Section A-Research paper

Segmentation and Classification of Lung Nodule Histology using an Optimized Extreme Learning Machine Algorithm

Dr.V.Gowri[1], Dr.V.Vijaya Chamundeeswari[2]

[1]Assistant Professor, SRM Institute of Science and Technology, Ramapuram Campus, Chennai- 600089, India
[2]Professor & Dean, Saveetha Engineering College, Chennai- 602105, India

Abstract — The development of an automated and accurate disease identification model to examine the stage, type, and severity of lung malignancy from computed tomography (CT) images is the more perplexing and onerous task, even for skilled radiologists, owing to the various shapes, texture, and structures of the lung nodules. Though several machine and deep learning techniques are employed to improve the effectiveness of lung nodule identification models, the predictive performance of these models still necessitates substantial enhancement to meet the real-time demands of healthcare applications. This study proposes a novel column-wise optimized extreme learner (COEL) to differentiate the CT images efficiently. The proposed COEL exploits the concept of column-wise parameter tuning whose key idea is that the tuning of the multifaceted resultant matrix is divided into fine-tuning of the matrix with a single column vector. The difficult orthogonal classification task is converted into simple least squares regression with orthogonal restraints, which can extract useful knowledge from the attribute vector of the extreme learner to create an output vector, these make COEL more regression exploration and classification aptitude. Besides, this study proposes a weighted iteratively mean-separated equalizer to enhance the quality of the scan and a boosted superpixel grouping algorithm to segment the affected regions from the CT scan and COEL to classify the sample into benign (normal) and malicious (cancerous). The performance of the intended system is cautiously analyzed on real-life databases such as LUNA16 (Lung nodule analysis 2016) dataset. The experimental outcomes gained from the COEL are compared with other existing classification models with respect to evaluation measures. The results of a comprehensive empirical analysis of the LUNA16 database divulge that the intended COEL considerably outdoes other classifiers with improved accuracy, sensitivity, specificity, precision, and Jaccard similarity score (JSS) of 98%, 97.2%, 90.8%, 96.9%, and 96.5%, correspondingly. From comprehensive experimental results, we can conclude that the proposed COEL is the better diagnostic system as compared with advanced lung cancer classification models.

Keywords—data classification; decision forest; feature selection; healthcare data; heart disease prediction; penalizing attributes;

I. INTRODUCTION

Lung malignancy is a very common and serious fatal disease with an increased mortality rate across the globe [1]. It is more dangerous than breast, brain, bladder, stomach, prostate, and colorectal melanomas [2, 3]. Indeed, lung cancer resulted in around 1.8 million new patients and approximately 1.4 million deaths worldwide yearly [4, 5]. The visual appearance of lung nodules varies with subtle distinctiveness in form, texture, and dimension [6]. Particularly, lung tumors with a dimension equal to or larger than 3 mm are called nodules, and bigger ones seem to be malevolent [7]. The major etiology of lung cancer is smoking, but it is essential to consider other threats including chemicals and workplace risks, air pollution, exposure to radon gas, family history of lung cancer, etc. On the other hand, the lung malignancy death rate is decreased when cases can be identified and treated in the initial stages. Though modern oncology developments enable significant progress in cancer survival much work is still required in the field of diagnosis systems for classifying lung malignancy.

Automatic isolation and classification of lung nodules would unswervingly affect the development cycle of clinical practice in radiation oncology, one of the most commonly employed approaches of therapy for malignancy [8]. Of late, machine learning approaches enable medical organizations to procure, process, and analyze data at exceptional speeds without deteriorating the classification performance [9]. Machine learning-based lung tumor analytical models have made substantial breakthroughs in the domain of healthcare; therefore, several effective diagnostic systems have been developed. ELM is a predominant and constructive classifier used in various applications due to its fast convergence and better generalization characteristics.

Numerous studies indicate the extreme learning machine (ELM) is an advanced machine learning algorithm to train lung cancer diagnostic models and classify new images into benign and malignant based on significant features. ELM is characterized by the arbitrary selection of input parameters without optimization for the single hidden layer feed-forward neural networks (SLFN) [10]. More precisely, the ELM is a superior artificial neural network, whose internal parameters are produced arbitrarily and set, in order to obtain an exclusive least-squares solution of the output parameter, increasing the effectiveness of the system [11]. Such

enhancement significantly eliminates the complexity of parameter optimization due to back-propagation processes and therefore ensures the fast learning of extreme learning machines.

The characteristics of ELM includes a better ability to resolve issues related to convergence, generalization, local optimum, over-fitting, and variable tuning as compared to conventional deep learning models. Consequently, these inherent features make the ELM greater [12]. Furthermore, the training phase of the ELM is very easy to realize. The learning process in ELM is carried out by following steps: (i) some core variables of the hidden layer (e.g., internal parameters to link the visible (input) and hidden layers, the number of neurons in the hidden layer, and so on.) are engendered arbitrarily, which are constant during the entire procedure; (ii) the non-linear mapping function is used to relate the input sample to the attribute vector, and by studying the actual and anticipated outcome, the main parameter (the parameter relating the hidden and the output layer) can be unswervingly calculated, overlooking iterative optimization. Hence, its learning rate is significantly faster than that of the other deep learning models selection algorithm to select [13].

When performing complex image processing tasks, the superficial structure of ELM makes it impractical to have good enactment when unprocessed scan records are unswervingly fed as input. Therefore, several models developed to utilize efficient segmentation and feature selection algorithms to select important features before using an ELM classification algorithm. On the other hand, the utilization of too many/few hidden neurons may lead to overfitting/ underfitting problems in classification. To address these issues we introduce orthogonal limits on the transformation matrix of ELM. As the vital information holding capacity is potentially related to the classification ability of the algorithm, the integration of orthogonal constraint with ELM provides improved classification performance than conventional ELM.

When exploiting ELM to identify lung cancer, efficient preprocessing, segmentation, and feature selection algorithms need to be applied, which makes it practical to realize improved enactment. Suitable attributes can be designated using an appropriate attribute extraction algorithm using the radiologist's expertise, or other approaches. In segmentation, a boosted superpixel grouping algorithm is employed to segment the affected regions from the histopathology scans based on their features. In the classification process, this study proposes a novel COEL algorithm to differentiate the CT images efficiently. The proposed COEL exploits the concept of column-wise parameter fine-tuning whose key idea is that the tuning of the multifaceted resultant matrix is divided into fine-tuning of the matrix with a single column vector. The images are preprocessed to remove noise artifacts in clinical images.

This study uses a new incremental grey model (IGM) for removing noise artifacts from the clinical image. Besides, this work proposes an improved version of the histogram equalization method, called weighted iteratively mean separated equalizer which will improve the quality of the scan in consort with reduction of the image artifacts. This technique retains vital information and intensity considerably. Segmentation is the most important process in designing any image-based diagnostic system to isolate the affected area from better-quality scans. In this work, we use a super-pixel grouping method, known as the boosted superpixel grouping (BSG) algorithm, to segment the lung nodule from the CT scan effectively. The performance of the intended system is cautiously analyzed on real-life databases such as LUNA16 (Lung nodule analysis 2016) dataset. The experimental outcomes gained from the COEL are compared with other existing classification models with respect to evaluation measures. The results of a comprehensive empirical analysis of the LUNA16 database divulge that the intended COEL considerably outdoes other classifiers with improved classification performance. From extensive experimental results, it is concluded that the proposed COEL is the better diagnostic system as compared with advanced lung cancer classification models.

The remaining sections of this paper are structured as follows. Section 2 discusses relevant studies and various state-of-the-art methods employed for lung cancer diagnostic systems. In Section 3, we describe the theoretical context of the ELM classifier. Section 4 presents the proposed column-wise optimized extreme learner in detail. Section 5 discusses the implementation details of the proposed classifier for lung cancer diagnosis using the LUNA16 dataset. We evaluate the performance of the intended diagnostic model by relating its enactment with other germane classifiers in terms of performance measures in Sections 6 and 7. In Section 8, we conclude the manuscript with intended future work.

II. LITERATURE SURVEY

Early identification of tumors from clinical images is a cutting-edge approach to recognizing and classifying lung nodules. Numerous works have exploited machine learning and deep learning algorithms for lung malignancy identification and to increase the performance of the classifier and its speed. In all these studies, experts have endeavored to increase classification performance through different phases of the diagnostic system including preprocessing, attribute selection, isolation, and classification. Xu et al. developed a CNN model-based classification algorithm for lung nodule identification [14]. In this method, the affected area is first isolated from the given CT images using a convolutional neural network (CNN) model that contains a convolutional unit, one max pooling unit, and two completely linked units. Then, the k-means clustering algorithm is used to categorize the data input.

Liu et al. developed an effective classification approach through the superpixel technique with an integrated system using the CNN model and random forest algorithm. This approach contains various stages including preprocessing, isolation, and

optimization [15]. Initially, a quality-enhanced lung image is isolated using superpixel and then a classifier is used to identify lung malignancy. Xie et al. introduced a classification algorithm using deep CNN (DCNN) to recognize precise positions of a lung nodule in medical images quickly [16]. This approach uses a swift area-based CNN to integrate 3D lung information, integrate the results of three different slices, and vote out the outcome. This work exploits a performance-enhancing method of 2D CNN to reduce the rate of false alarms.

Nakrani et al. developed a lung malignancy recognition model by applying ResNet to CT images. This approach contains the preprocessing and the classifier [17]. ResNet uses morphological features for isolating lung tumors and CNN for classifying lung cancers. Zhang et al. proposed a revised form of ResNet and employed it to isolate the region of interest from CT images with non-small cell nodules [18]. Jiang et al. developed a multiple-resolution residually linked model to integrate attributes at various image resolutions and realized improved classification outcomes in different databases [19].

