Journal of Theoretical and Applied Information Technology
31st May 2023. Vol.101. No 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

INNOVATIVE METHOD OF CLASSIFICATION OF PULMONARY NODULES USING 3D CNN ARCHITECTURE

KISHORE SEBASTIAN[1], S. DEVI[2]

[1]Research Scholar, Department of Computer Science and Engineering, PRIST Deemed to be University, Thanjavur 613403, India
[2]Department of Electronics and Communication Engineering, Sree Venkateswara College of Engineering, Nellore, Andhra Pradesh-52316, India

E-mail: sebastiankishore9@gmail.com

ABSTRACT

Pulmonary nodules are spots or lesions that are diagnosed in the lungs and can have both benign and malignant causes, mainly related to lung cancer. According to statistics, lung cancer is in seventh place and is the most lethal. Various international medical institutions are working to improve the diagnosis of lung cancer, since the main cause of death is late diagnosis. For this reason, the analysis of pulmonary nodules is a challenge in the processing of medical images to determine the appropriate treatment, new methodologies and techniques are proposed not only by experts in medicine but also in by other scientists. In this work 3D Convolutional Neural Network is used to classify pulmonary nodules in CT images. The proposed architecture results, with an accuracy of 0.9076, kappa of 0.7773, sensitivity of 0.8483, and specificity of 0.9321 and AUC of 0.8903, in the classification of pulmonary nodules.

Keywords: *Lung Cancer, Lung Nodule Classification, Machine Learning*

1. INTRODUCTION

At the end of the 20th century, lung cancer became one of the most preventable causes of death as long as its detection is carried out early. In India, lung cancer is the second most common and the first worldwide. The computed tomography (CT) examination is, according to [1], the method preferred by specialists to perform non-invasive screening of patients with the disease.

Lung cancer is a global health problem. Approximately eight thousand people die from it in Mexico each year, with late diagnosis being the main problem [2]. In 99% of cases cancer is detected in stage III or IV, there are few initiatives to detect it earlier [3]. Early detection can prolong survival and make it easier for people to recover. In addition, the exact diagnosis is key to adequate treatment, so it is necessary to implement timely diagnostic measures.

Medical images are very important, as they help make diagnoses, determine treatment and assess changes that occur over time [4], examples of which are radiography, CT (CT) and ultrasound. CT is a radiological study that uses X-rays to generate multiple cross-sectional images of the interior of the body. For the detection of lung cancer, a radiologist interprets a CT of the chest, through his experience and medical knowledge, said interpretation is received by the onc radiologist to issue the lung cancer treatment, and however, viewing multiple scans can tire the radiologist and lead to errors.

By means of automatic learning, patterns can be detected and predicted [5], it has been shown to be effective in applications in the medical field, such as: early detection of diseases, surgical planning, an analysis of pathologies, oncology and radiology. Machine learning is one of the disciplines of artificial intelligence, which uses statistics, probability, and optimization techniques to understand complicated rules and high-dimensional features; in the subject of lung cancer, it has been applied to detect it [2], classify the severity of the cancer [3], predict patient survival [4], classify lung cancer [5] and classify pulmonary nodules [6–14].

The analysis of CT images is still a challenging task, as the density of nodules can be similar to other lung structures and, in addition, specialists have a large number of images to analyze. in a repetitive process. A pulmonary nodule is

defined as a single, spherical radiological opacity, up to 3 cm in diameter, circulated through the lung parenchyma. Lesions larger than 3 cm are often classified as malignant [12]. For the extraction of features from the nodules and their classification as malignant or benign, we propose a CNN 3D; finally, to evaluate the results, we used the metrics accuracy, kappa index, sensitivity, specificity and area under the ROC curve (AUC).

2. LITERATURE REVIEW

The authors [6] use the multiview neural network architecture, which takes multiple views of each node entered, allowing binary classification of nodules (benign and malignant). To validate their architecture they use the LIDC-IDRI dataset.

