Healthcare Analytics 3 (2023) 100196

Contents lists available at ScienceDirect

Healthcare Analytics

journal homepage: www.elsevier.com/locate/health

A novel machine learning-based feature extraction method for classifying intracranial hemorrhage computed tomography images

Santwana Gudadhe [a], Anuradha Thakare [a,*], Ahmed M. Anter [b,c]

[a] Computer Engineering, Pimpri Chinchwad College of Engineering, Pune, India
[b] Egypt-Japan University of Science and Technology (E-JUST), Alexandria, 21934, Egypt
[c] Faculty of Computers and Artificial Intelligence, Beni-Suef University, Benisuef, 62511, Egypt

ARTICLE INFO

Keywords:
Machine learning
Intracranial hemorrhage
Computed tomography
Cosine transform
Wavelet transform
Classifiers
Gray level co-occurrence matrix

ABSTRACT

One of the most serious forms of brain stroke is intracranial hemorrhage (ICH). When an artery bursts, the brain and the tissue around the artery start bleeding. This study proposes a joint feature selection strategy to classify computed tomography (CT) images of intracranial hemorrhage. The joint feature set is composed of transform and texture features. Joint features are constructed from a combination of grey level co-occurrence matrix (GLCM) features, discrete wavelet features (DWT), and discrete cosine features (DCT). Brain hemorrhage CT image classification uses ensemble-based machine learning (ML) techniques. On the training dataset, a Synthetic Minority Over-Sampling Technique (SMOTE) is applied to treat the problem of oversampling by adding fresh data. Additionally, the sequential forward feature selection technique is used to obtain feature subsets. The classification accuracy is further examined for varied feature vector sizes. Confusion matrix, precision, and recall in categorization are employed as performance evaluation measurements. The ML-based ensemble classifiers can produce highly accurate results with the aid of the proposed novel feature extraction mechanism. When taking into consideration a crucial feature set consisting of six features, it can be seen that Random Forest obtained the greatest accuracy, which is 87.22%.

1. Introduction

Patients' lives can be drastically changed by a brain hemorrhage. In the medical field, a correct diagnosis is crucial to accelerating the patient's recovery. The condition known as intracranial hemorrhage (ICH) is diverse. It comprises subarachnoid hemorrhage and intracerebral hemorrhage [1]. Subarachnoid hemorrhage results from bleeding in the area around the brain. On the other hand, intracerebral hemorrhage is brought on by bleeding within the actual brain tissue. The primary methods for diagnosing ICH are a patient's physical examination and medical history. 50% of patients die within 24 h, making accurate and prompt diagnosis essential. 35%–52% of patients who enter the crucial area may die in a month [2], with mortality increasing to 60% after 30 days. First, a clinician uses non-contrast CT to make a diagnosis. The site of the brain hemorrhage is determined and pinpointed using CT image analysis. A stroke is an urgent disorder of the brain that needs to be treated right away. Accurate ICH diagnosis, a lengthy decision-making process, occasional insufficient expertise, and possible inadequacies in the decision-making process are just some of the many challenges associated with dealing with it [3].

Traditional methods include looking at a patient's body and looking at CT scans [4,5]. Treatment is administered based on the specialist radiologists' understanding. The radiologist needs to be backed up by an automated computerized diagnostic tool that helps them make decisions about patient care more swiftly. In order to speed up clinical workflow and shorten diagnosis times, contemporary research investigations have mainly focused on the analysis of hemorrhage. The classification technique specifically helps medical experts correctly evaluate the images and distinguish between hemorrhage and normal images. To classify CT images of intracranial hemorrhages, researchers tried out a variety of feature extraction strategies based on texture, shape, morphs, statistical features, etc. Most common approaches to categorization were texture analysis, transform based and statistical features. For transform-based classification, very little work has been noticed. Considering the transform and texture based method, to take advantage of the strengths of both methods, this work aims to develop a joint feature set from a texture and transform-based perspective and then use that set to classify intracranial hemorrhage CT images using a variety of machine learning classifiers. The GLCM technique is used to generate the texture features, while the cosine and wavelet

* Corresponding author.
E-mail addresses: santwana.gudadhe@pccoepune.org (S. Gudadhe), anuradha.thakare@pccoepune.org (A. Thakare), Ahmed_Anter@fcis.bsu.edu.eg (A.M. Anter).

https://doi.org/10.1016/j.health.2023.100196
Received 31 January 2023; Received in revised form 26 April 2023; Accepted 11 May 2023

transforms are used to obtain the transform feature. Through the use of the GLCM, wavelet, and cosine transforms, we are able to generate joint feature vectors. This study introduces a novel machine learning-based feature extraction method for classifying intracranial hemorrhage CT scans. Finally, classification accuracy is calculated for logistic regression (LR) [6], K-nearest neighbor (KNN) [7], Gaussian naive Bayes (GNB) [8], support vector machine (SVM) [6,9,10], decision tree (DT) (entropy), Random Forest (RF) [11,12], AdaBoost, and Gradient boost [13] machine learning classifiers.

Some of the most salient features of the proposed combined feature selection strategy for classifying intracranial hemorrhage CT images using machine learning methods are as follows: The accurate extraction of features aids in differentiating abnormal from normal images, so this is a new twist on the feature extraction method. To reliably extract features from CT images of intracranial hemorrhage, a texture-based characteristic feature extraction method is suggested. The experiments and comparisons using various machine learning classifiers verify the suggested work. The proposed method can distinguish between hemorrhage and regular CT scans. This article's goal is to explain how to extract cerebral hemorrhage CT imaging features using a transform approach, a texture-based strategy, and a machine learning classification algorithm's performance evaluation for the joint feature selection method.

This paper has the following structure: 2. is concerned with a review of the research. Materials and methods were discussed in 3. Proposed joint feature method is exhaustively discussed in Section 4. Section 5 discusses the results analysis, while Section 6 contains the discussion. Section 7 then covers the conclusion and future scope.

2. Literature survey

According to current studies, image analysis is the most important aspect of medicine. Radiologists can make diagnoses more precisely by using image analysis. Hemorrhage identification, categorization, and prediction are receiving increased attention in the investigation of brain strokes recently. The ability of patients to recover critically depends on an accurate diagnosis of cerebral bleeding. The categorization process is significantly aided by the process of feature extraction. Obtaining optimal and precise extracted feature sets can hasten the classification process. One of the most promising techniques for precise diagnosis is the selection of texture features. To correctly classify the input images, researchers used the texture-based Local Binary pattern [14], GLCM [15], and first-order order features [16], and different classifiers were evaluated. The shape features and GLCM texture data from ROI images were extracted [17] in order to identify hemorrhage images in contrast to the normal images using the RF classifier. Wavelet-based methods for feature extraction with a gray-level run time were first presented. Support vector machine were used to segment and identify soft tissues from images. Optimal features were achieved by merging the novel algorithm approach with the distance measure (Bhattacharyya). Utilizing feature extraction from ROI to reduce processing time speeds up execution and diagnostics. Rapid detection of intracranial hemorrhages using pretrained deep learning models was performed in [18]. Along with pretrained model, hybrid approach of GLCM and local binary pattern with AlexNet was also proposed [18].

