Journal of Computer Science

Original Research Paper

Nonnegative Matrix Factorization Features Extraction and Ensemble Methods Classifier for ECG Image Classification

[1]Hussain K. Ibrahim, [1]Shawkat K. Guirguis and [2]Hend A. Elsayed

[1]*Department of Information Technology, Institute of Graduate Studies and Researches, Alexandria University, Egypt*
[2]*Department of Electrical Engineering, Faculty of Engineering, Damanhour University, Damanhour, Egypt*

Article history
Received: 09-01-2025
Revised: 15-01-2025
Accepted: 06-02-2025

Corresponding Author:
Hend A. Elsayed
Department of Electrical Engineering, Faculty of Engineering, Damanhour University, Egypt
Email: hend.alielsayed@yahoo.com

Abstract: This study is a novel proposed classification system for ECG image classification. This proposed system uses nonnegative matrix factorization for feature extraction and ensemble methods for classification, and the results are compared with the features extracted by the principal component analysis, kernel principal component analysis, and independent component analysis. Although algorithms find image classification challenging, humans typically find it easy. Data processing might raise new issues during the decision-making process due to the huge and quick development of computers and information technology. In the subject of image classification, the researchers have encountered various challenges, particularly in identifying the image's best features that can provide a high degree of classification accuracy. The ultimate goal of this research is to optimize the image categorization process. There are six types of ensemble methods classifiers used in the proposed classification system. The six classification methods are the Adaboost M1 "Adaptive Boosting", Bagging Mea-Estimator, Logitboost, Gentleboost, Robustboost and Subspace. The results made for different numbers of feature extraction and As demonstrated by the experimental results, the bagging mea-estimator outperforms other ensemble methods for classification and the performance of the classifiers using features extracted from the nonnegative matrix factorization outperforms the performance using other feature extraction methods such as principal component analysis, kernel principal component analysis and independent component analysis. The performance of the classifiers using these methods is dependent on the number of features used.

Keywords: Nonnegative Matrix Factorization (NMF), Principal Component Analysis (PCA), Kernel Principal Component Analysis (KPCA) and (ICA) Independent Component Analysis, Ensemble Methods Classifier, Features Extraction

Introduction

Both the actual data and a model of the data can be used in classification to assign an object to a class based on how similar it is to earlier examples of other objects. Image classification groups data into categories by examining the numerical characteristics of different image features (Aggarawal, 2015; Vendrow *et al*., 2021). Among the most fundamental issues in computer vision, image classification has not changed. Simultaneously, owing to the widely available Internet, the amount of image data has significantly expanded (Gonzalez, 2009). Its extensive application value is seen in its widespread use in machine intelligence, medical research, criminal investigation and detection, access control systems, video surveillance pattern recognition and image processing (Fang *et al*., 2012). How to achieve computational and memory efficiency in large-scale image classification without sacrificing classification accuracy is a major challenge (Chen *et al*., 2015). Typically, classification algorithms use two processing phases: Training and testing. During the first training phase, features that are characteristic of normal images are isolated, and each classification category (or training class) is given a unique description based on these features. Classifying image features using these feature-space partitions is done in the following testing process (Revathy *et al*., 2015). With inputs from the factoring matrix of nonnegative

Nonnegative Matrix Factorization (NMF), the system suggests using a novel approach to categorize image band methods for Categories of Ensemble Methods for Classification. The outcomes are then compared with inputs from Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Kernel Principal Component Analysis (KPCA).

Materials and Methods

Nonnegative Matrix Factorization (NMF) and Ensembles (Fit ensemble)

Nonnegative Matrix Factorization (NMF)

The dimension-reduction method known as Nonnegative Matrix Factorization (NMF) is predicated on a low-rank feature space approximation. The NMF technique has attracted a lot of interest in the bioinformatics field due to its ability to extract interpretable portions from high-dimensional datasets but provide fewer features overall. Furthermore, statistical independence was successful in a number of clustering and classification tasks (Wu *et al*., 2019; 2021; MacMillan and Wilson, 2017). NMF only permits additive partnerships. Based on this methodology, we may map the low-level image features into the additive combination of latent semantics for the clustering, which has been shown to be closer to how humans perceive and understand the data (Fang *et al*., 2012; Wu *et al*., 2021). Discussion is held regarding the evolution and convergence qualities of hybrid methods that rely on both smoothness and sparsity requirements for the generated nonnegative matrix factors (MacMillan and Wilson, 2017).