The authors [7] presents the Nodular-Deep system that allows classifying pulmonary nodules, contemplating three stages: (i) feature extraction using a convolutional neural network (ii) feature optimization using a neural network recurrent, which is a supervised learning technique (iii) classification of nodules optimized through the softmax linear classifier. Due to the presented architecture, Nodular-Deep requires no pre-processing or post-processing. To validate its results, it uses the LIDC-IDRI dataset.

The authors [1] propose to use a convolutional neural network to classify pulmonary nodules, likewise, they implement three neural networks: (i) 2D at the division level (ii) 2D at the n level nodules (iii) 3D at nodule level. To compare the neural networks, they use the LIDC-IDRI dataset, obtaining better results in the 3D network at the node level.

The authors [8] show an automatic diagnosis system, which consists of four stages (i) preprocessing, in which they use a fuzzy filter (ii) segmentation using the Watershed algorithm (iii) feature extraction using a gray level co-occurrence matrix (iv) classification using support vector machines.

The authors [9] describe the development of a computer-aided diagnostic system to classify large nodules. The methodology they use is (i) extraction of lung regions by thresholding (ii) extraction of the most relevant features (iii) selection of features, at this stage they choose the most relevant features important in differentiating between normal and cancerous tissues (iv) classification of normal and cancerous tissues using support vector machine with different kernels. They use a dataset of 12 CT scans retrieved from Cornell University.

The authors [10] classify pulmonary nodules, the methodology they use is (i) segmentation through morphological operations and the OTSU method (ii) they extract statistical parameters from the regions segmented images of the lungs (iii) use neural networks for the classification of cancerous and non-cancerous nodules.

The authors [11] apply the focal loss function in the training process to increase the classification accuracy of the machine learning model. To validate their results, they use the LUNA16 dataset.

It is worth noting that one of the main challenges encountered in this work was the need to balance the set of images used for training the 3D CNN, as the amount of segmented benign nodules was much greater than the amount of malignant nodules. Thus, data augmentation techniques were used in order to improve the efficiency of the proposed algorithm.

3. MATERIALS AND METHODS

The proposed methodology is divided into four stages: image acquisition, segmentation, feature extraction, classification and evaluation of results. A summary of the procedures that are part of this methodology is presented in Figure 1.

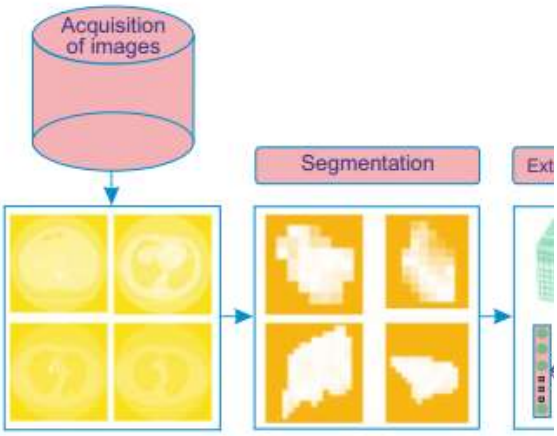

Figure 1. Flowchart Of The This Work

3.1. Acquisition of images

The LIDC-IDRI image collection consists of the diagnosis and screening of lung cancer in CTs with marked lesions. The set contains 1018 patterns [12]. Figure 2 shows five slices from the LIDC-IDRI-0001 patient pattern, where the average number of slices per pattern in this database is 248.

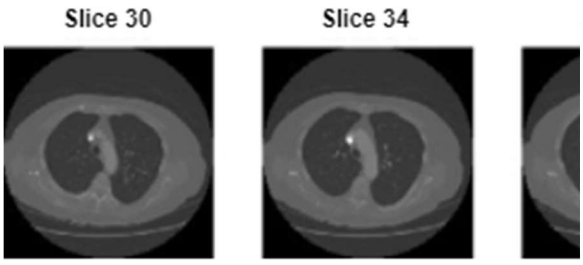

Figure 2. Slices Of The Patient Pattern "LIDC-IDRI-0001".

