HYBRID APPROACH OF MEDICAL IMAGE CLASSIFICATION USING GROUP SPARSITY REGULARIZATION AND BAT ALGORITHM

SHASHWATI MISHRA[1] AND MRUTYUNJAYA PANDA[2]

[1,2]Department of Computer Science and Applications, Utkal University, Vani Vihar, Bhubaneswar, Odisha, India

e-mail: shashwati.mishra@gmail.com[1], mrutyunjayapanda@yahoo.in[2]

Abstract - The classification has gained popularity in the research community due to its wide variety of applications in different fields. The complexity in the extraction of information from images and the use of this information for classification task has made image classification a tedious job. The Proposed method of classification uses the SIFT algorithm for extracting rotation and scale invariant features. Using the concept of Deep Neural Network, sparse coding technique is used for generating codebooks from the extracted features. Sparse codes give an intermediate representation between local codes and dense codes. These codes can extract information at different levels and with varying amounts according to the type of input. Instead of using the L1 norm, another popular regularizer is used in this paper, which maintains group sparsity for non-overlapping groups. Pooling operation is applied to the sparse coded features. Bat algorithm is used on these pooled features for the classification of medical images. Experimental results prove the fruitfulness of our proposed method in medical image classification.

Keywords -SIFT, Group Sparsity, Bat Algorithm, Spatial Pyramid Matching, Classification

(Received June 1st, 2020/ Accepted June 11st, 2020)

1. Introduction

The automated analysis of medical images using computer programs has gained its popularity from the day when it became possible to load and store images on the computer. Processing of low-level pixels, creating compound rule-based systems using mathematical models were used extensively during the period from the 1970s to 1990s. Towards the end of 1990s popularity of rule-based artificial intelligence gradually decreased and the concept of supervised techniques emerged as a new technique for image analysis. Algorithms were developed for the extraction of features from training dataset and statistical classification methods were used for object detection, pattern recognition, classification etc. Now-a-days these techniques are also very popular in which feature extraction step is an important component and greatly affects the final result. A later development on these feature extraction technique is the use of models having more than one hidden layer which converts a given input to an output. This process of learning features at different levels using the neural network is known as deep learning which gained its popularity around 2006[1].

Deep learning is one of the popular techniques of detection of anomalies and outputs best for unstructured data like images, text, audio etc. The capability of estimating new features from the set of training features is one of the greatest advantages of deep learning as compared to other neural network algorithms. This characteristic has also motivated the researchers to apply the deep learning technique on more complex feature sets. Deep learning works as a robust method for extraction of information from the images which are normally affected by various parameters like product reflectance, intensity of light, distortion in the lens etc. Unlike other machine learning techniques, deep learning techniques can detect irregular patterns and shapes in the image. Deep Neural Network (DNN) adopts different strategies like supervised (Convolutional Neural Network (CNN), Recurrent Neural Network (RNN)) and unsupervised (Auto-encoders (AE), Restricted Boltzmann Machine (RBM), Sparse Coding (SC)) for training the network.

Sparse coding is one of the unsupervised deep learning techniques used for the extraction of sparse features from the given input. Using the corresponding basis the data can again be reconstructed from the learned features. Principal Component Analysis (PCA) helps to extract a complete set of basis vectors whereas sparse coding considers the concept of sparsity and extracts an over-complete set of basis vectors effectively [2]. This paper is based on the application of the sparse coding technique of deep learning on the SIFT extracted features.

Classification is a process of categorizing or diving or separating information for better analysis, identification, and visualization of information content. It helps in object detection and extraction in medical science, classification of organisms in biological science, classification of documents, statistical analysis and many more. In this paper, the Bat algorithm is used on spatially pooled sparse features for image classification. Bat algorithm is one of the flexible and simple algorithms and helps in giving good optimal results. Implementation of this algorithm is easy and also gives promising solutions in case of complicated and nonlinear problems. Unlike some other metaheuristic algorithms, the Bat algorithm does not use a fixed parameter. With an increase in the number of iterations the loudness and pulse rate emission parameters vary in the Bat algorithm. This characteristic of controlling the values of parameters helps in switching from exploration to exploitation stage. Gradually it converges towards the optimal solution. The region where the solution can be found is zoomed automatically and search operation for a solution is started in that region. The idea of frequency tuning used in the Bat algorithm is similar to the key concept of Particle Swarm Optimization and Harmony Search. So, the Bat algorithm also possesses the advantages of these algorithms [3].

1.1 *Contribution:-*

Motivated by the use of the Bat algorithm in different image processing applications, this algorithm is applied for grouping images according to their contents. Another major contribution is the use of Group regularization for sparse coding and pooling instead of L1-norm regularization. The accuracies obtained are also analysed considering different pooling techniques like max pooling, sum pooling and average pooling. It is found that the max poling technique gives the best result in less time as compared to the other two. The comparison with the SVM classifier proves that, Bat algorithm based technique gives better classification accuracy than SVM.

1.2 *Organization of the paper:-*

Section 2 gives an overview of the work done by the researchers using SIFT, deep learning techniques and the Bat algorithm. Section 3 contains the proposed methodology along with the descriptions about the methods used in the paper. The experimental observations and discussions are given in section 4 followed by the concluding remarks in section 5.

2. Related Work

Feature extraction plays a very important role in different image processing applications. Feature acts like the basic building block in an image that contains all the information related to the image. To perform any type of analysis and processing activity those features where maximum image related information is stored should be selected and extracted first. All type of processing activities is carried out on these features to reduce the time and space complexity. So the accuracy of any analysis depends on the extraction of suitable features which depend on feature extraction algorithms. One of the main problems with images is that the size and orientation of a particular object may vary from image to image. The brightness of one image may be different from another and so on. Scale Invariant Feature Transform (SIFT) algorithm extracts rotation and scale invariant features from an image. These features are also not affected by the change in illumination which is very much essential for the extraction of appropriate image features. SIFT descriptors are used for different purposes like action recognition [4], image matching [5], remote sensing image registration [6], object recognition [7], Content Based Image Retrieval (CBIR) [8], image watermarking [9] and many more. Several variants of SIFT are also proposed for solving various real-world problems. Sparse coding techniques are also used for obtaining a sparse representation of extracted features which can be helpful for efficient image classification, image representation and different processing activities [10] [11] [12] [13].

One of the primary areas of image processing where deep learning techniques have maximum contribution is image classification. Transfer learning techniques either use a pre-trained network or fine-tune a trained network to analyse the input data. A detailed analysis of both strategies has performed to obtain a comparative overview of both [1]. Unsupervised network architectures like Deep Belief Networks are applied to brain MRI images for manifold learning [14]. Latent features are extracted from MRI images using Sparse Auto-encoders for diagnosis of brain disease [15]. A. Payan and G. Montana used sparse autoencoders along with 3D convolutional neural networks for predicting the status of Alzheimer's disease from brain MRI scan images [16]. E. Hosseini-Asl et al. proposed a technique for diagnosing Alzheimer's disease using 3D convolutional autoencoder for feature extraction and Deeply Supervised Adaptive 3D Convolutional Neural Network for classification [17].

