iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 Issue: XII Month of publication: December 2017

DOI:

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 5 Issue XII December 2017- Available at www.ijraset.com

Lung Image Classification using Hill Climbing Approach on Gabore Filter Optimization

Mr. Ganesh D [1], Dr. S.K .Mahendran MCA[2]
[1]*Research Scholar, Bharathiar Univerisity, Coimbatore, India.*
[2]*Assistant Professor, Government Arts College, Udhagamandalam, India.*

Abstract: Lungs are essential organs which are involved in respiration and disorders involving the lung can prove to be mortal. This work involves with detecting and classifying lung diseases by effective feature extraction through Gabor filter, feature selection through Minimum Redundancy Maximum Relevance (MRMR) and the results are classified by the Instance Based Learning (IBL) and Random Forest (RF) classifiers. Noises are removed through preprocessing techniques and removal of useful features in the image is done through feature extraction and the top ranking features are optimized through feature selection technique and images are classified through classifiers and the measures of performance were found for the same. A new hybrid technique called Shuffled Frog Leaping Algorithm (SFLA) - Hill Climbing (HC) is projected, which presented HC to SFLA through the combination of fast search technique of HC and global search plan of SFLA. The result shows that better performance is got through the proposed technique.
Keywords: Image Classification, Lung Diseases, Gabor Filter, Minimum Redundancy Maximum Relevance (MRMR), Instance Based Learning (IBL), Random Forest (RF), Shuffled Frog Leaping Algorithm (SFLA) and Hill Climbing (HC).

I. INTRODUCTION

The process of classification include image sensor and pre-processing, object recognition and segmentation, feature extraction and object classification. Classification system is made up of database which has predefined pattern that compares the object that is detected and classify it into a proper category. In multiple application domains such as biomedical imaging, biometry, video surveillance, vehicle navigation, industrial visual inspection, robot navigation, and remote sensing, image classification is essential and a challenging task.

The process of classification takes place in the following steps:

1) Preprocessing [2] – atmospheric correction, noise removal, transforming image, main component analysis and so on are included in this.
2) Detection and extraction of an object – Detection of position is included in this and other features of object image that moves are obtained from camera; in extraction, from the object that is detected, the trajectory of the object is estimated in the image plane.
3) Training – This includes the virtue of selection with pattern description to its best.
4) Classification of the object – These classifies the objects that are detected into classes that are predefined with the use of suitable technique that compares the pattern of images with the target patterns.

Spectral information is represented through image classification via digital numbers in one or more spectral bands and entire pixel is categorized in a digital image into one of several land cover classes. Every pixel is allocated in the image to specific classes of themes for instance, water, coniferous forest, deciduous forest, corn, wheat, etc. Image classification's aim is identification as an exclusive gray level (or color), the characteristics which occur in an image with effect to the object or type of land cover these characters that actually represent on the ground. There are three types of classification:

a) Supervised – here both the principles of spectrum and classes are utilized in training samples;
b) Unsupervised – Difference in spectral value are determined here;
c) Hybrid – Both supervised and unsupervised classifications are used together [3].

In general, image classification is a two-stage process. Initially, from the training data, the features encapsulate image information that are removed. Elucidation of feature extraction is projected from image to feature space. Such features exhibit two features.

5) Low dimensionality – In real world image classification issues, there is common conflict of high dimensional data. With respect to effectiveness of computation, image information are compressed through feature removal techniques of much smaller dimension by taking advantage of redundancy in the information of image pixel intensity.

6) Discriminability: This is a characteristic which can classify of class of images from another group. A good understanding is required of the realm of application and imaging physics underneath. Various types of features are recognized using different techniques and application domain.

II. METHODOLOGY

This section uses attribute extraction with Gabor filter technique, MRMR method is used for feature selection and IBL and RF for classification. Additionally, hybrid SFLA-HC optimization technique is presented.

