*Nat. Volatiles & Essent. Oils, 2021; 8(5): 4377 - 4387*

# An Effective Wrapper Fs Based On Binary Swallow Swarm Optimization With Score-Based Criteria Fusion For Image Segmentation And Feature Selection

**[1]Mr.N.Prabhu, 2Dr.M.Shanthakumar, [3]Dr. Shanmugasundaram A, [4]Dr.K.Sathishkumar**

*[1]Assistant Professor, Department of Computer Science(Autonomous), Gobi Arts & Science College, Gobichettipalayam, Tamilnadu, India*
*[2]Assistant Professor, Department of Computer Science, Kamban College of Arts and Science, Sulthanpet, Coimbatore.*
*[3]Assistant Professor, Department of Computer Science, College of Computer Science and Information Technology, Jazan University, Kingdom of Saudi Arabia.*
*[4]Assistant Professor, Department of Computer Science(Autonomous), Gobi Arts & Science College, Gobichettipalayam, Tamilnadu, India*

**Abstract:**
This paper investigates an efficient image segmentation method for medical imaging with the goal of reducing clinicians' interpretation of computer tomography (CT) scan images. Large pictures generated by modern medical imaging modalities are particularly difficult to manually evaluate. The results of segmentation algorithms are determined by their precision and convergence time. There is a pressing need to investigate and develop novel evolutionary algorithms to overcome the challenges related with medical picture segmentation at the moment. Lung cancer is the most often diagnosed cancer in males all over the world. Early identification of lung cancer leads to appropriate treatment, which saves lives. An Effective Wrapper FS Based on Binary Swallow Swarm Optimization with Score-Based Criteria Fusion for Image Segmentation and Feature Selection scheme is proposed in this article. The experimental outcome of the proposed scheme is compared with existing approaches and the proposed lung cancer detection scheme outperforms the existing approaches.

**Keywords:** Lung cancer, feature selection, segmentation, optimization, classification and wrapper scheme.

## Introduction

Image processing is incorporated in various fields to carry out a certain operations on the digital image to enrich the quality of the image or to uncover the needed information's from the images. Image processing is a kind of signal processing, whereas it takes input as an image and produces output as a needed element that is associated with an image. The chief intent behind image processing is enhancing the visualization of needed content by the progression of substantial parameters. The key applications of image processing are renovating the contrast, resolution, brightness, and the level of noise in the image. The image processing approach is widely used in the medical field for the effective diagnosis of disease.

Lung cancer is an important cause of death and it is reported in many of the developed as well as underdeveloped countries with low survival rates after diagnosis. But, lung cancer has a high chance of being cured effectively, if it is treated early rather than letting the disease progressed. The diagnosis of lung cancer is mostly based on the identification of the lung nodule, commonly known as the pulmonary nodule. Lung nodules are a form of pathological tissue growth that may cause lung cancer. They are usually round or spherical in structure with a diameter of up to 30mm. Early lung nodule detection is a significant form of effective therapeutic diagnosis and prevention of lung cancer. Therefore, Computed Tomography (CT) scans must be the major decision for detecting lung nodules. Analysis of CT images via physical processing is a time utilization process for radiologists since there are hundreds of slices in one scan and less than 100 voxels in a single nodule.

Malignant growth therapy is according to the clinical and actual examination just as characterizes toclinical history. Nonetheless, this treatment devours additional time. It will be extremely late for restoring the patient if a growth is found in its essential stage. Also, it is exceptionally critical for clinical researches for planning treatment processes as indicated by the reasonable information which comprises of adequate for recognizing the ailment.

Sifting and covering strategies are extensively applied for picking the critical elements. Sifting strategies choose credits during pre processing with no arrangement calculations. However, the covering strategies utilize the accuracy of a strategy for evaluating the likely subgroup of qualities and picking the subgroup of characteristics that accomplishes the greatest accuracy. Despite the fact that covering methods ordinarily comprise more noteworthy accuracy contrasted with the channel draws near, they are profoundly cost and troublesome of over fitting. Improvement strategies are regularly embraced due to their computational productivity.

The significant reason for this article is to foster an improved flexible covering highlight decision technique for anticipating malignancy through KNN, SVM and RNN classifiers. It wires two capacities like SCF and exactness for estimating the comparability among qualities and class, also as estimating the repetition in the picked ascribes. The significant arrangement is to procure the characteristics which are to a great extent connected with the class and independent of each other.

