PROCEEDINGS OF SCIENCE

A machine learning approach for the feature extraction of pulmonary nodules

Cecilia Irene Loeza Mejía*
Departamento de Posgrado e Investigación, Tecnológico Nacional de México, Campus Misantla Misantla, Veracruz, México
E-mail: ceciliaIoeza@yahoo.com

R. R. Biswal*
Tecnologico de Monterrey, Escuela de Ingenieria y Ciencias
E-mail: rroshanb@tec.mx

Gregorio Fernandez Lambert
Departamento de Posgrado e Investigación, Tecnológico Nacional de México, Campus Misantla Misantla, Veracruz, México
E-mail: gfernandezl@misantla.tecnm.mx

In recent times, computational studies have emerged as a viable alternative for complementing the efforts of experienced radiologists in disease diagnosis. Computed tomography (CT) studies are a common way of predicting the lung nodule malignancy for the early diagnosis and treatment of lung cancer in patients. Early detection of the type of nodule is the key to determining the appropriate treatment, thus increasing patient survival. Feature extraction is an important stage in classifying benign and malignant nodules in chest CT scans. However, determining the type of nodule in CT scans is a challeging task in medical imaging, since CT images cannot be evaluated as an average or generic image. Hence, this study is based on the application of machine learning techniques, specifically convolutional neural networks (CNNs) and transfer learning, for the feature extraction and identification of tomograms with pulmonary nodules on a public, Lung TIME dataset. Pretrained CNNs architectures (VGG, ResNet, MobileNet, Xception, NASNet and DenseNet) and a proposed CNN architecture were used, thereby obtaining a minimum training accuracy in pretrained architectures of 70.11% and minimum test accuracy of 33.91%. In contrast, in the proposed CNN, 95.25% and 94.21% respectively were obtained. These results show that the transfer learning is not always feasible in medical applications and architectures focused on the problem to be solved are usually most effective.

Artificial Intelligence for Science, Industry and Society, AISIS2019
October 21-25, 2019
Universidad Nacional Autónoma de México, Mexico City, México

*Speaker.

PoS(AISIS2019)024

1. Introduction

Massive progress has been observed in the successful application of machine learning in the identification, prediction, image segmentation, optimization and classification tasks. Motivated by the success of machine learning and evolution of biomedical images, algorithms for the medical field are increasingly used in the early identification and classification of diseases [1, 2, 3, 4, 5, 6, 7]. In machine learning research, deep learning techniques are considered to facilitate feature extraction [8, 9, 10, 11, 12], which involves using high-level representations learned by a network with the aim of obtaining interesting features of new samples [13]. Also, it is an important stage to obtain reliable results of a vision system [14]. Through deep learning techniques, different types of biomedical images have been analyzed [6, 15], such as X-rays, ultrasound [16], magnetic resonance imaging (MRI) [17], and computed tomography (CT). This has allowed the characterization of organs and pathologies, among them cancer [18]. Convolutional neural networks (CNN) have been used in several investigations of the state-of-the-art on lung cancer using CT thorax scans, showing that they can learn directly from the data [19, 20, 21, 22, 23, 24, 25, 26] instead of using hand-crafted features, showing comparable results in the existing domain of the expert. In addition, using CNN, transfer learning can be performed, which consists of learning discriminatory features in a knowledge domain and applying it in a related domain. Transfer learning is mainly used in applications where little training data is available [27]. However, transfer learning has been little explored in research related to lung cancer, in addition to the comparison (both in accuracy and at running time) between several CNN architectures applied to chest CT analysis. This investigation focuses in comparing different CNN architectures for the feature extraction of tomograms with pulmonary nodules in the Lung Test Images from Motol Environment (Lung TIME [28]) dataset and to identify whether the tomogram has nodules. Different CNN architectures with transfer learning such as the likes of VGG [29], ResNet [30], MobileNet [31], Xception [32], NASNet [33] and DenseNet [34] as well as a proposed CNN architecture were used. These architectures were chosen because they have had an excellent performance in tasks of identification and classification of images. All CNN architectures were analyzed both using the Dropout layer with a ratio of 0.0005 and without it. This paper is organized as follows, in the second section the materials and methods used are shown, in the third section comparisons in terms of accuracy and training time in CNN architectures are provided and in the last section the conclusions and future work of the investigation are presented.

2. Dataset and methods

The feature extraction and identification of tomograms with pulmonary nodules were performed by using CNNs. Pretrained CNN architectures (VGG, ResNet, MobileNet, Xception, NASNet and DenseNet) were compared using the weights of the ImageNet [35] dataset with a proposed CNN architecture. The dataset Lung TIME was used to carry out the experiments, which contains 157 computed tomography (CT) images of the thorax in DICOM 3.0 format with annotated nodules in the XML format. 2003 tomograms with nodules and 4158 without nodules were used, keeping the fact that, each CT consists of large number of tomograms with 512x512 pixels in size. Later, the tomograms were resized to 64x64 pixels comprising three color channels, with the purpose of being able to use them as input in the pretrained CNN architectures. To analyze the performance of CNNs, 70% of the tomograms were randomly selected for training and the rest for validation. Figure 1 shows the underlying methodology for the study, which consists of a tomogram with 512x512 pixels as the input, subsequently resized to 64x64 pixels which is received as the input to a CNN architecture; both for the transfer learning process as well as for the evaluation of the proposed architecture. As a result, the identification of tomograms with nodules or without nodules is obtained.

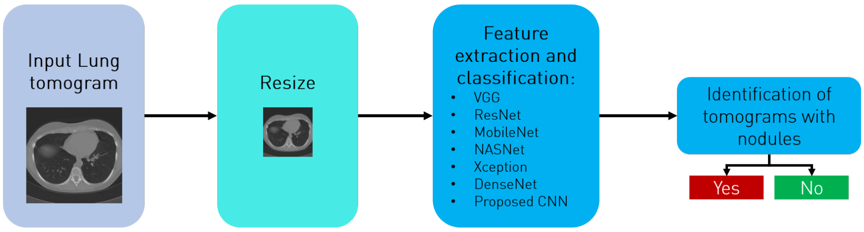

Figure 1: Outline of the Methodology

2.1 CNN architectures used

2.1.1 Proposed CNN

Figure 2 shows our CNN, which consists of five successive layers of convolution (with ReLU activation) and max-pooling, followed by a flattening layer, which is followed by fully connected layer, and a final fully connected softmax layer. The architecture was developed incorporating the Dropout (with a rate of 0.0005) layer and without incorporating it.

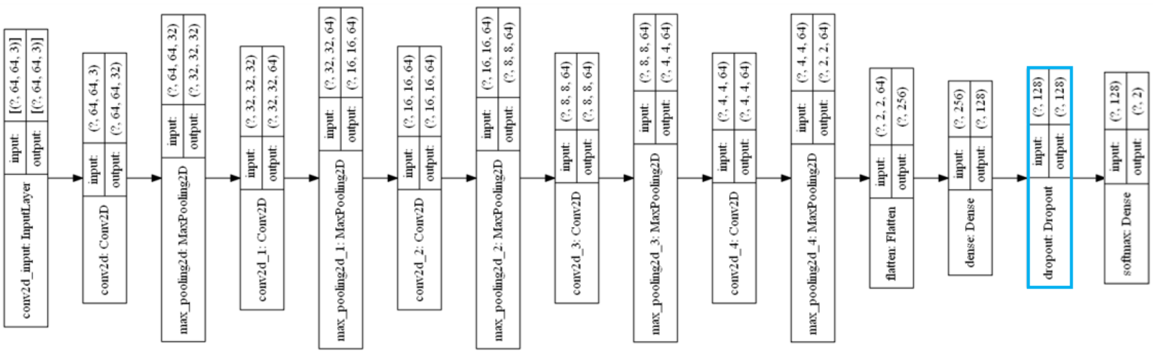

Figure 2: Proposed CNN architecture

2.1.2 Transfer learning with pretrained CNNs

Through the transfer learning, CNN architectures can be used in many domains and different tasks than those for which they were trained [36]. In addition, they can be used when there is not enough training data [27]. So, it was decided to use them to observe their behavior in the identification of tomograms with pulmonary nodules. Table 1 shows the pretrained CNN to perform the feature extraction and identification of tomograms with nodules, which were chosen, given the competitive performance they have achieved in identification and classification tasks [29, 30, 31, 32, 33, 34]. Figure 3 presents the pipeline of the CNN network using transfer learning, which consists of an input layer, then the pretrained model (without the last predicting layer) layer followed by a Flatten layer and three successive Dense layers. The experiments were conducted incorporating the Dropout (with a rate of 0.0005) layer and without incorporating it.

Table 1: Pretrained CNNs used

Architecture	Variants	Description
VGG	VGG16 VGG19	Achieve state-of-the-art accuracy on ILSVRC classification and localization tasks [29].
ResNet	ResNet101 ResNet152 ResNet50	Useful in detection, location, classification and segmentation tasks [30].
MobileNet	MobileNet MobileNetV2	It requires less computing power. Useful in mobile and embedded applications [31].
Xception	–	Useful in image classification tasks [32].
NASNet	NASNetMobile NASNetLarge	Achieve state-of-the-art in the CIFAR-10 and ImageNet datasets [33].
DenseNet	DenseNet121 DenseNet169 DenseNet201	May be good feature extractors in various computational vision tasks [34].

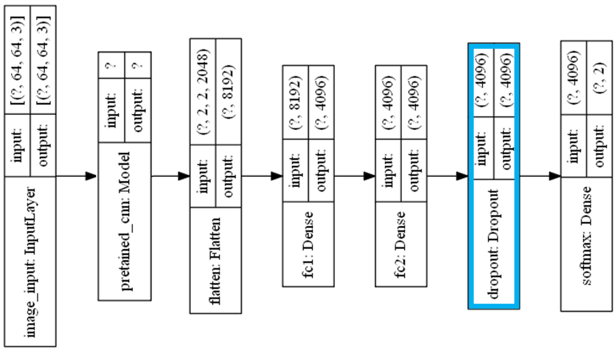

Figure 3: CNN architecture incorporating transfer learning

2.2 Experimental setup

In all architectures, the Nadam [37] optimizer with a learning rate 0.001 and batch size of 32 was used. The loss function used was sparse categorical crossentropy. Training was done in 5 epochs. Experiments were performed with a Dropout rate of 0.0005 and without Dropout. To implement the CNNs, Tensorflow 2.0 was used in Python 3.6.7. The library that was used to read the tomograms was imageio. The experiments were conducted on a computer with Windows 10 64-bit (Build 18362), Intel(R) Core(TM) i3-5015U CPU @ 2.10GHz (4 cores), 2.1GHz and 16 GB RAM.

3. Experimental results & discussion

Table 2 reports training accuracy, test accuracy and runtime in minutes with/without Dropout, obtained using CNN. In both the VGG16 and VGG19 architectures, it was observed that when using Dropout, the accuracy of training decreased and the training time increased. In pretrained CNNs, the greater training accuracy of 97.15% without Dropout was observed in ResNet50. In contrast, when using Dropout rate of 0.0005 on MobileNet, a 97.26% was obtained. However, the test accuracy was less than 70%, which is not acceptable in medical applications. A shorter runtime was observed in MobileNet. and ResNet152 got the longest training time. So, in applications that use few computing resources, it is recommended to use a MobileNet architecture instead of RestNet. While the best test accuracy using transfer learning was achieved by the VGG16 architecture, a comparable accuracy of 73.23% and 75.77% was obtained in each case, with and without using Dropout. It was also noted that the use of Dropout did not always improve accuracy, also, it increases the execution time. Overall, the performance of the proposed CNN architecture was better than the pretrained CNNs, obtaining a training accuracy of 96.08% and a test accuracy 94.75% without using dropout, while 95.25% and 94.21% respectively when used. In addition, the proposed CNN was trained in less time, which means that in medical applications specifically where chest CT scans are used, the use of transfer learning can give imprecise results than an architecture specifically focus on the problem to be solved.

Table 2: Results with pretrained architectures and proposed CNN

	No Dropout			With Dropout		
CNN	Training ACC	Test ACC	Training time	Training ACC	Test ACC	Training time
VGG16	0.7419	0.7323	30.07	0.7252	0.7577	32.16
VGG19	0.7087	0.7161	34.90	0.7011	0.7090	37.81
ResNet101	0.9668	0.6609	130.31	0.9692	0.6609	126.78
ResNet152	0.9647	0.6609	158.46	0.9643	0.6609	160.68
ResNet50	**0.9715**	0.3391	51.20	0.9719	0.3391	23.11
MobileNet	0.9682	0.6598	13.63	**0.9726**	0.6733	9.85
MobileNetV2	0.9372	0.4072	12.00	0.9330	0.5008	10.67
Xception	0.8583	0.4024	18.54	0.8581	0.3802	20.04
NASNetMobile	0.8882	0.6609	16.14	0.8836	0.6522	16.37
NASNetLarge	0.9332	0.5349	46.19	0.9246	0.5641	50.90
DenseNet121	0.9583	0.4732	14.77	0.9578	0.5479	18.47
DenseNet169	0.9620	0.6230	19.54	0.9585	0.6701	20.89
DenseNet201	0.9636	0.6382	20.93	0.9647	0.6111	23.04
Proposed CNN	0.9608	**0.9475**	**7.23**	0.9525	**0.9421**	**5.68**

4. Conclusions and future directions

In this study different CNN architectures were used for the feature extraction and identification of tomograms with pulmonary nodules. It was observed that in case of pretrained CNN architectures, MobileNet requires less computational power, on the contrary, ResNet uses greater computing power. In some cases of feature extraction and identification of tomograms with nodules, although the training accuracy is greater than 90%, the test accuracy is found to be less than 70%, which for the purpose of transfer learning, is not an feasible option for CT scans. This is because, these architectures were trained for problems not focusing on medical applications. In general, with the proposed CNN architecture, better results were obtained, with a training accuracy of 96.08% and a test accuracy of 94.75% without using Dropout. In contrast, when using Dropout, the results were slightly lower 95.25% and 94.21% respectively. The proposed architecture can be used as part of a pulmonary nodule detection system. Nonetheless, it was also shown that the use of Dropout does not always improve accuracy, but affects adversely, the execution time in most cases. When using Dropout, only in the architectures ResNet101, ResNet50, MobileNet, MobileNetV2 and in the proposed CNN the training time was reduced. Thus, further research in the medical field is required for evaluating architectures that use transfer learning, mainly for improving test accuracy and parameter optimization, to be able to provide shorter training times. In addition, appropriate Dropout rates can also be studied in order to reduce overfitting. Finally, it is equally necessary to perform a correlation of the medical records with images, in order to improve accuracy and to obtain reliable results.

References

[1] A. Criminisi, *Machine Learning for Medical Images Analysis*, *Medical Image Analysis* **33** 2016

[2] C. Liew, *The future of radiology augmented with Artificial Intelligence: A strategy for success*, *European Journal of Radiology* 2018

[3] M. L. Giger, *Machine Learning in Medical Imaging*, *Journal of the American College of Radiology* **15** (3) 2018

[4] P. Meyer, *Survey on deep learning for radiotherapy*, *Computers in Biology and Medicine* **98** (1) 2018

[5] I. Balki *et al.*, *Sample-Size Determination Methodologies for Machine Learning inMedical Imaging Research: A Systematic Review*, *Canadian Association of Radiologists Journal* **70** 2019

[6] Z. Zhang, E. Sejdić, *Radiological images and machine learning: Trends, perspectives, andprospects*, *Computers in Biology and Medicine* **108** 2019

[7] G. Currie, *Machine Learning and Deep Learning in Medical Imaging: Intelligent Imaging*, *Journal of Medical Imaging and Radiation Sciences* **50** 2019

[8] H. Yang *et al.*, *Deep Learning for the Classification of Lung Nodules*, *arXiv* 2016 [1611.06651v2].

