Learning and Nonlinear Models - Journal of the Brazilian Society on Computational Intelligence (SBIC), Vol. 20, Iss. 2, pp. 5-20, 2022

TUNING OF DATA AUGMENTATION HYPERPARAMETERS TO COVID-19 DETECTION IN X-RAY IMAGES WITH DEEP LEARNING

Pedro Rici , Samara Oliveira Silva Santos , André Luiz Carvalho Ottoni
Technologic and Exact Center, Federal University of Recôncavo da Bahia, Cruz das Almas, Brazil
pedrorici@ieee.org, oss.samara@gmail.com, andre.ottoni@ufrb.edu.br

Abstract – The Covid-19 pandemic has been declared in 2020 by the World Health Organization. One of the most relevant aspects of this respiratory disease is the fact that the infection caused by the new coronavirus has a high rate of spread. Thus, rapid and accurate diagnosis can contribute to reducing the transmission rate. In this aspect, in the literature, Deep Learning techniques are studied for application in the detection of this disease through X-ray images of the patient's lung. However, one of the challenges in this area is the training of Convolutional Neural Network models with a database with few samples. One possibility is the generation of artificial images through Data Augmentation techniques. Thus, the objective of this work is to propose a careful methodology for the tuning of Data Augmentation hyperparameters for the classification of lung X-ray images in Covid-19 detection with Deep Learning. The proposed method consists of analyzing the accuracy of 36 Data Augmentation transformations applied to generate new images for training with balanced and unbalanced database. After the selection of hyperparameters, the classifier system achieved accuracies up to 100% on the testing stage, both for combinations and individual transformations with balanced database. Therefore, it is recommended to use a balanced database with the use of zoom, rotation, brightness in combination or individually, for Covid-19 versus Normal and Covid-19 versus Pneumonia classification.

Keywords – Deep Learning, Data Augmentation, Covid-19, Hyperparameter Tuning, X-Ray Images.

1 Introduction

Covid-19 is an infectious respiratory disease caused by the new coronavirus (SARS-CoV-2) [1–4]. The new disease was first identified in 2019 in Wuhan in China and quickly spread around the world, until on March 11, 2020, the World Health Organization (WHO) declared a Covid-19 Pandemic [1]. To identify Covid-19 contamination, there are tests that, combined with clinical symptoms, can confirm the diagnosis [5]. One way to perform detection is through molecular testing, when the presence of genetic material from the virus is detected in the patient; or serological, when antibodies are detected [2].

Rapid and accurate diagnosis of the Covid-19 virus helps to decrease the rate of transmission and also to monitor the patient's condition [6]. However, the lack of tests, the waiting time for results, and the stress on health professionals make it difficult to make a rapid diagnosis [7]. So, alternatives to help detect coronavirus and other respiratory diseases such as pneumonia are emerging to speed up and lower the cost [7–9]. One such approach is the use of Machine Learning to detect Covid-19 using X-ray images of patients with symptoms [10–16]. Along these lines, the study of [17] develops a new architecture to classify X-ray images into three groups of patients: healthy, viral pneumonia, Covid-19 pneumonia. Another example is the work of [14], which conducts an extensive literature review and brings together the main advances in the field fo Covid-19 detection in images with Deep Learning.

However, one of the challenges in this area is the hyperparameter tuning and neural architecture selection of Machine Learning methods. [10–13]. Along this line, some works have already performed comparative studies of different Deep Learning architectures for Covid-19 detection in X-ray images [10, 12, 15]. In this sense, the paper of [10] analyzes two different Convolutional Neural Network architectures (ResNet50 and ResNet101) and obtains an accuracy of 97.77% in detecting Covid-19 in X-ray images. The study of [12], in turn, compares five different Deep Learning architectures (DenseNet, InceptionV3, MobileNet, ResNet50, VGG16 and VGG19), achieving accuracy of 98.81% for the binary classification and 91.68% for the multi-class classification.

Following this aspect of defining good conditions for Deep Learning experiments, an important area is Data Augmentation [18–21]. This technique is used to reduce overfitting, generating artificial training images through random transformations [22]. The application of this method becomes important especially in cases where the number of samples for training is relatively small [18, 21, 22]. In this context, most of the studies on the application of Deep Learning in the diagnosis of Covid-19 via images are found, since they perform experiments with databases with few examples [16, 23, 24]. Within this aspect, some papers discuss the role of Data Augmentation in Covid-19 classification in X-ray images [16, 23, 24]. However, little attention has been focused in the literature on Data Augmentation hyperparameter analysis (image transformations) when it comes to Covid-19 diagnosis using X-ray images [16].

The objective of this work is to propose an experimental methodology to tuning of Data Augmentation hyperparameters in the classification of X-ray images for Covid-19 detection with Deep Learning. To do this, it is necessary to study and understand how Deep Learning techniques work, apply a Machine Learning architecture for this type of

classification and analyze the accuracy for different combinations of hyperparameters. A Convolutional Neural Network discussed in the literature was selected [25, 26], and thus performed the classification from two experimental groups: (1) between X-ray images of normal patients *versus* patients with Covid-19; (2) between X-ray images of patients with pneumonia *versus* Covid-19 patients.

This paper is an extended version of the study [27], published in the XV Brazilian Congress of Computational Intelligence and presents the following advances compared to the previous version:

- Classification of x-ray images for Covid-19 detection with an unbalanced dataset;
- Comparison between the accuracy in the testing step of a balanced and unbalanced dataset;
- Analysis of Data Augmentation hyperparameters individually, including two new hyperparameters: *width shift range* and *height shift range*;
- New results, highlighted in Section 4.1, 4.3.2 and 4.3.3;
- Addition of new references;
- Increase of new suggestions for future work;
- Paper written in English;

The present work is divided into six sections. Section II presents the theoretical background. Section III presents the database, the architecture and hyperparameters, and the methods used in the work. Section IV presents the results obtained. Section V presents a comparison between the results obtained in this paper and other related work. Finally, Section VI presents the contributions made by this work, as well as suggestions for future work.

2 Background

Artificial Neural Networks (ANNs) are machine learning methods widely discussed in the literature. [28–30]. In this aspect, ANNs have several applications, such as: facial recognition [31, 32], stock exchange [33–35], autonomous vehicles [36–38] and building construction [26, 39].

The inspiration for creating the ANN came with the concept of the human nervous system [28], where there are similarities between the physical structures and functions of the biological neuron with the artificial neuron. An example of an ANN can be represented by Figure 1.

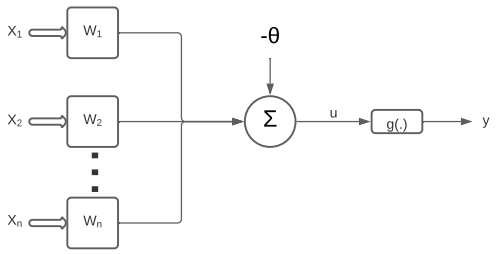


Figure 1: Artificial Neuron. Source: based on [28].

where x_1, x_2 and x_n are the input signals, w_1, w_2 and w_n are the synaptic weights, $\sum$ is the linear combiner, θ is the activation threshold, u is the activation potential, g is the activation function, and y is the output signal. In this sense, Equations (1) and (2) summarize the output produced by the artificial neuron [28]:

$$u = \sum_{i=1}^{n} w_i.x_i - \theta, \tag{1}$$

$$y = g(u). \tag{2}$$

Convolutional Neural Networks (CNN) are a type of ANN [25, 40]. A CNN is characterized by its efficiency and robustness in handling information from a large number of layers and neurons [25, 28, 40, 41]. CNN has 4 basic operations [25, 40, 41]:

1. Convolution operator (Kernel): The convolution matrix is multiplied by the image matrix to apply image effects. The result of this multiplication is a feature map, where the most notable features are stored, decreasing the size of the original matrix for easier processing. After the feature map is obtained, the activation function Relu is usually applied.