Wang et al. developed a DCNN, called VGG16-T with enhancing approach to classify lung cancers in their early stage from CT images [20]. This work uses an ensemble of several VGG16-T techniques and trains them with a boosting method. This model achieves improved performance in identifying lung cancer using a joint voting technique to integrate the results. Ali et al. developed and analyzed a reinforcement learning approach through DCNN for the early identification of lung cancer [21]. Kamal et al. developed Recurrent 3D-DenseUNet, which includes a 3D encoder to select significant attributes, a recurrent convolutional unit to gain spatio-temporal data, and a 3D decoder unit to create isolation masks [22]. Several researchers have achieved their goal effectively. However, their effectiveness in terms of accuracy, sensitivity, specificity, and precision is frequently not the best. Hence, this work develops an ELM-based classifier, called COEL for classifying lung cancer with improved performance. The proposed model exploits appropriate preprocessing, image segmentation, feature selection, and COEL classifier. This study uses an IGM for removing noise artifacts from the clinical image and an improved version of the histogram equalization method, called WIMSE which will improve the quality of the scan. Besides, we use a BSG algorithm to segment the affected region from the CT scan. Finally, COEL is applied to classify the CT scan samples into benign and malignant.

III. THEORETICAL BACKGROUND OF ELM

ELM is an advanced learning technique for the single hidden layer feed-forward neural networks (SLFNs), introduced by Huang et al. [10], that is pigeonholed by the local arbitrary variables caused without optimization. More precisely, the ELM is a superior artificial neural network approach, whose input parameters are produced arbitrarily and defined, in order to obtain an exclusive least-squares solution of the output parameter, increasing the effectiveness of the system [23]. Such enhancement significantly reduces the complexity of parameter optimization due to the back-propagation process and therefore ensures fast ELM learning. The characteristics of ELM include a better ability to resolve issues related to generalization, convergence, local optimum, over-fitting, and variable tuning as compared to conventional deep learning models. Consequently, these inherent features make the ELM a superior approach for image classification in the healthcare industry [12]. Furthermore, the training phase of the ELM is very easy to realize.

The learning process in the ELM model is carried out in three steps. Initially, some of the hidden layer parameters are engendered arbitrarily, which are preserved during the entire classification task. Then, a non-linear mapping function is used to relate the input sample to the attribute space, and by examining the actual outcome and anticipated outcome, the significant factor can be calculated directly, overlooking recurrent optimization. Thus, its learning performance is significantly improved than that of the other deep learning model [24].

For S number of discrete instances, $(i_k, o_k)_{k=1}^{S}$ where $i_k \in \mathbb{R}^n$ is the input space, $o_k \in \mathbb{R}^m$ is the anticipated output vector, and the anticipated outcome of the k^{th} instance is o_k, $o_k = [o_{k1}, o_{k2}, \ldots . o_{km}]^T$, $k = 1,2, \ldots S$. For a particular activation function, if the actual outcome of the SLFNs is identical to the anticipated outcome o_k, we define the SLFN using Equation (1).

$$f(i_j) = \sum_{k=1}^{H} \gamma_k \, g(\varphi_{k,} i_{j,} \alpha_k) = o_j \quad (1)$$

$$\sum_{k=1}^{H} \gamma_k \, g(\varphi_{k,} i_{j,} \alpha_k) = \sum_{k=1}^{H} \gamma_k \, g(\varphi_k i_j + \alpha_k) = o_j \quad (2)$$

where $\varphi_k = [\varphi_{k1}, \varphi_{k2}, \ldots . \varphi_{kn}]^T$ is the input parameter mapping the visible layer to the k^{th} neuron of the hidden layer, α_k is the predefined value of the k^{th} hidden node, $\gamma_k = [\gamma_{k1}, \gamma_{k2}, \ldots . \gamma_{km}]^T$ is the output parameter relating the k^{th} neuron in the hidden layer to the output layer, and H is the number of the hidden nodes as given in Figure 1.

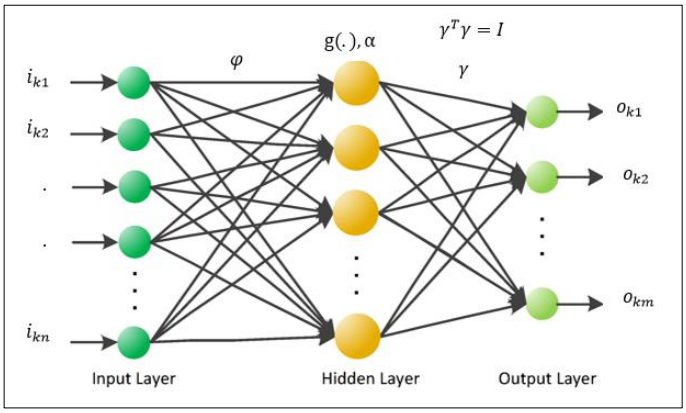

Figure 1: General architecture of ELM model.

Equation (2) can be revised as given in Equation (3).

$$A\gamma = O \qquad (3)$$

where

$$A(\varphi_{k,} i_{j,} \alpha_k) = \begin{bmatrix} g(\varphi_1 . i_1 + \alpha_1) & : & g(\varphi_{\dot{H}} . i_1 + \alpha_H) \\ : & : & : \\ g(\varphi_1 . i_S + \alpha_1) & : & g(\varphi_{\dot{H}} . i_S + \alpha_{\dot{H}}) \end{bmatrix}_{S\times H}$$

$$\gamma = \begin{bmatrix} \gamma_1^T \\ : \\ \gamma_{\dot{S}}^T \end{bmatrix}_{H\times m} \text{ and } O = \begin{bmatrix} o_1^T \\ : \\ o_{\dot{S}}^T \end{bmatrix}_{S\times m} \qquad (4)$$

here A denotes the output matrix (transformation matrix) of hidden nodes. For instance, the k^{th} column of A represents the output of the k^{th} hidden nodes for input data points. Therefore, according to the concept of ELM, the optimum value of Equation (3) is defined by Equation (5).

$$\gamma = A^{\dagger}O \qquad (5)$$

where $A^{\dagger}$ is the Moore–Penrose inverse of the Matrix H, $A^{\dagger} = (A^T A)^{-1}A^T$. Indeed, the ELM model targets to achieve the minimum learning error of the output weight matrix. Generally, the controlling factor is added to ELM to improve the model precision. Hence, the optimum parameter is obtained using Equation (6).

$$\min_{\gamma} \left(\frac{1}{2}\|A\gamma - O\|^2 + \frac{R}{2}\|\gamma\|^2\right) \qquad (6)$$

where $R > 0$ is the controlling factor, which is employed to balance the experimental risk and physical risk. According to the Karush–Kuhn–Tucker constraint, the optimum value of γ is calculated as given in Equation (7).

$$\gamma = (A^T A + RI)^{-1}A^T O \qquad (7)$$

IV. PROPOSED COLUMN-WISE OPTIMIZED EXTREME LEARNER

In order to increase the locality-maintaining ability of hidden layer representation (attribute space) to the output layer (label space), this research introduces orthogonal limits on the transformation matrix. As the classification ability of the algorithm is based on its locality-preserving capacity, the integration of orthogonal constraint with ELM provides improved classification performance

than conventional ELM. By adding the orthogonal limit, the objective function of optimized ELM is formulated as given in Equation (8).

$$\min_{\gamma^T\gamma = \mathrm{I}} (\|\mathrm{A}\gamma - \mathrm{O}\|^2) \qquad (8)$$

where$A \in \mathbb{R}^{S\times H}$, $\gamma \in \mathbb{R}^{H\times c}$ and $\mathrm{O} \in \mathbb{R}^{S\times c}$. Here c is the number of class labels. By applying the orthogonal limit, we can map the data to an orthogonal vector in which the structure of the data metric can be well-preserved. As $H > c$, the objective function given in Equation (8) is a matrix approximation problem which is challenging to find out the result unswervingly owing to the orthogonal limit. In this study, we develop a recurrent approach to fine-tune Equation (8) using the subsequent theorem.

