

Original paper

An improvement of the CNN-XGboost model for pneumonia disease classification

Yousra Hedhoud^{1,A,B,C,D,E,F}, Tahar Mekhaznia^{1,2,A,C,D,E}, Mohamed Amroune^{2,A,C,D}

¹Tebessi University, Tebessa, Algeria

²LAMIS Laboratory, Cheikh Larbi Tebessi University, Tebessa, Algeria

Abstract

Purpose: X-ray images are viewed as a vital component in emergency diagnosis. They are often used by deep learning applications for disease prediction, especially for thoracic pathologies. Pneumonia, a fatal thoracic disease induced by bacteria or viruses, generates a pleural effusion where fluids are accumulated inside lungs, leading to breathing difficulty. The utilization of X-ray imaging for pneumonia detection offers several advantages over other modalities such as computed tomography scans or magnetic resonance imaging. X-rays provide a cost-effective and easily accessible method for screening and diagnosing pneumonia, allowing for quicker assessment and timely intervention. However, interpretation of chest X-ray images depends on the radiologist's competency. Within this study, we aim to suggest new elements leading to good interpretation of chest X-ray images for pneumonia detection, especially for distinguishing between viral and bacterial pneumonia.

Material and methods: We proposed an interpretation model based on convolutional neural networks (CNNs) and extreme gradient boosting (XGboost) for pneumonia classification. The experimental study is processed through various scenarios, using Python as a programming language and a public database obtained from Guangzhou Women and Children's Medical Centre.

Results: The results demonstrate an acceptable accuracy of 87% within a mere 7 seconds, thereby endorsing its effectiveness compared to similar existing works.

Conclusions: Our study provides a model based on CNN and XGboost to classify images of viral and bacterial pneumonia. The work is a challenging task due to the lack of appropriate data. The experimental process allows a better accuracy of 87%, a specificity of 89%, and a sensitivity of 85%.

Key words: classification, viral pneumonia, convolutional neural network, chest X-ray, bacterial pneumonia, extreme gradient boosting.

Introduction

Pneumonia is a kind of respiratory infection caused by bacteria, viruses, or fungi in lungs, with symptoms ranging from mild to severe including fever, nausea, phlegm, cough, and chest pain [1]. Pneumonia is viewed as a serious illness, especially against children less than 5 years old. The World Health Organization (WHO) affirms that

pneumonia caused the deaths of more than 800,000 children in 2017 and in 2018, a child's death occurs every 39 seconds. It is estimated that 11 million children will die from pneumonia by 2030. South Asia has the highest number of deaths, where India recording more than 158,176 deaths in 2016 and 808,694 in 2018 [2]. For this reason, early diagnosis seems vital for ensuring effective care and increasing survival rates.

Correspondence address:

Yousra Hedhoud, Tebessi University, Tebessa, Algeria, e-mail: yousra.hadhoud@univ-tebessa.dz

Authors' contribution:

A Study design · B Data collection · C Statistical analysis · D Data interpretation · E Manuscript preparation · F Literature search · G Funds collection

Bacterial pneumonia is a kind of pneumonia infection caused by various external factors such as inadequate nutrition, old age, excessive coldness, or complications of various other viruses such as seasonal influenza or COVID-19. First aid consists of antibiotics; however, supportive care is required for viral pneumonia [3,4].

Pneumonia is detected by a blood or sputum analysis, but in practice, tests based on imaging interpretation are the most requested [5]. They comprise computed tomography of the lungs (CT), ultrasound, magnetic resonance imaging (MRI), and X-ray of the chest (CXR). The CXR are images produced by X-ray beams after passing through the human body; the light areas within image show parts of dense tissue, while the dark areas denote regions with less compact tissue. Despite such interpretation, many factors influence the identification of an anomaly in a given tissue, if one exists. Hence, misinterpretations occur in regions of superimposed structure when using inappropriate material or a bad viewing angle, or due to patient motion during exposure to X-rays. Furthermore, pneumonia symptoms share similar features with other chest diseases, leading to problems with image interpretation.

Given the significant differences in specialists' opinions in image interpretation, which decreases its professional acceptability area limits, it seems essential to alternate toward automatic solutions avoiding human errors and helping radiologists to diagnose patients more accurately. Within this context, various studies have been carried out, including some that utilised deep learning models (DL) for pneumonia detection and classification, which is still the most effective and extensively used method for diagnosing diseases in general, and radiology specifically. DL often refers to Artificial Neural Networks (ANNs), which consist of a set of algorithms inspired by the human brain system. They are based on a learning concept that enriches a knowledge database with the environment experience acquired through execution. Among the best-known ANNs architectures are the Convolutional Neural Networks (CNN); they outperform human capabilities in some visual recognition tasks, and this potential is applied in medicine to detect and classify diseases. It leads to accurate interpretation of pathologies and enhances the quality of the healthcare system [6].

In this paper, we present a comparative study of detection and classification of pneumonia disease using CXR radiography. The image processing is performed by a CNN model for extracting X-ray image features and the XGboost tool for classification. The objective of the work is to emphasize the distinction between CXRs of healthy bodies and those with viral or bacterial pneumonia. The experimental process would ensure an effective diagnosis of patients and help doctors to proceed to the appropriate treatment.

The paper is organized as follows: after a brief introduction in section 1, an overview of related works is high-

lighted in section 2; the proposed approach is described in section 3; and section 4 comprises further details of the approach and a detailed experimental study. The conclusion of the work and future prospects are presented in section 4.