Ensembles

Ensembles have been demonstrated to have greater flexibility in the functions they may represent, making them a useful tool for merging less accurate classifiers to create highly accurate ones (Zhang and Ma, 2012; Zhang *et al*., 2014). When there is a considerable degree of model variation, ensembles typically produce superior outcomes. The classifier can concurrently promote individual accuracy and diversity within the ensemble by training in several new regions (Jukic *et al*., 2020). The fit ensemble includes Adaboost m1 "Adaptive Boosting", bagging mea-estimator, GentleBoost, RobustBoost and LogitBoost. The objective is to use the inputs to anticipate the values of the outputs, even though every learning algorithm will tend to fit some issue types better than others and will often have many distinct parameters and configurations to be tweaked before obtaining optimal performance on a dataset. Supervised learning is the name of this exercise, which makes use of the more recent machine learning language. The term "predictors," which we will use interchangeably with "inputs," is frequently used in the statistical literature to refer to the inputs (Stamp, 2017; Freund and Schapire, 1999; Hastie *et al*., 2021).

Supervised learning is the name of this exercise, which makes use of the more recent machine learning language. The term "predictors," which we will use interchangeably with "inputs," is frequently used in the statistical literature to refer to the inputs. (Partial search algorithm) that, in certain cases, can offer a good enough solution to an optimization problem, particularly when dealing with limited computer resources or imprecise or partial data.

The process of creating several copies of a predictor and combining them to create an aggregated predictor is known as bagging predictors. For numerical predictions, the aggregation averages across versions; for class predictions, it uses a plurality vote (Liang *et al*., 2011). Modified training sets for Bagging are created by resampling the original training set, and classifiers built with these training sets are then merged by voting. Combining several versions via arcing or Bagging greatly lowers variation (Breiman, 1998).

Proposed Methods

This section presents the proposed method for how the system makes preprocessing for the images and ensemble methods for classification of images with inputs obtained from factoring matrix of nonnegative matrix factorization (NMF) and compares the output from the Independent Component Analysis (ICA), Kernel Principal Component Analysis (KPCA) and Principal Component Analysis (PCA) with the findings. Figure (1) shows the main steps of the proposed system. This proposed system contains five major steps. The first step is to feed the images into the system. Next, preprocess the images by resizing them according to the measured rate of (256*256). Finally, convert the images from an RGB two-dimensional matrix to a grayscale two-dimensional matrix by the measured rate of (256*256). In the second stage, features are extracted from training and testing data using an NMF method, and the features are then used in the classification. In the third step, the features extracted are input to the fit ensemble classifier. The fit ensembles include Adaboost m1 "Adaptive Boosting", bagging mea-estimator, GentleBoost, RobustBoost, LogitBoost and Subspace.

The Fourth Step consists of a great model depending on the previous steps where learning this model about all features of each image and its correct classification. The fifth and final step is to use this model to measure the performance examination of the testing data accuracy by using the confusion matrix function and the amount of error.

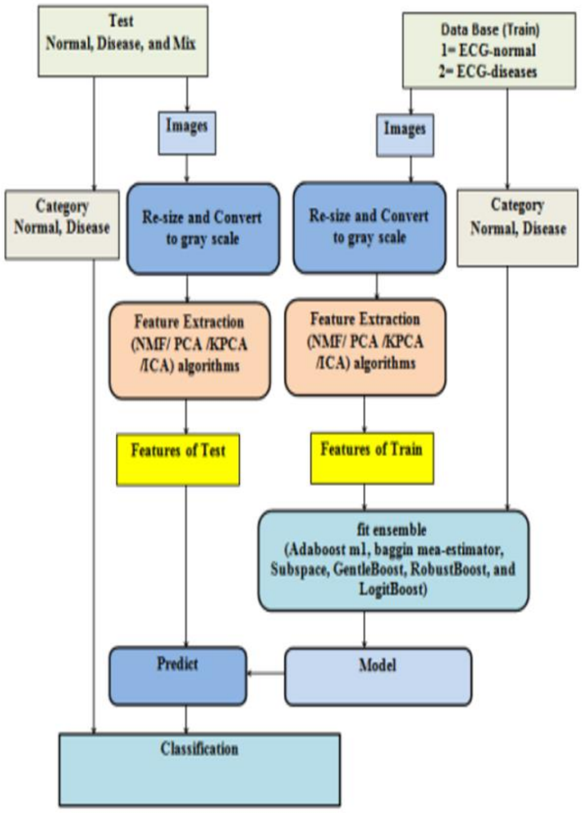

Fig. 1: The main steps for the proposed system

Image Preprocessing

This is the first step of the proposed system. When input into the system, preprocessing will be applied to images, which are resized of the image in accordance, after which the data is transformed from RGB to grayscale at the determined rate of (256*256) two-dimensional matrix (Dual Matrix dimension by the rate of (256*256) (Fig. 2). Resize image down samples an image with a factor of two from 512 by 512-256 by 256 in the vertical and horizontal directions with the aid of Eq. (1):

$$fd(m,n) = f(2m,2n) \tag{1}$$

The sampled image is denoted by $fd(m, n)$, and the original continuous image is represented by $f(x, y)$. While there are methods to convert a full-color image to grayscale using Eq. (2), grayscale algorithms all follow the same fundamental three-step procedure:

