Development of a Convolutional Neural Network for detection of Lung Cancer based on Computed Tomography Images

1st Gabriela Narvaez
School of Biological Sciences and Engineering, Universidad Yachay Tech
Urcuquí, Ecuador
gabriela.narvaez@yachaytech.edu.ec

2nd Andrés Tirado-Espín
School of Mathematical and Computational Sciences, Universidad Yachay Tech
Urcuquí, Ecuador
Universidad de Otavalo
Otavalo, Ecuador
ctirado@yachaytech.edu.ec

3rd Carolina Cadena-Morejon
School of Mathematical and Computational Sciences, Universidad Yachay Tech
Urcuquí, Ecuador
ccadena@yachaytech.edu.ec

4th Fernando Villalba-Meneses
School of Biological Sciences and Engineering, Universidad Yachay Tech
Urcuquí, Ecuador
gvillalba@yachaytech.edu.ec

5th Jonathan Cruz-Varela
School of Biological Sciences and Engineering, Universidad Yachay Tech
Urcuquí, Ecuador
jcruz@yachaytech.edu.ec

6th Gabriela Villavicencio Gordón
Language Departament, Universidad Yachay Tech
Urcuquí, Ecuador
lvillavicencio@yachaytech.edu.ec

7th César Guevara
Centro de Mecatrónica y Sistemas Interactivos, Universidad Indoamérica
Quito, Ecuador
Instituto de Ciencias Matemáticas, ICMAT-CSIC
Madrid, España
cesarguevara@uti.edu.ec

8th Omar Alvarado-Cando
Escuela de Electrónica, Universidad de Azuay
Cuenca, Ecuador
oalvarado@azuay.edu.ec

9th Diego Almeida-Galárraga
School of Biological Sciences and Engineering, Universidad Yachay Tech
Urcuquí, Ecuador
Universidad de Otavalo
Otavalo, Ecuador
dalmeida@yachaytech.edu.ec

***Abstract*—Lung cancer is a disease that generates an impact worldwide as it is one of the cancers with the highest incidence and mortality among all types of cancer. Lung cancer usually does not show symptoms until it is in an advanced stage of the cancer, lowering the survival rate of those affected. At the national level, lung cancer is recognized as the second type of cancer with the highest incidence and the first with the highest mortality, so an early and effective diagnosis can increase the survival rate. Currently there are many techniques for the diagnosis of cancer in early stages such as Computed Tomography (CT). CT scans are widely used due to their high sensitivity in the detection of pulmonary nodules without the need to be invasive, which avoids physical damage to the tissues. In addition, there has currently been a great growth in advanced computational techniques that, together with CT, can extract useful characteristics that support radiologists in the detection process. One of the most used techniques for this purpose are neural networks, which are designed to identify patterns and classify image objects. For this reason, in this project the methodology presented consists of training two convolutional neural networks with the VGG16 and DenseNet networks, which have been used in other projects with good results. Both methods were trained with the LIDC-IDRI database, which has a large number of CT scans of patients diagnosed with lung cancer. An evaluation of the models was carried out based on the metrics of accuracy, precision, sensitivity and specificity where the values of 0.992, 0.859, 0.952 and 0.994 were obtained for the VGG16 network and for the DenseNet network the values 0.995, 0.909, 0.959 and 0.996 were obtained respectively. Demonstrating the relevance of this technique in the medical field, serving as support for radiologists and later serving as an alternative for lung cancer detection.**

Keywords—Computed-aided diagnostic system, Neural networks, VGG16, DenseNet, Computed tomography, LIDC-IDRI

I. INTRODUCTION

Cancer is a group of diseases which is characterized by uncontrolled cell growth, where this type of cells ignores the signals that normally induce apoptosis, also cancer cells are characterized by cell invasion and dissemination from the primary site to others body sites [1]. Cancer has evolved through the mechanism of genomic alteration of cells, so it is a disease of the genome at the cellular level [2]. Because cancer is a group of diseases, it has been possible to classify at least 100 types of cancer where the distinctive characteristics of the type of cancer are given by the tissue of origin. Approximately 85% of cancers occur in epithelial cells and are referred to as carcinomas. On the other hand, cancers in mesodermal sites, such as bone or muscle, are referred to as sarcomas while cancers of glandular tissue are referred to as adenocarcinomas [1].

According to the World Cancer Research Fund International [3], by 2020 approximately 18.1 million new cases were diagnosed worldwide, where 9.3 million of the total cases were men and 8.8 million were diagnosed cases in women. In addition, the most common cancers being breast, lung, and colorectal cancers, representing 12.5%, 12.2%, and 10.7% of the total number of new cases diagnosed for that year, respectively. Fig.1 represents the most common cancers worldwide in more detail where Lung cancer is defined as the second type of cancer with the newest cases worldwide and the leading cancer in terms of cancer deaths worldwide [4]. Lung cancer resulted in more deaths than breast, prostate, colorectal, and leukemia cancers in men over 40 years of age and women over 60 years of age [5,6].

