DOI: 10.21917/ijsc.2022.0358

GENE SELECTION AND MODIFIED LONG SHORT TERM MEMORYNETWORK-BASED LUNG CANCER CLASSIFICATION USING GENE EXPRESSION DATA

V. Yuvaraj, G. Pandiyan and G. Purusothaman

School of Computer Studies, Rathnavel Subramaniam College of Arts and Science, India

Abstract

Lung cancer is one of the fatal forms of cancer worldwide. Genetic variability has been identified as influencing a person vulnerability to lung cancer in epidemiologic research. A new study undertaken by a team of experts from the United States National Cancer Institute on 14,000 Asian women discovered that Asian women, regardless of whether they smoke or not, are more likely to acquire cancer owing to genetic abnormalities. Early detection of this lethal disease is a novel clinical application of microarray data. Recent research establishes a model for the early diagnosis of lung cancer. Additionally, multilayer perceptron, random subspace, and Sequential Minimal Optimization (SMO) approaches are used for classification. While information acquisition is typically a good indicator of an attribute significance, it is not perfect. A noticeable issue develops when knowledge gain is applied to qualities that might take on many distinct values. This paper provides an efficient gene selection model based on the Improved Whale Optimization Algorithm (IWOA) to address these concerns. It saves time and identifies relevant genes from gene expression data, increasing lung cancer categorization accuracy. Then, a Modified Long Short-Term Memory (MLSTM) Network is used to classify lung cancer. It accepts specified genes as inputs and determines which class they belong to, such as lung cancer or normal subjects. As demonstrated by empirical observations, the suggested model is effective in precision, recall, accuracy, and f –measure.

Keywords:

Lung Cancer, Early Stage, Developing Cancer, Genetic Variations, Feature Selection, Information Gain Attribute, Whale Optimization, Long Short Term Memory

1. INTRODUCTION

It is estimated that lung cancer is responsible for more than a quarter of all cancer-related deaths each year. This equates to the mortality of 1.6 million people and the emergence of 1.8 million cases in 2012. Both men and women in the United States die from it, and it kills more people than colorectal, breast and prostate cancers put together. There is only a 17.8% survival rate for patients with lung cancer that could have been much higher if the disease had been discovered earlier. Only 15% of the cases were caught in the early stages, though. This demonstrates the significance of early detection in fight against lung cancer [1] [2].

Lung cancer is the leading cause of cancer death in smokers. As a result, it has been claimed that cancer may result from an individual genetic propensity inherited from the family. In other words, some people are genetically predisposed to developing lung cancer because of genetic mutations or flaws in a gene. However, many studies have revealed that relatives of people who have had lung cancer are more likely to develop the disease themselves. In addition, there is evidence that genetic variation in the individual response to carcinogens may influence the risk of lung cancer [3] [4].

The innovation in DNA microarray technology has allowed us to measure thousands of genes in a single cell or tissue. Using microarray technology, researchers searched for cancer categorization and resulted in prediction markers in a wide range of tumours [5]-[7].

In recent work introduces a model for lung cancer detection at an early stage. In which feature selection was made by using the information gain attribute ranking method. And classification using multilayer-perceptron, Random subspace and SMO (Sequential Minimal Optimization) methods. However, while information gain is typically a good measure for determining the significance of a feature, it is not ideal. When information gain is applied to qualities that potentially have many distinct values, a significant issue arises.

This work first proposes an efficient gene Selection model using an improved whale optimization algorithm (IWOA) to overcome these issues. It reduces time consumption and provides significant genes from the gene expression data to enhance the lung cancer classification performance. An MLSTM Network is employed for lung cancer classification. It takes selected genes as inputs and classifies them to which class it belongs, like lung cancer or normal subjects.

The paper is divided into the five sections listed below. Section 1 discusses lung cancer categorization using gene expression data. Section 2 provides a discussion of the various strategies for classifying lung cancer. Section 3 develops the design technique for the suggested lung cancer categorization utilizing a gene expression data model. Section 4 outlines the experimental study and contains several findings analyses. Finally, section 5 concludes the work and discusses future plans.