1. Obtain a pixel's red, green and blue values
2. To convert those values into a single grey value, use complex math
3. Put the new grey value in lieu of the previous red, green and blue values

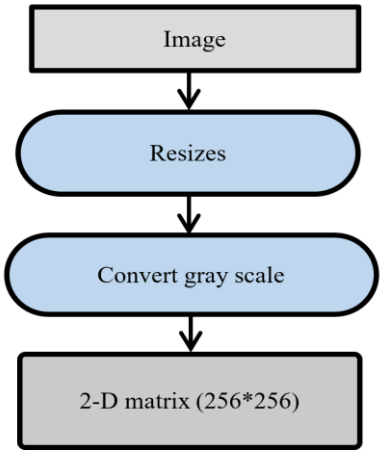

Fig. 2: Image conversion and downsizing

In my explanation of grayscale algorithms, I'll concentrate on step 2: Applying math to convert colour data to grayscale values. Thus, when you come across a formula such as this one:

$$Gray = (Red + Green + Blue)/3 \tag{2}$$

Acknowledge that the real code implementing this kind of algorithm looks like:

- For Every Pixel in the image {
- Red = Pixel.Red Green = Pixel. Green Blue = Pixel.Blue
- Gray = (Red + Green + Blue) / 3
- Pixel.Red = Gray
- Pixel.Green = Gray
- Pixel.Blue = Gray}

Feature Extraction Implementation

This process includes choosing the best features by (Principal component analysis (PCA), Kernel Principal Component Analysis (KPCA), Independent Component Analysis (ICA), and Nonnegative Matrix Factorization (NMF)) After image preprocessing, which restores the image size and then converting to grey, where is the introduction of these features on the third step for training model for the purpose of completing the classification process.

Non Negative Matrix Factorization (NMF)

A class of multivariate analysis and linear algebraic techniques known as nonnegative matrix approximation divide a matrix V into (typically) two matrices, W and H, with the characteristic that none of the three matrices contains a negative element. The resultant matrices are simpler to examine because of this non-negativity. Using the NMF technique to extract features, enter the features with the Function's correct image classification (Fit ensemble (Bagging)) and create a model that includes the proper image classification and features (Fig. 3).

DOI: 10.3844/jcssp.2025.918.927

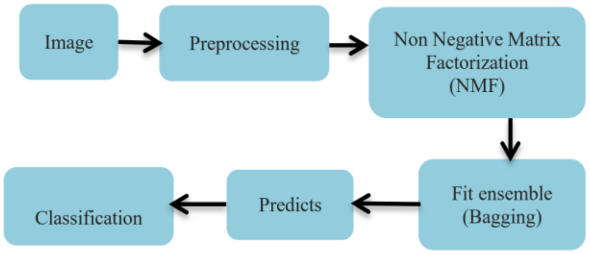

Fig. 3: Extracting features by NMF and completing the classification process

Principal Component Analysis (PCA)

Principle component analysis (PCA) is a statistical technique that turns a series of observations of potentially correlated variables into a set of values of linearly uncorrelated variables known as principal components via an orthogonal transformation. The number of original variables is either equal to or less than the number of primary components. The process of extracting features involves the use of the PCA technique. The features are then entered into the Fit Ensemble (Bagging) Function, which correctly classifies the image. This allows the model to be configured with the features and image classification shown in Fig. (4).

Kernel Principal Component Analysis (KPCA)

Principal Component Analysis (PCA) is expanded upon by Kernel Principal Component Analysis (KPCA), which makes use of kernel approaches. Utilize the KPCA technique to extract features, which will then be entered into the Fit ensemble (Bagging) and training model to ensure accurate picture categorization and feature extraction.

The goal is to extract features using the KPCA technique, enter those features with the Function's correct image classification (Fit ensemble (Bagging)) and create a model that includes the features and correct image classification (Fig. 5).

Independent Component Analysis (ICA)

The process of identifying underlying factors or components from multivariate (multi-dimensional) statistical data is known as independent component analysis or ICA. The way that ICA looks for components that are non-Gaussian and statistically independent sets it apart from other approaches. Use the ICA technique to extract features, then enter the features with the Function's correct image classification (Fit ensemble (Bagging)) and create a model that includes the features and correct image classification (Fig. 6). Using the ICA technique to extract features, enter the features with the Function's correct image classification (Fit ensemble (Bagging)) and create a model that includes the features and correct image classification (Fig. 6).