Theorem [25]: The optimum value of the matrix approximation problem defined by Equation (8) is equal to the optimum solution to the problem given in Equation (9).

$$\min_{\hat{\gamma}^T\hat{\gamma} = \mathrm{I},\, \mathrm{O}_1} (\|\mathrm{A}\hat{\gamma} - [\mathrm{O}, \mathrm{O}_1]\|^2) \qquad (9)$$

where $\hat{\gamma} = [\gamma, \gamma_1] \in \mathbb{R}^{H\times H}$, $\gamma_1 \in \mathbb{R}^{H\times(H-c)}$ and $\mathrm{O}_1 \in \mathbb{R}^{S\times(H-c)}$. From the objective function given in Equation (9), we know that the optimum value of O_1 is $\mathrm{A}\gamma_1$. Substituting the value of O_1in Equation (9), we can get the optimum value as defined by Equation (10).

$$\min_{\hat{\gamma}^T\hat{\gamma} = \mathrm{I}} (\|\mathrm{A}\hat{\gamma} - [\mathrm{O}, \mathrm{A}\gamma_1]\|^2) \qquad (10)$$

which is equal to Equation (8). Based on this theorem, we can resolve the objective function given in Equation (9) rather than directly solving the matrix approximation problem given in Equation (8). But, simplifying Equation (9) cooperatively regarding $\hat{\gamma}$ and O_1 is challenging and therefore we interchangeably apprise one parameter while setting the other. After setting O_1, $\mathrm{A} \in \mathbb{R}^{S\times H}$, and $[\mathrm{O}, \mathrm{O}_1] \in \mathbb{R}^{S\times H}$, Equation (8) will be simplified as given in Equation (11).

$$\min_{\hat{\gamma}^T\hat{\gamma} = \mathrm{I}} (\|\mathrm{A}\hat{\gamma} - [\mathrm{O}, \mathrm{O}_1]\|^2) \qquad (11)$$

In this work, the above equation is resolved by the singular value decomposition (SVD) approach. The SVD of $\mathrm{A}^T[\mathrm{O}, \mathrm{O}_1]$ is defined by Equation (12).

$$\mathrm{A}^T[\mathrm{O}, \mathrm{O}_1] = \mathcal{U}\mathcal{E}\mathcal{V}^T \quad (12)$$

where $\mathcal{U} \in \mathbb{R}^{H\times H}$, $\mathcal{E} \in \mathbb{R}^{H\times H}$, $\mathcal{V} \in \mathbb{R}^{H\times H}$ and the solution to the problem defined by Equation (11) is shown in Equation (13).

$$\hat{\gamma} = \mathcal{U}\mathcal{V}^T \qquad (13)$$

After calculating $\hat{\gamma}$, the output weight matrix γ can be constructed by the first c columns of matrix $\hat{\gamma}$. After setting $\hat{\gamma}$, the problem defined by Equation (9) can be simplified to the following problem.

$$\min_{\mathrm{O}_1} (\|\mathrm{A}[\gamma, \gamma_1] - [\mathrm{O}, \mathrm{O}_1]\|^2) \qquad (14)$$

Clearly, the solution to Equation (14) is

$$\mathrm{O}_1 = \mathrm{A}\gamma_1 \qquad (15)$$

where γ_1 is the last $H - \mathrm{c}$ columns of matrix W.

V. IMPLEMENTATION OF COEL IN LUNG CANCER CLASSIFICATION

The proposed diagnostic model can efficiently recognize lung malignancy using appropriate preprocessing, image segmentation, feature selection, and COEL classifier. The CT images are preprocessed to remove noise artifacts. This study uses IGM for removing noise artifacts from the CT scans image. Besides, this work proposes WIMSE to improve the quality of scan in consort with the reduction of visual artifacts. This technique retains vital information and the quality of the image considerably. Segmentation is the most important process in developing any image-based diagnostic system to isolate the affected area from better-quality scans. In this work, we use the BSG algorithm to segment the affected region from the input scans. Then, we use the isolated candidates to train a classification algorithm. This study uses principal component analysis (PCA) for minimizing the dimension of the attribute space by selecting vital attributes to recognize lung malignancy. Then, the proposed COEL is used to classify the data samples into normal or cancerous. The all-inclusive depiction of each process is given in the following subsections.

5.1 Data preprocessing

The digital images obtained from clinical laboratories are often tainted by artifacts in the cohort and communication techniques, which considerably disturbs the performance of the segmentation and classification processes. Recently, some artifact removal techniques have been proposed to confiscate noises from medical images. This study proposes an effective artifact removal technique, known as the incremental grey model for removing noise from the CT scans. Noise artifacts in a clinical scan are often defined a discrete data points or a group of pixels to provide a better-quality visual effect. The proposed IGM can effectively eliminate noise in input images. Consider a positive sequence $(S^{(0)})$ and the corresponding k-order incremental operator sequence $(S^{(k)})$ given in Equations (16) and (17), respectively.

$$S^{(0)} = \left\{s^{(0)}(1), s^{(0)}(2), \ldots . s^{(0)}(n)\right\}, n > 4 \quad (16)$$

$$S^{(k)} = \left\{s^{(k)}(1), s^{(k)}(2), \ldots . s^{(k)}(n)\right\} \quad (17)$$

where $s^{(k)}(l)$ is calculated at any point (m) using Equation (18).

$$s^{(k)}(m) = \sum_{i=1}^{m} \binom{k-1+m-i}{m-i} s^{(k)}(i) \quad (18)$$

For two points S_a and S_b, we calculate their 1D k-order incremental operator sequences $(S_a^{(k)})$ and $(S_b^{(k)})$ using Equations (19) and (20), correspondingly.

$$S_a^{(k)} = \left\{s_a^{(k)}(1), s_a^{(k)}(2), \ldots . s_a^{(k)}(n)\right\} \quad (19)$$

$$S_b^{(k)} = \left\{s_b^{(k)}(1), s_b^{(k)}(2), \ldots . s_b^{(k)}(n)\right\} \quad (20)$$

where $k = \frac{3}{4}$. The similarity measure (χ_{ab}) between S_a and S_b is calculated using Equation (21).

$$\chi_{ab} = \frac{1 + |\Gamma_a| + |\Gamma_b|}{1 + |\Gamma_a| + |\Gamma_b| + |\Gamma_a - \Gamma_b|} \quad (21)$$

where Γ_a and Γ_b are calculated by Equations (22) and (23), respectively.

$$\Gamma_a = \left| \sum_{l=2}^{n-1} s_a^{(0)}(m) + \frac{1}{2} s_a^{(0)}(n) \right| \quad (22)$$

$$\Gamma_b = \left| \sum_{l=2}^{n-1} s_b^{(0)}(m) + \frac{1}{2} s_b^{(0)}(n) \right| \quad (23)$$

The artifacts removal value of a pixel *j* in the CT scan *H*is calculated using Equation (24).

$$H(j) = \frac{\sum_{p \in Y} H(p) \,.\, \delta_{ip} + H(j)}{\sum_{p \in R} \delta_{ip} + 1} \qquad (24)$$

where Y is the m-adjacent pixels p. In this work, we select m=8. Figure 2 displays some samples of enhanced pictures gained from our IGM method after removing artifacts. From the pictures shown in Figure 2(b), it is observed that the IGM can provide a better-quality inputs for further processing phases.

(a)

(b)

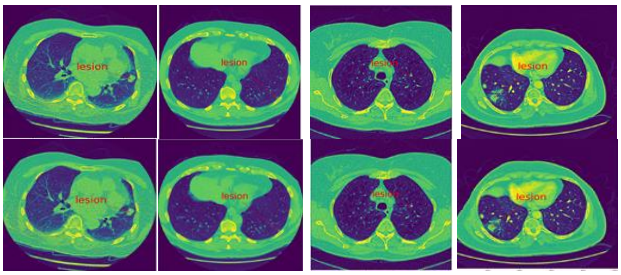

Figure 2: Artifacts removal using IGM (a) Original images (b) Scans after removing noise artifacts

The quality of CT images, as in several medical image processing, is defined with respect to noise, artifacts, and contrast of the image. Generally, the quality of CT scans is based on both scanner design factors and clinical protocol factors. Poor scanner causes low-quality images with deprived perceptibility and distortion. Histogram equalization is the traditional technique for improving the quality of a scan, but it generates undesirable visual relics, which provide over-contrasted images, loss of important information, and lack of natural appearance of the scan. Iteratively mean separated equalizer (IMSE) is one of the variations of histogram equalization to solve the problems related to the conventional histogram equalizer. This technique will decrease the image noise and improve the quality of the scan but still, there is a lack of vital information. In this context, this work proposes an improved version of IMSE, called weighted iteratively mean separated equalizer (WIMSE) which will improve the quality of the scan by decreasing artifacts. This technique retains vital information and quality considerably.

In this WIMSE technique, the histogram (refer Figure 3) of a scan is alienated into four sub-histograms according to the value of weighted mean and then uses a cutting boundary to every sub-histogram followed by equalization of every sub-histogram. Before increasing the image contrast, every input sample is recorded against the predefined value and the existing din in the scan is substituted by the mean value. This procedure is iterative till eliminates the whole din available in the CT scan. Then the quality of the image is enhanced using the weighted mean histogram equalizer.

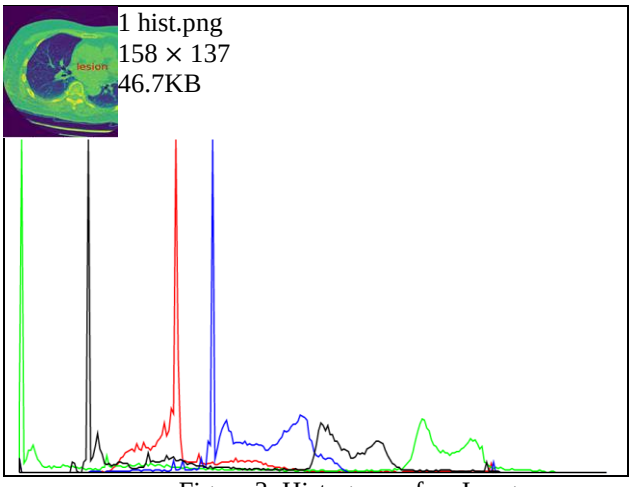

Figure 3: Histogram of an Image

Consider the scan H which comprises l individual gray level pixels defined by $h_0, h_1, \ldots \ldots h_{l-1}$. Once defining these values, the contrast measure and combined distribution of each picture element are computed. The brightness of each pixel is correlated to a specific range of data points. Based on the value of brightness, the pixel quality is measured, and the probability density function $f_d(h_0)$ is calculated by Equation (25).

$$f_d(h_0) = \frac{h^x}{h} \qquad (25)$$

where $x = 0,1,2 \ldots\ldots (l-1)$; h^x signifies the value of the pixel that is related to the value of h. The term h represents the total pixels of H. The cumulative density function ($C_d(h_0)$) of gray level 'l' is defined by Equation (26).

$$C_d(h_0) = \sum_{l=h_0}^{h_{l-1}} f_d(h_0)$$

By applying $C_d(h_0)$, we calculate the transfer function which relates the input picture to the output picture for the whole range of gray levels. The transfer function $f_t(h_0)$ is defined by Equation (27).

$$f_t(h_0) = (h_{l-1}) \times C_d(h_0) + lower\ gray\ level \qquad (27)$$

After computing the probability density function of pixels, the weighted mean value (W_{mean}) of each pixel is calculated using Equation (28).

$$W_{mean} = \frac{\sum_{g=u}^{v} g \times f_c(g)}{\sum_{g=u}^{v} f_c(g)} \qquad (28)$$

where $f_a(g)$ is the combined distribution function and it is defined by Equation (29).

$$f_a(g) = \sum_{i=1}^{p} f_d(g_i) \qquad (29)$$

Based on the efficient study using weighted mean value and pixel-wise histogram the stable pixel is selected. An effective investigation is performed on each pixel and the procedure of sub-image-based histogram estimation improves the quality and also eliminates the artifacts from the CT image. After increasing the quality of the scan, the nodule is isolated by applying a boosted superpixel grouping algorithm.