Related works

The use of CNN for pneumonia detection has been the subject of various studies in recent years, due to its ability to solve complicated problems. In [5] the authors developed a cheXNet model based on a pretrained CNN; this model was performed using CXR images extracted from a dataset of various lung pathologies [7]. The model achieved an acceptable performance of about 76.8% regarding the area under the receiver operating characteristic (AUCROC) for the case of pneumonia. Jiang *et al.* [8] used variants of pretrained CNNs for pneumonia detection; the trained process is accomplished by freezing the bottom layers and fine tuning the top layers. The accuracy reaches 94.2%. Also, in [9], Kermany *et al.* applied transfer learning to train a CNN model in order to differentiate between normal, viral, or bacterial pneumonia CXR images using a binary classification, and the experimental results reached an accuracy of 92.8%. An approximately similar result is obtained by a shallow CNN model proposed by [10]. In [11], the authors proposed a computer-aided diagnosis (CAD) system to detect pneumonia using 3 classification methodologies including a hybrid CNN (VGG-16 and VGG-19); the system offered an accuracy of 98%. In [12], the authors proposed a DL model to detect and localize abnormalities in the CXR images, the related algorithm is based on 3 classifier options and 4 model detectors using the VinDr-CXR dataset, which is augmented with Albumentations [13] in order to solve the class imbalance; the model's performance measured approximately 29% in terms of mean average precision (MAP). Likewise, in [14] the authors proposed a set of 2 CNN models: Mask R-CNN and RetinaNet, for the prediction and localization of pneumonia in small areas of CXR images. The model was tested on a database from Kaggle pneumonia challenge [15], and it achieved a recall of 79.3%. In [16], the authors developed a deep learning model for classification of CXR images into consolidation (alveolar pneumonia) and non-consolidation. The system allows visualization of areas of interest (AOI), which is used to obtain the outcome, and it achieved a result of 92% in terms of AUC. In [17], the authors proposed an approach of CNN with 3 different optimizers (Adam, SGD, and RMSProp). Hidden layers from 1 to 5 were used in the design of the CNN models and trained without any transfer learning technique. As a result, they found that CNN with 4 hidden layers and SGD achieved the highest test accuracy of 91%. In [18], the authors proposed a fine-tuned pre-trained deep CNN to classify CXR images of various pneumonia diseases. The experimental results showed 93.3% overall performance in terms of accuracy.

In terms of analysis, the presented approaches showed an overall effective performance, whereas the majority have been addressed in the CXR images to only predict the presence of pneumonia without considering its nature (viral or bacterial). Also, these works used various databases such as ChexPert [19], Chest x-ray14 [7], RSNA [15], etc. which include samples with just ‘pneumonia’ and ‘normal’ labels, which cannot render more details about the infection type. To the best of our knowledge, the only dataset with ‘normal’, ‘viral’, and ‘bacterial’ labels can be found at the Guangzhou Women and Children’s Medical Centre dataset [20], which is mostly considered for binary classification [21]. We notice also that fewer studies have addressed the multi-class classification of pneumonia disease; this is due to the complexity of the task in terms of processing time. In [5], the authors extended the cheXNet model for the detection of all diseases available in the related dataset. The model’s performance varies between 73% and 93% based on the F1-score metric, which depends on the disease. In [22], the authors used a customized VGG-16 and achieved an accuracy of 91.8%, and in [23], the authors applied a multi-classification to distinguish between ‘COVID-19’, ‘viral pneumonia’, and ‘normal’ using a wide number of samples in which the experiment results reached an accuracy of 98%. In [24], the authors trained a CNN and proved that the accuracy of the model is inversely proportional to the number of classes in the output layer.

Proposed model

A CXR image of an infected person is characterized by the presence of white hazy patches. Significant fluid in the lungs denotes bacterial pneumonia. Such anomalies are, in both cases (bacterial and viral), observable as white spots, but without enough precision about the infection nature, which makes prediction more difficult. The proposed model allows us to deal with this issue. It consists of a diagnosis and classification of CXR images using the power of CNNs in feature extraction and the XGboost ability for classification. CNNs are renowned for their ability to handling unseen data based on their hierarchical feature extraction process. XGboost can improve the predictions given by CNN, and the model’s generalization abilities are improved by this combination.

Convolutional neural network

CNNs are a type of deep neural network that can be used in situations involving large numbers of unprocessed data. Based on its capability to extract features automatically, CNN is viewed as a powerful tool in the computer vision field. It uses images as input data and considers just local regions instead of the whole image area. This reduces widely the number of parameters and processing time.

CNNs are built based on 3 types of layer: (a) convolutional layers, which allow extraction of the input image features (lines, edges, etc.), (b) the pooling layers, which decreases the input map’s dimensionality and retracts only important information, and (c) the fully connected layers that unroll all the values into a single vector in order, leading to the final classification decision.

XGboost

Extreme Gradient Boosting (XGboost) is a machine learning technique that can resolve problems involving classification or regression predictive modelling. It was developed by T. Chen and C. Guestrin [25] for accurate implementation of gradient boosting [26]. It also employs sequentially constructed shallow decision trees to provide reliable results and a scalable training strategy that prevents overfitting.

The idea of boosting consists of generating a strong model by combining a number of weak models. The success of boosting algorithms in general is due to their capacity to handle model and variable selection during the fitting process. When compared to other gradient boosting implementations, XGboost is typically faster [27], and it enables column parallelism at several levels leading the distribution of workload among various CPU cores or threads.

Dataset

To deal with the classification, we used the dataset proposed by [9], available at [20]. It is built based on 3 folders (Train, Val, and Test); each one involves 2 subfolders (Pneumonia and Normal). The dataset encloses 5863 images of paediatric patients in the age range 1–5 years old, obtained from Guangzhou Women’s and Children’s Medical Centre. The distribution of images is shown in Figure 1. Initially, all CXR images were pretreated by removing low quality or unreadable scans to ensure a suitable control quality. Then, they were arranged and labelled for training process (Figure 2) by 2 expert physicians before being used to train the AI system.

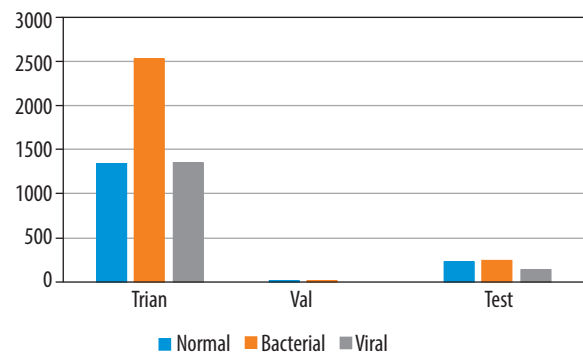


Figure 1. Dataset image distribution

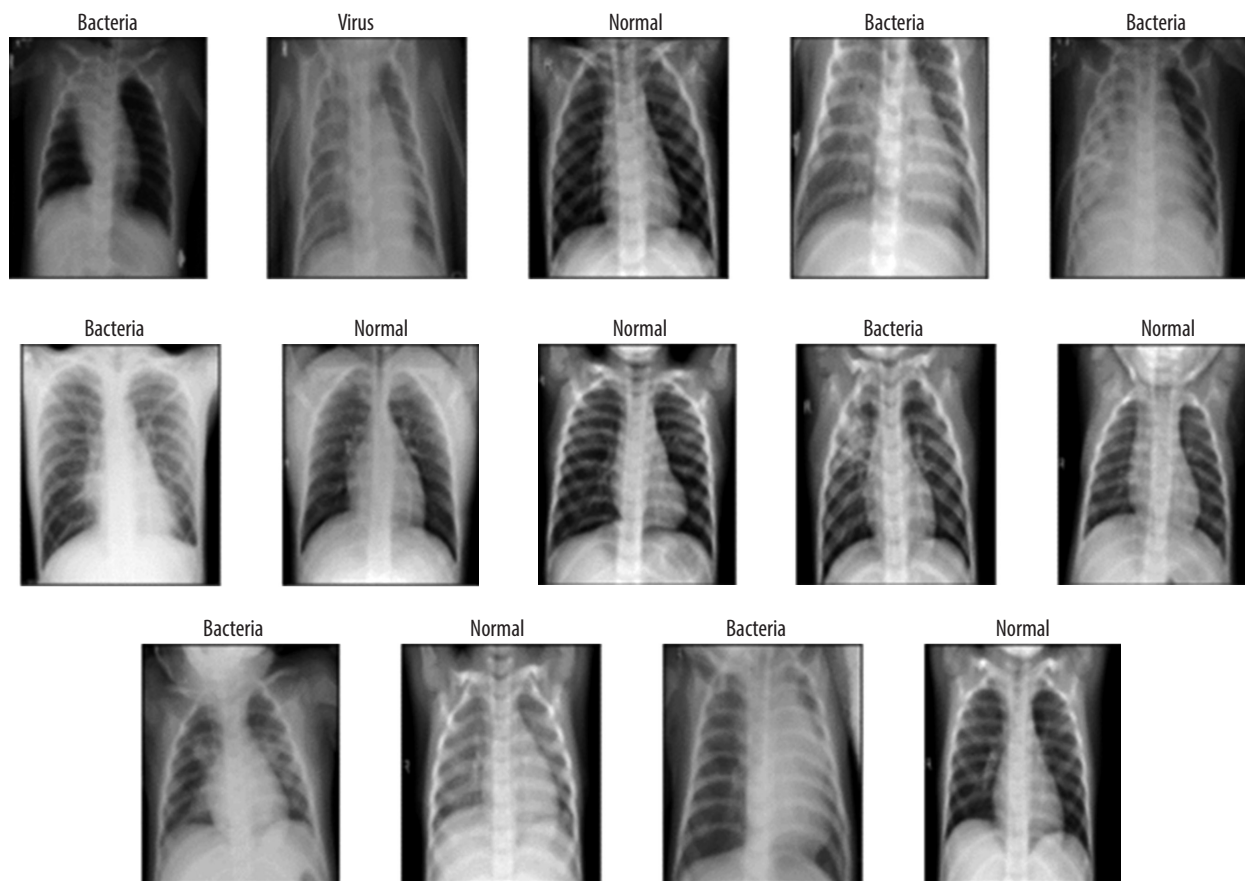


Figure 2. Various samples of the considered dataset

Data preprocessing and augmentation

It is evident that the power of CNN depends on the dataset size and its content quality. In another way, it seems difficult to acquire or to share an appropriate dataset in the medical field regarding ownership rights of such resources. For this reason, the proposed model makes use of data augmentation techniques to increase the volume of data. This is accomplished by managing the imbalance of data, hence reducing the overfitting. Unlike classic augmentation approaches, which are limited to simple geometric transformations, we used the Albumentations tool, which allows image transformation operations [28].

Using this tool, a operations transform pipeline was applied: horizontal flip, random degree rotation, and ran-

dom Gaussian noise (Figure 3). The first and second operations introduce a major transformation while keeping the underlying data that does not impair the diagnosis and analysis, allowing the model to generalize to unseen data more effectively, whereas the Gaussian noise helps the model to be more robust to noise and artifacts that might be present in the data by focusing on the underlying features rather than memorizing specific pixel values [29].

Table 1 shows the parameters used in these operations. This augmentation was applied after resizing all the images to 64×64 pixels and converting them from BGR format to RGB format, followed by a normalization that consists of collapsing the inputs to the range 0 and 1 in order to enhance the performance of the model.

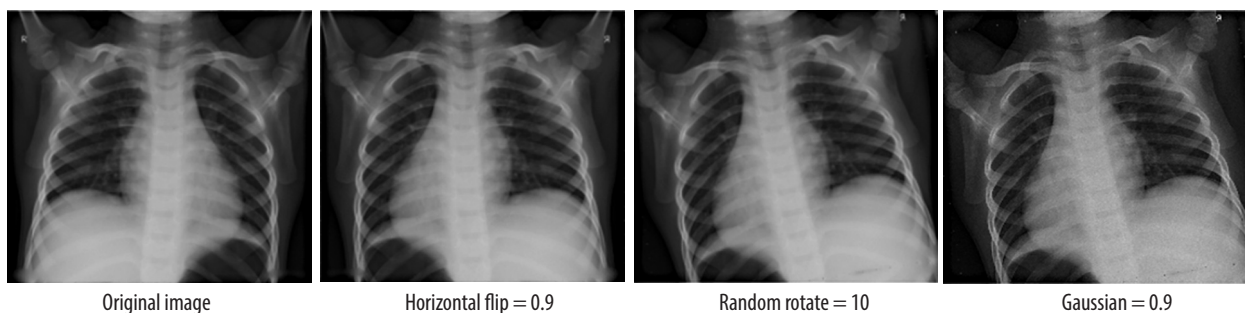


Figure 3. The data augmentation process

Table 1. List of applied augmentation pipeline with parameter values

List of augmentations	Parameters
Horizontal flip	$p = 0.9$
Random degree rotation	Limit = 10
Gaussian noise	$p = 0.9$

The model architecture

The proposed architecture is built upon a shallow CNN of 5 blocks of convolutional layers, followed by a fully connected layer with 128 neurons. The output is generated using another fully connected layer of 3 neurons representing normal, viral, and bacterial pneumonia. To prevent the overfitting problem during the training,

a normalization process is performed after each convolutional block. In addition, a dropout layer was inserted in the model to prevent the memorizing of training data (Figure 4). The features extracted from the training of the CNN are fed to the XGboost classifier, which is evaluated using the test data (Figure 5).

Experimental study

Model training

The training experiments were processed using various hyper-parameters. We trained the CNN model firstly to fix initial values that lead to better results. Different epoch sizes were considered. Experiments showed that a set

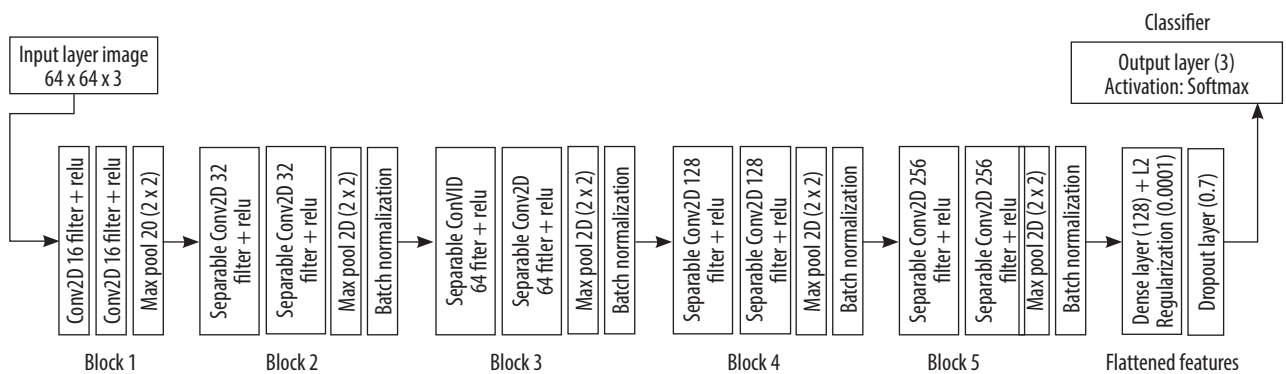


Figure 4. The proposed CNN model architecture

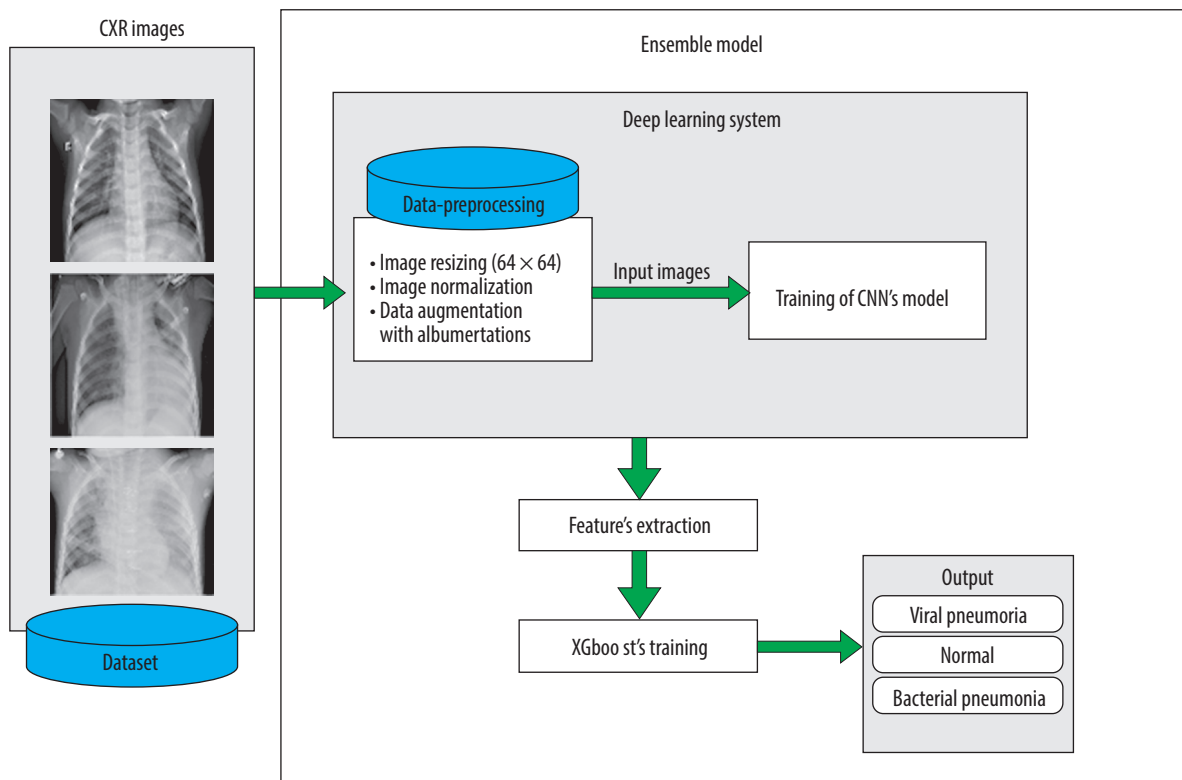


Figure 5. Block diagram of the proposed model architecture

Table 2. Optimal experiment results

Hyperparameter	Value	Hyperparameter	Value
Image size	64 × 64	Output activation function	SoftMax
Number of convolutional layers	10	Loss function	Categorical-cross-entropy
Number of dense layers	1	Optimizer	Adam
Kernel size	3	No-epochs	40
Pooling size	2	Batch size	128
Initial learning rate	0.0001	Dropout	0.7
Layer's activation function	Relu	L2 normalization	0.0001

of 40 epochs returns an optimal result (Table 2). Beyond that, the algorithm's performance seems to be flat, and the model loss stopped decreasing. Hence, an *early stopping* process was added to prevent the overfitting. For training, the *Adam optimizer algorithm* was used with a learning rate value of 0.0001. The loss function was handled by a *categorical-cross-entropy*, and for the classification we used the *softmax* activation function in the output layer conjointly with L2-regularization optimizers. Table 2 shows details about the considered hyperparameters.

The XGboost classifier handles the second stage of training; it is fed a feature vector of length 52,163 that was extracted from CNN training by using a learning rate in the range 0.00001–0.1, which is considered within a wide variety of estimators. For the learning rate and number of estimators, experiments allowed for the best performance with values of 0.001 and 200, respectively.

As it is shown in Figure 1, the validation data are insufficient to adjust the model's hyperparameters. So, we split the training data by 20% to validate this model.

Experimental process

Experiments were processed using Google Colab pro [30] on a A100 NVIDIA premium GPU with 27.3 Gb in the Python environment. We also used TensorFlow 2.11.0 and keras 2.11.0 as backend libraries and some other auxiliary tools such as sklearn for calculation of metrics and Numpy for image manipulation.