2. Pooling: This is the layer responsible for emphasizing the key features resulting from the application of the convolutional layer, reducing overfitting and unnecessary noise.

3. Flattening: This is the layer that transforms the resulting matrix from Pooling into a vector, thus being able to start the application in the dense neural network.

4. Fully connected layer (Dense): this is the net with neurons like the net formed by the traditional ANN, with sets of neurons and fully connected layers that together define the output classes.

To assist in the development of a CNN, programming language libraries can be used. An example is Keras, a library based on TensorFlow, free and available in Python or R [22, 25].

3 Proposed Method

The methodology proposed in this work is composed of five steps: (1) database definition, (2) image pre-processing, (3) CNN architecture definition, (4) training and validation phase, and (5) test phase.

3.1 Database

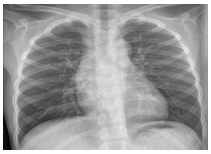
The selected image database is available on the Kaggle[1] platform. This database contains images obtained from three different sources[2][3][4]. According to the information provided on the Kaggle website, all sources are freely licensed and can be used for research, evaluation and commercial purposes.

The database has a total of 6,432 images, divided into three classes:

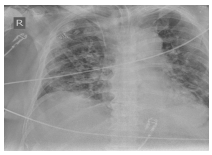
- Covid-19: 576 lung X-ray images from patients diagnosed with Covid-19.
- Pneumonia: 4,273 lung X-ray images of patients diagnosed with pneumonia.
- Normal: 1,583 lung X-ray images from patients diagnosed without Covid-19 and pneumonia.

Figure 2 demonstrates some examples of images made available in the database.

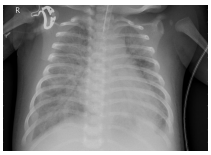
(a) Class Normal.



(b) Class Normal.



(c) Class Covid-19.



(d) Class Pneumonia.

Figure 2: Example images, for each class (Normal, Covid-19 and Pneumonia) made available in the database.

3.2 Pre-Processing

After the database was defined, the images were reorganized in two dataset types: balanced and unbalanced. For each dataset type, the images were organized into two experimental groups: Covid-19 *versus* Pneumonia and Covid-19 *versus* Normal. Moreover, the images were separated separated into three sets (training, validation, and test) as follows:

[1] https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia
[2] https://github.com/agchung
[3] https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
[4] https://github.com/ieee8023/covid-chestxray-dataset

3.2.1 Unbalanced

- Covid-19 versus Pneumonia
  - Training: 432 images of the Covid-19 class and 3205 images of the Pneumonia class.
  - Validation: 72 images of the Covid-19 class and 534 images of the Pneumonia class.
  - Test: 72 images of the Covid-19 class and 534 images of the Pneumonia class.
- Covid-19 versus Normal
  - Training: 432 images of the Covid-19 class and 1187 images of the Normal class.
  - Validation: 72 images of the Covid-19 class and 198 images of the Normal class.
  - Test: 72 images of the Covid-19 class and 198 images of the Normal class.

Thus, the Covid-19 versus Pneumonia group has 4849 images and the Covid-19 versus Normal group has 2159. Furthermore, the images were divided into 75% for training, 12.5% for validation and 12.5% for testing.

3.2.2 Balanced

- Covid-19 versus Pneumonia
  - Training: 432 images from the Covid-19 class and 432 images from the Pneumonia class.
  - Validation: 72 images from the Covid-19 class and 72 images from the Pneumonia class.
  - Test: 72 images from the Covid-19 class and 72 images from the Pneumonia class.
- Covid-19 versus Normal
  - Training: 432 images from the Covid-19 class and 432 images from the Normal class.
  - Validation: 72 images from the Covid-19 class and 72 images from the Normal class.
  - Test: 72 images from the Covid-19 class and 72 images from the Normal class.

Thus, the reorganized database now has 1728 images. Furthermore, the images were divided into 75% for training, 12.5% for validation and 12.5% for testing. The Pneumonia and Normal images were selected in the original numbered order, i.e. the first 576 images of each class were selected.

3.2.3 Data Augmentation

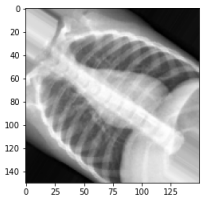
The amount of images contained in the database is considerably small. Thus, the proposed methodology aims to apply the technique of Data Augmentation [18, 22] to generate new images and consequently increase the amount of data for training. For this, the function $ImageDataGenerator()$ provided by the library Keras [22] was adopted, where it generates new images, as needed, online, which means that the new batches of image data are generated in real time. The number of artificially created images depends on the training settings: *batch_size*, *steps_perhaps_epochs* and *epoch*. For example, for this work, these parameters are set in the first phase of experiments as: $batch\_size = 32$; $steps\_per\_epochs = 25$; and $epoch = 10$. In this sense, for each simulation approximately 8000 images are randomly generated for training from the original images. Thus, 4000 images artificially generated are added for each class.

The proposed methodology selected six Data Augmentation transformations to be analyzed:

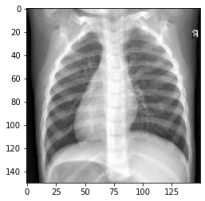
- rotation range: represented by an integer value that identifies the maximum rotation value the image can have.
- brightness range: a list or tuple that represents the brightness range that the image can have.
- zoom range: a value of the float type or a vector [min, max] that determines the zoom range that the image can get.
- horizontal flip: boolean type where the image can be inverted horizontally or not.
- width shift range: a value of type integer or float that determines the range of horizontal shift that the image can get.
- height shift range: a value of type integer or float that determines the vertical shift range the image can get.

Figure 3 exemplifies some images generated from using the $ImageDataGenerator()$ function with the following hyperparameter values:

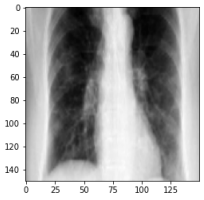
- *rotation_range*: 60;
- *brightness_range*: [0.2, 0.8];
- *zoom_range*: [0.3, 1.0];
- *horizontal_flip*: True;
- *width_shift_range*: 0.6;
- *height_shift_range*: 0.4.



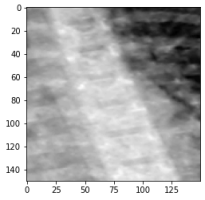
(a) Class Normal using only rotation.



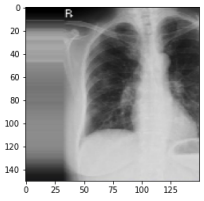
(b) Class Normal using only horizontal flip.



