Performance Evaluation of Hybrid Machine Learning Algorithms for Medical Image Classification

N. T. Renukadevi

Abstract Medical imaging is the process of creating images of parts of human body for diagnosis and treatment purposes. These images are collected from traditional X-ray based methods like Mammography and Computed Tomography (CT). Some advanced sources of images include Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET). Since large volumes of digital medical images are deposited in repository, it is a humongous task to access these voluminous images. To access the required image representation from data store, a technique called Content Based Image Retrieval (CBIR) is currently in use. Content-based image retrieval (CBIR) is an assembly that can overcome the problem mentioned above as it is based on the visual analysis of contents that are part of the query image. CBIR retrieves the images which are needed based on its visual contents. CBIR includes Feature extraction and Feature matching. In feature extraction, information like colour, texture and shape known as feature vectors are retrieved through various extraction methods. Similarly, in feature matching the extracted features are compared between normal and abnormal images for classification. The major challenge in CBIR is implementing flexible methodologies to process the different images of different characteristics like colour, shape and pattern. At the same time, applications for retrieving images for proper indexing is done through Picture Archiving and Communication Systems (PACS). In this chapter, retrieving medical images from different data stores and performance of various machine learning classifiers such as Support Vector Machine (SVM) and Deep Learning methodology are focussed to improve the classification accuracy.

Keywords Support vector machine · Deep learning · Particle swarm optimization · Genetic algorithm · Grasshopper optimization

N. T. Renukadevi (✉)
Department of CT-UG, Kongu Engineering College, Erode, India

S. Dash et al. (eds.), *Advanced Soft Computing Techniques in Data Science, IoT and Cloud Computing*, Studies in Big Data 89, https://doi.org/10.1007/978-3-030-75657-4_12

1 Introduction

In images, features like contrast, brightness, entropy may influence the effectiveness of the entire system. Features like colour, texture, and shape orientation do not match with label manual or human interface. The semantics of images can thus be obtained. Semantic gap occurs due to the variations among these features. In order to make connectivity between the local and global features CBIR faces lot of challenges. Amongst which identifying the required storage area for accumulating the images is the greatest task of CBIR. Next, the performance of the system may also degrade due to high cost incurred in computation of the images.

Medical databases contain digital medical images which includes information about patient's record and image description. These descriptions are utilized for both research and academic processes. Since large volume of digital medical images are created and accumulated in repository, retrieval of particular image is a major risk. Hence, CBIR aids in such retrieval process based on user's query. Previously several researches have been done in CBIR with excellent feature extraction and classification methods like [6, 22]. Research has been done to a large extent to manage particular type of images like mammography, brain tumour [3], lung cancer detection [28] or some blood related diseases [39].

Global features play a vital role in identifying the required images for retrieval [32] when compared to low level features. In this chapter, retrieving medical images from different data bases is focussed.

The key idea of the investigation is to find out the efficiency of machine learning algorithms like Support Vector Machine (SVM) hybridized with optimization techniques such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Deep Learning Techniques.

Further sections focus on:

- Evaluation of the performance of SVM with Radial Basis Function (RBF) Kernels for classifying medical images with Coiflet wavelets and Moment Invariant (MI).
- Implementation of Principal Component Analysis (PCA) for reducing the features of an image.
- Hybrid implementation of optimization algorithms GA and PSO with SVM and evaluating the classification accuracy for multi class medical images.
- Implementation of Deep Learning Technique with Grasshopper Optimization (GOA) algorithm for identification of cancer in liver images.

2 Methodologies

2.1 *Efficiency of SVM-RBF Kernels for Medical Image Classification*

The medicinal information system in current era shows tremendous changes in making verdict of health conditions about a patient with the aid of images from various scanning devices. In these circumstances, CBIR is very helpful for the retrieval of required data based on the query given by the physicians or medical practitioners. This leads to propose a research work with appropriate machine learning technique SVM to classify medical images obtained through CT scans.

Coiflet wavelets, MI methods, PCA and Kernel PCA feature reduction method along with SVM are needed for classifying the images. Even if the images underwent transformations like Rotation, scaling and Translation, they do not change the performance of wavelets and MI methods for feature extraction. But it is really a tedious task in spatial domain because of the natural characteristics of medical images. But wavelet transformation methods are able to accomplish this inadequacy by removing pixel coefficients of high frequency and observing the features of low frequency coefficients at various resolution levels. In order to extract the shape vectors of an image, MI is used since the image differs in shapes.

Coiflet wavelet is used to extract the energy coefficients because of the orthogonal property displayed by this wavelet of X number of coefficients that leads to a shorter filter with one-third of X diminishing instants and one-third of X − 1 scale functions. The wavelet function (Ψ) has twice the X instants leads to 0 and the scale function (φ) has $\Psi - 1$ instant leads to 0. Both Ψ and φ hold up a length of six times the $X - 1$ values.

Let p represents a standard uninterrupted instant signal, coefficient for maximum value of i is $<p, \varphi_{i,k}>$ is computed by $2^{-\frac{i}{2}}\, p\left(2\varphi_k^i\right)$ Where p is a multinomial of degree n, the scale function is equal when $n \leq X - 1$. This characteristic is helpful for finding the variation over the function φ with values i, k for the sample signal given.

