Bulletin of Electrical Engineering and Informatics
Vol. 13, No. 3, June 2024, pp. 2009~2018
ISSN: 2302-9285, DOI: 10.11591/eei.v13i3.6828

Automated 3D convolutional neural network architecture design using genetic algorithm for pulmonary nodule classification

Kamel Hussein Rahouma[1], Shahenda Mahmoud Mabrouk[2], Mohamed Aouf[2]
[1]Department of Electrical Engineering, Faculty of Engineering, Minia University, Minia, Egypt
[2]Department of Biomedical Engineering, Higher Technological Institute, 10th of Ramadan City, Cairo, Egypt

Article Info

Article history:

Received May 29, 2023
Revised Sep 28, 2023
Accepted Oct 12, 2023

Keywords:

Convolutional neural network
Deep neural networks
Dense block
Evolutionary algorithm
Genetic algorithm
Lung nodule classification

ABSTRACT

Cancer of the lungs is considered one of the primary causes of death among patients globally. Early detection contributes significantly to the success of pulmonary cancer treatment. To aid the pulmonary nodule classification, many models for the analysis of medical image utilizing deep learning have been developed. Convolutional neural network (CNN) recently, has attained remarkable results in various image classification tasks. Nevertheless, the CNNs performance is heavily dependent on their architectures which still heavily reliant on human domain knowledge. This study introduces a cutting-edge approach that leverages genetic algorithms (GAs) to automatically design 3D CNN architectures for differentiation between benign and malignant pulmonary nodules. The suggested algorithm utilizes the dataset of lung nodule analysis 2016 (LUNA16) for evaluation. Notably, our approach achieved exceptional model accuracy, with evaluations on the testing dataset yielding up to 95.977%. Furthermore, the algorithm exhibited high sensitivity, showcasing its robust performance in distinguishing between benign and malignant nodules. Our findings demonstrate the outstanding capabilities of the proposed algorithm and show an outstanding performance and attain a state of art solution in lung nodule classification.

Corresponding Author:

Shahenda Mahmoud Mabrouk
Department of Biomedical Engineering, Higher Technological Institute
Industrial Area2, 10th of Ramadan City, Ash Sharqia Governorate, Egypt
Email: shahenda-mahmoud@hotmail.com

1. INTRODUCTION

Pulmonary cancer ranks among the primary causes of death worldwide. In 2020 the estimated number of new cases diagnosed with lung cancer reaches about 2.2 million new cases worldwide and these numbers are expected to increase in near future [1]–[3]. Cancer is an uncontrolled malignant growth in the lung tissues [4]. It may also cause metastasis, the spread of cancer from the lung into other organs which increases the risk of death. That's why early detection and treatment significantly boost the success of treatment and decrease mortality rates in diagnosed patients [5], [6]. Computed tomography (CT) screening has proven effective in early detection of lung nodules, offering a potential solution to mitigate this situation and decreases lung cancer mortality rates [7]. However, the manual nodule detection process is laborious and time-consuming for radiologists since it requires a long time owing to the fact that they review sheer volume of scans in a day, which may affect their capacity to accurately identifying and classifying tumors [2], [8]. Even expert radiologists sometimes faces difficulty detecting and diagnosing lung nodules in CT scans [9]. Also, the accuracy of a radiologist's diagnosis are heavily influenced by the individual clinician's experience

[10]. Recently, computer-aided diagnosis (CAD) systems have emerged as a valuable tool for easing the burden on radiologists by providing objective prediction with non-invasive solution to aid radiologists to diagnose pulmonary nodules [3], [7], [9]. Typically, CAD systems for lung nodule detection involve five stages: i) image acquisition, ii) preprocessing, iii) lung segmentation, iv) nodule detection, and v) classification [6]. This study specifically focuses on the classification method for lung nodule. The existing CAD systems can be divided into two categories: handcrafted feature extraction classification models and deep neural network classification models with automated feature extraction [7], [10].

The approaches in the first category usually assess the radiological characteristics, like nodule shape, texture, and size then it uses a classifier to indicate the status of malignancy. The processing processes carried out before the feature extraction stage have some bearing on how well a CAD system doing Handcraft or standard feature extraction performs. As a result, CT scan images are subjected to image enhancement and image segmentation techniques. Approaches used throughout the improvement and segmentation rounds are sometimes manually and carefully tweaked in order to prepare the region of interest (RoI) for feature extraction [7]. Rustam *et al.* [8] used feature extraction techniques including gray level co-occurrence matrices (GLCM), lung nodule size, and local binary pattern (LBP) to extract features from the SPIE-AAPM-NCI Lungx Challenge in 2015. Fuzzy kernel C-means and fuzzy C-means were employed for classifying 2D lung nodules images as benign or malignant. Using Python 3.7, fuzzy kernel C-means achieved the highest accuracy, recall, and f1 score with 74.1%, 69.2%, and 72% respectively. However, fuzzy C-means achieved the highest specificity of 87.5%, and both classifiers achieved the best precision of 80%. Asuntha and Srinivasan [11] developed a novel method named as "fuzzy particle swarm optimization convolution neural network (FPSOCNN)". It applies fuzzy particle swarm optimization (FPSO) to select optimal features from various geometric, texture, intensity, and volumetric features extracted using different techniques. Then these features are classified using 2D CNN to classify the lung cancer type. The system reduces the computational complexity of CNN and achieves accuracy of 95.62% on the dataset of the lung image database consortium and image database resource initiative (LIDC-IDRI). Despite the great performance of the handcrafted feature extraction classification models, the gathering and choosing a valid features subset for diagnosis of pulmonary nodules is time-consuming and risky, as it introduces measurement errors that affect the classifier performance and accuracy [10].