The leftover segments of this paper is organized as follows: Section 2surveys the past works related with the cellular breakdown in the lungs discovery, Section 3 portrays the proposed philosophy for highlight choice and division, Section 4 shows the investigational results of the proposed conspire and Section 5 closes the proposed work.

## Literature survey

Kim et al. [11] proposed an original Multi-scale Gradual Integration CNN (MGI-CNN) structure for learning highlight portrayals of multi-scale inputs with a steady component extraction strategy. In this structure, three essential procedures were utilized, for example, using multi-scale inputs with various degrees of logical data, utilizing conceptual data inborn in various information scales with progressive coordination and learning multi-stream include mix in a start to finish way. In any case, bogus positive rate was exceptionally high.

Li et al. [12] proposed a combination calculation by consolidating hand tailored highlights and the elements learned at the yield layer of a 3D profound CNN. At first, unique handmade elements were separated including force, mathematical and surface elements based on dim level co-event network. From that point onward, 3D CNNs were prepared for removing the CNN highlights learned at the yield layer. For every 3D CNN, the CNN highlights melded with the hand tailored elements were utilized as the contribution for the SVM combined with the successive forward include determination technique for picking the ideal element subset and building the classifiers. In any case, the power of this calculation was not compelling.

Cao, P., et al [13] had proposed a multi-piece based strategy for choosing the highlights and learning the imbalanced information in lung knob CAD. In this strategy, a multi-portion include determination was accomplished based on pairwise likenesses from the component level and a multi-bit over-testing for the imbalanced information learning. Be that as it may, the computational intricacy of this technique was high.

An original pneumonic knob location technique [14] was proposed based on Deep CNN (DCNN). At first, a deconvolutional structure was acquainted with Faster Region-based CNN (Faster R-CNN) for identifying lung knob applicants on pivotal cuts. From that point forward, a 3D DCNN was proposed with the guide of dropout plot for lessening the bogus positive in competitor identification. Nonetheless, it doesn't consider the data between the little fixes that were removed in a huge fix.

## Proposed methodology

In covering FS process is as indicated by the particular streamlining to fit on a given dataset. It follows a BWSSO by considering every one of the potential blends of elements in contrast to the assessment. SSO is a populace based metaheuristic. Toward the start of each cycle, the population(features) is arranged by the wellness work figured from SCF and accuracy. Then, at that point, the accompanying jobs are appointed:

1. Head Leader (HL) is a molecule with the ideal upsides of not really settled from SCF and characterization precision;
2. Neighborhood Leaders(LL) are l particles which follow the HL as per wellness figured from SCF and order exactness;
3. Capricious Particles (aip) are k particles with the most noticeably terrible wellness esteem;
4. Different particles are voyagers.

During present cycle, HL doesn't move, filling in as signals for traveler particles that observe the hunt space between the nearest LL and HL. Wayfarer particles change their areas by means of underneath Eqns. (1-4):

$\theta_{exp}(t+1) = \theta_{exp}(t) + V(t+1)$ (1)

$V(t+1) = V_{HL}(t+1) + V_{LL}(t+1)$ (2)

$V_{HL}(t+1) = V_{HL}(t) + rnd(0,1)\left(\theta_{exp}^{best}(t) - \theta_{exp}(t)\right) + rnd(0,1)\left(\theta_{HL}(t) - \theta_{exp}(t)\right)$(3)

$V_{LL}(t+1) = V_{LL}(t) + rnd(0,1)\left(\theta_{exp}^{best}(t) - \theta_{exp}(t)\right) + rnd(0,1)\left(\theta_{LL}(t) - \theta_{exp}(t)\right)$(4)

where $\theta_{exp}, \theta_{HL}$ and $\theta_{LL}$are the location of the explorer, HL and LL closest to the explorer, $\theta_{exp}^{best}$refers to the ideal area, V, $V_{HL}$ and $V_{LL}$ are the speed of the molecule, molecule moving to the HL and molecule moving to the storage room LL. The Eq. to change the areas of floating particles since the genuine Eq. may influence particles for gathering at the edges of the quest space for picking the ideal qualities. It limits the probability of this conduct and works with voyager particles. For adjusting the area of floating particles to pick the ideal qualities, the beneath Eqns. (5-6) are applied:

$\theta_{\mathrm{aip}}(t+1) = rnd(0.5,2).V_{FS}$ (5)

$V_{FS} = \frac{\sum_{j=1}^{N-k} \theta_{exp}^{j}(t+1)}{N-k}$ (6)

Where $\theta_{\mathrm{aip}}$ and $\theta^{j}$are the location of drifting particle, jth particle, N refers to the overall amount of particles in the populaces (attributes from every GEM dataset), and k defines the amount of ‘aip’. After, it terminates if the termination criteria is reached.