The characteristics of images extracted are organized through SVM for their efficacy [9, 19, 37].

2.1.1 Moment Invariants

MI is applied as an input characteristic in various research areas like processing images, identifying diverse shapes and remote sensing. Moments endow with important description of an entity which unambiguously symbolizes its shape. Identification of constant shapes is done by the categorization in n-dimensional feature space. Conventionally, MI is calculated based on the details inferred by marginal and internal area of the shapes. The moments applied to create invariants are calculated as discrete values practically [9, 19, 35].

Let $f(a, b)$ be a function, then the standard instants are explained by [21] in "(1)":

$$MI_{ab} = \iint a^m b^n f(a,b) \partial a \partial b \tag{1}$$

where MI_{ab} represents the 2D instance of the function $f(a, b)$. The instance sequence is defined as $(m + n)$ and m, n here are natural numbers. This can be shown in "(2)" as distinct values like:

$$MI_{m,n} = \sum_a \sum_b a^m b^n f(a,b) \tag{2}$$

The translation process in an image is normalized with the help of central pixels called as centroid. The pixel positions of an image centroid are computed using the following Eqs. "(3–6)":

$$\bar{a} = \frac{MI_{10}}{MI_{\infty}} \tag{3}$$

$$\bar{b} = \frac{MI_{01}}{MI_{\infty}} \tag{4}$$

The centroid positions are interpreted as:

$$\mu_{mn} = \sum_a \sum_b (a - \bar{a})^m (b - b)^n \tag{5}$$

In order to transform the scaling points, the instances are regularized as

$$\eta = \frac{\mu_{mn}}{\mu'_{\infty}} \tag{6}$$

The normalized centroids are calculated as a group of seven moment invariants which are explained as in "(7)" and are independent of rotation transformation [23].

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} + \eta_{02})^2 + 4\eta^2 11 \\
\phi_3 &= (\eta_{30} + \eta_{12})^2 + (\eta_{03} - \eta_{21})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \\
\phi_5 &= (3\eta_{30} - \eta_{12})(\eta_{30} + \eta_{12})^2[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + 3(\eta_{21} + \eta_{03})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\phi_6 &= (\eta_{20} - \eta_{20})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + 3(\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned} \tag{7}$$

2.1.2 Support Vector Machines

The classification technique is employed in several domains like image analysis, text categorization, and bio-informatics. In order to design a good model using SVM, the challenge is the selection of proper features. Limited number of input features in a classifier produces good computational model. The essential task here is to identify right values of parameters along with proper selection of features, which helps in enhancing the performance of SVM in classification. For classification, kernel function should be chosen and kernel parameters and soft margin constant C must be determined. In this work, the kernel function which is used is RBF Kernel function. Thus, the cost parameter and the kernel parameter should be optimized [24, 26].

Let (p_i, q_i) be a training set with $i, j \in 1, 2 \ldots l$ *where* $p_i \in R^n$ and $q_i \in (1, -1)$. The optimization problem is solved by using SVM as given in "(8)" [17]

$$\begin{aligned} &\min\left(\frac{1}{2}\right) we^T we + c\sum_{i=1}^{l} \varepsilon_i \\ &\text{subject to } y_i(we^T \varphi(x_i) + b) \geq 1 - \varepsilon_i \text{ and } \varepsilon_i \geq 0 \end{aligned} \tag{8}$$

where x_i vector is mapped in elevated dimensional space with the function φ. The margin constant c and the positive slack variable ε_i are crucial in reducing the training errors. The value of c should be a positive whole number. To compute the efficient model for SVM with linear separable hyper plane, Lagrangian method is used by maximizing the objective function which is shown in "(9)":

$$\begin{aligned} &Max_{\infty} = L_D(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j < x_i x_j > \\ &\text{subject to : } 0 \leq \alpha_i \leq C \; i = 1, \ldots, m \text{ and } \sum_{i=1}^{m} \alpha_i y_i = 0 \end{aligned} \tag{9}$$

where the constants α_i are called Lagrange multipliers.

The best possible hyper plane is found by maximizing the objective function α_i subject to the constraints $\sum_{i=1}^{m} \alpha_i y_i = 0$ and value of α_i lies between 0 and c. The user should examine the maximum limit on α_i. The product of ordered pair $\alpha(x_i x_j)$ for input data is computed using the kernel function K which is given as "(10)".

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|)^2 \; \textit{where } \gamma > 0 \tag{10}$$

2.2 Feature Reduction using PCA

The most acceptable linear method for reducing the dimensionality of image features is PCA. It increases the variance and reduces the data from high dimension feature to lower one. It reduces all the aspect of features with the support of vector space which works well for even complex data sets. In real datasets, empirical zero mean assumed in PCA may not be possible [5]. To overcome this issue, a modification of the PCA kernel is proposed. The data $x_1, x_2, \ldots x_m \in X$ is given and is assumed as a vector area. The principal components $\varphi(x_i), \ldots, \varphi(x_n)$ have been figured out by using Kernel PCA. In general, a problem in PCA should be altered with the requisites of kernel.

2.2.1 Kernel PCA

For a non-empty set N and a non-negative specific kernel k, a function $(N \times \mathrm{N}) \epsilon$ R has a characteristic that a function $\varphi : n \to H$ is mapped for all $n, n' \in N$ with $(n, n') = k(n, n')$. K can be viewed as a non-linear similarity measure for the kernel methods [12].