Numerous researchers have looked into the extraction of discriminative features from ROI brain CT scans as well as classification utilizing artificial neural networks (ANN), DT, and SVM [19]. GLCM and Length Matrix were used to extract texture features along with Hu moments in [20]. These features were optimized and further classified using a machine-learning classifier. Various texture feature learning-based approaches are explored in [21], along with the performance metrics and traits that they exhibit. CT image classification with ANN, KNN, and SVM classifiers, as shown in [22]. The values of the discrete wavelet transform were used as features, and principal component analysis was performed. Applying the SVM, Bayesian network, and DT classifiers allowed us to detect and categorize cerebral bleeding in a CT image of the brain. Shape-based features have a crucial role [23]. The distinction between pathological and normal brain bleeding is covered in [24]. The proposed work focuses mostly on image segmentation and classification. Minimal angular local binary patterns and GLCM-level features were considered for classification; brain bleeding (normal and abnormal) is accomplished using a naive Bayes-probabilistic kernel classifier.

GLCM features and ML algorithms like SVM were outlined in [25] as a method for texture image classification. Ten GLCM characteristics were extracted and used for SVM classification. Both the training and assessment phases of the suggested method make use of datasets. Limited research work has been observed in brain image classification using the cosine and wavelet transform approaches. Pass-band DCT, k-means clustering, and SVM classification were proposed for brain tumor classification [26]. A two-stage reduction in dimension was achieved by combining a smaller number of DCT coefficients with the k-means clustering method. The linear kernel was selected to be used for SVM categorization. In [27], the authors propose a technique for feature selection that uses a combination of features. Wavelet transform pulled features from images, and principal component analysis reduced feature space.

Intracranial hemorrhage detection from imaging includes accurate diagnosis of acute ICH in 3D CT scans, which was achieved by using a symmetry-based detection method [28]. An imaging-based machine learning algorithm was developed in [29] with the purpose of functional outcome prediction from ICH patients. The proposed method, which used quantitative image features from non-enhanced CT scans, distinguished between ICH patients with good functional outcomes and those with poor functional outcomes at different mRS cut-off values. For the automatic identification and treatment of intracerebral hemorrhage in CT scans, see [30], which proposes a method for doing so. The approach accurately differentiates between normal and ICH images by using supervised learning methods and the extraction of entropy features.

3. Material and methods

3.1. Dataset description

Kaggle [31] is used to download the common intracranial hemorrhage CT imaging collection for brain stroke. 30 hemorrhage patient records and 50 records for healthy patients are included in the collection. Averages of 30 CT imaging slices are provided for each subject. Intracranial hemorrhage CT imaging datasets are subjected to feature extraction. The extracted feature vectors are kept and provided to classifiers as input. Fig. 1 shows an example image dataset for brain strokes.

The experiment includes 1000 normal images and 825 abnormal images. Every image has a.jpg file extension. The original photos have a resolution of 640 by 640. Because of the imbalance in the training data, SMOTE was used. SMOTE helps to create new data, which is then oversampled for the aberrant pictures [32–34]. It has been scaled down to 256 × 256 sizes for processing. For preprocessing, a Gaussian filter [35] is applied. Joint feature selection and classification by using different machine learning classifiers is evaluated and discussed below in the subsection.

3.2. Texture feature

Texture feature extraction is one of the promising ways. The feature extraction technique based on image texture is called Grey Level Co-occurrence Matrices [36,37]. It makes use of the notion of pixel intensity distribution. The main application is the analysis of X-rays. The homogeneity value for each pixel is obtained, and if a change is observed, it indicates that there has been a significant change in

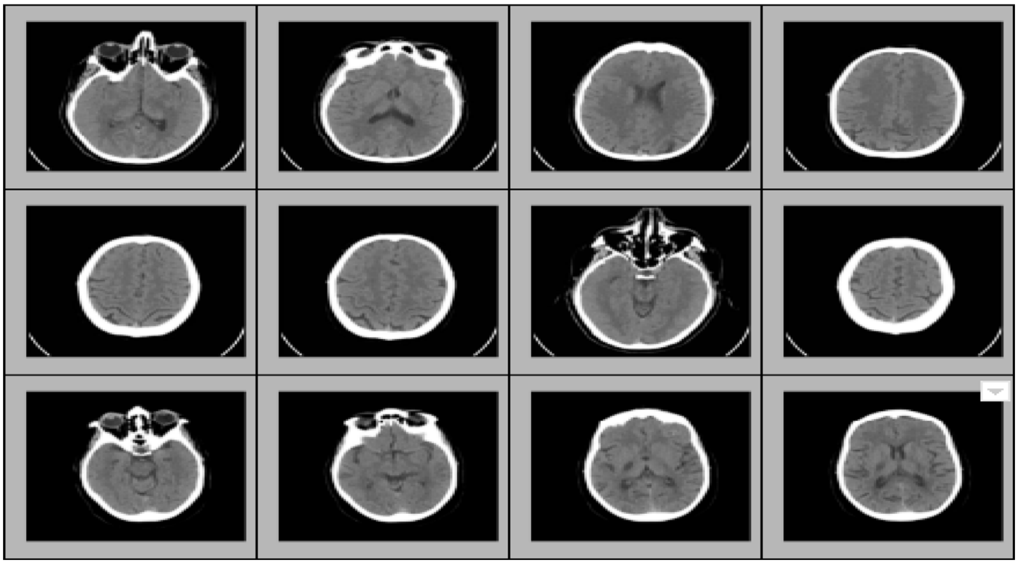

Fig. 1. Sample set of images from Dataset.

the distribution of various textures. The abnormal region of the brain is characterized by a distinctively different texture from that of the surrounding region. The likelihood of finding an anomalous region is higher when there is a major move in the matrix value [38,39]. Gray Level Co-occurrence Matrix characteristics based on images are used in this paper. Extraction Calculation of some of the features in the following list [18]

- Contrast — It is measurement of how the intensity contrast changes between a pixel and its adjoining pixel has changed. Using Eq. (1), contrast can be determined.

$$Contrast = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \tag{1}$$

- Energy - Energy — It determines how many square elements there are overall in the GLCM matrix. Its range of values range from 0 to 1. It can be calculated with Eq. (2).

$$Energy = \sum_{i,j=0}^{N-1} P_j(i,j)^2 \tag{2}$$

- Homogeneity — Homogeneity is a measurement of how evenly the grayscale levels are dispersed throughout the image. Homogeneity can be determined with Eq. (3)

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{(1+(i-j))^2} \tag{3}$$

- Correlation — Is a measurement of grey level correlation within the neighboring pixel. Eq. (4) is a formula for calculating correlation.

$$Correlation = \sum_{i,j=0}^{N-1} \left(\frac{(i-\mu)(j-\mu)}{\sigma^2} \right) \tag{4}$$

3.3. Transform features

3.3.1. Cosine transform

DCT-based feature extraction transforms images using the cosine transform [40]. Low-energy locations are given high energy coefficients in the transform image process. In a 2D image after the cosine transform is applied, a relatively small number of lower-order DCT coefficients store the majority of the energy. The energy composition of these DCT coefficients would serve as the basis for the quantization of these coefficients. These higher-order coefficients can only be fully condensed using a smaller number of bits, whereas the lower-order coefficients can use an increased bit count.