It applies a specified combine (cmb) function for updating the particles location to choose the features. Since its input, it accepts 2 feature vectors **A** and **B**, including the amount pr, which computes the effect of **A**on **B**, between 0 and 1. It combines the feature vector **C**every component as: if $A_i$and $B_i$overlap, then they have similar range; or else, $C_i$ assigns $A_i$ range with the likelihood pr or $B_i$ with the likelihood (1–pr).

$$\mathrm{cmb}(A,B,pr)_i = \begin{cases} A_i if rnd(0,1) < pr, i = 1,..D \\ B_i, else, i = 1,..D \end{cases} (7)$$

On the off chance that a property accomplishes more noteworthy exactness and SCF, this quality is profoundly fitting. Or disaster will be imminent, it is exceptionally commotion. In this way, it works with an upgraded answer for having more characteristics which are perhaps fitting and eliminating more traits which are conceivably clamor. It signifies the picked highlights as vectors FS which encode ascribes. At beginning advance, a gathering of vectors is haphazardly made. The measure of vectors in the general population is fixed. For each element vector, precision and not really settled freely through KNN, SVM and RNN. During each cycle, each component vectors are positioned in dropping request of precision and SCF. Wayfarer particles alter their ideal component areas as per the ideal element areas of the leaders. The following Eqns. (8-11)are for wayfarers:

$$FS_{exp}(t+1) = \mathrm{cmb}\left(V(t+1), FS_{exp}(t), pr_{v_{exp}}, we_{exp}\right) (8)$$

$$V(t+1) = \mathrm{cmb}\left(V_{HL}(t+1), V_{LL}(t+1), pr_{v_{\mathrm{hl}}}, we_{\mathrm{hl}}\right) (9)$$

$$V_{HL}(t+1) = \mathrm{cmb}\left(\mathrm{cmb}\left(FS_{HL}(t), FS_{exp}(t), pr_{\mathrm{hexp}}, we_{\mathrm{hexp}}\right), rnd(0,1)^D, pr_{\mathrm{hexpr}}, we_{\mathrm{hexpr}}\right)(10)$$

$$V_{LL}(t+1) = \mathrm{cmb}\left(\mathrm{cmb}\left(FS_{LL}(t), FS_{exp}(t), pr_{\mathrm{lexp}}, we_{\mathrm{lexp}}\right), rnd(0,1)^D, pr_{\mathrm{lexpr}}, we_{\mathrm{lexpr}}\right)(11)$$

Where $FS_{HL}, FS_{LL}$ and $FS_{exp}$ are the feature locations of the HL, LL and explorer, $pr_{v_{exp}}, pr_{v_{\mathrm{hl}}}$are the likelihood of the velocity vector on the feature location of the explorer and HL, $pr_{\mathrm{hexpr}}$ refers to the likelihood of HL and explorer, $pr_{\mathrm{lexp}}$ stands the likelihood of LL with explorer and $pr_{\mathrm{lexpr}}$ defines the likelihood of LL and explorer. $we_{exp}$represents the weight of the velocity vector on the feature location of the explorer, $we_{\mathrm{hl}}$ defines the weight of the velocity vector regarding HL and LL, $we_{\mathrm{hexp}}$ stands for the weight of HL and explorer, $we_{\mathrm{hexpr}}$ represents the weight of HL and explorer, $we_{\mathrm{lexp}}$ refers to the weight of LL and explorer, and $we_{\mathrm{lexpr}}$ denotes the likelihood of LL and explorer.

## FINTESS COMPUTATION

Wellness is processed by consolidating exactness and SCF. In case both are higher than the highlights are chosen from the GEM dataset. Not really settled from KNN, SVM, and RNN. Chosen highlights are given to classifier, and afterward exactness will likewise be considered with wellness alongside SCF rules.