In contrast, the second category automatically learns from the CT scans the nodule features by utilizing deep neural network-based models. These are high performing models with great prediction accuracy for lung nodule classification. This sort of system often requires less complicated denoising and image enhancement techniques, instead employing image preprocessing methods such as cropping, zooming or image rotation to enhance diagnostic efficiency [7]. Silva *et al.* [5] presented a system utilizing deep learning in conjunction with the GA to classify lung lesions on CT scans. The system does not need feature extraction or selection which reduces the system computing complexity. The system used images of size 28×28 dimension as input. This system was tested on the dataset of LIDC-IDRI, achieving an accuracy of 94.78%, specificity of 95.14%, sensitivity of 94.66%, and an area under the curve (AUC) of 0.949. Where Dey *et al.* [10] presented four two-pathway 3D CNNs to classify CT image of pulmonary nodules as benign or malignant. Each pathway accepted images with the size 50×50 pixels×5 slices and 100×100 pixels×10 slices respectively. Accordingly, the "multi-output DenseNet (MoDenseNet)" outperformed the other three models with AUC of 0.9548 and accuracy of 90.40%. The model size and total number of parameters was equal to 34.8 million. While Pfeffer and Ling [12] adopted the evolutionary algorithm presented in [13], and implement it to classify whether pulmonary nodules visible on CT scans are malignant or benign. The algorithm is the implementation of GA for 2D CNN architecture design. The authors used images with size 32×32 pixels×1 input channel. The algorithm named CNN-GA presented a model with accuracy equal to 91.3%. Another classification technique called bilinear convolutional neural network (BCNN) was introduced in [14]. The (VGG16 and VGG19) two-stream CNNs were used in BCNN as feature extractors, to reduce false positives before using a support vector machine (SVM) as classifier. The public LUNA16 database was used to evaluate the method's performance. The system used images with size 50×50 for training and validation. The BCNN attained a 91.99% accuracy rate and a 0.959 AUC rate. Jiang *et al.* [15] developed a lung nodule classifier known as "NASLung" using neural architecture search (NAS). This approach automatically seeks a 3D architecture network with high speed and accuracy. The use of the convolutional block attention module (CBAM) and SoftMax loss during training helped learn representations of angularly discriminative features. Using the dataset of LIDC-IDRI, the model reached 90.77% acc. Additionally, in a study by Al-Huseiny [16], GoogLeNet was employed for transfer learning using the IQ-OTH/NCCD lung cancer dataset. The system promises to isolate the lungs area and eliminate any leftover clutter and unnecessary surrounds using inexpensive pre-processing process. According to experimental findings, the trained model has improved its overall accuracy on the validation data by 94.38%.

Despite this approach can give a promising performance, but it still demands the expertise of deep learning hyper-parameters incorporation with the knowledge of the problem space [13]. Also, studies have shown that model performance can vary significantly to data sets with different domains even if it has the same network architectures and hyper-parameter [17], [18]. Furthermore, the majority of the current deep learning models use 2D CNN networks to solve the lung nodule classification problem [18] which may cause the discard of some important features of the original image. For these reasons, there is a demand for a smart and automated 3D CNN architecture design that can be related to the entered data and does not necessitate the presence of experts.

In order to fulfill these demands, we propose in this study, an efficient 3D CNN architectural design with the deployment of evolutionary algorithm GA to automate the CNN design to classify the pulmonary nodules whether it is malignant or benign from 3D CT lung images. The algorithm's importance lies in discovering the best 3D CNN architecture for the input data automatically, without any manual interference through the evolutionary search so there is no need for extracting or selecting the proper features or need of experts to build the proper CNN model for entered data. The algorithm employs GA to evolve and select the best CNN architecture by optimizing the network parameters such as: i) the number of CNN basic building blocks, ii) the value of feature map for the convolutional layers, and iii) the learning rate value of the model. The proposed algorithm used the LUNA16 dataset [19] for training and evaluation. The subsequent sections detail the paper's structure as follows: section 2 outlines the main steps and details of the proposed algorithm. Section 3 discusses the obtained results, and section 4 presents the concluding remarks and outlines.

