



Performance Analysis of Machine Learning and Deep Learning Architectures on Early Stroke Detection Using Carotid Artery Ultrasound Images

S. Latha¹, P. Muthu^{2*}, Khin Wee Lai^{3*}, Azira Khalil^{4*} and Samiappan Dhanalakshmi¹

OPEN ACCESS

Edited by:

Sang-Bing Tsai,
Wuyi University, China

Reviewed by:

N. Venkateswaran,
SSN College of Engineering, India
Selvaraju Narayanasamy,
Indian Institute of Technology
Guwahati, India

*Correspondence:

P. Muthu
muthup@srmist.edu.in
Khin Wee Lai
lai.khinwee@um.edu.my
Azira Khalil
azira@usim.edu.my

Specialty section:

This article was submitted to
Alzheimer's Disease and Related
Dementias,
a section of the journal
Frontiers in Aging Neuroscience

Received: 03 December 2021

Accepted: 28 December 2021

Published: 27 January 2022

Citation:

Latha S, Muthu P, Lai KW, Khalil A
and Dhanalakshmi S (2022)
Performance Analysis of Machine
Learning and Deep Learning
Architectures on Early Stroke
Detection Using Carotid Artery
Ultrasound Images.
Front. Aging Neurosci. 13:828214.
doi: 10.3389/fnagi.2021.828214

¹ Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Chennai, India, ² Department of Biomedical Engineering, SRM Institute of Science and Technology, Chennai, India, ³ Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia, ⁴ Faculty of Science and Technology, Universiti Sains Islam Malaysia, Nilai, Malaysia

Atherosclerotic plaque deposit in the carotid artery is used as an early estimate to identify the presence of cardiovascular diseases. Ultrasound images of the carotid artery are used to provide the extent of stenosis by examining the intima-media thickness and plaque diameter. A total of 361 images were classified using machine learning and deep learning approaches to recognize whether the person is symptomatic or asymptomatic. CART decision tree, random forest, and logistic regression machine learning algorithms, convolutional neural network (CNN), Mobilenet, and Capsulenet deep learning algorithms were applied in 202 normal images and 159 images with carotid plaque. Random forest provided a competitive accuracy of 91.41% and Capsulenet transfer learning approach gave 96.7% accuracy in classifying the carotid artery ultrasound image database.

Keywords: carotid artery, ultrasound image, machine learning, deep learning, stroke

INTRODUCTION

Every year, in India, 26% of people die due to cardiovascular diseases, stroke because of artery stenosis is 75%, and heart attack is 42%. In the United States, one of the 19 deaths is due to stroke (Farah, 2018). Risk factors that may lead to stroke are physical inactivity, being obese, heavy drinking, use of illegal drugs, family history having a stroke and other cardiovascular diseases, cholesterol, high blood pressure, diabetes, and smoking. Other factors with increased stroke risk are race-, sex-, age-, and hormones-related problems.

Stroke is the third prominent reason for death in many developed countries (Benjamin et al., 2019). The common cause of stroke is the formation of atherosclerotic plaque in the carotid artery that can grow large enough to block blood flow leading to stenosis or rupture causing clots in the artery. Progressive intimal accumulation of protein, lipid, and cholesterol makes medium- and large-sized arteries, causing atherosclerosis. Atherosclerosis may be existing in body parts, such as infernal aorta, coronary artery, superficial femoral artery, and the common carotid artery

bifurcation region. Strain in the arterial wall causes variance in clinical, mechanical, and molecular levels in the artery. The plaque formation is compensated by artery enlargement with no changes in the lumen region, where blood flows.

The mapping of features to any one of the classes in a computer-assisted diagnostic system is called classification. Machine learning algorithms that are used for biomedical image classification are neural network, backpropagation, support vector machine (SVM), adaptive binary tree-based SVM, decision trees, such as linear regression, logistic regression, random forest, k-nearest neighbor (KNN), k-means, Boltzmann machine, mean shift clustering, Markov statistics nonparametric techniques, and fuzzy-based classification methods.

Stimulated by the function and structure of the brain, an artificial neural network (ANN) was developed. A subset of machine learning, called deep learning, performs classification tasks directly from the images. The accuracy of deep learning sometimes exceeds human performance. The model extracts all the necessary features by itself and performs the classification. Transfer learning is a kind of deep learning which uses the learnt knowledge from some other data and uses that for the application in hand. Some of the transfer learning algorithms are Alexnet, Mobilenet, Imagenet, Capsulenet, etc.

Carl Azzopardi et al. (2020) used a deep neural network (DNN) to delineate lumen-intima boundary (LIB) and media-adventitia boundary (MAB) with a fully automatic segmentation technique. For the network stochastic gradient descent optimization problem, a new objective function was formulated. The invariant intensity data input was given to the network with a bimodal synthesis of amplitude and phase congruency. The performance in MAB and LIB detection was 96.2 and 92.5%, respectively. The study was made with just 15 images in each stenosis category which is not a sufficient number for deep learning-based segmentation. Images from different sources were not considered for learning, missing generalizability (Azzopardi et al., 2020).

Roy-Cardinal et al. (2019) extracted noninvasive vascular ultrasound elastography (NIVE) and ultrasound features, such as homodyned-K (HK), Nakagami parametric maps, log-compressed images. The algorithm identified large lipid area, calcification, ruptured fibrous cap presence, differentiation of nonvulnerable and vulnerable plaques, and confirming symptomatic and asymptomatic patients using a random forest classifier. The study population was 91, and only 5 cases with fibrous caps were involved. A balanced dataset may give better classification performance. Based on elastography and B mode gray-level features, the AUC obtained was 0.90 (95% CI 0.80–0.92, $p < 0.001$). The area of calcification accuracy obtained was 0.95 (95% CI 0.94–0.96, $p < 0.001$), performed using the above features. Area under the curve variation for other tasks varied between 0.79 and 0.97 (Roy-Cardinal et al., 2019).

Loizou et al. (2017) studied the texture variability in the ultrasound video to identify the presence of vulnerable plaque. The videos were intensity normalized, denoised, IMT segmented, and texture feature learned to find systole and diastole states. The texture was visibly variable for diastolic and systolic states. More gray-scale average was recorded for systole compared to

diastole. Plaque structures had variable textures in both the states. Systole and diastole features combined gave better results. Borders of type 1 plaque were not identified by this method. Acoustic shadowing was produced in type V plaque and was not recognizable. The state diagram was improper for 2% of cases (Loizou et al., 2017).

Lekadir et al. (2017) proposed a CNN classification model for the different plaque constituents. Lipid core, calcified tissues, and fibrous caps were detected with a correlation of 0.90 related to clinical results. Based on the patch batched technique, 56 images were converted into 90,000 patches for the process. SVM with predefined image features gave an accuracy of 78.5%. The testing time taken for classifying each image was 52 ± 13 ms, and changes in accuracy were reduced by 0.003 by changing the patches between 9×9 , 11×11 , 13×13 , and 15×15 (Lekadir et al., 2017). Pazinato et al. (2016) used the features of neighboring pixels for carotid image classification. On a dataset with calcium, lipids, muscles, fibrous, and blood tissues texture, gradient, statistical, and local binary pattern (LBP) features were used. Pixel-based machine learning classification was carried out on the normalized image following multiscale description. The method was computationally complex and did not focus on any particular machine learning algorithm. The technique applied in ultrasound tissue engineering achieved a classification accuracy of 73%, and was statistically verified (Pazinato et al., 2016).