Medical science has made significant advancements in 3D imaging. 3D ultrasound images were used to diagnose syndromes on babies before birth [18]. E. Vezzetti et al. [19] also suggested a technique of diagnosing and formalizing prenatal cleft lip using representative key points and identifying defect types in 3D ultrasonography [19].

Besides brain MRI images deep learning techniques are also used in the analysis of other medical images like Computed Tomography (CT) of lungs, retinal images etc. A. Esteva et al. also applied Convolutional Neural Network for classification of skin cancer [20]. Z. Gao et al. classified human epithelial-2 cell images using Deep Convolutional Neural Networks (CNNs) [21].

C. Zhang et al. proposed a technique of image classification using non-negative sparse coding followed by multi-scale max pooling on the extracted features. Low rank and sparse matrix recovery techniques were used to separate feature vectors according to their class from which a low-rank matrix and a sparse error matrix were constructed. Locality-constrained Linear Coding (LLC) is applied to these matrices and finally, linear SVM classifier is used to get the final classified result [12]. J. Yang et al. applied sparse coding on the extracted SIFT features. Then spatially pooled features are classified using linear SVM (Support Vector Machine) classifier [22].

T.-H. Chan et al. developed PCANet (Principal Component Analysis Network) for image classification having a training stage and a classification stage. The training stage involves three layers one for filtering followed by processing and then pooling the features. PCA is used for filtering followed by the process of binary hashing. Then blocks are combined to generate histogram plots. Linear Support Vector Machine is applied to pooled features for classification [23].

Z. Akata et al. used attribute information for embedding class labels. A function is used to measure the compatibility between the embedded inputs and output classes [24]. M. Korytkowski et al. applied fuzzy classifiers for image classification and boosting meta learning is used to extract local features [25]. J. G. Serra et al. used logistic regression based on a probabilistic approach for multispectral image classification [26]. L. He et al. applied Gabor filtering for hyperspectral image classification [27]. A lot of research has also been done for classification using artificial neural network [28] [29].Y. Wang and H. Li used artificial neural network for remote sensing image classification [30]. E. Maggiori et al. applied Convolutional Neural Networks (CNNs) for remote sensing image classification [31]. E. H. Aria et al. applied Back Propagation Neural Network (BPNN) for satellite image classification [32]. Z. Li et al. considered locality and label information to learn dictionary for image classification [33]. G. Sharma et al. used spatial saliency maps for classifying images [34].

Classification is one of the important image processing techniques and has a wide variety of applications in medical science. A large number of researches have already been done for selecting an appropriate classifier [35]. The popularity of nature- inspired algorithms has increased its use in finding an optimal solution. Cuckoo search [36], Firefly algorithm [37], Harmony search [38], Ant Colony optimization [39], Artificial Bee Colony [40], Honey Bee [41], Monkey Search [42], Bat algorithm [3] are some examples of these nature-inspired algorithms. These algorithms can also be applied to generate an optimal model for classification.

Bat algorithm was combined with fuzzy classifier to obtain optimized rules and membership functions for classifying medical data. The accuracy obtained was 75.21% for Lung cancer data and 76.67% for Indian Liver data [43]. Bat algorithm is applied as an image enhancer to detect hairline bone fracture in medical x-ray images [44]. Using Bat algorithm a method called Bat-Active Contour Method (BA-ACM) was suggested for segmenting medical images [45]. Microarray data contains redundant, irrelevant and noisy information. O. A. Alomari et al. [46] were proposed a technique of gene selection in microarray data for accurate diagnosis of cancer. In the wrapper stage of their proposed technique Bat algorithm and SVM were used [46]. Bat algorithm was also used to obtain optimal peak signal to noise ratio value for medical image watermarking in wavelet domain [47].

3. Proposed Methodology

The process of image classification involves some of the basic steps like feature extraction, building dictionary from the extracted features, generating codebook, pooling the coded information, generating feature vector from the pooled information and finally use of a suitable technique for classification. In this paper Scale Invariant Feature Transform(SIFT) algorithm is used for feature extraction, sparse coding is applied for generating sparse codes, followed by pooling of sparse codes. In the final step, the Bat algorithm is used to generate an optimal model for classification from the training set of images. The test image set is classified using the same model which gives very good accuracy in classifying the input images. Figure 1. gives a graphical representation of the proposed method.

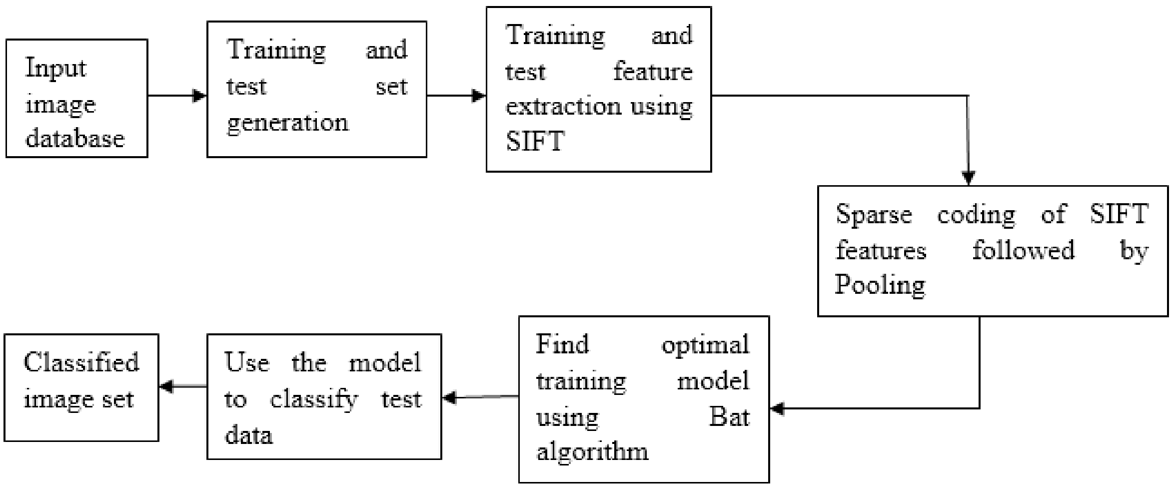

Fig. 1. Proposed Methodology

3.1*Scale Invariant Feature Transform (SIFT):-*

SIFT (Scale Invariant Feature Transform) algorithm plays a vital role in the extraction of interesting and necessary information from an image. SIFT algorithm helps in finding scale and rotation invariant features that are not affected by noise and illumination changes also. The algorithm was developed by David Lowe from UBC (University of British Columbia) [48]. Scale-space extrema detection, key-point localization, orientation assignment and key-point descriptor creation are the four basic steps of the SIFT algorithm. A detailed description of this algorithm along with the four steps can be found in [48]. The outputs of this algorithm are the feature vectors having 128 dimensions.