3.2. Segmentation

In order to find the region of interest in the images, in the work of [13] the authors performed a segmentation of the nodules' regions. In this procedure, an algorithm based on quality threshold clustering was used. 800 patterns were selected from the LIDC-IDRI database and the threshold of 90 was used for the entire methodology. After image segmentation, 1,011 benign and 394 malignant nodules were obtained. We emphasize that in the step of extracting features and classifying the nodules, the segmented regions proposed by [14].

We observe that the number of examples of each one of the classes is well unbalanced, that is, we have more examples of the benign class in relation to the malignant class. This effect can produce bad results, mainly for the detection of elements of the class with a smaller number of examples, which is the most important in our problem. Therefore, we used data augmentation techniques (using rotation and application of random noise) during the training of the 3D CNN used in the next step.

3.3. Extraction/Classification

For feature extraction and nodule classification, we used a 3D CNN with two blocks of convolutional layers. Each block containing two convolutional layers and followed by a pooling layer. Then there are two fully connected layers and a softmax layer that calculates the network output. Figure 3 presents a complete diagram of the 3D CNN architecture used in this work.

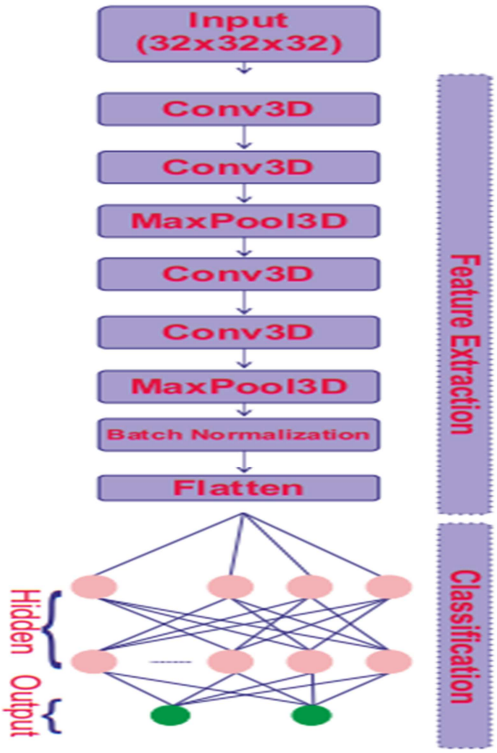

Figure 3. 3D CNN Architecture For Classifying Pulmonary Nodules

Table 1 shows the main parameters of the architecture used. The training was carried out for 2000 epochs, with a learning rate of 0.0001, dropout of 0.4 and batch size of 128.

Table 1. Parameters Of The Used Architecture.

Parameters	Corresponding value
Input layer size	(32, 32, 32)
Number of convolutional layers	4
Number of filters per layer	(8, 16, 32 e 64)
Activation function	relu
Size of convolutional kernels	(3, 3, 3)
Size of pooling kernels	(2, 2, 2)
Regularizer(L2) kernel value	0.0001
Fully connected layers	2 (256 and 128 neurons)
Softmax	2 neurons

3.4. Evaluation of results

To evaluate the results, the following metrics were used: Accuracy, Kappa coefficient, Sensitivity, Specificity and Area Under Curve (AUC). Before understanding each of these, it is necessary to introduce some elementary concepts used in their compositions, these are shown in Table 2.

Table 2. Metrics, With The Positive Class Being Benign And The Negative Class Being Malignant.

Term	True Corresponding Value
True Positives (TP)	Correct classification of Positive class.
True Negatives (TN)	Correct Class Negative Classification.
False Positives (FP)	Error that occurs because the algorithm predicted Positive result when expected would be Negative.
False Negatives (FN)	Error that occurs because the algorithm predicted a Negative result when the expected result would be Positive.