2. LITERATURE REVIEW

Salem et al. [8] introduced a novel technique that uses gene expression profiles to categorise human cancer disorders. Information Gain (IG) and the Standard Genetic Algorithm (SGA) are combined in the Standard Genetic Algorithm (SGA). Genetic Programming (GP) and Genetic Algorithms (GA) are used first to pick features and subsequently to reduce the number of features. The proposed system is evaluated by analysing seven cancer datasets and comparing their classifications to the most recent techniques. When applied to cancer datasets, the suggested system outperforms all other machine learning methods except for the Genetic Algorithm, which generally improves the classification performance of different classifiers.

Motieghader et al. [9] proposed a hybrid meta-heuristic approach by combining Genetic Algorithms and Learning Automata (GALA) for gene selection. The time complexity of GALA is $O(G.m.n^3)$, and it has higher accuracy and performance. All AML, SRBCT, MLL Tumors_9 and Tumors_11 cancer datasets were used to test GALA performance. The GALA

algorithm outperformed other recently suggested algorithms on each dataset, according to the evaluation method.

Lu et al. [10] developed a hybrid feature selection algorithm integrating the mutual information maximization (MIM) and the adaptive genetic algorithm (AGA). According to the experiments, the proposed MIMAGA-Selection approach decreases the size of gene expression data and eliminates redundant information for categorization. Compared to typical feature selection algorithms, the reduced gene expression dataset delivers the best classification accuracy. Furthermore, four different classifiers are employed in the reduced dataset to show this suggested technique strength.

Berger et al. [11] developed an expression-based variant-impact phenotyping (eVIP) approach to differentiate effectual from normal somatic mutations. eVIP found that 69% of the mutations tested had a significant impact, whereas 31% had no effect. Some of the considerable modifications cause xenograft tumour growth in mice and confer resistance to cellular EGFR inhibition. Several uncommon somatic, clinically actionable mutations, such as EGFR S645C, ARAF C and S214F, ERBB2 S418T, and various BRAF variations, are included in this list, illustrating that even rare mutations can have a significant impact on cancer function and treatment.

Alshamlan et al. [12] designed a new hybrid technique called Genetic Bee Colony (GBC) algorithm by integrating Genetic Algorithm (GA) with Artificial Bee Colony (ABC) algorithm. Extensive tests were undertaken to evaluate the correctness of the suggested method. Colon, leukaemia, and lung are three of the three binary microarray datasets used. SRBCT, lymphoma, and leukaemia are among the other three multi-class microarray datasets used. When mRMR is used in conjunction with the Artificial Bee Colony algorithm, the GBC algorithm results are compared (mRMR-ABC). Both the mRMR-GA and mRMR-PSO algorithms were compared to see how well they worked together. The GBC algorithm outperforms the competition with the highest classification accuracy and the lowest number of selected genes. Thus, using the GBC method to solve the gene selection problem in binary and multi-class cancer classification is a promising technique.

Ghaddar and Naoum-Sawaya [13] designed a new approach based on iteratively adjusting a bound on the $l1$-norm of the classifier vector to urge the chosen characteristics to unite with a predefined maximum limit. Analysis of real-world classification issues using high dimensional features. Cancer categorization based on gene expression is used for the first time in the medical diagnosis of cancers using microarray data. Amazon, Yelp, and IMDb reviews are the subject of the second case study, which focuses on sentiment classification. The suggested approach is computationally tractable and yields minimum error values crucial for developing sophisticated decision-support systems, as shown by the results.

Jain et al. [14] developed a two-phase hybrid model for cancer classification using Correlation-based Feature Selection (CFS) and improved-Binary Particle Swarm Optimization (iBPSO). Using a Naive–Bayes classifier, this model chooses a low-dimensional list of prognostic genes for binary and multi-class cancer samples. Eleven cancer microarray datasets were employed to evaluate the accuracy of the suggested approach. Our model performed better in classification accuracy, and the number of genes picked for study compared to known techniques. In addition, seven datasets with a very tiny prognostic gene subset (up to 1.5 %) achieved up to 100% categorization accuracy.

