



# A new model for classification of medical CT images using CNN: a COVID-19 case study

Pedro Moises de Sousa<sup>1</sup> · Pedro Cunha Carneiro<sup>1</sup> · Gabrielle Macedo Pereira<sup>1</sup> · Mariane Modesto Oliveira<sup>1</sup> · Carlos Alberto da Costa Junior<sup>1</sup> · Luis Vinicius de Moura<sup>2</sup> · Christian Mattjie<sup>2</sup> · Ana Maria Marques da Silva<sup>2</sup> · Túlio Augusto Alves Macedo<sup>3</sup> · Ana Claudia Patrocínio<sup>1</sup>

Received: 6 April 2022 / Revised: 18 November 2022 / Accepted: 10 December 2022 /

Published online: 19 December 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

SARS-CoV-2 is the causative agent of COVID-19 and leaves characteristic impressions on chest Computed Tomography (CT) images in infected patients and this analysis is performed by radiologists through visual reading of lung images, and failures may occur. In this article, we propose a classification model, called Wavelet Convolutional Neural Network (WCNN) that aims to improve the differentiation of images of patients with COVID-19 from images of patients with other lung infections. The WCNN model was based on a Convolutional Neural Network (CNN) and wavelet transform. The model proposes a new input layer added to the neural network, which was called Wave layer. The hyperparameters values were defined by ablation tests. WCNN was applied to chest CT images to images from two internal and one external repositories. For all repositories, the average results of Accuracy (ACC), Sensitivity (Sen) and Specificity (Sp) were calculated. Subsequently, the average results of the repositories were consolidated, and the final values were ACC = 0.9819, Sen = 0.9783 and Sp = 0.98. The WCNN model uses a new Wave input layer, which standardizes the network input, without using data augmentation, resizing and segmentation techniques, maintaining the integrity of the tomographic image analysis. Thus, applications developed based on WCNN have the potential to assist radiologists with a second opinion in the analysis.

**Keywords** Chest CT images · Convolutional neural networks · COVID-19 · Wavelet · WCNN

## 1 Introduction

Recently, tests have detected a type of virus in patients' lung fluids, and this led to the discovery of a new coronavirus (CoV). Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus which belongs to the *Coronaviridae* family. It can

---

✉ Pedro Moises de Sousa  
pedrosousa@ufu.br

cause respiratory, enteric, hepatic, and neurological diseases both in domestic animals and people [52, 57]. They also have a phylogenetic relationship with the coronaviruses that cause Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) [54, 57]. COVID-19, SARS and MERS were zoonotic in origin and their viruses were transmitted by bats, civets, and camels, respectively.

Due to the global impact of the COVID-19 pandemic, international efforts have been made to simplify researchers' access to viral data through repositories such as the 2019 Novel Coronavirus Resource (2019nCoV) [53] and National Center for Biotechnology Information (NCBI) [43]. The more accessible the information, the more likely it is that a set of medical countermeasures will be rapidly developed to control the disease worldwide, as has happened with other diseases on other occasions [37, 38, 45]. In this context, computed tomography (CT) has been performed as an initial modality for screening patients, as it allows the visualization of abnormal anatomy [5, 33].

CT images show similarities between patients with COVID-19 and others type of viral pneumonia, such as SARS and MERS. Nevertheless, analysis of CT images evolves a one-by-one tedious procedure. The radiologists who analyze, sometimes, thousands of image slices per patient faces a high environmental risk of human error events occurrence. In this period of the COVID-19 pandemic, the pressure for results and the radiologists' time limitations are elements that make analyses more error-prone.

As the use of deep learning is not new in radiology CT imaging research [3, 21, 28, 50], as soon as the pandemic was declared by the WHO in March 2019, researchers around the world started creating computer models that processed radiological images of patients with COVID-19 [3, 5, 33, 56]. Chest CT images have common features that may show a specific pattern for COVID-19, however manual analysis is time consuming for the radiologist. To speed up the analysis and reduce the probability of error, we assume that deep learning can be effective in analyzing large volumes of data generated by CT images [4, 6, 18, 28, 32, 35, 41, 56].

In this paper we describe a new model for classifying CT images based on deep learning, called WCNN. The aim is to improve the differentiation of images from patient with COVID-19. In this paper, images of patients diagnosed with COVID-19 comprises the COVID-19+ image base. The images of patients with other lung infections or inflammatory diseases such as pneumonia, cardiomegaly, pleural effusion, atelectasis, and consolidation are inserted in the COVID-19- image base.

WCNN was created using CNN approaches well established in literature [19, 48, 55]. It avoids pre-processing operations as image resizing and data augmentation. Otherwise, we proposed an additional layer to the convolutional network, which we called wave layer. The proposed layer uses the wavelet transform to decompose the image and extract its characteristics, being responsible for the pre-processing and generation of the output image that will be processed by other model's layers.

The main contributions our research has to offer to our field of research can be divided in two categories: wave layer and overall WCNN model. Regarding the wave layer contributions, the follow ones stand out:

- development entirely based on TensorFlow, replacing the Keras input layer;
- automated data entry standardization;
- capability of noise reduction filters application for medical images;
- capacity to embody various feature extraction techniques;
- use of crops to obtain the best region of the organ under study.

On the other hand, the most relevant contributions of the WCNN are the following ones:

- test, in the same research, of two image databases for model development, being one public and another private;
- use of an external database aimed to independently validate the model;
- creation and application of objective criteria for the inclusion and exclusion of images from the image bases;
- use of images in 16 bits, keeping the necessary information for the characterization of the disease;
- neither data augmentation nor image resize were required to accurately discriminate the disease; and above all;
- WCNN has great potential to reduce the clinical workload of radiologists, serving as a first or second analyst.

We organized this paper in sections, being the second one (Related Works) responsible to present related works to our research and the third one describing basic topics about wavelet transforms. Fourth, fifth, sixth, seventh and eighth sections describes the model functioning workflow, the process of image base creation, WCNN model core foundations, ablation tests and training parameters and model evaluation metrics, respectively. After this, Section 9 section presents the results by WCNN model application, Section 10 discusses the research outfits and, finally, Section 11 concludes the paper with the final considerations and future works ideas.

## 2 Related works

As quickly as the pandemic spread, governments, supranational organizations, research institutes, universities and corporations mobilized unprecedented amounts of human and financial resources to end the crisis. In a flash, both SARS-CoV-2 and COVID-19 were raised to the height of interest of the world scientific research community in the most diverse areas i.e., medicine, infectiology, biochemistry, information technologies, applied mathematics, artificial intelligence. Our research was conducted in this context, motivated by the intrinsic urgency of ending the pandemic; and the same happened with the initiative of other researchers, whose results are presented and discussed in this section.

The first work analyzed proposes an algorithm based on transforms and CNN for CT image recognition [6]. The authors present a solution with two branches: Trans-CNN model and Transformer module. The Trans-CNN model uses CNN's local resource extraction capability and Transformer's global resource extraction capability. The survey consisted of 194,922 chest CT images of 3745 patients aged from 0 to 93 years, extracted from the COVIDx-CT database. Images include patients i) healthy, ii) sick with COVID-19; and iii) patients with other lung diseases. The base was expanded by 15°, 45°, 90° and 180° rotations. The values obtained for accuracy, sensitivity and specificity were 0.9673, 0.9776 and 0.9602, respectively.

The COVID-CT-Mask-Net model also uses CNN [41]. It is performed in two steps: *i)* the Mask R-CNN network is trained to locate and detect regions of Ground Glass Opacity lesions on CT images; and *ii)* images of these lesions are merged to classify the input image. The experiment used 3000 chest CT images from the COVIDx-CT database, whose patients can be *i)* healthy, *ii)* sick with COVID-19; or, *iii)* patients with other pulmonary infections. The metric

values of accuracy, sensitivity and specificity were calculated and resulted in 0.9673, 0.9776 and 0.9602, respectively.

Another relevant work describes a hybrid model that combines *queezeNet* and *ShuffleNet*. It uses 1252 COVID-19+ CT images and 1230 COVID-19- images from the public SARS-COV-2 Ct-Scan database, which were collected from real patients from hospitals in São Paulo, Brazil [35]. The data were expanded by performing random operations of *i*) rotation of  $\pm 5^\circ$ ; *ii*) change in intensity value of  $\pm 20$ ; and *iii*) shear of  $\pm 20^\circ$ . In addition to random operations, *i*) blurring; *ii*) inversion; and *iii*) resizing to  $224 \times 224$  spatial resolution, were also performed on the images. The respective results for accuracy, sensitivity and specificity are 0.9781, 0.9615 and 0.9608.