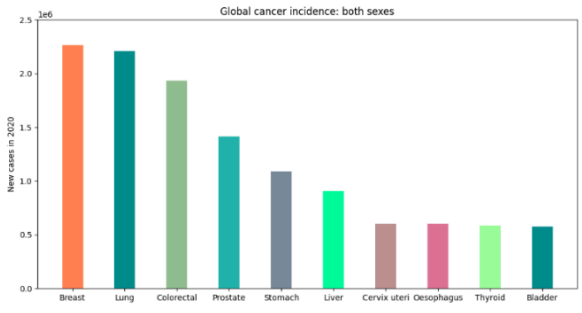

Fig. 1. Top 10 most common types of cancer worldwide.

A. *Lung Cancer in Ecuador*

Ecuador has a population of 17.2 million people according to the National Institute of Statistics and Censuses (INEC) [8] regarding the last census carried out in 2020. In addition, the association in charge of keeping statistics on cancer in this country is the Society Fight Cancer (SOLCA) [10] by its acronym in Spanish. Lung cancer in Ecuador, unlike world statistics, is in seventh place among the most common types of cancer at the national level in terms of men, while for women it is in eighth place with 6.3 cases and 5.7 cases per 100,000 inhabitants, respectively.

By 2020, according to INEC statistics, Quito had the largest population in Ecuador with 2.7 million inhabitants. For this reason, different statistical studies of the country are usually carried out in this city. Thus, according to the study carried out by SOLCA, it has been determined that the incidence of cancer in Quito is higher than the national average.

Therefore, in Quito, an average of 120 new cases are diagnosed each year and 1,100 new cases nationwide [11]. These diagnosed cases show the high incidence that exists in Quito since to reach the estimated number of new cases for the national level, each province should have an average of 45 new diagnostic cases.

In addition, approximately 79% of the diagnosed cases of lung cancer in Quito were diagnosed at stage IV (Fig. 2) when treatment alternatives are limited. Therefore, survival is very low, making mortality high, with values similar to those of diagnosed cases [10]. At least in Ecuador, early detection of lung cancer is difficult because people with smoking cases often do not have a culture of prevention and routine controls are not carried out [11].

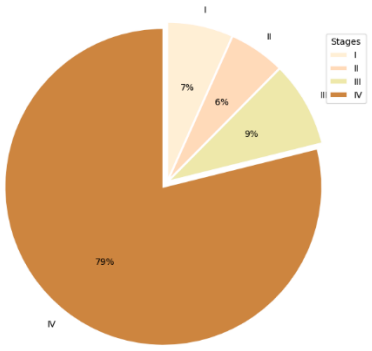

Fig. 2. Stages of lung cancer diagnosis in Quito.

Thus, it has been determined that attention should be paid to identifying patients who require screening and whether they meet the requirements to be candidates for this type of technique. As an example, Low-Dose Computed Tomography (LDCT) can be used for lung cancer screening in high- risk populations such as persons of 55 to 74 years of age with a relatively small smoking history [7,9]. In this way, since lung cancer is a heterogeneous disease, a misdiagnosis can lead to an increase in mortality since the treatment to be followed would be incorrect. Currently, due to technological advances, these tools can be used to perform better diagnoses and improve patient mortality [9].

B. *Economic cost of cancer*

Cancer results in an economic burden for both patients and health care systems, leaving countries with health care expenses and lost productivity due to morbidity and premature deaths [12, 14]. Due to the lack of data from certain countries, it is not possible to estimate costs worldwide, which is why most statistical studies take the United States as a reference due to the associations that are in charge of cancer statistics.

For example, for 2015 the national expenses directed to cancer are estimated to be 190.2 billion, while for 2020 the estimated expenses amount to 208.9 billion, increasing the figures by 10% [13, 15]. Specifically, the spending figures for Ecuador directed to cancer are estimated to be on average 13.84 million dollars. However, according to the estimates made by Simiao et. al [16], between the years 2020 and 2050 the expenses can rise to 28.7 million.

On the other hand, the economic impact caused by lung cancer is not fully known because there are no data available from some countries [12, 15]. Although there are different studies that have made approximations to estimate such figures, previous studies are 10 years old or more since they were carried out, so the estimates are outdated. In addition, most approaches have not taken into account dynamic changes in the population related to morbidity, mortality, and treatment costs, so new estimates for the economic impact of cancer have been needed [16, 17].

The diagnosis and treatment of cancer can be divided into 4 main stages of care: prediagnosis, initial, continuing and terminal [13]. The prediagnosis phase consists of a period of time from one year to 3 months before the diagnosis, where the medical expenses used are estimated if the patient had not developed lung cancer. In this phase, the 3 months prior to diagnosis are excluded so that the estimate is not affected by previous treatments in symptomatic patients [18].

The initial phase focuses on the first year after cancer diagnosis, while the terminal phase focuses on the patient's last year before death from cancer. Finally, there is the continuing phase that focuses on the period of time between the initial and terminal phases. The annual costs per patient are estimated to be higher in the terminal phase, followed by the initial, continuing and prediagnosis phase (Fig. 3), where medical services and oral prescription drugs are included [7,19].