2. METHOD

This section illustrates the proposed algorithm by means of a genetic algorithm as depicted in Figure 1. The framework encompasses a series of steps involving the collection, cropping, and classification of a total of 868 lung nodules. The training and validation of the model are performed using a 3D image dataset. The subsequent subsections elaborate on the distinct stages of this framework.

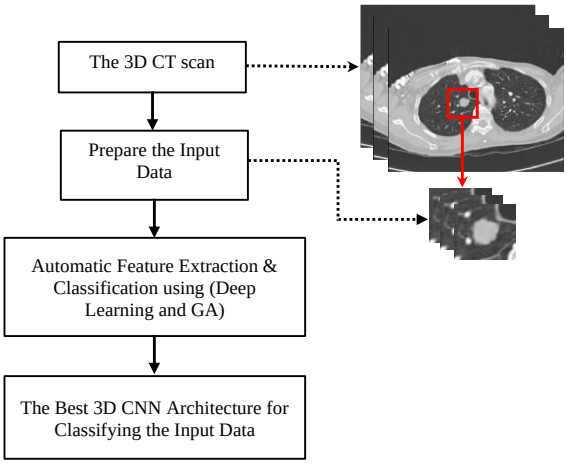


Figure 1. The proposed system framework

2.1. Dataset and pre-processing

The pulmonary nodule CT scans dataset used in this study is the LUNA16 dataset. It is a subset from LIDC-IDRI dataset [20]. The LIDC-IDRI is a public dataset that contains 1,010 CT scans with one or more pulmonary nodules can be seen in each image associated with annotations by one to four radiologists [12]. Annotations used included radiologists' ratings of nodules varying from one to five, with one indicating a small likelihood of malignancy and five indicating the highest likelihood of malignancy [21]. The heterogeneous scans in the LUNA16 dataset are filtered using by various criteria [19], [22]. A thin slice should be used since pulmonary nodules might be rather tiny. Consequently, scans having a slice thickness >2.5 mm were disregarded. Additionally, images that had missing slices or uneven slices spacing were disregarded. This resulted in 888 CT scans and a total of 1,186 nodules. Nonetheless, not all the scans were considered in this study. We applied the following criteria to determine the appropriate sample and its class membership whether it is malignant or benign. First, at least three radiologists should diagnose the lung

nodule. Then, the mean value of the annotated nodule was computed. If the calculated average value was greater than 3 then it is malignant. Conversely, if the average value was less than 3 then it is benign. Nodules have 3 as mean value, were excluded from the dataset. This process resulted with 434 malignant and 540 benign nodules. We performed further balancing to dataset so that the malignant data is equal to the benign one. Ultimately, we end up with 868 pulmonary nodules (50% malignant and 50% benign) for training, validation, and testing.

After applying these criteria, we proceed with the preprocessing of the CT scan images. Starting with the Hounsfield units (HU) values, we convert them from the range of (-1,200, 600) HU to a normalized range of [0, 1]. These normalized values are then scaled to an 8-bit image format within the range of [0, 255]. This transformation allows for better visualization and further processing of the CT images. Next, we resampled the data to 1×1×1 pixels. Then each patient's lung nodule coordinates are used to crop the picture from 512×512×no. of slices, to dimensions (48×48 pixels×8 slices×1 input channel) and saved in PNG format. In essence, the code performs several tasks, including processing input data, extracting nodules associated with each image, generating cropped nodule images, and saving slices of each case in PNG format in a separate folder categorized by their malignancy status. With this understanding of how the input data is handled, we can now proceed to describe the proposed algorithm in detail.

2.2. Convolutional neural networks

This subsection introduces the proposed algorithm CNNs basic building unit as shown in Figure 2, where its fundamental units are convolutional and pooling layers. The basic building unit used in the proposed network is adopted from dense block. A dense block is a CNN module that attaches all layers having feature-map sizes that match directly to one another. It was initially proposed as a structure of the DenseNet architecture [23]. To maintain the nature of feed-forward, every layer acquires from all previous layers' additional inputs, and passes on to all subsequent layers it is own feature-maps. DenseNet combines the features by concatenating them, in contrast to ResNets, which sums the features before passing them into a layer [23]. Each basic block used in the proposed algorithm contains no. of convolutional layers with a 3D filter of size 3×3×3. GA operators generate the blocks number in each individual as we illustrate in the next subsection. As the number of blocks increases, so does the number of connected convolutional layers within each block. Specifically, the first block contains 4 layers, the second block will contain 10 layers, and the third and subsequent blocks each will contain 20 layers. After each block a transition layer of a 3D max pooling with kernel size 2×2×1 is used, followed by a dropout layer to help prevent over fitting, during training. Padding is used with each convolution layer to maintain feature map size. Batch normalizing and Relu activation function were appended through each convolution layer. At the end of each individually generated model, we add another convolution and max pooling layers of size 2×2×2 before flattening the output. Since our dataset is only divided into two classes, the loss function was computed using binary cross entropy. This function measures the discrepancy between the predicted probability distribution and the actual probability distribution for a binary classification problem. In (1) represents the binary cross-entropy:

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \,.\, log\big(p(y_i)\big) + (1 - y_i) \,.\, log\,(1 - p(y_i)) \quad (1)$$

Where y_i represent the labels (0 or 1), $p(y_i)$ represent the probability of 1 and $(1 - p(y_i))$ represent the probability of 0, and N represent the samples number. For both target classes, the final classification was generated using a sigmoid classifier.