Another example of research in which neural networks other than CNN were used was described in the article [32]. The model proposed by the authors combines the VGG-16, GoogleNet and ResNet-50 networks; aims to detect COVID-19 in the initial phase. It obtained an accuracy of 0.9827, a sensitivity of 0.9893 and a specificity of 0.9760. 150 chest CT images belonging to the *Società Italiana di Radiologia Medica e Interventistica* were used. They gave rise to 3000 images, grouped into Subset-1, called “COVID-19”, and Subset-2, labeled “No findings”. The resolution of the subset images is  $16 \times 16$  and  $32 \times 32$ , respectively.

Research comparing the performance of various convolutional network architectures stands out [18]. It involves VGG16, DenseNet121, MobileNet, NASNet, Xception and EfficientNet networks. The study used chest CT images obtained from Kaggle, being 1958 from COVID-19+ patients and 1915 from COVID-19-. The image base was expanded by resizing to  $224 \times 224$  spatial resolution. The model was trained with 70% of the images, validated with 15% and evaluated with the other 15%. Of these architectures, VGG16 presented the best results, with an accuracy of 0.9768, sensitivity of 0.9579 and specificity of 0.9971.

Another related work describes the creation of an application for detecting pneumonia caused by COVID-19 through high resolution CT analysis [3]. It was created by staff at Renmin Hospital, University of Wuhan, China. The base model of the application uses architecture derived from UNet++. Application performance was measured using 46,096 anonymous images from 106 hospital patients, grouped into two groups. The first group, with 51 COVID-19+ patients; and the second, used as a control group, with 55 COVID-19- patients. In addition, the authors retrospectively used the images of twenty-seven patients seen before the start of the project to compare the effectiveness of the diagnosis made by experts with the effectiveness obtained by the application. The result of accuracy, sensitivity and specificity of the application were, respectively, 0.9524, 1.0000 and 0.9355. Considering the twenty-seven previous image patients, the accuracy, sensitivity, and specificity achieved values of 0.9885, 0.9434, and 0.9916, respectively. So, there was demonstrated that application’s performance is compatible with medical experts obtained results.

The COVID-19-CNN model combines the use of previously trained CNNs [4]. Training and performance testing of this model used images from 405 COVID-19+ patients and 397 COVID-19- patients. 612 images were used for training, ninety-nine for validation and ninety-one for testing. The database was not expanded, but the image was scaled to a spatial resolution of  $224 \times 224$ . The COVID-19-CNN model had an accuracy of 0.9670, sensitivity of 0.9780 and specificity of 0.9556.

Feature extraction is an important stage of the process, thus, diverse robust CNN architectures are implemented, for example: DenseNet, VGGNet, InceptionV3 and ResNet [15]. Regarding deep learning based feature extraction, recent studies have been using different methods to deal with a variety of problems [14, 22, 23, 31]. In one of these studies, the goal was to segment objects from relational visual data [26]. For feature extraction, they used

convolution blocks of DeepLabV3 [9] that applies atrous convolution to extract dense feature maps with the use of upsampled filters. ResNet50 [42] and VGG16 [46] were used in order to assess the influence of backbone feature extraction networks in deep models for visual tracking [25].

There are also studies about video object segmentation and resource association. Those will train a prototypical siamese network in order to find the pixel or resource which is the closest association to the first frame or segmented frame, as well as the reference frame. Then, they provide their corresponding labels [24, 27].

According to the characteristics of state-of-the-art research, the main contributions given by our work include: *i*) systematization of the concomitant use of public and private image banks, in the training and testing phases of the network, and creation of an external test base for independent validation of the model; *ii*) application of objective criteria for the inclusion and exclusion of images from these databases; *iii*) use of images allocated in 16 bits, containing the necessary information for the characterization of the disease; and, *iv*) WCNN did not use data augmentation or image resizing to classify COVID-19+ images [4, 6, 18, 35].

### 3 Wavelet

In this section, we present the wavelet transform theory, inspired by small waves (wavelets) of varying frequency and limited duration. Related works have shown that the use of convolutional networks is common in image classification models. The use of Discrete Wavelet Transform (DWT) is also not uncommon, however, the use we made of the DWT gave WCNN characteristics that positively impacted its performance. The DWT is a sequence of numbers that shows a certain continuous function. In addition to its efficient and intuitive structure for representing and storing multi-resolution images, the wavelet transform provides insight into the spatial and frequency characteristics of an image [7, 8, 11, 29]. Let the image be  $f(a, b)$ , the discrete wavelet transform (DWT) of this image is defined in Eq. 1, as [7, 8, 11, 29]:

$$Wf(a, b) = \int_{-\infty}^{+\infty} f(t) \psi_{a,b}(t) dt \quad (1)$$

For a discrete  $n$ -point signal, the DWT integral can take the summation form, as in Eq. 2:

$$Wf(a, b) = \sum_{t=0}^{N-1} f(t) \psi_{a,b}(t) \quad (2)$$

The wavelet function  $\psi_{a, b}(t)$  is derived from the function  $\psi(t)$ , through the transformation shown in Eq. 3:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (3)$$

where  $a \in R^+$ ,  $b \in R$  and  $\psi\left(\frac{t-b}{a}\right)$  are the wavelet bases, where  $b$  represents the wavelet translation and  $a$  is the scale parameter associated with the width.

There is a great possibility of choice for the function  $\psi(t)$ , called mother wavelet, among which there are Daubechies, Symlets, Coiflet. The scaled and shifted versions of the mother wavelet correspond to bandpass filters with different bandwidths and time durations. The wavelet transform performs a transformation step on each line, producing a matrix; where the left side contains the down-sampled low-pass coefficients (L) of each row, and the right side contains the

high-pass coefficients (H) (Fig. 1a). Then, a step is applied to each column (Fig. 1b), resulting in four types of coefficients, as shown in (Fig. 1c) [7, 8, 11, 29]:

- Coefficients that result from a convolution with high pass in both directions (HH) represent diagonal features of the image.
- Coefficients resulting from a high-pass convolution on the columns after a lowpass convolution on the rows (HL) correspond to the horizontal characteristics of the image.
- Coefficients originated from high-pass filters in the rows followed by lowpass filters in the columns (LH) correspond to the vertical characteristics of the image; and,
- Coefficients from lowpass filters in both directions (LL) correspond to the approximate characteristics of the image.

## 4 Workflow and model

Classification models based on deep learning are a constituent part of workflows that aim to detect characteristic patterns in images from a database. The workflow that we used in this research has steps for creating the base, executing the model, consolidating the result, and calculating the metrics. Figure 2 illustrates the workflow of which the WCNN model execution is part.

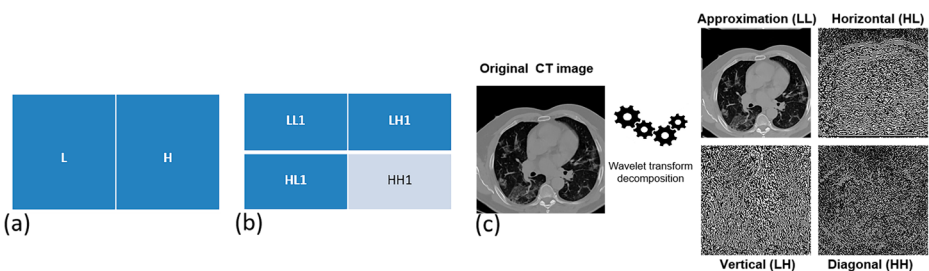
The workflow was divided in two major components/processes, *Materials* and *WCNN*. The *Materials* notably contains the selection and distribution of the datasets and images, which are needed for *WCNN*. On other hand, *WCNN* encapsulates all process regarding the wave layer, feature extraction, flatten and fully connected layer. In the subsequent sections, all workflow process and respective steps will be described, starting with datasets creation.

## 5 Image base creation

This section presents the image inclusion and exclusion criteria and the datasets used in the research.