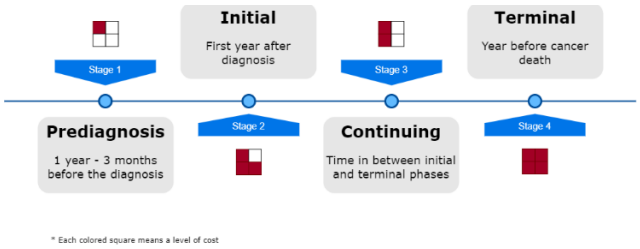

Fig. 3. Phases of care for cancer.

Therefore, as treatment costs such as medical costs such as hospitalizations, surgeries, radiotherapy or chemotherapy and the prescription of cancer drugs increase, prevention and early detection of cancer become more cost-effective and can become cost-saving [12, 20].

C. *Related Works*

The proposed method by Choi and Choi [35] reached 99% accuracy and 97.5% sensitivity, significantly reducing the number of false positives in highly sensitive nodule candidates. Table I presents in a general way the related works that have been dedicated to the detection of lung cancer for a future comparison with the results obtained at the end of this work.

The method of Carvalho Filho et al. [36] reached a precision of 99.2%, a sensitivity of 98% and a specificity of 98.9% for the classification between nodules and non-nodules. Da Silva et al. [37] model yielded accuracy values of 97.62%, a sensitivity of 92.20% and a specificity of 98.21%.

TABLE I. RELATED WORKS FOR LUNG CANCER DETECTION

Method	Database	Accuracy	Sensitivity	Precision
[35]	LIDC-IDRI	0.990	0.975	-
[36]	LIDC-IDRI	0.992	0.98	0.989
[37]	LIDC-IDRI	0.976	0.922	0.982

II. MATERIALS AND METHODS

Image analysis software can help detect abnormalities, such as tumors, fractures, or other injuries, that may be difficult to see with the naked eye [21]. Image analysis is a critical tool in the medical field that helps doctors and healthcare professionals diagnose diseases earlier and more accurately, leading to better patient outcomes [22].

A. *Database*

The Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) database is a publicly available database of thoracic CT scans, which is intended to serve as a resource for the development, training, and evaluation of CAD algorithms for lung nodules [23, 24].

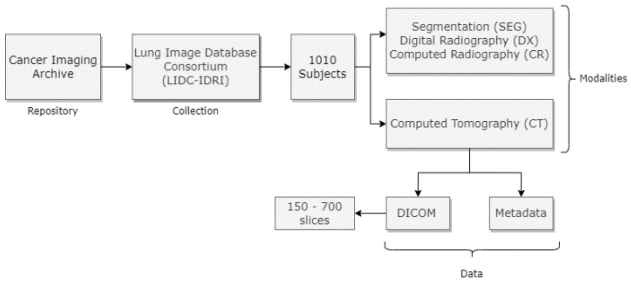

Fig. 4. Data acquisition from The cancer imaging archive.

The LIDC-IDRI database was acquired from The cancer imaging archive repository where the CT scans obtained were from 1010 patients divided into 3 different categories: nodules greater than 3 mm in diameter, where histological tests are required, minor nodules 3 mm, and no nodules. It also contains an XLM file with the annotations of each patient made by four radiologists.

Fig.4 shows the process of choosing the images that was carried out for the training of the neural network. The distribution of images that was carried out for the training of the network was random, allocating 70% for training and 15% for validation and 15% for testing. In this way, a total of 707 patients will be used for training, while validation and testing will use a total of 151 patients each.

B. *Preprocessing*

First, the computed tomographies were homogenized using the information provided by the .dcm format, such as pixel spacing and pixel array. In this way the values of each pixel in the image are converted to Hounsfield units (HU) used to measure the radiointensity of the image (Fig.5).

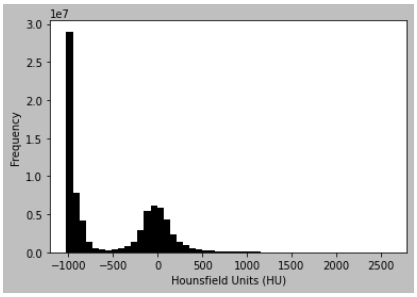

Fig. 5. Hounsfield units obteined from the CT scans.

A homogenization of the number of cuts that each patient has was carried out since this number depends on the tomograph with which it was performed. The number of slices chosen was 121 because it optimally focuses the region of the lungs that is the region of interest, thus eliminating slices that are out of range. In this case, the .dcm format type is helpful because the HU units specify the areas that each specific tissue belongs to.

Because the inputs of the neural network need to be the same, a homogenization is carried out in terms of size and orientation of each of the slices. The axial slices are resized to 124x124x1, these being the height and width of the image and the number of channels it has. Next, training data creation for nodule detection takes place, where both annotations and CT scans are used to obtain nodule coordinates and their diameter.

The data allocation for the training section, both for the section of nodules and non-nodules, is carried out randomly, complying with the percentages determined in the data distribution. The number of images used in both categories for training is summarized in the Table II.

TABLE II. CANDIDATES FOR TRAINING

Candidates selection for training	Quantity
Non-nodule	81305
Nodules	3535

C. Segmentation

Lung segmentation is the process of identifying and separating the lung regions from a medical image, such as a CT scan or X-ray [25, 26]. This task is essential for many medical applications, such as disease diagnosis, treatment planning, and image-guided interventions. There are several approaches to lung segmentation, including thresholding, region-growing, machine learning, and deep learning [27, 28].