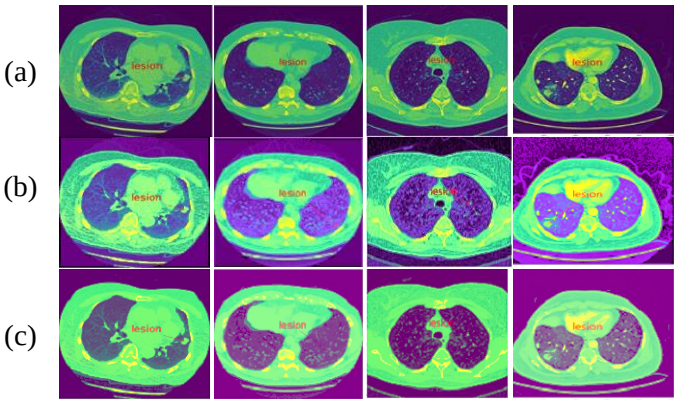

Figure 4: Enhanced images using IMSE (a) Denoised scans (b) Scans after applying histogram equalization technique (c) Scans after applying WIMSE technique

5.2 Image segmentation through boosted superpixel grouping algorithm

Segmentation is a vital process in designing a lung malignancy model to isolate the affected area from a quality-enhanced CT scan. This work uses a novel superpixel segmentation method, known as the boosted superpixel grouping approach, to segment the nodule from the CT scan. Superpixel is a group of connected picture elements which have related features (e.g., similar pixel

intensity). The key benefit of superpixel is to develop a more useful and perceptually important representation of the input scan. Currently, superpixel-based segmentation approaches are used as standard techniques to study the input images according to the resemblance of the pixels and divide the image into various sub-images for identifying lung cancers. This approach gathers a cluster of similar pixels to create visually important objects while deeply reducing the number of primitives for further processes. After grouping, the number of data points essential for processing should be considerably decreased.

The proposed BSG algorithm operates on two key steps: (i) analyzing each data point available in the image; and (ii) grouping the connected pixels into a group by differentiating the anomalous data points. In the segmentation process, each pixel is studied continually for measuring the resemblance of pixels through eigenvalues. The similarity metric includes the image quantitative measure that is used to form the cluster. The enhanced pixels are examined and each picture element is considered a data point in the feature space. An undirected graph is constructed for segmenting the lung tumor. The graph associated with the lung is represented as $G = \langle N, E \rangle$ where N is the node (i.e., input sample) and E is the relation among the nodes. Every connection has a fixed weight $w_{ab}(a, b) \in E$. Here, E is the resemblance function between nodes a and b. After defining the undirected graph, the segmentation is carried out where the graph excruciating problem is resolved by piercing the node into a high and low resemblance group. Consider P and Q are two subdivisions of nodes and the related normalized cut of the image is calculated by Equation (30).

$$w(P, Q) = \sum_{a \in P, b \in Q} w_{ab} \qquad (30)$$

Once computing the parameters of the nodes, the n-cut process should be performed as defined by Equation (31).

$$ncut(P, Q) = \frac{w(P, Q)}{w(P, N)} + \frac{w(P, Q)}{w(Q, N)} \qquad (31)$$

According to the $ncut$ process, the resemblance between the samples or pixels is examined in different subdivisions of the image. Then, the total similarity among the samples is measured using Equation (32).

$$nasso(P, Q) = \frac{w(P, P)}{w(P, N)} + \frac{w(Q, Q)}{w(Q, N)} \qquad (32)$$

In similarity estimation, the NP-hard problem is resolved by defining the value of the minimum cut using Equation (33).

$$d(a) = \sum_{b} w_{ab} \qquad (33)$$

D$(n \times n)$ is the diagonal matrix with size d. Then symmetric matrix is defined by W $(n \times n)$. Then, the minimum cut is measured using Equation (34).

$$\min ncut(P, Q) = \min_z \frac{z^T (D - W) z}{z^T D y} \qquad (34)$$

where, $z_i \in \{1, -y\}$; y is constant, and $z^T D = 0$. Therefore, the resemblance between the samples is calculated by reducing the value of the $ncut$. Then, the grouping process is performed based on the number of groups. The grouping problem is resolved by the k-means approach which works on an enormous volume of samples in the feature vector and maps the samples successfully. Once resolving the grouping problem, germane pixels are gathered into one group and irrelevant pixels are grouped into another to support the segmentation of the region of interest. Besides, the BSG algorithm runs based on the undirected graph which relates each sample. This makes an effective classification of the lung nodule.

5.3 Feature Engineering

In lung nodule diagnostic systems, feature engineering performs a vital role in deciding classification accuracy. It selects significant attributes about the position of the lung from CT images by forming an attribute space. Then, this space is employed for the nodule classification. This work only considers significant attributes of nodules such as forms, edges, and their brightness. The attribute associated with the form of cancer delivers information about the shape and size of tumors. The attribute of boundaries

offers useful information about the edge of the identified tumors. The features related to intensity define the brightness variation of each pixel in the scan. The attribute selection (reduction) is employed to reduce the dimension of the attribute vector without ignoring the fundamental disease information. The feature reduction algorithm targets to designate the vital features from the original feature space. The present work exploits principal component analysis (PCA) for reducing the dimension of the attribute vector. The major concept of PCA is to reduce the size of the attribute vector comprising irrelevant data and it is identified by transforming various factors known as the principal component. In PCA, the attributes of the image are transformed into a lower level of feature space by means of the principal eigenvectors of the relationship matrix.

3.1 Image classification

To classify the segmented lung nodules into normal or cancerous, this study exploits the COEL model. The proposed COEL is a completely trained ELM and it can learn attributes on the database. It comprises one hidden and one output unit. The hidden nodes learn the attributes and the output layer learns the mapping function. This algorithm classifies the dataset in an improved manner through a simple network structure. It is one of the balanced classification algorithms for particular class tags with reduced computational complexity. COEL is used to improve the predictive performance of the lung nodule recognition system. COEL increases the reliability of the classification process. A reduced number of input-output data points are used for training the diagnostic system. In order to preserve the vital information in hidden layer representation, this study introduces orthogonal limits on the transformation matrix. As the classification ability of the algorithm is based on its locality-preserving capacity, the integration of orthogonal constraint with ELM provides improved classification performance than conventional extreme learning algorithms.

This work uses the standard classification error (SCE) to calculate the prediction uncertainty of the model. Each input sample is verified by Equation (35).

$$SCE_i = \sqrt{(1 + L_i) \times MSE\_C_{BC}} \qquad (35)$$

In the above equation, $L_i = X_i^T (X^T X)^{-1} X_i$ represents the regularization factor for input image i. The bias-corrected mean square error of testing (MSE_C_{BC}) is calculated using Equation (36).

$$MSE\_C_{BC} = \frac{1}{1 - J - 2} \sum_{i=1}^{I} (\hat{o}_i - o_i - bias_c)^2 \quad (36)$$

where $bias_c = \frac{1}{I_C} \sum_{i=1}^{I_C} (\hat{o}_i - o_i)^2$, $C = 0,1$. Then, we calculate the value of $p(\hat{y}|w_c) = \frac{1}{I_C} \sum_{i=1}^{I_C} g_i(\hat{o}), c = 0,1$. The term $g_i(\hat{o})$ denotes $f_d(h_0)$ of every testing image i for tags w_0 and w_1 with the form of the Gaussian curve centered at $\hat{o}_i$ and standard deviation of SCE_i. It is measured using Equation (37).

$$g_i(o) = \frac{1}{SCE_i \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\hat{o} - \hat{o}_i}{SCE_i}\right)^2} \quad (37)$$

Algorithm 1: COEL for lung nodule classification

Input: Learning data samples $N = \{((i_k, t_k|)x_k \in R^m, k = 1, \ldots, N)\}$, g(x), and the number of hidden nodes $(\widehat{N})$.

