A Probabilistic Neural Network Model for Brain Tumor Segmentation and Detection

Priyanka Kotwar[1] Prof. Priyanka Choudhary[2]

VITM, Indore, India

ABSTRACT

Artificial Intelligence (AI) and related computational tools are making their presence felt in various walks of life. However, AI in healthcare is grabbing more attention in latest trends of research with improved accuracy of I based techniques for different health related problems. This paper presents a technique for brain tumor classification using an amalgamation of image processing techniques and artificial intelligence. Brain tumors are often very difficult to classify into malignant and benign categories owing to the high level of similarity among the two categories of images. The proposed technique uses the discrete wavelet transform along with threshold based segmentation for separation and de-noising of brain tumor images. Further, feature extraction is performed followed by training a probabilistic neural network with the computed feature values. Principal component analysis is used for reduction of the dimensionality of the training data. It has been found that the proposed technique achieves a classification accuracy of 97.31% for the used data set. It is expected that the proposed approach can be useful for effective automated classification of brain tumor images.

Keywords

Artificial Intelligence, Brain tumor classification, threshold based segmentation, discrete wavelet transform (DWT), Principal Component Analysis (PCA), Probabilistic neural network (PNN).

1. INTRODUCTION

Brain tumor classification has been a non-trivial task for medical practitioners owing to the fact that cancerous and non-cancerous brain tumors show stark similarity in appearance. [Dahab (2012)]Hence, chances for inadvertent error on the part of physicians can't be ruled out. Off late, artificial intelligence based techniques have made their presence felt in different domains of healthcare. Artificial Intelligence can be formally defined as the design and implementation of computational techniques which can perform tasks generally needing human intervention thereby resembling the human intelligence mechanism. A practical way to implement artificial intelligence is the design of artificial neural network (ANN). Artificial neural network resemble the human intelligence process of:

(1) Parallel data processing capability

(2) Ability to learn and adapt

Artificial neural networks can often perform tasks needing human effort. Neural networks need to be trained with training data and subsequently can be tested for some task. In the proposed approach, artificial neural networks have been used for the classification of brain tumors into three categories v.i.z. normal, benign and malignant. The data used is in the form of MRI images of brain. [Hua(2012)] Prior to extracting numerical features from the images, different image pre-processing techniques have been used. The description of the neural network design and image pre-processing techniques has been explained in the subsequent section.

2. ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networks can be thought of as a practical technique to implement artificial intelligence. Artificial Neural Networks is kind of a system that makes the machine work intelligently and performs tasks that is required. The ANN mainly focuses on building the self learning or self training mechanism like humans. [Mala (2015)] .The human brain comprises of millions of neurons that are interconnected to each other. This helps the brain to perform complex tasks and think well attributing to the large network of interconnected neurons. But it is not the same with the ANN. An ANN has the ability of saving the previous data that was input to it. ANN works in a very different manner to come out as intelligent. It functions in a complex parallel process and works parallel. It exhibits

high performance of speed due to the parallel architecture. ANN tries to learn and adapt to the data that was input before. This is done through training and testing of the Neural Network. This is an extremely important part as the entire performance of the ANN depends on how well the neural network has been trained and tested. And also the amount and quality of input training dataset that is sued plays a pivotal role. The basic ANN topologies consists of Single layered Feed Forward networks, multilayered feed forward networks, re current networks and probabilistic neural networks. The main internal structure of ANN includes a Input Layer, Hidden Layer and an Output Layer. The mathematical model of the ANN is governed by the expression:

$$y = \sum_{i=1}^{n} X_i . W_i \ + \ \theta_i$$

(1)

Here,

X represents the parallel inputs

Y represents the output of the ANN

θ represents the bias

W represents the weights associated with the inputs received the ANN.

Here, a parallel stream of data X enters the neural network. The summation of the data streams take place at the neural network interface. The experiences created by the network based on the received inputs are designated as the weights of the ANN.

2.1 Probabilistic Neural Network (PNN)

PNN is built on the theory of Bayesian network and the estimation of probability density

function. This theory allows for cost function to represent the fact that it may be worse to misclassify a vector that is actually a member of class A but it's classified as a vector that belongs to class B. PNN is extensively used for classification problem. [Kulhalli (2012)] It works as a supervised classifier. A sample of our data set to it as an input and the data is then handled through numerous layers. The working of the PNN is based on the Baye's theorem of conditional probability given by:

$$P(A/B)=(P(B/A)P(A))/(P(B))$$

(2)

Here,

P(A) represents probability of any event A

P(B) represents probability of any event B

P(A/B) represents the probability of event A if event B is true or has already occurred

P(B/A) represents the probability of event B if event A is true or has already occurred.