### 5.1 Inclusion and exclusion criteria

The first step in creating the image base was defining the image inclusion and exclusion criteria for the COVID-19+ and COVID-19- bases. The inclusion criteria are i) the



**Fig. 1** a Transformation step on each line b Transformation step on each column c Diagram overview of the wavelet transform

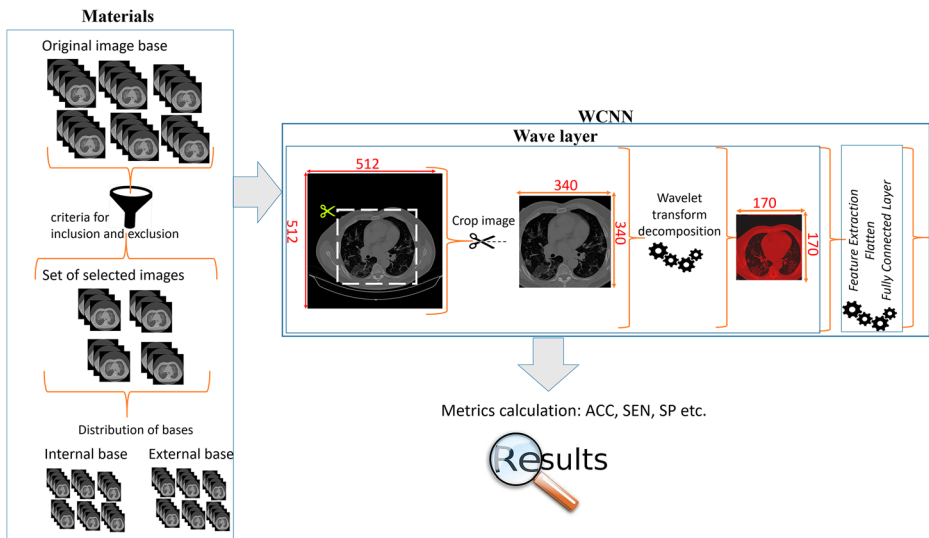


Fig. 2 Workflow overview and the WCNN template

tomographic reconstruction matrix formed by  $512 \times 512$  pixels; ii) patients over 18 years of age; and iii) images of patients testing positive for COVID-19.

As mentioned earlier, images of patients diagnosed with COVID-19 make up the COVID-19+ database and images of patients who evaluated negative are entered into the COVID-19-database. This is because, although the patient is not sick with COVID-19, he has other infectious or inflammatory lung diseases such as pneumonia, cardiomegaly, pleural effusion, atelectasis, consolidation.

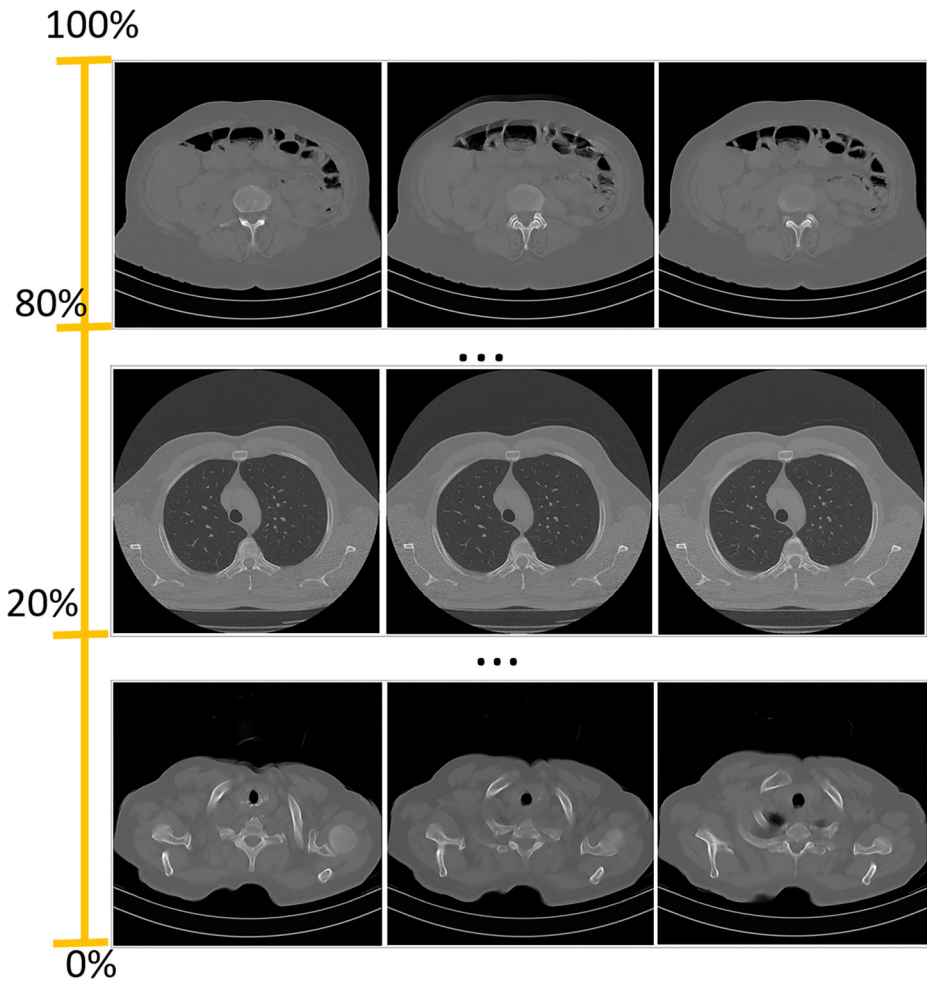
The exclusion criteria application was done with the help of a radiologist. He guided us to discard about 40% of the total amount of slides from each exam, being 20% of the initial slices and 20% of the final slices. This discard helps focus on the area of interest which is the lung, since the initial and final slices do not highlight it sufficiently, as illustrated in Fig. 3.

The following subsections describe the original image repositories as well as the datasets created using the criteria.

### 5.2 Dataset I

Dataset I contain images from the Valencian Region Medical Image Bank (BIMCV) [49] public repository. They were generated between 02/26/2020 and 04/18/2020 and this dataset is divided into:

- BIMCV-COVID-19+ images of patients testing positive for COVID-19, and includes CT radiographic findings and their respective reports, polymerase chain reaction (PCR) test, antibody diagnostic tests (immunoglobulin G-IgG and immunoglobulin M-IgM); and,
- BIMCV-COVID19- images of patients testing negative for COVID-19, including CT radiographic findings and their respective reports, including pathologies such as pneumonia, cardiomegaly, pleural effusion, atelectasis, consolidation.



**Fig. 3** Selection of CT slices

Between 50 and 400 CT slices were used for each exam, whose slice thickness varies from 1 mm to 7 mm. CT radiographs were performed using the following equipment: KONICA MINOLTA 0862; GMM ACCORD DR; Philips Medical Systems DigitalDiagnost; Philips Medical Systems PCR Elevate; SIEMENS SOMATOM; TOSHIBA Aquilion; Philips DigitalDiagnost; Philips Brilliance 16; Philips Medical Systems Essenta DR.

As for the distribution of patients from Dataset I, 174 patients were selected, of which 87 were BIMCV-COVID-19+ patients, constituting the COVID-19+ base, and 88 BIMCV-COVID19- patients, who composed the COVID-19- base. Of the 87 patients in the COVID-19+ database, 70% were used in the training phase, 15% in the testing phase and the remaining 15% in the validation phase. The same distribution was used in the COVID-19-base, as shown in Table 1. The patient images used in the training phase were not used again in the testing and validation phases, that is, the patients were disjointedly divided.



**Table 1** Dataset I image distribution

Sub dataset	Training Patients/Images	Testing Patients/Images	Validation Patients/Images
COVID-19+	62/5670	13/1032	13/1032
COVID-19-	62/5670	13/1032	13/1032

### 5.3 Dataset II

Dataset II is composed by CT images from the private repository of Hospital São Lucas of the Pontifical Catholic University of Rio Grande do Sul (HSL-PUCRS), generated between 03/03/2020 and 07/30/2020. To use the images from this repository, which is private, we submitted a request for use to the PUCRS evaluation committee, under number 30791720.5.0000.5336. The process followed the normal procedures, and, in the end, the demand was approved.