Gastouniotti et al. (2015) explained the importance of kinematic features for plaque analysis for a computer-aided diagnosis (CAD). Fisher discriminant ratio-based feature selection and SVM-based classification were performed. Applying texture features gave 80% of accuracy and kinematic features recorded 88% of accuracy. The accuracy of this proposed CAD has still lots of scope for improvement. AUC, specificity, and sensitivity improved by 0.70, 0.83, and 0.67, respectively (Gastouniotti et al., 2015). Vegas-Sánchez-Ferrero et al. (2014) defined a gamma mixture model (GMM) for the subsampled RF images, and their parameters are useful features to identify various plaque tissues. The method outperformed in terms of plaque echogenicity and characteristics. It achieved an accuracy of 95.16% for four-class classifications and 86.56% for three-class classification, which can still be improved (Vegas-Sánchez-Ferrero et al., 2014).

Saba et al. (2021) proposed a classification approach for carotid artery ultrasound images using four machine learning models, one deep learning model, and one transfer learning model. He used the scattering principle of the plaque, where the symptomatic ones are more scattered than the asymptomatic ones (Saba et al., 2021). He achieved stable results for the characterization and classification of the carotid artery ultrasound images.

Classification of the carotid artery images to identify the presence of plaque deposit is performed by machine learning algorithms, CART decision tree, random forest, and logistic regression. Convolutional neural network (CNN)-based deep learning classification and Mobilenet and Capsulenet transfer learning approaches are performed in the carotid artery image database. The performance of these classification methods is analyzed with the true values confirmed by three radiologists.

In this article, section 2 gives the methodology, section 3 describes the results and discussions, and section 4 concludes the article.

METHODOLOGY

This section defines the approach involved in the classification of the carotid artery ultrasound images. Feature extraction and selection are done to obtain the appropriate features. The selected features are given as input to the machine learning classification algorithms, CART decision tree, random forest, and logistic regression. The images are given as input to the CNN, transfer learning algorithms, Mobilenet and Capsulenet. The classification performance measures are used to identify the efficiency of the algorithms.

Figures 1A,B give the sample carotid artery ultrasound images with and without plaque deposit.

Database Creation

Ethical clearance is obtained from the SRM Medical College Hospital and Research Center, Kattankulathur, Tamil Nadu, India, to collect carotid artery ultrasound images. Database of the carotid artery ultrasound B mode images is collected from the Bharat Scans, Chennai and the SRM Medical College Hospital and Research Center, Kattankulathur, Chennai.

Feature Extraction

Machine learning involves high-dimensional data, where the analysis requires a considerable amount of data for learning and testing. The images obtained are denoised by curvelet decomposition to remove speckle and preserve useful edges. Feature reduction minimizes the effects of redundant variables by selecting feature subsets. Choosing the most significant features progresses the classification model performance and reduces over fitting.

Following preprocessing and segmentation of the images, 63 features are taken from the images in the database. A number of 33 texture features, 5 shape features, 10 histogram and correlogram features, and 15 morphology features are extracted from the images. Out of that, 22 most significant features are selected by principal component analysis (PCA) method (Parhizkar et al., 2021).

The most discriminant features from the extracted features are selected based on the following approach. Distance between two classes for every feature is computed as follows for mean m_1 , m_2 and standard deviation σ_1 , σ_2 .

$$\text{distance} = \frac{|m_1 - m_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad (1)$$

Features with more distance are those with more significance. From the 65 extracted features, 22 most significant features were selected for the classification task. PCA-based feature selection was performed in addition. The principal components are derived from the eigenvalues. A correlated feature set is

converted into uncorrelated ones called principal components by an orthogonal transformation.

The selected features are texture, spatial structure, skewness, kurtosis, histogram, correlogram, histogram of oriented gradient (HOG), Gabor wavelet, angular 2nd moment, shape, sharpness, length irregularity, mean probability density function, gray-scale median, multiregion histogram, arterial wall ROI's randomness, absolute gradient, radian and angular sum of discrete Fourier transform for Fourier power spectrum, coarseness, convexity, connectivity, and plaque volume. The potential features are given as input to the machine learning classification algorithms.

Classification by Machine Learning Algorithms

Proper data preparation, automation and iterative learning, testing, scalability, and ensemble modeling are necessary for a classification algorithm. The classification of the carotid artery images database is performed with the machine learning algorithms, CART decision tree, logistic regression, and random forest algorithm.

Machine learning is to develop a mathematical model built by training the inputs. The inputs are the features selected from the ultrasound image dataset of the carotid artery. The learning experience is generalized so that it can give the correct output for the new image which is not in the database. The generalization of the model is improved by applying a validation set to the trained model. The resulting output and error are given as feedback to the input so that training of the model improves. After many iterations of tuning and training of the model, the trained model is used with new unseen test data to find the performance of the approach (Lundervold and Lundervold, 2019; Latha et al., 2020).

CART Decision Tree

The decision tree is a prediction-based machine learning model with parameters represented in the branches and target outputs represented in the form of leaves. Branch labels are represented by leaves and feature conjunctions that lead to the leaves are represented as branches. Target with continuous values is called regression trees. Classification and regression tree (CART) is a nonparametric decision tree algorithm (Seera and Lim, 2014). Information gain defines how to quantify the quality of the split. For attributes p and q , the information gain I is represented as

$$I(p, q) = -\frac{p}{p+q} \log_2 \left(\frac{p}{p+q} \right) - \frac{q}{p+q} \log_2 \left(\frac{q}{p+q} \right) \quad (2)$$

To create a tree from the available attributes, entropy is computed. It depends on how much variance the data has.

$$E(A) = \sum_{i=1}^v \frac{p_i + q_i}{p+q} I(p, q) \quad (3)$$

The training sets each attribute that is found from the gain. It is the variance between entropy and information gain.

$$\text{Gain} = I(p, q) - E(A) \quad (4)$$

Decision trees can identify the nonlinearity in the dataset and adapt accordingly. The data need not be standardized because a

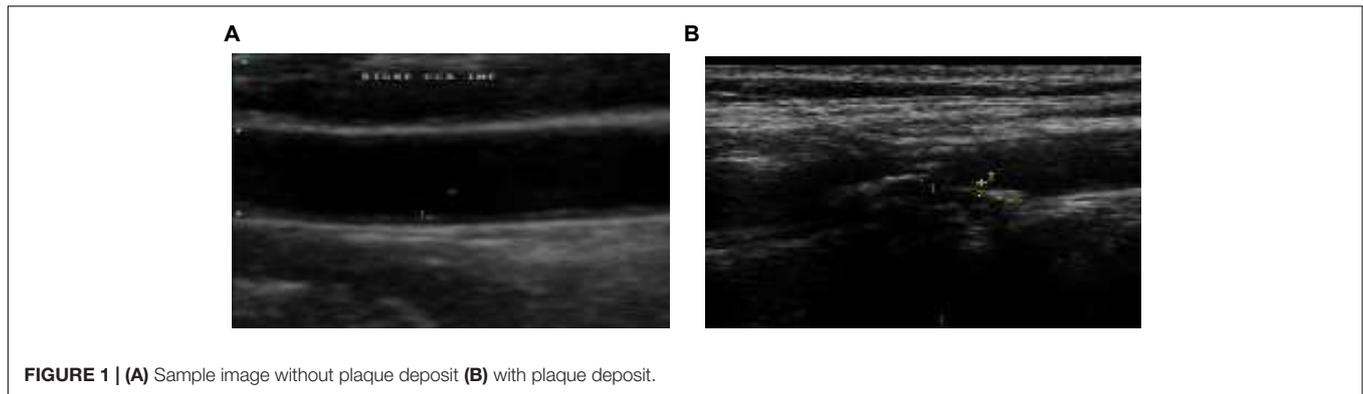


FIGURE 1 | (A) Sample image without plaque deposit (B) with plaque deposit.