The PPN assumes the training data to be true and then evaluates the probability of a given testing sample to belong to a particular category. The category with the highest probability is adjudged a class. [Sridhar (2013)] The PNN serves as an effective classifier if the deciding boundary between the different classes is not clear and a probabilistic approach is more suitable for the classification.

3. IMAGE PRE-PROCESSING AND FEATURE EXTRACTION

Prior to training the neural network, it is necessary to pre-process the raw data and then compute the features. The different steps executed are:

3.1 RGB to Grayscale Conversion

The first step is RGB to grayscale conversion implemented as:

The conversion of the RGB the RGB to the gray scale image is given by:

$$I_y = 0.333 f_r + 0.5 f_g + 0.1666 f_b$$

(3)

Here,

Fr is the red component of the pixel

Fg is the green component of the pixel

Fb is the blue component of the pixel

The expression evaluating the grayscale value of the image pixel is given by:

$$I_y = 0.333*190 + 0.5*183 + 0.1666*175 \quad = 183.925$$

(4)

The above relation is used for the scaling of RGB images into the grayscale counterpart. The expression yields a simple value for each pixel value.

3.2 Segmentation

The segmentation is an extremely critical step in the final classification process since the segmentation process separates the tumor segment from the entire image. [Manzes(2015)]The various types of segmentation techniques are explained as under:

(1) Thresholding: In this technique the segmentation is carried out based on the following condition:

$$X_i \epsilon C_i \ if \ val(X_i) \geq T$$
$$else, X_i \epsilon \ C_2$$

(5)

Here,

Xi represents the data value 'i'

T represents the segmentation threshold

C1 represents category1

C2 represents category2

(2) Edge Based Segmentation: In this category of segmentation, regular shapes such as lines or curves are used to segment out the data. The regular shapes generally tend to have a clear boundary.

(3) Region Based Segmentation: In this technique, the segmentation is done based on the difference in pixel values of different areas. The process works only when the difference in metrics of different regions is bear stark differences.

In the proposed work, threshold based segmentation is used since in MRI images, there may not exist clear boundaries which yield a regular shape segmentation.

3.3 Discrete Wavelet Transform

Often, MRI images are corrupted by effects of noise thereby rendering inaccurate feature values. The discrete wavelet transform acts as an effective filter to remove the effects of noise and disturbance from the images. [Hasan (2016)] The discrete wavelet transform can be thought of as a sampled version of the continuous wavelet transform and tries to remove the local disturbances using down samplings.

The scaling feature can be defined as:

$$W\Phi \ (Jo, \ ok) \ = \ \psi(X\{x,t\})$$

(6)

The Wavelet feature may be defined as:

$$W \ = \ F_N(\psi)$$

(7)

Where,

F_N is Normalizing term

x is the space variable

t is the time variable

ψ is the transform

3.4 Feature Extraction

Since artificial neural networks can accept and process numerical data, hence it becomes mandatory to compute mathematical features of parameters which yield differences the different data categories. [Popescu (2014)] The computed features in this case are contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, Entropy, RMS value, variance, smoothness, kurtosis and skewness. The features are necessary to train a neural network as a neural network doesn't recognize image or non-numerical data sets. [Kolge(2012)]

3.5 The Principal Component Analysis (PCA)

The principal component analysis (PCA) is especially useful as it tries to remove the redundancies of a given dataset. It maps a redundant correlated data set (X) to an uncorrelated data set (Y) with significant

dimensional reduction. It is often useful for effective training and reducing the time complexity of the system.

3.6 Features Extraction

Since artificial neural networks can accept and process numerical data, hence it becomes mandatory to compute mathematical features of parameters which yield differences the different data categories. [Popescu (2014)] The computed features in this case are contrast, Correlation, Energy.

4. PROPOSED METHODOLOGY

The proposed methodology can be explained using the following steps:

(1) Extract data set. The data set used in the proposed approach is the mathworks.com/brain tumor data set.

(2) Divide the data set in the ratio of 70:30 as training and testing sets.