Within this context, various performance metrics were also used to evaluate the model including accuracy, precision, recall, and F1-score. The accuracy is defined

Table 3. Classification performance for the test set

	Accuracy	Precision	Recall	F1-score	Training time
CNN	0.85	0.86	0.83	0.85	14 min
CNN-XGboost	0.87	0.89	0.85	0.87	7 s
CNN-Catboost	0.81	0.83	0.75	0.78	2 min 13 s
CNN-LightGBM	0.79	0.80	0.68	0.73	57 s

as the ratio between the number of correctly predicted samples and the total number of samples. The precision, also known as specificity, estimates how well the model classifies a sample as positive. However, the recall or sensitivity, gauges the model's ability to detect positive samples. In addition, the F1-score offers a pertinent blend of recall and precision. These evaluation metrics are represented by the following equations:

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

$$Precision \text{ (or specificity)} = \frac{TP}{(TP + FP)} \tag{2}$$

$$Recall \text{ (or sensitivity)} = \frac{TP}{(TP + FN)} \tag{3}$$

$$F1\text{-score} = 2 \times \frac{precision \times recall}{precision + recall} \tag{4}$$

where *TP*, *TN*, *FP*, and *FN* denote, respectively, the number of samples correctly classified as positive, the number of samples correctly classified as negative, the number of samples wrongly classified as positive, and the number of samples wrongly classified as negative.

Table 3 shows the performances obtained by the various mentioned tools during the training process including other gradient boosting classifiers such as Catboost [31] and LightGBM [32], which are trained based on default hyperparameters.

The learning curves of the proposed model with respect to epochs are shown in Figure 6. We noticed that no overfitting occurred to impact the model accuracy. The Figure 7 illustrates the performance analysis of the CNN model training and its combination with the different gradient boosting classifiers.

To evaluate the performance of the proposed model, we trained the model for a binary classification by combining the viral pneumonia images with the bacterial pneumonia images as one class. The results are showed in Figures 8 and 9.

We also tested the efficiency of the used data augmentation technique. The difference between the model's per-

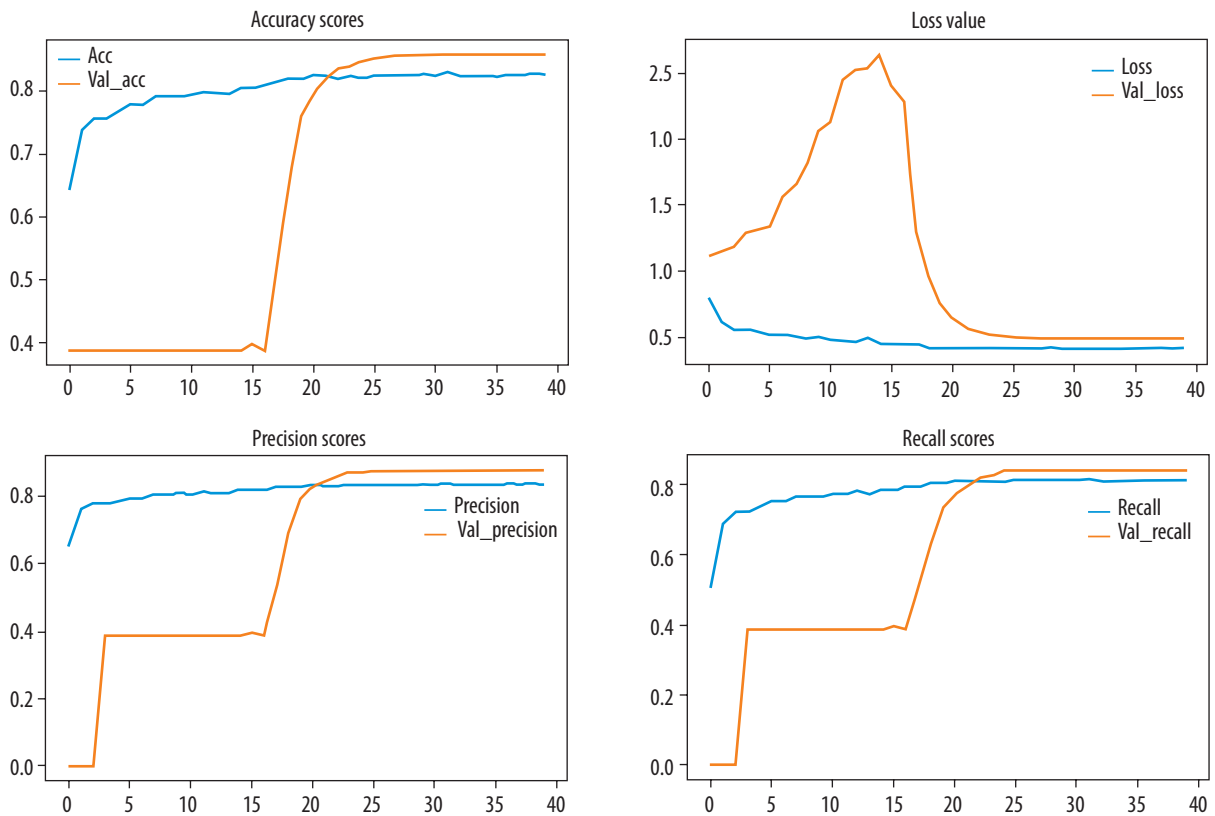


Figure 6. Curves demonstrating the performance of the CNN model for 3-way classification

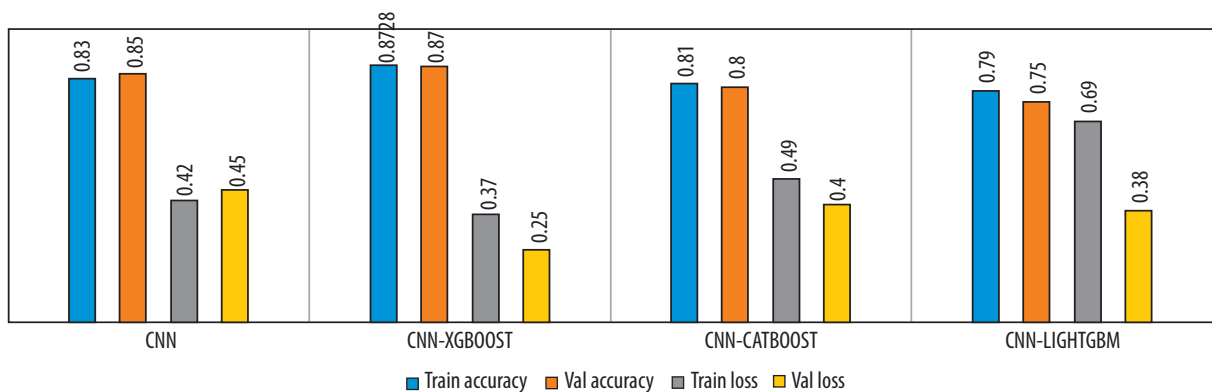


Figure 7. Comparison of the model's performance in term of accuracy and loss for 3-class classification

formance with and without data augmentation is given in Figure 10.