Grouping exactness is a small portion of the measure of definitively dispensed classes to the general measure of substances to be anticipated. In SCF, the Symmetrical Uncertainty (SU) and Relief F are coordinated through combination instrument i.e., either score-or rank-based [13]. Here, a score-based combination is utilized. Along these lines, each major rules makes the score vectors having upsides of each attribute and melded as a solitary vector [14]. Then, at that point, ascribes are positioned dependent on the resultant score vector. The combination technique joins 2 score vectors by means of increasing the weight factor. It is portrayed as [23]:

$$R_{i,cl} = \mu SU_{i,cl} + (1-\mu)W_i (12)$$

In Eq. (12), the weight factor$\mu \in [0,1]$creates a trade-off between SU and Relief F, and it estimates the fundamental conditions to the estimated relevance. SU is adopted for compensating the bias of mutual data. It is defined as [23]:

$SU(X,Y) = \frac{2I(X,Y)}{H(X)+H(Y)}$ (13)

In Eq. (13), SU∈ [0, 1] and symmetric. When SU is 1, it defines that the data of a random variable X is fully predict the data of Y, and if it is 0, then X and Y are sovereign. H(X) and H(Y) are margin entropy, $I(X,Y)$is mutual information [23].

Relief F is as per the properties wellness for differentiating close by information to gauge the highlights weight. At each iteration, a information R is picked haphazardly, and ReliefF searches KNN from one another information classes. For a trait, if the space among R and the contiguous hit is more modest contrasted with the space among R and the nearby miss M, then, at that point, it is successful to separate information from shifted classes and its weight must be expanded. Or the consequences will be severe, it is insufficient and its weight must be limited. For an information base having n ascribes, it is proceeded ntimes for assessing the weight W for every resultant characteristic [23].

Here, SCF score created by SU ∈[0, 1], when the score span delivered through Relief F is unsure. In this way, score regularization is directed for Relief F in advance of the score combination. It is given as [23]:

$u_i' = \frac{u_i - u_{min}}{u_{max} - u_{min}}$ (14)

A gathering of helpful qualities is extremely appropriate to classes. Here, the repetition in the picked property subgroup is assessed [29] which regularizes the shared information for disposing of its biasas:

$NI(f_i, f_s) = \frac{I(f_i, f_s)}{min\{H(f_i), H(f_s)\}}$(15)

$D_i = \frac{1}{|S|} \sum_{f_i, f_s \in S} NI(f_i, f_s)$(16)

Where |S| is the amount of chosen attributes and its reciprocal is applied for balancing similarity and redundancy. It creates an OSCF as:

$G = R_{i,cl} - D_i$ (17)

Eventually, an appropriate and less redundant attribute subgroup is decided via increasing the cost factor (17). Observe that when$\mu = 0$and $\mu = 1$, respectively Eq. (17) becomes Eqns. (18) and (19) which are the OSCF variants. These are considered as redundancy to SU and ReliefF.

$G_1 = W_i - D_i$(18)

$G_2 = SU_{i,cl} - D_i$(19)

Also, the search policy of OSCF uses an incremental forward choice. Later, the selected features are classified with the assistance of CNN.

## Results and Discussion

The prediction accuracy and the classification performance of the proposed WFS-CNN and the existing MGI-CNN is compared and the outcomes are evaluated [15-26].

### Accuracy

Accuracy is defined as the closeness of a given value from a set of classified examples. The accuracy shows systematic flaws and statistical bias. It is the calculation of the true value as well as the identification (both TP and TN values) of the predicted classes amongst count. Variation between the resultant and genuine resultant values occurs when the lowest accuracy occurs. It's the proportion of accurate disease detection to the total number of cases examined. It's calculated as follows:

$$Acy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative}$$

**Precision**

The closeness of the measurement and the importance among the values identified are indicated by the positive analytical value or precision. Random errors are expressed as precision, which is calculated using statistical factors. Precision and accuracy are phrases that are interchangeable. Precision is directly proportional to the percentage of positive values in the whole population. The count of true positive values is the precision value for a particular problem in the classification process (i.e. the count of the item correctly labelled as positive classes). The algorithm with maximum precision achieves more necessary information than extraneous data as a consequent value. It is calculated as,