3.2*Sparse coding:-*

Sparse coding is a group of unsupervised algorithms useful in obtaining clear representations of essential information from input data. Extracting higher level features from the input with the help of basis functions is computationally complex. L1-regularized least squares problem, L2-constrained least squares problem are some of the efficient algorithms used in sparse coding. The basis vectors obtained using sparse coding have the same characteristics as the receptive fields of the neurons present in the visual cortex of the brain. Sparse coding can learn bases greater than the dimensions of inputs, which cannot be learned by some unsupervised techniques like PCA. The extracted features can be used as input to different classification algorithms [2].

Suppose F represents the feature vectors extracted using the SIFT algorithm. D is the dimension of the extracted feature vector. N is the number of extracted features. This can be represented mathematically as

$$F = \begin{bmatrix} f_1 \\ f_2 \\ . \\ . \\ f_N \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} \cdots\cdots f_{1D} \\ f_{21} & f_{22} \cdots\cdots f_{2D} \\ .. & .. \quad .. \quad .. \\ .. & .. \quad .. \quad .. \\ f_{N1} & f_{N2} \cdots\cdots f_{ND} \end{bmatrix} \Rightarrow F \in R^{N \times D} \tag{1}$$

$f_1, f_{2,\ldots\ldots,} f_N$ represent individual feature descriptors each of having dimension D [22].

If the number of classes is K, then for each class one codebook can be generated. Let C be the set of all codebooks. This can be written as,

$$C = \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ c_K \end{bmatrix} \tag{2}$$

Vector Quantization (VQ) technique uses L_2norm and tries to minimize the difference between the feature descriptors and cluster centres by using the k-means clustering concept [22].

$$\min_{C} \sum_{n=1}^{N} \min_{k=1,...,K} \| f_n - c_k \|^2 \tag{3}$$

This will work like the k-means algorithm and group similar features in one group. Each feature descriptor can have a membership value, which indicates the probability of belonging to a particular class. If M represents the membership indicators,

$$M = \begin{bmatrix} m_1 \\ m_2 \\ . \\ . \\ m_N \end{bmatrix} \tag{4}$$

then Eq. (3) can be rewritten as

$$\min_{M,C} \sum_{n=1}^{N} \| f_n - m_n C \|^2 \tag{5}$$

with the constraints that, only one element of m_n is nonzero ($card(m_n) = 1$) ,all elements of m_n are positive ($m_n \geq 0$) and L_1 norm of m_n , $|m_n| = 1$ [22]. When the optimization process is over, the index of the nonzero element in m_n represents the cluster to which f_n belongs.

The sparse coding technique reduces the restriction that, only one element of m_n is nonzero. m_n is allowed to have a small number of nonzero elements by adding a L_1 norm regularization on m_n multiplied by a constant term λ . So, in sparse coding we have

$$\min_{M,C} \sum_{n=1}^{N} \| f_n - m_n C \|^2 + \lambda |m_n| \tag{6}$$

with the constraints $\|c_K\| \leq 1,\ \forall K = 1,2,.....,k$.

Generally, the number of codebooks is greater than the dimension of the features, that means $K > D$. Since the sign of m_n is not important, the constraint $m_n \geq 0$ is removed and both positive and negative values for C and m_n are considered [22].

The first term of the objective function is reconstruction term or data fitting term and the second term is the penalty or regularization term. The reconstruction term helps to find a good representation of the input feature vector. The objective of the penalty is to obtain a sparse representation of the input vector [2].

In this paper instead of the L_1 norm, another popular regularizer for sparsity is used which maintains group sparsity without overlaps. This sparsity constraint is particularly useful for linear models where groups are generally non-overlapping in nature [49]. Group sparsity without overlaps is the sparsity constraint used to regularize the groups of features. This helps in regularizing the structure of sparsity and helps in maintaining some sparsity pattern. This structure sparsity regularization method assumes an a priori partition of coefficients in different groups which are non-overlapping in nature. By regularizing the features in non-overlapping groups, this sparsity constraint provides some prior knowledge about the groups or clusters which helps in better classification than L1 norm.

The regularization term can be represented as in Eq. (7)

$$R = \begin{cases} (|m_n| - \lambda \dfrac{|m_n|}{\|m_n\|}), & if\ \|m_n\| > \lambda \\ 0, & if\ \|m_n\| \in [-\lambda, \lambda] \\ (|m_n| + \lambda \dfrac{|m_n|}{\|m_n\|}), & if\ \|m_n\| < -\lambda \end{cases} \tag{7}$$

So, using this sparsity constraint Eq. (6)can be converted to Eq. (8) [22] [49]

$$\min_{M,C} \sum_{n=1}^{N} \| f_n - m_n C \|^2 + R \tag{8}$$

Sparse coding involves two phases, one training phase and another coding phase. During the training phase, Eq. (8) is used to generate the codebook C . The codebook C is generated by solving Eq. (8) with respect to M and C . In the coding phase, Eq. (8) is optimized with respect to M only to obtain sparse coding codes [22].

Sparse coding is less restrictive as compared to vector quantization coding. So, sparse coding has lower reconstruction error than vector quantization coding. In comparison to vector quantization coding, sparse coding gives a better representation of an image by capturing its salient properties [22]. The Lagrange dual is used to learn the bases as proposed by H. Lee at al. [2]. Better classification accuracy can be obtained if the groups are non-overlapping in nature. The group sparsity constraint without overlaps is used to obtain better accuracy which can also be verified from the experimental results.

3.3*Pooling:-*

After coding the next step is to pool the sparse coded information to obtain the features that best represent the image. Different pooling techniques used on images are average pooling, max pooling and sum pooling, among which max pooling is most popular. and gives good results. It has been found that sparse

Coding with max pooling gives better classification results and also helps in accurate retrieval of images from large databases. In this paper, the features are pooled from each image separately using sparse coded dictionary. The local features are extracted and linear spatial pyramid matching is used to pool the sparse coded features. Finally, a set of pooled features and corresponding class labels are obtained which are given as input to the Bat algorithm for classification.

3.4*Bat Algorithm:-*

Xin-she Yang in 2010 developed one of the popular nature-inspired algorithms called the Bat algorithm, which is popularly used in colour image segmentation [50], image matching [51] and multilevel image thresholding [52]. Firefly, cuckoo search and Bat algorithms are some of the recent nature-inspired algorithms. Firefly algorithm has a high convergence rate whereas cuckoo search is simple to implement. Bat algorithm is more accurate and efficient than these two algorithms. The characteristics like auto zooming, controlling the parameter values and frequency tuning make this algorithm more effective for optimization. Bat algorithm also helps in getting a globally optimal solution. For classifying images in a database a globally optimal solution is desired. The feature is the basic element of an image that is analysed and explored for obtaining an optimal result. Auto zooming helps in providing better classification results of image pixels. These factors inspired us to apply the Bat algorithm for image classification.

Bat use SONAR (Sound Navigation and Ranging) echoes during echolocation for detecting and avoiding the obstacles. The waves are reflected back from obstacles and the time gap between emission and reflection of the wave affects the movement of the bat. Using its own pulse bat determines the time gap and finds the distance of the pray from itself.