Discussion

Figures 6-9 and Table 3 demonstrate, in terms of performance and training time, that the hybridization of CNN and XGboost tools provide better results than each tool separately. However, the results obtained by CNN-LightGBM were not satisfactory. Thus, the same algorithm with 2-way classification achieved a higher accuracy compared with 3-way classification.

Table 3 indicates that the precision of the model outperforms the recall, which means that the model was able to correctly classify healthy CXR images compared with infected CXR images. This is due to the low number of images in the dataset and the strong resemblance between viral images and bacterial images. In the case of binary classification, the model yields a high recall, which means that this algorithm successfully identified a large number of pneumonia cases, which is very important for early diagnosis.

In regard to the literature, good performance was presented by [22], with an accuracy of 91.8% in distinguishing

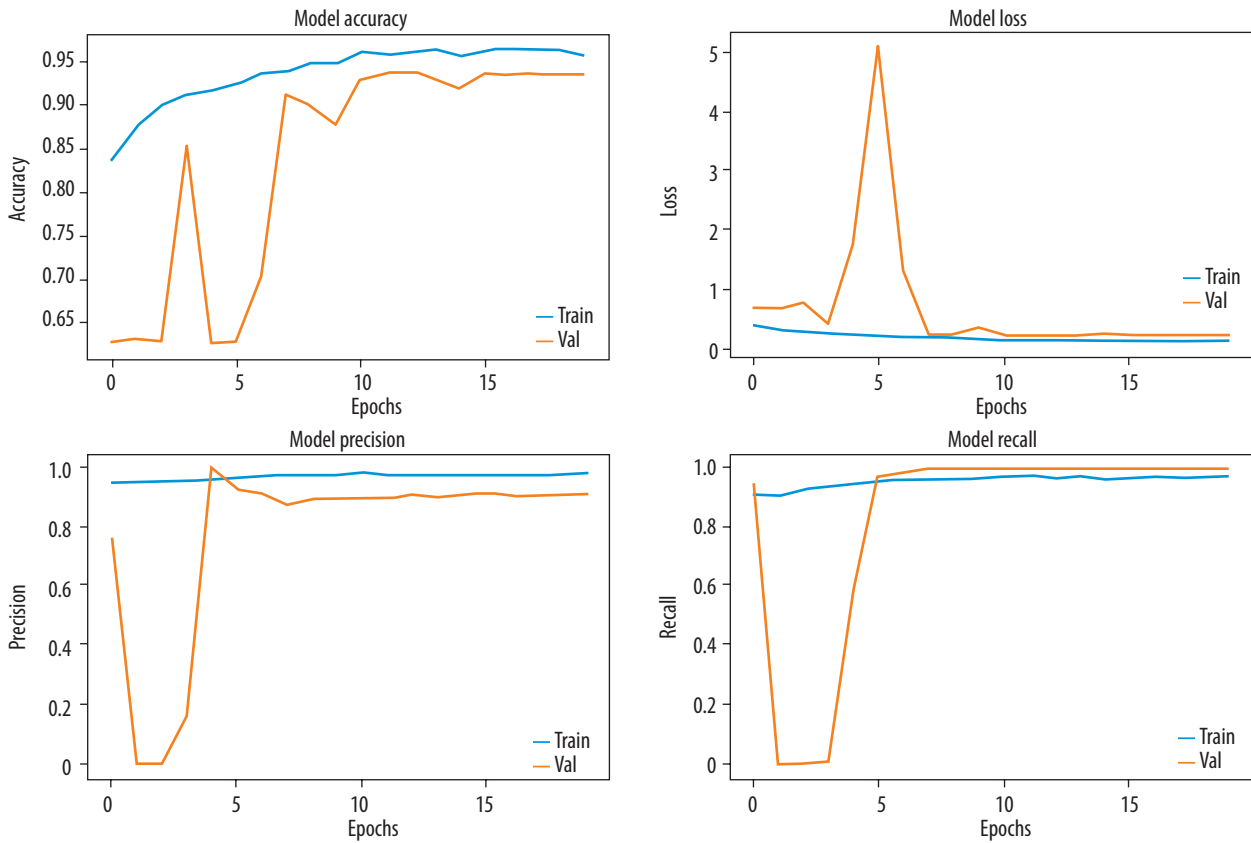


Figure 8. Curves demonstrating the performance of the CNN model for 2-class classification

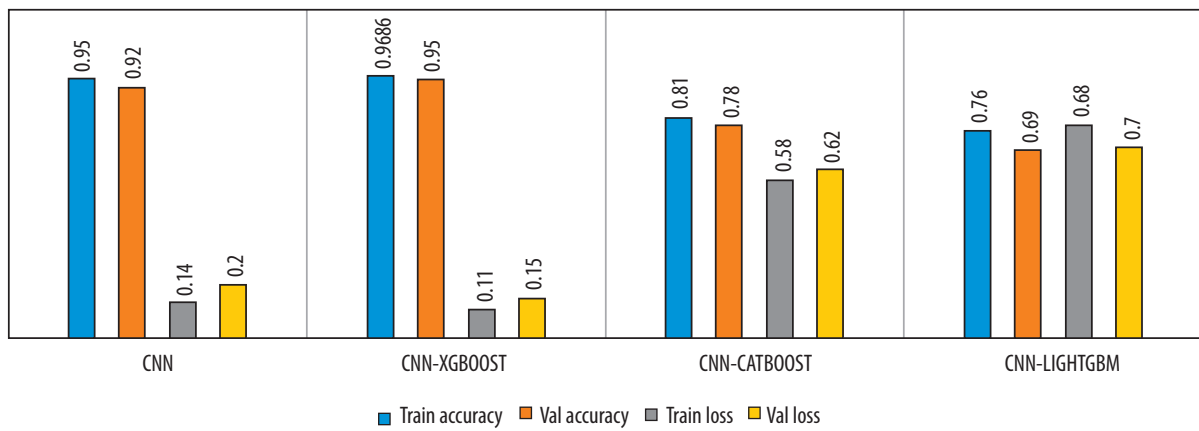


Figure 9. Comparison of the model's performance in term of accuracy and loss for 3-class classification