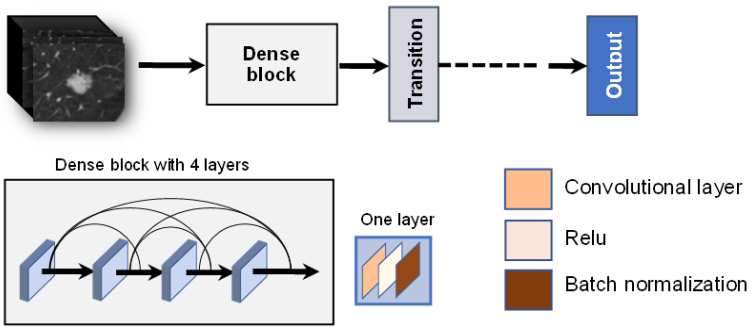


Figure 2. The proposed algorithm CNN basic building unit

2.3. Genetic algorithms

GAs employs bio-inspired, processes like selection, crossover, and mutation to obtain the most effective solution to any presented problem [24]. This algorithm mirrors the natural selection process where the fittest individuals are chosen for reproduction, resulting in offspring of the next generation. The illustration in Figure 3 depicts a general concept for GAs based on the individuals' selection for the offspring, randomized cross-over, mutation that leads to the proper individuals in order to obtain the most effective solution for the given problem. Each candidate is assessed and assigned a fitness value, often a function of the decoded bits in their chromosome. These candidates are chosen for reproduction based on their fitness values, guided by selection criteria. These selected individuals, referred to as parents, contribute to the subsequent generation of the population. The crossover mechanism combines two parents to generate children for the next generation, while the mutation mechanism introduces random changes to individual parents to create children.

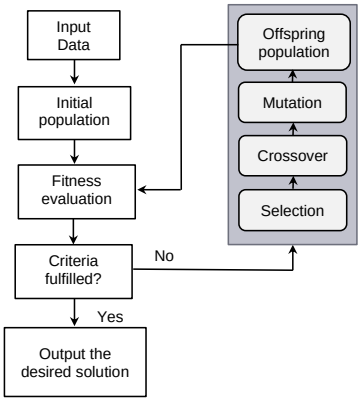


Figure 3. The flow chart of genetic algorithm

Regarding image classification, GAs can thus aid in the automatic and efficient discovery of optimum solutions for CNNs hyper-parameter settings, in addition to the overall design of the CNNs architectures [12]. Lately, GA has shown great performance in term of generating CNNs architecture automatically to classify 2D images from the dataset of CIFAR10 with accuracy up to 96.78%, which is higher than most manually tuned models by deep learning experts [13]. By exploiting the GAs advantages, we were able to introduce the proposed algorithm to classify a 3D lung nodule images. Algorithm 1 demonstrates the proposed algorithm framework. In the beginning the algorithm starts by entering a predefined parameter such as: i) the basic building block of CNNs that we illustrated earlier; ii) the 3D cropped lung nodule images; iii) the size of the population; and iv) GA maximum generation number. The first population is generated randomly with different depth using the predefined parameters and GA operator, in which the GA operator generates: i) the number of basic building blocks of CNNs in each individual and this step determines the depth of each individual; ii) the value of feature map for convolutional layers in each individual; and iii) the individual models learning rate value. Regarding the building blocks value, the algorithm can choose from a set of numbers ranging from 1 to 10. In relation to feature map the algorithm can choose from {8, 10, 12, 24, 32, 48, 64}, as for the learning rate value the algorithm can choose from a set of numbers ranging from 0.00001 to 0.001. In Table 1 we demonstrate the parameters limit used will training the proposed algorithm, in which the GA operator can randomly select from, to generate different population in each generation. Now, the present generation's counter is set to zero. Then, the algorithm starts to traverse a series of evolutionary processes in order to obtain the optimum CNNs architecture. Throughout evolution, each individual's fitness is assessed, and the best fitness individuals are selected as parents. New offspring are generated from the parents through genetic mechanisms encompassing selection, crossover, and mutation processes.

After training the initial generation, each individual is subjected to validation and testing, allowing for a comparison of their final accuracies. Subsequently, a selection process is initiated, wherein the five best-performing models are chosen. Crossover and mutation operations are then applied to these selected models, giving rise to a new population of offspring models that will be trained, validated, and tested in the subsequent iterations.