Accuracy (Equation 1), considered by several researchers to be one of the simplest metrics to assess the quality of a classification, was used in this work to measure the number of predictions correct actions regardless of positives and negatives.

$$\text{Accuracy (AC)} = (T\,P + T\,N)/(T\,P + T\,N + FP + FN) \quad (1)$$

Sensitivity (Equation 2), or rate of true positives, is the metric computed by the relationship between true positives and all positive cases.

$$\text{Sensitivity (SE)} = T\,P/(T\,P + FN) \quad (2)$$

Specificity (Equation 3), or false positive rate, corresponds the ratio of false positives to all other negative data.

$$\text{Specificity (S)} = T\,N/(T\,N + FP) \quad (3)$$

The Kappa index (Equation 4), is a very useful metric when working with problems of several classes or when the classes of the problem are unbalanced. Although there is no standard for interpreting their data, [15] present a way to achieve understanding, which makes it possible

to verify the degree of agreement between evaluators, where: a value less than 0 indicates no agreement, between 0 and 0.20 mild agreement, between 0.21 and 0.40 as fair agreement, between 0.41 and 0.60 as moderate agreement, between 0 .61 and 0.80 as substantial agreement and between 0.81 and 1 as almost perfect agreement.

$$Kappa(K) = \frac{P_0 - P_e}{1 - P_e} \quad (4)$$

Where,

$$P_0 = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$P_e = \frac{[(TP+F\ )(TP+FP)]+[(TN+\ \ )(TN+FP)}{(TP+TN+FP+F\ \ )^2} \quad (6)$$

Finally, Area under ROC curve is used when the classification is a binary solution problem, as is the case in this work. In practice, it corresponds to the probability that a positive example chosen randomly is higher than that of a negative example being chosen also randomly, and demonstrates the contrast in prediction.

4. RESULTS AND DISCUSSION

The segmented nodules were divided into three sets: training, validation and testing, containing 70%, 10% and 20% respectively of each class (benign and malignant). We perform a normalization of the data and the image intensities form scales between 0 and 1. The nodules that had only one slice were discarded, in the end we obtained 956 benign and 389 malignant. Of these, 939 were used for training (668 benign and 271 malignant), 135 for validation (96 benign and 39 malignant) and 271 for testing (192 benign and 79 malignant).

To evaluate the effect of data augmentation using the different scenarios in Table 3, we used these techniques to generate four different training sets: the first training set was all reconstructed nodes (no data augmentation), 668 benign and 271 malignant; Second, we duplicate (make a copy of) the malignant nodules, obtaining a total of 542 malignant nodules, leaving 668 benign nodules; For the third set, we randomly selected 668 malignant nodules by rotating the malignant nodules (45, 90, 135 and 180 degrees) using random noises. This set contains the same number of instances of both classes. Finally, we doubled the number of benign examples using rotation (45, 90, 135, and 180 degrees) and random noise to create a fourth set. Next, we apply data multiplication to the malignant class and using the same operations, generate 1336 of this class as well. Note that data augmentation was applied to the training set only. In addition to the discussed techniques, other possibilities can be used: differences in color space, geometric changes, and others .

Table 3 shows the results obtained with the application of the proposed method and in Figure 4 the results of the confusion matrix are presented. The metrics used: Accuracy (AC), Kappa (K), Sensitivity (SE), Specificity (S) and AUC, were presented in section 3.4. The results presented are for the test set, containing 192 benign and 79 malignant nodules.

Table 3. Results Obtained After The Application Of The Proposed Method.

Experiment	Training Images	AC K SE	S	AUC
EX1 - Segmented Images	939	0.9003 .7465 .7468	**0.9635** 0 0	0.8551
EX2 - Replicated images (malignant)	1210	0.9003 .7578 .8227	0.9322 0 0	0.8775
EX3 - Rotation and noise (malignant)	1336	**0.9077** **.7774** **.8481**	0.9322 **0** **0**	**0.8901**
EX4 - Rotation and noise (both)	2672	0.8966 .7441 .7848	0.9427 0 0	0.8637

	Experiment EX1	PREDICTED	
		Malignant	Benign
REAL	Malignant	58 TP	8 FP
	Benign	22 FN	183 TN
	Experiment EX2	Malignant	Benign
	Malignant	67 TP	11 FP
	Benign	15 FN	178 TN
	Experiment EX3	Malignant	Benign
	Malignant	68 TP	12 FP
	Benign	11 FN	180 TN
	Experiment EX4	Malignant	Benign
	Malignant	61 TP	12 FP
	Benign	16 FN	182 TN

Figure 4. Confusion Matrix For Four Experiments, Where We Present The Actual Values And Predicted Values For Each Class.