3. PROPOSED METHODOLOGY

Lung cancer categorization is presented in this section. Two-step process including gene selection and categorization utilizing modified LSTM. The Fig.1 depicts the overall design of the method.

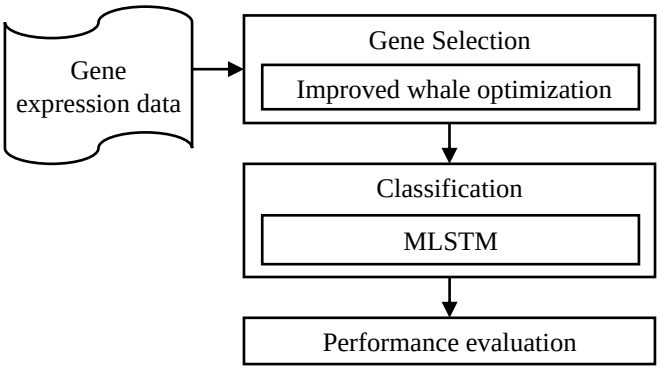

Fig.1. Overall architecture of the proposed model

3.1 GENE SELECTION USING IMPROVED WHALE OPTIMIZATION (IWOA)

Only a tiny subset of the input gene expression data show differential profiles for various classes or samples for this study. Therefore, improved Whale Optimization (IWOA) is employed in this study to aid in selecting genes. The WOA is a new stochastic optimization technique that was recently created based on nature-inspired population dynamics. An optimization problem can be solved using the WOA search agents. The WOA employs a method known as bubble-net hunting to mimic the behaviour of humpback whales in their quest for prey. Encircling the prey, attacking with a bubble net, and searching for the best prey are the three basic steps of the WOA.

The central mathematical part of the WOA is shown in Eq.(1) and Eq.(2).

$$X(t+1) = X^*(t)-A.|C.X^*(t)-X(t)| if p<0.5 \quad (1)$$

$$X(t+1)|C.X^*(t)-X(t)|.e^{bl}\cos(2\pi t)+X^*(t) if p\geq 0.5 \quad (2)$$

The best solution since then is X^*, in which X is a vector representing all of the whales positions, t indicates time, and a constant value defines the shape of the logarithmic spiral, b, which is set to 1 to maintain a linear reduction in a coefficient vector a, such that $A=2a.(r-a)$; $C=2.r$, from 2 to 0 throughout repetitions. The random vector, r, has values between 0 and 1. A random value between -1 and 1 is utilised to switch between Eq.(1) and Eq.(2) while updating the positions of the whales; in Eq.(1) and Eq.(2), probabilities are 50% and 50%, respectively, meaning whales choose either path randomly and equally during the optimization process. Vector A has a random value of [-1, 1] when in the bubble-net phase, but during the search phase, the value of vector A can be more or less than 1. Eq.(3) shows the search mechanism in action.

$$X(t+1) = X_{rand}-A.|C.X_{rand}-X(t)| \quad (3)$$

For the WOA algorithm to do a worldwide search, this random search technique is used. First, random solutions are generated at the start of the WOA search. Then, the method is used to update these solutions, iteration by iteration. Until a predetermined maximum number of iterations have been achieved, the search will continue.

3.2 IMPROVED WHALE OPTIMIZATION ALGORITHM

An optimization algorithm needs a solid balance between exploration and exploitation for attaining good results. Therefore, step size reduces as iterations increase in WOA. This step size is controlled by the parameter *A*. However, it has been discovered that while WOA iterates, weak divergence causes the traps to be limited to the local optimum.

An enhanced whale optimization approach is used in this work to address these difficulties. To change the value of *A*, the levy fly function. This improves WOA exploration and exploitation capabilities at the same time.