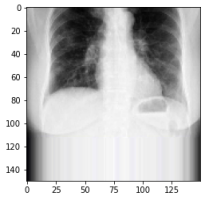
(c) Class Covid-19 using only zoom.



(d) Class Pneumonia using four hyperparameters.



(e) Class Normal using only width shift range.



(f) Class Normal using only height shift range.

Figure 3: Examples of images generated with data augmentation for the Normal, Pneumonia and Covid-19 classes.

3.3 Convolutional Neural Network Architecture

The CNN architecture used in this work was proposed by [25] and used in other works in the literature [26, 39]. This architecture has 12 layers, four convolutional, four Max Pooling, one Flatten, one Dropout, a dense layer with 512 neurons and the output layer with the classifier neuron. The architecture used can be represented by Figure 4.

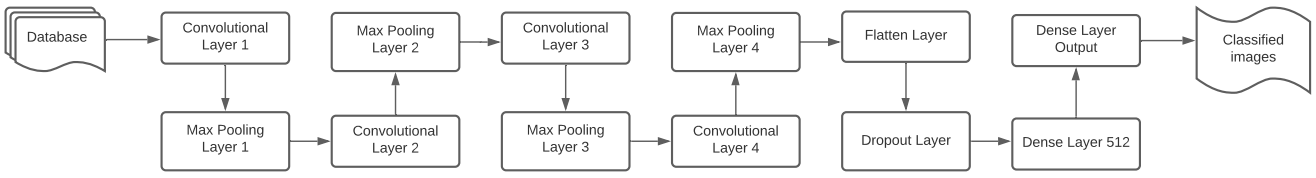


Figure 4: Flowchart for representing the CNN-12 Convolutional Neural Network adopted in this work. Based on [25] and [26].

In addition to the data augmentation hyperparameters mentioned above, fixed hyperparameters were used for the experimental stages of the CNN: number of filters in each convolutional layer (32, 64, 128, 128), kernel size in the convolutional layers (3x3), pool size in the Max Pooling layers (2x2), and RELU activation functions in the convolutional/dense and sigmoid layers of the classification layer. The CNN architecture was coded from the keras.models.Sequential() command, available from the Keras library of the Python language [25]. Also set the dropout rate at 0.2, used the Relu and Sigmoid activation functions, and the Adagrad optimizer with a learning rate of 0.001.

3.4 Training and Validation

The training and validation experiments aim to evaluate the hyperparameter combinations of Data Augmentation. In this aspect, the steps proposed by the methodology proposed in this study are represented in Figure 5.

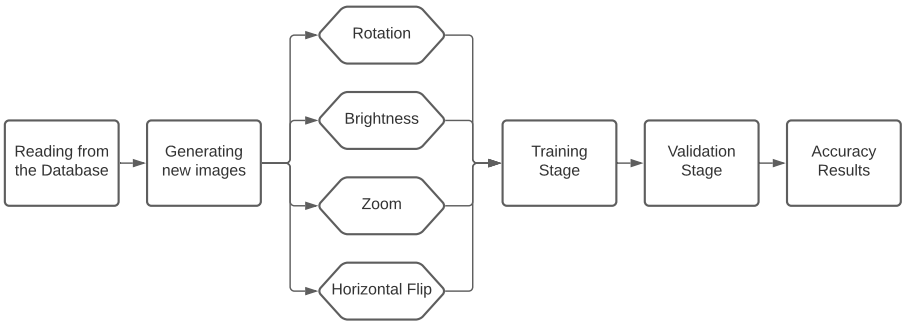


Figure 5: Flowchart of the methodology for hyperparameter tuning of Data Augmentation - training and validation.

The experiments were divided into two phases. The first phase, Data Augmentation Combined Analysis (DACA), consists in testing how combinations of Data Augmentation interfere with the accuracy. Already in the second phase, Data Augmentation Single Analysis (DASA), of experiments, hyperparameters are evaluated individually, varying their experimental levels at standard intervals.

In the DACA phase of the study, two experimental levels were defined for analysis of the Data Augmentation hyperparameters:

- *rotation_range*: 60 e 140;
- *brightness_range*: [0.3, 1.0] e [0.2, 0.8];
- *zoom_range*: [0.3, 1.0] e [0.2, 0.8];
- *horizontal_flip*: True e False.

Thus, Table 1 presents the 16 combinations of hyperparameters analyzed in DACA step.

Table 1: Hyperparameter combinations of Data Augmentation used for Phase 1 (DACA).

Combination	Rotation	Brightness	Zoom	Horizontal Flip
1	60	[0.3 1.0]	[0.3 1.0]	True
2	60	[0.2 0.8]	[0.3 1.0]	True
3	60	[0.3 1.0]	[0.2 0.8]	True
4	60	[0.2 0.8]	[0.2 0.8]	True
5	60	[0.3 1.0]	[0.3 1.0]	False
6	60	[0.2 0.8]	[0.3 1.0]	False
7	60	[0.3 1.0]	[0.2 0.8]	False
8	60	[0.2 0.8]	[0.2 0.8]	False
9	140	[0.3 1.0]	[0.3 1.0]	True
10	140	[0.2 0.8]	[0.3 1.0]	True
11	140	[0.3 1.0]	[0.2 0.8]	True
12	140	[0.2 0.8]	[0.2 0.8]	True
13	140	[0.3 1.0]	[0.3 1.0]	False
14	140	[0.2 0.8]	[0.3 1.0]	False
15	140	[0.3 1.0]	[0.2 0.8]	False
16	140	[0.2 0.8]	[0.2 0.8]	False

On the other hand, Table 2 presents the 12 levels of hyperparameters analyzed for the DASA Phase. In addition to the combinations cited in Table 2, other hyperparameters will be used for Covid-19 versus Normal analysis: width shift range and height shift range (see Table 3).

In this sense, in both phases (DACA and DASA), each of the combinations shown in Table 1 and 2 were simulated in 3 repetitions with 10 epochs. Furthermore, for training the CNN models the function *fit_generator()* from Keras library was used [25].

Table 2: Hyperparameter combinations of *Data Augmentation* used for Phase 2 (DASA).

Combination	Rotation	Brightness	Zoom
1	20	-	-
2	40	-	-
3	60	-	-
4	80	-	-
5	-	[0.0 0.2]	-
6	-	[0.2 0.4]	-
7	-	[0.4 0.6]	-
8	-	[0.6 0.8]	-
9	-	-	[0.0 0.2]
10	-	-	[0.2 0.4]
11	-	-	[0.4 0.6]
12	-	-	[0.6 0.8]

Table 3: Other Data Augmentation hyperparameter combinations used for Phase 2 (DASA).

Combination	Width Shift Range	Height Shift Range
1	0.2	-
2	0.4	-
3	0.6	-
4	0.8	-
5	-	0.2
6	-	0.4
7	-	0.6
8	-	0.8

In the validation stage, in turn, CNN performance analysis was performed for each combination of transformation Data Augmentation. This evaluation was performed by observing the metric of image classification accuracy, according to Eq. (3):

$$A_c = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{3}$$

where, A_c is the accuracy value, T_P is the true positives, T_N is the true negatives, F_P is the false positives, and F_N is the false negatives.

3.5 Test

Following the proposed methodology, the three combinations of hyperparameters that obtained the best average accuracy values in the validation stage for the DACA Phase with the balanced database were selected for the test stage. These three combinations went through the testing step for the balanced and unbalanced database. For the DASA Phase, the best experimental hyperparameter level was chosen for each type of transformation. In this new step, the analyzed images are distinct from the training and validation steps, according to the database division described in Section 3.2.

The CNN architecture was maintained and for each hyperparameter combination 3 repetitions with 30 epochs were performed. Therefore, the goal of the testing step is to select which combination of hyperparameters results in the best accuracy when analyzing a new set of images. For this, the *evaluate_generator()* command from the *Keras* library was used for the accuracy evaluation in the test step [25].

3.5.1 Comparison between Balanced and Unbalanced

To determine the influence of database balancing on the experiments, first a test step was performed with the presented architecture. Three repetitions with 30 epochs each were made for each experimental group, that is, for the Covid-19 versus Pneumonia and Covid-19 versus Normal groups, with the database balanced and unbalanced.

For this step, the proposed Data Augmentation hyperparameters were not used, since the goal is to test the influence of database balancing for this architecture on the classification of lung x-ray images for Covid-19 detection.

4 Results

In this section, the results of the training, validation, and testing phases of the CNN application are presented. In each of the steps, for both the DACA Phase and the DASA Phase, hyperparameters were selected from Data

Augmentation with the best accuracy metrics for image classification in the two experimental groups: (1) Normal versus Covid-19 and (2) Pneumonia versus Covid-19.

4.1 Results of Training and Validation

Training and Validation are the first stages of experiments. Based on the results obtained in this stage, the combinations and transformations for the next phase (test experiments) will be determined.

4.1.1 Results of Data Augmentation Combined Analysis (DACA)

Table 4 shows the accuracy values for in the classification process of the first experimental group (Normal × Covid-19) for the DACA phase.

Table 4: Accuracy results (%) for validation in the DACA phase of combinations of Data Augmentation for Normal × Covid-19.

Rot.	Brig.	Zoom	Hor. Flip	Ac_1(%)	Ac_2(%)	Ac_3(%)	Avg.(%)
60	[0.3 1.0]	[0.3 1.0]	True	78.1	43.8	75.0	65.6
60	[0.2 0.8]	[0.3 1.0]	True	90.6	96.9	84.4	**90.6**
60	[0.3 1.0]	[0.2 0.8]	True	71.9	43.8	43.8	53.1
60	[0.2 0.8]	[0.2 0.8]	True	84.4	68.8	65.6	72.9
60	[0.3 1.0]	[0.3 1.0]	False	96.9	81.3	96.9	**91.7**
60	[0.2 0.8]	[0.3 1.0]	False	90.6	90.6	78.1	**86.5**
60	[0.3 1.0]	[0.2 0.8]	False	50.0	81.3	56.3	62.5
60	[0.2 0.8]	[0.2 0.8]	False	71.9	78.1	90.6	80.2
140	[0.3 1.0]	[0.3 1.0]	True	59.4	53.1	90.6	67.7
140	[0.2 0.8]	[0.3 1.0]	True	31.3	75.0	53.1	53.1
140	[0.3 1.0]	[0.2 0.8]	True	56.3	81.3	59.4	65.6
140	[0.2 0.8]	[0.2 0.8]	True	40.6	50.0	43.8	44.8
140	[0.3 1.0]	[0.3 1.0]	False	62.5	65.6	46.9	58.3
140	[0.2 0.8]	[0.3 1.0]	False	62.5	87.5	71.9	74.0
140	[0.3 1.0]	[0.2 0.8]	False	40.6	62.5	75.0	59.4
140	[0.2 0.8]	[0.2 0.8]	False	65.6	84.4	59.4	69.8

Rot.: Rotation; Brig.: Brightness; Hor. Flip: Horizontal Flip.

Therefore, in Table 4 it is possible to observe the influence of the hyperparameters of data augmentation in the accuracy values. In this sense, we highlight the average accuracies of 91.7%, 90.6% and 86.5% as the best results obtained, respectively, by the combinations 2, 5 and 6. These three combinations with the highest average accuracy results were selected for the testing stage of the first experimental group (Covid-19 × Normal). It is important to note that the rotation of 60° and zoom into a range between 30% and 100% are present in the three best accuracy results. In this aspect, it is also worth noting the worst performances (low accuracy), as is the case of 44.8% for combination 12 and 53.1% for combinations 10 and 3.

Table 5 presents the accuracy results from the methodology for the classification of the second experimental group (Pneumonia *versus* Covid-19).

From Table 5, it can also be seen that Data Augmentation hyperparameters directly interfere with CNN performance. For example, the average accuracy value ranged from 55.2% to 88.5%. It is noteworthy that the combinations that achieved the highest average accuracies were: combination 1 (88.5%); combination 6 (87.5%); and combination 7 (81.3%). Thus, these 3 combinations were selected for the testing stage of the Pneumonia × Covid-19 experiments.

4.1.2 Results of Data Augmentation Single Analysis (DASA)

Continuing the analysis of the first experimental group (Normal × Covid-19), Table 6 and 7 presents the DASA Phase results, where the hyperparameters are observed on an individual basis.

Thus, it is possible to perceive the individual influence of the hyperparameters where the rotation and brightness reach accuracy values of 100%, when the applied rotation rate was 20%, 40%, 60% or 80%. As for the brightness at rates of 20% to 40%, 40% to 60% or 60% to 80%. With this analysis it is also seen that the zoom hyperparameter has a low accuracy when compared to the previously mentioned ones, where its maximum value was 84.38% when applied a range between 0% and 20%.

Table 7 shows the accuracy results for the DASA phase validation of combinations for the Pneumonia × Covid-19 group.

Table 5: Accuracy results (%) for validation in the DACA phase of Data Augmentation combinations for Pneumonia × Covid-19.

Rot.	Brig.	Zoom	Hor. Flip	Ac_1(%)	Ac_2(%)	Ac_3(%)	Avg.(%)
60	[0.3 1.0]	[0.3 1.0]	True	90.6	84.4	90.6	**88.5**
60	[0.2 0.8]	[0.3 1.0]	True	75.0	84.4	78.1	79.2
60	[0.3 1.0]	[0.2 0.8]	True	68.8	59.4	71.9	66.7
60	[0.2 0.8]	[0.2 0.8]	True	78.1	68.8	71.9	72.9
60	[0.3 1.0]	[0.3 1.0]	False	84.4	81.3	62.5	76.0
60	[0.2 0.8]	[0.3 1.0]	False	87.5	90.6	84.4	**87.5**
60	[0.3 1.0]	[0.2 0.8]	False	84.4	78.1	81.3	**81.3**
60	[0.2 0.8]	[0.2 0.8]	False	62.5	65.6	62.5	63.5
140	[0.3 1.0]	[0.3 1.0]	True	81.3	65.6	71.9	72.9
140	[0.2 0.8]	[0.3 1.0]	True	81.3	71.9	65.6	72.9
140	[0.3 1.0]	[0.2 0.8]	True	59.4	59.4	62.5	60.4
140	[0.2 0.8]	[0.2 0.8]	True	46.9	65.6	53.1	55.2
140	[0.3 1.0]	[0.3 1.0]	False	50.0	71.9	59.4	60.4
140	[0.2 0.8]	[0.3 1.0]	False	56.3	81.3	78.1	71.9
140	[0.3 1.0]	[0.2 0.8]	False	59.4	59.4	50.0	56.3
140	[0.2 0.8]	[0.2 0.8]	False	75.0	46.9	62.5	61.5

Rot.: Rotation; Brig.: Brightness; Hor. Flip: Horizontal Flip.

Table 6: Accuracy results (%) for DASA phase validation of combinations of Data Augmentation for Normal × Covid-19.

Rot.	Zoom	Brig.	Ac_1(%)	Ac_2(%)	Ac_3(%)	Avg.(%)
20	-	-	100.00	100.00	96.88	**98.96**
40	-	-	100.00	100.00	59.38	86.46
60	-	-	100.00	90.62	100.00	96.87
80	-	-	100.00	96.88	93.75	96.88
-	[0.0 0.2]	-	62.50	84.38	43.75	63.54
-	[0.2 0.4]	-	53.12	59.38	75.00	62.50
-	[0.4 0.6]	-	50.00	50.00	78.12	59.37
-	[0.6 0.8]	-	71.88	65.62	59.38	**65.63**
-	-	[0.0 0.2]	96.88	96.88	90.62	94.79
-	-	[0.2 0.4]	96.88	90.62	100.00	95.83
-	-	[0.4 0.6]	100.00	93.75	96.88	96.88
-	-	[0.6 0.8]	100.00	100.00	96.88	**98.96**

Rot.: rotation; Brig.: brightness.

Table 7: Accuracy results (%) for DASA Phase validation of combinations of Data Augmentation for Pneumonia × Covid-19.

Rot.	Zoom	Brig.	Ac_1(%)	Ac_2(%)	Ac_3(%)	Avg.(%)
20	-	-	96.88	96.88	93.75	95.84
40	-	-	93.75	93.75	96.88	94.79
60	-	-	100.00	90.62	81.25	90.62
80	-	-	100.00	96.88	96.88	**97.92**
-	[0.0 0.2]	-	40.62	53.12	31.25	41.66
-	[0.2 0.4]	-	43.75	59.38	46.88	50.00
-	[0.4 0.6]	-	59.38	75.00	59.38	64.59
-	[0.6 0.8]	-	84.38	81.25	78.12	**81.25**
-	-	[0.0 0.2]	90.62	93.75	96.88	93.75
-	-	[0.2 0.4]	100.00	96.88	96.88	**96.88**
-	-	[0.4 0.6]	96.88	93.75	96.88	95.84
-	-	[0.6 0.8]	90.62	93.75	100.00	94.79

Rot.: rotation; Brig.: brightness.

From Table 7, it can be seen that for the individual Rotation and Brightness transformations maximum accuracies of up to 100% were obtained, highlighting the maximum mean value of 97.92% for rotation at 80 degrees and 96.88% for brightness from [0.2 0.4]. As for the zoom transformation levels, one notices lower values when compared to the other transformations, with the maximum mean value of 81.25% for a zoom from [0.6 0.8].

4.2 Results of Test

With the results obtained in the Training and Validation step and with the analysis of the best averages, it is possible to perform the Test experiments for the balanced and unbalanced database and obtain the final results.

4.2.1 Results of Data Augmentation Combined Analysis (DACA) - Balanced

Table 8 presents the accuracy results referring to the second step (test) of the methodology for the classification of the first experimentation group (Normal × Covid-19).

Table 8: Accuracy results (%) for test in the DACA - Balanced phase of combinations of Data Augmentation for Normal × Covid-19.

Rot.	Brig.	Zoom	Hor. Flip	$Ac_1(\%)$	$Ac_2(\%)$	$Ac_3(\%)$	Avg.(%)
60	[0.3 1.0]	[0.3 1.0]	False	96.9	93.7	96.9	**95.8**
60	[0.2 0.8]	[0.3 1.0]	True	93.7	90.6	**100.0**	94.8
60	[0.2 0.8]	[0.3 1.0]	False	90.6	93.7	90.6	91.7

Rot.: Rotation; Brig.: Brightness; Hor. Flip: Horizontal Flip.

The influence of the hyperparameters can be mentioned again, where in the test step for the first experimental group (Normal × Covid-19) a maximum mean value of 95.8% was obtained for combination (60; [0.3 1.0]; [0.3 1.0]; False) and a minimum mean value of 91.7% for combination (60; [0.2 0.8]; [0.3 1.0]; False). It is also worth mentioning that an accuracy of 100% was obtained for one of the tests (60; [0.2 0.8]; [0.3 1.0]; True).

Table 9 shows the accuracy results referring to the second step of the methodology for the second experimental group (Pneumonia × Covid-19) classification. Here we have the accuracy for the three best combinations of Data Augmentation hyperparameters obtained in stage 1.

Table 9: Accuracy results (%) for test in the DACA - Balanced phase of combinations of Data Augmentation for Pneumonia × Covid-19.

Rot.	Brig.	Zoom	Hor. Flip	$Ac_1(\%)$	$Ac_2(\%)$	$Ac_3(\%)$	Avg.(%)
60	[0.3 1.0]	[0.3 1.0]	True	90.6	78.1	**100.0**	**89.6**
60	[0.2 0.8]	[0.3 1.0]	False	87.5	84.4	87.5	86.5
60	[0.3 1.0]	[0.2 0.8]	False	84.4	81.3	87.5	84.4

Rot.: Rotation; Brig.: Brightness; Hor. Flip: Horizontal Flip.

Once again, the experiments revealed the importance of the hyperparameters of *Data Augmentation* in the results of a CNN. It can be seen that combination (60; [0.3 1.0]; [0.3 1.0]; True) presented the best accuracy value on average, with about 89.6%, and also obtained a maximum accuracy of 100%.

4.2.2 Data Augmentation Combined Analysis (DACA) - Unbalanced

Table 10 presents the accuracy results for the second step of the methodology for the classification of the Normal versus Covid-19 experimentation group with the unbalanced database.

Table 10: Accuracy Results (%) for test in the DACA - Unbalanced phase of combinations of Data Augmentation for Normal × Covid-19.

Rot.	Brig.	Zoom	Hor. Flip	$Ac_1(\%)$	$Ac_2(\%)$	$Ac_3(\%)$	Avg.(%)
60	[0.3 1.0]	[0.3 1.0]	False	78.1	75.0	53.1	**68.8**
60	[0.2 0.8]	[0.3 1.0]	True	59.4	78.1	56.3	64.6
60	[0.2 0.8]	[0.3 1.0]	False	81.3	68.8	53.1	67.7

The highest accuracy value for the experimental group Covid-19 versus Normal for an unbalanced database shown in Table 10 is 81.3%. When compared to the same test step for the balanced database, this value is lower than the

minimum accuracy value obtained (90.6%), as shown in Table 8. This comparison demonstrates the influence of the database and how balancing contributes to a significant increase in the accuracy value.

Table 11 shows the accuracy results referring to the second step of the methodology for the classification of the Covid-19 versus Pneumonia experimentation group with the unbalanced database.

Table 11: Accuracy results (%) for test in the DACA - Unbalanced phase of combinations of Data Augmentation for Pneumonia × Covid-19.