A. Feature Extraction Using Gabor Filter

Gabor filters have found its application in various computer related techniques and gets its ingenuity from biological findings on the similarity of 2D Gabor filters and receptive fields of neurons in the visual cortex. In both spatial and frequency domains there is optimal joint localization. Various image structure examination applications find its usage in edge detection, texture analysis, image coding, handwritten number recognition, face recognition, detection of vehicle and image retrieval [18]. The general functionality of the 2-D Gabor filter family can be represented as a Gaussian function modulated by a complex sinusoidal signal. Specifically, a 2-D Gabor filter $g(x,y)$ can be formulated as (1 & 2):

$$g(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{\tilde{x}^2}{\sigma_x^2}+\frac{\tilde{y}^2}{\sigma_y^2}\right)\right]\exp\left[2\pi jW\tilde{x}\right] \quad (1)$$

$$\begin{cases}\tilde{x}=x\ \cos\theta+y\ \sin\ \theta\\ \tilde{y}=-x\ \sin\theta+y\ \cos\theta\end{cases} \quad (2)$$

Where σ_x and σ_y are the scaling parameters of the filter and determine the effective size of the locality of a pixel in which the weighted summation takes place. $\theta(\theta\in[0,\pi])$ define the orientation of the Gabor filters. W is the radial frequency of the sinusoid. A filter will answer stronger to a bar or an edge with a standard parallel to the orientation θ of the sinusoid. The Fourier transform of the Gabor function in (1) is given by (3):

$$G(u,v)=\exp\left[-\frac{1}{2}\left(\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)\right] \quad (3)$$

Where $\sigma_u=\frac{1}{2\pi\sigma_x,\sigma_v}=\frac{1}{2\pi\sigma_y}$. The Fourier domain representation in (3) specifies the amount by which the filter modifies each frequency part of the input image.

B. Minimum Redundancy Maximum Relevance (MRMR) Feature Selection

Optimal features are selected for this classification so as to address the issue in exploring MRMR method. For a feature set S with n_0 features $\{x_i\}$, (i = 1... n_0). Maximum relevance is to search for features such that the mutual information values between individual feature and target should be maximized [20]. Let D(S, y) be the mean of the mutual information between individual features and target y. It is formally defined in (4):

$$\max D(S,y)=\frac{1}{|S|}\sum_{x_{i\in S}}I(x_i,y) \quad (4)$$

While considering target class, there might be strong separability between on the two target class and it might not be desirable to include them if they are very correlated. Features should be extracted with the idea of minimum redundancy so that they are unlike maximally. Let R(S, y) be the mean of the mutual information between set of features in S. It is formally defined in (5),

$$\min R(S) = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i, x_j) \qquad (5)$$

The criterion combining the above two constraints is called MRMR. The MRMR feature set is obtained by maximizing D(S, y) and minimizing R(S) simultaneously, which requires combining the two measures into a single criterion function.

C. Instance Based Learning (IBL) Classifier

IBL belongs to a family of machine learning algorithms and is also called memory-based and case-based learning. Only when there is request for prediction, IBL algorithms store and process the data and so it is called lazy learning method [21]. Based on the stored samples, predictions are derived and accomplished through nearest-neighbor estimation principle. Let the distance measure Δ (·), i.e., Δ (x, x_0) is the distance between instances x, x' . Usually Euclidean distance is used and attributes are normalized. Distance between two instances x_i *and* x_j is defined as (6):

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^{m} \left(a_r(x_i) - a_r(x_j)\right)^2} \qquad (6)$$

Attribute are normalized by (7):

$$a_r = \frac{v_r - \min v_r}{\max v_r - \min v_r} \qquad (7)$$

Where v_r is the actual value of the attribute is the output space and $< x, y > \in X \times Y$ is called a labelled instance, a case, or an example. In classification, Y is a finite (usually small) set comprised of m classes $\{\lambda_1, ..., \lambda_m\}$, whereas Y = R in regression.

The number of training instances stored is reduced by IBL to a small set of representative example. Another advantage of IBL is it can be used in problems other than classification.

D. Random Forests (RF) Classifier

Number of trees is present in RF multi-way classifier where there is some form of randomization for each tree grown. Every tree is labelled by estimates of posterior distribution over image classes. A test is contained in each internal node which best splits the data that is to be classified. By sending down every tree, an image is classified and the reached leaf distributions are aggregated. At two points, randomness can be injected – in subsampling the training data so that each tree is grown using a different subset; and in selecting the node tests [22].

E. Proposed SFLA-HC Algorithm

In this section, the proposed SFLA-HC approach is described. Gabor filter optimization corresponds to selecting the proper values for each of the four parameters in the parameter set $\Phi = \{\theta, W, \sigma_x, \sigma_y\}$.

The basis for SFLA is multitude intelligence heuristic computing technology, which produce high computing performance and good global search capability. Frogs which can communicate with each other forms the population. Each frog is a memetic vehicle and the evolution of frog population take place through population communication. The original algorithm has some advantages such as simple steps, a few parameters, fast speed and easy realization. However, some drawbacks are also existed in the original algorithm, such as non-uniform initial population, slow convergent rate, limitations in local searching and adaptive evolution in all population has ability and premature convergence [23]. During the evolution process of SFLA, without flog leaping length constraint, frogs with the worst fitness learns information from the memeplex best X_b, or the best frog of entire population X_g, or is substituted by

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 5 Issue XII December 2017- Available at www.ijraset.com

a randomly generated solution. And a lumber process occurs after each memeplex evolving for a specific iteration number in parallel.