At the end, covariance matrix of H will be written as follows in "(11–15)" even in the case of unrestricted dimensions also,

$$CM : \frac{1}{n} \sum_{i=1}^{n} \varphi(x_i)^T \tag{11}$$

where $\varphi(x_i)^T$ is the sequential representation of mapping a vector v to the function

$$\sum_{i=1}^{n} \varphi(x_i)^T \text{ and } Cv = \tau v \tag{12}$$

with $\tau \neq 0$ reclined in the distance of φ training images. Hence

$$v = \sum_{i}^{n} \alpha_i \varphi(x_i)^T \tag{13}$$

and this reduces the problem of finding α_i. And α_i is represented in the way of

$$n\tau\alpha = K\alpha_i \tag{14}$$

where $\alpha = (\alpha_i^T)$ and $K = k(x_p, x_q)$. Taking into consideration of these factors in eigen value τN, the zth extractor of the feature will be

$$(v^z, \varphi(x)) = \frac{1}{\sqrt{\tau^z}} \sum_{i=1}^{n} \alpha_i^z k(x_p, x_q) \tag{15}$$

And formulated by finding the multiplication of sample data point $\varphi(x)$ and the zth eigen vector in search space, then $\frac{1}{\sqrt{\tau^z}}$ makes sure that the ordered pair $(v^z, v^z) = 1$.

Thus, KPCA derives zth feature values which is comparatively equal to the coefficient values α_n^z.

Thus, from the Eqs. "(11–15)", a conclusion is drawn that there will be zero mean values for a data in feature space which is shown in "(16)"

$$M_n = \frac{1}{n} \sum_{i} \varphi(x_i) \tag{16}$$

The difference between mean value Mn and all data points are calculated, which leads to a dissimilar eigen value and its diagonal values can be used as in "(17)"

$$K = (I - ee^T)K(I - ee^T) \text{with } e \text{ as } \frac{1}{n}(T) \text{ rather than } k. \tag{17}$$

2.3 Hybrid Algorithm Based on PSO, GA with Local Search

Among the different amalgamation of parameters, the suitable parameters are selected and retrieved through computations which will be applied for the needed dataset. In order to automate this process, several research work have been carried over on retrieval and optimization techniques for images [20]. And, in continuation to the review, this chapter focuses on hybridizing PSO and GA with various combinations of parameters C and γ, in SVM RBF Kernel function.

GA is an optimization method involved in finding the best possible solutions of search problems. It is a nature inspired evolutionary algorithm to find the optimized solution using heuristics. It produces the probable results for optimization problems by working with small genetic chromosome like data configuration. GA includes stages like selection, crossover, mutation which are applied over these data structures to maintain crucial information [17].

PSO works is based on the communal behaviour of bees or bird congregates/fish groups. It is uncomplicated to implement and produces the results easily. Here, every unique element i denotes mixture of parameters representing the location of element i in exploration area. The velocity of a particle or element shows the route of searching the food items and keep informed the location and velocity of element at regular intervals of iteration. This leads in locating the most excellent area in search space [20]. The location and speed of a particle are computed as given in the Eqs. "(18, 19)"

$$s_i^{di+1} = w.s_i^{di} + ct_1.rnd_1(p_i^{di} - x_i^{di}) + ct_2.rnd_2(p_g^{di} - x_i^{di}) \tag{18}$$

$$x_i^{di+1} = x_i^{di} + s_i^{di} \tag{19}$$

where w is mass weight; di characterizes the number of iterations; i represents the extent of inhabitants; the two "most excellent" values—p_i^{di} means the finest way out accomplished by particle up to the current iteration and p_g^{di} means the finest way out accomplished by any of the particles in the group of inhabitants and will be shared among them. The parameters ct_1 and ct_2 are non-negative constants which evaluate the significance of comprehensive learning of the flock. Also, rnd_1 and rnd_2 are arbitrary values with a range of 0 and 1.

The evaluation of SVM along with RBF kernel depends on two constraint values such as *C and* γ. The values of C and γ are inversely proportional to each other since the increase in value of γ leads to better accuracy and at the same time produce unfair result. Similarly, the increase of C leads to poor accuracy but produce low bias result. Thus, SVM is influenced by these two parameters [4, 17]. In order to make a model of SVM to fit with the training data, optimized values of *C and* γ should be utilized. Such optimization can be done with the support of PSO which looks for the best combination of those parameters with the intention to reduce Root Mean Squared Error (RMSE).

In PSO, the efficiency of optimization is influenced by the parameters *w,* ct_1, ct_2, rnd_1 and rnd_2 [1, 7, 14]. The weight parameter w manages the searching process and it lies between 0 and 1 so that the elements can congregate. While w increases nearer to 1, it may lead to global exploration while it decreases to 0 from 0.5 it leads to local exploration. The parameters rnd_1 and rnd_2 lies in the range from 0 to 1. The values of ct_1, ct_2 are almost equivalent to each other and they lay in between 0 and 4. The Fig. 1 explores the PSO optimization with SVM.