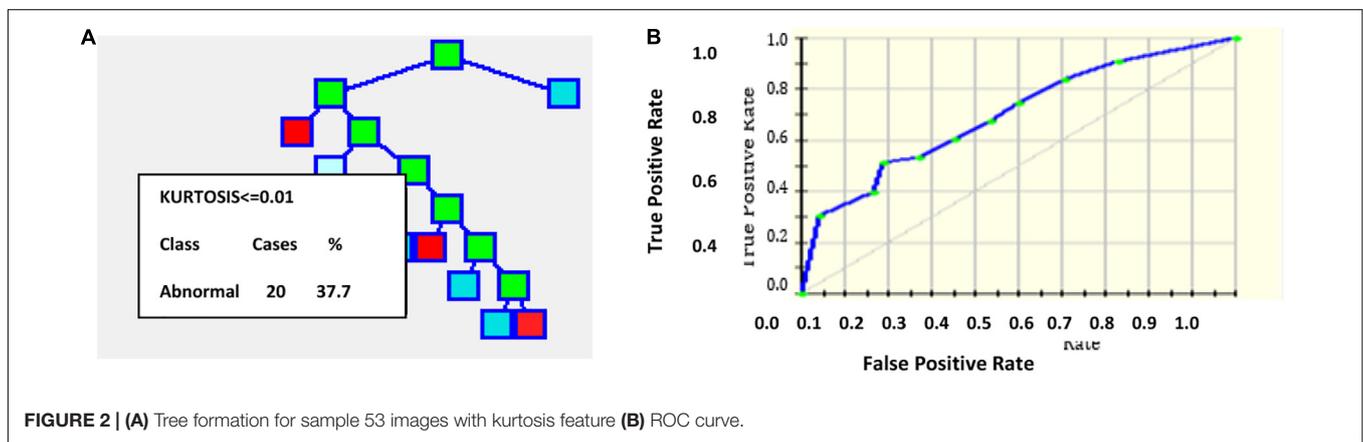


FIGURE 2 | (A) Tree formation for sample 53 images with kurtosis feature (B) ROC curve.

distance measure is not involved in the classification. Sigmoid activation is used to get the optimum classification result. The rules of CART and other decision trees are as follows:

1. Based on a variable’s value, the splitting criteria for a node are formulated.
2. The stopping criteria are decided when to stop splitting a tree.
3. Final target variable at the end of each node is calculated.

An output of one implies the presence of plaque, and zero represents the absence of plaque in the image with a threshold of 0.5. **Figure 2** gives the results of applying the CART decision tree for the carotid artery ultrasound image database. Using the kurtosis feature, the tree formation for sample 53 images is projected in **Figure 2A**. $Kurtosis \leq 0.01$ is separated and branches are formed from that node. **Figure 2B** is the ROC curve for which the AUC is 83.53%, which implies that CART is suitable for disease classification in the carotid artery.

Classification and regression tree is nonparametric and hence is independent on the distribution kind of the input data. The algorithm is not affected by the outliers in the input data. Without strictly following the stopping rule, the tree can be overgrown and can be pruned back to the optimal solution. Fit can be improved using a test set and validation sets. The input variable set can be selected by combining CART with other prediction methods. The drawbacks of

CART include variance in the model when a small change in the database is made and imbalanced class data lead to underfit trees.

Logistic Regression

Binary logistics is more suitable for categorical targets with linear or nonlinear decision boundaries, with a threshold fixed. It applies the logistic or sigmoid function. For the curve’s maximum value L , steepness parameter or growth rate k and x_0 being the midpoint of x , the logistic function is given by

$$f = \frac{L}{1 + e^{-k(x-x_0)}} \tag{5}$$

Assuming threshold 0.5, for probability 0.5, class = 1 is assigned. For probability < 0.5, class = 0 is assigned (Barui et al., 2018). The cost function J used is crossentropy since sigmoid activation is used.

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m \text{cost}(h_{\theta}(x^i), (y^i)) \tag{6}$$

Where $\text{cost}(h_{\theta}(x), y) = -\log(h_{\theta}(x))$ for $y = 1$ and $\text{cost}(h_{\theta}(x), y) = -\log(1 - h_{\theta}(x))$ for $y = 0$. The natural log of odds called logit which transforms the line into the logistic curve is

$$\log\left(\frac{p(x)}{1 - p(x)}\right) = \beta_0 + \beta_1(x) \tag{7}$$

The logistic regression coefficients are found by maximum likelihood estimation. Highly correlated inputs from the database are removed after calculating the pair-wise correlation of the features. It is done to prevent overfit because of multiple highly correlated inputs. The sparsity of the data is also reduced so that the likelihood estimation does not prevent target convergence (Zhang et al., 2018; Javeed et al., 2019; Zhang and Han, 2020). **Figures 3A,B** project the ROC curve and the number of trees with AUC 87.55%.

Random Forest

Random forest is an ensemble classification approach, protecting the structure from being affected by overfitting problems, introduced by Ho in 1995. The tree learners of the random forest follow bootstrap aggregation bagging. Without increasing, bias bootstrapping reduces the variance of the model. The trees are uncorrelated so the prediction of the average of many trees is not noise-sensitive. Bootstrapping gives different input sets for each training time. A forest is created randomly with root, internal, and terminal nodes. Algorithm efficiency improves for a bigger tree. Unlike other decision tree algorithms, random forest decides the root and other nodes randomly.

The classifier is efficient enough to handle missing values and is more suitable for categorical classification. Random forest is created first, and predictions are made from the created forest (Javeed et al., 2019; Wu et al., 2020). Sigmoid activation function is used. Using the random nodes, incorrect labeling can be identified using Gini impurity given by

$$I_G(n) = 1 - \sum_{i=1}^j (p_i)^2 \quad (8)$$

The algorithm for random forest creation is as follows.

1. From a total of m feature sets, K features are randomly selected $k < m$.
2. Find node from features after best split point.
3. From the best divided, segregate child node.
4. The above steps are repeated until 1 number of nodes is achieved.
5. Repeat the above steps for n times to achieve n nodes.

The prediction that forms the created random forest is done by the below procedure.

1. For each test feature, the rules of the model are applied to get the target.
2. For each predicted target, the votes are estimated.
3. The more voted target is considered the outcome.

Figure 4A projects the error rate which is least for nearly 85 number of trees, then increases, becomes constant, and the next drop is marked in nearly 920 trees. **Figure 4B** gives the ROC curve with AUC 90.63%.

Random forest combines individual tree's decisions and considers the maximum voted one, which makes it one of the best machine learning algorithms. Trees are modeled more diversely, thus implementing all possible models, and obtaining all possible

outcomes improves model efficiency. Kernel-induced random forest (KIRF) is followed where trees are built till error no longer reduces. Out of bootstrap (OOB) samples are applied to get the error rate of the random forest by taking the mean of the error from all the bags using all the available features. The drawbacks of the random forest include model complexity, more time consuming than other decision trees, and less intuitive for large decision trees.

Deep Learning Algorithms

Deep learning, which is a class of ANN, extracts the semantic from the images directly, resulting in better classification performance. The deep learning model is built with multisource labeled data and provides more generalized results. The carotid artery ultrasound image classification is performed with a deep learning approach, CNN.