After that, the high-energy and low-energy coefficients are used to generate feature vectors through the process of determining their standard deviation and mean in order to acquire the feature values through the cosine transform. Eqs. (5) and (6) below provide a mathematical representation of a 2-D DCT image of dimension M x N [41]. An image is a 2-D pixel matrix where each position (u,v) stands for a pixel value for that specific position. Consequently, to convert an image into its equivalent 2-D DCT is the DCT matrix we employ Eqs. (5) & (6).

$$F(x,y) = \alpha(x)\,\alpha(y) \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} \cos\left[\frac{\pi x}{2N}(2u+1)\right] \cos\left[\frac{\pi y}{2M}(2v+1)\right] f(x,y) \tag{5}$$

$$\alpha(x)\,\alpha(y) = \left\{ \sqrt{\frac{1}{N}}\ for\ x,y \neq 0;\ \sqrt{\frac{2}{N}}\ for\ x,v=0 \right\} \tag{6}$$

f(u,(v) is pixel intensity with image size $M \times N$.

3.3.2. Wavelet transform

The Haar transform is used in wavelet transforms due to its clarity. Using consecutive low-pass and high pass filters. This 2-D transform divides an input image into several frequency bands before returning the approximation coefficient matrix and detail coefficient matrices. These coefficients are used to calculate mean, entropy, and standard deviation, and these values are regarded as feature vectors that have been DWT-based altered. Eqs. (7) and (8) define the mathematical DWT calculation [16,42].

$$W_\varphi(j_0,k) = \frac{1}{\sqrt{M}} \sum_x f(x)\,\varphi_{j_{0,k}}(x) \tag{7}$$

$$W_\Psi(j_0,k) = \frac{1}{\sqrt{M}} \sum_k f(x)\Psi_{j,k}(x) \tag{8}$$

Where, $f(x)\,\varphi_{j_{0,k}}(x)$ and $\Psi_{j,k}(x)$ are functions of variables x.

4. Proposed joint feature extraction approach

The suggested intracranial hemorrhage classification system takes into account transform and texture properties. In addition to intracranial hemorrhage CT image classification, the proposed work covers the joint feature selection approach and its evaluation. Fig. 2 shows the proposed joint feature extraction approach. The suggested method

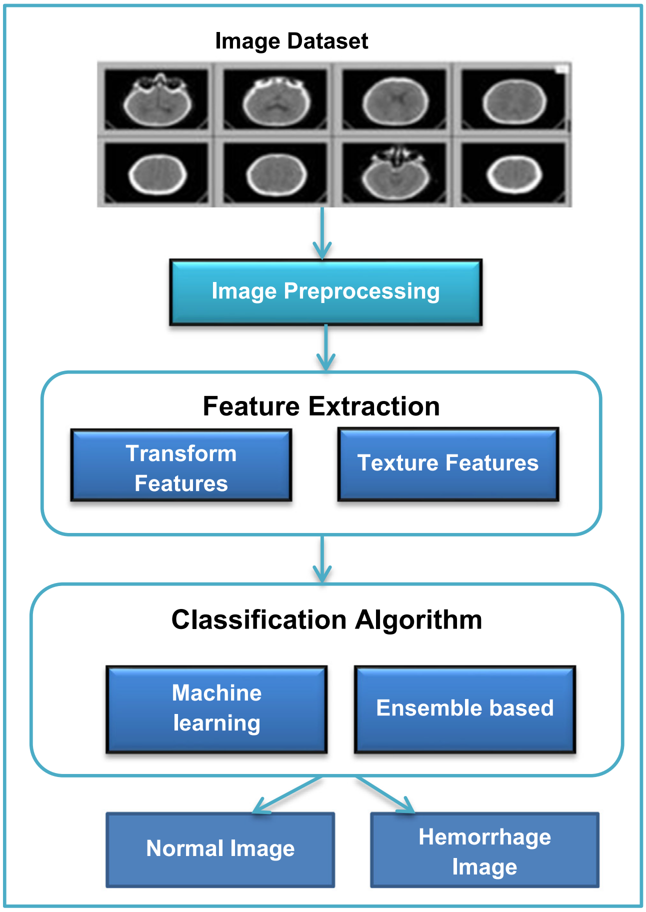

Fig. 2. The proposed Joint feature extraction classification model.

for feature extraction combines a transform-based and texture-based approach. Images are categorized using the GLCM method based on texture. The frequency coefficients of the images are obtained using transform-based methods. As a result, getting images containing important information from the energy coefficients is straightforward. These two methods are combined to create a number of joint feature vector sets that can be used to quickly and accurately classify images using machine learning models. Tuning the hyperparameters is an important part of controlling how a machine learning model works. The goal of hyperparameter tuning is to find the best set of hyperparameter values for a learning method and then use this set of values on a dataset. This set of hyperparameters makes the model work as well as possible by minimizing a fixed loss function to get better results with fewer mistakes. In this work, before a model is trained, hyperparameters are chosen. The models are trained with random hyperparameters and combinations. This makes machine learning work better and be more accurate. The classification algorithms experimented for the proposed work are GNB, LR, KNN, SVM, DT (entropy), RF, Gradient boost, and AdaBoost machine learning classifiers.

(A) Transform based image features

Cosine and wavelet transforms are used to construct the transform characteristics of images. The first two-dimensional cosine transform is used to turn a preprocessed input image into a matrix of high-frequency coefficients and low-frequency coefficients. Additionally, while creating feature vectors, these DCT coefficients are taken into account. The DCT-transformed feature vector is created by performing further processing on DCT-transformed images. The converted image's mean and standard deviation are thought to be its DCT-affected features. Additional discrete wavelet transforms are applied for the construction of DWT

Table 1
Joint feature set for feature selection.

Different feature selection methods
Cosine with GLCM
(high and low Frequency Coefficients and GLCM features)
Wavelet with GLCM
(Approximation , detailed coefficient and GLCM features)
Cosine
Wavelet
GLCM

coefficients. The feature vector that results from the DWT transformation of an image undergoes additional processing. Images' mean and standard deviation are thought to be represented via DWT-transform attributes.

(B) Texture based Features

Characteristics that are image-based make use of GLCM characteristics. Homogeneity, contrast, adjustment, and energy are aspects that are taken into consideration.

(C) Joint Feature set

In this part, a joint approach to feature selection is covered, including features with an energy coefficient and features with a texture. For the formation of a collection of feature sets, cosine, wavelet transforms (DCT, DWT), and GLCM features are employed. And finally, in order to classify normal and abnormal images, these feature vectors are passed to classifiers. The joint technique for feature selection is shown in Table 1. The cosine and GLCM features make up the first joint method under consideration. Wavelet and GLCM features are combined in the second consideration. Algorithm 1 for joint feature selection is described below. After joint feature selection, feature vectors are given to wrapper sequential forward feature selection (SFS) to find important features. Fig. 3 shows the SFS approach for GLCM, DCT, DWT, and joint approaches.

Algorithm 1: Joint feature selection approach

Procedure Joint feature selection (N, s, F, MLCF, V) (N= Dataset having s CT images, F= no. of features, MLCF= machine learning classifier set, V= feature vector)

i. Extract Joint feature set (s)
ii. Put s through the Gaussian filter.
iii. Obtained a transform image features by using cosine and wavelet transform
iv. Derive the high-frequency and low-frequency coefficients. Mean and standard deviation of transform image used as a features
v. Acquired GLCM features, namely one that measures contrast, correction, energy, and homogeneity
vi. Implement the joint feature selection method and produce feature vectors.
vii. Apply wrapper (SFS approach) to find the key features
viii. Apply the classifier
ix. Return MLCF with best accuracy measure Accuracy.