Rarely these parameter values may have sudden convergence speed, and in order to prevent from such impulsive convergence the PSO algorithm is customized with GA. Since GA already has several phases like selection, crossover and mutation, these phases coordinate with constraints in PSO to improve the performance and to reduce the immature convergence of the particles. In GA, the individual inputs are encoded using binary encoding with two points cross over rate and random mutation with Roulette wheel selection. As GA reaches the end, those parameter coefficients are modified and given as an input to PSO algorithm.

In this moment, GA and PSO are hybridized to become an evolutionary technique as it prevents the impulsive convergence of elements in population. GA does extremely well in routing problem and PSO is excellent in fuzzy control systems and neural networks due to its grouping characteristics. Here, both C and γ are optimized to develop inhabitants based on these optimization algorithms. Also, these exhibits effective appearance while evaluation. During each step, the particles are split into two groups and evaluated with GA and PSO individually. Later, all the particles are pooled into a new population and the previous step is done for next iteration. The iteration continues until the best solution is reached [2, 21, 41].

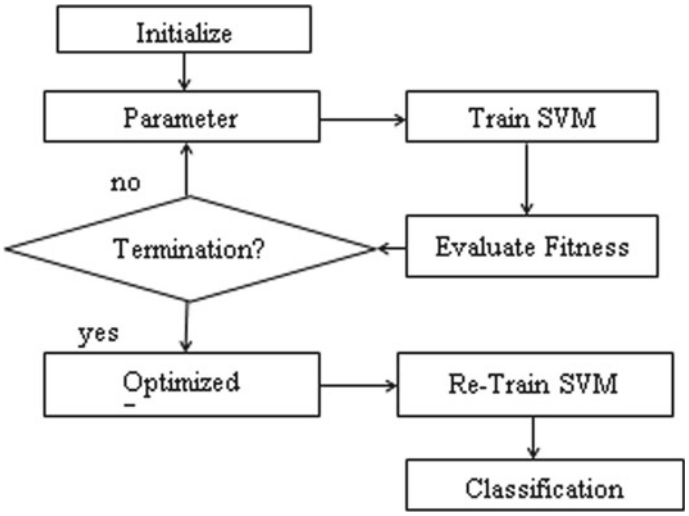

Fig. 1 Flowchart—PSO optimizing SVM

Local Search is the foundation of several heuristics' techniques used to solve hybridized optimization problems. It is a method with frequent repetition of steps that helps to identify solution with good estimation of parameter coefficients. It includes the search space with the parameters as nodes of a graph and objective function as edges. The local search scans the search area by traversing the nodes down the edges.

The important factor to be considered here is the conditions followed by local search algorithm. The conditions are (i) a solitary sequence of data items apart from group-based particles; and (ii) explore the enhanced results in minimized search area defined by self-learning approaches. In general, local search is a repeatedly applied heuristic method whichever optimization algorithm can utilize it [30].

2.4 Deep Learning

Even though SVM is good in classification of medical images both linearly and non-linearly, finding appropriate kernel for complex datasets consumes long time for training. In order to overcome these conditions, Deep Learning technique, which is a compartment of Machine Learning, is being used. It is an inspiration of Neural Networks which is similar to the functioning of human brain and works well on prediction.

A neural network possesses layers like input layer, an intermediary hidden layer, and the last output layer, where each layer consists of "nodes". The medical images collected from different modalities like CT, Mammogram, MRI, PET and Ultrasound

are converted to pixels or picture elements which are keyed into the nodes in input layer. Some weights along with input values are fed into hidden layer for computation, and finally the prediction outcome can be obtained from output layer.

Deep learning automatically learns to generate and reduce the pixel characteristics of images, and, if needed, create new features. In addition to this, deep learning method predicts the occurrences of diseases with statistical measures and benefits the practitioners to make decision at the earliest [24, 27]. This chapter also includes an analysis of wide-ranging applications in the field of medical image classification.

2.4.1 Applications of Deep Learning in Medical Image Analysis

Massive computational power and challenges in the investigation of medical imaging is accomplished all the way through deep learning. In the recent years, image detection, segmentation and disease classification are highly involved with deep learning. CNN plays a dynamic role along with selected features among the different deep learning methods.

Clinical practice is improved through deep learning and the illustration is increasing on a daily basis. Deep learning is applied in treatment of disease in the form of radiation [31]. Scanning of images are done through PET or MRI to get detailed pictures of anatomy of body [11, 25, 29], in methods that extracts large volume of features from radiography [13, 34], and in the field which combines diagnostic test with therapy in neurosurgical imaging [10, 15].

When deep learning is applied in healthcare industries, it provides preferable solutions to a variety of problems like diagnosis of diseases, suggestions for personalized treatments etc. A good amount of data is generated through various methods of radiological imaging. But still there are undersized amount of significant data which are needed to be included by means of deep learning model.

2.4.2 Deep Learning in Liver Disease Identification

Nowadays, the frequency and death rate of liver disease are greater than ever. Liver is infected by many types of diseases like Cirrhosis, the scar in liver; Hepatitis, the inflammation in cells; Cholestasis, the obstruction of bile flows, etc., Identifying the disease through CT images is done effectively with the help of deep learning algorithm. In this work, Deep Learning Neural Network DBN incorporated with Grasshopper optimization algorithm is applied for the classification of livers from diseased one [38].

