	58.0	58.0	58.0
	Sum	60.5	60.5	58.5	61.5	57.5	56.5	55.5
	Single	64.0	63.0	62.0	60.5	54.0	53.5	53.0
	Multi	60.0	56.5	56.5	55.5	55.0	54.0	54.0
Polynomial SVM	Concat	50.0	50.0	50.0	50.0	50.0	50.0	50.0
	Sum	50.0	50.0	50.0	50.0	50.0	50.0	50.0
	Single	50.0	50.0	50.0	50.0	50.0	50.0	50.0
	Multi	50.0	50.0	50.0	50.0	50.0	50.0	50.0



a. CNN1 features map



b. CNN2 features map

Fig. 5 Illustrates the first layer features representation maps.

3) Effectiveness of Feature Fusion: In the experiments, we investigate the fusion modules in four different ways, namely, concat, sum, single and multi-operators strategy. Fig.6 demonstrates the performance analysis by SVM and PCA with different dimensions. Using the CNN1 and CNN2 off-the-shelf representation with RBF support vector machine training significantly outperforms 50 features. We use the CNN1 and CNN2 architectures as a baseline. This is shown as a comparison of each classification model to see how they perform with features from CNN1 and CNN2. Table IV shows the proposed approach's performance by using the most common kernel functions, including polynomial and RBF. The proposed technique is evaluated using performance metrics, namely accuracy, sensitivity, recall, and F-measure. The performance metrics are examined based on the confusion matrix. This can be described by the terms as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) as follows:

- True Positive (TP) - abnormal class images are predicted correctly from the model,
- True Negative (TN) – normal class images are predicted correctly from the model,
- False Positive (FP) - normal class images are predicted incorrectly as an inclusion to the abnormal class from the model, and
- False Negative (FN) - abnormal class images are predicted incorrectly as an inclusion to the normal class from the model.

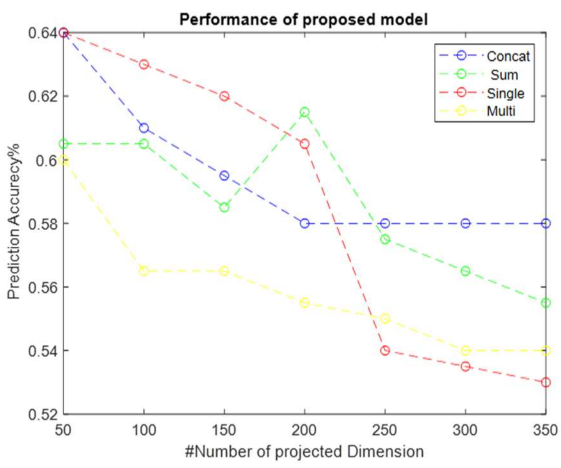
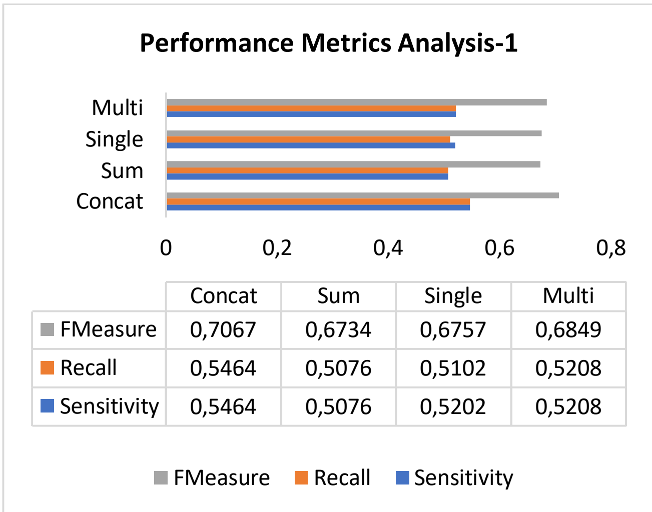


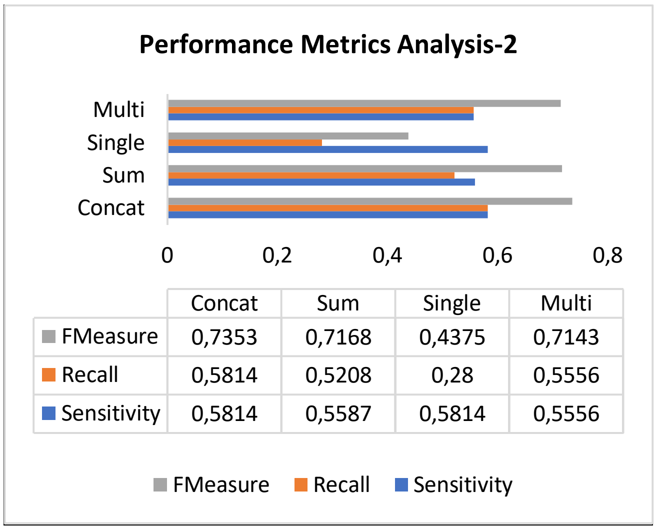
Fig. 6 Performance of SVM with various Dimensions.