(D) Classification

The proposed work is evaluated using logistic regression, K-nearest neighbor, Gaussian naive Bayes, support vector machine, decision tree (entropy) machine learning algorithm along with Random Forest, AdaBoost, and Gradient boost ensemble based machine learning algorithm. The goal of ensemble learning, a general Meta approach to machine learning, is to improve the performance of predictions by mixing predictions from various models. In proposed work, most popular ensemble based RF, AdaBoost and Gradient boost are evaluated. AdaBoost is a boosting ensemble method where straightforward decision trees are constructed consecutively, in contrast to Random Forest. In case of Gradient Boosting have a similar workflow with AdaBoost except rather than using sample weight to direct the next Decision Stump, uses the Decision Tree's residual to guide the next tree [43].

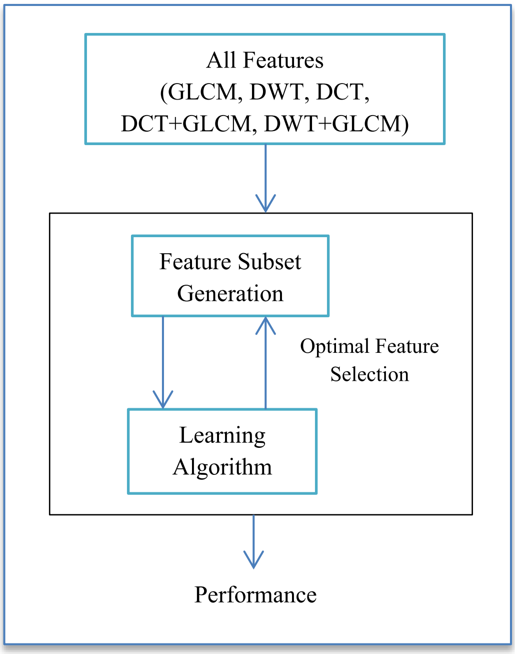

Fig. 3. SFS selection approach for optimal feature selection.

Table 2
Cosine transform based classification accuracy.

Classifiers	Training accuracy	Testing accuracy	Recall	Precision
Logistic Regression	56.22	58.94	85.47	58.86
Gaussian Naive Bayes	57.94	60.76	81.84	60.78
K-nearest neighbor	61.23	57.48	68.31	60.17
Support vector machine	54.58	55.29	100	55.29
Decision Tree	100	57.66	56.43	63.09
Random Forest	100	65.51	66.66	69.65
Bagging	100	64.23	66.66	68.01
AdaBoost	64.44	59.85	63.36	63.78
Gradient boost	64.44	59.85	63.36	63.78

Table 3
Wavelet Transform based classification accuracy.

Classifiers	Training accuracy	Testing accuracy	Recall	Precision
Logistic Regression	61.00	56.38	67.00	58.52
Gaussian Naive Bayes	52.62	50.00	31.31	57.05
K-nearest neighbor	66.32	59.12	71.71	66.33
Support vector machine	62.09	57.29	65.31	59.59
Decision Tree	100	73.17	75.08	75.33
Random Forest	100	77.37	80.47	78.36
Bagging	100	77.37	80.80	78.17
AdaBoost	67.26	60.03	71.38	61.27
Gradient boost	79.40	67.33	75.75	67.77

5. Experimental results

Transform and texture-based feature extraction is performed using cosine, GLCM, and wavelet transforms. A combined feature-based approach of the cosine transform with GLCM and the wavelet transform with GLCM is generated. For the proposed joint feature selection method, extracted features are sent to classifiers, and classification performance for the training and testing classes is observed. The use of an ensemble-based algorithm has also been investigated for additional classification. The research evaluates the GLCM against the cosine and wavelet transform. Performance is measured with matrices that include training accuracy, assessment accuracy, precision, and recall [14,44]. Performance has been assessed for all classifiers, taking various feature combinations into account in order to acquire the key feature. The paragraph below has a discussion of the results from various methodologies.

5.1. Experimentation 1 — Cosine Transform -based classification

Several machine learning and ensemble-based classifiers have used discrete cosine transform-based feature extraction and classification. Table 2 shows the classification accuracy for the different classifiers evaluated. When compared to other classifiers, RF classifiers perform best in terms of accuracy (both for the testing and training classes), precision (65.51%), recall (66.66%), and overall quality of predictions (68.01%). Fig. 4 shows the cosine transform based testing accuracy plot for different classifiers. Utilizing the forward selection wrapper strategy, an important optimal feature set is obtained. The classification performance was evaluated using a variety of features.

Primarily, RF exhibits the highest classification accuracy among all classifiers; the forward feature selection strategy is assessed for the RF classifier, taking into account several optimal feature sets. In discrete cosine transform-based feature extraction, high energy coefficients are considered.

The classification effectiveness of a RF classifier has been assessed with regard to different feature sizes. At first, there was only one feature taken into consideration when constructing a feature vector. Subsequently, one by one, features were added to the feature vector set, and after each addition, the correctness of the entire feature set was evaluated. 53.58% for a single feature (entropy), 57.30% for 2 features (entropy and standard deviation), and 65.60% for 3 features (entropy, standard deviation, and mean). All three features are significant, and the result demonstrates strong performance in terms of classification accuracy.

5.2. Experimentation 2 — Wavelet Transform based classification

Classification accuracy, precision, and recall are calculated for discrete wavelet transform features for different machine learning techniques. Table 3 shows the different classifiers evaluations. The feature vector is the result of computing the mean, standard deviation, and entropy of the estimated detail coefficient matrix. We use different machine learning models to figure out how well DWT-based methods for extracting features work. RF and bagging ensemble-based classifiers show better results in terms of how well they classify training data, which is 100% accurate, and testing data, which is 77.37% accurate. Fig. 5 shows the wavelet transform based testing accuracy plot for different classifiers.

Precision for RF is 80.47 percent and recall is 78.33 percent. Similarly, the accuracy obtained for the bagging classifier is 80.80& and the recall is 78.17%. Given that RF outperforms on other classifiers, the forward feature selection method is assessed for RF classifiers.

Single-feature (entropy) accuracy is 60.59%, two-feature accuracy is 66.44%, and three-feature accuracy is 77.42%. Strong performance in terms of classification accuracy is shown, indicating the importance of all three characteristics.