In this case, thresholding method is used for the lung segmentation due to it involves selecting a threshold value to separate the lung regions from the rest of the image. The goal of lung segmentation is to find the region of interest that lies within the lungs. Therefore,a mask must be created to separate the region of the lungs from other areas involved in the image. The segmentation process is specified in Fig.6 where the first step to follow is binarization so that the image formed by pixels becomes a binary image.

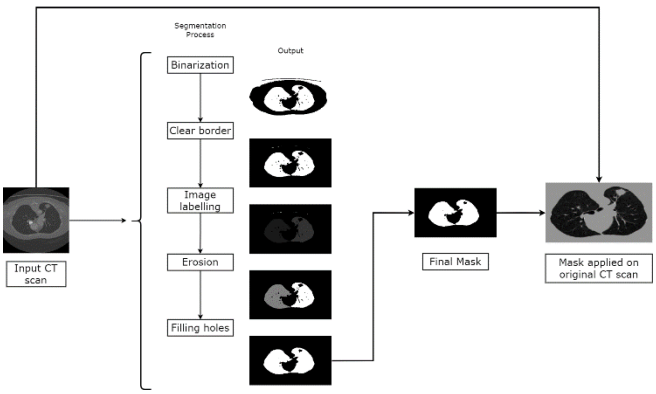

Fig. 6. Process of Lung segmentation.

The segmentation process works with the skimage library, specifically the segmentation module that provides different tools for the process. The image is converted to a grayscale to correctly select the binarization threshold. Using the HU values, it is determined that the lung tissue is at -500 HU. Therefore, the value of -400 HU is used as a threshold to separate the lung tissue from the other elements located in the image.

A border removal is done, where the spots that connect to the limit of the lungs are removed. This is done with the clear_border tool that will be in charge of suppressing the structures that are clearer than the environment and that are connected to the edge of the image. Afterwards, an image labeling is carried out, where the nodules are assigned a label while the rest of the image is assigned another type of label.

Consequently, an erosion is performed, which is a commonly used technique for binary morphological images. The boundaries of the white pixel regions are eroded, thus reducing their size, and the holes within those areas become larger. This step is fundamental because it makes a differentiation between the pulmonary nodules and the blood vessels and alveoli.

The gaps in the regions with the initial sizes are filled due to the goal is to keep the droplets connected to the lung walls. Lastly, hole filling is done, where the holes inside the binary mask are filled by means of the binary_fill_holes tool. After the segmentation process, the obtained mask is multiplied with the original CT scan for all tomographic slices.

Fig.7 shows a comparison between the input image with the output image after having gone through the segmentation process to later serve as input in the classification CNN.

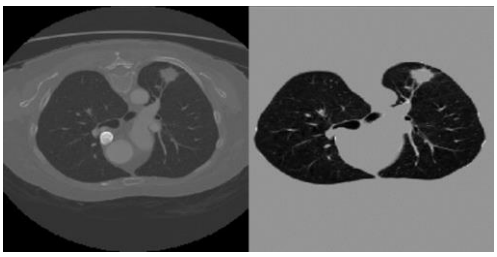

Fig. 7. Application of mask on lung CT Scan after the segmentation process.

D. Convolutional Neural Network architectures for classification

CNN consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers [30]. CNNs are typically trained using backpropagation, which adjusts the weights of the filters in each layer based on the error between the predicted and true labels. This process of training the model involves optimization techniques such as stochastic gradient descent or its variants [29, 30].

Recent advances in CNN architecture include the use of skipconnections, residual blocks, and attention mechanisms to improve performance and reduce training time. These improvements have led to the development of state-of-the-art models such as VGG16,ResNet, DenseNet, and EfficientNet, which have achieved top performance in several computer vision benchmarks [31].

1) VGG16 model

The VGG16 architecture consists of 16 convolutional layers, each with a 3x3 filter and a ReLU activation function. The output of each convolutional layer is passed through a 2x2 max pooling layer with stride 2, which reduces the spatial dimensions of the output while preserving important features. After the convolutional layers, the output is flattened and passed through 3 fully connected layers, which act as a classifier to predict the presence of cancer in the input image. The final layer is a softmax layer, which produces a probability distribution over the different classes (Fig.8)

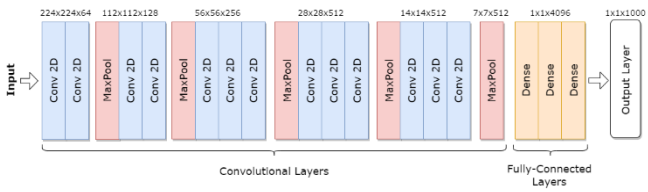

Fig. 8. VGG16 architecture.

The value 16 within the VGG16 symbology refers to the layers that have weights since, as can be seen in Fig.8, there are 13 convolutional layers, 5 Max Pooling layers and 3 dense layers, adding a total of 21 layers, however only 16 layers have learnable parameters. The network receives a tensor input size of

224x224 with a single channel. It focuses on the fact that the VGG16 has convolutionlayers with 3x3 filters with a stride of 1, taking into account that the same padding is always used and the MaxPooling layer that has a 2x2 filter with a stride of 2.