Dataset II is composed of patients who evaluated positive or negative for COVID-19, the latter that has other lung diseases such as pneumonia, cardiomegaly, pleural effusion, atelectasis, consolidation. CT scans were performed using the following equipment: Siemens; GE Medical Systems, Philips Medical Systems, Toshiba. Each exam consists of 50 to 400 CT slices, and each slice has a thickness that varies from 1 mm to 5 mm.

Regarding the distribution of patients from the HSL-PUCRS, sixty patients with a positive test for COVID-19 were selected from the base of Hospital São Lucas at PUCRS. Seventy percent were used in the training phase, 15% in the testing phase and the remaining 15% were used in the validation phase, as shown in Table 2. The patient images used in the training phase were not used again in the testing and validation phases, or that is, the patients were disjointedly divided.

### 5.4 Dataset III

Dataset III is composed of CT images obtained from the private repository of the Hospital de Clínicas of the Federal University of Uberlândia (HC-UFU), Brazil. They were generated between 04/08/2020 and 10/12/2020. Dataset III was used to validate the model with images different from those used in training and testing. Images were collected from twenty patients positive for COVID-19, in a total of 2300 images; and twenty patients negative for COVID-19 but positive for viral pneumonia, also in a total of 2300 images. The same inclusion and exclusion criteria were used in the HC-UFU database. In addition, information was anonymized. The images were obtained by the Toshiba CT Scanner equipment. The scanning parameters were defined as follows: the lung window reconstruction matrix,  $512 \times 512$ ; cutting thickness, 1 mm–7 mm. Table 3 shows the tabulated data.

**Table 2** Dataset II image distribution

Sub dataset	Training Patients/Images	Testing Patients/Images	Validation Patients/Images
COVID-19+	42/3767	9/750	9/740
COVID-19-	42/3767	9/750	9/740

**Table 3** Dataset III image distribution

Sub dataset	Validation patients/images
COVID-19+	20/2300
COVID-19-	20/2300

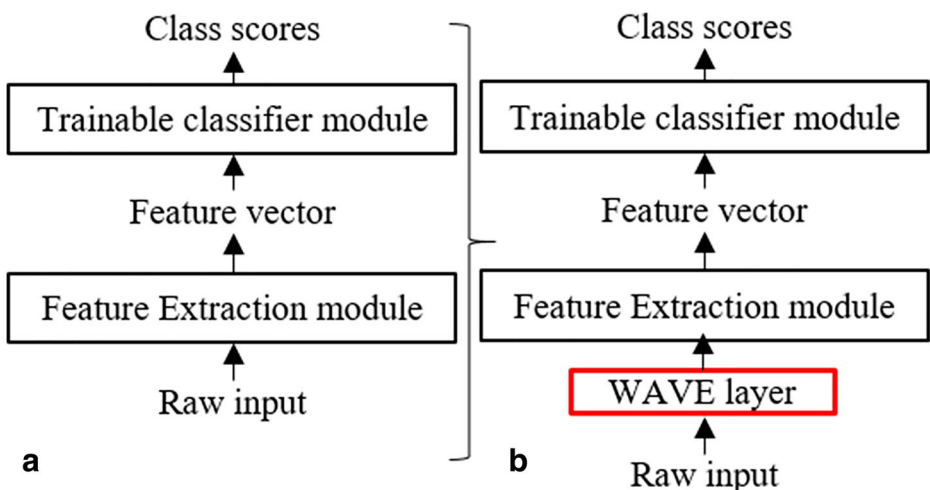
## 6 WCNN model

Convolutional Neural Networks (CNNs) were proposed to assess image data. The name comes from the convolution operator, a straightforward way of doing complex operations using the convolution kernel [36]. Many variations of the CNN were already proposed such as AlexNet [19], Clarifai [55], GoogleNet [48]. WCNN is also a CNN variation and embody the CNN basic architecture as well the customized layer [20].

As illustrated in Fig. 4a, the conventional CNN architecture contains is composed by two modules: a resource extractor, which processes the raw input, and a trainable classifier, which generates the class scores (Adapted from [20]). Otherwise, Fig. 4b is a representation of the architecture of our customized CNN, which contains the same modules of the conventional CNN architecture as well the highlighted new layer we created.

WCNN is composed of four stages: wave layer, feature extraction, flatten layer and fully connected layer. In the CNN, the pooling and convolution layers acts as a stage for feature extraction, whereas the classification stage is made of one or more fully connected layers followed by a sigmoid function layer [51].

Figure 5 illustrates the WCNN classification scheme and the next subsection will detail it functioning and each of its elements.



**Fig. 4** a Conventional CNN architecture and b Architecture of our customized CNN highlighting the wave layer (Adapted from [20])

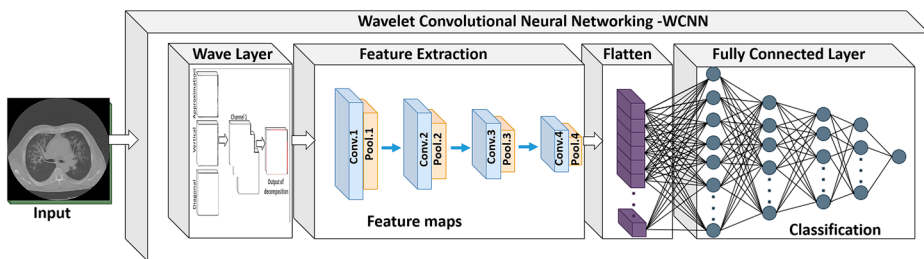


Fig. 5 WCNN classification scheme

## 6.1 Wave layer

The wave layer creation required both selection of the wavelet function as well its respective level of decomposition and the wavelet transform most relevant coefficients analysis. The mother wavelet and the coefficients were chosen after analyzing the available options and considering the best one. In this subsection we will describe, in detail, this analysis as well the processing mechanism that the wave layer performs when receiving the images. To perform all analyzes we used the same dataset and WCNN parameters configuration as further described in Section 7: “Ablation Tests”.

### 6.1.1 Mother wavelet selection analysis

The decision to use the discrete wavelet transform of the Coiflets 5 family was made, partially, based on the work of [12, 13], in which the authors tested the wavelet transforms of the Daubechies, Symlets, Coiflets, Fejer-Korovkin and dMeyer families. Among them, the Coiflets 5 family showed the best noise reduction results in dense breast radiography images.

To ensure that Coiflets 5 would also be the most suitable family for the object of our research, we analyzed a set of discrete wavelet families, whose result reiterated in the finding of [12, 13]. To perform this analysis, we selected six discrete wavelet families implemented in Python’s Pywavelet library and considered the start and end tags of each one. The selected families were Biorthogonal, Coiflets, Daubechies, Discrete FIR approximation of Meyer, Reverse biorthogonal and Symlets. The analysis consolidated data results that are presented in Table 4.

**Table 4** Accuracy of the initial and final tags of each analyzed wavelet family

Families	Tags	Accuracy
Biorthogonal	Bior 1.3	0.9777
	Bior 6.8	0.9714
Coiflets	Coif 5	0.9821
	Coif 17	0.9801
Daubechies	Db 2	0.9733
	Db 20	0.9777
Discrete FIR approximation of Meyer	Dmey	0.9404
	Rbio 1.3	0.9767
Reverse biorthogonal	Rbio 6.8	0.9781
	Sym 2	0.9743
Symlets	Sym 20	0.9782

As shown in Table 4, tag 5, from the Coiflets family, obtained the best result and because of it, Coiflets 5 was chosen for our model.

### 6.1.2 Decomposition level definition

The decomposition level was set to one to avoid loss of information that might be necessary for image classification. Other levels have been assessed but degraded the image.

### 6.1.3 Decomposition coefficients selection analysis

The decomposition coefficients selection analysis was performed with 1000 images, being 500 images both of COVID-19+ and COVID-19-. As the data are heterogeneous, independent, and non-parametric, we used the BioEstat statistical analysis software, version 5.3 to run the Friedman test using data entry, hypothesis, and significance tests, considering  $\alpha = 0.05$ .

Our intent with this test was *i*) to analyze the significance between the groups [10, 58] and *ii*) to verify if they present statistically similar values among themselves, in relation to the approximate, horizontal, vertical, and diagonal coefficients. The standard deviation of the coefficients was used as an attribute for the significance test between the COVID-19+ and COVID-19- bases and evidenced the existence of significant statistical differences for the approximate, vertical, and diagonal coefficients.