1: Initialize the values of bias terms α_i and input weights w_i randomly using $f_d(h_0)$.
2: Calculate the matrix A of the hidden node.
3: Calculate the weight vector γ using Equation (5).
4: Compute $f_d(h_0)$ for every output label. The COEL classification for the given data is defined by $\hat{O} = A\gamma$ and computes the SCE.
5: The learning sample of class (c) is averaged to calculate $f_d(h_0)$ of class $w_c, c = 0,1$. Then, calculate $p(\hat{o}|w_c)$ and $g_i(\hat{o})$
6: Compute $p(w_c|\hat{o}_u)$
7: Categorize new input image by Equation (40).

The Bayes formulation for the probability with the classification $\hat{o}_u$ fits into the tag w_c is defined by Equation (38).

$$p(w_c|\hat{o}_u) = \frac{p(\hat{o}_u|w_c) \times p(w_c)}{p(\hat{o}_u|w_0) \times p(w_0) + p(\hat{o}_u|w_1) \times p(w_1)} \quad (38)$$

The prior probability $p(w_c)$ is converted into posterior probability $p(w_c|\hat{o}_u)$ by means of the prediction value of $\hat{o}_u$. The expectation range lies between the values as shown in Equation (39).

$$R = \{\hat{o}_{u,1} \leq \hat{o}_u \leq \hat{o}_{u,r}\} \quad (39)$$

where the maximum and minimum bounds are fixed using the expressions $\hat{o}_{u,1} = o_u - k.SCE_u$ and $\hat{o}_{u,r} = o_u + k.SCE_u, k = 1$ and 2. A new sample is categorized using Equation (40).

$$p\{w_c|\hat{o}_{u,1} \leq \hat{o}_u \leq \hat{o}_{u,r}\} = \frac{\int_{\hat{o}_{u,1}}^{\hat{o}_{u,r}} p(\hat{o}_u|w_c)}{p(\hat{o}_u)d\hat{o}_u} p(w_c) \quad (40)$$

In the above equation, the numerator defines the area under the curve $p(\hat{o}_u|w_c) \times p(w_c)$ in the interval $\{\hat{o}_{u,1} \leq \hat{o}_u \leq \hat{o}_{u,r}\}$ as defined by Equation (41).

$$Area_{u,c} = p(w_c) \int_{\hat{o}_{u,1}}^{\hat{o}_{u,r}} p(\hat{o}_u|w_c)\, d\hat{o}_u, c = 0,1 \quad (41)$$

Algorithm 1 illustrates the pseudocode for intended COEL in lung nodule classification.

VI. PERFORMANCE EVALUATION

To assess the effectiveness of the proposed system, we carry out a comprehensive experimental analysis on NVIDIA TESLA K80 GPU with 12 GB graphics memory in PyTorch. The performance of the proposed classifier is evaluated by comparing the empirical results with five related state-of-the-art classifiers, namely AlexNet [26], ResNet [17], TumorNet [27], DFCNet [28], and CMixNet [29].

4.1 Dataset Preparation

To assess the performance of any machine learning algorithm in the healthcare sector, we need a large dataset that enables a better result. In this work, we use a corpus of labeled LUNA16 (Lung nodule analysis) dataset [30]. It contains 1,186 CT images. The images with missing slices or unpredictable slice spacing are also omitted. This resulted in 888 images, with a total of 36,378 interpretations created by several radiologists. In this database, only the interpretations characterized as nodules greater than 3 mm are considered related, as the other interpretations (nodule size less than 3 mm and non-nodules) are not considered pertinent for lung nodule classification.

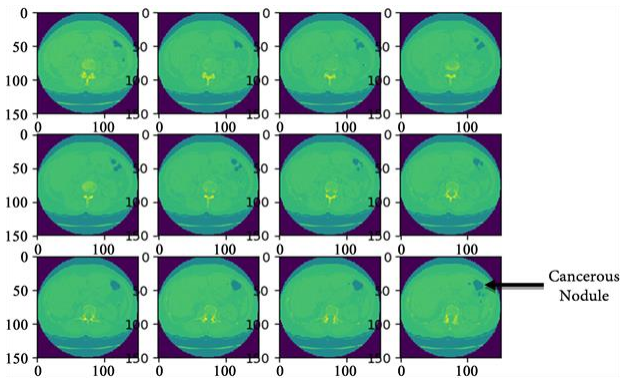

Figure 5: Sample LUNA16 scans

Nodules observed by various experts that are closer than the summation of their radii are combined. Now, the locations and diameters of these combined interpretations are averaged. This led to a group of 2290, 1602, 1186, and 777 nodules interpreted by at least 1, 2, 3, or 4 physicians, correspondingly. In this study, we decisively split the existing samples into 70% for learning and 30% of samples for testing. Some sample LUNA16 images are displayed in Figure 5. Also, since we used 10-folds cross-validation (10-FCV), the entire dataset is split into 10 fragments (each of 10%). Then, one fold (10%) is used for testing, while the remaining 90% instances are used for learning. The usage of 10-FCV ensures that each image in the dataset gets to be in a test exactly once.

4.2 Evaluation measures

The efficiency of the BSG segmentation method is evaluated by relating its performance with other existing segmentation algorithms in terms of selected measures such as the volume error (VE) and dice similarity coefficient (DSC). These measures are calculated by observing the difference between the segmentation fallouts and a physically annotated standard reference. The volume error is computed by Equation (42).

$$VE = \frac{2 \times (S - G)}{(S + G)} \qquad (42)$$

where G is the gold standard and S is the segmentation result gained from the BSG segmentation algorithm. For any clinical practice, $\text{volume error} < 5\%$ is more likely tolerable [31]. DSC is mostly used to state the performance of the segmentation algorithm. DSC is a resemblance measure calculated between two data points. DSC is defined as the fitness degree between the input scan and the isolated scan. The value of DSC is always in [0, 1] and it is calculated by Equation (43).

$$\text{DSC} = \frac{2 \times |G \cap S|}{|G| + |S|} \qquad (43)$$

The performance of the proposed lung cancer diagnostic model is evaluated through five performance measures including accuracy, sensitivity, specificity, precision, and JSS. These measures need to be bigger to improve the performance of the classifier. The utility of the model is calculated with respect to the accuracy (ACC) and it is defined by Equation (44). In this equation, T^+ represents the true positives which represent the number of instances that are correctly classified as cancerous; F^- denotes the false negatives that define the number of malevolent instances that are incorrectly classified as normal. T^-represents true negatives which state that the number of samples are correctly pigeonholed as benign and F^+ represents the false positives which state the number of normal samples that are erroneously classified as malevolent. Sensitivity, specificity, and precision define how well the algorithm tags malevolent and benign samples.

Sensitivity (SEN) defines the T^+ rate of the classifier. It is the capacity of an algorithm to appropriately classify an instance as malevolent. Specificity (SPE) is the estimate of the T^- rate of the algorithm. It is the capacity of the algorithm to correctly tag normal samples. Precision (PRE) is the capacity of an algorithm to classify only the related samples. It describes the number of malicious samples that really fit into the malicious class. JSS is a performance measure employed to calculate the performance of any classification algorithm. Given a dataset, the JSS reflects the resemblance between the calculated data and the gold standard for a sample available in the dataset. Equations (44) – (48) define the abovementioned performance measures.

$$\text{ACC} = \frac{T^- + T^+}{T^- + T^+ + F^- + F^+} \qquad (44)$$

$$\text{SEN} = \frac{T^+}{T^+ + F^-} \qquad (45)$$

$$\text{SPE} = \frac{T^-}{T^- + F^+} \qquad (46)$$

$$PRE = \frac{T^+}{T^+ + F^+} \qquad (47)$$

$$\text{JSS} = \frac{T^+}{T^+ + F^- + F^+} \qquad (48)$$

VII. RESULTS AND DISCUSSION

The developed diagnostic model is realized with appropriate preprocessing, image segmentation, feature selection, and COEL algorithms using the machine learning toolbox in MATLAB R2018b software. The database has been standardized in [−1, +1] before processing. To achieve more accurate results, the 10-FCV method is employed. As a result, the entire dataset is split into 10 fragments. For every run, one portion is used for testing, and the other portions are used to train the classification algorithm. Now, the average value of all the trials is considered for assessment.

7.1 Results from Segmentation

An extensive analysis of our segmentation outcomes reveals the potential and challenges of our proposed system. In most cases, regardless of the size of the standard reference, the proposed approach isolates the affected region accurately. Figure 6 illustrates sample input scans employed for assessment, isolated regions acquired by our diagnostic system, and their resultant ground truth. Although the nodules are in dissimilar arbitrary sections of the lung and appear in a different dimensions, the isolated nodules seem to overlap flawlessly. The proposed BSG achieves very low volume error (0.822±0.0012) and high DSC (0.976±0.014) when applied to LUNA16 dataset. By applying the concept of a superpixel, the BSG develops a more useful and perceptually important representation of the input scan. However, high-level features extracted by this model frequently do not include adequate high-resolution edge information of the input sample, causing amplified uncertainty in which high-resolution edges mostly affect the ultimate segmentation results.

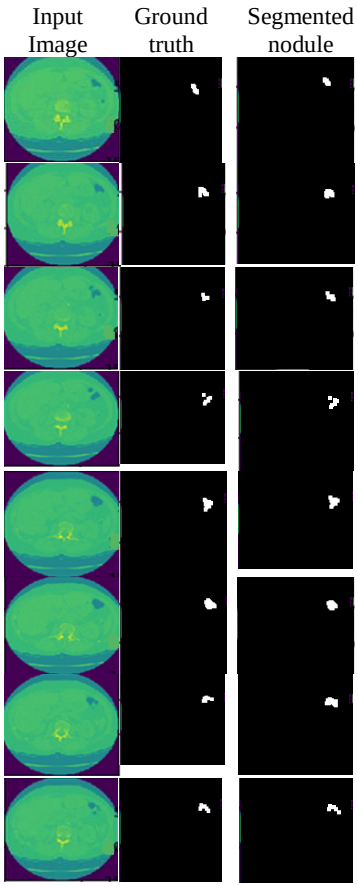

Figure 6: Segmentation results

7.2 Results from a classification

To prove the better performance of the COEL model, we relate the enactment of this model to other lung nodule diagnostic models found in the literature. The complete results gained from the developed diagnostic model are given in Table 1.