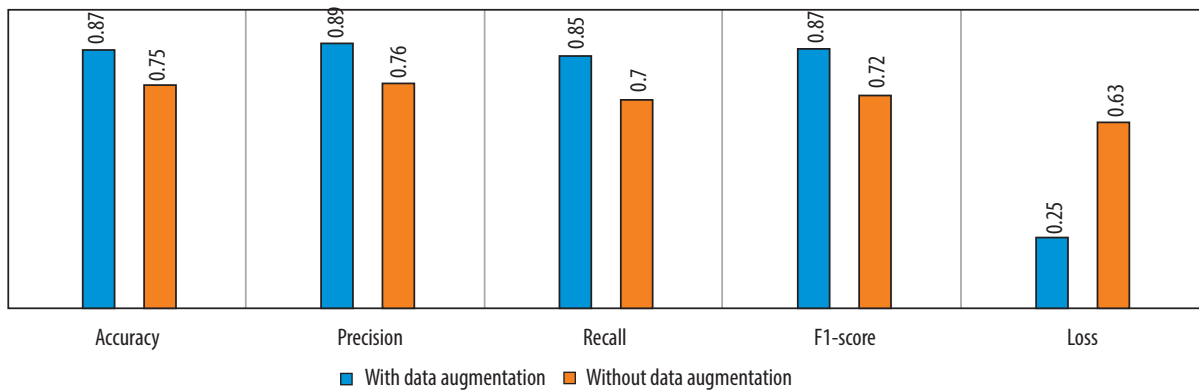


Figure 10. Comparison of CNN-XGboost for 3-class classification vs. data augmentation

Table 4. Summary of different results achieved with regard to related works

Authors	Pathology	Dataset	Number of classes	Model architecture	Results
Rajpurkar <i>et al.</i> [5]	Pneumonia	Chest x-ray 14 [7]	14	Pretrained DenseNet 121	AUCROC = 76.8%
Jiang <i>et al.</i> [8]	Pneumonia	Children pneumonia dataset [20]	2	Inception ResNet V2	Accuracy = 94.2%
Kermany <i>et al.</i> [9]	Pneumonia	Children pneumonia dataset [20]	2	Transfer learning	Accuracy = 92.8%
Shah <i>et al.</i> [10]	Pneumonia	Children pneumonia dataset [20]	2	Shallow CNN	Accuracy = 92.8%
Yaseliani <i>et al.</i> [11]	Pneumonia	Children pneumonia dataset [20]	2	Hybrid CNN VGG16+VGG19	Accuracy = 98%
Pham <i>et al.</i> [12]	Chest abnormalities	VinDr-CXR [35]	15	ResNet and efficientNet for classification / FasterRCNN, yolov5 and efficientDet for abnormality localisation	MAP = 29%
Sirazitdinov <i>et al.</i> [14]	Pneumonia	Children pneumonia dataset [20]	2	Mask RCNN and RetinaNet	Recall = 79.3%
Liz <i>et al.</i> [16]	Alveolar pneumonia and non-consolidation	Children pneumonia dataset [20] + X-ray pediatric pneumonia Xrpp dataset	2	CNN from scratch with 5 convolutional layers	AUC = 92%
Wang <i>et al.</i> [18]	Pneumonia and COVID-19	CovidX dataset [18]	3	Pretrained deep CNN	Accuracy = 93.3%
Rajaraman <i>et al.</i> [22]	Pneumonia	Children pneumonia dataset [20]	3	Customized VGG16	Accuracy = 91.8%
Verma <i>et al.</i> [23]	Viral pneumonia and COVID-19	Children pneumonia dataset [20] + BIMCV COVID-19 dataset [36]	3	Fine-tuned pretrained VGG16	Accuracy = 98%
Gielczyk <i>et al.</i> [34]	COVID-19	Hospital database of the Antoni Jurasz University hospital	2	CNN-LightGBM	Accuracy = 100%
Our	Pneumonia	Children pneumonia dataset [20]	3	Hybrid model CNN-XGboost	Accuracy = 87%

ishing between normal, viral, and bacterial pneumonia. Within the Kaggle challenge for pneumonia detection in the spring of 2022, the best result was 87.018% [33]. Both works used fine-tuned, pre-trained CNNs (such as VGG-16, ResNet-50, or DenseNet), which allow general learned features to enhance the model's performance. However, this kind of model encloses various convolutional layers, leading to complex treatment, and consequently requires more processing time. On the other hand, a similar approach was proposed by [34], which is based on CNN for feature extraction and XGboost, Catboost, LightGBM, and random forest for the classification of COVID-19 CXR images; LightGBM achieved the highest accuracy of 100%. A sparse number of 60 real data obtained from Antoni Juraz University was used (30 from healthy individuals and 30 with COVID-19) for which we

cannot ensure the generalizability of the model to other data.

The use of a combination of CNN-XGboost models allowed us to take advantage of the power of CNN in the feature's extraction and parallelization process of XGboost, which reduced the execution time. Table 4 shows an overview of results achieved by works developed for pneumonia classification compared to our work.

Conclusions and future perspectives

The main goal of this study was to develop a high-precision model for real-time automated diagnosis using CXR images, which could help clinicians to anticipate the pneumonia treatment process. The paper provides a model based on convolutional neural network and extreme

gradient boosting classifier to detect and classify images of viral and bacterial pneumonia diseases. The work was a challenging task due to the lack of appropriate data. The experimental process allows a greater accuracy of 87%, a specificity of 89%, and a sensitivity of 85%.

In future, we aim to enhance the model performance using a larger dataset, either by augmenting the data using more robust tools of data augmentation such as Genera-

tive Neural Networks (Gans) or by collecting more CXR images of pneumonia cases. We can also attempt experimentation with more models to obtain greater accuracy.

Conflict of interest

The authors report no conflict of interest.