Depending on the test result, we partially analyzed the Wavelet coefficients, with the following configurations, keeping the approximate coefficient in all of them, because it contains more information about the image: *i*) approximate, horizontal, and vertical coefficients; *ii*) approximate, vertical, and diagonal coefficients and *iii*) approximate, horizontal, and diagonal coefficients. The accuracies obtained in this test are shown in Table 5.

Considering the result of the statistical test and the partial analysis of the coefficients, the approximate, vertical, and diagonal coefficients were selected for the creation of the WCNN.

### 6.1.4 Wave layer processing

The Wave layer receives a CT image, with  $512 \times 512$  spatial resolution. It goes through steps in this layer, which are described below:

The first step is responsible for reducing the impact of the background, where each image is cropped up and down by 172 pixels. The reason we did not perform lung segmentation on the selected images is to avoid removing areas of the lesion at the lung boundaries. Cropping results in a  $340 \times 340$ -pixel image, according to Fig. 14.

In the second step, the image is normalized to remove any variations caused by different CT equipment. Its characteristics are extracted from the standard normal distribution, considering the mean  $\mu = 0$  and variance  $\sigma^2 = 1$ . From there, the mean  $\mu$  and the variance  $\sigma^2$  are calculated, as in Eq. 4 and Eq. 5, respectively. The image  $I$  is formed by  $m$  rows and

**Table 5** Accuracy of the initial and final tags of each analyzed wavelet family

Coefficients configuration	Accuracy
Approximation-Horizontal-Vertical	0.9733
Approximation-Vertical-Diagonal	0.9820
Approximation-Horizontal-Diagonal	0.9675

$n$  columns, which is denoted by  $I_{0,0}, I_{0,1}, \dots, I_{m,n}$  [1].  $I_{Normalized}$  is calculated according to Eq. 6.

$$\mu = \frac{\sum_1^m \sum_1^n I_{i,j}}{m * n} \tag{4}$$

$$\sigma^2 = \frac{\sum_1^m \sum_1^n (I_{i,j} - \mu)^2}{m * n} \tag{5}$$

$$I_{Normalized} = \frac{I_{i,j} - \mu}{\sigma} \tag{6}$$

In the third step the image is processed through the wavelet transform decomposition, in a single level of decomposition, using the Coiflets 5 mother wavelet. Of the four generated coefficients (approximate, horizontal, vertical, and diagonal), in this work only three are used to render the digital image. A digital image is composed of the Red, Green and Blue (RGB) space, so the R channel receives the approximate coefficient, the G channel receives the vertical coefficient, and the B channel receives the diagonal coefficient, forming an output of decomposition that will be used by the layers, as shown in Fig. 6. The image cropped by the first step results in a region of interest of the lung that has a spatial resolution of  $340 \times 340$ . Thus, after the wavelet transform decomposition the output of decomposition results in an image with a spatial resolution of  $170 \times 170$ , shown in Fig. 6.

### 6.2 Feature extraction

The convolution operation was established for the convolutional layer, in which a kernel is used to map the activations from one layer into the next. The convolution operation places the

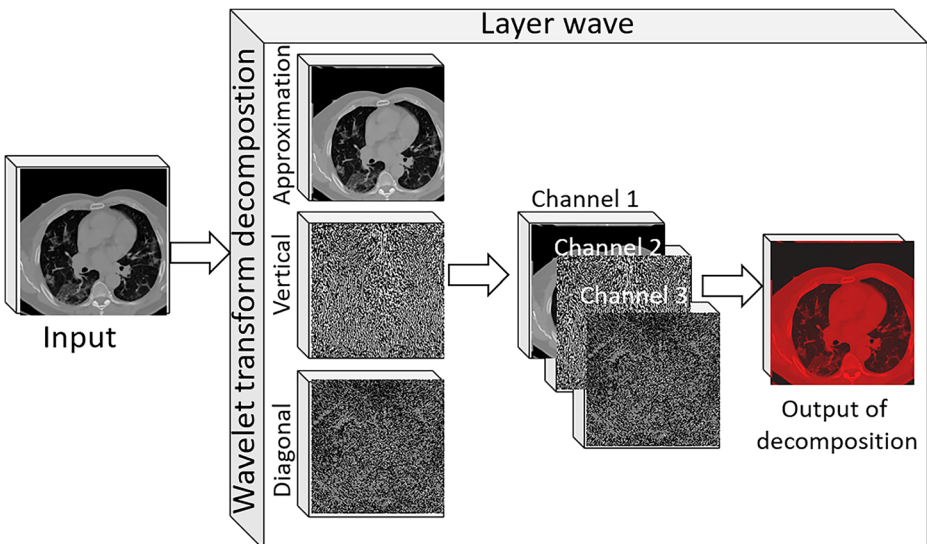


Fig. 6 WCNN Wave Layer

kernel in each position in the image (or hidden layer) so that the kernel overlaps the entire image and executes a dot product between the kernel parameters and its corresponding receptive field – to which the kernel is applied – in the image. The convolution operation is executed in all the regions of the image in order to define the next layer, in which activations keep their spatial relations in the previous layer [1, 21, 34]. There may more than one kernel in the convolutional layer. Every kernel uncovers a feature, such as an edge or a corner. During the forward pass, each kernel is slid to the width and the height of the image (or hidden layer), thus generating the feature map layer [1, 2, 21, 34].

The pooling layer is used to reduce the receptive field's spatial size, thus reducing the number of network parameters. The pooling layer selects a reduced sample of each convolutional layer feature map. Max-pooling was the technique used for this work; it generates the maximum value in the receptive field. The receptive field is  $2 \times 2$ , therefore, max pooling will issue the maximum of the four input values [51].

### 6.3 Flattening layer

After the convolution and pooling processes, the next step is flattening, which converts all feature maps into a one-dimensional matrix, creating an input vector for the fully connected layer [51].

### 6.4 Fully connected layer

In this layer, each neuron from the previous layer is connected to each neuron from the subsequent layer, and all values contribute to predict how strongly a value correlates with a given class [51]. Fully connected layers can be layered on top of each other to capture even more sophisticated combinations of features. The output of the last fully connected layer is fed by an activation function that generates the class scores. WCNN uses the sigmoid activation function, whose output value varies in the range [0, 1]. WCNN entries with an output value above 0.5 are classified as COVID, and those with output below 0.5 relate to other lung diseases [51].

WCNN uses Adaptive Moment Estimation (ADAM), an adaptive optimization technique which saves an exponentially decaying average of previous squared gradients  $v_t$ . In addition to that, ADAM also computes the average of the second moments of the gradients  $m_t$ . [17, 51]. Average and non-centered variance values  $m_t$  are presented in Eq. 7 and Eq. 8, respectively:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)gt \quad (7)$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)gt^2 \quad (8)$$

ADAM updates exponential moving averages of the gradient and the squared gradient where the hyperparameters  $\beta_1, \beta_2 \in [0, 1]$  control the decay rates of these moving averages (Eq. 9) and (Eq. 10):

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \quad (9)$$

$$\widehat{v}_t = \frac{v_t}{1-\beta_2^t} \quad (10)$$

The final equation for update is (Eq. 11):

$$w_{t+1} = w_t - \frac{\alpha \cdot \widehat{m}_t}{\sqrt{\widehat{v}_t + \epsilon}} \quad (11)$$

where  $\alpha$  is the learning rate and  $\epsilon$  is a constant added to the denominator for quick conversion methods in order to avoid the division by 0 [17, 51].

WCNN uses the Dropout technique, the most popular technique to reduce overfitting. Dropout refers to dropping out neurons in a neural network during training. Dropping out a neuron means temporarily disconnecting it, as well as all its internal and external connections, from the network. Dropped-out neurons neither contribute to the forward pass nor do they contribute to the backward pass. By using the dropout technique, the network is forced to learn the most robust features as the network architecture changes with every input [2, 51].