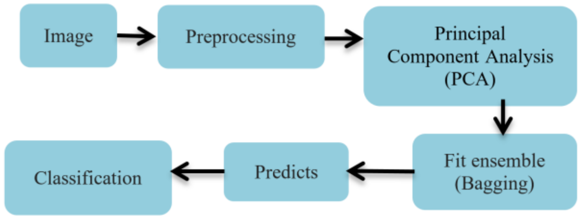

Fig. 4: Extracting features by PCA and completing the classification process

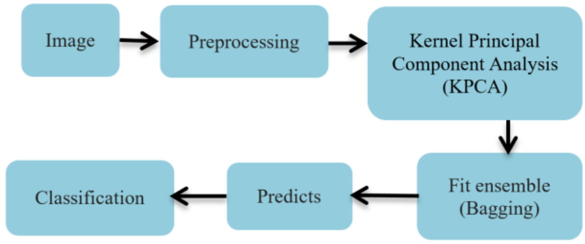

Fig. 5: Extracting features by KPCA and completing the classification process

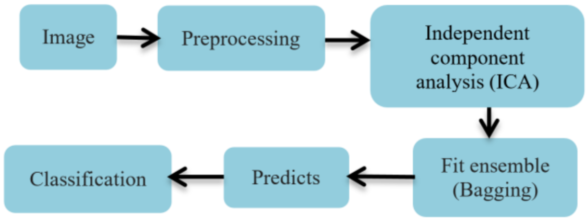

Fig. 6: Extracting features by ICA and completing the classification process

Construction of Classifiers

This is the main step in the proposed system, which is the model building, feature extracted input to the Fit ensemble classifier; Fit ensemble includes Adaboostm1 "Adaptive Boosting", bagging mea-estimator, GentleBoost, RobustBoost and LogitBoost.

Framework for Ensemble

This part of the model, called the classifier, used the ensemble method in this work. An ensemble classifier is a technique that combines two or more classifiers to improve prediction and data voting. Ensemble approaches, which are used in machine learning and statistics, combine several learning algorithms to produce a predicted performance that is higher than that of any one of the individual learning algorithms. A machine learning ensemble, in contrast to a statistical ensemble in statistical mechanics, which is frequently unlimited, refers exclusively to a concrete, finite set of possible models, although it usually permits a far more flexible structure to exist among those options.

Effective ensemble approaches combine the predictions of several individual classifiers, but they are best characterized by the accuracy and diversity of each individual classifier. More accurate predictions should be produced by an accurate classifier than by a random classifier, whereas independent predictions should be produced by a varied classifier. Creating a number of trained models is the first stage in creating an ensemble classifier. Classifiers form the basis of the models. Every base classifier is trained using training data that has been altered in a different way than the original dataset. The basis classifiers are combined in the second phase. An unweighted or weighted vote is used to combine all of their forecasts into a final prediction.

Fit Ensemble Classifier Implementation

The features extracted from one of the algorithms above are input to the (Fit ensemble (Bagging)) and used this image's attributes to train the model for accurate image classification.

Adaboost m1 "Adaptive Boosting"

Another widely used ensemble technique is called boosting, which learns a succession of "weak" classifiers, each one with the goal of improving on the mistakes made by the one before it. Boosting is now among the best generic inductive classification methods. Base classifiers are created by resampling a training data set, much like Bagging. On the other hand, boosting employs an alternative resampling method. In order to train samples of the previous iteration that were mistakenly classified, the method samples a training data set for a base class. By increasing their sensitivity to instances of inaccurate classification, the prediction performance of base classifiers is improved. Adaboost is the most popular boosting implementation.

To train the model for accurate image classification, upload a specific image to the system, resize it to 256 by 256 and convert it to grayscale. Next, the NMF technique is used to extract the image's features, which are then fed into the fit ensemble (Adaboost m1) classifier.

Bagging (Bootstrap Aggregating)

One of the most well-known ensemble techniques is Bagging, a machine learning ensemble meta-algorithm created to increase the stability and accuracy of machine learning algorithms used in statistical classification and regression. Every member classifier in the Bagging algorithm is built using a distinct set of training data. A uniformly random replacement sample is taken from the original data set to create each training data set. When the dimensionality of the feature vector is more than the number of data samples, the feature subset technique performs better and lowers variance. A specific image is entered into the system, resized to (256*256) and then converted to grayscale. Next, features are extracted from the image using the NMF technique, which is then input to a classifier (fit ensemble (Bagging)) and the model is trained using the accurate classification of the image and its features.

LogitBoost

Another popular ensemble technique is called LogitBoost. Resampling a training data set yields base classifiers. But boosting makes use of an alternative resampling technique. In order to train erroneously categorized samples of the previous iteration, the process samples a training data set for a base class. By increasing their sensitivity to cases that are erroneously identified, logit boost improves the prediction performance of base classifiers. A specific image should be entered into the system, resized to 256 by 256 and then converted to grayscale. Next, the image should be extracted using the NMF technique, and this feature should be inputted to a classifier (Fit ensemble (LogitBoost)) so that the model can be trained on the correct classification of the image and its features.