References

1. Chang CC, Cheng AC, Chang AB. Over-the-counter (OTC) medications to reduce cough as an adjunct to antibiotics for acute pneumonia in children and adults. *Cochrane Database Syst Rev* 2014; 3: CD006088. doi: 10.1002/14651858.CD006088.PUB4/FULL.
2. World Pneumonia Day – Stop Pneumonia/Every Breath Counts. Available at: <https://stoppneumonia.org/latest/world-pneumonia-day/> (Accessed: 02.04.2023).
3. Peterson R (ed.). *Pneumonia: Symptoms, Types and Treatment*. Foster Academics; 2020.
4. Pneumonia – Treatment | NHLBI, NIH. Available at: <https://www.nhlbi.nih.gov/health/pneumonia/treatment> (Accessed: 28.08.2023).
5. Rajpurkar P, Irvin J, Zhu K, et al. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv*: 1711.05225; 2017. doi: <https://doi.org/10.48550/arXiv.1711.05225>.
6. Han B. Application of deep learning in medical imaging. 2021 2nd International Conference on Computing and Data Science (CDS); 2021, pp. 49-56. doi: 10.1109/CDS52072.2021.00016.
7. Wang X. NIH Chest X-ray dataset of 14 Common thorax disease categories. *Proc IEEE Conf Comput Vis Pattern Recognit* 2017, pp. 2097-2106.
8. Jiang Z. Chest X-ray Pneumonia detection based on convolutional neural networks. 2020 International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE); 2020, pp. 341-344. doi: 10.1109/ICBAIE49996.2020.00077.
9. Kermany DS, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 2018; 172: 1122-1131.e9. doi: 10.1016/j.cell.2018.02.010.
10. Shah S, Mehta H, Sonawane P. Pneumonia detection using convolutional neural networks. 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT); 2020, p. 933-939.
11. Yaseliani M, Hamadani AZ, Maghsoodi AI, Mosavi A. Pneumonia detection proposing a hybrid deep convolutional neural network based on two parallel visual geometry group architectures and machine learning classifiers. *IEEE Access*, vol. 10, pp. 62110-62128, 2022. doi: 10.1109/ACCESS.2022.3182498.
12. Pham V, Tran C, Zheng S, et al. Chest X-ray abnormalities localization via ensemble of deep convolutional neural networks. 2021 International Conference on Advanced Technologies for Communications (ATC); 2021, pp. 125-130.
13. Albumentations Documentation. Available at: <https://albumentations.ai/docs/> (Accessed: 02.02.2023).
14. Sirazitdinov I, Kholiavchenko M, Mustafaev T, et al. Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. *Comput Electr Eng* 2019; 78: 388-399. doi: 10.1016/j.compeleceng.2019.08.004.
15. RSNA Pneumonia Detection Challenge | Kaggle. Available at: <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/> (Accessed: 02.02.2023).
16. Liz H, Sánchez-Montañés M, Tagarro A, et al. Ensembles of Convolutional Neural Network models for pediatric pneumonia diagnosis. *Futur Gener Comput Syst* 2021; 122: 220-233.
17. Chowdhury YS, Dasgupta R, Nanda S. Analysis of various optimizer on CNN model in the application of pneumonia detection. 2021 3rd International Conference on Signal Processing and Communication (ICSPC); 2021, 417-421. doi: 10.1109/ICSPC51351.2021.9451768.
18. Wang L, Lin ZQ, Wong A. Covid-net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest x-ray images. *Sci Rep* 2020; 10: 19549. doi: 10.1038/s41598-020-76550-z.
19. Irvin J, et al. CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. *Proceedings of the AAAI Conference on Artificial Intelligence* 2019; 33(01): 590-597. doi: 10.1609/aaai.v33i01.3301590.
20. Chest X-Ray Images (Pneumonia) | Kaggle. Available at: <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia/versions/1> (Accessed: 01.11.2022).
21. Gu X, Pan L, Liang H, Yang R. Classification of bacterial and viral childhood pneumonia using deep learning in chest radiography. *ICMIP '18: Proceedings of the 3rd International Conference on Multimedia and Image Processing*; 2018, pp. 88-93. doi: 10.1145/3195588.3195597.
22. Rajaraman S, Candemir S, Kim I, et al. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. *Appl Sci* 2018; 8: 1715. doi: 10.3390/app8101715.
23. Verma DK, Saxena G, Paraye A, et al. Classifying COVID-19 and viral pneumonia lung infections through deep convolutional neural network model using chest X-Ray images. *J Med Phys* 2022; 47: 57-64. doi: 10.4103/jmp.jmp_100_21.
24. Tiwari A, Sharan TS, Sharma S, Sharma N. Deep learning-based automated multiclass classification of chest X-rays into COVID-19, normal, bacterial pneumonia and viral pneumonia. *Cogent Eng* 2022; 9. doi: 10.1080/23311916.2022.2105559.
25. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 2016; pp. 785-794. doi: 10.1145/2939672.2939785.

26. He Z, Lin D, Lau T, Wu M. Gradient boosting machine: a survey. arXiv:1908.06951; 2019. doi: <https://doi.org/10.48550/arXiv.1908.06951>.
27. Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms. *Artific Intell Rev* 2021; 54: 1937-1967. doi: [10.1007/s10462-020-09896-5](https://doi.org/10.1007/s10462-020-09896-5).
28. Buslaev A, Iglovikov VI, Khvedchenya E, et al. Albumentations: fast and flexible image augmentations. *Information* 2020; 11: 125. doi: [10.3390/info11020125](https://doi.org/10.3390/info11020125).
29. Kusk MW, Lysdahlgaard S. The effect of Gaussian noise on pneumonia detection on chest radiographs, using convolutional neural networks. *Radiography* 2023; 29: 38-43. doi: [10.1016/j.radi.2022.09.011](https://doi.org/10.1016/j.radi.2022.09.011).
30. Welcome To Colaboratory – Colaboratory. Available at: <https://colab.research.google.com/> (Accessed: 27.05.2023).
31. Dorogush AV, Ershov V, Gulin A. CatBoost: gradient boosting with categorical features support. arXiv:1810.11363; 2018. doi: <https://doi.org/10.48550/arXiv.1810.11363>.
32. Ke G, Meng Q, Finley T, et al. LightGBM: a highly efficient gradient boosting decision tree. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; 2017, pp. 3149-3157.
33. Detect Pneumonia (Spring 2022) | Kaggle. Available at: <https://www.kaggle.com/competitions/detect-pneumonia-spring-2022/leaderboard> (accessed May 19, 2023).
34. Gielczyk A, Marciniak A, Tarczywska M, et al. A novel lightweight approach to COVID-19 diagnostics based on chest X-ray images. *J Clin Med* 2022; 11: 5501. doi: [10.3390/jcm11195501](https://doi.org/10.3390/jcm11195501).
35. Nguyen HQ, Lam K, Le LT, et al. VinDr-CXR: an open dataset of chest X-rays with radiologist's annotations. *Sci Data* 2022; 9: 429. doi: [10.1038/s41597-022-01498-w](https://doi.org/10.1038/s41597-022-01498-w).
36. de la Iglesia Vayá M, Saborit JM, Montell JA, et al. BIMCV COVID-19+: a large annotated dataset of RX and CT images from COVID-19 patients. arXiv:2006.01174; 2020. doi: <https://doi.org/10.48550/arXiv.2006.01174>.