The output of each convolutional layer is fed by an activation function. The activation function layer consists of an activation function which uses the feature map produced by the convolutional layer and generates the activation map as the output. The activation function is used to change a neuron activation level in an output signal. Thus, it performs a mathematical operation and generates the neuron activation level at a specific interval, for instance, 0 to 1 or  $-1$  to 1 [51]. The functions used were the following:

1. Sigmoid / Logistic activation function: The sigmoid function  $\sigma(x) = \frac{1}{1+e^{-x}}$  is a curve shaped like an S [34].
2. The activation function  $f(x) = \max(0, x)$  is called Rectified Linear Unit – ReLU [34] and generates a non-linear activation map.

The WCNN detailed architecture is depicted in Table 6. Furthermore, a rectified linear unit (ReLU) activation function is used after each convolution layer (1st, 3rd, 5th, and 7th) and dense layers (9th, 10th, 11th, and 12th). To reduce the possibility of overfitting, a dropout rate of 20% was implemented to the first four fully connected layers (9th, 10th, 11th, and 12th).

Once the main components of the WCNN architecture have been already presented, the next section describes the series of tests to which the model was submitted.

## 7 Ablation tests

In artificial intelligence (AI), particularly machine learning, ablation is the removal of a component from an AI system. An ablation study investigates the performance of the AI system by changing or removing certain components to understand its contribution to the totality of the system [30]. The term ablation is an analogy with biology, as it consists of altering or removing components from an organism to determine how the individual behaves [30].

**Table 6** WCNN Architecture. The network contains Wave Layer (W), Convolutional Layer (C), Max-Pooling Layer (M) Fully Connected Network Layer (F)

WCNN			
Layer	Operation	Kernel Input/Output size	Dimension
0	W		512×512
1	C	5×5×256	296×296×256
2	M	2×2	148×148×256
3	C	3×3×128	146×146×128
4	M	2×2	73×73×128
5	C	3×3×64	71×71×64
6	M	2×2	35×35×64
7	C	3×3×32	33×33×32
8	M	2×2	16×16×32
9	F	16×16×32 ×256	1×256
10	F	1×1×256×128	1×128
11	F	1×1×128×64	1×64
12	F	1×1×64×32	1×32
13	F	1×1×32×1	1×1

The ablation tests were performed on the WCNN network, which was developed in Python with TensorFlow library [49] running in machine I7-8750H Intel processor, 2.21GHz CPU, 16.0 GB RAM and a GeForce GTX 1060 graphic card with Max-Q Design.

The WCNN network was configured with the following parameters for these tests: *i*) weights were randomly initialized; *ii*) initial learning rate,  $\alpha = 0.001$ , reduced by a factor of 10; *iii*) the number of epochs was 200; *iv*) batch size of 32 and *v*) applied in the dataset I-BIMCV.

## 7.1 Optimization techniques tests

The use of the gradient descent technique is a slow process of convergence, as it depends on parameters chosen at random. In the case of neural networks, this randomness falls on the initial choice of weights. Optimization methods can help an algorithm to converge faster. As aforementioned, the SGD, RMSprops and ADAM techniques were tested [17, 51], the results of which are presented in Table 7.

Once the most suitable optimization method for our research been determined, we conduct the pooling test, detailed in the next subsection.

## 7.2 Pooling test

As the pooling layer makes a reduced sampling from the feature map of the convolutional layer, this test consisted of the use of pooling techniques, associated with WCNN, to identify

**Table 7** Test results using the different optimizing methods

Technique	Accuracy in the validation base
SGD	Memory overflow
RMSprop	97.38%
ADAM	98.2%



which one would return the best accuracy in a specific set of images. The techniques considered were [1, 21, 34]: i) Max pooling, which samples the maximum of each feature map; ii) Min pooling samples the minimum of each feature map; and iii) Avg pooling samples the average of each feature map.

The techniques were assessed considering the configuration of the WCNN standard architecture, presented in section 7, Ablation tests. The results obtained are shown in Table 8.

The Max pooling technique obtained an accuracy of 98% and Avg pooling, 97%. The Min pooling technique was not used as it would result in an activation map at or close to zero. This is because, when repeating Min pooling, the activation values will be zero, so the network cannot be trained once all useful information would have been lost.

## 8 Training configuration parameters

Having the results of the ablation tests in hand, in this section we will describe the training configuration parameters for our neural network. For this, *i*) the weights were randomly initialized; *ii*) the optimizer used was ADAM; *iii*) the standard parameters were set as  $\beta_1 = 0.9$  and  $\beta_2 = 0.999$  [17]; *iv*) the initial learning rate was defined as  $\alpha = 0.001$ ; *v*) reduction factor defined as 10; *vi*) the training consisted of 200 epochs; *vii*) batch size equal to 32; *viii*) pooling technique was max pooling with filter ( $2 \times 2$ ) and, *ix*) 20% for dropout rate. Once the configuration of training parameters has been described, the next section will address the metrics that will be used to evaluate the WCNN performance.

## 9 WCNN evaluation metrics

The follow set of metrics evaluates the WCNN model performance:

1. Accuracy (ACC): accurate classification rate as per the total number of elements.
2. Recall/Sensitivity (Sen): true positive rate.
3. Specificity (Sp): true negative rate.
4. F1-score: weighted average of precision and recall.

They are commonly used to assess the performance of classification algorithms [16, 40, 47]. There is a standard way to show the number of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) to be more visual. This method is called confusion matrix. The confusion matrix allows determining the following metrics algorithms [16, 40, 47] as per Table 9.

**Table 8** Test results with various pooling techniques

Technique	Accuracy in the validation base
<i>Max pooling</i>	0.9801
<i>Min pooling</i>	–
<i>Avg pooling</i>	0.9700

**Table 9** Model Evaluation Metrics

Accuracy (ACC)	Recall/Sensitivity (Sen)	Specificity (Sp)	F1-score (F1)
$ACC = \frac{TN+TP}{TN+TP+FN+FP}$	$Sen = \frac{TP}{TP+FN}$	$Sp = \frac{TN}{TN+FP}$	$F1 = \frac{2 \cdot (Recall \times Precision)}{(Recall + Precision)}$

It is also possible to generate the Receiver Operating Characteristic Curve (ROC curve) if it was necessary. ROC analysis is often called the ROC accuracy ratio, a common technique for judging the accuracy of default probability models [44].

Once the WCNN datasets, architecture and workflow, tests and metrics have been described in the previous sections, section 10 presents the results obtained by using our novel neural network in many contexts.

## 10 Results of WCNN model

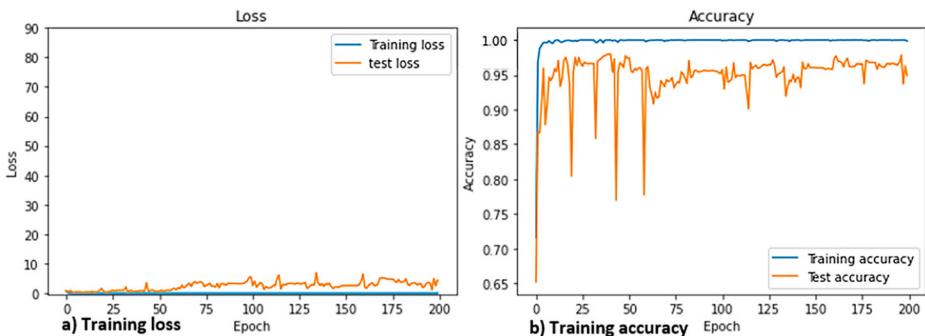
In this section we present the results of use of the WCNN classification model on both internal and external datasets. The obtained results considered two approaches to evaluate the model performance being internal and external validation. Internal validation assessed Dataset I-BIMCV and Dataset II-BIMCV. On the other hand, external validation evaluated Dataset III-BIMCV.

### 10.1 Dataset I result

The training done with Dataset I, consisted of two hundred epochs and generated the results shown in graphs **a) Training Loss** and **b) Training Accuracy**, which make up Fig. 7.

The confusion matrix calculated by validating the internal Dataset I base, whose distribution is presented in Table 1, and shown in Fig. 8.

In Fig. 8 we can see that *i*) true positives (TP) = 1029; *ii*) true negatives (TN) = 994; *iii*) false positives (FP) = 3; and *iv*) false negatives (FN) = 38. Using the parameters TP, TN, FP and FN, the metrics of accuracy, sensitivity and specificity were calculated and are presented in Table 10.