GentleBoost

By increasing their sensitivity to cases that are erroneously classified, a gentle boost increases the prediction performance of base classifiers. A training data set is resampled to create base classifiers, much as bagging. The resampling method used by GentleBoost, on the other hand, is distinct and works well with objects that have more symmetrical and rigid features. Using the NMF technique, extract the image's features and input them into a classifier (Fit ensemble (GentleBoost)) by entering the image in the system, resizing it to 256 by 256 and then converting it to grayscale. Finally, the model will be trained on the accurate classification of the image and its features.

Rusboost

Another popular ensemble method is Rusboost. Base classifiers are created by resampling a training data set, much like Bagging. Nevertheless, Rusboost employs an alternative resampling technique. In order to train samples of the previous iteration that were mistakenly classified, the method samples a training data set for a base class. Rusboost improves the prediction performance of base classifiers by increasing their sensitivity to instances of wrong classification. Rusboost is the most widely used implementation of the Technique; however, some newer algorithms are said to perform better. Enter a specific image into the system, resize it to 256 by 256 and then convert it to grayscale. Next, use the NMF technique to extract the image's features, which you then feed into a classifier (Fit Ensemble (Rusboost)) to train the model on the accurate classification of the image and its features.

DOI: 10.3844/jcssp.2025.918.927

Subspace

Subspace uses a different resampling mechanism. The mechanism samples a training data set for a base class to train incorrectly classified samples of the previous iteration. Subspace prediction performance of base classifiers by making them more sensitive at incorrectly classified instances. The most common implementation of the subspace is that some newer algorithms have been reported to achieve better results.

Enter a specific image into the system, resize it to 256 by 256 and then convert it to grayscale. Next, use the NMF technique to extract the image's features, which you then feed into a classifier (Fit Ensemble (Subspace)) to train the model on the accurate classification of the image and its features.

Predictive Models

Step four consists of a model containing a number of features of each image as well as the correct classification of the image. Through the use of prediction between the model and features of test images. Finally, measure performance accurately through the Function called (Confusion Matrix).

Results and Discussion

The type of data selected is the electrocardiogram (ECG) images, which consist of two types of images: Normal ECG images and disease ECG images for training. The testing that has been selected is for normal diseases or both. The data set used in the experiment is relevant to two categories: One of them is the normal class, and the other is the disease class. Eight experiments were implemented in this work. One hundred and forty images for training are used, whereas the first seventy images are used for normal ECG images, and the second seventy images are used for disease ECG images. Four experiments were conducted to test the accuracy and efficiency of the same training images, and the other four experiments were conducted to test the accuracy and efficiency of the different images from outside the training images, called testing images. Twenty images of them are normal ECG images, and twenty images are disease ECG images.

In the first experiment, the input of one hundred and forty images for training is used, where the first seventy images are used for normal ECG images and the second seventy images for disease ECG images. Selecting all these one hundred and forty images for testing, the NMF has been used for feature extraction, and output has been fed into the fit ensemble using (Adaboost m1), Bagging (mea-estimator), GentleBoost, RobustBoost, LogitBoost and Subspace for classification and compares the accuracy between them using a different number of features. The results are shown in Table (1).

In the second experiment, the input one hundred and forty images for training are used, where the first seventy images are used for normal ECG images and the second seventy images for disease ECG images. Selecting all one hundred and forty images for testing, the NMF has been used for feature extraction, and output has been fed into the fit ensemble using Bagging (mea-estimator) for classification and comparing the accuracy with the results from PCA, KPCA, and ICA uses different numbers of features. These results are shown in Table (2).

Table 1: Classification results using NMF with Adaboost m1, Bagging (mea-estimator), GentleBoost, RobustBoost and LogitBoost for classification using hundred and forty same training images as test images

Features	Performance (Accuracy)					
Number	Adaboost m1	Bag	Logit boost	Gentle boost	Robust boost	Subspace
1	83.5714	100	83.5714	88.5714	67.8571	100
2	100	100	100	100	98.5714	100
3	100	100	100	100	98.5714	100
4	100	100	100	100	98.5714	100
5	100	100	100	100	98.5714	100
10	100	100	100	100	97.1429	100
15	100	100	100	100	95.7143	100
20	100	100	100	100	95.7143	100
25	100	100	100	100	90.7143	100
30	100	100	100	100	90	100
35	100	100	100	100	90.7143	100
40	100	100	100	100	89.2857	100
45	100	100	100	100	82.1429	100
50	100	100	100	100	85.7143	99.2857
55	100	100	100	100	80	100
60	100	100	100	100	82.8571	100
65	100	100	100	100	82.1429	100
70	100	100	100	100	73.5714	100
100	100	100	100	100	79.2857	98.5714