In short, first conv layer has 64 filters, second conv layer has 128 filters, third conv layer has 256 filters, fourth conv layer and fifth conv layer has 512 filters. The three fully connected layers (FC) follow a convolutional layer stack:the first two layers have 4096 channels each, the third layer performs 1000-way classification and thus contains 1000 channels (one channel per layer).

2) DenseNet model

The CNN is composed of several densely connected blocks, where each block contains a set of convolutional layers followed by a batch normalization and a ReLU activation function. The output of each block is concatenated to the input of every subsequent block in the network, forming a dense connectivity pattern [32].

The DenseNet, apart from the convolutional and pooling layers, consists of two important blocks which are the dense blocks and the transition layers. The architecture begins with a basic layer of convolution and pooling. Next follows a sequence of dense blocks followed by their respective transition layer. Finally, a dense block is followed by the classification layer.

The first convolution block has 64 filters of size 7x7 in a stride of 2 followed by the MaxPooling layer with a pool of 3x3 and a step of 2. As can be seen in Fig.9, each dense block has 3 convolutions with 1x1 and 3x3 sized kernels. In the transition layers the number of channels is reduced to half of the existing channels. There is a 1x1 convolutional layer and a 2x2 average pooling layer. The Keras backend (K) was used to take the initial tensor and return it as a tuple of dimension x. Finally, there is a GlobalAveragePooling followed by a final output layer.

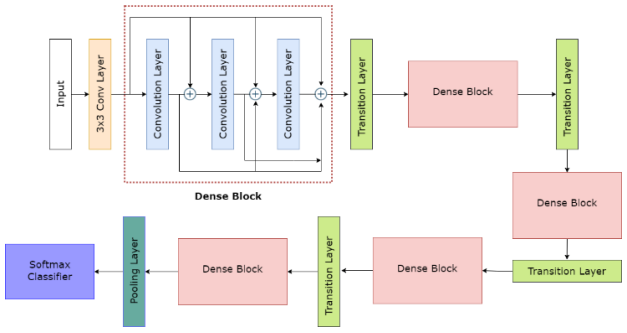

Fig. 9. DenseNet architecture.

E. Evaluation metrics

For this work, 4 types of metrics will be implemented, which are common when evaluating a CNN. These metrics are determined with the confusion matrix of the CNN model and will be further described below. As can be seen in Fig.10, confusion matrix reveals the number of true positives, false positives, true negatives and false negatives, which are used to determine the metrics mentioned before [33,34].

		Observed		
		n	1(+)	0(-)
Modeled	1(+)	a	b	
	0(-)	c	d	

a = true positives
b = false postitives
c = false negatives
d = true negatives
n = a+b+c+d

Fig. 10. Principle of confusion matrix.

Metrics used to determine the performance of the model are Accuracy, Precision, Sensitivity, and Specificity which are described as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

Representing into the equations TP: True positive, TN: True negative, FP: False positive and FN: False negative

III. Results and Discussion

A. Evaluation Metrics

The main metrics considered for the evaluation of the model were obtained from the confusion matrix applied to the training, validation and testing sets. Table III shows in a summarized way the values obtained in the confusion matrix.

The values presented in the confusion matrix for each set were obtained by separating the data with a specific percentage for each set. Thus, for the training set, 70% of the total data obtained from the LIDC-IDRI database was allocated, while for the validation and testing sets, 15% of the total images were allocated for each one. Obtaining a total of 84,840 images for the training set, 18,392 images for the validation set, and 18,271 images for the testing set.

TABLE III. Confusion Matrix for Training, Validation and Testing Sets Obtained from the Proposed Models.

		Training		Validation		Testing	
		(+)	*(-)*	*(+)*	*(-)*	*(+)*	*(-)*
VGG-16	*(+)*	3142	157	704	56	671	89
	(-)	516	81025	237	17395	208	17424
DenseNet	*(+)*	3273	139	701	59	684	76
	(-)	327	81101	151	17481	160	17351

All the metrics were evaluated for the VGG16 and DenseNet models, where the training was carried out with 20, 35 and 50 epochs. However, for the analysis only the models trained at 35 epochs were taken into account. In this way, the results of the metrics established in the training, validation and testing sets were obtained, where the highest value obtained in each section was chosen for the analysis. As can be seen in Table IV, all values obtained during training, validation, and testing are summarized, where the difference between the two

proposed models in terms of established metrics can be observed.

In the model accuracy section, it can be seen that the DenseNet model with a training process with 35 epochs achieved an accuracy of 99.5%, surpassing the VGG16 model, which has an accuracy of 99.2%. For the validation and testing sets, a decrease in accuracy was observed for both models, this is largely due to the reduced data set compared to the training set.

TABLE IV. EVALUATION ON RESULTS USING THE METRICS ACCURACY, PRECISION, SENSITIVITY, SPECIFICITY ON THE LIDC-IDRI DATABASE.