**Fig. 7** WCNN training loss and training accuracy of Dataset I

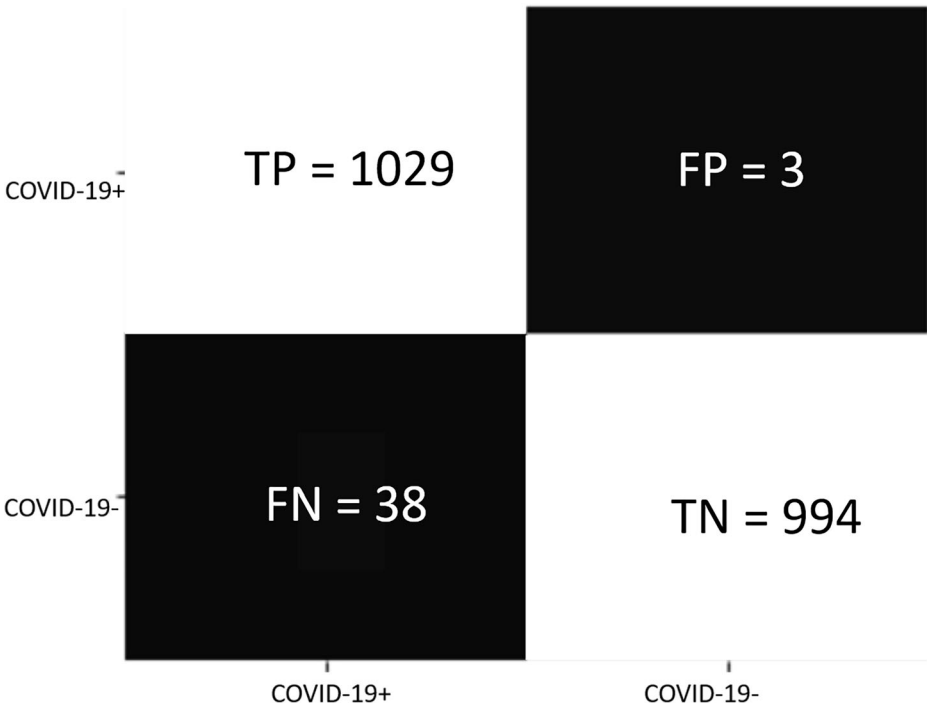


Fig. 8 Dataset I confusion matrix

Using the values from Table 10, the ROC curve was calculated with  $(1-Sp) = 0.0031$  and  $Sen = 0.9643$ , as  $x$  and  $y$ , respectively. Based on the ROC analysis, the area (AUC) was calculated to be 0.98, as shown in Fig. 9.

### 10.2 Dataset II results

The Dataset II training, conducted in two hundred epochs, generated the results shown in graphs a) Training Loss and b) Training Accuracy, which make up Fig. 10.

The confusion matrix calculated by validating the internal Dataset II base, whose distribution is presented in Table 2, and shown in Fig. 11.

In Fig. 11 we can see that i) true positives (TP) = 733; ii) true negatives (TN) = 737; iii) false positives (FP) = 7; and iv) false negatives (FN) = 3. Using the parameters TP, TN, FP and FN, the metrics of accuracy, sensitivity and specificity were calculated and are presented in Table 11.

Table 10 Metrics Results for BIMCV Dataset I

Class	Accuracy	Recall /sensitivity	Specificity	F1-score
COVID-19+	0.9801	0.9643	0.9969	0.98

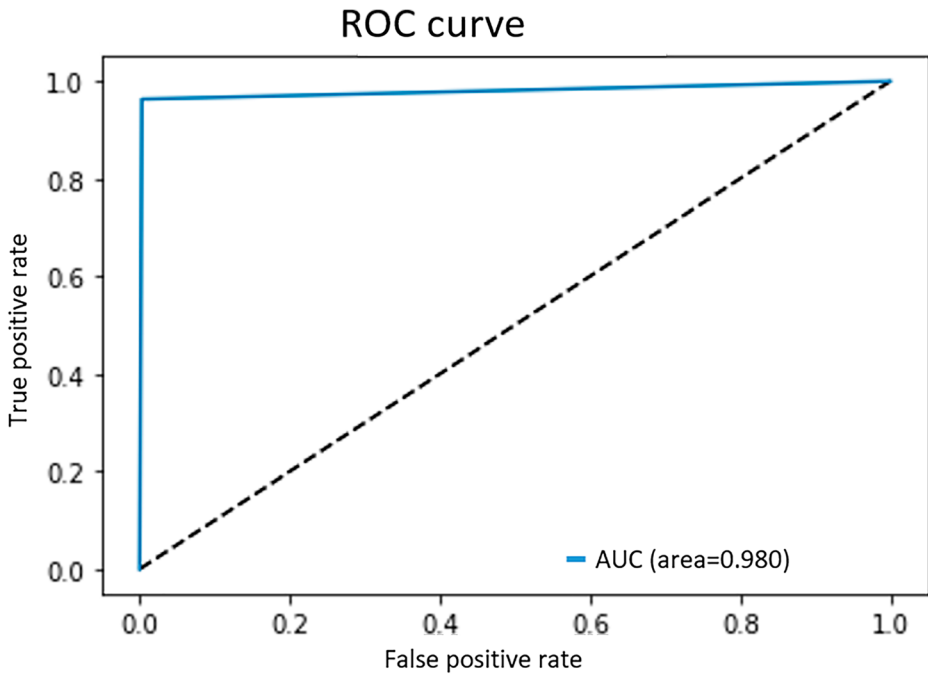


Fig. 9 Dataset I ROC curve

According to the values presented in Table 11, the ROC curve was calculated with  $(1-Sp) = 0.0094$  and  $Sen = 0.9959$ , as  $x$  and  $y$ , respectively. Based on the ROC analysis, the area (AUC) was 0.993, as shown in Fig. 12.

### 10.3 Dataset III results

Dataset III images were submitted to the WCNN model in two scenarios: i) WCNN was executed with Dataset I train weights; and ii) WCNN was run with Dataset II training weights, which generated the results documented in Table 12.