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Table 1. Comparison of WFS-CNN Accuracy and Precision with MGI-CNN

| Iteration | Accuracy | | Precision | |
|---|---|---|---|---|
| | MGI-CNN | WFS-CNN | MGI-CNN | WFS-CNN |
| 50 | 0.904 | 0.933 | 0.887 | 0.943 |
| 100 | 0.9404 | 0.964 | 0.884 | 0.945 |
| 150 | 0.952 | 0.973 | 0.883 | 0.948 |
| 200 | 0.974 | 0.994 | 0.885 | 0.923 |

In Table 1, the outcomes of the proposed approach WFS-CNN and existing approach MGI-CNN is given. In this comparison, the accuracy and the precision values of the algorithms are given.

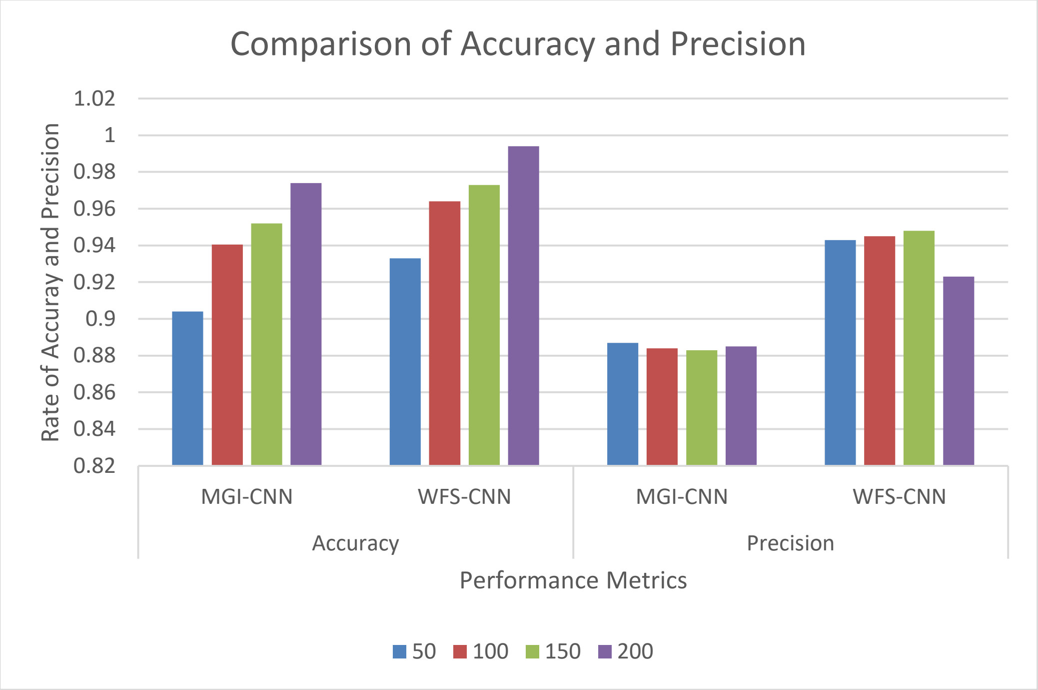

**Figure 1. Comparison of Accuracy and Precision**

From the Figure 1, it is identified that the prediction accuracy of the proposed approach WFS-CNN is high when compared to the existing approach.

### Sensitivity

The ratio of genuine negative values in all samples is known as specificity, and it has no defined condition. The accurate detection of real positive test readings is referred to as sensitivity. It aids in the determination of sample categorization accuracy. The sensitivity is equated as,

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

### Specificity

The ratio of genuine negative values in all samples is known as specificity, and it has no defined condition. The accurate detection of real positive test readings is referred to as sensitivity. It aids in the determination of sample categorization accuracy. The specificity is equated as,

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive}$$

Table 2. Comparison of WFS-CNN Sensitivity and Specificity with MGI-CNN

| Iteration | Sensitivity | | Specificity | |
|---|---|---|---|---|
| | MGI-CNN | WFS-CNN | MGI-CNN | WFS-CNN |
| 50 | 0.781 | 0.829 | 0.799 | 0.819 |
| 100 | 0.819 | 0.879 | 0.895 | 0.899 |
| 150 | 0.862 | 0.916 | 0.882 | 0.875 |
| 200 | 0.91 | 0.929 | 0.899 | 0.901 |

In Table 2, the outcomes of the proposed approach WFS-CNN and existing approach MGI-CNN is given. In this comparison, the sensitivity and the specificity values of the algorithms are given.