The mathematical formula of the Levy distribution is as follows.

$$L(s,\gamma,\mu)=\begin{cases}\sqrt{\frac{\gamma}{2\pi}}\exp\left[-\frac{\gamma}{2(s-\mu)}\right]\frac{1}{(s-\mu)^{1.5}} & if\ 0<\mu<\infty\\ 0 & if\ s\le 0\end{cases} \quad (4)$$

Here μ, γ, and s denote the position parameter, the scale parameter managing the scale of distribution and the collection of samples, respectively.

Algorithm for Improved Whale Optimization

Step 1: **Start**
Step 2: import data
Step 3: Set the locations of the whale population ***X***
Step 4: calculate the fitness of every whale
Step 5: Set ***a*** and ***r***, Compute ***A*** and ***C***
Step 6: Set ***X**** as the best hunter whale location
Step 7: Set $t = 1$
Step 8: **while** $t \le$ max iterations **do**
Step 9: **for** each hunting whale, **do**
Step 10: **if** $p < 0.5$
Step 11: **if** $|\boldsymbol{A}| < 1$
Step 12: upgrade the present hunting whale location using Eq.(1)
Step 13: **else if** $|\boldsymbol{A}| \ge 1$
Step 14: randomly choose different search agent
Step 15: upgrade the current hunting whale location using Eq.(3)
Step 16: **end if**
Step 17: **else if** $p \ge 0.5$
Step 18: upgrade the present hunting whale location using Eq.(2)
Step 19: **end if**
Step 20: **end for**
Step 21: upgrade ***X**** if there is the best solution
Step 22: $t = t + 1$
Step 23: **end while**
Step 24: output **X***
Step 25: **End**

3.3 LUNG CANCER CLASSIFICATION USING MODIFIED LONG SHORT-TERM MEMORY

It necessary to categorise these genes to identify lung cancer patients. For categorization, this study makes use of Modified LSTM. With the LSTM architecture, temporal sequences and long-range dependencies can be described more precisely than traditional RNNs. An LSTM cell consists of an input gate, a forget gate and an output gate, as depicted in Fig.2. To activate the cell, these devices receive activation signals from various sources and use the built-in multipliers to regulate the signal [21]-[24].

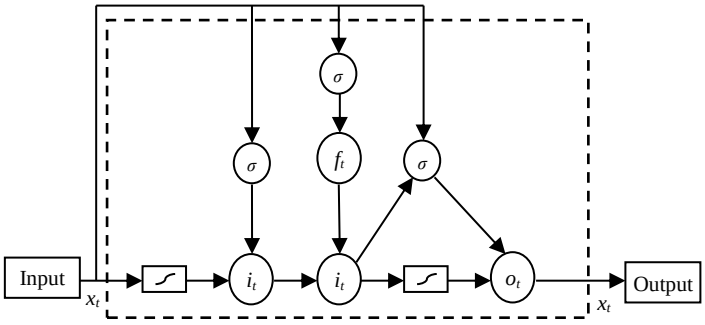

Fig.2. Long Short Term Memory Cell

The input gate of LSTM is described as:

$$i_t=\sigma(W_{xi}x_t+W_{hi}h_{t-1}+W_{ci}c_{t-1}+b_i \quad (5)$$

The forget gate is described as

$$f_t=\sigma(W_{xf}x_t+W_{hf}h_{t-1}+W_{cf}c_{t-1}+b_f \quad (6)$$

The cell gate is described as:

$$c_t=f_tc_{t-1}+i_c\tan h(W_{xc}x_t+W_{hc}h_{t-1})+b_c) \quad (7)$$

The output gate is defined as:

$$o_t=\sigma(W_{x0}x_t+W_{h0}h_{t-1}+W_{c0}c_t+b_0) \quad 8)$$

Finally, the hidden state is computed as:

$$h_t=o_t\tan h(c_t) \quad (9)$$

where

tan*h* - hyperbolic tangent activation function:

x_t - input at time t

W and b - network parameters (Weights and Biases)

3.4 MODIFIED LSTM

To keep the memory cells from being changed for an extended period, LSTM gates can be used. Networks with long-term memory, such as LSTMs, can store and transmit errors considerably longer than RNNs.