Analyzing the EX1 experiment, we can observe that the sensitivity value was lower than in the other experiments, on the other hand, the specificity value was higher. These results indicate that more examples of the benign class are being correctly classified, however, more false negatives are being obtained. This is because the EX1 training set has much more examples of the benign class, and this influences the training of the network, which tends to identify fewer malignant nodules.

One of the key measures for analyzing unbalanced data is the kappa index, as it gives better results when the prediction of the two classes is balanced. In the EX2 experiment, we can observe that the kappa value is increased compared to EX1 and EX4. This happened due to an increase in the number of examples of the malignant class in training, which reduced the number of false negatives generated by the network.

In experiment EX3, we performed rotation and noise addition operations on the malignant class examples. As a result, there was an improvement in the generalization capacity of the network due to the increase in the diversity of the nodes present in the training set. Thus, this experiment produced the lowest number of false negatives, which is reflected in the Kappa and Sensitivity index, which obtained the highest values among the tests performed. In addition, we also obtained the best accuracy and AUC with experiment EX3, and only lost in terms of specificity for experiments EX1 and EX4.

In experiment EX4, even increasing the number of data from both classes, the result did not improve. As the examples of the malignant class were increased by almost 5 times (from 271 to 1336), many images from this set had similar characteristics, which influenced the result obtained in this experiment, as the network did not obtain a greater capacity. of generalization. Comparing the EX4 experiment with the EX1 experiment, we obtained a better result in terms of Sensitivity and AUC, however the metrics Accuracy, Kappa and Specificity were inferior to the other experiments.

Therefore, the best result was produced by the EX3 set, in which the number of malignant nodules were increased by means of rotation operations and addition of random noise. rio until the training set had the same number of examples from both classes. Table 4 provides a comparison of the best result obtained by the proposed methodology with works present in the literature that perform the classification of pulmonary nodules into benign and malignant. It is important to mention that the results presented in each work were obtained with different evaluation criteria and for the number of different LIDC-IDRI images.

Table 4. Comparison Of The Results With Other Works

Work	AC	SE	S	AUC
.	0.9110	0.6978	0.9854	0.9695
.	0.9494	0.9461	0.9521	-
.	0.9318	0.9274	0.9334	-
[1]	0.873	0.8943	0.8523	0.9472
.	0.8842	-	-	0.9318
.	0.8412	0.9168	0.7315	-
.	0.8221	-	-	0.8773
.	0.9041	0.9046	0.9032	0.9547
Our work	**0.9078**	**0.8482**	**0.9321**	**0.8903**

5. CONCLUSION

In this work, the 3D CNN algorithm to classify pulmonary nodules can be used for the construction of CAD systems that help specialists in the interpretation of the pattern, providing a second opinion during the analysis of the same. The proposed methodology presents promising results, mainly with the use of data augmentation techniques. We observed that when balancing the classes in the training set, there was a significant increase in the prediction of examples of the malignant class, which contains the least amount of examples. The best results were achieved by balancing the training set by increasing the number of malignant nodules with rotation operations and adding random noise. In this work, we perform rotation and noise application operations at slice level (2D).

DECLARATION:

Ethics Approval and Consent to Participate:

No participation of humans takes place in this implementation process

Human and Animal Rights:

No violation of Human and Animal Rights is involved.

Funding:

No funding is involved in this work.

Data availability statement:

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study

Conflict of Interest:

Conflict of Interest is not applicable in this work.

Authorship contributions:

All authors are contributed equally to this work

Acknowledgement:

There is no acknowledgement involved in this work.

REFERENCES

[1]. X.Yan, J. Pang, H. Qi, Y.Zhu, C.Bai, X. Geng and M. Liu," Classification of lung nodule malignancy risk on computed tomography images using convolutional neural network: A comparison between 2d and 3d strate- gies," In *Asian Conference on Computer Vision*, 2016,pp.91–101.