Model	**Set**	**Metrics**			
		Accuracy	***Precision***	***Sensitivity***	***Specificity***
VGG16	***Training***	0.992	0.859	0.952	0.994
	Validation	0.984	0.748	0.926	0.987
	Test	0.984	0.763	0.883	0.988
DenseNet	***Training***	0.995	0.909	0.959	0.996
	Validation	0.989	0.823	0.922	0.991
	Test	0.987	0.810	0.90	0.991

For the precision, sensitivity and specificity metrics, it can be observed that a similar pattern is observed where the training sethas the best performance of the models compared to the validation and testing sets. In addition, in each of the metrics, the DenseNet model presented better overall performance compared to the VGG16 model.

Although there is no significant difference in the values obtained for both models, the precision and sensitivity values may indicate that the performance of the networks is not the best, since there may be a significant number of false positives and false negatives. Unlike the precision of both models, which almost reaches 1, the values of the other metrics are lower.

Especially the precision, where the lowest values are obtained, indicating that there are cases in which cancer is diagnosed when the patient is actually healthy, resulting in a false positive and a type 1 error. Although, as mentioned above, when there is a data imbalance between the categories, the precision value is notas reliable.

In more detail, as can be seen in the Table I, the images used during the training process that have cancer labels (3535) are considerably smaller than the images labeled as non-cancer (81305). Causing that both categories are not balanced with respect to each other. In the same way, both having greater amounts of data labeled as positive, the models tend to better detect images that do not have cancer, as can be seen in the confusion matrix, both models have a greater number of false positives than of false negatives.

Similarly, it can be observed with the evaluation metrics, since the specificity that determines the true negative rate is higher compared to the precision that is responsible for determining the true positive rate. Although as mentioned before, due to the imbalance between both categories, it is not a reliable enough metric.

B. Comparison between the proposed models

Regarding the performance of the models, Fig.11 shows the performance comparison of the model between the VGG16 and DenseNet networks.

The graph shows how there is a considerable difference between the two methods, since the DenseNet network needs a lower number of epochs to reach 99% accuracy, while the VGG16 network needs a higher numberof epochs to reach the same accuracy value. In addition, the curves of the DenseNet model have greater stability than the VGG16 model, demonstrating that the learning of the network is carriedout in an optimized way.

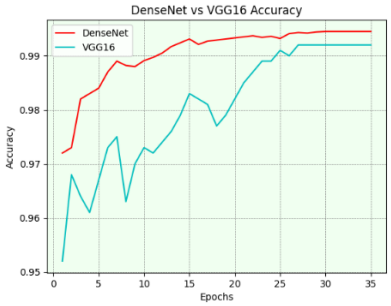

Fig. 11. Performance comparison between VGG16 and DenseNet models.

The confusion matrix and the stability of the values obtained during training become a determining factor when choosing the best diagnostic model, since the values of maximum precision do not make a great difference. Regarding the other metrics, there is a notable difference showing a notable range between the values of both networks.

On the other hand, regarding the loss curves of the models, they show how the loss function, in this case the Binary crossentropy, changes over time. The purpose of this function is to measure the difference between the predicted output and the target output, taking care of the optimization of the weights and biases of the network. For this comparison, the training process of 35 epochs was taken into account.

As can be seen in Fig.12, the curves for both models decrease as the epochs progress, which indicates that both models improved their precision. In addition, both models show stable values throughout the training, which shows that the model learns well with the data provided. However,it can be seen how the DenseNet model shows a better approximation with the objective output data by having smaller values compared to the VGG16 model.

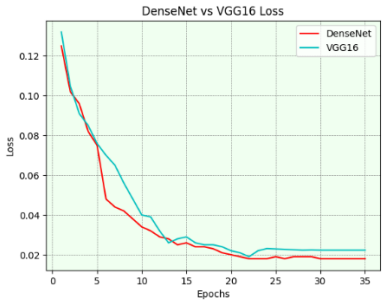

Fig. 12. Loss comparison between VGG16 and DenseNet models.

C. Comparison with literature

A comparison of the CNN models carried out throughout this work with other authors who were previously mentioned was made in order to validate the proposed models. The database used in all the works was the LIDC-IDRI database due to the large number of images it has on lung cancer.

Table V summarizes the results obtained from the comparison. Compared with the methods carried out by other authors, the values of the metrics do not differ to a great extent despite having carried out very different processes from each other.

Since the authors use pretrained networks and also perform a combination of a densely connected network with a sparsely connected network, good network performance has been recorded, thus avoiding overfitting of the model. However, regarding the models proposed in this project, it is shown that in comparison with the other models, the network has a good performance for the detection of pulmonary nodules.

Regarding the sensitivity and precision metrics, the DenseNet model has high values. On the other hand, the VGG16 model has good precision, but it is the model with the least sensitivity, with a value of 0.952, meaning that the model has the lowest rate of true positives. Because the models made by the authors do not have specificity values or training times, comparisons cannot be made in terms of these sections.

TABLE V. RESULTS OF THE PROPOSED METHODOLOGY FOR LUNG CANCER DETECTION COMPARED TO SIMILAR PREVIOUS WORKS.