Because weighted regression is used in this study to prevent these problems, the suggested work employs Modified Long Short-Term Memory, which weights features based on how relevant observation is to an estimation point. In addition, three-dimensional distance metrics can be used to calculate the weights assigned to various features.

The form of the weight function determined is as follows:

$$\omega = \begin{cases} \left(1-\left(\frac{d}{h}\right)^3\right)^3 & if\ |d| \le h \\ 0 & if\ |d| > h \end{cases} \quad (10)$$

where,

ω-denotes the weight,

d-indicates the distance between the observation and the estimation points

h- half window width.

Here *i*, *f*, *o* and *c* are the input gate, forget gate, output gate, and cell state of the logistic sigmoid function σ. Weight matrices for peephole connections are W_{ci}, W_{cf}, and W_{co} [25, 26, 27]. The flow of information is controlled using three gates (*i*, *f*, *o*). The input gate determines the amount of input that can be received. This ratio affects the equation used to calculate cell state (Eq.(7)). The forget gate determines whether or not the prior memory ht-1 is accessed and whether or not it is necessary to keep the information in memory. For Eq.(6), the preceding memory ratio is calculated, then multiplied by the current memory ratio (Eq.(7)). Finally, passing or rejecting data from the memory cell is determined by this gate. This process is depicted in Eq.(9). The recurrent hidden layer is replaced by an LSTM cell in the LSTM-RNN architecture.

4. RESULTS AND DISCUSSION

This section compares the experimental findings of the proposed MLSTM and existing MIMAGA and SMO models in terms of precision, recall, accuracy, and f-measurement. Kent Ridge Bio-Medical Dataset uses MATLAB to implement the proposed MLSTM model. Tissues of 86 primary lung adenocarcinomas and ten non-neoplastic lung samples were analysed for gene expression data. There are 7129 genes in all samples. In addition, there is a 70% training set and 30% test set for the complete dataset. Both the training and testing datasets are used to verify the results.

Table.1. Performance Comparison Results

Metrics	Methods		
	MIMOGA	SMO	MLSTM
Accuracy	82.75	86.20	89.65
Precision	92.00	95.83	96.00
Specificity	33.33	66.66	66.66
F Measure	90.19	92.00	94.11

Precision: It is computed as,

$$\text{Precision} = TP/(TP+FP) \quad (11)$$

It is the evaluation of accuracy or quality, while recall is a calculation of completeness or quantity. Algorithms with high precision are generally more relevant than those with low precision. For example, precision may be calculated by dividing its number of true positives by its total number of false positives.

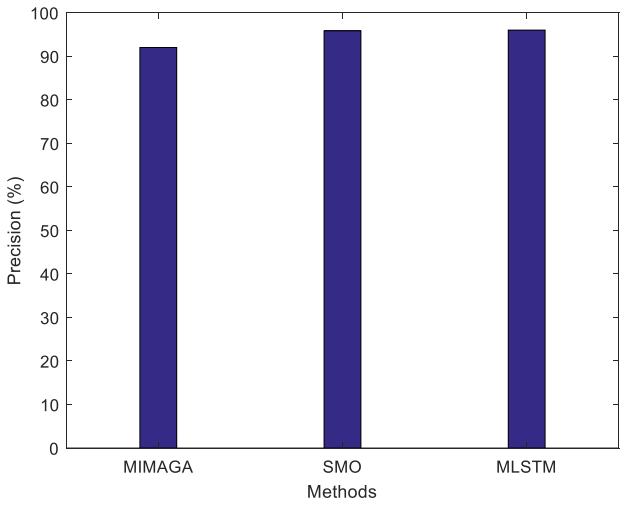

Fig.3. Precision

The Fig.3 compares the newly designed approach precision values and traditional techniques like MIMAGA and SMO. Again, processes are represented on the x-axis, and precision is indicated on the y-axis. The findings depict that the MLSTM algorithm delivers higher precision values of 96%, while existing MIMAGA and SMO techniques produce 92% and 95.83%, respectively.