5.3. Experimentation 3 — Gray-level co-occurrence matrix based classification

For proposed system, performance of different classifiers is measured by using the GLCM to create features. The evaluation of classifiers is shown in Table 4. RF and bagging ensemble-based classifiers demonstrate improved performance across the board in terms of recall, accuracy, and precision. Fig. 6 shows the GLCM based testing accuracy plot for different classifiers. These improvements are shown across all classifiers. Results for the RF and bagging classifiers were 98.35% during training, 81.20% during testing, 83.27% during precision, and 83.00% and 82.78% during recall, respectively. The second-order features are utilized to create the feature vector. Classification accuracy is

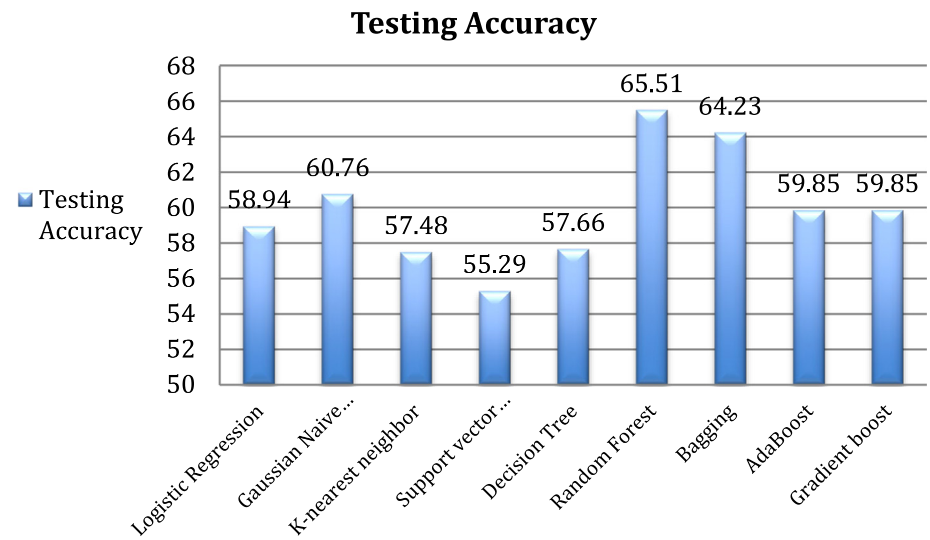

Fig. 4. Cosine transform based testing accuracy plot for different classifiers.

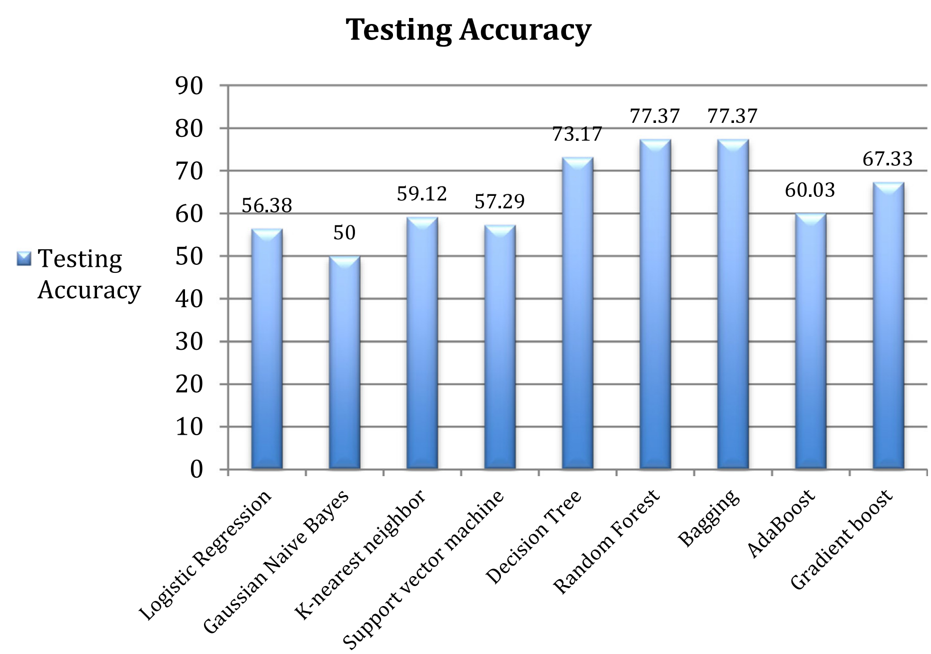

Fig. 5. Wavelet transform based testing accuracy plot for different classifiers.

examined using forward feature selection for second-order feature sets. The findings obtained for a single feature (contrast) are 64.76 percent accurate; for two features (contrast and energy), they are 74.2 percent accurate; for three features (contrast, energy, and homogeneity), they are 77.79 percent accurate; and for four features (contrast, energy, homogeneity, and correlation), they are 80.1 percent accurate.

5.4. *Experimentation 4 — Proposed Joint feature set of cosine and GLCM classification*

Cosine transforms and GLCM-based classification are done with different machine-learning classifiers. The proposed fusion comb set combines DCT and GLCM features. For several machine learning and ensemble-based classifiers, classification performance is computed in addition to accuracy and recall.

The classification accuracy with training and testing data, precision, and recall using different classifiers are all shown in Table 5. In all classifiers, ensemble-based bagging classifiers do better in terms of accuracy for testing and training classes, precision, and recall, which are 100%, 84.40%, 88.15%, and 86.45%, respectively. Fig. 7 shows the Joint feature set of cosine transform and GLCM based testing accuracy plot for different classifiers. RF classifiers are better than bagging classifiers in terms of accuracy, precision, and recall, with

Table 4
Classification accuracy for Gray level co-occurrence matrix based features.

Classifiers	Training accuracy	Testing accuracy	Recall	Precision
Logistic Regression	54.65	54.74	98.03	55.26
Gaussian Naive Bayes	51.99	51.27	31.14	62.50
K-nearest neighbor	85.66	75.00	77.70	77.45
Support vector machine	54.42	55.65	100	55.65
Decision Tree	98.51	75.36	77.70	77.96
Random Forest	98.35	81.20	83.27	83.00
Bagging	98.35	81.20	83.27	82.78
AdaBoost	66.09	63.13	64.26	67.82
Gradient boost	79.48	72.62	72.45	77.00

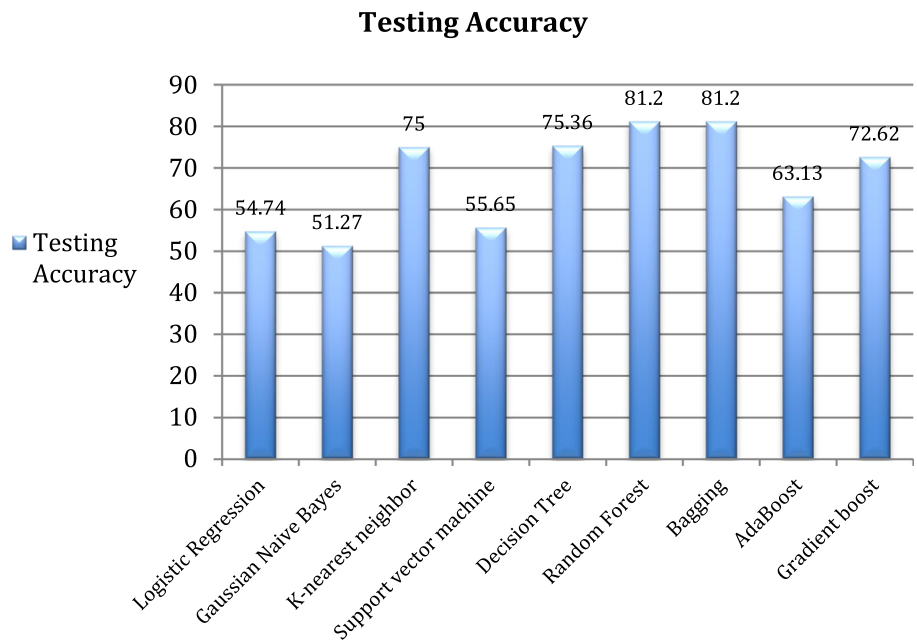

Fig. 6. GLCM based testing accuracy plot for different classifiers.

Table 5
Proposed Joint feature set of cosine transform and GLCM Classification accuracy.