Method	Accuracy	Sensitivity	Precision	Specificity
[35]	0.990	0.975	-	-
[36]	0.992	0.98	0.989	-
[37]	0.976	0.922	0.982	-
Proposed VGG16	0.992	0.952	0.859	0.994
Proposed DenseNet	0.995	0.959	0.909	0.99

IV. CONCLUSION AND LIMITATIONS

The models demonstrated a good degree of classification, having a high specificity of 0.994 for the VGG16 networks and 0.996, which could minimize misdiagnoses. In addition, the curves obtained demonstrated that the learning of the networks was correct, finding the correct weights in the identification of patterns. However, since the works of the other authors do not show specificity data or training time, it cannot be ensured that the proposed models are optimized in terms of the previously proposed networks for the detection of lung cancer.

Therefore, the work carried out has managed to meet most of the objectives proposed at the beginning, managing to create CNN models that can detect lung cancer from CT images, surpassing previously proposed models in the same field, which can optimize the detection of cancer at an early stage, improving the survival of diagnosed patients. In addition, the correct validation of the networks was carried out, finding the best parameters to avoid excessive computational consumption without diminishing their effectiveness.

Among the main limitations that were found during the work is that the final objective could not be fully met because with the proposed models it was not possible to perform lung cancer staging or perform a multimodal classification. This is mainly due to the segmentation process, since several details of the nodules were lost in the process of creating the masks.

In the same way, these masks lost the dimensionality of the nodules, so that extremely small nodules were lostduring the segmentation, which indicates that the networks could not perform a correct detectionof the nodules in an early stage of cancer. Another important limitation to take into account was the computational capacity of the device where the network training was carried out, since the hardware plays an important role in the process.

V. FUTURE WORK

As for the future work to be carried out, to improve the convolutional models, an investigation ofbetter segmentation techniques can first be carried out and in this way improve the results to be obtained. Another important issue to consider is to seek to homogenize the data in each category,with a similar number of images in each category, the results of the metrics can be improved and thus avoid false positives and false negatives.

Furthermore, the architecture of both models can also be improved so that a multimodal classification can be performed. Thus, improving the type of diagnosis of lung cancer, showing the stage of the tumor and what kind of tumor is being diagnosed, in the same way, it will also seek to improve and reduce computational consumption and training time for the optimization of the networks.

Finally, another prospect to take into account is the incorporation of previous models for later comparison. In this way, the optimization of the proposed networks can be guaranteed both in time and in computational consumption. In this way, these models could serve as a support tool for health professionals in Ecuador, since they could improve diagnoses without the need for many resources.

REFERENCES

[1] L. Pecorino, *Molecular Biology of Cancer*. Oxford: Oxford University Press, 2016.

[2] R. Paulmurugan, "Introduction to cancer biology," *Molecular Imaging Probes for Cancer Research*, pp. 3–27, 2012.

[3] "Worldwide cancer data: World cancer research fund international," WCRF International, https://www.wcrf.org/cancer-trends/worldwide-cancer-data/.

[4] R. Sharma, "Mapping of global, regional and national incidence, mortality and mortality-to-incidence ratio of lung cancer in 2020 and 2050," *International Journal of Clinical Oncology*, vol. 27, no. 4, pp. 665–675, 2022.

[5] B. C. Bade and C. S. Dela Cruz, "Lung cancer 2020," *Clinics in Chest Medicine*, vol. 41, no. 1, pp. 1–24, 2020.

[6] M. Alexander, S. Y. Kim, and H. Cheng, "Update 2020: Management of Non-Small Cell Lung Cancer," *Springer Nature*, vol. 198, pp. 897–907, 2020.

[7] D. Almeida-Galarraga *et al.*, "Glaucoma detection through digital processing from fundus images using MATLAB," *2021 Second International Conference on Information Systems and Software Technologies (ICI2ST)*, 2021.

[8] Instituto Nacional de Estadística y Censos, "Población Y Demografía," Instituto Nacional de Estadística y Censos, https://www.ecuadorencifras.gob.ec/censo-de-poblacion-y-vivienda/.

[9] O. P. Yanchatuña *et al.*, "Skin lesion detection and classification using convolutional neural network for deep feature extraction and support

vector machine," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 11, no. 3, p. 1260, 2021.

[10] W. Tarupi, A. Velasco, O. González, H. Caballero, and P. Cueva, Solca Núcleo de quito boletín epidemiológico, https://solcaquito.org.ec/wp-content/uploads/2022/04/boletinEpidemiologico02_2022.pdf

[11] E. Médica, "79% de Casos de Cáncer de Pulmón son diagnosticados en Quito," Edición, https://www.edicionmedica.ec/secciones/salud-publica/79-de-casos-de-cancer-de-pulmon-son-diagnosticados-en-quito-98318.

[12] J. D. Suquilanda-Pesántez *et al.*, "Prediction of parkinson's disease severity based on gait signals using a neural network and the fast fourier transform," *Artificial Intelligence, Computer and Software Engineering Advances*, pp. 3–18, 2021.

[13] P. A. Vásquez-Ucho *et al.*, "Analysis and evaluation of the systems used for the assessment of the cervical spine function: A systematic review," *Journal of Medical Engineering & Technology*, vol. 45, no. 5, pp. 380–393, 2021.