[2]. L.Zhuo ,Y. Chenhui ,Y.Hang and W. Taihua ," Deep Reinforcement Learning with Its Application for Lung Cancer Detection in Medical," *Internet of Things*, vol. 97,2019,pp.1–9. https://doi.org/10.1016/j.future.2019.02.068

[3]. S.Bergquist, G. Brooks, N. Keating, M. Landrum and S. Rose, "Classifying Lung Cancer Severity with Ensemble Machine Learning in Health Care Claims Data," *In:*

2nd Machine Learning for Healthcare Conference,2017, pp. 25–38 .

[4]. D.N.V.S.L.S. Indira, R.K., Ganiya, P. Ashok Babu, A. Xavier, L. Kavisankar, S. Hemalatha, V. Senthilkumar, T. Kavitha, A. Rajaram, K. Annam and A.Yeshitla, " Improved artificial neural network with state order dataset estimation for brain cancer cell diagnosis." *BioMed Research International*, *2022*. https://doi.org/10.1155/2022/7799812

[5]. K.Pradeep, N. Naveen, " Lung Cancer Survivability Prediction based on Performance Using Classification Techniques of Support Vector Machines, C4.5 and Naive Bayes Algorithms for Healthcare Analytics," *Proceeded Computer Science*, vol. 132,2018, pp. 412–420 https://doi.org/10.1016/j.procs.2018.05.162

[6]. A.Barchuck, " Evaluation of Machine Learning Algorithm Utilization for Lung Cancer Classification Based on Gene Expression Levels," *Asian Pac J Cancer Prev. vol.*17,2016, 835-838. https://doi.org/10.7314/APJCP.2016.17.2.835

[7]. A.Rajaram and K.Vinothkumar, " Efficient Multipath Location Aware Routing Protocol for Mobile Ad Hoc Networks," *Journal of Theoretical and Applied Information Technology, Scopus Indexed*, Vol.*59,2019,* pp.130-38.

[8]. K.Liu andG. Kang, " Multiview Convolutional Neural Networks for Lung Nodule Classification,"*International Journal of Imaging Systems and Technology* ,vol.27,2017, pp. 12–22 .

[9]. Q.Abbas, " Nodular-Deep: Classification of Pulmonary Nodules using Deep Neural Network." *International Journal of Medical Research & Health Sciences* ,vol.6,2017, pp. 111-–118 (2017)

[10]. C.R.Rathish and A.Rajaram, " A New Load Balanced Routing Protocol for Wireless Sensor Networks,"2016.

[11]. C.R.Rathish and A.Rajaram, " Hierarchical Load Balanced Multipath Routing Protocol for Wireless Sensor Networks," 10,2015,pp.16521-16534.

[12]. P.Ashok Babu, L. Kavisankar, J. Xavier, V. Senthilkumar, G. Kumar, T. Kavitha, A. Rajendran, G. Harikrishnan, A. Rajaram and A.G.Adigo, "Selfish Node Detection for Effective Data Transmission Using Modified Incentive Sorted Pathway Selection in Wireless Sensor Networks,"*Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2022/9359135

[13]. N.Shukla, A. Narayane, A. Nigade, K. Yadav and H. Mhaske, " Lung cancer detection and classification using Support Vector Machine," *International Journal of Engineering and Computer Science* vol.4,2015,pp. 14983—14986 .

[14]. M.Mabrouk, A. Karrar and A. Sharawy, " Support Vector Machine Based Computer Aided Diagnosis System for Large Lung Nodules Classification," *Journal of Medical Imaging and Health Informatics*, vol. 3,2013, pp. 214–220, https://doi.org/10.1166/jmihi.2013.1154 .

[15]. J.Kuruvilla and K. Gunavathi, " Lung cancer classification using neural networks for CT images," *Computer Methods and Programs in Biomedicine, vol.*113,2014,pp. 202-209. https://doi.org/0.1016/j.cmpb.2013.10.011 .