Algorithm 1. The proposed algorithm

Inputs A set of predefined parameters; the basic building block of CNNs, the 3D cropped lung nodule images, the size of population, maximum number of generations.

Output The 3D CNN best architecture for classification.

$\mathbf{P}_0 \leftarrow$ Generate the initial population with different (block numbers, learning rate, and CNN feature map size).

$\mathbf{g} \leftarrow 0$

while $\mathbf{g}$ < maximum generation number:

In $\mathbf{P}_g$, train and evaluate each individual fitness.

Select the individuals with the highest accuracies.

$\mathbf{S}_g \leftarrow$ Apply the crossover and mutation operation, for offspring generation from the selected parent individuals.

$\mathbf{P}_g+1 \leftarrow$ selection from $\mathbf{P}_g \cup \mathbf{S}_g$

$g \leftarrow g +1$

end

Return The individual in P_g with the best fitness.

Table 1. The parameters limit used by the algorithm

Parameter	Minimum value	Maximum value
Basic building blocks	1	10
Learning rate	0.00001	0.001
Feature map	8	64

In the crossover processes the algorithm selects a random integer "n" as a crossover point from the set {0, 1, 2} corresponding to the number of basic building blocks of CNNs, the value of feature maps for convolutional layers, and the learning rate value respectively. It performs the crossover operation by swapping the genetic material between the selected parent chromosomes at the designated crossover point. The resulting offspring chromosomes are collected in a new list, and any duplicate chromosomes are removed. Once the number of offspring chromosomes in the list is equal to the crossover rate, the new offspring is returned as the output. In the mutation processes the algorithm selects a random integer "n" from the set {0, 1, 2} to use as the mutation point. Based on the mutation rate the algorithm performs the mutation operation on the offspring chromosomes by generating a random value at the mutation point. Upon the completion of these operations, the generation's counter is incremented by one, ensuring that the evolutionary process continues until the counter reaches the algorithm's maximum number of generations. The evolutionary process concludes at the end of the last generation, and the proposed algorithm presents the performance of each population categorization across all generations. Utilizing the test data, each individual's categorization performance is assessed, ultimately yielding the highest-performing model as the output.

3. RESULTS AND DISCUSSION

Our research is implemented utilizing Python code using Keras and TensorFlow as the backend [25]. We initiated the proposed algorithm with a population count and generation number both set to n=10. The maximum iteration for training each population in each generation is set to 300 epochs with an early stopping condition to stop training when the validation accuracy stopped improving. All generated models underwent training using Adam optimizer and L2 regularizer. The evaluation of the LUNA16 dataset was conducted using a 4-fold cross-validation approach. The dataset was divided into three subsets for training and validation and one subset for testing. Specifically, 80% of the data was allocated for training, while the remaining 20% was designated for testing. Within the training set, 20% was further set aside for validation purposes. Upon completion of the final generation, the evolutionary process was terminated. The proposed algorithm displayed the classification performance of each individual population across all generations. Performance evaluation for each individual was performed using the test data. Ultimately, the best-performing model was saved as the outcome of the algorithm.

The proposed algorithm achieved the ability to generate models with accuracies evaluated on the testing dataset of up to 95.977%. Figure 4 presents the average performance of all populations in each generation. It is evident from the blue line, which represents the average accuracy of each generation, that the algorithm's performance continually improves as the generations progress.

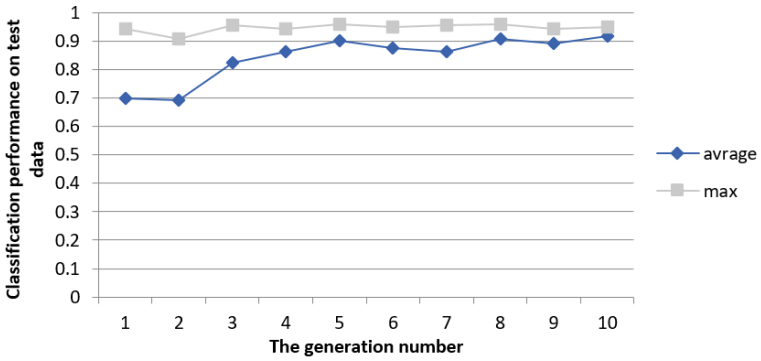


Figure 4. The average accuracy and top performance CNNs models for all populations in each generation during the evolutionary process