Pulse rate 0 indicates no emission and maximum emission is indicated by pulse rate 1. Generally when a bat finds its pray loudness decreases. With a decrease in loudness, pulse rate emission increases [3].

For searching a pray, bats fly randomly with a velocity (say,v_i) from a position (say,x_i). Suppose the initial frequency isf_{min}, wavelength is λ and loudness is A_0. The bat measures the proximity of the target and accordingly adjusts the frequency (or wavelength) of the emitted pulses as well as the rate of pulse emission (say, r_i). Bats move from one position to another position, generate new solutions and adjust their velocities and positions using the Eq.(9), Eq. (10) and Eq. (11):

$$f_i = f_{min} + (f_{max} - f_{min})\beta \tag{9}$$

$$v_i^t = v_i^{t-1} + (x_i^t - x^*)f_i \tag{10}$$

$$x_i^t = x_i^{t-1} + v_i^t \tag{11}$$

where, $v_i^t$is the velocity of the bat at iteration t in a d-dimensional search space, x_i^t is the location or position of the bat at iteration t in a d-dimensional search space,$v_i^{t-1}$is the velocity of the bat at iteration (t-1) in a d-dimensional search space, x_i^{t-1} is the location or position of the bat at iteration (t-1) in a d-dimensional search space, $\beta \in [0,1]$ is a random vector drawn from a uniform distribution, x^* is the current best location or solution [3].

The value of loudness A_i and pulse emission rate r_i changes during the iteration steps. As the bat moves closer to its prey loudness gradually decreases whereas the rate of pulse emission increases gradually. Finally when the bat finds its prey loudness becomes zero and emission of sound stops temporarily. Assuming α and γ as constants, loudness and pulse rate emission updating rules can be represented mathematically as in Eq. (12) and Eq. (13) respectively [3].

$$A_i^{t+1} = \alpha A_i^t \tag{12}$$

$$r_i^{t+1} = r_i^0[1 - \exp(-\gamma t)] \tag{13}$$

When $t \to \infty$ for $0 < \alpha < 1$ and $\gamma > 0$,

$$A_i^t \to 0 \tag{14}$$

$$r_i^t \rightarrow r_i^0 \tag{15}$$

In this paper, the Bat algorithm is used for grouping the features according to their similarity. To obtain an optimal value entropy function is used and the objective function is calculated by maximizing the entropy. Entropy measures the randomness which helps in characterizing the texture of an image and can be defined as in Eq. (16).

$$E = -sum(p.*\log 2(p)) \tag{16}$$

where, p represents the histogram count which indicates the probability of occurrence of a particular intensity. Each bin corresponds to a particular intensity value. A histogram of an image represents the tonal distribution of the image graphically.

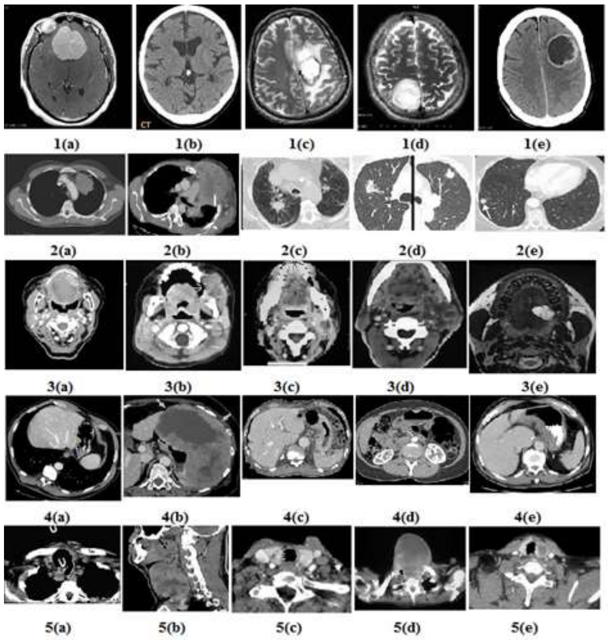

Fig. 2. Some examples of images considered for applying the proposed methodology brain cancer images (1(a) to 1(e), lung cancer images (2(a) to 2(e), oral cavity cancer (3(a) to 3(e), stomach cancer (4(a) to 4(e), thyroid cancer (5(a) to 5(e)

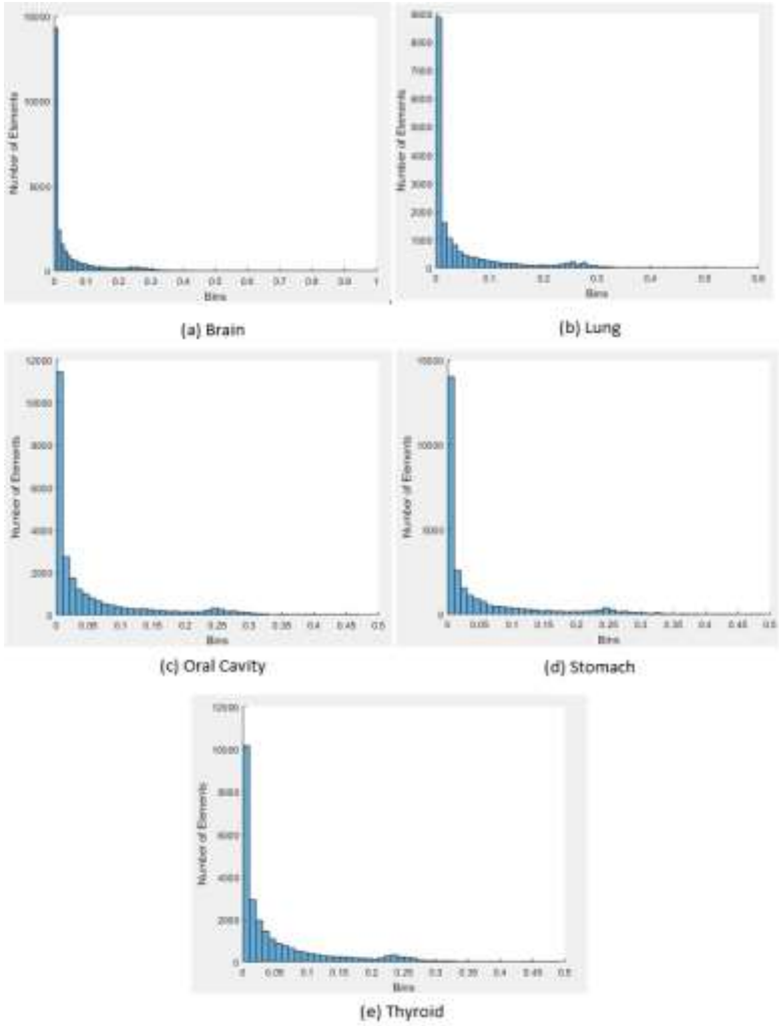

Fig. 3. Histogram plot of extracted SIFT features

4. Experimental Observations and Discussions

For experimental analysis of the proposed method, a medical image set is prepared which comprised of various types of cancer images. The images are selected from publicly available medical image repositories. The images are selected considering different sizes, shapes and orientations of organs. It also includes images taken from different angles and at different levels. The size and shape of tumors in the images are also different from each other. The image set comprised of brain cancer, lung cancer, oral cavity cancer, stomach cancer, thyroid cancer images. These five categories of cancer images are assigned class labels 1, 2, 3, 4 and 5 depending on the class (brain or lung or oral cavity or stomach or thyroid) to which the image belongs. Figure 2 shows some of these images used for applying the proposed approach.