Rot.	Brig.	Zoom	Hor. Flip	Ac_1(%)	Ac_2(%)	Ac_3(%)	Avg.(%)
60	[0.3 1.0]	[0.3 1.0]	True	65.6	68.8	75.0	**69.8**
60	[0.2 0.8]	[0.3 1.0]	False	71.9	71.9	53.1	65.6
60	[0.3 1.0]	[0.2 0.8]	False	65.6	51.3	56.3	58.3

When comparing the accuracy values obtained in Table 9 and Table 11 it is noted that the accuracy for unbalanced database is lower (maximum of 75.0%). In contrast, the maximum accuracy for the balanced database was 100.0% (see Table 9) . This comparison again demonstrates how database balancing influences the classification of x-ray images for Covid-19 detection.

4.2.3 Results of Data Augmentation Single Analysis (DASA)

Table 12 presents the results of the DASA Phase in the Test stage for the first experimental group (Normal × Covid-19).

Table 12: Accuracy results (%) for Testing in the DASA Phase of combinations of Data Augmentation for Normal × Covid-19.

Rot.	Zoom	Brig.	Ac_1(%)	Ac_2(%)	Ac_3(%)	Avg.(%)
20	-	-	96.88	100.00	100.00	98.96
-	[0.6 0.8]	-	53.13	34.38	75.00	54.17
-	-	[0.6 0.8]	100.00	100.00	100.00	**100.00**

Rot.: rotation; Brig.: brightness.

The best accuracy values of the training and validation steps were selected for each hyperparameter. Thus, we have again the highlight for the rotation, which presented accuracy of 100%, with an average of 98.96%, and for the brightness, where the average was 100%. For zoom, the results were not adequate, since its average was 54.17%, with a maximum value in the repetitions of 75%.

Table 13 shows the results obtained with the combinations of width shift range and height shift range.

Table 13: Accuracy results (%) for Testing in the DASA Phase with transformations of shift range (width and height) for Normal × Covid-19.

Wid.	Hei.	Ac_1(%)	Ac_2(%)	Ac_3(%)	Avg.(%)
0.2	-	93.7	100.0	100.0	97.9
0.4	-	100.0	93.7	100.0	97.9
0.6	-	96.9	100.0	96.9	97.9
0.8	-	96.9	96.9	96.9	96.9
-	0.2	100.0	96.9	100.0	99.0
-	0.4	96.9	100.0	100.0	99.0
-	0.6	100.0	100.0	100.0	100.0
-	0.8	100.0	100.0	93.7	97.9

Wid.: width shift range; Hei.: height shift range.

On the other hand, Table 14 shows the accuracy values for test phase of the second experimental group (Pneumonia × Covid-19).

It can be seen that both the rotation of 80 degree and brightness from [0.2 0.4] obtained maximum values of 100% in one of the experiments and got an equal maximum average of 96.88%. The accuracy of the zoom transformation from [0.6 0.8] had a maximum average accuracy of 87.50%.

4.3 Comparison between Balanced and Unbalanced

In this section, the adoption of balanced and unbalanced databases is compared. For this, data augmentation is not applied. The results of this step can be seen in Table 15.

Table 14: Accuracy Results (%) for test in the DASA phase of combinations of Data Augmentation for Pneumonia × Covid-19.

Rot.	Zoom	Brig.	$Ac_1(\%)$	$Ac_2(\%)$	$Ac_3(\%)$	Avg.(%)
80	-	-	100.00	100.00	90.63	**96.88**
-	[0.6 0.8]	-	90.63	81.25	90.63	87.50
-	-	[0.2 0.4]	100.00	93.75	96.88	**96.88**

Rot.: rotation; Brig.: brightness.

Table 15: Accuracy Results (%) for Testing for Covid-19 × Normal and Covid-19 × Pneumonia.

		$Ac_1(\%)$	$Ac_2(\%)$	$Ac_3(\%)$	$Avg.(\%)$
Covid-19 × Normal	Balanced	100.0	100.0	100.0	100.0
	Unbalanced	46.9	65.6	84.4	65.6
Covid-19 × Pneumonia	Balanced	93.8	96.9	93.8	94.8
	Unbalanced	71.9	37.5	37.5	49.0

Table 15 shows the accuracy values for the test step for the group Covid-19 versus Normal and Covid-19 versus Pneumonia for the balanced and unbalanced database. Through the same it is possible to observe that the accuracy values for the balanced database is higher than the unbalanced one, having an average of 100.0% while for the unbalanced database of 65.6%. Similarly as before, in the Covid-19 versus Pneumonia the accuracy values for the balanced tests were higher compared to the unbalanced, with an average of 94.8% for the balanced while the unbalanced with 49.0%. It is then possible to state, that database balancing contributes to an improvement in accuracy for Covid-19 detection on lung X-ray images without the use of Data Augmentation.

4.4 Summary of Main Results and Recommended Configurations

Therefore, the best results for all the phases presented in this paper are are summarized in Table 16.

Table 16: Summary of main results For accuracy results (%) for testing: Covid-19 × Normal and Covid-19 × Pneumonia.

		With Combinations of D.A	With Individual D.A	Without D.A
Covid-19 × Normal	Balanced	95.8	100.0	100.0
	Unbalanced	-	68.8	65.6
Covid-19 × Pneumonia	Balanced	89.6	96.88	94.8
	Unbalanced	-	69.8	49.0

D.A: Data Augmentation.

As is apparent from the Sections 4.2.1, 4.2.2, 4.2.3, and 4.3, the configurations that used the balanced dataset obtained the best results for Covid-19 × Normal and Covid-19 × Pneumonia, realized when Data Augmentation was used and in its absence. The highlights are Covid-19 × Normal balanced with individual analyses of data augmentation and without Data Augmentation, which obtained an average accuracy of 100%, and Covid-19 × Pneumonia balanced with individual use of Data Augmentation with an average accuracy of 96.88%.

For the results using combinations of Data Augmentation, the combination of (60; [0.3 1.0]; [0.3 1.0]; False) in Covid-19 × Normal is highlighted, where it is possible to perceive an average accuracy of 95.8%, and for Covid-19 × Pneumonia the combination (60; [0.3 1.0]; [0.3 1.0]; True), which obtained an average accuracy of 89.6%, being Rotation, Brightness, Zoom and Horizontal Flip the sequence of the hyperparameters.

5 Comparison with other papers

In this section, Table 17 presents the comparison between the present proposal and other works using Deep Learning in Covid-19 classification on lung X-ray images: I [10], II [11], III [42] and e IV [23].

Table 17 reveals that studies I, II and III are dedicated to analyzing how the change of CNN architecture interferes with accuracy. On the other hand, work IV and the proposal of this paper aim to analyze the influence of Data Augmentation. In this sense, an important contribution of this work is the application of Deep Learning in the classification of lung X-ray images for Covid-19 diagnosis, considering the influence of Data Augmentation hyperparameters.

It is also worth noting that the present proposal adopts a methodology to select the relevant Data Augmentation hyperparameters in the training and validation stages of a CNN model : *rotation*, *brightness*, *zoom*, *flipping*, *width shift* and *height shift*. The other studies reviewed in Table 17 use only some of these hyperparameters.

Table 17: Comparison of the present proposal (Prop.) with other works using *Deep Learning* in Covid-19 classification on lung X-ray images: I [10], II [11], III [42] e IV [23].