Classifiers	Training accuracy	Testing accuracy	Recall	Precision
Logistic Regression	58.33	55.65	70.39	58.31
Gaussian Naive Bayes	52.62	51.64	31.57	62.74
K-nearest neighbor	67.50	59.12	73.68	60.86
Support vector machine	57.71	55.47	69.07	58.33
Decision Tree	100	76.09	77.30	79.12
Random Forest	100	84.89	87.34	85.24
Bagging	100	85.40	88.15	86.45
AdaBoost	69.69	63.32	76.64	64.18
Gradient boost	85.35	76.27	78.94	78.43

Table 6
Proposed Joint feature set of wavelet transform and GLCM Classification accuracy.

Classifiers	Training accuracy	Testing accuracy	Recall	Precision
Logistic Regression	59.90	58.21	79.72	57.71
Gaussian Naive Bayes	57.71	58.02	51.89	62.65
K-nearest neighbor	61.55	57.11	69.41	58.04
Support vector machine	57.08	57.11	94.50	55.66
Decision Tree	100	82.29	85.56	81.90
Random Forest	100	87.22	91.40	85.53
Bagging	100	87.02	92.09	85.07
AdaBoost	64.68	66.24	79.03	64.97
Gradient boost	82.22	74.08	81.09	73.06

values of 100%, 84.15, 87.34%, and 85.24%, respectively, for testing and training classes, precision, and recall. DCT-based features and GLCM-based features, for a total of seven feature values, are used to form the feature vector. For RF classifiers with different sizes of feature vectors, the forward feature selection method is tested to see how well it works. Results are 64.69% for one feature (entropy), 74.46% for two features (entropy and standard deviation), 79.09% for three features (entropy, standard deviation, and mean), and 81.50% for four features (contrast, entropy, standard deviation, and mean). 84.93 percent for the first five features (contrast, energy, entropy, standard deviation, and mean), 85.01% for the next six (energy, entropy, contrast, homogeneity, standard deviation, and mean), and 84.89% for the final seven features (energy, entropy, contrast, homogeneity, correlation, standard deviation, and mean).

5.5. *Experimentation 4 — Proposed Joint feature set of Wavelet transform and GLCM classification*

The GLCM and wavelet transform features are used to figure out how well different classification methods work. Table 6 shows the precision, recall, testing data, and accuracy of classification for each machine learning classifier that was looked at. According to the results, ensemble-based classifiers perform better. Fig. 8 shows the accuracy plot for testing data with different classifiers.

In comparison to all other classifiers, RF ensemble-based classifiers perform better in terms of accuracy for testing and training classes, precision, and recall, with respective values of 100%, 87.22%, 91.40%, and 85.53%. Applying a forward feature selection strategy while taking different feature vector sizes into account yields additional outcomes. The findings obtained are 64.69% for a single feature (entropy), 74.26% for two features (entropy and standard deviation), 79.82% for three features (entropy, standard deviation, and mean), and 84.70% for four features (contrast, entropy, standard deviation, and mean). Contrast, entropy, energy, standard deviation, and mean are the first five features with an accuracy of 86.1%, followed by 87.22% for six features (energy, entropy, contrast, homogeneity, correlation, and standard deviation), and 86.5% for seven features (energy, contrast, entropy, homogeneity, correlation, standard deviation, and mean). Considerable characteristics include energy, entropy, homogeneity, correlation, contrast, standard deviation, and mean.

6. Discussion

For the proposed system, we explore cosine and wavelet-based features as well as GLCM-based features. Classifiers are contrasted using performance matrices for accuracy, precision, and recall. The efficiency of a categorization model employing transform features is first evaluated, with attention paid to the discrete cosine and discrete wavelet transforms. The low-frequency and high-frequency coefficients, the estimate coefficient matrix, and the detail coefficient matrix are all used to make the feature vector. As a result, analysis shows that features based on discrete wavelet transforms have better classification accuracy than other features. In addition, categorization model accuracy is measured using a GLCM feature. For optimal performance, GLCM-based

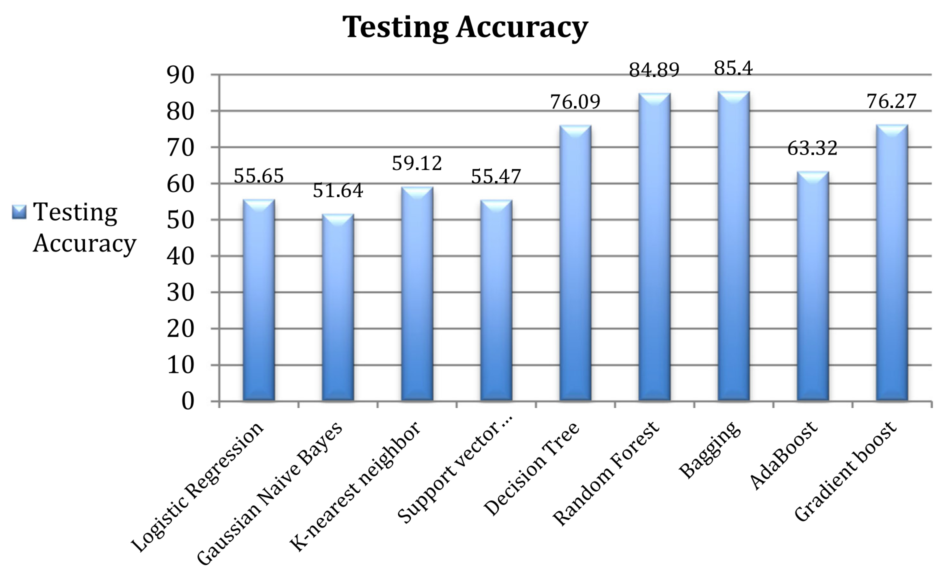

Fig. 7. Joint feature set of cosine transform and GLCM based testing accuracy plot for different classifiers.

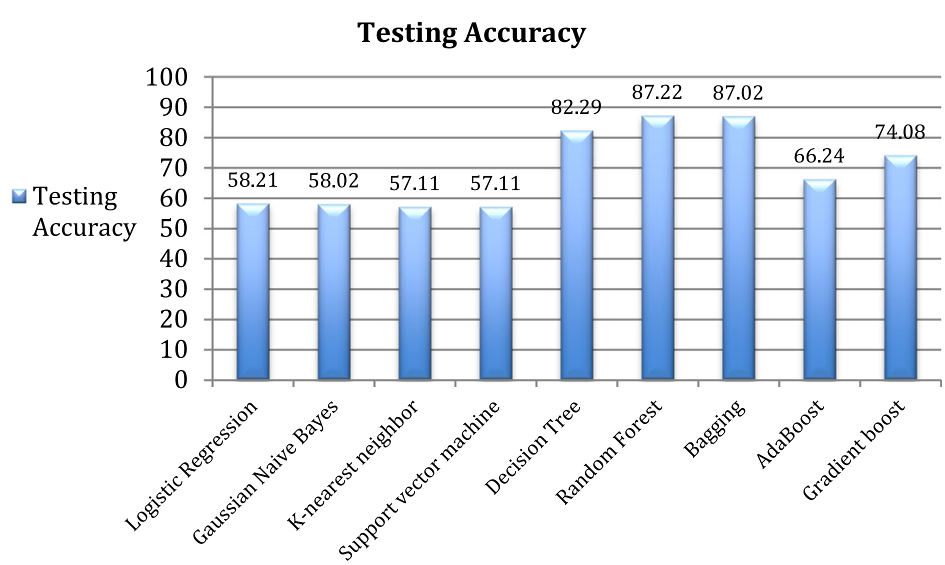

Fig. 8. Joint feature set of wavelet transform and GLCM based testing accuracy plot for different classifiers.