HC algorithm is a loop which continuously moves towards increasing value. When the peak is reached, it stops. This is one of the easy procedures in implementing heuristic search. The main objective of HC is that if one has to find the top of the hill then he has to go up the direction from wherever he is. In this heuristic, both depth first and breath first searches are understood in a single method. The situation of the person climbing the hill is the source of the name derivation. The person will move in the direction of the top of the hill. The crest of the hill is the highest value of heuristic function and his movement stops as he reaches the peak. Most of the unproductive search space is eliminated through the knowledge of the local terrain which acts as a useful heuristic. It is a branch by a local evaluation function. The HC is a variant of generate and test in which direction the search should proceed. At each point in the search path, a descendant node that appears to reach for exploration [24].

III. RESULTS AND DISCUSSION

In this work the Gabor filter parameters are optimized using SFLA-HC for efficient feature extraction to classify lung Images. Three types of lung images are classified after feature extraction namely normal (200 CT images), bronchiectasis (100 CT images) and pleural effusion (75 CT images). The SFLA-HC Gabor-IBL, SFLA-HC Gabor-RF, SFLA-HC Gabor-MRMR – IBL and SFLA-HC Gabor-MRMR- RF methods are used. The classification accuracy for normal, bronchiectasis and pleural effusion as shown in table 1 and figure 1.

Table 1 Classification Accuracy

Techniques	Classification accuracy
SFLA-HC Gabor-IBL	93.02
SFLA-HC Gabor-RF	94.29
SFLA-HC Gabor-MRMR - IBL	95.24
SFLA-HC Gabor-MRMR- RF	96.19

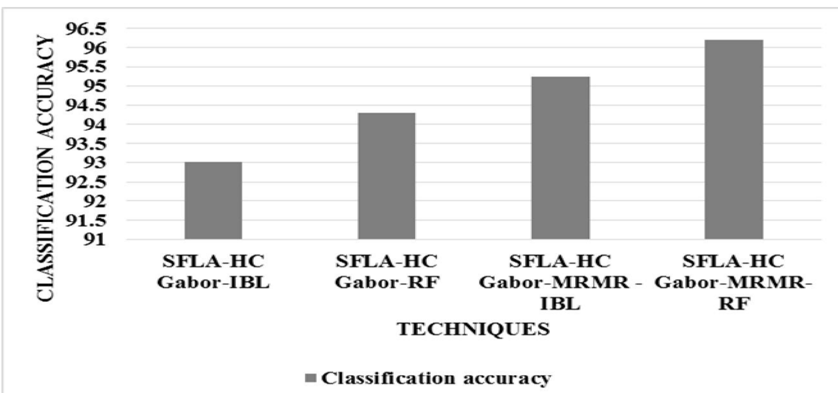


Figure 1 Classification Accuracy

From the figure 1, it can be observed that the SFLA-HC Gabor-MRMR- RF has higher classification accuracy by 3.35% for SFLA-HC Gabor-IBL, by 1.99% for SFLA-HC Gabor-RF and by 0.99% for SFLA-HC Gabor-MRMR - IBL.

IV. CONCLUSION

Many factors may affect image classification. Diverse diseases are referred to by the disorders that affect the lungs such as asthma, COPD, infections such as tuberculosis, influenza, lung cancer, pneumonia and other respiratory issues. Feature extraction using Gabor's theory implies that the information can be presented by the amplitudes of functions that are localized in both space and frequency. This work involves a hybridized optimization technique called SFLA-HC. There is no limit for frog jump over in SFLA through which frog gets out of local optimum. After shuffling is over, frog with the worst fitness is replaced by the best that is searched by HC that helps in the evolution of population efficiently. Results show that the SFLA-HC Gabor-MRMR- RF has higher classification accuracy by 3.35% for SFLA-HC Gabor-IBL, by 1.99% for SFLA-HC Gabor-RF and by 0.99% for SFLA-HC Gabor-MRMR - IBL.

REFERENCES

[1] Nath, S. S., Mishra, G., Kar, J., Chakraborty, S., & Dey, N. (2014, July). A survey of image classification methods and techniques. In Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014 International Conference on (pp. 554-557). IEEE.