The proposed method is applied on the medical image set consisting of 52 cancer images and in the initial step features are extracted using SIFT algorithm. Figure 3 contains the histogram plot of the extracted features, one from each category of cancer image. In the next step, the sparse coding technique is used followed by max pooling. The SIFT feature descriptors are of 128 dimensions and for sparse coding also we have considered the number of bases as 128. So, the learned dictionary size is 128x128. The sparsity regularization parameter is considered as 0.2 and smoothing regularization parameter as 0.05. For comparative analysis, we have experimented using different numbers of samples like 2000, 2500 and 3000. Depending on this number of samples the size of the sparse coded matrix also varies. The objective function value is calculated on the basis of the difference between actual data values and sparse coded values. The sparsity of the sparse matrix also affects the value of the objective function. Table 1 shows the objective function values for 20 iterations considering different numbers of samples like 2000, 2500 and 3000. A graphical representation of these values for different number of samples is also given in Figure 4.

Table 1. Variation in objective function values with change in number of iterations and number of samples

Dimension-128 Number of iterations-20 Gamma=0.2 Beta=0.05	No. of Iterations	Objective function (2000 Samples)	Objective function (2500 Samples)	Objective function (3000 Samples)
	1	0.4836	0.4841	0.4843
	2	0.3309	0.3311	0.3301
	3	0.3075	0.3076	0.3085
	4	0.3025	0.3020	0.3031
	5	0.3002	0.2997	0.3006
	6	0.2988	0.2985	0.2992
	7	0.2980	0.2978	0.2983
	8	0.2973	0.2973	0.2976
	9	0.2968	0.2969	0.2972
	10	0.2964	0.2966	0.2968
	11	0.2962	0.2963	0.2965
	12	0.2959	0.2961	0.2963
	13	0.2958	0.2959	0.2961
	14	0.2956	0.2958	0.2960
	15	0.2955	0.2957	0.2959
	16	0.2954	0.2956	0.2958
	17	0.2953	0.2955	0.2957
	18	0.2952	0.2954	0.2956
	19	0.2951	0.2954	0.2956
	20	0.2951	0.2953	0.2955

After coding, the features are pooled using max pooling by computing spatial bins at different pyramid levels. The proposed method uses group regularization instead of L1-norm regularization which maintains group sparsity without overlaps. A comparative analysis of results obtained using L1-norm regularization and group regularization for different pooling techniques is given in Table 2. The graphical representation of these results for max, mean and sum pooling are added as Figure 5, Figure 6 and Figure 7 respectively.

For classifying the images Bat algorithm is applied on the pooled features. The image set is divided into training and test image sets, considering 4 images from each category as training images. Out of a total of 52 images of five categories such as brain cancer, lung cancer, oral cavity cancer, stomach cancer, thyroid cancer 20 are considered as training images and 32 are considered as test images. The optimal objective function value for grouping the images into different classes is obtained using entropy function in

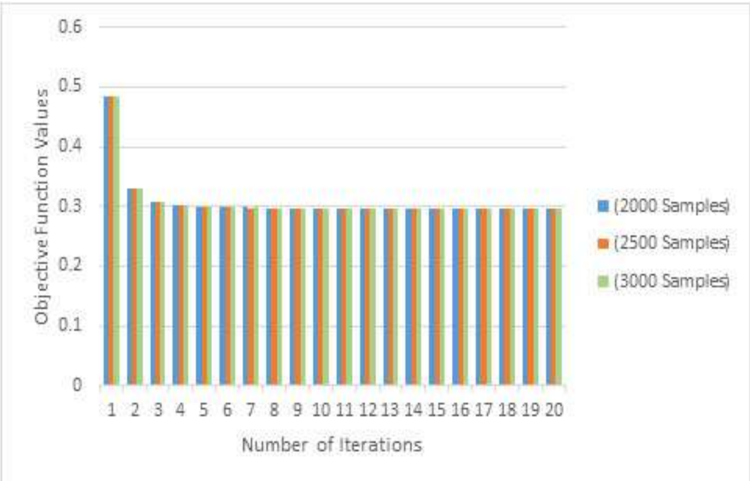

Fig. 4. Change in objective function with change in number of iterations and number of samples

the Bat algorithm.

The publicly available cancer image sets are classified using Bat algorithm and a comparative analysis is made with the classification results obtained using the SVM algorithm. For comparative analysis, the SVM classifier is used in the paper. The SVM used is a linear multiclass classifier. To reduce the number of computations linear kernel has been applied. It also uses differentiable quadratic hinge loss so that gradient based optimization methods can be used as proposed by J. Yang et al. [22]. Comparison is made for different numbers of samples like 2000, 2500 and 3000 for getting a better conclusion. The comparative analysis is shown in Table 3 and a graphical representation of the same is given in Figure 8. From the table, it can be concluded that in each case the Bat approach works better than SVM and the highest accuracy obtained is 87.50 for 2500 number of samples.

Table 2. Comparative analysis of accuracy obtained using L1-norm and Group regularization with different pooling methods and Bat algorithm for classification

No. of Samples	Accuracy Using Bat Algorithm					
	Max Pooling		Mean Pooling		Sum Pooling	
	L1	Group	L1	Group	L1	Group
2000	65.63	71.88	56.25	62.50	62.50	62.50
2500	68.75	87.50	65.63	68.75	50.00	75.00
3000	68.75	75.00	62.50	65.63	59.38	68.75

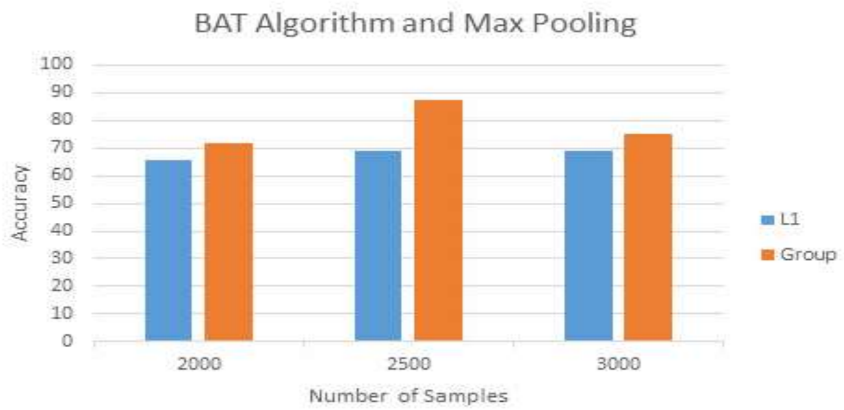

Fig. 5. Comparative analysis of accuracy of classification using L1-norm and Group regularization with max pooling and Bat algorithm