Deep learning is a promising machine learning field that can unravel artificial intelligence problems efficiently. It uses a DNN where the solution depends on the database. Deep learning is superior in terms of nonlinearity, generalization, harmony, fault tolerance, parallelism, and learning. There are undisclosed neural network layers that perform the learning for the available data. Each layer holds a relationship with the next and the previous layers. Deep learning absorbs features and useful representations directly from the raw image bypassing the feature extraction step. This automatic learning of feature representation and learning both happen in the layers.

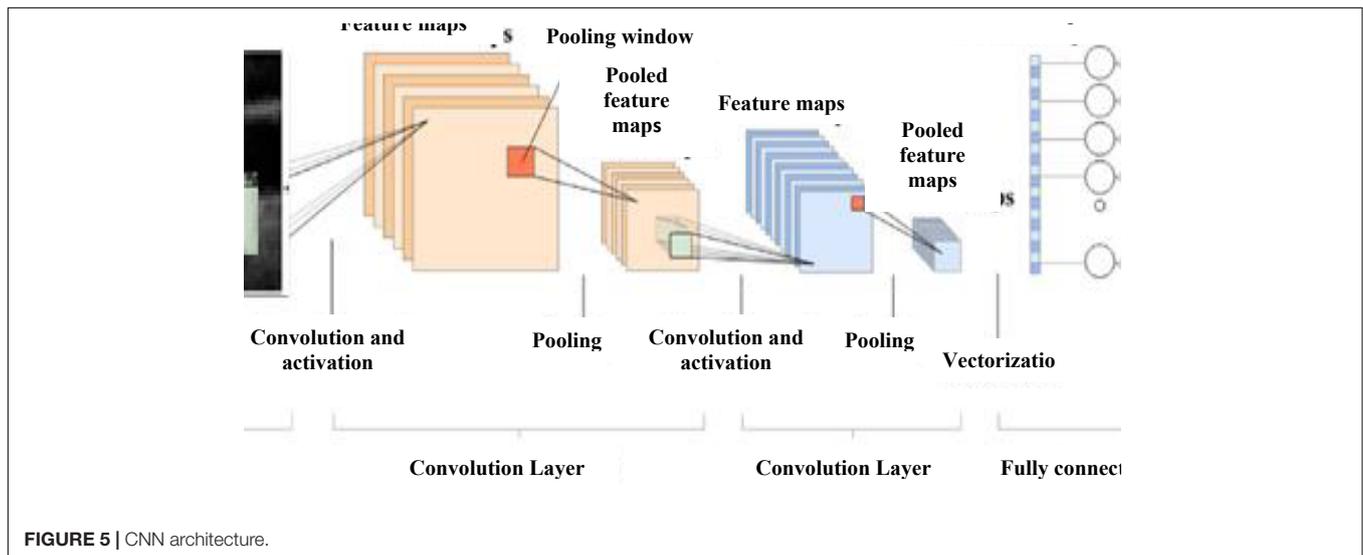
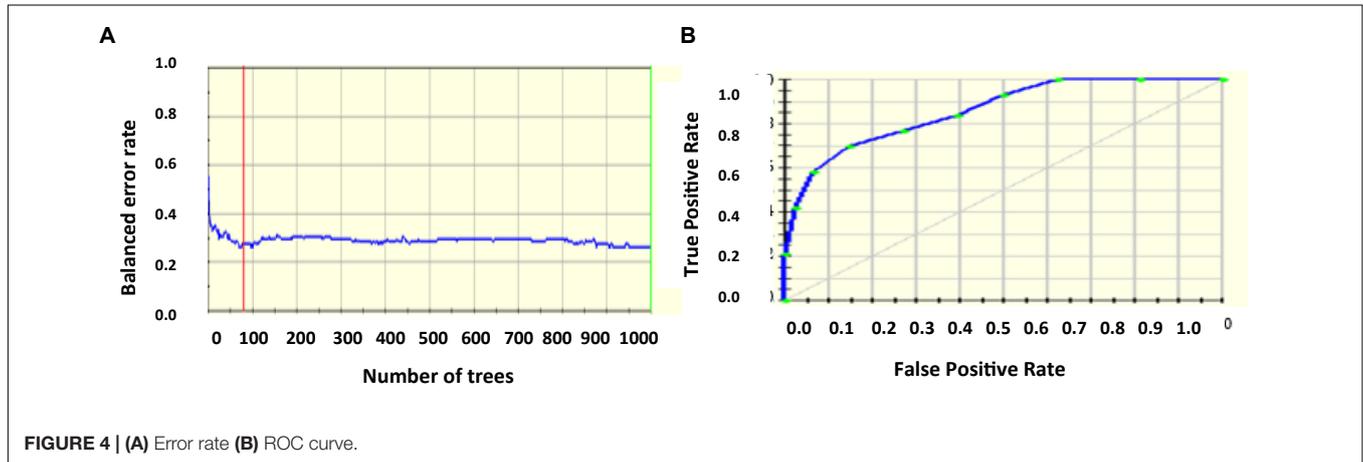
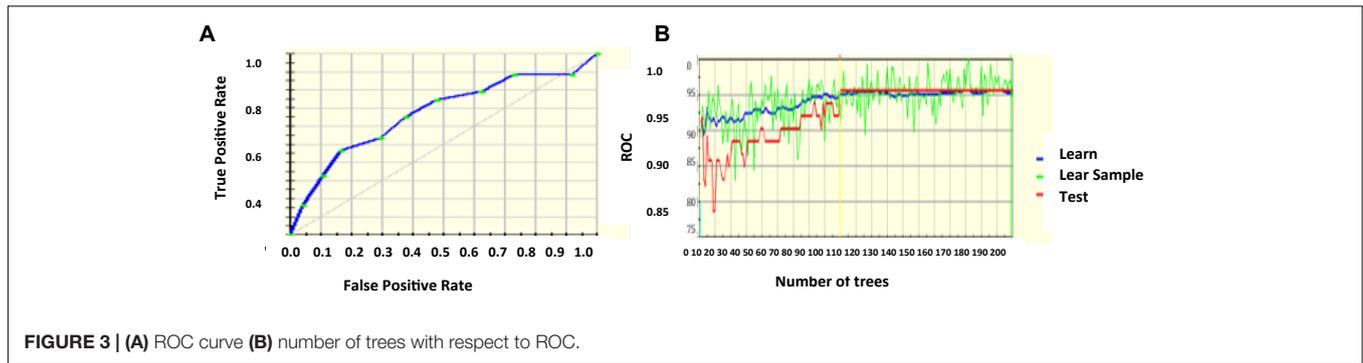
Due to complexity, the importance of the subject, carotid image analysis using machine learning is not efficient enough and needs a model learnt from a huge number of images. The analysis does not depend on the features extracted manually. The data may be patient-dependent and expert-dependent which may influence the outcomes. Deep learning extracts the hidden feature representations of the images and helps in efficient diagnosis. For example, deep learning algorithms are CNN, DNN, DBM, LSTM networks, and generative adversarial networks (GANs), each having their pros and cons which does not require any preprocessing of data. The extension of CNN called transfer learning algorithms, such as Alexnet, Leenet, Googlenet, and Resnet, has proved their efficiency in the testing phase to a huge extent in terms of complexity.

Deep learning stacks many neuron layers constructing a hierarchical feature representation. The layer count in the model is over 1,000 creating a gigantic model memorizing all features and thus makes more intelligent classification.

Deep learning executes feature engineering on its own by combining and correlating the necessary attributes of the image. Deep learning solves the classification problem end-to-end, which makes the model better than other machine learning approaches. There is a lot of scope of development of deep learning with emerging techniques, such as transfer learning. Other challenges of deep learning are interpretability, trust, data, regulations, and workflow integration.