Initially, pre-processing work is done with feature extraction and feature reduction process. For extraction of texture features, Gabor filter along with Local binary Pattern (LBP) is used. Gabor filter detects edge based on frequency values in the image around the point of analysis. LBP works by classifying the pixels based on threshold value of neighbourhood pixels and the output is a binary value. Colour features can be extracted using measures like mean, variance, skewness and kurtosis.

Feature reduction process takes place using PCA which is explained above [36]. Gray Level Co-occurrence Matrix (GLCM) feature extraction method represents the number of incidences of grey intensity values in an image. Features like contrast, entropy, homogeneity, etc., were extracted [36, 38, 42]. Mean value of each region is calculated by using intensity value as expressed in Eq. "(20)"

$$Mean = \frac{1}{NXP} \sum_{A=1}^{N} \sum_{B=1}^{P} f(A, B) \tag{20}$$

Contrast is the variation of colours in an object that makes it distinguishable from others. Equation "(21)" is of the form

$$Con = \sum_{A=1}^{n} \sum_{B=1}^{n} (A - B)^2 C_{AB} \tag{21}$$

where *A, B* represents the luminance and intensity of an object and C represents the Co-occurrence matrix values. If the value of difference of *A* from B is 0 then there will be no contrast, if the difference increases, then the contrast also increases exponentially.

Entropy is the random distribution of intensity values, and when the same values are repeated for certain patterns, entropy is said to be uniform as mentioned in Eq. "(22)". If the entropy is low, then randomness is also low.

$$Entr = \sum_{i}^{n} \sum_{j}^{n} p_{ij} \log p_{ij} \tag{22}$$

where p_i is the intensity value of pixel in a grey scale image.

Homogeneity shows nearness of the pixels of an image as represented in Eq. "(23)"

$$Homo = \sum_{p,q=1}^{n} H_{pq} \left(\frac{1}{1 + (p - q)^2} \right) \tag{23}$$

After the feature extraction and reduction processes, the liver images undergo classification process using Grasshopper Algorithm based Deep Belief Network (DBN). It is a new model generated from Restricted Boltzmann Machines (RBM). Here, the nodes in each layer are connected to previous and subsequent layers, and are used to extract the features of images and classify it [18, 40].

DBN works by taking probability values as input and produces outcome by implementing unsupervised learning algorithm. DBN consists of nodes in input layer, hidden layer and output layer. The input layer includes $I \in \{0, 1\}^a$ with parameter values *m* of any real numbers. The hidden layer contains of $J \in \{0, 1\}^b$ with parameter values *n* of any real numbers. And, *I, J* represent the count of input and hidden nodes

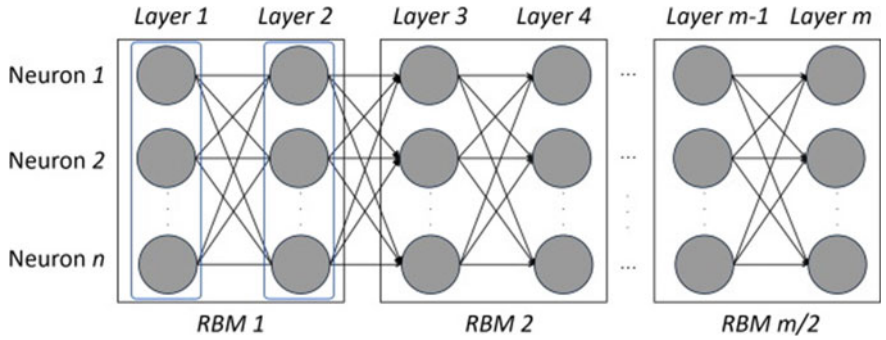

Fig. 2 Architectural diagram of DBN

with weight *Wab*. The main point to note here is, there is no relationship between nodes in a layer on its own. The energy can be obtained by training the parameters *m, n, W*. The formula for computing energy is given below as Eq. "(24)". The architectural diagram of DBN is also depicted in Fig. 2 [8, 18].

$$E(I, J) = -\sum_{a}^{x} m_a I_a - \sum_{b}^{x} n_b J_b - \sum_{i}^{x} \sum_{j}^{x} I_a W_{ab} J_b \tag{24}$$

At first, the initial weights are identified by unsupervised pre-training process in DBN which shows error detection and good optimization results. The best possible values of count of layers and nodes are selected based on the data set type and still there is a deficient in global optimum value. This research work overcomes this problem by implementing Grasshopper Optimization Algorithm (GOA).

GOA is an innovative and efficient optimization algorithm that reduces the movement of grasshoppers to solve optimization problem. Generally, the insects fly as a group and each one has a meticulous space to others and their progress is calculated as in Eq. "(25)"

$$X_i = C_i + F_i + AS_i \tag{25}$$

where X_i is the position of *i*th grasshopper and C_i the shared communication, F the force of attraction, and AS_i the air stream advection on the *i*th grasshopper respectively [16]. The random behaviour of the equation can be represented in Eq. "(26)" and $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$ are the random numbers with weight between 0 and 1 [33]

$$X_i = C_i r_i + F_i r_2 + AS_i r_3 \tag{26}$$

GOA includes the following steps in Algorithm 1 [16, 38]:

1. Initializing the random population of RBM parameters with minimum, maximum values and no. of iterations. 2. For each iteration:
a. Compute the fitness solution of each grasshopper g_i b. Renew the location of existing search particle. c. Standardize the distance between grasshoppers d. Update the position of grasshopper g_i with respect to others e. If g_i *is* beyond the limit, correct its position
3. If the solution reaches, update the area of attraction.
Algorithm 1. GOA with DBN

GOA does better than PSO and Genetic Algorithm in speed of convergence and can solve complex optimization techniques with less time for good optimum solution [33].