Table 1: Results obtained by the COEL on the LUNA16 dataset

Fold	ACC	SEN	SPE	PRE	JSS
#1	0.965	0.978	0.878	0.969	0.969
#2	0.975	0.958	0.897	0.995	0.961
#3	0.983	0.946	0.938	0.962	0.974
#4	0.977	0.969	0.957	0.969	0.968
#5	0.996	0.968	0.929	0.986	0.966
#6	0.987	0.967	0.910	0.942	0.961
#7	0.993	0.991	0.899	0.976	0.972
#8	0.976	0.991	0.885	0.962	0.942
#9	0.974	0.974	0.904	0.942	0.974
#10	0.969	0.978	0.880	0.979	0.961
Mean	0.980	0.972	0.908	0.969	0.965
SD	0.010	0.014	0.026	0.017	0.010

From Table 1, we can find that the intended COEL has realized better classification performance regarding ACC of 98.0%, SEN of 97.2%, SPE of 90.8%, PRE of 96.9%, and JSS of 96.5%. Besides, it is interesting to observe that the SD values gained by the COEL classifier are very small. It achieves performance measures in terms of SD as ACC of 1.0%, SEN of 1.4%, SPE of 2.6%, PRE of 1.7%, and JSS of 1.0%. This divulges the reliability and robustness of the intended system. Figure 7 demonstrates the efficiency of the proposed model in terms of performance measures.

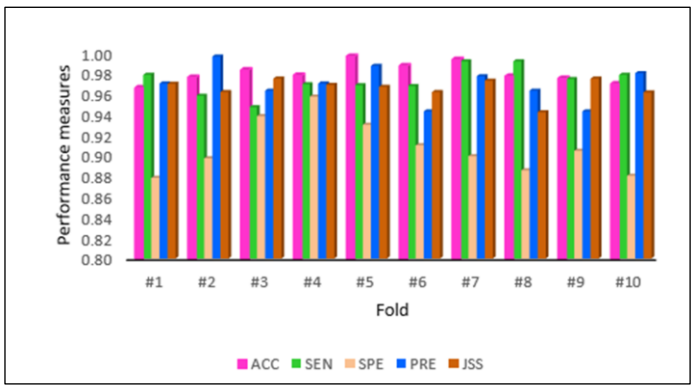

Figure 7: Performance measures gained by the COEL on the LUNA16 dataset

Table 2 illustrates the results obtained from various lung nodule classifiers. From this table, the AlexNet generates nominal classification outcomes with the mean value of 83.4% ACC, 88.6% SEN, 83.1% SPE, 94.8% PRE, and 92.6% JSS related to ground truth references. With the competence of generating high-level feature depictions from input scans, ResNet generates improved results than the AlexNet classifier with respect to classification ACC (85.5%), SEN (90.1%), SPE (83.8%), PRE (86.6%), and JSS (91.1%).

Table 2: Effectiveness of various classifiers in terms of performance measures

Algorithm	Criteria	ACC	SEN	SPE	PRE	JSS
AlexNet	Mean	0.834	0.886	0.831	0.948	0.926
	SD	0.034	0.070	0.009	0.044	0.048
ResNet	Mean	0.855	0.901	0.838	0.866	0.911
	SD	0.041	0.082	0.028	0.038	0.036
ELM	Mean	0.861	0.927	0.841	0.896	0.922
	SD	0.036	0.033	0.037	0.046	0.040
TumorNet	Mean	0.866	0.953	0.845	0.924	0.921
	SD	0.038	0.031	0.027	0.036	0.022
DFCNet	Mean	0.892	0.953	0.874	0.856	0.901
	SD	0.049	0.051	0.027	0.032	0.010
CMixNet	Mean	0.914	0.954	0.875	0.962	0.955
	SD	0.033	0.029	0.028	0.023	0.009
COEL	Mean	0.980	0.972	0.908	0.969	0.965
	SD	0.010	0.014	0.026	0.017	0.010

By applying identity mapping to handl the vanishing gradient problem, the ResNet classifier achieves 85.5% of ACC, 90.1% of SEN, 83.8% of SPE, 86.6% of PRE, and 91.1% of JSS. However, it exhibits deprived performance regarding SD values owing to the biases and weights of this model selected arbitrarily to classify the input samples. It provides 4.1% of ACC, 8.2% of SEN, 2.8% of SPE, 3.8% of PRE, and 3.6% of JSS. The conventional ELM classifier enables better performance compared with AlexNet and ResNet classifiers with 86.1% ACC, 92.7% SEN, 84.1% SPE, 89.6% PRE, and JSS of 92.2%. Since this classification algorithm is based on the random generation of the initial population it provides exclusive least-squares solution of the output parameter and increases the effectiveness of the system.

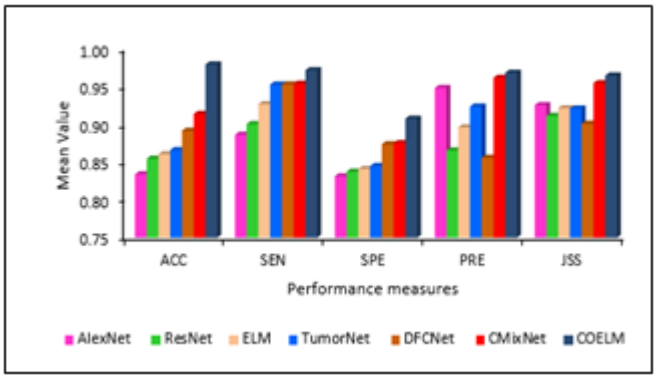

Figure 8: Output of COEL on the LUNA16 dataset regarding the mean value of performance measures

By applying a multi-view convolutional neural network with a Gaussian process, TumorNet achieves a similar performance as the ELM classifier. It provides 86.6% ACC, 95.3% SEN, 84.5% SPE, 92.4% PRE, and JSS of 92.1%. At the same time, it provides improved results regarding SD values with 3.8% ACC, 3.1% SEN, 2.7% SPE, 3.6% PRE, and 2.2% JSS related to other classifiers. Similarly, DFCNet provides better results in terms of the mean value of performance measures (ACC (89.2%), SEN (95.3%), SPE (87.4%), PRE (85.6%), and JSS (90.1%)) and deprived result in terms of SD values (ACC (4.9%), SEN (5.1%), SPE (2.7%), PRE (3.2%), and JSS (1.0%)). The CMixNet provides improved performance in terms of the mean value (ACC (91.4%), SEN (95.4%), SPE (87.5%), PRE (96.2%), and JSS (95.5%)) and better result in terms of SD values (ACC (3.3%), SEN (2.9%), SPE (2.8%), PRE (2.3%), and JSS (0.9%)).

The proposed COEL model outperforms all other prevailing classification algorithms in terms of performance metrics with 98% classification accuracy, 97.2% sensitivity, 90.8% specificity, 96.9% precision, and 96.5% similarity index. Besides, COEL

provides much better fallouts regarding SD values with 1% of ACC, 1.4% of SEN, 2.6% of SPE, 1.7% of PRE, and 1.0% of JSS as compared with other classification algorithms. It is noteworthy that the COEL has achieved improved results as compared to all other advanced classification models. Also, the SD values obtained from the COEL are less than that of all other classifiers which demonstrates that the COEL can provide more dependable fallouts.

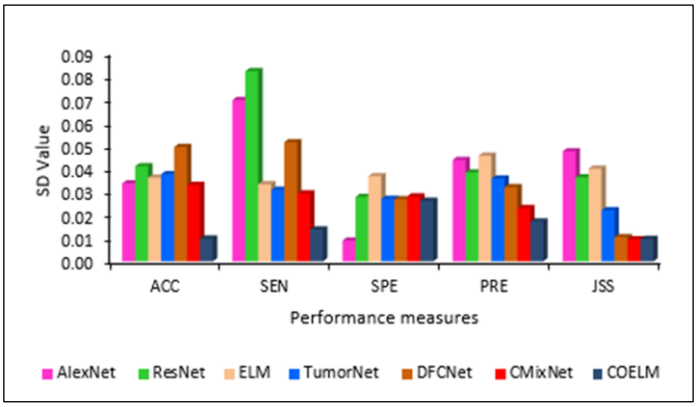

Figure 9: Output of COEL on the LUNA16 dataset regarding the SD value of performance measures

Figures 8 and 9 display the performance measures gained by all the classifiers including COEL designated for performance evaluation. The capacity of BSG provides an efficient segmentation approach for the COEL classifier. The COEL realized improved predictive performance. Similarly, it is interesting to see that the SD values of the measures gained by the COEL classifier are lower than that of all other classifiers which indicate that the COEL classifier can deliver more reliable and robust performance.