feature extraction techniques are recommended over transform-based ones. The classification accuracy of ensemble-based machine learning methods has been shown to be better than that of traditional ML algorithms. With respect to the accuracy, precision, and recall indices for classification, RF outperformed every ensemble-based model. In the case of the combined cosine and wavelet transforms using GLCM features. The effectiveness of a joint feature group in classification has been tested experimentally. Compared to other feature sets, the results indicate that the wavelet transform using the GLCM set achieves the best results. To obtain meaningful characteristics, we combine multiple factors, including the high-frequency coefficients as well as the estimate and detail coefficients, as well as the values of the GLCM. Classification models were tested using forward feature selection with varying feature sets and feature sizes. The primary objective is to evaluate the RF classifier's forward feature selection strategy, taking into consideration all feature set categories. A RF classifier has performed well when compared to other techniques of classification. Table 7 compares the accuracy of the suggested feature extraction methods for the RF classifier. Important contributions worth highlighting include the following:

- A new method for classifying intracranial hemorrhage CT scans based on machine learning is suggested.
- Proposed method showed accuracy improvement as compared to only transform based or image based features.
- Highest accuracy observed is 87.22% for wavelet transform and GLCM feature set for RF classifier with contrast, energy, homogeneity, entropy, standard deviation and mean features.
- Ensemble based models have shown improved performance as compared to standard machine learning algorithms.

Table 7
Classification accuracy for RF considering proposed feature set.

Feature group	Accuracy
Cosine with GLCM	84.15
Wavelet with GLCM	87.22
Cosine	65.51
Wavelet	77.37
GLCM	81.20

7. Conclusion and future work

In this work, intracranial hemorrhage CT images were classified using different machine learning classifiers. Texture-based GLCM and transform features are combined to create a joint feature-based method. The proposed joint feature approach helps to classify CT images more accurately than the texture-based and transform-based approaches. The transform-based feature technique relies on CT image energy coefficients. The wavelet transform compresses images more effectively than the cosine transform, resulting in optimal feature extraction. It is examined how accurate the joint feature-based categorization is using cosine with GLCM and wavelet with GLCM. The highest classification accuracy for the RF classifier is 87.22 percent, according to the wavelet transform with the GLCM joint feature group. After that, cosine with GLCM is able to reach the highest possible classification accuracy of 85.40 percent for bagging ensemble-based classifiers. When compared to GLCM and transform feature approaches, the joint features-based technique performs better, according to the outcome analysis. Experimentation shows that RF is able to achieve better accuracy. Future joint feature sets for sizable datasets can add more intricate feature groups.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Evie Marcolini, Christoph Stretz, Kyle M. DeWitt, Intracranial hemorrhage and intracranial hypertension, Emerg. Med. Clin. 37 (3) (2019) 529–544.
[2] S.C. Juan, C. Steren, S. Carolina, et al., Convolutional neural networks for detection of intracranial hemorrhage in CT images, in: Proceedings of the 4th Congress on Robotics and Neuroscience, Valparaíso, Chile, 2020, p. 2564.
[3] Hasan Ayaz, Meltem Izzetoglu, et al., Early diagnosis of traumatic intracranial hematomas, J. Biomed. Opt. 24 (5) (2019) 1–10.
[4] R. Badenes, F. Bilotta, Neuro critical care for intracranial haemorrhage: A systematic review of recent studies, Br. J. Anaesth. 115 (2) (2015) 1168–1174.
[5] Bruce Lee, Andrew Newberg, Neuroimaging in traumatic brain imaging, NeuroRx: J. Am. Soc. Exp. NeuroTherapeutics 2 (2) (2005).
[6] Syed Irfan Ullah Nazish, Abdus Salam, Wajid Ullah, Muhammad Imad, COVID-19 lung image classification based on logistic regression and support vector machine, in: 2021, Artificial Intelligence Systems and the Internet of Things in the Digital Era, 239, ISBN: 978-3-030-77245-1, 2021.
[7] Daniel T. Larose, Chantal D. Larose, Discovering Knowledge in Data: An Introduction to Data Mining, Wiley-Interscience, Hoboken, NJ, USA, 2004.
[8] Hajer Kamel, Dhahir Abdulah, Jamal M. Al-Tuwaijari, Cancer classification using Gaussian Naive Bayes algorithm, in: International Engineering Conference, IEC, Erbil, Iraq, 2019, pp. 165–170.
[9] D.V. Kecman, Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models, MIT Press, Cambridge, MA, USA, 2001.
[10] A.M. Anter, M. Ali, Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems, Soft Comput. 24 (3) (2020) 1565–1584.
[11] Agata Sage, Pawel Badura, Intracranial hemorrhage detection in head CT using double-branch convolutional neural network, support vector machine, and random forest, Appl. Sci. 10 (21) (2020) 1–12.
[12] Mandeep Kaur, Sanjay Kadam, Naeem Hannoon, Multi-level parallel scheduling of dependent-tasks using graph-partitioning and hybrid approaches over edge-cloud, Soft Comput. 26 (2022) 5347–5362.
[13] Pritika Bahad, Preeti Saxena, Study of AdaBoost and gradient boosting algorithms for predictive analytics, in: International Conference on Intelligent Computing and Smart Communication, 2019, pp. 235–244.
[14] Santwana S. Gudadhe, Anuradha D. Thakare, Diego Oliva, Classification of intracranial hemorrhage CT images based on texture analysis using ensemble-based machine learning algorithms: A comparative study, Biomed. Signal Process. Control. 84 (2023).
[15] A. Humeau-Heurtier, Texture feature extraction methods: A survey, in: IEEE Access, 7 (2019) 8975–9000.
[16] Namita Aggarwal, R.K. Agrawal, First and second order statistics features for classification of magnetic resonance brain images, J. Signal Inf. Process. 3 (2012) 146–153.
[17] A. Padma Nanthagopal, R. Sukanesh, SVM based classification of soft tissues in brain CT images using wavelet based dominant gray level run length texture features, Middle East J. Sci. Res. 13 (2013) 883–888.
[18] B.A. Mohammed, E.M. Senan, et al., Multi-method diagnosis of CT images for rapid detection of intracranial hemorrhages based on deep and hybrid learning, Electronics 11 (15) (2022) 2460.
[19] Anh-Cang Phan, Van-Quyen Vo, Thuong-Cang Phan, Automatic detection and classification of brain hemorrhages, intelligent information and database systems, in: ACIIDS 2018, in: Lecture Notes in Computer Science, vol. 10752, 2018, pp. 417–427.
[20] V. Vidhya, U. Raghavendra, A. Gudigar, et al., Automated intracranial hematoma classification in Traumatic Brain Injury (TBI) patients using meta-heuristic optimization techniques, Informatics 9 (1:4) (2022) 1–14.
[21] V. Vidhya, Anjan Gudigar, U. Raghavendra, et al., Automated detection and screening of Traumatic Brain Injury (TBI) using computed tomography images: A comprehensive review and future perspectives, Int. J. Environ. Res. Public Health. 18 (12) (2021) 6499.
[22] Rajini N. Hema, R. Bhavani, Automatic classification of computed tomography brain images using ANN and K-NN and SVM, AI & Soc. 29 (2013) 97–102.
[23] Yong-Hong Li, Liang Zhang, et al., Automatic subarachnoid space segmentation and hemorrhage detection in clinical head CT scans, Int. J. Comput. Assist. Radiol. Surg. 7 (2012) 507–516.
[24] Nita Kakhandaki, S.B. Kulkarni, Identification of normal and abnormal brain hemorrhage on magnetic resonance images, in: Cognitive Informatics, Computer Modeling, and Cognitive Science, Academic Press, 2020, pp. 71–91.
[25] R. Anand, T. Shanthi, others Sabeenian, GLCM feature-based texture image classification using support vector machine, in: 3rd EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing, in: EAI/Springer Innovations in Communication and Computing, Springer, 2022, pp. 143–154.
[26] V. Anitha, S. Murugavalli, Brain tumor classification based on clustered discrete cosine transform in compressed domain, J. Comput. Sci. 10 (2014) 1908–1916.
[27] Yudong Zhang, Shuihua Wang, Genlin Ji, et al., An MR brain images classifier system via particle swarm optimization and kernel support vector machine, Sci. World J. 2013 (2013) 1–9.
[28] Y.H. Foo, J.H.D. Wong, et al., Identification of acute intracranial bleed on computed tomography using computer aided detection, J. Phys. Conf. 1497 (1) (2020) 012–019.
[29] J. Nawabi, H. Kniep, S. Elsayed, C. Friedrich, P. Sporns, et al., Imaging-based outcome prediction of acute intracerebral hemorrhage, Transl/. Stroke Res. 12 (6) (2021) 958–967.
[30] U. Raghavendra, The-Hanh Pham, Anjan Gudigar, V. Vidhya, B. Nageswara Rao, Sukanta Sabut, Joel Koh En Wei, Edward J. Ciaccio, U. Rajendra Acharya, Novel and accurate non—linear index for the automated detection of haemorrhagic brain stroke using CT images, Complex Intell. Syst. 7 (2021) 929–940.
[31] 2022. https://www.kaggle.com/intracranial-brain-hemorrhage-ct-images. (Last Accessed on 20 January 2022).
[32] M. Kaur, S. Sakhare, Kirti Wanjale, Early stroke prediction methods for prevention of strokes, Behav. Neurol. (2022) 1–9.
[33] Ahmed M. Anter, Diego Oliva, et al., AFCM-LSMA: New intelligent model based on Lévy slime mould algorithm and adaptive fuzzy C-means for identification of COVID-19 infection from chest X-ray images, Adv. Eng. Inf. (2021) 49.
[34] Ahmed M. Anter, Gan Huang, et al., A new type of fuzzy-rule-based system with chaotic swarm intelligence for multiclassification of pain perception from fMRI, IEEE Trans. Fuzzy Syst. 28 (6) (2020) 1096–1109.
[35] A.K. Jain, F. Farrokhnia, Unsupervised texture segmentation using gabor filters, IEEE Trans. Syst. Man. Cybern. Syst., Los Angeles, CA, USA (1990) 14–19.
[36] J. Nawabi, H. Kniep, R. Kabiri, et al., Neoplastic and non-neoplastic acute intracerebral hemorrhage in CT brain scans: Machine learning-based prediction using radiomic image features, Front. Neurol. 11 (2020).
[37] Punal M. Arabi, Gayatri Joshi, N. Vamsha Deepa, Performance evaluation of GLCM and pixel intensity matrix for skin texture analysis, Perspect. Sci. 8 (2016) 203–206.
[38] Mohamed Abu ElSoud, Ahmed M. Anter, Computational intelligence optimization algorithm based on meta-heuristic social-spider: Case study on CT liver tumor diagnosis, Int. J. Adv. Comput. Sci. Appl. 7 (4) (2016) 455–475.