Specificity: It is computed as below:

$$\text{Specificity} == TN/(TN+FP) \quad (12)$$

Specificity is defined as the proportion of actual negatives, which got predicted as the negative (or true negative).

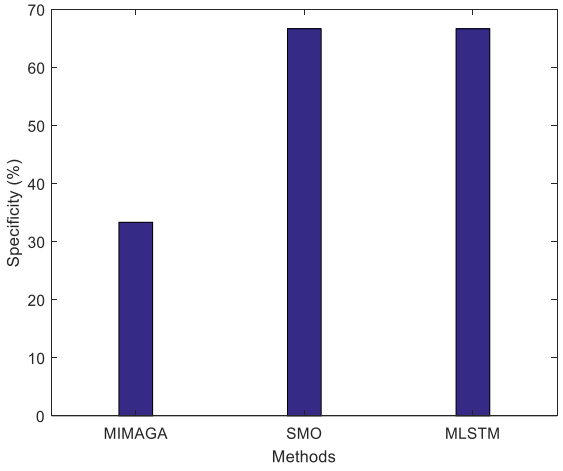

Fig.4. Specificity

Fig.4 compares the newly designed approach recall values and traditional techniques like MIMAGA and SMO. Again, processes are represented on the x-axis, and recall is indicated on the y-axis. The findings depict that the MLSTM algorithm delivers higher recall values of 66.66%, while existing MIMAGA and SMO techniques produce 33.33% and 66.66%, respectively.

F-measure: F1-Score is described as:

$$\text{F1-score}=(2\times precision\times recall)/(precision+recall) \quad (13)$$

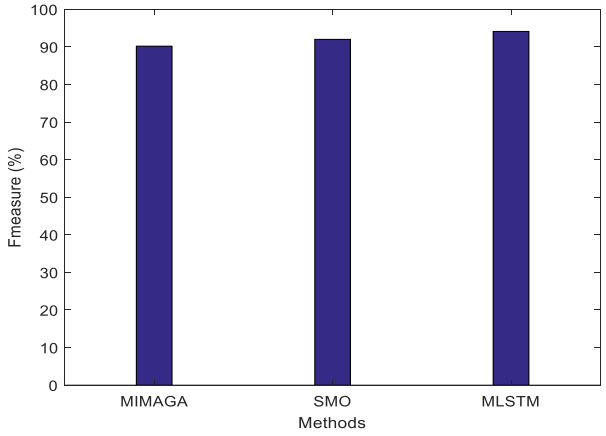

Fig.5. F-measure

The Fig.5 compares the newly designed approach F –measure values and traditional techniques like MIMAGA and SMO. Again, processes are represented on the x-axis, and F-measure is indicated on the y-axis. The findings depict that the MLSTM algorithm delivers higher F-measure values of 94.11%, while existing MIMAGA and SMO techniques produce 90.19% and 92%, respectively.

Accuracy: It is defined as the overall correctness of the method and is calculated as the total actual classification parameters (*TP+TN*) divided by the sum of the classification parameters (*TP+TN+FP+FN*).

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN) \quad (14)$$

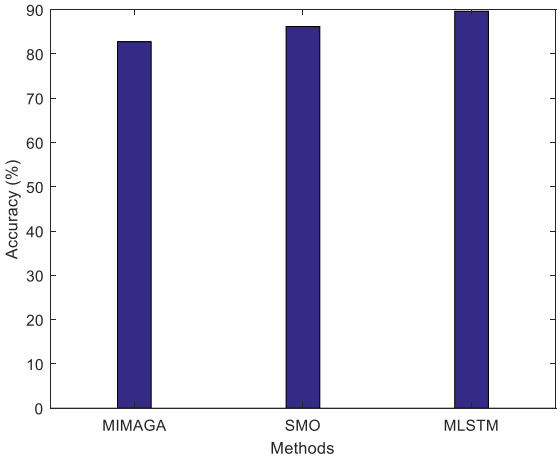

Fig.6. Accuracy

The Fig.6 compares the accuracy values of the newly designed approach and traditional approaches like MIMAGA and SMO. Again, processes are represented on the x-axis, and accuracy is indicated on the y-axis. The findings depict that the MLSTM algorithm delivers a higher accuracy result of 89.65 %, while existing MIMAGA and SMO techniques produce 82.75% and 86.20%, respectively.