Table 3: COEL Vs other classifiers regarding the accuracy

Fold	Alex Net	Res Net	ELM	Tumor Net	DFC Net	CMix Net.	COEL
#1	0.851	0.859	0.812	0.816	0.824	0.912	0.965
#2	0.831	0.884	0.887	0.887	0.876	0.938	0.975
#3	0.817	0.901	0.850	0.882	0.887	0.965	0.983
#4	0.784	0.854	0.892	0.892	0.893	0.885	0.977
#5	0.844	0.799	0.909	0.909	0.885	0.962	0.996
#6	0.865	0.798	0.814	0.817	0.890	0.864	0.987
#7	0.860	0.888	0.825	0.825	0.852	0.890	0.993
#8	0.771	0.885	0.849	0.849	0.857	0.896	0.976
#9	0.869	0.882	0.868	0.868	0.972	0.906	0.974
#10	0.848	0.799	0.900	0.917	0.981	0.921	0.969
Mean	0.834	0.855	0.861	0.866	0.892	0.914	0.980
S.D	0.034	0.041	0.036	0.038	0.049	0.033	0.010

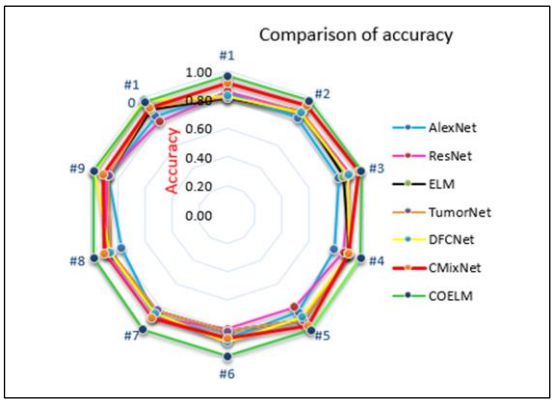

Figure 10: COEL Vs other classifiers regarding the accuracy

Tables 3 - 7 provide the outcomes of all the classifiers for different folding. The mean and SD values achieved by every classification model are listed. It is observed that the evaluation metrics acquired by the COEL are greater than all other classifiers in most cases. Therefore, the outcomes prove that the amalgamation of BSG and COEL provides superior outcomes related to all other models employed in this work. This divulges that the amalgamation of BSG and COEL significantly increases the classification performance.

Table 4: COEL Vs other classifiers regarding sensitivity

Fold	Alex Net	Res Net	ELM	Tumor Net	DFC Net	CMix Net.	COEL
#1	0.817	0.832	0.841	0.884	0.816	0.931	0.978
#2	0.785	0.977	0.947	0.968	0.938	0.938	0.958
#3	0.960	0.981	0.962	0.960	0.988	0.937	0.946
#4	0.966	0.824	0.907	0.947	0.945	0.916	0.969
#5	0.869	0.752	0.939	0.941	0.968	0.920	0.968
#6	0.815	0.944	0.942	0.964	0.987	0.994	0.967
#7	0.908	0.972	0.927	0.924	0.955	0.958	0.991
#8	0.951	0.966	0.932	0.972	0.970	0.973	0.991
#9	0.838	0.835	0.930	0.988	0.978	0.981	0.974
#10	0.955	0.922	0.940	0.980	0.986	0.992	0.978
Mean	0.886	0.901	0.927	0.953	0.953	0.954	0.972
S.D	0.070	0.082	0.033	0.031	0.051	0.029	0.014

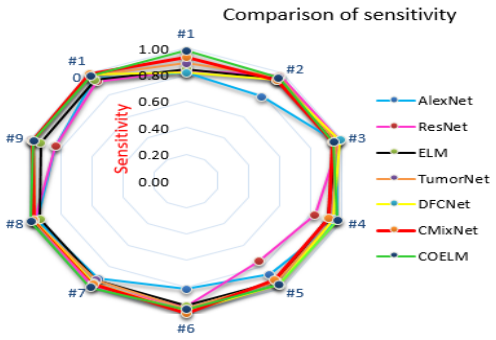

Figure 11: COEL Vs other classifiers regarding sensitivity

Figures 10 – 15 reveal the improved performance of the developed COEL classifier. The integration of BSG with COEL shows improved results related to other lung image classifiers used in this study. Similarly, it is significant that COEL outperforms other classifiers in most cases regarding SD values. This reveals that the combination of the BSG with COEL has extensively improved the enactment of the classifier.

Table 5: COEL Vs other classifiers regarding specificity

Fold	Alex Net	Res Net	ELM	Tumor Net	DFC Net	CMix Net.	COEL
#1	0.826	0.875	0.824	0.889	0.905	0.910	0.878
#2	0.826	0.862	0.884	0.876	0.878	0.899	0.897
#3	0.844	0.861	0.795	0.855	0.843	0.885	0.938
#4	0.832	0.862	0.920	0.843	0.834	0.920	0.957
#5	0.845	0.804	0.827	0.804	0.887	0.872	0.929
#6	0.817	0.825	0.850	0.839	0.843	0.862	0.910
#7	0.838	0.801	0.806	0.815	0.894	0.837	0.899
#8	0.822	0.806	0.829	0.820	0.867	0.875	0.885
#9	0.828	0.844	0.837	0.858	0.877	0.854	0.904
#10	0.835	0.837	0.840	0.851	0.911	0.841	0.880
Mean	0.831	0.838	0.841	0.845	0.874	0.875	0.908
S.D	0.009	0.028	0.037	0.027	0.027	0.028	0.026

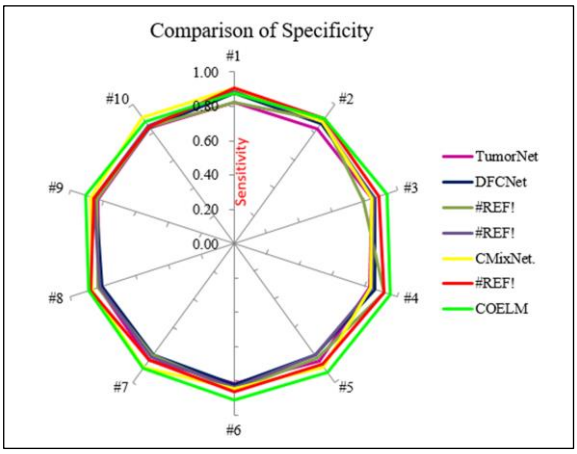

Figure 12: COEL Vs other classifiers regarding specificity

Table 6: COEL Vs other classifiers regarding precision

Fold	Alex Net	Res Net	ELM	Tumor Net	DFC Net	CMix Net.	COEL
#1	0.888	0.800	0.885	0.866	0.855	0.961	0.969
#2	0.902	0.837	0.926	0.926	0.900	0.980	0.995
#3	0.989	0.848	0.966	0.985	0.894	0.944	0.962
#4	0.968	0.901	0.879	0.971	0.849	0.938	0.969
#5	0.965	0.899	0.807	0.895	0.881	0.975	0.986

#6	0.997	0.884	0.925	0.927	0.854	0.922	0.942
#7	0.901	0.903	0.935	0.896	0.869	0.953	0.976
#8	0.905	0.909	0.912	0.912	0.844	0.999	0.962
#9	0.972	0.827	0.876	0.915	0.812	0.978	0.942
#10	0.994	0.847	0.856	0.943	0.802	0.969	0.979
Mean	0.948	0.866	0.896	0.924	0.856	0.962	0.969
S.D	0.044	0.038	0.046	0.036	0.032	0.023	0.017

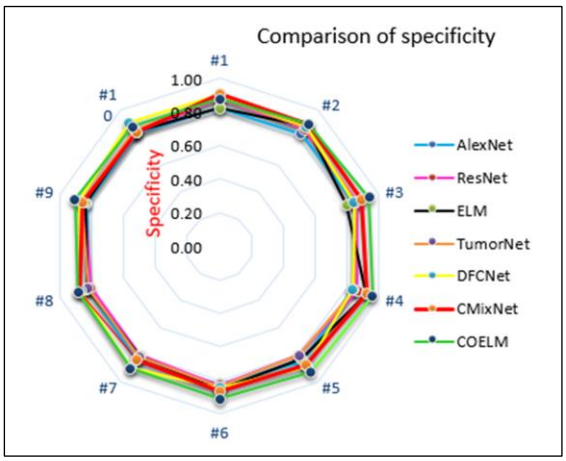

Figure 13: COEL Vs other classifiers regarding precision

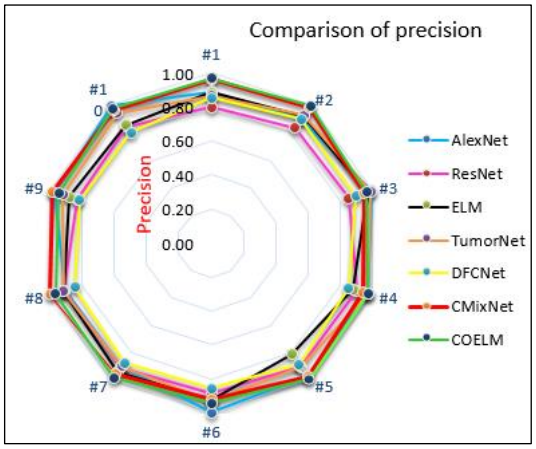

Figure 14: COEL Vs other classifiers regarding precision

Table 7: COEL Vs other classifiers regarding JSS

Fold	Alex Net	Res Net	ELM	Tumor Net	DFC Net	CMix Net.	COEL
#1	0.851	0.876	0.907	0.881	0.888	0.962	0.969
#2	0.888	0.867	0.875	0.906	0.901	0.951	0.961
#3	0.873	0.899	0.889	0.905	0.907	0.965	0.974
#4	0.888	0.880	0.877	0.918	0.905	0.966	0.968