[2] Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. International journal of Remote sensing, 28(5), 823-870.

[3] Humayun Khan, P., & Kumar, S. (2014). Survey on Remotely Sensed Image Classification Techniques using Support Vector Machines and Swarm Intelligence. International Journal of Emerging Technology and Advanced Engineering, 4 (1), 507-511.

[4] Rijal, O. M., Ebrahimian, H., Noor, N. M., Hussin, A., Yunus, A., & Mahayiddin, A. A. (2015). Application of phase congruency for discriminating some lung diseases using chest radiograph. Computational and mathematical methods in medicine, 2015.

[5] Mahersia, H., Zaroug, M., & Gabralla, L. (2015). Lung cancer detection on CT scan images: A review on the analysis techniques. International Journal of Advanced Research in Artificial Intelligence (IJARAI), 4(4).

[6] Bhuvaneswari, C., Aruna, P., & Loganathan, D. (2014). Classification of Lung Diseases by Image Processing Techniques Using Computed Tomography Images. International Journal of Advanced Computer Research, 4(1), 87.

[7] Zhou, X., & Wang, J. (2015). Feature selection for image classification based on a new ranking criterion. Journal of Computer and Communications, 3(03), 74.

[8] Singh, H., Verma, S., & Marwah, G. K. (2015). The new approach for medical enhancement in texture classification and feature extraction of lung MRI images by using Gabor filter with wavelet transform. Indian Journal of Science and Technology, 8(35).

[9] O'Neil, A., Shepherd, M., Beveridge, E., & Goatman, K. (2017, July). A Comparison of Texture Features Versus Deep Learning for Image Classification in Interstitial Lung Disease. In Annual Conference on Medical Image Understanding and Analysis (pp. 743-753). Springer, Cham.

[10] Vu, T. H., Mousavi, H. S., Monga, V., Rao, G., & Rao, U. A. (2016). Histopathological image classification using discriminative feature-oriented dictionary learning. IEEE transactions on medical imaging, 35(3), 738-751.

[11] Du, P., Samat, A., Waske, B., Liu, S., & Li, Z. (2015). Random forest and rotation forest for fully polarized SAR image classification using polarimetric and spatial features. ISPRS Journal of Photogrammetry and Remote Sensing, 105, 38-53.

[12] Nayak, D. R., Dash, R., & Majhi, B. (2016). Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests. Neurocomputing, 177, 188-197.

[13] Ramamoorthy, S., Kirubakaran, R., & Subramanian, R. S. (2015). Texture feature extraction using MGRLBP method for medical image classification. In Artificial Intelligence and Evolutionary Algorithms in Engineering Systems (pp. 747-753). Springer, New Delhi.

[14] Veeramani, S. K., & Muthusamy, E. (2016). Detection of abnormalities in ultrasound lung image using multi-level RVM classification. The Journal of Maternal-Fetal & Neonatal Medicine, 29(11), 1844-1852.

[15] Bhuvaneswari, P., & Therese, A. B. (2015). Detection of cancer in lung with k-nn classification using genetic algorithm. Procedia Materials Science, 10, 433-440.

[16] Ladgham, A., Hamdaoui, F., Sakly, A., & Mtibaa, A. (2015). Fast MR brain image segmentation based on modified Shuffled Frog Leaping Algorithm. Signal, Image and Video Processing, 9(5), 1113-1120.

[17] Jothi, G. (2016). Hybrid Tolerance Rough Set–Firefly based supervised feature selection for MRI brain tumor image classification. Applied Soft Computing, 46, 639-651.

[18] Sun, Z., Bebis, G., & Miller, R. (2005). On-road vehicle detection using evolutionary Gabor filter optimization. IEEE Transactions on Intelligent Transportation Systems, 6(2), 125-137.

[19] Bhuvaneswari, C., Aruna, P., & Loganathan, D. (2014). A new fusion model for classification of the lung diseases using genetic algorithm. Egyptian Informatics Journal, 15(2), 69-77.

[20] Zhuo, Z. (2012). Automatic Glaucoma Diagnosis with mRMR-based Feature Selection. Journal of Biometrics & Biostatistics.