DOI: 10.3844/jcssp.2025.918.927

Table 2: Classification for one hundred and forty ECG images using NMF, PCA, KPCA and ICA with Bagging (mea-estimator)

Features number	Performance (Accuracy)			
	NMF	PCA	KPCA	ICA
1	100	100	100	100
2	100	100	100	100
3	100	100	100	100
4	100	100	100	100
5	100	100	100	100
10	100	100	100	100
15	100	100	100	100
20	100	100	100	100
25	100	100	100	100
30	100	100	100	100
35	100	100	100	100
40	100	100	100	100
45	100	100	100	100
50	100	100	100	100
55	100	100	100	100
60	100	100	100	100
65	100	100	100	100
70	100	100	100	100
100	100	100	100	100

In the third experiment, selecting the seventy normal ECG images was used for testing only, the NMF was used for feature selection, and output was fed into the fit ensemble using Bagging (mea-estimator) for classification and comparing the accuracy with the results from PCA, KPCA and ICA. Table (3) shows the results using different numbers of features.

The fourth experiment uses only the seventy disease ECG images for testing; the NMF has been used for feature selection, and output has been fed into the fit ensemble using Bagging (mea-estimator) for classification and comparing the accuracy using different numbers of features with the results from PCA, KPCA and ICA as shown in Table (4).

In the fifth experiment, the input one hundred and forty images for training are used, where the first seventy images are used for normal ECG images and the second seventy images for disease ECG images. The testing images used forty images, with twenty from the normal ECG images and twenty from the disease ECG images from outside of the training images. All forty images are used for testing, the NMF has been used for feature extraction, and output has been fed into the fit ensemble using (Adaboost m1), Bagging, mea-estimator, GentleBoost, RobustBoost, LogitBoost and Subspace for classification and comparing the accuracy between them using a different number of features. The results are shown in Table (5).

In the sixth experiment, the input of one hundred and forty images for training is used, where the first seventy images are used for normal ECG images and the second seventy images for disease ECG images. The testing images used forty images, with twenty from the normal ECG images and twenty from the disease ECG images from outside of the training images. All These forty images are used for testing; the NMF has been used for feature extraction, and output has been fed into the fit ensemble using Bagging (mea-estimator) for classification and comparing the accuracy with the results from PCA, KPCA and ICA using a different number of features. These results are shown in Table (6).

Table 3: Classification for seventy normal ECG images using NMF, PCA, KPCA and ICA with Bagging (mea-estimator)

Features Number	Performance (Accuracy)			
	NMF	PCA	KPCA	ICA
1	100	100	100	100
2	100	100	100	100
3	100	100	100	100
4	100	100	100	100
5	100	100	100	100
10	100	100	100	100
15	100	100	100	100
20	100	100	100	100
25	100	100	100	100
30	100	100	100	100
35	100	100	100	100
40	100	100	100	100
45	100	100	100	100
50	100	100	100	100
55	100	100	100	100
60	100	100	100	100
65	100	100	100	100
70	100	100	100	100
100	100	100	100	100

Table 4: Classification of seventy disease ECG images using NMF, PCA, KPCA and ICA with Bagging (mea-estimator)

Features Number	Performance (Accuracy)			
	NMF	PCA	KPCA	ICA
1	100	100	100	100
2	100	100	100	100
3	100	100	100	100
4	100	100	100	100
5	100	100	100	100
10	100	100	100	100
15	100	100	100	100
20	100	100	100	100
25	100	100	100	100
30	100	100	100	100
35	100	100	100	100
40	100	100	100	100
45	100	100	100	100
50	100	100	100	100
55	100	100	100	100
60	100	100	100	100
65	100	100	100	100
70	100	100	100	100
100	100	100	100	100

Table 5: Classification result using NMF with Adaboost m1, Bagging (mea-estimator), Gentle Boost, Robust Boost and Logit Boost for classification using forty different images for test

Features	Performance (Accuracy)					
Number	Adaboost m1	Bag	Logit Boost	Gentle Boost	Robust Boost	Subspace
1	80	70	77.5	80	75	70
2	100	100	100	100	100	100
3	100	100	100	100	100	97.5
4	100	100	100	100	100	90
5	50	100	100	100	50	100
10	100	100	100	100	97.5	92.5
15	100	100	100	100	95	90
20	100	100	100	100	87.5	75
25	100	100	100	100	92.5	70
30	100	100	100	100	95	67.5
35	97.5	100	97.5	100	82.5	62.5
40	100	100	100	100	90	65
45	100	100	100	100	82.5	70
50	97.5	100	100	100	67.5	72.5
55	97.5	100	100	100	60	55
60	100	100	100	100	92.5	52.5
65	95	100	92.5	92.5	60	60
70	95	100	97.5	100	70	55
100	97.5	100	100	95	77.5	50

Table 6: Classification of forty ECG images using NMF, PCA, KPCA and ICA with Bagging (mea-estimator)