(3) Apply RGB to gray scale conversion as per equation. 3

(4) Apply threshold based segmentation as per

$$X_i \epsilon\ R1\ if\ val(X_i) \geq 70\ pd$$

else

$$X_i \epsilon\ R2$$

(8)

Here,

X is the image sample

R1 is region 1

R2 is region 2

70 pd is the pixel difference

(5) Apply DWT to the segmented image to remove the noise effects in the wavelet domain. Here, the image is transformed to the time frequency domain using the wavelet transform. Mathematically:

$$X(z) \leftarrow T\ \{x(t)\}$$

(9)

Here,

T is the transform

x(t) is the data in one domain

X(z) is the signal in the transform domain

T is the transform operator

(6) Apply principal component analysis to remove correlation redundancy and convert the data to an uncorrelated data set with reduced dimensionality. Mathematically,

$$Y(z) \leftarrow X(x)$$

(10)

Here,

Y(z) is the uncorrelated data set.

(7) Apply the data to the Neural Network for training.

(8) Repeat steps 1-7 for testing

(9) Compute accuracy of classification.

Due to the size and complexity of the data, neural network design to address the problem becomes almost invariable. In such a case, a neural network model is needed which can address the classification problem with high accuracy. The mathematical model of the neural network is shown in figure 1.
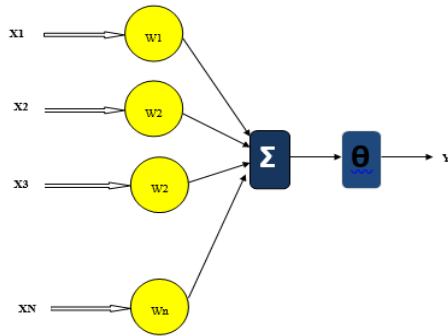
Fig.1 Mathematical Model of ANN

Fig.2 Ada-Boost Architecture

Here,

X represents the parallel input data stream to the ANN
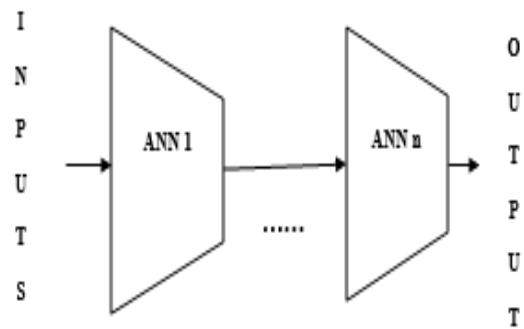
Y represents the output of the ANN

W represents the dynamic weights of the system

θ represents the logic for analysis or bias of the network [9].

Next, for the purpose of image forgery detection, two critical aspects need to be understood.

(i)   Neural networks need to be fed with numerical features which have to be extracted from the raw data set.

(ii) The network should be able to train or learn from the highly uncorrelated data

(iii) The neural classifier needs to have a holistic approach in regards to classification.

For this purpose, the cascade ada-boost neural architecture is chosen. It is depicted in the figure below:

It can be seen from the above figure that the neural networks are connected in cascade and the output of one neural network is fed as the input to the other neural network. Such a strategy is employed for complex classification problems, where the data format is different from the extracted feature values.

It can be clearly seen that the proposed approach uses two ANN architectures for the final classification. The first one is the scaled conjugate gradient (SCG) and the other is the Probabilistic Neural Network (PNN). The essence of the SCG is the fact that it uses least space complexity for implementation and is also fast. The PNN is a very effective probabilistic classifier. They are discussed subsequently. The weights are updated as follows

$$x_{k+1} = x_k + \alpha_k g_k \qquad (11)$$

Where, α_k is the determined step size and

$$P_k = -g_k + \beta_k P_{k-1} \qquad (12)$$

Here p is search direction vector and g is gradient direction vector. For SCG, β_k factor calculation and direction of the new search can be shown as in following equations.

$$\beta_K = \frac{(|g_{k+1}|^2 - g_{k+1}{}^T g_k)}{g_k{}^T g_k} \qquad (13)$$

$$P_{k+1} = -g_{k+1} + \beta_k P_k \qquad (14)$$

The probabilistic neural network works on the principle of the Baye's Theorem pertaining to conditional probability of associated random events. The theorem is defined as:

$$P\left(\frac{A}{B}\right) = \frac{P\left(\frac{B}{A}\right) P(A)}{P(B)} \qquad (15)$$

Here:

P(A) is the individual probability of event A

P(B) is the individual probability of event B

P(A/B) is the probability of A given B is true

P(B/A) is the probability of B given A is true

The probabilistic neural network is specifically used for applications where there is a sense of uncertainty in computing the output. The PNN computes the cost function of the event A with respect to event B.

$$P_A C_A f_A > P_B C_B f_B \qquad (16)$$

Here,

P_A represents probability of A

P_B represents probability of B

C_A represents the cost function for event A

C_B represents the cost function for event B

F_A represents the probability function for event A

F_B represents the probability function for event B

The PNN is used for the final classification.