5. CONCLUSION AND FUTURE WORK

Each day, the number of people dying from lung cancer around the world continues to rise. Lung cancer diagnosis and therapy can only be successful if a precise and reliable categorization is used. Genomic biomarkers for cancer diagnosis and prognosis can be discovered using a high-throughput platform such as a gene expression microarray. First, an effective gene selection model based on the IWOA is presented in this study. It speeds up the process while also providing valuable information from the gene expression data that aids in the more accurate classification of lung cancer. Second, for lung cancer categorization, an MLSTM Network is introduced. Genes are inputs into the system, classifying them based on whether they belong to a specific class, such as a cancer patient or a healthy individual. For precision, recall, accuracy, and f-measure, MLSTM outperform the previous MIMAGA and SMO techniques. However, LSTMs will have to be used in conjunction with other classification approaches in the future because they are prone to overfitting.

REFERENCES

[1] J.X. Liu, Y. Xu, C.H. Zheng and Z.H. Lai, "RPCA-Based Tumor Classification using Gene Expression Data", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 12, No. 4, pp. 964-970, 2014.

[2] M. Hosni, J.M. Carrillo-De-Gea, A. Idri and J.L. Fernandez-Aleman, "A Mapping Study of Ensemble Classification Methods in Lung Cancer Decision Support Systems", *Proceedings of International Conference on Medical and Biological Engineering and Computing*, pp.1-17, 2020.

[3] P. Tripathi, S. Tyagi and M. Nath, "A Comparative Analysis of Segmentation Techniques for Lung Cancer Detection", *Pattern Recognition and Image Analysis*, Vol. 29, No. 1, pp. 167-173, 2019.

[4] H. Azzawi, J. Hou and R. Alanni, "Multiclass Lung Cancer Diagnosis by Gene Expression Programming and Microarray Datasets, *Proceedings of International Conference on Advanced Data Mining and Applications*, pp. 541-553, 2017.

[5] S. Vanjimalar, D. Ramyachitra and P. Manikandan, "A Review on Feature Selection Techniques for Gene Expression Data", *Proceedings of IEEE International Conference on Computational Intelligence and Computing Research*, pp. 1-4, 2018.

[6] J.C. Ang, A. Mirzal, H. Haron and H.N.A. Hamed, "Supervised, Unsupervised, and Semi-Supervised Feature Selection: A Review on Gene Selection", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 13, No. 5, pp. 971-989, 2015.

[7] C.S. Seah, S. Kasim, M.F.M. Fudzee and M.A. Ismail, "An Enhanced Topologically Significant Directed Random Walk in Cancer Classification using Gene Expression Datasets", *Saudi Journal of Biological Sciences*, Vol. 24, No. 8, pp. 1828-1841, 2017.

[8] H. Salem, G. Attiya and N. El-Fishawy, "Classification of Human Cancer Diseases by Gene Expression Profiles", *Applied Soft Computing*, Vol. 50, pp. 124-134, 2017.

[9] H. Motieghader, A. Najafi, B. Sadeghi and A. Masoudi-Nejad, "A Hybrid Gene Selection Algorithm for Microarray Cancer Classification using Genetic Algorithm and Learning Automata", *Informatics in Medicine Unlocked*, Vol. 9, pp. 246-254, 2017.