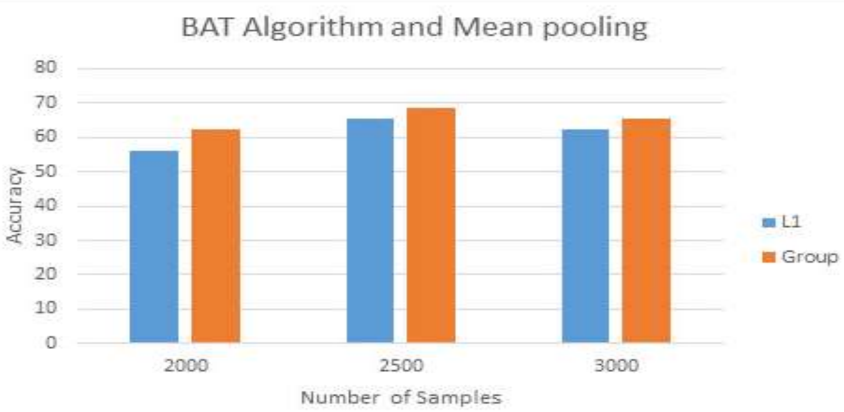

Fig. 6. Comparative analysis of accuracy of classification using L1-norm and Group regularization with mean pooling and Bat algorithm

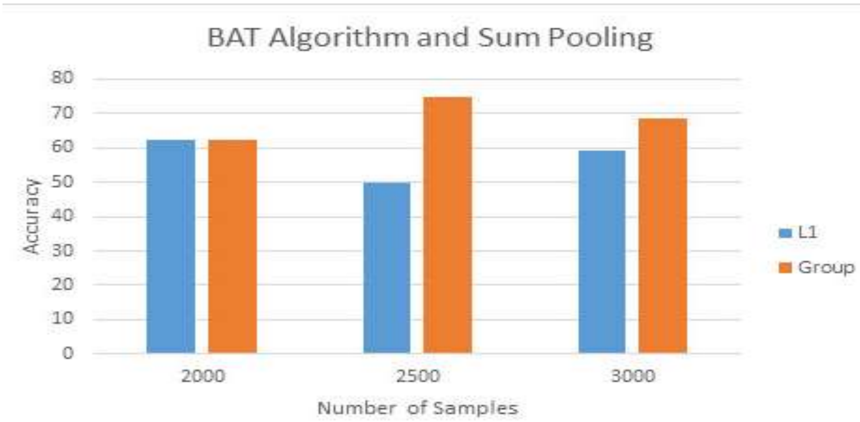

Fig. 7. Comparative analysis of accuracy of classification using L1-norm and Group regularization with sum pooling and Bat algorithm

Table 3. Comparative analysis of accuracy obtained using SVM and Bat algorithm for classification

No. of Samples	Accuracy	
	SVM	Bat
2000	67.83	71.88
2500	81.50	87.50
3000	73.83	75.00

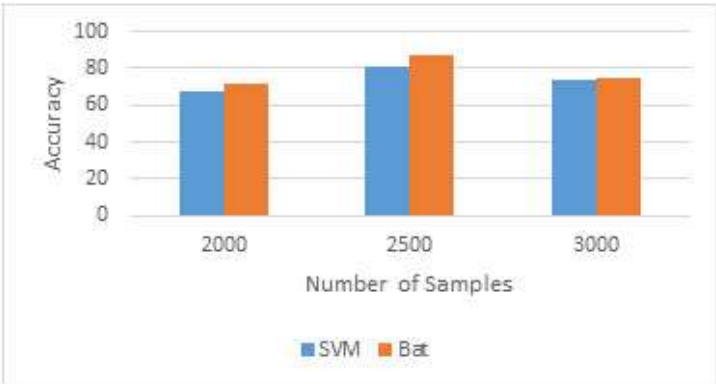

Fig. 8 Comparative analysis of accuracy of SVM and Bat for different number of samples

Confusion matrices calculated using the Bat algorithm for each sample size are given in Table 4 (for 2000 samples), Table 5 (for 2500 samples) and Table 6 (for 3000 samples). For each category of images, a class label is assigned. The image set consists of five different categories of images. Accordingly, our class labels are named as class 1, class 2, class 3, class 4 and class 5. The confusion matrices are shown in Table 4, Table 5 and Table 6. These matrices contain the class labels on one side as predicted labels and on the other side as actual labels. Table 4 says that 6 images from class 1, 5 from class 2, 2 from class 3, 5 from class 4 and 5 from class 5 are correctly classified. Similarly, 1 from class 2, 4 from class 3, 3 from class 4 and 1 from class 5 are incorrectly classified. Similar representations are also given in Table 5 and Table 6 for 2500 and 3000 samples respectively.

Table 4. Confusion matrix for 2000 samples

Confusion Matrix for 2000 Samples						
	Predicted Class Labels					
Actual Class Labels		1	2	3	4	5
	1	6	1	3	3	0
	2	0	5	0	0	0
	3	0	0	2	0	0
	4	0	0	1	5	1
	5	0	0	0	0	5

Table 5. Confusion matrix for 2500 samples

Confusion Matrix for 2500 Samples						
	Predicted Class Labels					
Actual Class Labels		1	2	3	4	5
	1	6	0	0	0	1
	2	0	5	0	0	0
	3	0	1	6	0	2
	4	0	0	0	8	0
	5	0	0	0	0	3

Table 6. Confusion matrix for 3000 samples

Confusion Matrix for 3000 Samples						
	Predicted Class Labels					
Actual Class Labels		1	2	3	4	5
	1	5	0	2	1	2
	2	0	5	0	0	0
	3	1	0	3	0	0
	4	0	1	1	7	0
	5	0	0	0	0	4

The proposed technique is implemented in Matlab R2015a with a system specifications as Intel core i5 processor, 2.30 GHz speed, 4 GB RAM, 64-bit operating system, 1 TB hard disk. Table 7 gives the time taken by the proposed technique for execution which is measured in seconds.