5. RESULTS

The results have been simulated on MATLAB 2020a. The sequential execution of the steps in the proposed algorithm yields the following results. The images are segmented and classified in 3 categories which are:

(1) Normal

(2) Benign

(3) Malignant

A sequential description of the obtained results has been shown in the figures below.
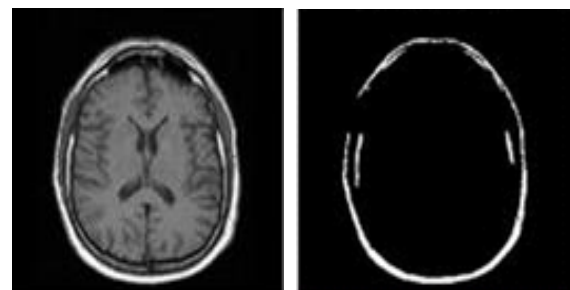
1(a)   1(b)

Fig.1(a) Original Image and Fig.1 (b) Segmented Version of Normal Image

The normal case image will exhibit not separate region or lesion for the tumor part while the benign and malignant will exhibit the tumor part. The tumor part which is the actual affected area of the image is separated out using the

thresholding technique. The 12 feature values are then shown below.

Table 1. Feature Values for Normal Image

	MRI type	Normal
1	Contrast	0.423863636
2	Correlation	0.080305851
3	Energy	0.702853822
4	Homogeneity	0.912159091
5	Mean	0.003935541
6	Standard Deviation	0.106556652
7	Entropy	2.743369367
8	RMS	0.106600358
9	Variance	0.011361868
10	Smoothness	0.879123104
11	Kurtosis	6.961505164
12	Skewness	0.711892665

The feature extraction of the normal images helps to separate the normal images from the images with tumors. However, the images with tumors can be sub-divided into two categories:

1) Benign (non cancerous tumors)

2) Malignant (cancerous tumors)

The following section illustrates the segmentation and feature extraction of benign and malignant tumor images. It should also be noted that the benign and malignant images have an identical visual representation and classifying them based on observation is nearly impossible. Hence feature extraction and classification using the PNN is necessary.


2(a)   2 (b)
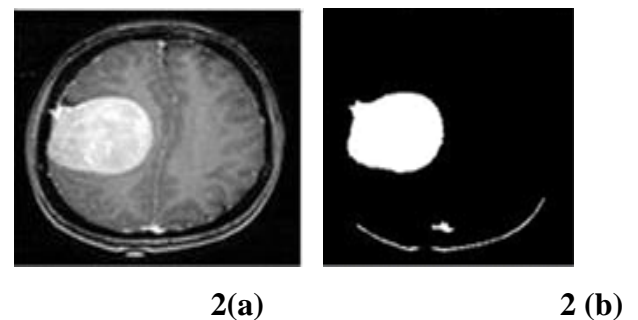
Fig.2(a) Original Image and Fig.1 (b) Segmented Version of Benign Image

Table 2. Feature Values for Benign Image

	MRI type	Benign
1	Contrast	0.376136364
2	Correlation	0.114973113
3	Energy	0.704254261
4	Homogeneity	0.914289773
5	Mean	0.002934793
6	Standard Deviation	0.106588795
7	Entropy	3.202367414
8	RMS	0.106600358
9	Variance	0.011340392
10	Smoothness	0.844321551
11	Kurtosis	6.766791088
12	Skewness	0.574506554

Table 3. Feature Values for Malignant Image

|	MRI type	Malignant
1	Contrast	0.378409091
2	Correlation	0.156222476
3	Energy	0.663764205
4	Homogeneity	0.902547348
5	Mean	0.007867343
6	Standard Deviation	0.106338423
7	Entropy	3.536728304
8	RMS	0.106600358
9	Variance	0.011233896
10	Smoothness	0.935645171
11	Kurtosis	7.950905337
12	Skewness	0.726189966
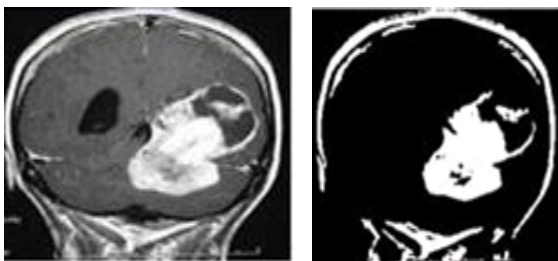
3 (a)   **3 (b)**