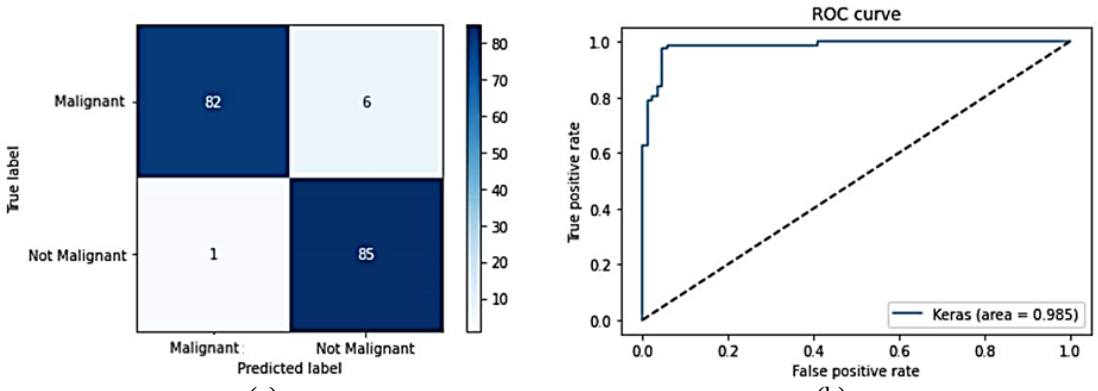
To assess the performance of the generated 3D CNN architectures, the confusion matrix [3], [11] was employed. Given the two classes of malignant and benign nodules, a 2×2 confusion matrix was used, as depicted in Figure 5(a). Furthermore, sensitivity, specificity, and precision were calculated using the confusion matrix, as shown in (2) to (4):

$$Sensitivity = \frac{TP}{TP+FN} \quad (2)$$

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

Where TP represents true positive, FN represents false negative, TN represents true negative, and FP represents false positive. The selected model had sensitivity equal to 98.8%, specificity equal 93.4% and precision equal 93.2%. The results of these measurements, which are provided here, demonstrate the validity of the suggested framework for categorizing data of a comparable modality since increasing the sensitivity, which represents the success or hit rate, while keeping the specificity, which signifies true negative rate, is a major goal of classification models [16]. We can see the receiver operating characteristic (ROC) curve of the network architecture with the best performance in Figure 5(b), with an AUC equal to 0.985. Figure 6 depicted the network architecture with the best performance. The model consists of four basic building blocks with convolutional layers equal to 4, 10, 20 and 20, respectively. Where, each convolution layer has a 3D filter of size 3×3×3 and feature map value equal 10. The model learning rate value is equal to 5.600000000000001e-4. The size of the model and parameters total number is 1 million and 176,211 thousand.



(a) (b)

Figure 5. The output CNN model classification performance; (a) the output CNN architecture confusion matrix and (b) the output CNN architecture ROC curve

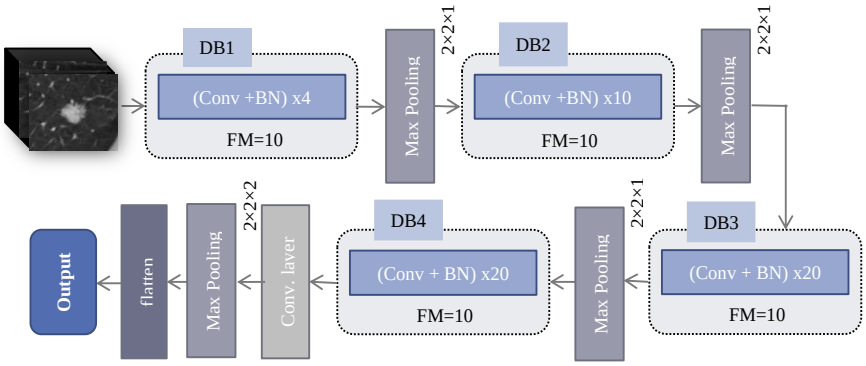


Figure 6. Architectures of the best performing network, where; DB: dense block; conv: convolutional layer; BN: batch normalization FM: feature map

Table 2 demonstrates comparison of our model with others such as the ones presented in [5], [8], [10]–[12], [14]–[16]. It is clearly shown that the proposed method performs comparably better in different performance criteria such as the model size and total number of parameters, and classification accuracy. Because we use fewer parameters, our model is extremely comparable to earlier models. It is evident that the proposed method outperforms other methods used with the same dataset and others with different dataset. However, caution should be exercised when drawing firm conclusions about the performance of this algorithm in comparison to others until evaluated on equal grounds.