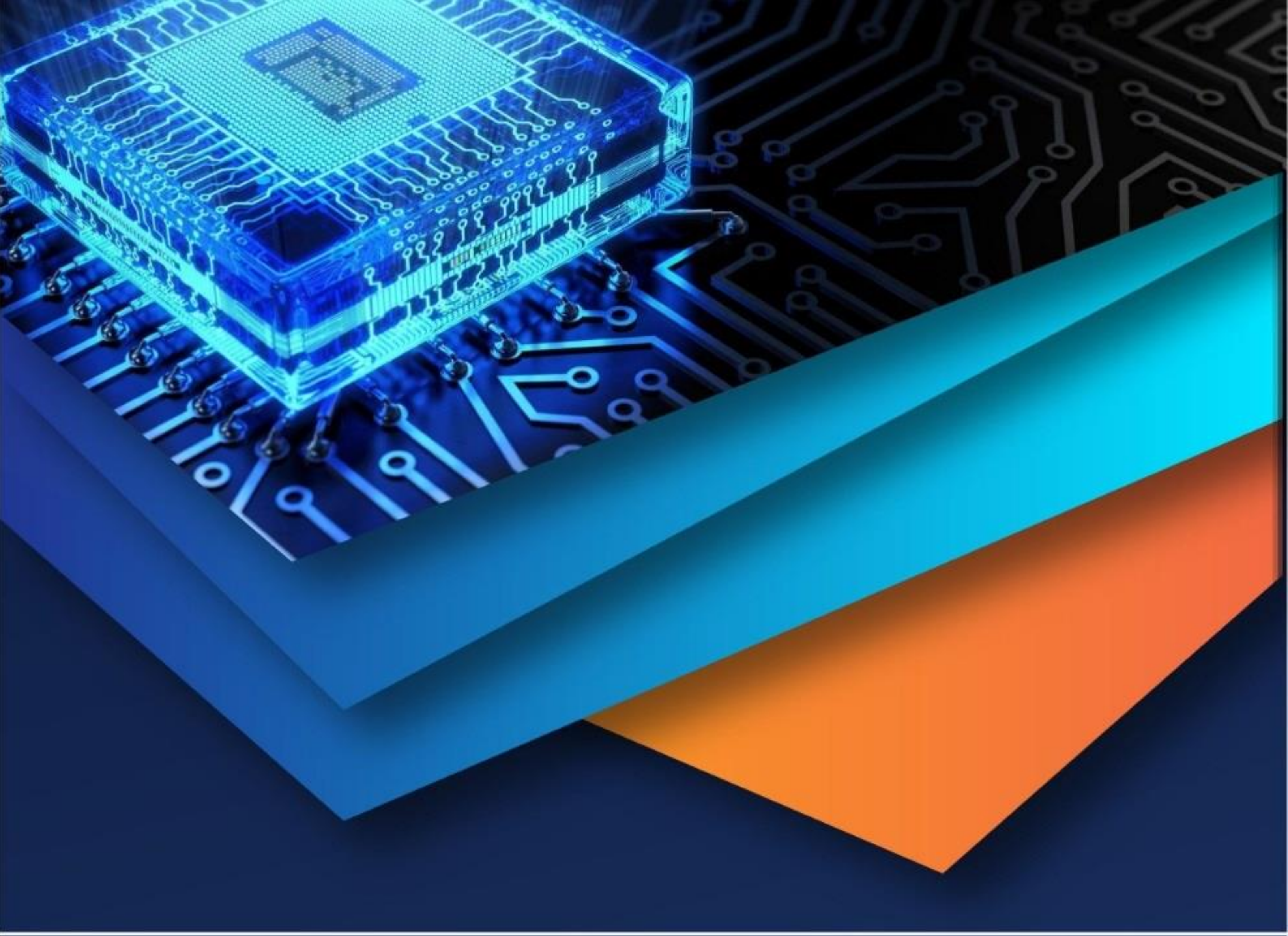
[21] Enireddy, V., & Reddi, K. K. (2012). Application of CART and IBL for Image Retrieval. International Journal of Computer Science and Telecommunications, 3 (12), 62-66.

[22] Bosch, A., Zisserman, A., & Munoz, X. (2007, October). Image classification using random forests and ferns. In Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on (pp. 1-8). IEEE.

[23] Qiu-Hong, S. H. I. (2013). An improved shuffled frog leaping algorithm. Journal of Chongqing University of Technology (Natural Science), 5, 019.

[24] Selman, B., & Gomes, C. P. (2006). Hill-climbing Search. Encyclopedia of Cognitive Science.

[25] Lin, M. J., Luo, F., Xu, Y. G., & Luo, L. (2013). A Novel Hybrid Optimization Method of Shuffled Frog Leaping Algorithm and Particle Swarm Optimization. In Advanced Materials Research (Vol. 717, pp. 433-438). Trans Tech Publications.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY