ToraxIA: Virtual Assistant for Radiologists Based on Deep Learning from Chest X-Ray

Martin Carnier¹, Raul Albertti¹, Liliana Gavidia², Erika Severeyn³, and Alexandra La Cruz¹

¹ Alumbra.ai™, Santiago, Chile
{martincarnier,raulalbertti}@alumbra.ai
² GBBA, Grupo de Bioingeniería y Biofísica Aplicada, Caracas, Venezuela
lgavidia@usb.ve
³ Department of Thermodynamics and Transfer Phenomena, Universidad Simón Bolívar, Caracas, Venezuela
severeynerika@usb.ve
⁴ Faculty of Engineering, Universidad de Ibagué, Ibagué, Tolima, Colombia
alexandra.lacruz@unibague.edu.co
https://www.alumbra.ai

Abstract. Misdiagnosis of pulmonary pathology may have several causes. Some of them are related to the unavailability of radiology specialists, the increasingly overwhelming number of Chest X-Ray images generated daily, or the respiratory disease by itself. Consequently, it might lead to radiologist burnout. Furthermore, an unusual location of pathological lung signs may lead to a misdiagnosis of severe pulmonary disease. Several pathological signs are observed from a Chest X-Ray and on different tissues: Heart and Pericardium, Hil, Mediastinum, Lungs, Pleura, and Chest wall. In the present study, a Deep Learning model is presented, which identifies several tissues and abnormal signs that may require a deeper evaluation. The model was trained using a total of 240,000 Chest X-Ray images, which can automatically identify several structures observed on a Chest X-Ray and some pathological signs of respiratory disease. The model reached 97% of accuracy. These results suggest that our model can be used as an automatic identifier of tissues and pathological signs from Chest X-Ray images. In consequence, it constitutes a meaningful and essential tool for preliminary studies of the severe respiratory disease since it can potentially decrease the number of false positives in the most common radiological examination ordered for patient evaluation. It also could be used for screening purposes to detect findings that may be useful in any pathology assessment, and in the decision-making process of conducting further image modality evaluations, such as CT.

Keywords: Severe respiratory disease · Artificial intelligence · Medical validation · Convolutional neural networks · Deep learning
1 Introduction

Chest X-ray evaluation remains the primary tool by excellence for the screening of respiratory disease [4, 23]. Therefore, it should be readily available to all patients with respiratory symptoms. The diagnosis of respiratory diseases requires specialized training for proper evaluation. Currently, the number of specialists evaluating respiratory disease is insufficient, and the few available are generally suffering fatigue due to excessive work. Besides, there is a highly increasing incidence of lung disease due to different unfavorable environmental factors and particularly under the COVID-19 Coronavirus pandemic. On the other hand, the overwhelming amount of clinical data and radiological images causes medical decision-making in uncertain conditions, which in turn leads to inaccuracy in medicine due to burnout in radiologists [6]. This seriously compromises patient safety and might shorten life expectancy, as well as increase health costs. Besides, it produces fatigue of the radiologist [6, 17], and delays in report generation to proper diagnosis and treatment. In addition, the increasing demand for radiological services has been worsened with the COVID-19 pandemic.

Convolutional neural networks (CNN) have become an increasingly popular and effective method for analyzing diagnostic images [22]. In general, machine learning techniques have been used in the diagnosis of lung diseases. Particularly, the convolutional neural networks have been used in the diagnosis of tuberculosis, obtaining satisfactory results in the classification of subjects with tuberculosis with a sensitivity of 99% and specificity of 100% [12]. CNN has also been used in the classification of subjects with pneumonia [20] and with viral pneumonia due to COVID-19 [3]. CNN generally takes a set of training images, learns meaningful features from the data, and optimizes a set of parameters such as weights and bias to be able to distinguish and classify the data by their type [19]. Diagnostic image evaluation of respiratory disease is commonly carried out directly by an expert who observes if there are irregular patterns that do not exist in healthy lungs. This process consists of a visual recognition, based on the experiences of the specialist, the presence of pathological patterns suggestive of lung disease, and the comparison with the form, proportion, location, and definition of the visible organs considered as healthy. Since the information is unstructured, it is not possible to perform its analysis using traditional statistical models. As it is well known, CNN has become actively used for solving this kind of issue efficiently. In [21], several CNN were implemented and compared for differentiation between normal and abnormal frontal chest radiographs, including normal versus lung opacity classification, with best-reported values of AUC of 0.9804 ± 0.0032 (with an accuracy of 94.71 ± 0.32%, a sensitivity of 92.20 ± 0.34% and a specificity of 96.34 ± 0.31%).

This study proposes an application using a convolutional neural network model to operate as a predictive model for diagnosis support as a virtual assistant. This model allows for the detection of pathological signs suggestive of respiratory diseases from Chest X-rays. Some of the pathological signs detected by the application are pleural effusion, interstitial pattern, consolidation,
atelectasis, pneumothorax, cardiomegaly, and nodule-mass. This model of artificial intelligence (AI) provides seamless integration with the pre-existing workflow without disrupting the regular procedure of the physicians, enhancing the specialized care that sometimes cannot always become easily accessible due to the limited availability of radiologists. The highly audited algorithm presented will help demystify the misleading or erroneous information physicians sometimes have about AI as a black-box, or that it constitutes a menace to their job. It is important to emphasize that AI will not replace radiologists. Radiologists who use AI will replace radiologists who do not. It is the opposite. Toraxia is a tool designed to be used as a virtual assistant that provides augmented radiology, making medical expertise transcend to function as an extension of themselves.

2 Methodology

All datasets used are available in the following public databases: PadChest [5], Chexpert [9], and SIIM-ACR [14]. The model was trained in two phases. The first training phase used 200,000 Chest X-Ray images from the open databases mentioned before. The second training phase considered COVID-19 cases, then 40,000 images were added, from the database MIMIC [10] and some others were from several health institutions from Latin America. Some cases where patients with COVID-19. In the first training phase, the result was not enough for localizing the area of the image where the pathological sign needed to be identified. We realized there were some mismatches in the annotations. Consequently, several radiologists and medical specialists were invited to participate in a manual annotation process, to audit the data that would be used in the second phase of the training. The dataset annotation process is explained in Sect. 2.2, while the model is described in Sect. 2.3.

2.1 Data Set

At an early stage, full datasets from PadChest, SIIM-ACR, and Chexpert, were used to train the model with 200,000 images. In the second stage, the database MIMIC was added to the training data set. In addition, other Chest x-ray images from local healthcare institutions in Chile, Venezuela, Colombia, Nicaragua, and Ecuador were added (see Table 1). A total of 240,000 images were used for training, validation, and testing: 70–73% of them were used for training, 20% for model validation, and 7–10% for testing (see Table 2). All these data were previously annotated by the expertise of a group of medical specialists and radiologists, in addition to the available annotations of the open databases already mentioned. The reports of the radiologists and the clinical history of the patients were used as well. The dataset was classified and annotated by three experts into the following classes: Interstitial pattern, Consolidation, Atelectasis, Pneumothorax, Pleural Effusion, Cardiomegaly, and Nodule/mass (see Table 3).
Table 1. Dataset source and annotation

| Source of data                      | Image quantity | Annotation          |
|-------------------------------------|----------------|---------------------|
| Health centers from Chile, Venezuela, Colombia, Nicaragua, and Ecuador | 10,000          | By experts          |
| PadChest                            | 120,000        | Already annotated   |
| CheXpert                            | 35,000         | Already annotated   |
| MIMIC                               | 30,000         | Already annotated   |
| SIIM-ACR [1] https://www.kaggle.com/c/siim-acr-pneumothorax-segmentation | 45,000         | Already annotated   |

Table 2. Dataset distribution per pathological sign for training, validation, and testing, used in the first and second training phases

| Pathological signs                  | First training phase | Second training phase |
|-------------------------------------|----------------------|-----------------------|
|                                     | Total used | Training | Validation | Testing | Dataset used | Training | Validation | Testing |
| Interstitial lung pattern           | 43,000     | 30,100   | 8,600      | 4,300   | 8,000       | 5,840    | 1,600      | 564     |
| Consolidation pattern               | 34,000     | 23,800   | 6,800      | 3,400   | 8,000       | 5,840    | 1,600      | 564     |
| Atelectasis                         | 15,000     | 10,500   | 3,000      | 1,500   | 4,000       | 2,920    | 800        | 282     |
| Pneumothorax                        | 45,000     | 31,500   | 9,000      | 4,500   | 4,000       | 2,920    | 800        | 282     |
| Effusion                            | 15,000     | 10,500   | 3,000      | 1,500   | 6,000       | 4,380    | 1,200      | 423     |
| Cardiomegaly                        | 30,000     | 21,000   | 6,000      | 3,000   | 6,000       | 4,380    | 1,200      | 423     |
| Nodulo/mass                         | 18,000     | 12,600   | 3,600      | 1,800   | 4,000       | 2,920    | 800        | 282     |
| **Total**                           | **200,000** | **140,000** | **40,000** | **20,000** | **40,000** | **29,200** | **8,000** | **2,818** |
|                                     | **70%**    | **20%**  | **10%**    |          | **73%**    | **20%**  | **7%**     |          |
2.2 Annotation Process

Datasets from PadChest, Chexpert, MIMIC, and SIIM-ACR are already annotated. These annotations are regarded as our first expert. Several institutions and collaborators have provided collections of images with the respective radiology reports (see Acknowledgement section). This big dataset not annotated was then analyzed in a blinded study by several radiology experts. A group of two radiology experts was formed and each of them visualized a set of images without prior knowledge of the radiology report nor the classification. A set of images was then classified and annotated by at least three experts: two groups of specialists and the open annotated databases (PadChest, MIMIC, Chexpert, and SIIM-ACR) considered as the third expert. A total of twenty (20) collaborators between radiologists and specialists in the area of diagnosis of respiratory diseases participated in the annotation process. All of them between five (5), ten (10), and (15) years of experience analyzing Chest X-Ray and identifying pathological signs, and from different countries in Latin America.

The segmentation-based training process requires a mask annotation in addition to the previously available classification. This is why the annotation consisted of creating a mask that covered the pathological sign identified in the image. In addition, due to the existence of high variance resulting from these indirect classification methods, it was necessary to perform an audit process in order to select the gold standard dataset for testing. This process was carried out by physicians and radiologists. The image database, and its previous classification, were made available by a web site from Toraxia, so that the radiologist collaborators could access them for the auditing process.

Each image in the database was analyzed by two physicians. For them, the findings identified by at least two of the classifying elements (pre-sorting, physician 1, physician 2) have been determined to be definitive. In cases of positive signs, the final mask for each image and sign has been determined as the intersection of the masks generated by the individuals participating in the annotation process, resulting in a number of masks per pathological sign as shown in Table 3.

Table 3. Segmented and annotated images with pathological signs, selected as gold standard for testing the model

| Pathological sign          | Number of segmented and annotated images | Number of images used |
|---------------------------|------------------------------------------|-----------------------|
| Interstitial lung pattern | 4,800                                    | 4,300                 |
| Consolidation pattern     | 3,750                                    | 3,400                 |
| Atelectasis               | 1,600                                    | 1,500                 |
| Pneumothorax              | 4,800                                    | 4,500                 |
| Effusion                  | 1,600                                    | 1,500                 |
| Cardiomegaly              | 3,700                                    | 3,000                 |
| Nodule/mass               | 1,800                                    | 1,800                 |
2.3 Deep Learning Model

U-Net [18] is a convolutional neural network used for biomedical image segmentation, that has been successfully used for semantic segmentation, and shown to exhibit great performance on medical images. It comprises encoding and a decoding stage. The encoder consists of a set of blocks that perform a max-pooling operation, in which the resolution of the image is halved, and the number of channels is doubled. The decoder encompasses transposed convolution blocks that are now used to reverse the process. A distinctive feature is that each reduction stage is concatenated with its equivalent in maximization, managing to maintain the spatial information that would otherwise be lost.

Res-Net [8] is a neural network architecture inspired by the pyramidal cells of the cerebral cortex that make use of shortcuts so that the layers can jump over others, as a solution to the problem of gradient vanishing. In this model, the binary cross-entropy function, a loss function specialized for binary classification problems, is often used and shown to be very effective for balanced classes. However, in our case, the level of balancing is not available for the location of pathological signs, so the use of this function was not encouraged. Therefore, the Dice loss function [26] was used instead, since it maximizes the values for the least present classes and faces the problem of imbalanced datasets.

The U-Net architecture allows us to integrate a sub-network into the encoder/decoder process [25]. In order to assist in the training process, several architectures can be implemented in specific segments of the model. For this use case, Res-Net18, Res-Net50, and VGG11 (AlbuNet) were considered for integration to U-Net. The following results were obtained for the detection of signs of interstitial and consolidation patterns during the training, validation, and evaluation process, with the following accuracy for each architecture (see Table 4). These results lead to consider Res-Net50 as the best coder option in this study.

Table 4. Results of testing different network architectures integrated to U-Net

| Model     | Parameters | Validation process (Accuracy) | Evaluation process (Accuracy) |
|-----------|------------|-------------------------------|------------------------------|
| Res-Net18 | 15M        | 0.83                          | 0.85                         |
| Albu-Net  | 34M        | 0.78                          | 0.79                         |
| Res-Net50 | 34M        | 0.88                          | 0.93                         |

The information system training and development process was carried out with all these implemented components, as detailed in the following process. To prepare the image training set, images are first selected by class, then a manual segmentation is carried out and a mask is generated. This mask is used to guide the semantic segmentation in our learning model. On the other hand, the training process is performed in three steps: the first one involves the first training with a learning rate (Lr) equal to 0.0001; then a second training with a smaller Lr is
done. Finally, a validation process is carried out to estimate the model precision. Results from our model U-Net using Res-Net50 as a coder is shown in Table 5. The accuracy values obtained for segmentation per pathological sign are shown. These values are obtained from the Dice Loss Function [26] used for the coder in the validation process of the model. The results suggest that currently, our model is still not a very good identifier of Nodule/mass (0.69 of accuracy), as can be observed in Table 5.

| Pathological sign       | Dice loss function |
|-------------------------|--------------------|
| Interstitial lung pattern | 0.87               |
| Consolidation           | 0.81               |
| Atelectasis             | 0.72               |
| Pneumothorax            | 0.84               |
| Pleural effusion        | 0.84               |
| Cardiomegaly            | 0.94               |
| Nodule/mass             | 0.69               |

3 Experimental Results

3.1 Library Performance

Throughout the implementation phase of the proposed model, several combinations of platforms and hardware were evaluated for optimal training and testing process performance. Initially, the model was developed using Tensorflow [1] with Keras [7], a high-level library, on a Python Anaconda environment. The model was also developed using Pytorch library [16], and performance compared with each other. Table 6 presents the accuracy results obtained from each ResBlock in the validation and the evaluation (testing process). Results in Table 6 showed that Pytorch constituted the best selection to carry out the deep learning tasks related to the project and deployment to production. In the beginning, it was planned to use Ubuntu as the operating system to carry out the development and deployment of the project, but after a documentary review of the support offered by Nvidia for both operating systems and the hardware capabilities in terms of the driver options for each platform, it was concluded that Windows Server 2016 in an IBM cloud computing service was the most stable platform for the installation of all the required dependencies necessary for the deployment of the comprehensive system.
Table 6. Comparison of Keras/Tensorflow vs Pytorch processing

| Process            | CPU (AMD Ryzen 2700X) | GPU (Nvidia Tesla K80) | CPU/GPU Maximum consumption |
|--------------------|-----------------------|------------------------|-----------------------------|
| Training with Keras/Tensorflow | 86 h                  | 4 h                    | 100%/70%                    |
| Testing Keras/Tensorflow   | 45 s                  | 8 s                    | 100%/70%                    |
| Training with Pytorch    | 170 h                 | 2.5 h                  | 100%/100%                   |
| Testing Pytorch          | 190 s                 | 2.5 s                  | 100%/100%                   |

3.2 Deep Learning Model Results

Table 7 shows the confusion matrix obtained from the first and the second training phase. The first training phase used 200,000 images for training, where 10% (20,000) were used for testing the model. The second training phase used 40,000 images for training, where 7% (20,000) were used for testing the model. Several statistical metrics were calculated from the confusion matrix for the first and the second training phase, respectively (see Table 8).

Table 7. The confusion matrix obtained in the first training phase with 20,000 images for testing. The second Confusion matrix is from the second training phase where there are two (2) classes: Positive, for pathological sign detection suggestive of COVID-19, Negative for no COVID-19 signs detected

| First training phase | Second training phase |
|----------------------|-----------------------|
| Predicted | True | Predicted          | True |
| Positive | 8,788 | 1,606   | Positive | 266  | 43   |
| Negative | 1,512 | 8,094 | Negative | 30   | 2,479 |

Table 8. Statistical metrics obtained from the first and the second training phases

| Statistical metric     | First training phase | Second training phase |
|------------------------|----------------------|-----------------------|
| Accuracy               | 0.8441               | 0.9741                |
| Precision              | 0.8455               | 0.8608                |
| Sensitivity (Recall)   | 0.8986               | 0.8986                |
| Specificity            | 0.8384               | 0.9830                |
| Negative predicted value | 0.8426          | 0.9880                |
| F1                     | 0.8493               | 0.8793                |

As we can observe in Table 8, all values, including Accuracy, Precision, Recall, Negative Predicted Value, and F1, in both phases of training reached an excellent
value over 0.8, and the Accuracy reached a value over 0.9 in the second training phase, which are good indicators of an excellent CNN model. Besides, every pathological sign identified can be visible and available to the specialists for deeper analysis.

Figure 1 and 2 show the results obtained from four (4) patients evaluated using Toraxia™. Figure 1 shows the patients previously and randomly selected. Each image is passed through our model, and consequently, the model generates a color map image for each pathological sign detected. In Fig. 2, every row corresponds to a pathological sign identified for each patient. The signs found are listed in the following order, from row one to row six: effusion, atelectasis, consolidation, cardiomegaly, pneumothorax, and interstitial pattern, and reported for each patient, which correspond to the pathological signs better classified by our model. An example of a report generated by Toraxia™ is shown in Fig. 3. Colors are also a codification of the severity of the signs, and this represents a piece of important visual information for the radiologist. However, this is still under study. Currently, there are no official reports or in-depth studies about it.

![Fig. 1. Chest X-Ray of 4 patients with different pathological signs, which are shown in Fig. 2 as a result from Toraxia™](image)

### 3.3 COVID-19 Dataset

Experimentally, our model was evaluated for COVID-19 disease, then, a dataset of patients with COVID-19 was introduced to Toraxia™. A patient with COVID-19 typically presents an interstitial pattern or consolidation sign [24]. Therefore, we trained our model in order to detect both signs. This training was performed in the second training phase (see Table 2). For this, the data were separated into two groups, one with COVID-19 images and the other with non-COVID-19 images. For our model, a suggested positive patient presents an interstitial pattern or consolidation sign and a negative otherwise. The results of this binary classification are shown in Table 8 and the corresponding confusion matrix in Table 7. The area under the ROC curve gave a value of $\text{ROC}_{AUC} = 0.99$. In this case in particular, the binary cross-entropy loss function was used for the validation process, since the classification made was binary.
Fig. 2. Each column corresponds to a pathological sign identified as a result of applying Toraxia™ over each patient from Fig. 1
3.4 Toraxia™ Output

Toraxia™ automatically generates a report to radiologists and physician specialists (see Fig. 3). The report is divided into 4 sections (each section is indicated into a circle in Fig. 3): (1) the demographic data from the patient; (2) a detailed % of the area covered by the pathological sign in the left and the right lung; (3) an estimation of the detection accuracy, and finally in section (4), the color map of the signs identified in the image.

Fig. 3. Toraxia™ report generated for one patient. The study result is obtained in 3 s.
4 Discussion

The method applied in this work exhibit excellent results for the detection of pathological signs from Chest X-rays, with a high probability of detecting subjects with and without the disease, as can be seen in the high values of positive predictive value and negative predictive value. The method yields a few false positives and false negatives. These results corroborate the validity of using convolutional neural networks for the diagnoses of pulmonary disease, as suggested by other studies that reference the pertinent use of CNN for the detection of pathologies such as Tuberculosis and pneumonia [2,13,15]. A comparative evaluation, validation, and training based on different cost functions has resulted in Dice loss function to be the optimal cost function for detection of pathological signs in anteroposterior chest radiographs, and it has proven to be suitable for optimization of the model with unbalanced classes typically representing pathological signs that, when present, occupy a specific area of the image.

The highest precision indexes during the training and evaluation process were achieved with the Res-Net50 codifier, thereby continuing its use on the present operation of the system. This type of architecture has been used in the training of neural networks for the diagnosis of cancer and it has turned out to be the most suitable for these cases, which corroborates what was found in this research [11]. The use of segmented images for the training of the model resulted in a better orientation for model classification and localization of pathological signs, enabling optimal performance of the algorithm while requiring a reduced number of images. This facilitates the manual evaluation and audit process, encouraging the participation of healthcare professionals in the calibration of the system.

The Pytorch library resulted in the best for this implementation, regarding response time and system resource optimization on the hardware level. This enables expansion of the model to include additional functions without significantly affecting its performance on assistance in image analysis with an added value proposition, thus maintaining timely execution parallel to the workflow execution in today’s healthcare environment.

5 Conclusion and Future Work

Toraxia potentially constitutes an excellent virtual radiology assistant for automatic detection of pathological signs suggestive of respiratory disease, reducing the diagnostic error, increasing efficiency, and quality of care. It also reduces the burnout radiologists sometimes experience, when immersed in environments with limited availability of radiologists or highly increasing demand of radiology reporting, since it can provide precise quantitative information that the radiologist cannot usually calculate himself due to time constraints and the need for urgency evaluation. This information is derived from the automatic delineation of affected areas. In other words, it can reduce the time spent in identification, and characterization, which gains the radiologist valuable time for a more accurate interpretation and diagnosis. This represents a contribution to the degree
of ground-glass opacity progression of COVID-19 disease. Toraxia can assist the radiologist to expand coverage of practice without losing diagnostic quality, and ultimately assist in the expansion of radiology services. Using U-Net with ResNet50 offered the optimal combination for the detection of pathological signs through semantic segmentation. This implementation has demonstrated that it is possible to continuously include medical criteria in the calibration process of the system. Based on this, a more in-depth evaluation is currently undergoing targeting calibration and localization of the algorithm, considering phenotypes and particular characteristics of populations, local medical criteria, and local radiology services in Latin American countries.

Currently, Toraxia is being tested by several health institutions in Latin America. Furthermore, we are working on adding semantic information to our dataset, to create a knowledge graph and on top of it, apply a Machine learning process for better support and results.

Acknowledgments. Full acknowledgement is given to Alumbra.aiTM, CULDATE, Allm.net, IBM, IBM Cloud, InterSystemsTM, CORFO Openbauchef. Universidad de Chile, Centro Nacional en Sistemas de Información en Salud Chile, Inacap, Smart Salud, AGFATM. Special thanks to the following collaborators and their institutions, who either provided images to the training phase of the algorithm, time and effort segmenting and reviewing images used or are committed to the multinational study for calibration and localization of the algorithm considering phenotypes, local medical criteria and radiology services in Latin America. In Venezuela: Dr. Ivonne Rivas Gavidia. Radiologist. Instituto Clínico La Florida. President of Sociedad Venezolana de Radiología 2018–2020 (SOVERADI). Dr. Sergio Tovar. Radiologist. Vice President of the Board of Directors. Hospital de Clínicas Caracas. Dr. Oswaldo Ramos Núñez. Radiologist. Diagnostic Imaging Director. Centro Clínico La Sagrada Familia. President of Asociación Venezolana de Informática en Salud. Dr. Bernardo Lander. Radiologist. Chief of Radiology Service. Centro Médico de Caracas. Dr. Miguel Rocha and Dr María Elena Savino. Radiologists. Instituto Médico La Floresta, Dr. Anamel García. Radiologist. Instituto Clínico Unare. Dr. Eleonora Garcia Hernandez and Dr. Lisseth Diaz Liendo. Radiologists. Centro Médico de Caracas. Dr. Angel Riera García. Radiologist. Policlínica Metropolitana. Dr. Marialejandra Mendoza. Pulmonologist. Hospital Universitario de Los Andes. Researcher from the Department of Medicine. Universidad de Los Andes. Dr. Daniel Foldats. Internal Medicine. Inpatient Care Coordinator. Centro Médico Docente La Trinidad. In Colombia: Radiologists in Fundación Santa Fe de Bogotá: Dr. Javier Romero. President; Dr. María José Veloza and Dr. María Pinzón. In Hospital Universitario San Ignacio de Bogotá: Dr. Julio César Castellanos Ramírez. Specialist in Public Health and Healthcare Administration. Director. Dr. Margarita Manrique Andrade. Research Director. Dr. Carolina Gómez Rodríguez. Director of the Radiology Department. Dr. Ariel Emilio Martínez Martínez Researcher from Pontificia Javeriana de Bogotá. Dr. Nicolás Pérez Almanza. Public Health Specialist. Novatec Colombia. In Peru: Dr. Gustavo Adolfo Rondón Mendoza. Radiologist. Lima. in Chile: Benjamin Pizarro Galleguillos. PhD Candidate in Science and Radiology Specialty. Universidad de Chile.
References

1. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X.: TensorFlow: Large-scale machine learning on heterogeneous systems (2015), http://tensorflow.org/, software available from tensorflow.org

2. Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christie, A., Mougiakakou, S.: Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans. Med. Imaging 35(5), 1207–1216 (2016)

3. Apostolopoulos, I.D., Mpesiana, T.A.: Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys. Eng. Sci. Med. 43, 635–640 (2020)

4. Bordley, W.C., Viswanathan, M., King, V.J., Sutton, S.F., Jackman, A.M., Sterling, L., Lohr, K.N.: Diagnosis and testing in bronchiolitis: a systematic review. Arch. Pediatr. Adolesc. Med. 158(2), 119–126 (2004)

5. Bustos, A., Pertusa, A., Salinas, J.M., de la Iglesia-Vayá, M.: Padchest: a large chest x-ray image dataset with multi-label annotated reports. Med. Image Anal. 66, 101797 (2020). https://doi.org/10.1016/j.media.2020.101797, http://www.sciencedirect.com/science/article/pii/S1361841520301614

6. Chetlen, A.L., Chan, T.L., Ballard, D.H., Frigini, L.A., Hildebrand, A., Kim, S., Brian, J.M., Krupinski, E.A., Ganeshan, D.: Addressing burnout in radiologists. Acad. Radiol. 26(4), 526–533 (2019). https://doi.org/10.1016/j.acra.2018.07.001, http://www.sciencedirect.com/science/article/pii/S1076633218303416

7. Gulli, A., Pal, S.: Deep Learning with Keras. Packt Publishing Ltd., Birmingham (2017)

8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 (2016)

9. Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghgoo, B., Ball, R., Shpanskaya, K., Seekins, J., Mong, D., Halabi, S., Sandberg, J., Jones, R., Larson, D., Langlotz, C., Patel, B., Lungren, M., Ng, A.: Chexpert: a large chest radiograph dataset with uncertainty labels and expert comparison (2019)

10. Johnson, A.E., Pollard, T.J., Shen, L., Li-Wei, L.H., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L.A., Mark, R.G.: Mimic-iii, a freely accessible critical care database. Sci. Data 3(160035) (2016). https://doi.org/10.1038/sdata.2016.35

11. Kocher, M., Ruge, M.I., Galldiks, N., Lohmann, P.: Applications of radiomics and machine learning for radiotherapy of malignant brain tumors. Strahlentherapie und Onkologie: Organ der Deutschen Rontgengesellschaft (2020)

12. Lakhani, P., Sundaram, B.: Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 284(2), 574–582 (2017)

13. Liu, C., Cao, Y., Alcantara, M., Liu, B., Brunette, M., Peinado, J., Curioso, W.: Tx-cnn: detecting tuberculosis in chest x-ray images using convolutional neural network. In: 2017 IEEE International Conference on Image Processing (ICIP), pp. 2314–2318. IEEE (2017)
14. Norweck, J.T., Seibert, J.A., Andriole, K.P., Clunie, D.A., Curran, B.H., Flynn, M.J., Krupinski, E., Lieto, R.P., Peck, D.J., Mian, T.A., Wyatt, M.: Acr-napmssim technical standard for electronic practice of medical imaging. J Dig. Imaging 26, 38–52 (2013). https://doi.org/10.1007/s10278-012-9522-2
15. Pasa, F., Golkov, V., Pfeiffer, F., Cremers, D., Pfeiffer, D.: Efficient deep network architectures for fast chest x-ray tuberculosis screening and visualization. Sci. Rep. 9(1), 1–9 (2019)
16. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S.: Pytorch: an imperative style, high-performance deep learning library. In: Wallach, H., Larochelle, H., Beygelzimer, A., Alché-Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems, vol. 32, pp. 8024–8035. Curran Associates, Inc. (2019). http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf
17. Reiner, B.I., Krupinski, E.: The insidious problem of fatigue in medical imaging practice. J. Dig. Imaging 25(1), 3–6 (2012)
18. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). LNCS, vol. 9351, pp. 234–241. Springer, Heidelberg (2015). http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a, (available on arXiv:1505.04597 [cs.CV])
19. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR abs/1409.1556 (2015)
20. Sirazitdinov, I., Kholiavchenko, M., Mustafaev, T., Yixuan, Y., Kuleev, R., Ibragimov, B.: Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. Comput. Electric. Eng. 78, 388–399 (2019)
21. Tang, Y., Peng, Y., Yan, k., Bagheri, M., Redd, B., Brandon, C.J., Lu, Z., Han, M., Xiao, J., Summers, R.M.: Automated abnormality classification of chest radiographs using deep convolutional neural networks. npj Digit. Med. 70(3) (2020). https://doi.org/10.1038/s41746-020-0273-z
22. Taylor, E., Haven, K., Reed, P., Bissielo, A., Harvey, D., McArthur, C., Bringans, C., Freundlich, S., H., I.R.J., Perry, D., F., W., Milne, D., Modahl, L., Huang, Q.S., Gross, D., Widdowson, M.A., Grant, C.: A chest radiograph scoring system in patients with severe acute respiratory infection: a validation study. BMC Med. Imaging 15(61) (2015). https://doi.org/10.1186/s12880-015-0103-y
23. Van Cleef, M., Kivihiya-Ndugga, L., Meme, H., Odhiambo, J., Klatser, P.: The role and performance of chest x-ray for the diagnosis of tuberculosis: a cost-effectiveness analysis in nairobi, kenya. BMC Infect. Dis. 5(1), 1–9 (2005)
24. Vancheri, S.G., Savietto, G., Ballati, F., Maggi, A., Canino, C., Bortolotto, C., Valentini, A., Dore, R., Stella, G.M., Corsico, A.G., Iotti, G.A., Mojoli, F., Perlini, S., Bruno, R., Preda, L.: Radiographic findings in 240 patients with covid-19 pneumonia: time-dependence after the onset of symptoms. Eur. Radiol. 30, 1–9 (2020). https://doi.org/10.1007/s00330-020-06967-7
25. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: Unet++: redesigning skip connections to exploit multiscale features in image segmentation. IEEE Trans. Med. Imaging 39(6), 1856–1867 (2020)
26. Zou, K.H., Warfield, S.K., Bharatha, A., Tempany, C.M.C., Kaus, M.R., Haker, S.J., William, M.W., Jolesz, F.A., Kikinis, R.: Statistical validation of image segmentation quality based on a spatial overlap index. Acad. Radiol. 11(2), 178–189 (2004). https://doi.org/10.1016/s1076-6332(03)00671-8