Features Number	Performance (Accuracy)			
	NMF	PCA	KPCA	ICA
1	70	95	72.5	47.5
2	100	97.5	100	67.5
3	100	100	100	62.5
4	100	100	100	70
5	100	100	100	55
10	100	100	100	42.5
15	100	100	95	75
20	100	100	95	65
25	100	100	100	75
30	100	100	100	60
35	100	100	95	72.5
40	100	100	95	72.5
45	100	100	95	72.5
50	100	100	100	72.5
55	100	100	100	52.5
60	100	100	95	60
65	100	100	95	70
70	100	100	95	80
100	100	95	95	65

In the seventh experiment, the selecting twenty normal ECG images, only used for testing, the NMF was used for feature selection, and output was fed into the fit ensemble using Bagging (mea-estimator) for classification and compares the accuracy with the results from PCA, KPCA and ICA. Table (7) shows the results using different numbers of features.

The eighth experiment uses only the twenty diseases, ECG images for testing, the NMF is used for feature selection, and output has been fed into the fit ensemble using Bagging (mea-estimator) for classification and comparing the accuracy using different numbers of features with the results from PCA, KPCA and ICA as shown in Table (8).

Table 7: Classification for twenty normal ECG images using NMF, PCA, KPCA and ICA with Bagging (mea-estimator)

Features Number	Performance (Accuracy)			
	NMF	PCA	KPCA	ICA
1	70	90	65	60
2	100	100	100	60
3	100	100	100	65
4	100	100	100	60
5	100	100	100	65
10	100	100	100	65
15	100	100	100	45
20	100	100	100	75
25	100	100	100	55
30	100	100	100	65
35	100	100	100	60
40	100	100	100	45
45	100	100	100	70
50	100	100	100	60
55	100	100	100	60
60	100	100	100	55
65	100	100	100	55
70	100	100	100	55
100	100	100	100	40

Table 8: Classification for twenty disease ECG images using NMF, PCA, KPCA and ICA with Bagging (mea-estimator)

Features Number	Performance (Accuracy)			
	NMF	PCA	KPCA	ICA
1	65	80	75	60
2	100	100	100	55
3	100	100	100	60
4	100	100	100	65
5	100	100	100	65
10	100	90	100	45
15	100	100	90	70
20	100	100	100	70
25	100	100	90	70
30	100	100	100	85
35	100	100	100	75
40	100	100	100	75
45	100	100	100	70
50	100	100	100	60
55	100	100	100	65
60	100	100	100	70
65	100	90	100	50
70	90	100	100	70
100	95	90	100	65

Proposed a classification model, Fit Ensemble, for large-scale image classification, which has high efficiency for image processing of every class. Ensembles are a method for the highly accurate classic; The Fit ensemble combines the advantages of both learning-based and Nonnegative Matrix Factorization (NMF) methods. NMF reflects the concept of "parts form an integral" The results of NMF decomposition are interpretable, and Nonnegative Matrix Factorization (NMF) is much better than Technique last (from the Principal Component Analysis (PCA), (ICA) Independent Component Analysis and Kernel Principal Component Analysis (KPCA)). Because NMF is able to extract parts of the exegesis of the high data set dimensions as well as the independence of the statistical, he has an advantage and the ability to reduce the distance between the point and the point of the key data features and get specific information. Increasing the classification accuracy and optimal use (Fit ensemble) because of the use of more than classificatory and combined them to give results much better than working individually where there is more flexibility in jobs that represent a show. The proposed classification scheme was verified using an ECG (Electrocardiogram).

The result of experiments (1, 2, 3) shows when NMF is used for feature selection, the Fit Ensemble (Bagging) classifier increases the classification accuracy by increasing the number of features, especially when choosing the number of features that are equal to the number of images. Through experiments (5, 6, 7), when using the experimental data from outside the training data, it was observed that NMF for feature selection and Fit ensemble (Bagging) classifier noticeably increased the classification accuracy by increasing the number of features especially. Experiments show that the classifier (Bag) gives the best results when using the NMF. Further, we are planning to improve our results with different Classification algorithms to improve the quality of the resultant output.

Conclusion

This study proposed a classification system, Fit Ensemble, for large-scale image classification, which has high efficiency for image processing of every class. Ensembles are a method for the highly accurate classic; The Fit ensemble combines the advantages of both learning-based and Nonnegative Matrix Factorization (NMF) methods. NMF reflects the concept of "parts form an integral" The results of NMF decomposition are interpretable, and Nonnegative Matrix Factorization (NMF) is much better than Technique last (from the Principal Component Analysis (PCA), (ICA) Independent component analysis and Kernel Principal Component Analysis (KPCA)). Because NMF is able to extract parts of the exegesis of the high data set dimensions as well as the independence of the statistical, he has an advantage and the ability to reduce the distance between the point and the point of the key data features and get specific information. Increasing the classification accuracy and optimal use (Fit ensemble) because of the use of more than classificatory and combined them to give results much better than working individually where there is more flexibility in jobs that represent a show. The proposed classification scheme was verified using an Electrocardiogram (ECG). Further, we are planning to improve our results with different Classification algorithms to improve the quality of the resultant output.