		Prop.	I	II	III	IV
Proposed Method	Data Augmentation Influence	✓	-	-	-	✓
	Architectural Influence	-	✓	✓	✓	-
Data Augmentation Hyperparameters	Rotation	✓	✓	✓	✓	✓
	Brightness	✓	-	-	-	-
	Zoom	✓	-	✓	-	✓
	Flipping	✓	-	-	✓	✓
	Width Shift	✓	-	-	-	-
	Height Shift	✓	-	-	-	-
Library	Keras	✓	-	✓	✓	✓
	FastAi	-	✓	-	-	-
Image Classes	Normal	✓	✓	✓	✓	✓
	Pneumonia	✓	✓	✓	✓	-
	Covid-19	✓	✓	✓	✓	✓
Max. accuracy	= 100%	✓	✓	-	-	✓

Also noteworthy is the use of the *Keras* library, which is used in the present paper and in most of the papers analyzed in Table 17, with the exception of paper of [10], which uses the *FastAi* library.

Finally, it is noteworthy that studies I, II and III, including the proposal presented, use the same classifications for the images: Normal, Pneumonia and Covid-19, except for work IV, which does not present the Pneumonia group. In addition, the selected CNN model achieved accuracy equal to 100% for both groups of experiments in the testing stage. This level of hit rate (= 100%) was only presented in the results of two of the analyzed studies: [10] and [23]. Thus, indicating a good fit of the proposed model in the task of recognizing Covid-19 in lung x-ray images.

6 Conclusion and Future Works

The aim of this work was to propose a methodology for hyperparameter tuning of Data Augmentation in the classification of lung X-ray images. Thus, the present study focuses its contributions on the following aspects: (i) application of Deep Learning model in the process of Covid-19 detection through lung X-ray images in databases with few samples; (ii) proposal of a methodology for tuning of Data Augmentation hyperparameters, in combination with each other and individually, for this application; (iii) proposal for a methodology to analyze the influence of database balancing with and without Data Augmentation (iv) maximum accuracy equal to 100% in the testing stage, in both Phases, Data Augmentation Combined Analysis and Data Augmentation Single Analysis, experimental after the adjustment of the hyperparameters.

The results found demonstrate the influence that the Data Augmentation hyperparameters have on the classification accuracy of X-ray images of: i) healthy patients, ii) patients with pneumonia and iii) patients with Covid-19. With the variation of 4 hyperparameters and a total of 16 combinations, accuracy values between 53.1% and 91.7% for the Normal versus Covid-19 classification and between 55.2% and 88.5% for the Pneumonia versus Covid-19 classification were obtained for the validation step. After choosing the three combinations with the highest accuracy and performing the testing step, the mean accuracy results increased, reaching 95.8% for the Normal versus Covid-19 classification and 89.6% for the Pneumonia versus Covid-19 classification. It is worth noting that in both classifications it was still possible to obtain accuracy values of 100%, demonstrating the effectiveness of the hyperparameters chosen for this application. In experiments with unbalanced database the average value found for accuracy was 68.8% (Covid-19 versus Normal) and 69.8% (Covid-19 versus Pneumonia). Thus, again stating the importance of database balancing for greater accuracy in this type of classification.

In the phase of Data Augmentation Single Analysis with the hyperparameters analyzed individually were obtained average accuracy results of 98.96% for angular rotation and brightness rate (training and validation steps) for the Normal versus Covid-19 classification. In the test stage, mean accuracy values of 98.96% and 100% were obtained, again for the angular rotation and brightness rate hyperparameters. For the Pneumonia versus Covid-19 group it is also highlighted the accuracy averages above 95%, again for rotation and brightness rate: train (97.92%), validation (96.88%) and test (96.88%).

With this, it is stated that the recommended method to classify lung X-ray images for Covid-19 detection using Data Augmentation is the use of a balanced database and combinations of hyperparameters such as rotation, brightness, zoom, and horizontal flip for Covid-19 versus Normal and Covid-19 versus Pneumonia classification. For the use of individual transformations, it is recommended for the Covid-19 versus Pneumonia group the use of rotation, zoom and brightness. Besides these, it is also recommended for Covid-19 versus Normal the transformations of height shift and weight shift.

For future work, it is suggested to apply the analysis to other Convolutional Neural Network models, analyzing how the change in architecture also influences. As well as analyze the influence of other Data Augmentation hyperparameters on this type of classification, checking their individual behavior and when combined with other transformations. Furthermore, it is suggested to analyze with other databases, in order to verify the model performance beyond the images initially used and to validate the accuracy hypothesis. Finally, it is recommended to use other statistical measures for analysis (standard deviation, median and variance) and apply other methodologies for hyperparameter tuning [21, 39].

REFERENCES

[1] A. Moreira and L. Pinheiro. "OMS declara pandemia de coronavírus", 2020.

[2] D. B. Ribeiro, R. A. X. Ferreira, A. B. L. Marins, P. J. B. da Silveira Júnior and A. C. P. de Mello Pinheiro. "COVID-19: UM ESCLARECIMENTO FRENTE AOS MÉTODOS DIAGNÓSTICOS". *Revista Transformar*, vol. 14, no. 2, pp. 205–215, 2020.

[3] T. Singhal. "A review of coronavirus disease-2019 (COVID-19)". *The indian journal of pediatrics*, vol. 87, no. 4, pp. 281–286, 2020.

[4] A. S. R. Souza, M. M. R. Amorim, A. S. d. O. Melo, A. M. Delgado, A. C. M. C. d. C. Florêncio, T. V. d. Oliveira, L. C. S. Lira, L. M. d. S. Sales, G. A. Souza, B. C. P. d. Melo *et al.*. "Aspectos gerais da pandemia de COVID-19". *Revista Brasileira de Saúde Materno Infantil*, vol. 21, pp. 29–45, 2021.

[5] V. d. C. Dias, M. Carneiro, L. Michelin, C. Vidal, L. Costa, C. Ferreira *et al.*. "Testes sorológicos para COVID-19: interpretação e aplicações práticas". *J. Infect. Control*, vol. 9, 2020.

[6] L. M. F. Vieira, E. Emery and A. Andriolo. "COVID-19: laboratory diagnosis for clinicians. An updating article". *Sao Paulo Medical Journal*, vol. 138, no. 3, pp. 259–266, 2020.

[7] L. Magno, T. A. Rossi, F. W. d. Mendonça-Lima, C. C. d. Santos, G. B. Campos, L. M. Marques, M. Pereira, N. M. d. B. L. Prado and I. Dourado. "Desafios e propostas para ampliação da testagem e diagnóstico para COVID-19 no Brasil". *Ciência & Saúde Coletiva*, vol. 25, pp. 3355–3364, 2020.

[8] R. Caetano, A. B. Silva, A. C. C. M. Guedes, C. C. N. d. Paiva, G. d. R. Ribeiro, D. L. Santos and R. M. d. Silva. "Desafios e oportunidades para telessaúde em tempos da pandemia pela COVID-19: uma reflexão sobre os espaços e iniciativas no contexto brasileiro". *Cadernos de Saúde Pública*, vol. 36, pp. e00088920, 2020.