Table 2. Comparisons of the previously mentioned state-of-the-art networks to the proposed algorithm

The classification systems	Dataset	TPR%	TNR%	ACC%	AUC	No. para. (M)
Deep learning-based CADx [5]	LIDC-IDRI	94.66	95.14	94.78	0.949	–
Fuzzy kernel C-means [8]	LUNGX SPIE AAPM	69.2	78.6	74.1	–	–
MoDenseNet [10]	LIDC-IDRI	90.47	90.33	90.40	0.9548	34.8
FPSOCNN [11]	LIDC-IDRI	97.93	96.32	95.62	–	–
CNN-GA [12]	LIDC-IDRI	–	–	91.3	–	–
BCNN [14]	LUNA16	91.85	92.27	91.99	0.959	_
NASLung [15]	LIDC-IDRI	85.37	95.04	90.77	–	16.84
GoogLeNet DNN [16]	IQ-OTH/NCCD	95.08	93.7	94.38	–	–
The proposed algorithm	LUNA16	98.8	93.4	95.977	0.985	1. 18

Where TPR., TNR, ACC., and No. para., denote true positive rate or sensitivity, true negative rate or specificity, accuracy, and the number of parameters, respectively

4. CONCLUSION

In this study, the objective is to introduce an automated algorithm for 3D CNNs architecture design based on GA to classify whether pulmonary nodules are malignant or benign and helping users with no expertise in CNNs to find the best 3D CNN architecture to classify pulmonary nodules. The goal was achieved successfully by developing a cutting-edge approach which exploits the GA evolutionary search to encode and design 3D CNNs with random depths. Using the dataset of LUNA16 the proposed algorithm was evaluated. In comparison with other models the proposed algorithm showed a cutting-edge performance regarding the model number of parameters and the classification accuracy without the need of expert knowledge to extract the proper features or to build the CNN architecture. In the future, the proposed algorithm may be deployed in various clinical classification applications.

REFERENCES

[1] Global Cancer Observatory, "Global cancer observatory," *Malaysia Cancer Statistics*, vol. 593. pp. 1–2, 2020, [Online]. Available: https://gco.iarc.who.int/media/globocan/factsheets/populations/458-malaysia-fact-sheet.pdf, date accessed: Nov. 08, 2022.

[2] H. Jung, B. Kim, I. Lee, J. Lee, and J. Kang, "Classification of lung nodules in CT scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method," *BMC Medical Imaging*, vol. 18, no. 1, pp. 48–58, 2018, doi: 10.1186/s12880-018-0286-0.

[3] K. H. Rahouma, S. M. Mabrouk, and M. Aouf, "Lung cancer diagnosis based on chan-vese active contour and polynomial neural network," *Procedia Computer Science*, vol. 194, pp. 22–31, 2021, doi: 10.1016/j.procs.2021.10.056.

[4] National Cancer Institute, "Non-small cell lung cancer treatment (PDQ®)–patient version-national cancer institute," *Nih*, 2021. https://www.cancer.gov/types/lung/patient/non-small-cell-lung-treatment-pdq, date accessed: Nov. 09, 2022.
[5] G. L. F. da Silva, O. P. da S. Neto, A. C. Silva, A. C. de Paiva, and M. Gattass, "Lung nodules diagnosis based on evolutionary convolutional neural network," *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 19039–19055, 2017, doi: 10.1007/s11042-017-4480-9.
[6] S. Sari, T. Sutikno, I. Soesanti, and N. A. Setiawan, "A review of convolutional neural network-based computer-aided lung nodule detection system," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 3, pp. 1044–1061, 2023, doi: 10.11591/ijai.v12.i3.pp1044-1061.
[7] O. R. Kadhim, H. J. Motlak, and K. K. Abdalla, "Computer-aided diagnostic system kinds and pulmonary nodule detection efficacy," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 5, pp. 4734–4745, 2022, doi: 10.11591/ijece.v12i5.pp4734-4745.
[8] Z. Rustam, A. Purwanto, S. Hartini, and G. S. Saragih, "Lung cancer classification using fuzzy C-means and fuzzy Kernel C-means based on CT scan image," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 2, pp. 291–297, Jun. 2021, doi: 10.11591/ijai.v10.i2.pp291-297.
[9] A. AbuBaker and Y. Ghadi, "A novel CAD system to automatically detect cancerous lung nodules using wavelet transform and SVM," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 5, pp. 4745–4751, 2020, doi: 10.11591/ijece.v10i5.pp4745-4751.
[10] R. Dey, Z. Lu, and Y. Hong, "Diagnostic classification of lung nodules using 3D neural networks," *Proceedings-International Symposium on Biomedical Imaging*, vol. 2018-April, pp. 774–778, 2018, doi: 10.1109/ISBI.2018.8363687.
[11] A. Asuntha and A. Srinivasan, "Deep learning for lung Cancer detection and classification," *Multimedia Tools and Applications*, vol. 79, no. 11–12, pp. 7731–7762, 2020, doi: 10.1007/s11042-019-08394-3.
[12] M. A. Pfeffer and S. H. Ling, "Evolving optimised convolutional neural networks for lung cancer classification," *Signals*, vol. 3, no. 2, pp. 284–295, 2022, doi: 10.3390/signals3020018.
[13] Y. Sun, B. Xue, M. Zhang, G. G. Yen, and J. Lv, "Automatically designing CNN architectures using the genetic algorithm for image classification," *IEEE Transactions on Cybernetics*, vol. 50, no. 9, pp. 3840–3854, 2020, doi: 10.1109/TCYB.2020.2983860.
[14] R. Mastouri, N. Khlifa, H. Neji, and S. Hantous-Zannad, "A bilinear convolutional neural network for lung nodules classification on CT images," *International Journal of Computer Assisted Radiology and Surgery*, vol. 16, no. 1, pp. 91–101, 2021, doi: 10.1007/s11548-020-02283-z.
[15] H. Jiang, F. Shen, F. Gao, and W. Han, "Learning efficient, explainable and discriminative representations for pulmonary nodules classification," *Pattern Recognition*, vol. 113, 2021, doi: 10.1016/j.patcog.2021.107825.
[16] M. S. AL-Huseiny and A. S. Sajit, "Transfer learning with GoogLeNet for detection of lung cancer," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 2, pp. 1078–1086, 2021, doi: 10.11591/ijeecs.v22.i2.pp1078-1086.
[17] T. M. Breuel, "The effects of hyperparameters on SGD training of neural networks," *arXiv*, 2015, doi: 10.48550/arXiv.1508.02788.
[18] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," *International Conference on Information Processing in Medical Imaging IPMI 2015: Information Processing in Medical Imaging*, 2015, vol. 9123, pp. 588–599, doi: 10.1007/978-3-319-19992-4_46.
[19] "Lung nodule analysis, grand challenge (lUNA16)." [Online]. Available: https://luna16.grand-challenge.org/, date accessed: Nov. 08, 2022.
[20] B. Vendt, "The lung image database consortium (LIDC) and image database resource initiative (IDRI)," *The Cancer Imaging Archive (TCIA) Public Access*. [Online]. Available: https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=1966254, date accessed: Nov. 08, 2022.
[21] K. Clark *et al*., "The cancer imaging archive (TCIA): Maintaining and operating a public information repository," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1045–1057, 2013, doi: 10.1007/s10278-013-9622-7.
[22] D. Riquelme and M. A. Akhloufi, "Deep learning for lung cancer nodules detection and classification in CT scans," *AI (Switzerland)*, vol. 1, no. 1, pp. 28–67, 2020, doi: 10.3390/ai1010003.
[23] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings-30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Jul. 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.
[24] A. E. Eiben, Raué, and Z. Ruttkay, "Genetic algorithms with multi-parent recombination Parallel Problem Solving from Nature-PPSN III," in *Parallel Problem Solving from Nature-PPSN III*, 1994, vol. 866, pp. 78–87.
[25] F. Chollet, "Keras: deep learning for humans," *Github*. 2015, [Online]. Available: https://github.com/keras-team/keras, date accessed: Nov. 21, 2022.