[9] J. Cheng *et al.*, *Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans*, *Nature* **6** (24454) 2016

[10] C. Wang *et al.*, *Lung nodule classification using deep feature fusion in chest radiography*, *Comput Med Imaging Graph.* 2016

[11] C. Cao *et al.*, *Deep Learning and Its Applications in Biomedicine*, *Genomics, Proteomics & Bioinformatics* **16** (1) 2018

[12] H. Polat, H. Mehr, *Classification of Pulmonary CT Images by Using Hybrid 3D-Deep Convolutional Neural Network Architecture*, *Appl. Sci.* **9**, 2019

[13] F. Chollet, *Deep Learning with Python*, Manning Publications, NY 2018

[14] S. Khan *et al.*, *A Guide to Convolutional Neural Networks for Computer Vision*, Morgan & Claypool, 2018

[15] G. Litjens *et al.*, *A survey on deep learning in medical image analysis*, *MedicalImageAnalysis* **42** 2017

[16] S. Liu, *Deep Learning in Medical Ultrasound Analysis: A Review*, *Engineering* **5** (2) 2019

[17] A.S. Lundervold, A. Lundervold, *An overview of deep learning in medical imaging focusing on MRI*, *Z Med Phys29* **29** 2019

[18] A. B. Levine *et al.*, *Rise of the Machines: Advances in Deep Learning for Cancer Diagnosis*, *Trends in Cancer* **5** (3) 2019

[19] A. A. A. Setio *et al.*, *Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks*, *EEE Transactions on Medical Imaging* **35** (5) 2016

[20] W. Li *et al.*, *Pulmonary Nodule Classification with Deep Convolutional Neural Networks on Computed Tomography Images*, *Computational and Mathematical Methods in Medicine* 2016

[21] A. Shen *et al.*, *Multi-crop Convolutional Neural Networks for lung nodule malignancy suspiciousness classification*, *Pattern Recognition* **61** 2017

[22] S. Hamidian *et al.*, *3D Convolutional Neural Network for Automatic Detection of Lung Nodules in Chest CT*, *Proc SPIE Int Soc Opt Eng* 2017

[23] G. Kang *et al.*, *3D multi-view convolutional neural networks for lung nodule classification*, *PLoS ONE* **12** (11) 2017

[24] X. Zhao *et al.*, *Agile convolutional neural network for pulmonary nodule classification using CT images*, *IJCARS* 2018

[25] N. Tajbakhsh, K. Suzuki, *Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs*, *Pattern Recognition* **63** 2017

[26] R. Dey *et al.*, *DIAGNOSTIC CLASSIFICATION OF LUNG NODULES USING 3D NEURAL NETWORKS*, *ISBI* 2018 [1803.07192v1]

[27] C. Tan *et al.*, *A Survey on Deep Transfer Learning*, *arXiv* 2018 [1808.01974v1]

[28] M. Dolejsi *et al.*, *The Lung TIME-Annotated lung nodule dataset and nodule detection framework*, *Proceedings of SPIE* Vol. 7260, 72601U, 2009

[29] K. Simonyan, A. Zisserman, *Very deep convolutional networks for large-scale image recognition ICLR* 2015 [1409.1556]

[30] K. He *et al.*, *Deep Residual Learning for Image Recognition* , *arXiv* 2015 [1512.03385v1]

[31] A. G. Howard, *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*, *arXiv* 2017 [1704.04861v1]

[32] F. Chollet, *Xception: Deep Learning with Depthwise Separable Convolutions*, *arXiv* 2017 [1610.02357v3]

[33] B. Zoph *et al.*, *Learning Transferable Architectures for Scalable Image Recognition*, *arXiv* 2017 [1707.07012].

[34] G. Huang *et al.*, *Densely Connected Convolutional Networks*, *arXiv* 2018 [1608.06993v5].

[35] J. Deng *et al.*, *ImageNet: A large-scale hierarchical image database*, *IEEE conference on computer vision and pattern recognition* 2009

[36] A. Nibali *et al.*, *Pulmonary nodule classification with deep residual networks*, *Int J CARS* **12** 2017

[37] T. Dozat, *Incorporating Nesterov Momentum into Adam ICLR* 2016