Fig.3 Original Image and Fig.1 (b) Segmented Version of Malignant Image
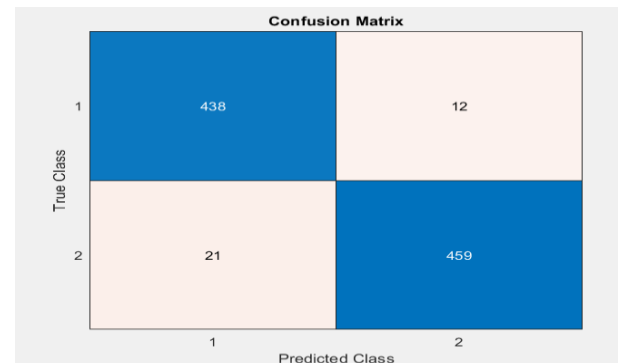
Fig.4 Confusion Matrix

It can be observed that the feature values vary very less for the different cases and hence a probabilistic approach for computing the class of data is necessary which is performed by the probabilistic neural network. The accuracy is computed according the following mathematical equation.

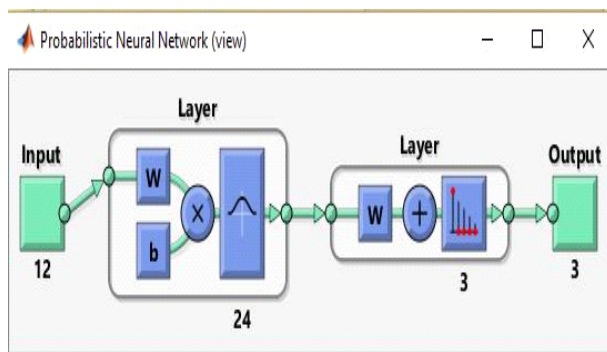$$Ac = \frac{TP+TN}{TP+TN+FP+FN} \qquad (11)$$

Here,

TP stands for true positive

TN stands for true negative

FP stands for false positive

FN stands for false negative

$$Ac = \frac{438+459}{438+459+12+21} = 97.5\% \qquad (12)$$

Fig.5 Designed Probabilistic Neural Network

The neural network designed has 12 neurons in the input layer corresponding to the 12 input parameters or features for training. The classification is in three categories and hence the number of neurons in the output layer is 3. A comparison has also been made with contemporary work, Yamazaki et al. [1] which attains an accuracy of 90.95% classification accuracy.

6. CONCLUSION

It can be concluded form the discussions, mathematical formulations and obtained results that the proposed system attains high level of classification accuracy. The rigorous image pre-processing techniques has resulted in accuracy enhancement in the feature extraction part. The PCA helped in de-correlating the data set so as to remove redundant data and render effective training for the neural network. The accuracy obtained using the data set is 98% which is relatively high compared to contemporary techniques.

7. REFERENCES

[1] N Le, K. Yamazaki, K. Guach (2022), "A multi-task contextual atrous residual network for brain tumor detection & segmentation", IEEE Transactions on Neural Networks And Learning Systems, Vol. 32, No. 12, pp.1-8.

[2] .N Ghassemi, A Shoeibi, M Rouhani (2021)," Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images" Biomedical Signal Processing and Control, Elsevier, 2020, Volume-57, pp-101678.

[3] Abiwinanda N., Hanif M., Hesaputra S.T., Handayani A., Mengko T.R. (2019) Brain Tumor Classification Using Convolutional Neural Network. In: Lhotska L., Sukupova L., Lacković I., Ibbott G. (eds) World Congress on Medical Physics and Biomedical Engineering 2019. IFMBE Proceedings, vol 68/1. Springer

[4] Hasan A., et al. (2018) **"Brain image classification based on discrete wavelet transform and probabilistic neural network"** International Journal of Scientific & Engineering Research, Volume 7, Issue 5

[5] Varuna Shree, N., Kumar, T.N.R. (2018) Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network. Brain Inf. 5, 23–30.

[6] Kolge V. and Kulhalli K.(2012), "PCA And PNN Assisted Automated Brain Tumor Classification", IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), PP: 19-23.