The Comparative model performance analysis-1 and analysis-2 of the proposed approach are shown in Fig 7. In fact, classification accuracy varies from 50.00 % to 64.00 %, with polynomial and radial basis function based on the selected features fusion dimensions after PCA. Table V shows that concatenation and single fusion methods have better accuracy (64.0%) compared to the original models, 50.00% for CNN1 and 52.55% for CNN2.

Finally, we apply our proposed method in two ways to test the classification accuracy for the entire test data set. First, 25% of the input images are for training, and 75% of the input images are for the test data set. Second, 50% of the input images are used to create each test and training data set. Further, SVM classifier achieved a classification accuracy of fusion approaches better than that of approaches, which use only CNN1 features or only CNN2 features. The best results obtained on different methods are tabulated in Table V. The experimental results demonstrated that fusion feature learning is superior.



	Concat	Sum	Single	Multi
FMeasure	0,7067	0,6734	0,6757	0,6849
Recall	0,5464	0,5076	0,5102	0,5208
Sensitivity	0,5464	0,5076	0,5202	0,5208



	Concat	Sum	Single	Multi
FMeasure	0,7353	0,7168	0,4375	0,7143
Recall	0,5814	0,5208	0,28	0,5556
Sensitivity	0,5814	0,5587	0,5814	0,5556

Fig. 7 Proposed model performance analysis.

In order to compare with state-of-the-art methods, we selected a reference paper [22], which had designed convolutional neural network (CNNs) models. We have used both CNN1 and CNN2 models in our proposed work. In addition, fusion strategies [23],[24] were also added and made an experimental setup. As a result, experiments carried out by feature-level fusions yield up to 64.00% (see table IV and V), which is competitive results without using many additional computational resources. The comparison is tabulated in Table VI.

TABLE V
COMPARISON OF BEST PERFORMANCES

Methods	#Feature	Acc (%)
CNN1_SVM	384	50.00
CNN2_SVM	384	52.55
Fusion: concat	768	58.50
Fusion: sum	384	51.50
Fusion: Single	384	55.00
Fusion: Multi	384	56.50
Concat_Fusion_PCA	50	64.00
Sum_Fusion_PCA	50	60.50
Single_Fusion_PCA	50	64.00
Multi_Fusion_PCA	50	60.00

TABLE VI
STATE-OF-THE-ART-METHOD COMPARISON

Method	Accuracy (%)
Proposed: Concat Fusion	64.00%
Proposed: Single Fusion	64.00%
Ref [22] CNN1 model	52.63%
Ref [22] CNN2 model	55.97%
Ref [22] CNN3 model	57.96%

IV. CONCLUSION

This paper proposes a technique for automatically classifying MRI slices as normal or abnormal with deep features extraction through designed discriminative CNN models and feature level fusion strategy. This approach combines CNN1 and CNN2 models used for feature extraction, and then we employed serial fusion and fusion operator's strategies. In addition to that, the principal component analysis (PCA) for diminishing the features and classifying MR images, the Support Vector Machine binary classifier, has been used. The experimental results provide a good classification accuracy of 64.00% by utilizing only as low as 50 features for the classifier input. Although the approach was created for only axial T2-weighted images, the same method can reasonably be applied to other types of MR images in the future.

REFERENCES

[1] Q. V. Le, J. Han, J. W. Gray, P. T. Spellman, A. Borowsky, and B. Parvin, "Learning invariant features of tumor signatures," *Proc. - Int. Symp. Biomed. Imaging*, pp. 302–305, 2012, doi: 10.1109/ISBI.2012.6235544.

[2] Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, and E. I. C. Chang, "Deep learning of feature representation with multiple instance learning for medical image analysis," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, no. 1, pp. 1626–1630, 2014, doi: 10.1109/ICASSP.2014.6853873.

[3] H. Y. Xiong *et al.*, "The human splicing code reveals new insights into the genetic determinants of disease," *Science (80-.).*, vol. 347, no. 6218, 2015, doi: 10.1126/science.1254806.

[4] M. M. Thaha, K. P. M. Kumar, B. S. Murugan, S. Dhanasekeran, P. Vijayakarthick, and A. S. Selvi, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *J. Med. Syst.*, vol. 43, no. 9, 2019, doi: 10.1007/s10916-019-1416-0.

[5] J. P. Horwath, D. N. Zakharov, R. Mégret, and E. A. Stach, "Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images," *npj Comput. Mater.*, vol. 6, no. 1, pp. 1–9, 2020, doi: 10.1038/s41524-020-00363-x.

[6] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Classification using deep learning neural networks for brain tumors," *Futur. Comput. Informatics J.*, vol. 3, no. 1, pp. 68–71, 2018, doi: 10.1016/j.fcij.2017.12.001.

[7] P. Dahal, "Learning Embedding Space for Clustering from Deep Representations," *Proc. - 2018 IEEE Int. Conf. Big Data, Big Data 2018*, no. April, pp. 3747–3755, 2019, doi: 10.1109/BigData.2018.8622629.

[8] L. M. Q. De Santana, R. M. Santos, L. N. Matos, and H. T. Macedo, "Deep Neural Networks for Acoustic Modeling in the Presence of Noise," *IEEE Lat. Am. Trans.*, vol. 16, no. 3, pp. 918–925, 2018, doi: 10.1109/TLA.2018.8358674.

[9] S. Spasov, L. Passamonti, A. Duggento, P. Liò, and N. Toschi, "A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease," *Neuroimage*, vol. 189, no. January, pp. 276–287, 2019, doi: 10.1016/j.neuroimage.2019.01.031.

[10] W. Wang, J. Lee, F. Harrou, and Y. Sun, "Early Detection of Parkinson's Disease Using Deep Learning and Machine Learning," *IEEE Access*, vol. 8, pp. 147635–147646, 2020, doi: 10.1109/ACCESS.2020.3016062.

[11] Y. Huang, Z. Wu, L. Wang, S. Member, and T. Tan, "(Pattern Analysis and Machine Intelligence, IEEE Transactions on 2013) Feature Coding in Image Classification A Comprehensive Study.pdf," pp. 1–15, 2013.

[12] N. Wahab, A. Khan, and Y. S. Lee, "Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathological images," *Microscopy*, vol. 68, no. 3, pp. 216–233, 2019, doi: 10.1093/jmicro/dfz002.

[13] S. Kido, Y. Hirano, and N. Hashimoto, "Detection and classification of lung abnormalities by use of convolutional neural network (CNN) and regions with CNN features (R-CNN)," *2018 Int. Work. Adv. Image Technol. IWAIT 2018*, pp. 1–4, 2018, doi: 10.1109/IWAIT.2018.8369798.

[14] C. Sun, A. Xu, D. Liu, Z. Xiong, F. Zhao, and W. Ding, "Deep Learning-Based Classification of Liver Cancer Histopathology Images Using Only Global Labels," *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 6, pp. 1643–1651, 2020, doi: 10.1109/JBHI.2019.2949837.

[15] P. K. Chahal, S. Pandey, and S. Goel, "A survey on brain tumor detection techniques for MR images," *Multimed. Tools Appl.*, vol. 79, no. 29–30, pp. 21771–21814, 2020, doi: 10.1007/s11042-020-08898-3.

[16] J. Liu *et al.*, "Applications of deep learning to MRI Images: A survey," *Big Data Min. Anal.*, vol. 1, no. 1, pp. 1–18, 2018, doi: 10.26599/BDMA.2018.9020001.

[17] A. M. Hasan, H. A. Jalab, F. Meziane, H. Kahtan, and A. S. Al-Ahmad, "Combining Deep and Handcrafted Image Features for MRI Brain Scan Classification," *IEEE Access*, vol. 7, pp. 79959–79967, 2019, doi: 10.1109/ACCESS.2019.2922691.

[18] K. Mao, R. Tang, X. Wang, W. Zhang, and H. Wu, "Feature Representation Using Deep Autoencoder for Lung Nodule Image Classification," *Complexity*, vol. 2018, no. d, 2018, doi: 10.1155/2018/3078374.

[19] Y. Liu, F. Nie, Q. Gao, X. Gao, J. Han, and L. Shao, "Flexible unsupervised feature extraction for image classification," *Neural Networks*, vol. 115, pp. 65–71, 2019, doi: 10.1016/j.neunet.2019.03.008.

[20] L. Zheng, Y. Yang, and Q. Tian, "SIFT Meets CNN: A Decade Survey of Instance Retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 5, pp. 1224–1244, 2018, doi: 10.1109/TPAMI.2017.2709749.

[21] J. Yang, W. Xiong, S. Li, and C. Xu, "Learning structured and non-redundant representations with deep neural networks," *Pattern Recognit.*, vol. 86, pp. 224–235, 2019, doi: 10.1016/j.patcog.2018.08.017.

[22] B. Athiwaratkun and K. Kang, "Feature Representation in Convolutional Neural Networks," pp. 6–11, 2015.

[23] J. Yang, J. Y. Yang, D. Zhang, and J. F. Lu, "Feature fusion: Parallel strategy vs. serial strategy," *Pattern Recognit.*, vol. 36, no. 6, pp. 1369–1381, 2003, doi: 10.1016/S0031-3203(02)00262-5.

[24] X. Yao, T. Huang, C. Wu, R.-X. Zhang, and L. Sun, "Federated Learning with Additional Mechanisms on Clients to Reduce Communication Costs," pp. 1–12, 2019.