[10] H. Lu, J. Chen, K. Yan and Z. Gao, "A Hybrid Feature Selection Algorithm for Gene Expression Data Classification", *Neurocomputing*, Vol. 256, pp. 56-62, 2017.
[11] A.H. Berger, A.N. Brooks and X. Wu, "High-Throughput Phenotyping of Lung Cancer Somatic Mutations", *Cancer Cell*, Vol. 30, No. 2, pp. 214-228, 2016.
[12] H.M. Alshamlan and G.H. Badr, "Genetic Bee Colony (GBC) Algorithm: A New Gene Selection Method for Microarray Cancer Classification", *Computational Biology and Chemistry*, Vol. 56, pp. 49-60, 2015.
[13] B. Ghaddar and J. Naoum-Sawaya, "High Dimensional Data Classification and Feature Selection using Support Vector Machines", *European Journal of Operational Research*, Vol. 265, No. 3, pp. 993-1004, 2018.
[14] I. Jain, V.K. Jain and R. Jain, "Correlation Feature Selection based Improved-Binary Particle Swarm Optimization for Gene Selection and Cancer Classification", *Applied Soft Computing*, Vol. 62, pp. 203-215, 2018.
[15] M.M. Ahmed, E.H. Houssein and E. Hassanien, "Maximizing Lifetime of Wireless Sensor Networks based on Whale Optimization Algorithm", *Proceedings of International Conference on Advanced Intelligent Systems and Informatics*, pp. 724-733, 2017.
[16] M. Sharawi, H.M. Zawbaa and E. Emary, "Feature Selection Approach based on Whale Optimization Algorithm", *Proceedings of International Conference on Advanced Computational Intelligence*, pp. 163-168, 2017.
[17] F.S. Gharehchopogh and H. Gholizadeh, "A Comprehensive Survey: Whale Optimization Algorithm and its Applications", *Swarm and Evolutionary Computation*, Vol. 48, pp. 1-24, 2019.
[18] Y. Ling, Y. Zhou and Q. Luo, "Levy Flight Trajectory-Based Whale Optimization Algorithm for Global Optimization", *IEEE Access*, Vol. 5, pp. 6168-6186, 2017.
[19] M.M. Mafarja and S. Mirjalili, "Hybrid Whale Optimization Algorithm with Simulated Annealing for Feature Selection", *Neurocomputing*, Vol. 260, pp. 302-312, 2017.
[20] A.G. Hussien, A.E. Hassanien and E.H. Houssein, "S-Shaped Binary Whale Optimization Algorithm for Feature Selection", *Proceedings of Recent Trends in Signal and Image Processing*, pp. 79-87, 2019.
[21] P. Rodriguez, G. Cucurull, T.B. Moeslund and F.X. Roca, "Deep Pain: Exploiting Long Short-Term Memory Networks for Facial Expression Classification", *IEEE Transactions on Cybernetics*, Vol. 23, No. 1, pp. 1-14, 2017.
[22] L. Yu, J. Chen, G. Ding and J. Sun, "Spectrum Prediction based on Taguchi Method in Deep Learning with Long Short-Term Memory", *IEEE Access*, Vol. 6, pp. 45923-45933, 2018.
[23] Y. Tian and L. Pan, "Predicting Short-Term Traffic Flow by Long Short-Term Memory Recurrent Neural Network", *Proceedings of IEEE International Conference on Smart City/SocialCom/SustainCom*, pp. 153-158, 2015.
[24] X. Yuan, L. Li and Y. Wang, "Nonlinear Dynamic Soft Sensor Modeling with Supervised Long Short-Term Memory Network", *IEEE Transactions on Industrial Informatics*, Vol. 16, No. 5, pp. 3168-3176, 2019.
[25] W. Lee, K. Kim, J. Park and Y. Kim, "Forecasting Solar Power using Long-Short Term Memory and Convolutional Neural Networks", *IEEE Access*, Vol. 6, pp. 73068-73080, 2018.
[26] J. Kim and H. Kim, "An Effective Intrusion Detection Classifier using Long Short-Term Memory with Gradient Descent Optimization", *Proceedings of International Conference on Platform Technology and Service*, pp. 1-6, 2017.
[27] X. Li and X. Wu, "Constructing Long Short-Term Memory based Deep Recurrent Neural Networks for Large Vocabulary Speech Recognition", *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4520-4524, 2015.