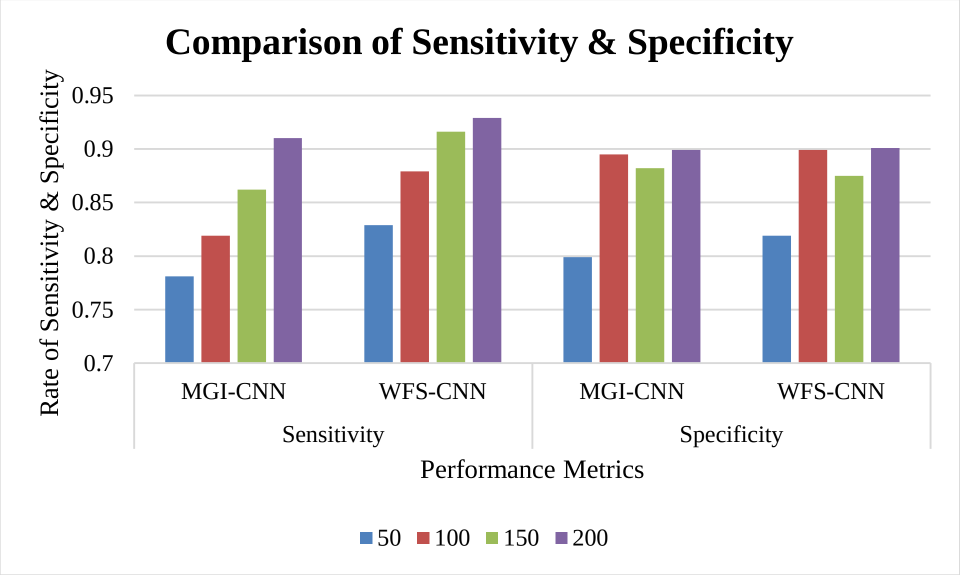

**Figure 2. Comparison of Sensitivity and Specificity**

From the Figure 2, it is identified that the prediction accuracy of the proposed approach WFS-CNN is improved with the acquired high sensitivity when compared to the existing approach. In the WFS-CNN, the sensitivity is high and specificity is low, which shows effective prediction of accuracy. In the MGI-CNN, the sensitivity is low and specificity is high that shows the contrast among the accuracy of prediction.

### Mean Absolute Error Rate

The incidence of error in the digital transformation is due to different variables such as noise, distortion, and disturbance in the data transformation. It's a performance rate ratio. The error rate is the percentage of patterns classified wrongly by the decision-making model. The error rate is calculated by multiplying the sum of the FP and FN values by the sum of the TP,TN, FP, and FN values. It is measured as:

$$Error\ rate = \frac{FP + FN}{TP + TN + FP + FN}$$

Table 3. Comparison of WFS-CNN mean absolute error rate with MGI-CNN

| Training Epochs | MGI-CNN | WFS-CNN |
|---|---|---|
| 100 | 4.0 | 3.7 |
| 200 | 3.3 | 3.0 |
| 300 | 3.0 | 2.9 |
| 400 | 3.8 | 3.3 |
| 500 | 3.6 | 3.4 |

In Table 3, the outcomes of the proposed approach WFS-CNN and existing approach MGI-CNN is given. In this comparison, the mean absolute error rate of the algorithms are given.