Convolutional Neural Network

Convolutional neural network is a proven traditional deep learning network based on its translation invariance property



and shared weights architecture. All nodes connected to all nodes in the other layers build a much complex system and may be inefficient. CNN uses the domain knowledge of the data preserving the spatial relationship, assembling complex patterns into small, simple patterns (Tajbaksh et al., 2016).

Rectified linear unit (ReLU) activation function is used for CNN activation. In convolution layer activation, previous layer activations are convolved with parameterized filters of size

3×3 . Learning the same weight reduces the complexity of weight calculation for each layer and node. The convolution layer outputs are pooled in a pooling layer. For small grids, the pooling layer provides single output by max-pooling or average pooling. Translational invariance is achieved after the pooling layer preventing a shift in activation maps because of the shift in the input. Increased stride length convolution leads to downsampled pooling reducing the model complexity.

Based on a stochastic sampling of the neural network, dropout regularization is performed. Different neurons are removed in different iterations leading to different outputs each time. Weights are updated each time to get more optimal results. Activation maps subtracted from the mean and divided by standard deviations for each training batch give batch normalized output (Lundervold and Lundervold, 2019). **Figure 5** gives CNN architecture. The image is directly fed as input to the model. The convolution layer extracts features, such as corners, edges, and colors from the input image. Deeper layers extract more deep features, such as plaque structure, kurtosis, texture of plaque, and nonplaque area. Dominant features from the restricted neighborhood are extracted in the pooling layer.

Max-pooling representation is used, which minimizes computational cost and provides translational in-variation to the internal representation. Alternate convolution and pooling layers are used to reduce the large feature space. Later, layers extract more disease-related features assisting the classification process and improve classification accuracy.

After the convolution and pooling, the data are converted into a column vector, suitable for multilevel fully connected architecture. It is followed by a feed-forward neural network and back-propagation architecture in successive training iterations. Dominant and low-level features are adequately identified and classification proceeds.

Transfer Learning Based on Mobile Network Architecture

A network pretrained on available images can be fine-tuned for the application to be performed. When the source and the target are nearly similar, transfer learning works best in terms of weight updating and optimization compared to random initializations.

Figure 6 gives Mobilenet architecture. The types of transfer learning are positive, negative, and neutral. Learning in a condition facilitating another condition is called positive transfer learning. Learning a task that makes learning another task harder is called negative learning. A learning which does not make a change in another learning is called neutral type of learning. A 1×1 convolution is associated with the depthwise convolution outputs in a pointwise convolution layer. In a single step, inputs and outputs are combined using a convolution filter. Using Mobilenet, computation and model size have drastically reduced. Transfer learning marks fast training, more accurate, and needs fewer data. The significant levels of transfer learning are

1. Full network adaptation—weights are updated from a pretrained network instead of arbitrary initialization and apprise them during the training phase (Wang et al., 2016).
2. Partial network adaptation—network parameters from the pretrained network are initialized and used as such for the first few layers and the last layers are updated for training (Zeng et al., 2017; Hesamian et al., 2019).
3. Zero adaptation—network parameters from a pretrained network are used and are not changed throughout.

Zero adaptation may not be suitable for medical images trained with other organs or general images because they may not have similar properties of the carotid image. In using this carotid

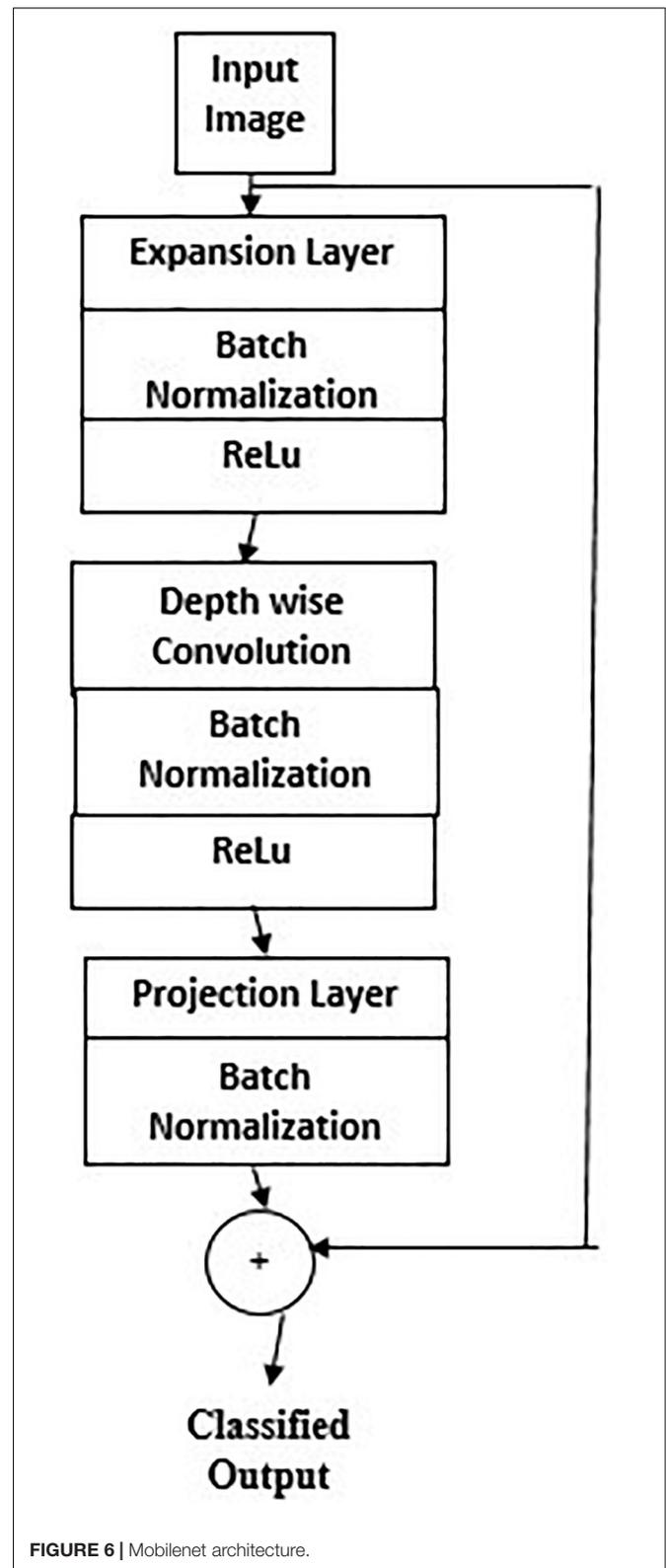


FIGURE 6 | Mobilenet architecture.

database for testing a pretrained network, since the available dataset is small than the training dataset, the following procedure is followed. Overfitting may be a concern because of the small

testing set (Akbarian et al., 2019; Latha et al., 2021). The extracted high-level features may not be similar to the target dataset. The key features of Mobilenet model compared with the CNN model are the following.

1. Most of the pretrained layers near the start of CNN are removed.
2. Instead, fully trained networks equal to the number of classes for the application are included.
3. The newly obtained weights are randomized and replaced instead of the removed network weights.
4. The network is trained to update the weights of the new fully connected layers.

Mobilenet is a family of mobile-first computer vision model for TensorFlow considering restricted data available and suited for embedded applications. The model is small, low latent, and

low power designed by google researchers. A width multiplier parameter is introduced to overcome the resource-accuracy tradeoff. The resolution multiplier term reduces the layers' internal structure. ReLU activation function is used.

Figure 7 gives the transfer learning with mobile net architecture, which provides training accuracy 100% and validation accuracy 95%. Though the training performance is less than that of CNN, the validation performance has improved drastically on using mobile net architecture.

Capsulenet

Geoffrey Hinton proposed Capsulenet in 2017, which is a better representation of capsules than convolution. The neuron activities also have a viewpoint variance in addition. CNN requires augmentation and depends more on texture features, which led to these transfer learning approaches. CNN's max-pooling may lose valuable information because of

```

I0211 08:02:39.756786 34236 retrain.py:704] 2020-02-11 08:02:39.756786: Step 490: Validation accuracy = 91.0% (N=100)
INFO:tensorflow:2020-02-11 08:02:40.463434: Step 499: Train accuracy = 100.0%
I0211 08:02:40.463434 34236 retrain.py:686] 2020-02-11 08:02:40.463434: Step 499: Train accuracy = 100.0%
INFO:tensorflow:2020-02-11 08:02:40.464450: Step 499: Cross entropy = 0.001643
I0211 08:02:40.464450 34236 retrain.py:688] 2020-02-11 08:02:40.464450: Step 499: Cross entropy = 0.001643
INFO:tensorflow:2020-02-11 08:02:40.543189: Step 499: Validation accuracy = 95.0% (N=100)
I0211 08:02:40.543189 34236 retrain.py:704] 2020-02-11 08:02:40.543189: Step 499: Validation accuracy = 95.0% (N=100)
    
```

FIGURE 7 | Transfer learning based on the Mobilenet architecture (snapshot of the obtained results).

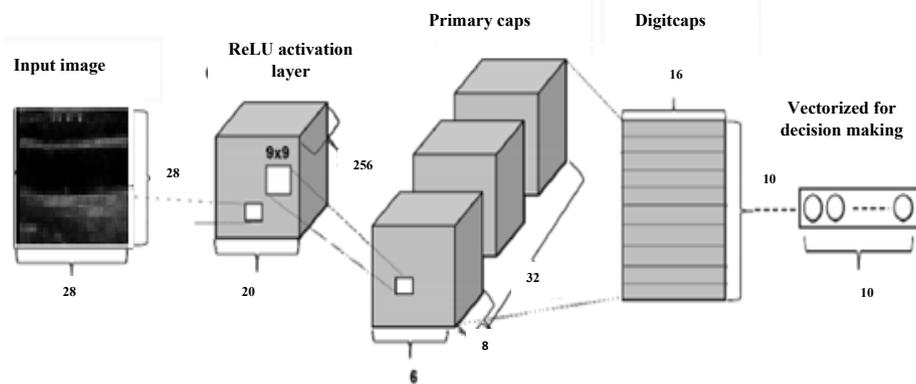


FIGURE 8 | Capsulenet architecture.

TABLE 1 | Confusion matrix of machine learning algorithms.

	CART Decision tree		Logistic regression		Random forest	
	Actual positive (1)	Actual negative (0)	Actual positive (1)	Actual negative (0)	Actual positive (1)	Actual negative (0)
Predicted positive (1)	123	34	120	27	132	23
Predicted negative (0)	23	181	14	200	8	198

```

Epoch 1/10
250/250 [=====] - 356s 1s/step - loss: 0.2015 - accuracy: 0.9121 - val_loss: 3.5977 - val_accuracy: 0.5500
Epoch 2/10
250/250 [=====] - 264s 1s/step - loss: 0.0047 - accuracy: 0.9997 - val_loss: 5.0452 - val_accuracy: 0.5500
Epoch 3/10
250/250 [=====] - 241s 964ms/step - loss: 5.2712e-04 - accuracy: 1.0000 - val_loss: 5.3359 - val_accuracy: 0.5500
Epoch 4/10
250/250 [=====] - 247s 989ms/step - loss: 2.2240e-04 - accuracy: 1.0000 - val_loss: 5.9795 - val_accuracy: 0.5500
Epoch 5/10
250/250 [=====] - 242s 966ms/step - loss: 1.2134e-04 - accuracy: 1.0000 - val_loss: 6.4811 - val_accuracy: 0.5500
Epoch 6/10
250/250 [=====] - 242s 968ms/step - loss: 1.2433e-04 - accuracy: 1.0000 - val_loss: 7.0387 - val_accuracy: 0.5500
Epoch 7/10
250/250 [=====] - 239s 957ms/step - loss: 5.6958e-05 - accuracy: 1.0000 - val_loss: 7.0544 - val_accuracy: 0.5500
Epoch 8/10
250/250 [=====] - 251s 1s/step - loss: 4.7950e-05 - accuracy: 1.0000 - val_loss: 7.0233 - val_accuracy: 0.5500
Epoch 9/10
250/250 [=====] - 289s 1s/step - loss: 3.1103e-05 - accuracy: 1.0000 - val_loss: 7.2318 - val_accuracy: 0.5500
Epoch 10/10
250/250 [=====] - 341s 1s/step - loss: 2.1839e-05 - accuracy: 1.0000 - val_loss: 7.1682 - val_accuracy: 0.5500
    
```

FIGURE 9 | CNN model applied to the carotid artery ultrasound image database (snapshot of the obtained results).

poor relationships between hierarchies of simple and complex objects. Capsulenet applies vector activation and outputs which encodes feature transformation information. ReLU activation function is used.

Figure 8 gives Capsulenet architecture with ReLU activation. Capsules are convolutions with block nonlinearity and routing. The iterations are slow but require few parameters than CNN. Inside the knowledge representations, Capsulenet builds a better model hierarchy. Capsule structures are added to the CNN model, and the outputs are reused to get more stable higher representations. Max-pooling is used instead of dynamic routing and hence achieves translation invariance. It improves the ability of the network to detect an object even wherever it lies in the image.

RESULTS AND DISCUSSION

Choice of performance measures to evaluate the machine learning algorithms gives hope for its practical use. An unsuitable incorrect measure will mislead to wrong results and a flawed model which is not suitable for the application. The available data are imbalanced, and thus, analyzing more number of metrics assists in proper model selection. It involves comparing the proposed model with an existing model or predicting the class label for a given image set.

Performance Metrics

The classification of a carotid artery ultrasound image as symptomatic or asymptomatic is a binary classification problem. The performance depends on the count of correctly classified samples to their class (true positive (TP)), not belonging to the class, correctly classified as (true negative (TN)), samples misclassified to that class (false positive (FP)), and those that are

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	(None, 200, 200, 1)	0	
conv1 (Conv2D)	(None, 192, 192, 256)	20992	input_1[0][0]
primarycap_conv2d (Conv2D)	(None, 92, 92, 256)	5308672	conv1[0][0]
primarycap_reshape (Reshape)	(None, 270048, 8)	0	primarycap_conv2d[0][0]
primarycap_squash (Lambda)	(None, 270048, 8)	0	primarycap_reshape[0][0]
digitcaps (CapsuleLayer)	(None, 1, 16)	34668544	primarycap_squash[0][0]
input_2 (InputLayer)	(None, 1)	0	
mask_1 (Mask)	(None, 16)	0	digitcaps[0][0] input_2[0][0]
capsnet (Length)	(None, 1)	0	digitcaps[0][0]
decoder (Sequential)	(None, 200, 200, 1)	41534016	mask_1[0][0]

Total params: 81,532,224
 Trainable params: 81,532,224

FIGURE 10 | Capsulenet implementation for the carotid artery database images (snapshot of the obtained results).

TABLE 2 | Performance comparison of carotid artery image classification using machine learning approaches.

Algorithm	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)	F score (%)	AUC (%)
CART Decision Tree	84.21	88.72	78.34	84.25	81.19	83.53
Logistic Regression	88.64	93.46	81.63	89.55	85.41	87.55
Random Forest	91.41	96.11	85.16	94.29	89.49	90.63

misrecognized as belonging to that category (false negative (FN)) (Sokolova and Lapalme, 2009). The overall effectiveness of the model is given by

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

The labels class agreement with positive labels in the algorithm is given by

$$\text{precision} = \frac{TP}{TP + FP} \quad (10)$$

Positive label identification efficiency is expressed by recall or sensitivity. The relevant data points are identified using. F score measures the relationship between the positive labeled data and that given in the classifier. Specificity explains how effective the model identifies a negative label. FPR is the false alarm probability and TPR is the recall parameter. The model's ability to identify false classification is derived from the area under the ROC curve (AUC). An AUC rate 1 is expected for an ideal classification model. These measures signify the classification model performance.

$$\text{recall} = \frac{TP}{TP + FN} \quad (11)$$

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

$$\text{F score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (13)$$

$$\text{specificity} = \frac{TN}{TN + FP} \quad (14)$$

$$\text{AUC} = \left(\frac{1}{2}\right) \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right) \quad (15)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

ReLU activation function is used in the classification models.

TABLE 3 | Performance comparison of carotid artery image classification by deep learning approaches.

Algorithm	Accuracy (%)
CNN	55
Mobilenet	95
Capsulenet Transfer Learning	96.7

Machine Learning

Table 1 gives the confusion matrix of the machine learning algorithms applied in the dataset containing 361 images, out of which 159 are abnormal and 202 are those without any disease indications.

The CART model gives an accuracy of 84.21%, specificity 88.72%, sensitivity 78.34%, and precision of 84.25%. The results prove that the model is useful in identifying the negative cases better than the positive ones. Logistic regression records an accuracy of 88.64% for the carotid database. The obtained specificity is 93.46%, sensitivity is 81.63%, and precision is 89.55%. More number of features added to the logistic regression model will increase the variance in the odds and may lead to overfitting. This reduces the generalization of the model fit. Based on the chi-square test, Hosmer–Lemeshow goodness-of-fit measure can improve model performance. The algorithm that assumes the data is noise-free. Outliers from the training data must be removed to prevent misclassification. Random forest gives an accuracy of 91.41%, specificity 96.11%, sensitivity 85.16%, and precision of 94.29%. The above results prove that random forest is a more accurate classifier than logistic regression and CART decision tree for classifying the carotid artery ultrasound images.

Deep Learning

Convolutional neural network model is applied on ultrasound image database for the classification of the images as with and without plaque deposit. The model achieved training accuracy of 100% and validation accuracy of 55% as given in **Figure 9**. **Figure 10** gives the result of the capsulenet implementation in the database.

Convolutional neural network requires a wide number of data for training the model. Because of the limited number of data, the validation performance is nearly half, though the training is efficient. To overcome this, transfer learning was introduced to perform a deep learning architecture with limited training dataset.

Capsules group neurons and thus require fewer parameters between layers. Pose matrix in Capsulenet defines the rotation and translation of an object, which represents its change in viewpoint. It makes the model better generalized to new viewpoints. The spatial relationship between part of the image and the whole is learnt which makes the image identification simple. It is a viewpoint-dependent neural activity which does not require image normalization and can also identify multiply transformed images (Samiappan and Chakrapani, 2016; Arun et al., 2019; del Mar Vila et al., 2020; Samiappan et al., 2020). Underfitting problem was seen in the classification problem by CNN, which has led to poor performance and generalization.

The carotid artery ultrasound image dataset is small and was not sufficient for a deep learning-based classification.

Initially, 300 training images and 61 validation images were used. Data augmentation methods, such as rotation, flipping, and translation were done to improve the classification accuracy.

Table 2 gives the performance of the three machine learning techniques applied for the carotid artery ultrasound image database. Random forest gives computationally faster and improved performance results compared to CART and logistic regression. Since the dataset was small (361 images), machine learning algorithms were not computationally complex, lags in accuracy of identification of the disease. Capsules group neurons and thus require fewer parameters between layers. Pose matrix captures rotated and translated versions as linear transformations, and so, Capsulenet is better generalized to new viewpoints. The spatial relationship between part of image and the whole is learnt, which makes the image identification simple. Capsulenet achieves accuracy of 96.7%, which is the highest for the carotid artery database images.

The images in the database were flipped to both plane axis rotated to $\pi/4$ axis. **Table 3** gives the performance of the three deep learning techniques applied in the carotid artery image database.

Proposed Capsulenet with max-pooling gives 12.91, 8.33, 5.47, 43.12, and 1.75% improvement in accuracy compared with a CART decision tree, logistic regression, random forest, CNN, and Mobilenet classification algorithms, respectively. Negative transfer is the interference of the previous knowledge in the new learning. It has not affected the classification performance of the carotid artery ultrasound images. It is proved with improved performance measures.

It is proved that deep learning approaches give improved accuracy of 95.7% for Capsulenet compared to other machine learning and deep learning algorithms reported in the literature.

CONCLUSION

A number of 361 images were processed to form a database with the help of radiologists. Extracted features from the database images are applied to the machine learning algorithms CART decision tree, random forest, logistic regression, CNN model, Mobilenet, and Capsulenet transfer learning algorithms for classifying the images as normal or abnormal. Machine learning algorithms were able to perform with an accuracy of 84.21, 88.64, and 91.41%, respectively, for CART, logistic regression, and random forest. Proposed Capsulenet transfer learning approach

REFERENCES

- Akbarian, S., Delfi, G., Zhu, K., Yadollahi, A., and Taati, B. (2019). Automated non-contact detection of head and body positions during sleep. *IEEE Access* 7, 72826–72834. doi: 10.1109/ACCESS.2019.2920025
- Arun, P. V., Buddhiraju, K. M., and Porwal, A. (2019). Capsulenet-based spatial-spectral classifier for hyperspectral images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 12, 1849–1865. doi: 10.1109/JSTARS.2019.2913097

eliminates the need for large amount of training data. Proposed Capsulenet with max-pooling gives 12.91, 8.33, 5.47, 43.12, and 1.75% improvement in accuracy compared with CART decision tree, logistic regression, random forest, CNN, and Mobilenet classification algorithms, respectively.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

Ethical clearances were obtained from SRM Medical College Hospital and Research Center, India. Ethics Clearance Number: 1736/IEC/2019. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SL performed conceptualization, methodology, design, data collection, data visualization, formal analysis, reviewing, and editing. PM carried out conceptualization, methodology, design, investigation, data collection, data analysis, and writing original draft preparation. KL done conceptualization, methodology, data collection, data visualization, formal analysis, reviewing, and editing. AK involved in data curation, critical analysis, writing, reviewing, and editing. SD contributed in formal analysis, reviewing, and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by The Institution of Engineers India, RDDR2016064.

ACKNOWLEDGMENTS

The authors are grateful to the SRM Institute of Science and Technology, the University of Malaya, and the Universiti Sains Islam Malaysia for supporting this research.

- Azzopardi, C., Camilleri, K. P., and Hicks, Y. (2020). Bimodal automated carotid ultrasound segmentation using geometrically constrained deep neural networks. *IEEE J. Biomed. Health Inform.* 24, 1004–1015. doi: 10.1109/JBHI.2020.2965088
- Barui, S., Latha, S., Samiappan, D., and Muthu, P. (2018). SVM pixel classification on colour image segmentation. *J. Phys. Conf. Ser.* 1000:012110. doi: 10.1088/1742-6596/1000/1/012110
- Benjamin, J., Muntner, P., Alonso, A., Bittencourt, M. S., Callaway, C. W., Carson, A. P., et al. (2019). Heart disease and stroke statistics- 2019 update a report from

- the American Heart Association. *Circulation* 139, e56–e528. doi: 10.1161/CIR.0000000000000659
- del Mar Vila, M., Remeseiro, B., Grau, M., Elosua, R., Betriu, A., Fernandez-Giraldez, E., et al. (2020). Semantic segmentation with DenseNets for carotid artery ultrasound plaque segmentation and CIMT estimation. *Artif. Intell. Med.* 103:101784. doi: 10.1016/j.artmed.2019.101784
- Farah, A. (2018). *Nutritional and Health Effects of Coffee*. Rio de Janeiro: Federal University of Rio de Janeiro. doi: 10.19103/AS.2017.0022.14
- Gastouniotti, A., Makrodimitris, S., Golemati, S., Kadoglou, N. P. E., Liapis, C. D., and Nikita, K. S. (2015). A novel computerized tool to stratify risk in carotid atherosclerosis using kinematic features of the arterial wall. *IEEE J. Biomed. Health Inform.* 19, 1137–1145.
- Hesamian, M. H., Jia, W., He, X., and Kennedy, P. (2019). Deep learning techniques for medical image segmentation: achievements and challenges. *J. Digit. Imaging* 32, 582–596. doi: 10.1007/s10278-019-00227-x
- Javed, A., Zhou, S., Yongjian, L., Qasim, I., Noor, A., and Nour, R. (2019). An intelligent learning system based on random search algorithm and optimized random forest model for improved heart disease detection. *IEEE Access* 7, 180235–180243. doi: 10.1109/ACCESS.2019.2952107
- Latha, S., Samiappan, D., and Kumar, R. (2020). Carotid artery ultrasound image analysis: a review of the literature. *Proc. Inst. Mech. Eng. H J. Eng. Med.* 234, 417–443. doi: 10.1177/0954411919900720
- Latha, S., Samiappan, D., Muthu, P., and Kumar, R. (2021). Fully automated integrated segmentation of carotid artery ultrasound images using DBSCAN and affinity propagation. *J. Med. Biol. Eng.* 41, 260–271. doi: 10.1007/s40846-020-00586-9
- Lekadir, K., Galimzianova, A., Betriu, A., Del Mar Vila, M., Igual, L., Rubin, D. L., et al. (2017). A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound. *IEEE J. Biomed. Health Inform.* 21, 48–55. doi: 10.1109/JBHI.2016.2631401
- Loizou, C. P., Pattichis, C. S., Pantziaris, M., Kyriacou, E., and Nicolaidis, A. (2017). Texture feature variability in ultrasound video of the atherosclerotic carotid plaque. *IEEE J. Transl. Eng. Health Med.* 5:1800509. doi: 10.1109/JTEHM.2017.2728662
- Lundervold, A. S., and Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Deep Learn. Med. Phys.* 29, 102–127. doi: 10.1016/j.zemedi.2018.11.002
- Parhizkar, T., Rafeipour, E., and Parhizkar, A. (2021). Evaluation and improvement of energy consumption prediction models using principal component analysis based feature reduction. *J. Clean. Prod.* 279:123866. doi: 10.1016/j.jclepro.2020.123866
- Pazinato, D. V., Stein, B. V., de Almeida, W. R., Werneck, R. O., Mendes Júnior, P. R., Penatti, O. A., et al. (2016). Pixel-level tissue classification for ultrasound images. *IEEE J. Biomed. Health Inform.* 20, 256–267. doi: 10.1109/JBHI.2014.2386796
- Roy-Cardinal, M., Destrempe, F., Soulez, G., and Cloutier, G. (2019). Assessment of carotid artery plaque components with machine learning classification using homodyned-K parametric maps and elastograms. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 66, 493–504. doi: 10.1109/TUFFC.2018.2851846
- Saba, L., Sanagala, S. S., Gupta, S. K., Koppula, V. K., Laird, J. R., Viswanathan, V., et al. (2021). A multicenter study on carotid ultrasound plaque tissue characterization and classification using six deep artificial intelligence models: a stroke application. *IEEE Trans. Instr. Meas.* 70, 1–12. doi: 10.1109/TIM.2021.3052577
- Samiappan, D., and Chakrapani, V. (2016). Classification of carotid artery abnormalities in ultrasound images using an artificial neural classifier. *Int. Arab J. Inf. Technol.* 13, 756–762.
- Samiappan, D., Latha, S., Rao, T. R., Verma, D., and Sriharsha, C. S. A. (2020). Enhancing machine learning aptitude using significant cluster identification for augmented image refining. *Int. J. Pattern Recognit. Artif. Intell.* 34:2051009. doi: 10.1142/S021800142051009X
- Seera, M., and Lim, C. P. (2014). Online motor fault detection and diagnosis using a hybrid FMM-CART model. *IEEE Trans. Neural Netw. Learn. Syst.* 25, 806–812. doi: 10.1109/TNNLS.2013.2280280
- Sokolova, M., and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* 45, 427–437. doi: 10.1016/j.ipm.2009.03.002
- Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Todd Hurst, R., Kendall, C. B., Gotway, M. B., et al. (2016). Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans. Med. Imaging* 35, 1299–1312. doi: 10.1109/TMI.2016.2535302
- Vegas-Sánchez-Ferrero, G., Seabra, J., Rodriguez-Leor, O., Serrano-Vida, A., Aja-Fernández, S., Palencia, C., et al. (2014). Gamma mixture classifier for plaque detection in intravascular ultrasonic images. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 61, 44–61. doi: 10.1109/TUFFC.2014.6689775
- Wang, J., MacKenzie, J. D., Ramachandran, R., and Chen, D. Z. (2016). “A deep learning approach for semantic segmentation in histology tissue images,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, eds S. Ourselin, L. Joskowicz, M. Sabuncu, G. Unal, and W. Wells (Cham: Springer), 176–184. doi: 10.1007/978-3-319-46723-8_21
- Wu, J., Chen, P., Lin, C., Chen, S., and Shung, K. K. (2020). Breast benign and malignant tumors rapidly screening by ARFI-VTI elastography and random decision forests based classifier. *IEEE Access* 8, 54019–54034. doi: 10.1109/ACCESS.2020.2980292
- Zeng, G., Yang, X., Li, J., Yu, L., Heng, P. A., and Zheng, G. (2017). “3D U-net with multi-level deep supervision: fully automatic segmentation of proximal femur in 3D MR images,” in *Proceedings of the International Workshop on Machine Learning in Medical Imaging*, eds Q. Wang, Y. Shi, H. I. Suk, and K. Suzuki (Cham: Springer), 274–282. doi: 10.1007/978-3-319-67389-9_32
- Zhang, C., Yao, L., Song, S., Wen, X., Zhao, X., and Long, Z. (2018). Euler elastica regularized logistic regression for whole-brain decoding of fMRI data. *IEEE Trans. Biomed. Eng.* 65, 1639–1653. doi: 10.1109/TBME.2017.2756665
- Zhang, Z., and Han, Y. (2020). Detection of ovarian tumors in obstetric ultrasound imaging using logistic regression classifier with an advanced machine learning approach. *IEEE Access* 8, 44999–45008. doi: 10.1109/ACCESS.2020.2977962

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Latha, Muthu, Lai, Khalil and Dhanalakshmi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.