[14] "Cancer statistics," National Cancer Institute, https://www.cancer.gov/about-cancer/understanding/statistics.

[15] S. Chen *et al.*, "Estimates and projections of the global economic cost of 29 cancers in 204 countries and territories from 2020 to 2050," *JAMA Oncology*, vol. 9, no. 4, p. 465, 2023.

[16] E. D. Aguiar Salazar *et al.*, "Design of a glove controlled by electromyographic signals for the rehabilitation of patients with rheumatoid arthritis," *Information and Communication Technologies*, pp. 3–11, 2020.

[17] "The burden," The Cancer Atlas, https://canceratlas.cancer.org/the-burden/.

[18] D. F. Sheehan *et al.*, "Lung cancer costs by treatment strategy and phase of care among patients enrolled in Medicare," *Cancer Medicine*, vol. 8, no. 1, pp. 94–103, 2018.

[19] O. Alvarado-Cando, H. Torres-Salamea, and D. A. Almeida, "Uda-µBioLab: Teaching microcontrollers with bioinstrumentation," *IFMBE Proceedings*, pp. 877–880, 2018.

[20] J. Pereira-Carrillo *et al.*, "Comparison between two novel approaches in Automatic Breast Cancer Detection and diagnosis and its contribution in military defense," *Smart Innovation, Systems and Technologies*, pp. 189–201, 2021.

[21] K. R. Yabroff, J. Lund, D. Kepka, and A. Mariotto, "Economic burden of cancer in the United States: Estimates, projections, and future research," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 20, no. 10, pp. 2006–2014, 2011.

[22] S. Armato *et al.*, "The lung image database consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed public database of CT scans for lung nodule analysis," *Medical Physics*, vol. 37, no. 6, pp. 3416–3417, 2010.

[23] D. A. Almeida-Galárraga, A. Ros Felip, F. Marco Martínez, and L. Serrano-Mateo, "Photoelastic analysis of shoulder arthroplasty: Current descriptive analysis of research in scientific journals," *IFMBE Proceedings*, pp. 713–717, 2018.

[24] J. D. Suquilanda-Pesántez *et al.*, "NIFtHool: An informatics program for identification of NifH proteins using deep neural networks," *F1000Research*, vol. 11, p. 164, 2022.

[25] C. Jacobs *et al.*, "Computer-aided detection of pulmonary nodules: A comparative study using the public LIDC/IDRI Database," *European Radiology*, vol. 26, no. 7, pp. 2139–2147, 2015.

[26] S. G. Armato *et al.*, "The Lung Image Database Consortium (LIDC): An evaluation of radiologist variability in the identification of lung nodules on CT scans," *Academic Radiology*, vol. 14, no. 11, pp. 1409–1421, 2007.

[27] J. Firdouse and M. Balasubramanian, "A Survey on Lung Segmentation Methods," *Advances in Computational Sciences and Technology*, vol. 10, no. 9, pp. 2875–2885.

[28] S. Guo and L. Wang, "Automatic CT image segmentation of the lungs with an iterative Chan-Vese algorithm," *2015 International Conference on Informatics, Electronics & Vision (ICIEV)*, 2015.

[29] J. H. Moltz *et al.*, "Advanced segmentation techniques for lung nodules, liver metastases, and enlarged lymph nodes in CT scans," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 122–134, 2009.

[30] T. Guo, J. Dong, H. Li, and Y. Gao, "Simple convolutional neural network on Image Classification," *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(*, 2017.

[31] Phung and Rhee, "A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets," *Applied Sciences*, vol. 9, no. 21, p. 4500, 2019.

[32] I. Rizwan I Haque and J. Neubert, "Deep Learning Approaches to Biomedical Image segmentation," *Informatics in Medicine Unlocked*, vol. 18, p. 100297, 2020.

[33] *Artificial Intelligence and Internet of Things: Applications in Smart*. S.l.: CRC PRESS, 2021.

[34] G. S. Handelman *et al.*, "Peering into the black box of artificial intelligence: Evaluation metrics of machine learning methods," *American Journal of Roentgenology*, vol. 212, no. 1, pp. 38–43, 2019.

[35] W.-J. Choi and T.-S. Choi, "Automated pulmonary nodule detection based on three-dimensional shape-based feature descriptor," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 37–54, 2014.

[36] A. O. de Carvalho Filho, A. C. Silva, A. C. de Paiva, R. A. Nunes, and M. Gattass, "Lung-nodule classification based on computed tomography using taxonomic diversity indexes and an SVM," *Journal of Signal Processing Systems*, vol. 87, no. 2, pp. 179–196, 2016.

[37] G. L. da Silva, T. L. Valente, A. C. Silva, A. C. de Paiva, and M. Gattass, "Convolutional Neural Network-based PSO for lung nodule false positive reduction on CT images," *Computer Methods and Programs in Biomedicine*, vol. 162, pp. 109–118, 2018.

[38] T. Sajja, R. Devarapalli, and H. Kalluri, "Lung cancer detection based on CT scan images by using Deep Transfer Learning," *Traitement du Signal*, vol. 36, no. 4, pp. 339–344, 2019.