3 Discussion of Experimental Results

Experiments were carried out with 750 images of 3 classes of CT scan medical images like colon, brain and liver and feature extraction was through use of Coiflet wavelet and MI. The images were acquired from National Biomedical Imaging Archive (https://imaging.nci.nih.gov/ncia). Experiments were conducted for 10-fold cross validations, SVM-RBF algorithm with different C and gamma parameters.

Features from Coiflet and MI were combined with the anticipated artifact imperative fusion procedure after obtaining the Median Absolute Deviation (MAD) among the two features.

Let Ci = {*ci, 1, ci, 2, ci, n*} be the Coiflet coefficient.

Let Mi = {*mi, 1, mi, 2,, mi, n*} be the MI coefficients.

The feature vectors are fused by normalizing the feature vector to obtain C*i* and M*i* using Median Absolute Deviation and taking the average of the same.

The formulae for classification accuracy, specificity, sensitivity and f-measure are given as follows in Eqs. "(27–30)":

$$CAcc = \frac{TPos + TNeg}{TPos + FPos + FNeg + TNeg} \tag{27}$$

$$Spec = \frac{TPos}{TPos + FNeg} \tag{28}$$

Table 1 Classification, accuracy of SVM with Coiflet and MI

Techniques used	Accuracy (%)		
	Coiflet	MI	Coiflet and MI
SVM-RBF with Cost C = 1, $\gamma = 0.1$	83.40	78.791	78.75
SVM-RBF with Cost C = 1, $\gamma = 0.05$	81.85	76.46	81.75
SVM-RBF with Cost C = 0.5, $\gamma = 0.5$	81.86	82.84	82.39

$$Sen = \frac{TPos}{TPos + FPos} \tag{29}$$

$$fscr = \frac{2 * Spec * Sen}{Sen + Spec} \tag{30}$$

where

True Negative (*TNeg*)—Count of relevant estimation that an occurrence is untrue
False Positive (*FPos*)—Count of irrelevant estimations that an occurrence is well-founded
False Negative (*FNeg*)—Number of irrelevant estimations that an occurrence is untrue
True Positive (*TPos*)—Number of relevant estimations that an occurrence is well-founded.

Tables 1 and 2 shows the feature fusion of Coiflet with MI for the SVM classifier produces different accuracies for various cost and gamma functions. Since Cost and gamma functions were selected randomly, the performance of the SVM classifier is not optimal. Hence, work needs to be done in the direction of feature reduction and SVM kernel optimization.

Tables 3 and 4 shows the performance of the classifier as it improved significantly due to feature selection as seen from the experimental results. Comparing these results with previous results when no feature reduction techniques were used, it is observed that the Kernel PCA improves the classification accuracy of the classifiers in the range of 2.00% as given in Table 5. Since there is considerable difference in the classification accuracy of the SVM classifier for various cost and gamma parameter, further investigation needs to be carried out to identify the ideal parameters for the RBF kernel.

From the above results, it is shown that Local Search along with SVM-GA-PSO gives better results when compared to others. Also, Kernel PCA gives classification accuracy of about 1.93% higher compared to PCA of the same method. Similarly, DBN with optimization algorithms like GA, PSO, and GOA are experimented. The classification accuracy calculated for 125 liver images alone and the results are tabulated in Table 6.

Table 2 Precision, Recall and F-Measures of SVM with Coiflet and MI

Methods	Precision			Recall			F-Measures		
	Coiflet	MI	Coiflet and MI	Coiflet	MI	Coiflet and MI	Coiflet	MI	Coiflet and MI
SVM-RBF C = 1, $\gamma = 0.1$	0.834	0.812	0.826	0.8332	0.7891	0.8125	0.8314	0.7688	0.8462
SVM-RBF C = 1, $\gamma = 0.05$	0.8325	0.7652	0.8323	0.8547	0.7653	0.8127	0.8334	0.7856	0.8224
SVM-RBF C = 0.5, $\gamma = 0.5$	0.8405	0.8234	0.8146	0.8236	0.8412	0.8432	0.8165	0.8218	0.8167

Table 3 Classification accuracy of SVM using KPCA

Methods	Classification accuracy (%)	
	PCA	KPCA
SVM-RBF with Cost C = 1, $\gamma = 0.1$	84.57	86.57
SVM-RBF with Cost C = 1, $\gamma = 0.05$	84.83	85.90
SVM-RBF with Cost C = 0.5, $\gamma = 0.5$	82.40	84.56