Acknowledgement

We would like to thank the Institute of Graduate Studies and Researches at Alexandria University for supporting our studies and work, especially the Department of Information Technology and also thank the reviewers for their comments.

Funding Information

The authors have not received any financial support or funding to report.

Author’s Contributions

Hussain K. Ibrahim: Participated in all experiments, coordinated the data analysis and contributed to the writing of the manuscript.

Shawkat K. Guirguis: Supervision.

Hend A. Elsayed: Participated in all experiments, coordinated the data analysis, contributed to the writing of the manuscript, designed the research plan, and organized the study.

Ethics

This article is original and contains unpublished material. There are no ethical issues involved.

References

Aggarwal, C. C. (2015). *Data Classification: Algorithms and Applications, Boca Raton CRC*.

Breiman, L. (1998). Arcing classifier (with discussion and a rejoinder by the author). *The Annals of Statistics*, *26*(3), 801–849. https://doi.org/10.1214/aos/1024691079

Chen, S., Yang, X., & Tian, Y. (2015). Discriminative Hierarchical K-Means Tree for Large-Scale Image Classification. *IEEE Transactions on Neural Networks and Learning Systems*, *26*(9), 2200–2205. https://doi.org/10.1109/tnnls.2014.2366476

Fang, W., Cheng, Z., & Yang, D. (2012). Face Recognition based on Nonnegative Matrix Factorization and Linear Regression. *International Journal of Advancements in Computing Technology*, *4*(16), 93–100. https://doi.org/10.4156/ijact.vol4.issue16.11

Freund, Y., & Schapire, R. (1999). A Short Introduction to Boosting. *Journal of Japanese Society for Artificial Intelligence*, *14*(5), 771–780.

Gonzalez, R. C. (2009). *Digital Image Processing*.

Hastie, T., Friedman, J., & Tibshirani, R. (2001). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. https://doi.org/10.1007/978-0-387-21606-5

Jukic, S., Saracevic, M., Subasi, A., & Kevric, J. (2020). Comparison of Ensemble Machine Learning Methods for Automated Classification of Focal and Non-Focal Epileptic EEG Signals. *Mathematics*, *8*(9), 1481. https://doi.org/10.3390/math8091481

Liang, G., Zhu, X., & Zhang, C. (2011). An Empirical Study of Bagging Predictors for Different Learning Algorithms. *Proceedings of the AAAI Conference on Artificial Intelligence*, *25*(1), 1802–1803. https://doi.org/10.1609/aaai.v25i1.8026

MacMillan, K., & Wilson, J. D. (2017). Topic supervised nonnegative matrix factorization. *ArXiv:1706.05084*. https://doi.org/10.48550/arXiv.1706.05084

Revathy, N., Srinivasan, H., Arun, B., & Manikandan, K. (2015). Instant fracture detection using air-rays. *International Journal of Computational Engineering Research (IJCER)*, *5*(4), 2250–3005.

Stamp, M. (2017). Boost Your Knowledge of AdaBoost. *San Jose State University*. https://www.cs.sjsu.edu/~stamp/RUA/ada.pdf

Vendrow, J., Haddock, J., Rebrova, E., & Needell, D. (2021). On a Guided Nonnegative Matrix Factorization. *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3265–32369. https://doi.org/10.1109/icassp39728.2021.9413656

Wu, W., Jia, Y., Wang, S., Wang, R., Fan, H., & Kwong, S. (2021). Positive and Negative Label-Driven Nonnegative Matrix Factorization. *IEEE Transactions on Circuits and Systems for Video Technology*, *31*(7), 2698–2710. https://doi.org/10.1109/tcsvt.2020.3027570

Wu, W., Kwong, S., Hou, J., Jia, Y., & Ip, H. H. S. (2019). Simultaneous Dimensionality Reduction and Classification via Dual Embedding Regularized Nonnegative Matrix Factorization. *IEEE Transactions on Image Processing*, *28*(8), 3836–3847. https://doi.org/10.1109/tip.2019.2907054

Zhang, C., & Ma, Y. (2012). *Ensemble Machine Learning: Methods and Applications*. https://doi.org/10.1007/978-1-4419-9326-7

Zhang, Y., Zhang, B., Coenen, F., Xiao, J., & Lu, W. (2014). One-class kernel subspace ensemble for medical image classification. *EURASIP Journal on Advances in Signal Processing*, *2014*(1), 14. https://doi.org/10.1186/1687-6180-2014-17