[9] L. O. Silva, L. S. Araújo, V. F. Souza, R. M. Barros Neto and A. Santos. "Comparative Analysis of Convolutional Neural Networks Applied in the Detection of Pneumonia Through X-Ray Images of Children". *Learning & Nonlinear Models*, vol. 18, no. 2, pp. 4–15, 2020.

[10] G. Jain, D. Mittal, D. Thakur and M. K. Mittal. "A deep learning approach to detect Covid-19 coronavirus with X-Ray images". *Biocybernetics and biomedical engineering*, vol. 40, no. 4, pp. 1391–1405, 2020.

[11] R. Jain, M. Gupta, S. Taneja and D. J. Hemanth. "Deep learning based detection and analysis of COVID-19 on chest X-ray images". *Applied Intelligence*, vol. 51, no. 3, pp. 1690–1700, 2021.

[12] P. Endo, I. R. Silva, G. L. Santos and D. Sadok. "Classifying COVID-19 positive X-ray using deep learning models". *IEEE Latin America Transactions*, vol. 100, no. 1e, 2020.

[13] S. Basu, S. Mitra and N. Saha. "Deep learning for screening covid-19 using chest x-ray images". In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 2521–2527. IEEE, 2020.

[14] O. L. de Sousa, D. M. Magalhães, P. d. A. Vieira and R. Silva. "Deep Learning in Image Analysis for COVID-19 Diagnosis: a Survey". *IEEE Latin America Transactions*, vol. 100, no. 1e, 2020.

[15] L. O. Hall, R. Paul, D. B. Goldgof and G. M. Goldgof. "Finding covid-19 from chest x-rays using deep learning on a small dataset". *arXiv preprint arXiv:2004.02060*, 2020.

[16] M. M. A. Monshi, J. Poon, V. Chung and F. M. Monshi. "CovidXrayNet: Optimizing data augmentation and CNN hyperparameters for improved COVID-19 detection from CXR". *Computers in Biology and Medicine*, vol. 133, pp. 104375, 2021.

[17] C. Ouchicha, O. Ammor and M. Meknassi. "CVDNet: A novel deep learning architecture for detection of coronavirus (Covid-19) from chest x-ray images". *Chaos, Solitons & Fractals*, vol. 140, pp. 110245, 2020.

[18] C. Shorten and T. M. Khoshgoftaar. "A survey on image data augmentation for deep learning". *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019.

[19] L. Taylor and G. Nitschke. "Improving deep learning using generic data augmentation". *arXiv preprint arXiv:1708.06020*, 2017.

[20] L. Perez and J. Wang. "The effectiveness of data augmentation in image classification using deep learning". *arXiv preprint arXiv:1712.04621*, 2017.

[21] A. L. C. Ottoni, R. M. de Amorim, M. S. Novo and D. B. Costa. "Tuning of data augmentation hyperparameters in deep learning to building construction image classification with small datasets". *International Journal of Machine Learning and Cybernetics*, pp. 1–16, 2022.

[22] F. Chollet and J. Allaire. *Deep Learning with R*. Manning Publications, 2018.

[23] A. Sedik, A. M. Iliyasu, A. El-Rahiem, M. E. Abdel Samea, A. Abdel-Raheem, M. Hammad, J. Peng, A. El-Samie, E. Fathi, A. A. A. El-Latif *et al.*. "Deploying machine and deep learning models for efficient data-augmented detection of COVID-19 infections". *Viruses*, vol. 12, no. 7, pp. 769, 2020.

[24] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman and P. R. Pinheiro. "Covidgan: data augmentation using auxiliary classifier gan for improved covid-19 detection". *Ieee Access*, vol. 8, pp. 91916–91923, 2020.

[25] F. Chollet. *Deep Learning with Python*. Manning Publications, 2017.

[26] A. L. C. Ottoni and M. Silva Novo. "A Deep Learning Approach to Vegetation Images Recognition in Buildings: a Hyperparameter Tuning Case Study". *IEEE Latin America Transactions*, vol. 19, no. 12, pp. 2062–2070, 2021.

[27] P. Rici, S. O. S. Santos and A. L. C. Ottoni. "Análise da seleção de hiperparâmetros de Data Augmentation na detecção de Covid-19 em imagens de raio-x com Deep Learning". In *Anais do 15 Congresso Brasileiro de Inteligência Computacional*, pp. 1–7, Joinville, SC, 2021. SBIC.

[28] D. H. S. I. N. d. Silva and R. A. Flauzino. *Redes neurais artificiais para engenharia e ciências aplicadas*. Artliber Editora, second edition, 2016.

[29] S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Pearson Education, third edition, 2010.

[30] C. Aggarwal. *Neural Networks and Deep Learning: A Textbook*. Springer International Publishing, 2018.

[31] D. V. Minatto, S. Coral and M. C. de Mattos Garcia. "Análise de Características em Redes Neurais Artificiais Aplicadas em Visão Computacional com Ênfase em Reconhecimento Facial". *Anais SULCOMP*, vol. 8, 2017.

[32] S. Peixoto, G. Cámara-Chávez, D. Menotti, G. Gonçalves and W. Schwartz. "Brazilian license plate character recognition using deep learning". In *Proceedings of the XI Workshop de Visão Computacional—WVC*, 2015.

[33] C. S. e Napoleão Galegale e Marcelo Okano e Celi Langhi. "Deep Learning Aplicado À Negociação De Ações Por Algoritmos: Uma Revisão Descritiva Da Literatura". *South American Development Society Journal*, vol. 6, no. 17, pp. 237, 2020.

[34] A. M. Moura. "Redes neurais recorrentes aplicadas à previsão de preços de ativos da BOVESPA: comparativo de modelos para regressão". *Matemática Bacharelado-Unisul Virtual*, 2019.

[35] D. G. Bernardes and O. L. V. Costa. "Carteiras de Black Litterman com Análises Baseadas em Redes Neurais". *Learning & Nonlinear Models*, vol. 18, no. 1, pp. 60–75, 2020.

[36] T. Okuyama, T. Gonsalves and J. Upadhay. "Autonomous driving system based on deep Q learnig". In *2018 International Conference on Intelligent Autonomous Systems (ICoIAS)*, pp. 201–205. IEEE, 2018.

[37] M. G. Bechtel, E. McEllhiney, M. Kim and H. Yun. "Deeppicar: A low-cost deep neural network-based autonomous car". In *2018 IEEE 24th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA)*, pp. 11–21. IEEE, 2018.

[38] F. X. Viana, G. M. Araujo, M. F. Pinto, J. Colares and D. B. Haddad. "Aerial Image Instance Segmentation Through Synthetic Data Using Deep Learning". *Learning & Nonlinear Models*, vol. 18, no. 1, pp. 35–46, 2020.

[39] A. L. C. Ottoni, M. S. Novo and D. B. Costa. "Hyperparameter tuning of convolutional neural networks for building construction image classification". *The Visual Computer*, pp. 1–15, 2022.

[40] M. Elgendy. *Deep Learning for Vision Systems*. Manning Publications, 2020.

[41] A. Rosebrock. *Deep Learning for Computer Vision with Python: Starter Bundle*. PyImageSearch, 2017.

[42] H. Panwar, P. Gupta, M. K. Siddiqui, R. Morales-Menendez and V. Singh. "Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet". *Chaos, Solitons & Fractals*, vol. 138, pp. 109944, 2020.