Table 4 Precision, Recall and F-Measures of SVM using KPCA

Methods	Precision		Recall		F-Measure	
	PCA	Kernel PCA	PCA	Kernel PCA	PCA	Kernel PCA
SVM-RBF Kernel C = 1, $\gamma = 0.1$	0.8652	0.8753	0.8432	0.8552	0.8543	0.8626
SVM-RBF Kernel C = 1, $\gamma = 0.05$	0.8234	0.8256	0.8569	0.8589	0.8385	0.8369
SVM-RBF Kernel C = 0.5, $\gamma = 0.5$	0.8434	0.8465	0.8456	0.8498	0.8435	0.8467

Table 5 Accuracy of SVM with optimization techniques

Methods	Classification accuracy (%)	
	PCA	Kernel PCA
SVM-GA	85.82	87.34
SVM-PSO	86.85	90.18
SVM-GA-PSO	90.56	91.54
SVM GAPSO with local search	92.62	94.53

Table 6 Accuracy of DBN with optimization techniques

Methods	Classification accuracy (%)
DBN-GA	95.25
DBN-PSO	94.36
DBN-GOA	97.38

4 Conclusion

In this chapter, medical images were retrieved through feature extraction techniques like Coiflet and MI. The features are reduced by PCA and Kernel PCA techniques and classified by SVM with RBF kernel functions. Since the parameters *C and Gamma* are randomly chosen, the classification accuracy is little poor in performance.

Hence, in order to optimize the values of *C and Gamma,* optimization algorithms are used such as GA and PSO. Experiments are evaluated with SVM and optimization algorithms individually and also as hybridized one. Results show that SVM with GA and PSO along with local search are also calculated. And, according to the current technological improvement in Deep Learning, DBN is also tried with these optimizers and with Grasshopper Optimization algorithm which shows comparatively better performance in classification.

References

1. Ahmed, M.: A survey on dynamic clustering based colour image segmentation using genetic algorithm. World J. Sci. Technol. **1**(12), 35–41 (2012)
2. Ali, A.F., Tawhid, M.A.: A hybrid particle swarm optimization and genetic algorithm with population partitioning for large scale optimization problems. Ain Shams Eng. J. (Ain Shams University) **8**(2), 191–206 (2017)
3. Arakeri, M.P., Ram Mohana Reddy, G.: An intelligent content-based image retrieval system for clinical decision support in brain tumor diagnosis. Int. J. Multimed. Inf. Retr. **2**(3), 175–188 (2013)
4. Ardjani, F., Sadouni, K., Benyettou, M.: Optimization of SVM multiclass by particle swarm (PSO-SVM. In: 2010 2nd International Workshop on Database Technology and Applications, DBTA2010—Proceedings, (December), pp. 32–38 (2010)
5. Avalhais, L.P., da Silva, S.F., Rodrigues, Jr., J.T.: Image retrieval employing genetic dissimilarity weighting and feature space transformation functions. In: Twenty-Seventh Annual ACM Symposium on Applied Computing, pp. 1012–1017 (2012)
6. Bhende, P., Cheeran, P.A.N.: Content Based Image Retrieval in Medical Imaging, pp. 10–15 (2013)
7. Cho, M.Y., Hoang, T.T.: Feature selection and parameters optimization of SVM using particle swarm optimization for fault classification in power distribution systems. Comput. Intell. Neurosci. (2017)
8. Dai, X., et al.: Deep belief network for feature extraction of urban artificial targets. Math. Probl. Eng. **2020**, 1–13 (2020)
9. Dixit, A., Majumdar, S.: Comparative analysis of coiflet and daubechies wavelet using global threshold for image de-noising. Int. J. Adv. Eng. Technol. **6**(5), 2247–2252 (2013)
10. Gulshan, V., et al.: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA, J. Am. Med. Assoc. **316**(22), 2402–2410 (2016)
11. Guo, L., Jiang, Q., Jin, X., Liu, L., Zhou, W., Yao, S., Min, W.: A deep convolutional neural network to improve the prediction of protein secondary structure. Curr. Bioinform. **15**, 1–11 (2020)
12. Ham, J.: A kernel view of the dimensionality reduction of manifolds. In: Twenty-first International Conference on Machine Learning, p. 47 (2004)
13. Han, W., et al.: Deep transfer learning and radiomics feature prediction of survival of patients with high-grade gliomas. Am. J. Neuroradiol. **41**(1), 40–48 (2020)
14. Hitam, N.A., Ismail, A.R., Saeed, F.: An optimized support vector machine (SVM) based on particle swarm optimization (PSO) for cryptocurrency forecasting. Procedia Comput. Sci. (Elsevier B.V.) **163**, 427–433 (2019)
15. Izadyyazdanabadi, M., Belykh, E., Mooney, M.A., Eschbacher, J.M., Nakaji, P., Yang, Y., Preul, M.C.: Prospects for theranostics in neurosurgical imaging: empowering confocal laser endomicroscopy diagnostics via deep learning. Front. Oncol. **8** (2018)