[7] Kulhalli K. and Kolge V(2014), "Primary Level Classification of Brain Tumor using PCA and PNN", International Journal on Recent and Innovation Trends in Computing and Communication Volume: 2 Issue.

[8] Mala K. and Alagappan V(2015), "Neural network based texture analysis of CT images for fatty and cirrhosis liver classification**"** Applied Soft Computing /Volume 32, , Pages 80–86.

[9] Menzes B.(2015), "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging, VOL. 34, NO. 10.

[10] Popescu L and Sasu I(2014),"Feature extraction, feature selection and machine learning for image classification: A case study" International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), IEEE.

[11] Sridhar D. and Krishna M(2013), "Brain Tumor Classification U sing Discrete Cosine Transform and Probabilistic Neural Network", International Conference on Signal Processing, Image Processing and Pattern Recognition [ICSIPR].

[12] B. Farahani, F. Firouzi, V. Chang, M. Badaroglu, N. Constant, and K.Mankodiya,''Towardsfog-drivenIoTeHealth:Promisesandchallenges of IoT in medicine and healthcare,'' Future Gener. Comput. Syst., vol. 78, pp. 659–676, Jan. 2018.

[13] P. M. Shakeel, S. Baskar, V. R. S. Dhulipala, S. Mishra, and M.M.Jaber,''Maintainingsecurityandprivacyi nhealthcaresystemusing learningbaseddeep-Q-networks,''J.Med.Syst.,vol.42,p.186,Oct.2018.

[13] G. S. N. Murthy and T. Veerraju, ''A novel approach based on decreased dimension and reduced gray level range matrix features for stone texture classification,'' Int. J. Elect. Comput. Eng., vol. 7, no. 5, pp. 2502–2513, 2017.

[14] F. J. Martinez-Murcia et al., ''On the brain structure heterogeneity of autism: Parsing out acquisition site effects with significance-

weighted principal component analysis,'' Hum. Brain Mapping, vol. 38, no. 3, pp. 1208–1223, 2017.

[15] B. H. Menze et al., ''The multimodal brain Tumor image segmentation benchmark (BRATS),'' IEEE Trans. Med. Imag., vol. 34, no. 10, pp. 1993–2024, Oct. 2015.

[16] M. Ji et al., ''Detection of human brain Tumor infiltration with quantitative stimulated Raman scattering microscopy,'' Sci. Transl. Med., vol. 7, no. 309, p. 309ra163, 2015.

[17] D. Wang, D. Miao, and C. Xie, ''Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection,'' Expert Syst. Appl., vol. 38, no. 11, pp. 14314–14320, 2011.

[18] F. N. Karameh and M. A. Dahleh, ''Automated classification of EEG signals in brain tumor diagnostics,'' in Proc. IEEE Amer. Control Conf., Chiago, IL, USA, Jun. 2000, pp. 4169–4173.

[19]S.Yang,L.F.Chen,T.Yan,Y.H.Zhao,andY.J.F an,''Anensembleclassification algorithm for convolutional neural network based on AdaBoost,'' in Proc. IEEE/ACIS 16th Int. Conf. Comput. Inf. Sci. (ICIS), May 2017, pp. 401–406.

[20] I. S. Samant, G. K. Kanungo, and S. K. Mishra, ''Desired EEG signals for detection brain tumor using LMS algorithm and feedforward network,'' Int. J. Eng. Trends Technol., vol. 3, no. 6, pp. 718–723, 2012.

[21] Y. Fu et al., ''Network topologies and dynamics leading to endotoxin toleranceandprimingininnateimmunecells,''P LoSComput.Biol.,vol.8, no. 5, p. e1002526, 2012.

[22] G. P. Rathi and S. D. Palani, ''Brain tumor MRI image classification with feature selection and extraction using linear discriminant analysis,'' Int. J. Inf. Sci. Techn., vol. 2, no. 4, p. 146, 2012.

[23] U. Orhan, M. Hekim, and M. Ozer, ''EEG signals classification using the K-means clustering and a multilayer perceptron neural network model,'' Expert Syst. Appl., vol. 38, no. 10, pp. 13475–13481, 2011.

[24] P. P. R. Filho, P. C. Cortez, A. C. da Silva Barros, V. H. C. Albuquerque, and J. M. R. S. Tavares, ''Novel and powerful 3D adaptive crisp active contour method applied in the segmentation of CT lung images,'' Med. Image Anal., vol. 35, pp. 503–516, Jan. 2017.