Table 7. Execution time using Bat algorithm for different number of samples

No. of Samples	CPU Time (in Seconds)
2000	5.107606
2500	4.976612
3000	5.110723

Table 8. Comparative analysis of different pooling methods taking sample size as 2500

Number of Samples=2500		
Pooling Methods	Accuracy	CPU Time (in seconds)
Mean	68.75	5.093417
Sum	75.00	5.046945
Max	87.50	4.976612

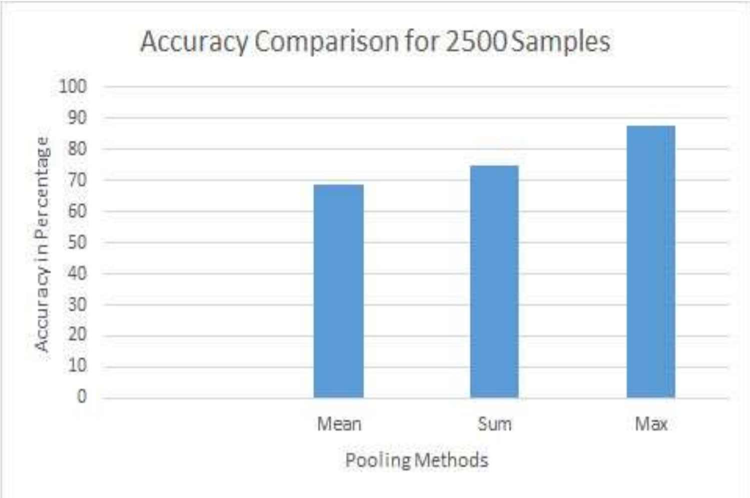

Fig. 9. Comparative analysis of accuracy using different pooling methods for 2500 samples

Fig. 10. Comparative analysis of CPU time using different pooling methods for 2500 samples

The maximum accuracy is obtained considering sample size as 2500 using Bat algorithm for classification. So, taking the same sample size accuracy and execution time for different pooling methods are also compared. The results obtained are shown in Table 8 and graphical analyses are given in Figure 9 and Figure 10. It is found that out of mean, sum and max pooling techniques, max pooling gives the best accuracy as 87.50% and less execution time as 4.976612 seconds.

R. Zhang et al. [53] applied multi-scale non-negative sparse coding technique for medical image classification. In their proposed approach medical images are first disintegrated into multiple scale layers which helps in extracting distinct visual information from different scale layers. To generate a discriminative sparse representation of the images, a non-negative sparse coding model along with fisher discriminant analysis is created for each scale layer. A multi-scale feature histogram is constructed from multi-scale non-negative sparse coding features. Then their technique uses SVM as the classifier for medical image classification. Their proposed technique is analysed using the ADL dataset and Neusoft NSR dataset.

Table 9. Performance analysis using ADL dataset

Technique Used	Accuracy for ADL dataset
R. Zhang et al. [ ]	80.3±0.4 (using Gauss-SVM), 81.6±0.2 (using CSK-SVM)
Our proposed technique	83.33%

For better comparative analysis the ADL dataset is also taken for applying our proposed methodology. The accuracy obtained using their approach and our suggested technique considering this data set is given in table 9. The results prove the effectiveness of our approach.

5. Conclusion

This paper proposes a method for image classification using concepts from the neural network and nature-inspired algorithms. The rotation, scale and illumination invariant nature of the SIFT algorithm helps in extracting accurate features. The extracted features are not affected by the change in size, illumination and rotation of images. Sparse codes which give an intermediate representation between local codes and dense codes are capable of extracting sparse features as per the input provided by the user. In classification, the features in a group have the same characteristics. Considering this factor, in this paper group sparsity is considered for generating codebooks. In the pooling stage, the pooled spatial features at different levels help in better extraction of image information. After sparse coding and pooling, the Bat algorithm is used for the classification of images. Bat algorithm considers the echolocation behaviour of bats and uses velocity, frequency, wavelength, loudness, pulse rate emission to obtain the optimal values after a certain number of iterations. This makes the process of feature extraction more refined and the extracted features become efficient and accurate. With the increase in the number of iterations loudness and pulse rate emission parameter values can be updated which helps in better estimation of these parameters. Searching in a location where the probability of finding the solution is more helps the process of searching simple and fast. The result obtained using the Bat algorithm is compared with that of the SVM classification result. The comparative study proves that the Bat approach gives better classification accuracy as compared to the SVM technique. The maximum accuracy obtained by our hybrid classification approach is 87.50% for 2500 number of samples. The proposed method can be improved for reducing the time taken for the classification of images.

References

1. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B. & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, *42*, 60-88.
2. Lee, H., Battle, A., Raina, R., & Ng, A. Y. (2007). Efficient sparse coding algorithms. In *Advances in neural information processing systems* (pp. 801-808).
3. Yang, X. S., & He, X. (2013). Bat algorithm: literature review and applications. *International Journal of Bio-inspired computation*, *5*(3), 141-149.
4. Scovanner, P., Ali, S., & Shah, M. (2007, September). A 3-dimensional sift descriptor and its application to action recognition. In *Proceedings of the 15th ACM international conference on Multimedia* (pp. 357-360).
5. Cheung, W., & Hamarneh, G. (2007, April). N-sift: N-dimensional scale invariant feature transform for matching medical images. In *2007 4th IEEE*

International Symposium on Biomedical Imaging: From Nano to Macro (pp. 720-723). IEEE.

6. Li, Q., Wang, G., Liu, J., & Chen, S. (2009). Robust scale-invariant feature matching for remote sensing image registration. *IEEE Geoscience and Remote Sensing Letters*, *6*(2), 287-291.
7. Lowe, D. G. (1999, September). Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision* (Vol. 2, pp. 1150-1157). IEEE.
8. Giveki, D., Soltanshahi, M. A., & Montazer, G. A. (2017). A new image feature descriptor for content based image retrieval using scale invariant feature transform and local derivative pattern. *Optik*, *131*, 242-254.
9. Li, L., Guo, B., & Shao, K. (2007). Geometrically robust image watermarking using scale-invariant feature transform and Zernike moments. *Chinese Optics Letters*, *5*(6), 332-335.
10. Dai, D., & Yang, W. (2010). Satellite image classification via two-layer sparse coding with biased image representation. *IEEE Geoscience and Remote Sensing Letters*, *8*(1), 173-176.
11. Mairal, J., Bach, F., Ponce, J., & Sapiro, G. (2009, June). Online dictionary learning for sparse coding. In *Proceedings of the 26th annual international conference on machine learning* (pp. 689-696).
12. Zhang, C., Liu, J., Tian, Q., Xu, C., Lu, H., & Ma, S. (2011, June). Image classification by non-negative sparse coding, low-rank and sparse decomposition. In *CVPR 2011* (pp. 1673-1680). IEEE.
13. Zheng, M., Bu, J., Chen, C., Wang, C., Zhang, L., Qiu, G., & Cai, D. (2010). Graph regularized sparse coding for image representation. *IEEE transactions on image processing*, *20*(5), 1327-1336.
14. Brosch, T., Tam, R., & Alzheimer's Disease Neuroimaging Initiative. (2013, September). Manifold learning of brain MRIs by deep learning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 633-640). Springer, Berlin, Heidelberg.
15. Suk, H. I., Lee, S. W., Shen, D., & Alzheimer's Disease Neuroimaging Initiative. (2015). Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. *Brain Structure and Function*, *220*(2), 841-859.
16. Payan, A., & Montana, G. (2015). Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. *arXiv preprint arXiv:1502.02506*.
17. Hosseini-Asl, E., Gimel'farb, G., & El-Baz, A. (2016). Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network. *arXiv preprint arXiv:1607.00556*.
18. Vezzetti, E., Speranza, D., Marcolin, F., Fracastoro, G., & Buscicchio, G. (2014). Exploiting 3D ultrasound for fetal diagnostic purpose through facial landmarking. *Image Analysis & Stereology*, *33*(3), 167-188.
19. Vezzetti, E., Speranza, D., Marcolin, F., & Fracastoro, G. (2016). Diagnosing cleft lip pathology in 3d ultrasound: A landmarking-based approach. *Image Analysis & Stereology*, *35*(1), 53-65.
20. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115-118.
21. Gao, Z., Wang, L., Zhou, L., & Zhang, J. (2016). HEp-2 cell image classification with deep convolutional neural networks. *IEEE journal of biomedical and health informatics*, *21*(2), 416-428.
22. Yang, J., Yu, K., Gong, Y., & Huang, T. (2009, June). Linear spatial pyramid matching using sparse coding for image classification. In *2009 IEEE Conference on computer vision and pattern recognition* (pp. 1794-1801). IEEE.
23. Chan, T. H., Jia, K., Gao, S., Lu, J., Zeng, Z., & Ma, Y. (2015). PCANet: A simple deep learning baseline for image classification?. *IEEE transactions on image processing*, *24*(12), 5017-5032.
24. Akata, Z., Perronnin, F., Harchaoui, Z., & Schmid, C. (2015). Label-embedding for image classification. *IEEE transactions on pattern analysis and machine intelligence*, *38*(7), 1425-1438.
25. Korytkowski, M., Rutkowski, L., & Scherer, R. (2016). Fast image classification by boosting fuzzy classifiers. *Information Sciences*, *327*, 175-182.
26. Serra, J. G., Ruiz, P., Molina, R., & Katsaggelos, A. K. (2016, September). Bayesian logistic regression with sparse general representation prior for multispectral image classification. In *2016 IEEE International Conference on Image Processing (ICIP)* (pp. 1893-1897). IEEE.
27. He, L., Li, J., Plaza, A., & Li, Y. (2016). Discriminative low-rank Gabor filtering for spectral–spatial hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(3), 1381-1395.