BIOGRAPHIES OF AUTHORS

Kamel Hussein Rahouma is a Professor at the Department of Electrical and Electronic Engineering, Minia University (Egypt). He earned his B.Sc. and M.Sc. in Communications and Electronic Engineering from the Faculty of Engineering, Cairo University in June 1984 and March 1988 respectively. He earned his first doctoral degree in Communications and Electronics Engineering jointly from the University of Kent at Canterbury (England) and Minia University (Egypt) in 1996. In May 2001, he earned his second doctoral degree in Computer Science from the University of Salzburg, Austria. His interests are multidisciplinary. This includes (but not limited to): cryptography and information security, artificial intelligence applications, biomedical application, and bioinformatics. He can be contacted at email: kamel_rahouma@yahoo.com.

Shahenda Mahmoud Mabrouk is a Teacher Assistant in the Department of Biomedical Engineering, Higher Technological Institute-10th of Ramadan city, Egypt, where she has been a faculty member since 2008. Shahenda graduated with a first-class honours' degree in Biomedical Engineering, from Higher Technological Institute, Egypt in 2007. She earned her M.S. Degree in system and biomedical engineering from the Faculty of Engineering, Cairo University, in 2015. she is working on receiving her Ph.D. in biomedical engineering from the Faculty of Engineering, Minia University, Egypt. Her current research interests include biomedical application, deep learning, and artificial intelligence applications. She can be contacted at email: shahenda-mahmoud@hotmail.com and shahenda.mabrouk@hti.edu.eg.

Mohamed Aouf is an Associate Professor at the Department of Biomedical Engineering, Higher Technological Institute, 10th of Ramadan City, Egypt, where he has been a faculty member since 2002. He received his B.Sc. in Biomedical Engineering, Higher Technological Institute, 10th of Ramadan City, Egypt in 2001. He earned his M.Sc. and Ph.D. in Systems and Biomedical Engineering from the Faculty of Engineering, Cairo University in 2010 and 2015 respectively. Currently he is the head of the biomedical engineering department in Higher Technological Institute. His interests are multidisciplinary. This includes (but not limited to): signal processing, wavelet implementation, artificial intelligence applications, and biomedical application. He can be contacted at email: maoufmedical@yahoo.com.