16. Jia, H., et al.: Hybrid grasshopper optimization algorithm and differential evolution for multilevel satellite image segmentation. Remote Sens. **11**(9) (2019)
17. Jrad, N., et al.: Sw-SVM: sensor weighting support vector machines for EEG-based brain-computer interfaces. J. Neural Eng. **8**(5) (2011)
18. Kamada, S., Ichimura, T., Harada, T.: Knowledge extraction of adaptive structural learning of deep belief network for medical examination data. Int. J. Semant. Comput. **13**(1), 67–86 (2019)
19. Kaur, S., Kaur, G., Singh, D.D.: Comparative analysis of haar and coiflet wavelets using discrete wavelet transform in digital image compression. Int. J. Eng. Res. Appl. **3**(3), 1–6 (2013)
20. Kaya, D.: Optimization of SVM parameters with hybrid CS-PSO algorithms for Parkinson's Disease in LabVIEW environment. Parkinson's Disease (2019)
21. Keyes, L., Winstanley, A.: Using moment invariants for classifying shapes on large-scale maps. Comput. Environ. Urban Syst. **25**(1), 119–130 (2001)
22. Kumar, A., et al.: Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data. J. Digit. Imaging **26**(6), 1025–1039 (2013)
23. Li, D.: Analysis of moment invariants on image scaling and rotation. In: Innovations in Computing Sciences and Software Engineering, pp. 415–419 (2010)
24. Litjens, G., et al.: A survey on deep learning in medical image analysis. Med. Image Anal. (Elsevier B.V.) **42**(Dec 2012), 60–88 (2017)
25. Liu, F., et al.: Technical developments. Deep learning Mr imaging-based attenuation correction for PeT/Mr imaging. Radiology **286**(2) (2018)
26. Lorena, A.C., De Carvalho, A.C.P.L.F.: Evolutionary tuning of SVM parameter values in multiclass problems. Neurocomputing **71**(16–18), 3326–3334 (2008)
27. Lundervold, A.S., Lundervold, A.: An overview of deep learning in medical imaging focusing on MRI. Zeitschrift fur Medizinische Physik. Elsevier B.V. **29**(2), 102–127 (2019)
28. Malviya, N., Choudhary, N., Jain, K.: Content Based Medical Image Retrieval and Clustering Based Segmentation to Diagnose Lung Cancer, vol. 10, no. 6, pp. 1577–1594 (2017)
29. Mehranian, A., Arabi, H., Zaidi, H.: Vision 20/20: magnetic resonance imaging-guided attenuation correction in PET/MRI: challenges, solutions, and opportunities. Med. Phys. **43**(3), 1130–1155 (2016)
30. Mesloub, S., Mansour, A.: Hybrid PSO and GA for global maximization. Int. J. Open Probl. Comput. Sci. Math. **2**(4), 597–608 (2009)
31. Meyer, P., et al.: Survey on Deep Learning for Radiotherapy. Computers in Biology and Medicine. Elsevier Ltd. (2018)
32. Nalini, P., Malleswari, B.L.: Local versus global features for medical image retrieval. Int. J. Mod. Trends Eng. Res. **4**(10), 1–6 (2017)
33. Nandal, D., Sangwan, O.P.: Bat algorithm, particle swarm optimization and grasshopper algorithm: a conceptual comparison. Int. J. Res. Anal. Rev. **5**(2), 2116–2120 (2018)
34. Oakden-Rayner, L.: Precision radiology: predicting longevity using feature engineering and deep learning methods in a radiomics framework. Sci. Rep. **7**(1) (2017)
35. Renukadevi, N.T., College, K.E.: Performance Analysis of Coiflet Wavelet and Moment Invariant Feature Extraction for CT Image Classification using SVM, vol. 2 no. 12, pp. 1–6 (2013)
36. Renukadevi, N.T., Karunakaran, S., Saraswathi, K.: Gray level cooccurrence matrix feature extraction and fuzzy based discriminative binary descriptor for medical image retrieval. Int. J. Comput. Sci. Eng. **7**(6), 62–70 (2019)
37. Renukadevi, N.T., Thangaraj, P.: Performance Evaluation of SVM-RBF Kernel for Medical Image Classification, vol. 13, no. 4 (2013)
38. Renukadevi, T., Karunakaran, S.: Optimizing deep belief network parameters using grasshopper algorithm for liver disease classification. Int. J. Imaging Syst. Technol. **30**(1), 168–184 (2020)
39. Sanghavi, J., Kayande, D.: Content based image retrieval (CBIR) system for diagnosis of blood related diseases. In: National Conference on Innovative Paradigms in Engineering & Technology (NCIPET-2013), (February), pp. 11–15 (2013)
40. Tumuluru, P., Ravi, B.: GOA-based DBN: Grasshopper optimization algorithm-based deep belief neural networks for cancer classification. Int. J. Appl. Eng. Res. **12**(24), 14218–14231 (2017)

41. Umamaheswari, J., Radhamani, G.: A hybrid approach for classification of DICOM image. World Comput. Sci. Inf. Technol. J. **1**(8), 364–369 (2011)
42. Varish, N., Pal, A.K.: Appl. Intell. **48**, 2930–2953 (2018)

Dr. N. T. Renukadevi received the M.C.A. and M.Phil. Degrees in Computer Science from Bharathiar University, India in 2003 and 2008 respectively. She received Ph.D. degree in Computer Applications from Anna University, India in 2014. Since 2008, she is working as an Assistant Professor in the Department of Computer Technology, Kongu Engineering College. She has been in the teaching profession for the past 15 years. She has published 17 papers in International Journals and Conferences. Her research interests include Data Mining and Medical Image Processing. She is also a reviewer in reputed journals.