28. Saravanan, K., & Sasithra, S. (2014). Review on classification based on artificial neural networks. *International Journal of Ambient Systems and Applications (IJASA)*, *2*(4), 11-18.
29. Zhang, G. P. (2000). Neural networks for classification: a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *30*(4), 451-462.
30. Wang, Y. G., & Li, H. P. (2010, August). Remote sensing image classification based on artificial neural network: A case study of Honghe Wetlands National Nature Reserve. In *2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering* (Vol. 5, pp. 17-20). IEEE.
31. Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2016). Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(2), 645-657.
32. Aria, E. H., Amini, J., & Saradjian, M. R. (2003). Back propagation neural network for classification of IRS-1D satellite images. In *Joint Workshop of High Resolution Mapping from Space, Tehran University, Iran*.
33. Li, Z., Lai, Z., Xu, Y., Yang, J., & Zhang, D. (2017). A locality-constrained and label embedding dictionary learning algorithm for image classification. *IEEE transactions on neural networks and learning systems*, *28*(2), 278-293.
34. Sharma, G., Jurie, F., & Schmid, C. (2012, June). Discriminative spatial saliency for image classification. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3506-3513). IEEE.
35. Mishra, S. & Panda, M. (2017, October). Selection of an Efficient Image Classifier-A Critical Analysis. CiiT International Journal of Digital Image Processing, 9(8), 169-176.
36. Yang, X. S., & Deb, S. (2010). Engineering optimisation by cuckoo search. *arXiv preprint arXiv:1005.2908*.
37. Yang, X. S., & He, X. (2013). Firefly algorithm: recent advances and applications. *arXiv preprint arXiv:1308.3898*.
38. Geem, Z. W., Kim, J. H., & Loganathan, G. V. (2001). A new heuristic optimization algorithm: harmony search. *simulation*, *76*(2), 60-68.
39. Dorigo, M., & Di Caro, G. (1999, July). Ant colony optimization: a new meta-heuristic. In *Proceedings of the 1999 congress on evolutionary computation-CEC99 (Cat. No. 99TH8406)* (Vol. 2, pp. 1470-1477). IEEE.
40. Karaboga, D., & Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of global optimization*, *39*(3), 459-471.
41. Karaboga, D. (2005). *An idea based on honey bee swarm for numerical optimization* (Vol. 200, pp. 1-10). Technical report-tr06, Erciyes university, engineering faculty, computer engineering department.
42. Mucherino, A., & Seref, O. (2007, November). Monkey search: a novel metaheuristic search for global optimization. In *AIP conference proceedings* (Vol. 953, No. 1, pp. 162-173). American Institute of Physics.
43. Binu, D., & Selvi, M. (2015). BFC: Bat algorithm based fuzzy classifier for medical data classification. *Journal of Medical Imaging and Health Informatics*, *5*(3), 599-606.
44. Das, G. (2013). Bat algorithm based Softcomputing Approach to Perceive Hairline Bone Fracture in Medical X-ray Images. *International Journal of Computer Science & Engineering Technology (IJCSET)*, *4*(04).
45. Isah, R. O., Usman, A. D., & Tekanyi, A. M. S. (2017). Medical image segmentation through bat-active contour algorithm. *International Journal of Intelligent Systems and Applications*, *9*(1), 30.
46. Alomari, O. A., Khader, A. T., Al-Betar, M. A., & Abualigah, L. M. (2017). Gene selection for cancer classification by combining minimum redundancy maximum relevancy and bat-inspired algorithm. *International Journal of Data Mining and Bioinformatics*, *19*(1), 32-51.
47. Kishore, P. V. V., Srivathsav, P. D., Manikanta, M., Venkatram, N., Reddy, L. S. S., Goutham, E. N. D., ... & Sastry, A. S. C. S. (2015). Medical image watermarking with PSNR optimization in wavelet domain based on bat algorithm. *Journal of Theoretical and Applied Information Technology*, *80*(3), 528.
48. Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, *60*(2), 91-110.
49. "*Regularization (Mathematics)*", https://en.wikipedia.org/wiki/Regularization_(mathematics).
50. Mishra, S., & Panda, M. (2018). Bat algorithm for multilevel colour image segmentation using

entropy-based thresholding. *Arabian Journal for Science and Engineering*, *43*(12), 7285-7314.

51. Zhang, J. W., & Wang, G. G. (2012). Image matching using a bat algorithm with mutation. In *Applied Mechanics and Materials* (Vol. 203, pp. 88-93). Trans Tech Publications Ltd.
52. Alihodzic, A., & Tuba, M. (2014). Improved bat algorithm applied to multilevel image thresholding. *The Scientific World Journal*, *2014*.
53. Zhang, R., Shen, J., Wei, F., Li, X., & Sangaiah, A. K. (2017). Medical image classification based on multi-scale non-negative sparse coding. *Artificial intelligence in medicine*, *83*, 44-51.