[39] Wahlang Imayanmosha, et al., Brain magnetic resonance imaging classification using deep learning architectures with gender and age, Sensors (Basel, Switzerland) 22 (5) (2022) 1766.
[40] S.D. Thepade, S.S. Gudadhe, Palm print identification using fractional coefficient of transformed edge palm images with cosine, Haar and Kekre transform, IEEE Int. Conf. Intell. Comput. Commun. Processing, India (2013) 1232–1236.
[41] B. Furht (Ed.), Discrete cosine transform (DCT), in: Encyclopedia of Multimedia, Springer, Boston, MA, 2006.
[42] Nilakshi Mule, Dipti D. Patil, Mandeep Kaur, A comprehensive survey on investigation techniques of Exhaled Breath (EB) for diagnosis of diseases in human body, Inform. Med. Unlocked 26 (2021) 100715.
[43] 2023. https://towardsdatascience.com/understanding-ensemble-methods-random-forest-adaboost-and-gradient-boosting-in-10-minutes-ca5a1e305af2. (Last Accessed on 20 April 2023).
[44] M. Hossin, M.N. Sulaiman, A review on evaluation metrics for data classification evaluations, Int. J. Data Min. Knowl. Manag. Process 5 (2) (2015) 1–11.

Santwana Gudadhe is a Assistant Professor in Department of Computer Science & Engineering (AI & ML) in Pimpri Chinchwad College of Engineering, Pune India. Santwana is pursuing her Ph.D from Savitribai Phule, Pune University. She has Completed Master of engineering (Computer Science) from Pimpri Chinchwad College of Engineering, Pune and B.E. (Information Technology) degrees from Government college of Engineering, Amravati University. Her area of research is Medical Image Processing and Artificial Intelligence. She published 20+ research papers in reputed Journals and Conferences with indexing in Scopus, SCI, Web of Science, IEEE.

Anuradha Thakare is a Professor in Department of Computer Engineering in Pimpri Chinchwad College of Engineering, Pune India. Anuradha received her Ph.D in Computer Science and Engineering from SGB Amravati University and M.E. degree in Computer Engineering from Savitribai Phule Pune University. She is serving for education and research from last 23 years. Her area of research is Evolutionary Computing, Artificial Intelligence, Machine Learning, Biomedical Engineering, Healthcare Analytics, High Performance Computing etc. She published 100+ research papers in reputed Journals and Conferences with indexing in Scopus, SCI, SCIE, Web of Science, ACM, Pubmed etc. She has authored and edited six books published by CRC Taylor & Francis, IGI Global, Wiley etc. She received Research grants from AICTE-AQIS, QIP-SPPU, BCUD-SPPU Pune, Maharashtra State Commission for Women and National Commission for Women. She worked as reviewer for Journal of International Blood Research, IEEE transactions and other Scopus indexed Journals.

Anuradha is PhD guide in Computer Engineering in SPPU, Pune. She has been a General Chair of IEEE International Conference ICCUBEA 2018 and Advisory member for International Conferences. Delivered 25+ expert talks on Machine Learning, Evolutionary Algorithms, Outcome Based Education etc. she worked with industries like DRDO, NCL etc for research projects. She is working as Subject Chairman for various Computer Engineering subjects under Savitribai Phule Pune University (SPPU). She contributed for SPPU syllabus Content designing and revision.

Ahmad M. Anter is working as an Associate professor, and his area of research is Biomedical Engineering, Image Processing and Analysis, Machine Learning, Metaheuristics and Optimization, Fuzzy and neutrosophic logic, Medical application systems.