#5	0.961	0.898	0.948	0.911	0.904	0.952	0.966
#6	0.981	0.952	0.931	0.920	0.885	0.935	0.961
#7	0.944	0.972	0.897	0.928	0.890	0.961	0.972
#8	0.984	0.908	0.924	0.942	0.902	0.947	0.942
#9	0.959	0.906	0.978	0.951	0.919	0.956	0.974
#10	0.926	0.955	0.988	0.949	0.908	0.951	0.961
Mean	0.926	0.911	0.922	0.921	0.901	0.955	0.965
S.D	0.048	0.036	0.040	0.022	0.010	0.009	0.010

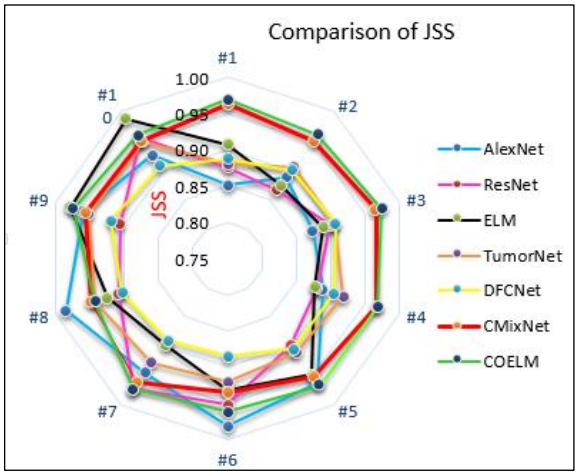

Figure 15: COEL Vs other classifiers regarding JSS

VIII.CONCLUSION

Precise isolation and classification of lung nodules from CT scans are of inordinate importance to developing an automatic diagnosis system for lung nodules. Numerous studies indicate that the ELM is an advanced machine learning algorithm to train lung cancer diagnostic models and classify the new images into benign and malignant based on significant features. Though ELM is employed to improve the effectiveness of lung nodule identification models, the predictive performance of this model still necessitates substantial improvement to satisfy the real-time demands of healthcare applications. In this context, this study proposes a novel COEL to differentiate the CT images efficiently. The proposed COEL exploits the concept of column-wise parameter tuning whose key idea is that the tuning of the multifaceted resultant matrix is divided into fine-tuning of the matrix with a single column vector. Besides, this study proposes a WIMSE to enhance the quality of the scan and a BSG algorithm to segment the affected regions from the input scan. The performance of this model is evaluated on the LUNA16 dataset by comparing results gained from the COEL with other existing classification models. The results of a comprehensive empirical analysis of the LUNA16 database divulge that the intended COEL considerably outdoes other classifiers with improved performance. From extensive experimental results, we can conclude that the proposed COEL is the better diagnostic system related to advanced lung cancer classification models.

REFERENCES

1. Prabhu S., Prasad K., Robels-Kelly A., Lu X. (2022). AI-based carcinoma detection and classification using histopathological images: A systematic review Comput. Biol. Med., Article 105209
2. Sori W.J., Feng J., Godana A.W., Liu S., Gelmecha D.J. (2021). DFD-Net: Lung cancer detection from denoised CT scan image using deep learning Front. Comput. Sci., 15 (2), pp. 1-13
3. Cao W., Wu R., Cao G., He Z.A (2020). Comprehensive review of computer-aided diagnosis of pulmonary nodules based on computed tomography scans, IEEE Access, 8, pp. 154007-154023

4. V. Gowri and V. V. Chamundeeswari (2023). "Classifying big medical data through bootstrap decision forest using penalizing attributes," Intelligent Automation & Soft Computing, vol. 36, no.3, pp. 3675–3690.
5. Pang S., Fan M., Wang X., Wang J., Song T., Wang X., Cheng X, (2020). VGG16-T: A novel deep convolutional neural network with boosting to identify pathological type of lung cancer in early stage by CT images, Int. J. Comput. Intell. Syst., 13 (1), p. 771
6. Shen S., Han S.X., Aberle D.R., Bui A.A., Hsu W (2019). An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification, Expert Syst. Appl., 128, pp. 84-95.
7. Li S., Liu D., (2021). Automated classification of solitary pulmonary nodules using convolutional neural network based on transfer learning strategy, J. Mech. Med. Biol., Article 2140002
8. Postmus, P. E. et al. (2017). Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann. Oncol., vol. 28, pp. 1-21.
9. Tanzila Saba, (2020). Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges, Journal of Infection and Public Health, vol. 13, no. 9, pp. 1274-1289.
10. Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. (2006). Extreme learning machine: Theory and applications. Neurocomputing vol. 70, 489–501.
11. Deo, R.C.; Tiwari, M.K.; Adamowski, J.F.; Quilty, J.M. Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. Stoch. Environ. Res. Risk Assess., 31, 1211–1240.
12. Huang, G.-B.; Chen, L. (2007). Convex incremental extreme learning machine. Neurocomputing, 70, 3056–3062.
13. Zhou, Z.; Chen, J.; Zhu, Z. (2018). Regularization incremental extreme learning machine with random reduced kernel for regression. Neurocomputing, 321, 72–81.
14. Xu, M., Qi, S., Yue, Y., (2019). 'Segmentation of lung parenchyma in CT images using CNN trained with the clustering algorithm generated dataset', Bio Medical Engineering, 18, (1), pp. 1–21.
15. Liu, C, Pang, M, Zhao, R, (2019). 'Novel superpixel-based algorithm for segmenting lung images via convolutional neural network and random forest,' IET Image Processing, vol. 14, no. 16, pp. 4340-4348.
16. Xie, H.; Yang, D.; Sun, N.; Chen, Z.; Zhang, Y (2019). Automated Pulmonary Nodule Detection in CT Images Using Deep Convolutional Neural Networks, Pattern Recognition, 85, 109–119.
17. Nakrani MG, Sable GS, Shinde UB (2020). ResNet based lung nodules detection from computed tomography images, International Journal of Innovative Technology Exploring Engineering, pp. 1711–1724. F.
18. Zhang, Q. Wang, H. Li, (2020). Automatic segmentation of the gross target volume in non-small cell lung cancer using a modified version of ResNet, Technology in Cancer Research & Treatment 19, 1–9.
19. J. Jiang, Y. C. Hu, C. J. Liu, D. Halpenny, M. D. Hellmann, J. O. Deasy, G. Mageras, H. Veeraraghavan, (2019). Multiple resolution residually connected feature streams for automatic lung tumor segmentation from CT images, IEEE Transactions on Medical Imaging 38 (1) 134–144.
20. Wang, X, Cheng, X, (2022). 'VGG16-T: A Novel Deep Convolutional Neural Network with Boosting to Identify Pathological Type of Lung Cancer in Early Stage by CT Images', International Journal of Computational Intelligence Systems, vol. 13, no. 1, pp. 771-780.
21. Ali, I.; Hart, G.R.; Gunabushanam, G.; Liang, Y.; Muhammad, W.; Nartowt, B.; Kane, M.; Ma, X.; Deng, J (2018). Lung Nodule Detection via Deep Reinforcement Learning, Frontier Oncology, vol. 8, pp. 213-221.
22. U. Kamal, A. M. Rafi, R. Hoque, J. Wu, M. K. Hasan (2020). Lung cancer tumor region segmentation using recurrent 3d-denseunet, in: J. Petersen, R. San Jos´e Est´epar, A. Schmidt-Richberg, S. Gerard, B. Lassen-Schmidt, C. Jacobs, R. Beichel, K. Mori (Eds.), Thoracic Image Analysis, Springer International Publishing, Cham, pp. 36–47.
23. Deo, R.C. Tiwari, M.K. Adamowski, J.F. Quilty, J.M. (2017). Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. Stoch. Environ. Res. Risk Assess. 31, 1211–1240.
24. Zhou, Z. Chen, J. Zhu, Z. (2018). Regularization incremental extreme learning machine with random reduced kernel for regression. Neurocomputing 321, 72–81.
25. Zhang, Z. Du, K (2006). Successive projection method for solving the unbalanced Procrustes problem. Sci. China Ser. A, 49, 971–986.
26. A. Agarwal, K. Patni and R. D (2021). Lung Cancer Detection and Classification Based on Alexnet CNN, 2021 6th International Conference on Communication and Electronics Systems (ICCES), pp. 1390-1397.
27. Hussein, S. Gillies, R.; Cao, K. Song, Q. Bagci, U. (2017). TumorNet: Lung nodule characterization using multi-view convolutional neural network with gaussian process. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, Australia, 18–21, pp. 1007–1010.
28. Masood, A. Sheng, B. Li, P. Hou, X. Wei, X. Qin, J. Feng, D (2018). Computer-Assisted Decision Support System in Pulmonary Cancer detection and stage classification on CT images. J. Biomed. Inform. 79, 117–128.
29. Ronneberger, O. Fischer, P. Brox, T (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention; Springer: Berlin/Heidelberg, Germany, Volume 9351, pp. 234–241.

30. Setio, A.A.A. Traverso, A. de Bel, T. Berens, M.S.N. van den Bogaard, C. Cerello, P. Chen, H. Dou, Q. Fantacci, M.E. Geurts, B. et al. (2017). Validation, Comparison, and Combination of Algorithms for Automatic Detection of Pulmonary Nodules in Computed Tomography Images: The LUNA16 Challenge. Med. Image Anal. 2017, 42, 1–13.
31. Kumar, S.S., Moni, R.S, Rajeesh, J., (2013). Automatic liver and lesion segmentation: a primary step in diagnosis of liver diseases, Signal, Image and Video Processing, vol. 7, pp.163–172.