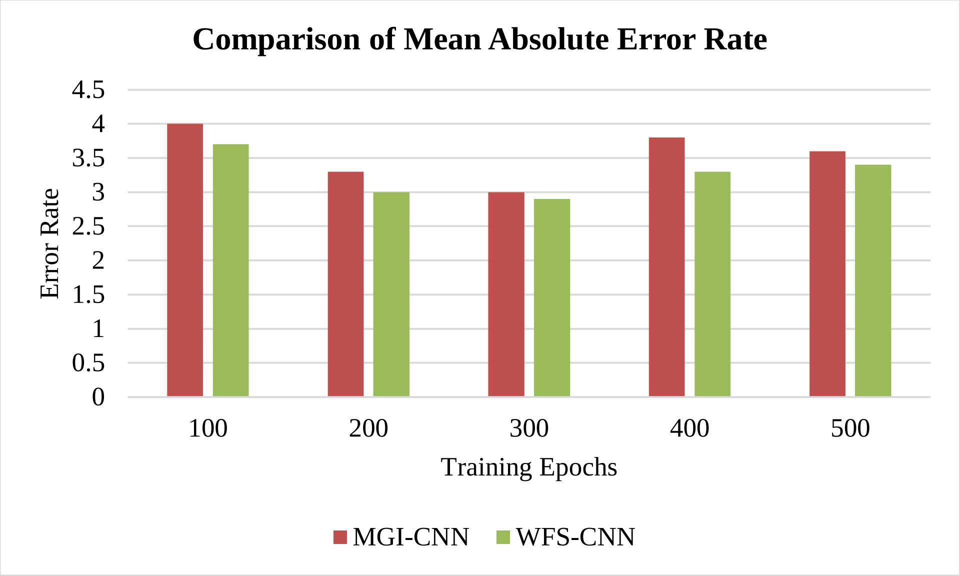

**Figure 3. Comparison of incidence of error rate**

In Figure 3, the error rate values for WFS-CNN and MGI-CNN models are shown. In this graph, x-axis denotes the number of training epochs and y-axis denotes the error rate values.

## Conclusion

This article proposes another Swallow Swarm Optimization with SCF (Optimized SCF) covering FS to anticipate the disease. It follows the course of Binary Weight Swallow Swarm Optimization (BWSSO) to precisely gauge the pertinence. It utilizes a specific consolidate (cmb) work for refreshing the molecule's regions for choosing the elements and weight esteem is likewise utilized for tracking down the component's significance. These two properties will expands the results of Optimized SCF calculation. Wellness of BWSSO calculation is determined from the classifiers CNN. Chosen highlights are given as contribution to classifier and afterward precision and furthermore SCF rules is considered for wellness assessment. At long last, investigational results depicted that the WFS-CNN strategy has preferred execution over the MGI-CNN dependent on accuracy, affectability, explicitness, exactness and mistake rate. The malignancy related highlights are recovered and ordered proficiently with the proposed plot.

## REFERENCES

1. Chander, M. P., Rao, M. V., &Rajinikanth, T. V. (2017). Detection of lung cancer using digital image processing techniques: a comparative study. International Journal of Medical Imaging, 5(5), 58.
2. Liao, X., Zhou, J., Zhang, R., & Zhang, Y. (2012). An adaptive artificial bee colony algorithm for long-term economic dispatch in cascaded hydropower systems. International Journal of Electrical Power & Energy Systems, 43(1), 1340-1345.
3. Yang, J., &Nanni, L. (Eds.). (2011). State of the art in Biometrics. BoD–Books on Demand.
4. Amutha, A., & Wahidabanu, R. S. D. (2013, April). Lung tumor detection and diagnosis in CT scan images. In 2013 International Conference on Communication and Signal Processing (pp. 1108-1112). IEEE.
5. Essays, UK. (November 2013). The History Of Image Processing Information Technology Essay. Retrieved from https://www.uniassignment.com/essay-samples/information-technology/the-history-of-image-processing-information-technology-essay.php?vref=1
6. Shinde, B. S., & Dani, A. R. The Origins of Digital Image Processing & Application areas in Digital Image Processing Medical Images. IOSR Journal of Engineering, 1(1), 066-071.
7. T. Acharya and P. S. Tsai. JPEG2UUU Standard for Image Compression:Concepts, Algorithms, and VLSI Architectures, Wiley, Hoboken, NJ, 2004.

8. Liu, Z. and Tan, M. “ROC-based utility function maximization for feature selection and classification with applications to highdimensional protease data, biometrics,” J. Int. Biometr. Soc., Vol. 64, No. 4, pp. 1155–1161, 2008.
9. Yamamoto, T., Ukai, Y., Kubo, M., Niki, N., Satou, H. and Eguchi, K. “Computer Aided Diagnosis System with Functions to Assist Comparative Reading for Lung Cancer Based on Helical CT Image”, International Conference on Image Processing, Vol. 1, pp. 180 - 183, 2000.
10. Demir, Ö., &YılmazÇamurcu, A. (2015). Computer-aided detection of lung nodules using outer surface features. Bio-medical Materials and Engineering, 26(s1), 1213-1222.
11. Kim, B. C., Yoon, J. S., Choi, J. S., & Suk, H. I. (2019). Multi-scale gradual integration CNN for false positive reduction in pulmonary nodule detection. Neural Networks, 115, 1-10.
12. Li, S., Xu, P., Li, B., Chen, L., Zhou, Z., Hao, H., ... & Jiang, S. (2019). Predicting lung nodule malignancies by combining deep convolutional neural network and handcrafted features. Physics in Medicine & Biology, 64(17), 175012.
13. Cao, P., Liu, X., Yang, J., Zhao, D., Li, W., Huang, M., &Zaiane, O. (2017). A multi-kernel based framework for heterogeneous feature selection and over-sampling for computer-aideds detection of pulmonary nodules. Pattern Recognition, 64, 327-346.
14. Ding, J., Li, A., Hu, Z., & Wang, L. (2017). Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks. arXiv preprint arXiv:1706.04303, 1-9.
15. Neovius, M. G., Linné, Y. M., Barkeling, B. S., &Rossner, S. O. (2004). Sensitivity and specificity of classification systems for fatness in adolescents. The American journal of clinical nutrition, 80(3), 597-603.
16. Huang, J., & Ling, C. X. (2005). Using AUC and accuracy in evaluating learning algorithms. IEEE Transactions on knowledge and Data Engineering, 17(3), 299-310.
17. Chicco, D., &Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC genomics, 21(1), 6.
18. Tuma, RS 2006, ‘Sometimes size does not matter: reevaluating RECIST and tumour response rate endpoints’, Journal of the National Cancer Institute, vol.98, no.18, pp.1272–1274.
19. K. Sathishkumar [2019], “Gene Expression Using Artificial Bee Colony besides Fuzzy C Means and Neuro-Fuzzy Discriminant Analysis”, Journal of International Pharmaceutical Research, ISSN: 1674-0440. (Accepted). (ELSEVIER Scopus Indexed).
20. K. Sathishkumar [2019], "Efficient Clustering Methods And Statistical Approaches For Gene Expression Data", Journal of Advanced Research in Dynamical and Control Systems, ISSN 1943-023X. (Accepted). (ELSEVIER Scopus Indexed).
21. Sathishkumar. K, V. Thiagarasu, E. Balamurugan, David Otto Arthur, [2018], “Clustering Mutual Outline for Multi Assessment Temporal Data and cancer Data”, International Journal of Computer Sciences and Engineering, Vol. 6, No.1, pp. 75-83.
22. Sathishkumar , Dr.V.Thiagarasu, Dr.E.Balamurugan, Dr.M.Ramalingam, [2018], "An Competent Artificial Bee Colony (ABC) and Fuzzy C Means Clustering Using Neuro-Fuzzy Discriminant Analysis from Gene Expression Data", International Journal of Science & Engineering Development Research, ISSN:2455-2631, Vol. 3, No. 4, pp. 24 - 28, April.
23. K. Sathishkumar, Dr.V. Thiagarasu, Dr.E. Balamurugan and Dr.M. Ramalingam, [2018], “Gene Selection for Cancer Classification on Microarray Gene Expression Data”, Journal of Advanced Research in Dynamical and Control Systems, ISSN 1943-023X, Vol. 10, 05-Special Issue, pp. 547-555. (ELSEVIER Scopus Indexed).
24. SathishkumarKarupusamy [2019], “Standard Weight and Distribution Function Using Glowworm Swarm Optimization for Gene Expression Data”, International Conference on Sustainable Communication

Network and Application [ICSCN 2019] held at Surya Engineering College, Tamil Nadu, India during 30-31 July 2019.

25. SathishkumarKarupusamy [2019], “Gene Expression Analysis Using Clustering Methods: Comparison Analysis”, International Conference on Sustainable Communication Network and Application [ICSCN 2019] held at Surya Engineering College, Tamil Nadu, India during 30-31 July 2019.
26. K. Sathishkumar, [2019] “Classification of Micro Array Gene Expression using k-NN, SVM and Naive Classifiers”, 9th International Conference and Exhibition on Spectrometry and Analytical Techniques”, held on April 17-18, 2019 in Osaka, Japan.