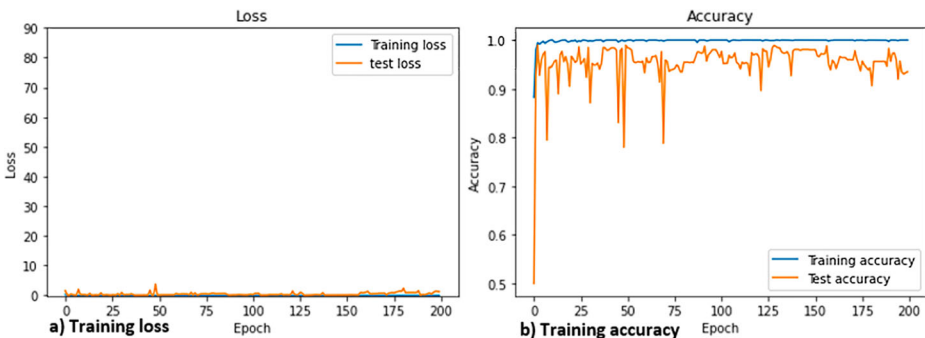


Fig. 10 WCNN training loss and training accuracy of Dataset II

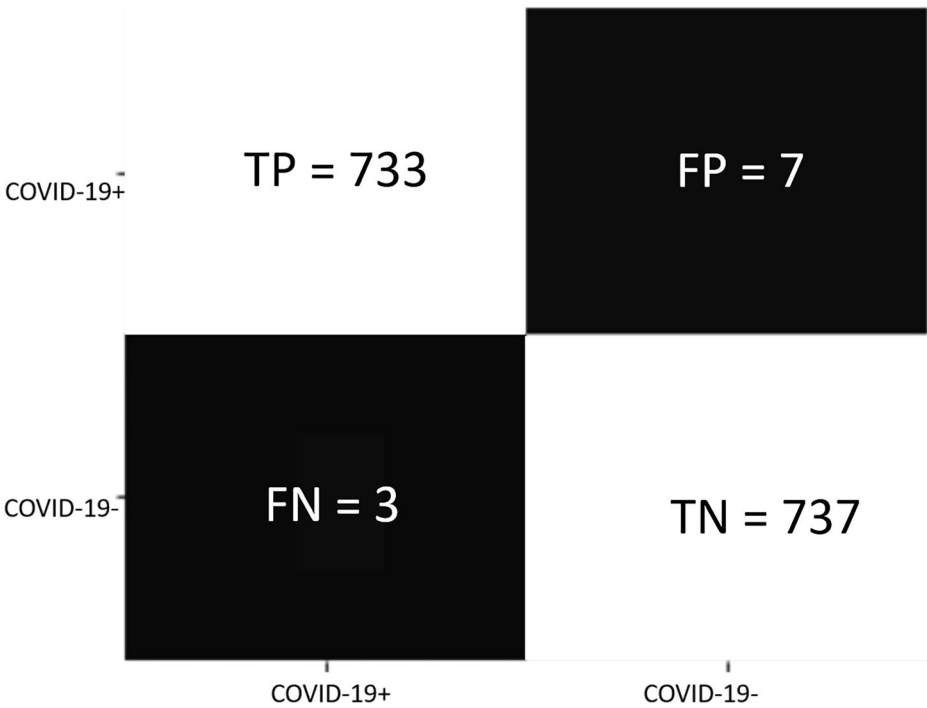


Fig. 11 Dataset II confusion matrix

### 10.4 Consolidated results

The results were consolidated according to environmental categories: internal and external. In Fig. 13, we can see that the validation of the internal datasets presented higher average accuracy than that found in external dataset validation.

Since there is a discrepancy between the results of the internal and external datasets, as shown in the graph in Fig. 13 and in the data in Table 13, we decided to consider the average accuracy found in these datasets to compare them with the research values, in the state of the art, to obtain a realistic scenario.

Table 14 shows the data used to compare our model Compare our model with that of the state of the arts.

As per data presented in Table 14, there exists limitations inherent to previous works, which do not exist in our research, such as:

1. The totality of the previous works makes the distribution of the images of the datasets by the phases of training, testing, and validation by image, and not by patient. Meanwhile,

Table 11 Metrics Results for Dataset II

Class	Accuracy	Recall /sensitivity	Specificity	F1-score
COVID-19+	0.9932	0.9959	0.9932	0.99

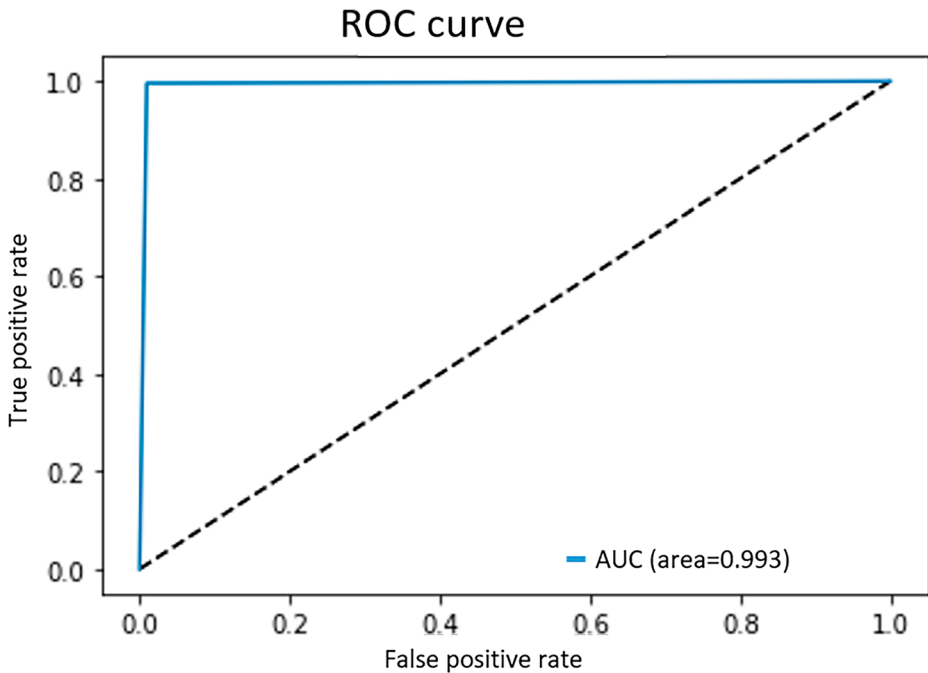


Fig. 12 Dataset II ROC curve

“authors have to ensure that images from the same patient were not included in the different dataset partitions” [39], e.g., training and testing.

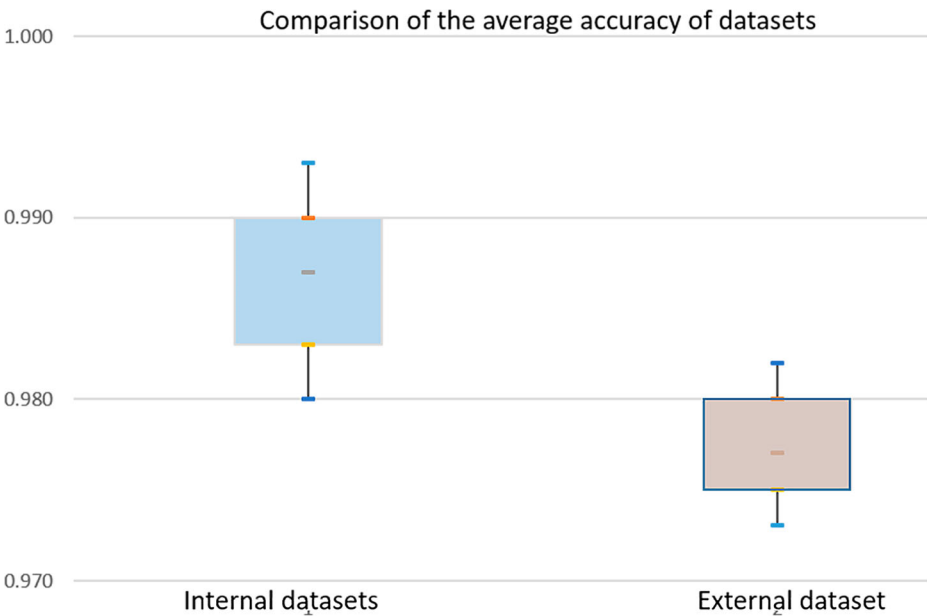
- Most works use data augmentation techniques and more than 50% of them use resizing technique to force the adequacy of the images to the input size defined by the networks. The big problem in using such techniques is that they can cause the loss of image relevant information, which helps in the classification of medical images [39].

## 11 Discussion

The development of this work consisted in the use of the WCNN model in three bases of chest CT images. There were two internal image bases, Dataset I and Dataset II, which are, respectively, public, and private, and the external database Dataset III. As expected in these cases, the databases are heterogeneous, have a variable number of patients and images, were obtained by diverse types of CT equipment and, originally, contain patients of different profiles.

Table 12 Metrics Results for HC-UFU Dataset III

Class	WCNN trained by	Accuracy	Recall/Sensitivity	Specificity	F1-score
COVID-19+	Dataset I	0.9726	0.9621	0.9835	0.97
COVID-19+	Dataset II	0.9821	0.9911	0.9735	0.98



**Fig. 13** Average accuracy for internal (Dataset I e Dataset II) and external (Dataset III)

So that the model’s performance results could be safely compared, we created criteria for exclusion of the images based in [39], with the help of a radiologist. The images that we used have a size of 512 × 512 pixels (original size provided by the institutions), and the images whose patients are under 18 years of age were all discarded, i.e., only adult patients were considered in this work. This exclusion was done based in [12].

The distribution of data for training, testing and validation was done by patient and not by image, which eliminates the risk of using the same image in training and testing, for example. Thus, the training, testing and validation phases had, respectively, the percentages of 70%, 15% and 15%. We emphasize that our work does not use data augmentation and resize resources, as in literature works (see Table 14). Our approach avoids the risk of information loss by artificially increasing the image.

In addition, a wave layer was created to standardize the images. The wave normalizes the images, calculates the decomposition output by means of wavelet transform, replacing the original RGB channels by the approximate, vertical, and diagonal coefficient channels and finally, composing a new digital image that is passed on to the following layers. This process helps to reduce the difference in the extraction of images by different equipment. Associated

**Table 13** Datasets Metrics Values Consolidated Result

Environment	Class	Dataset	Accuracy	Recall/Sen	Specificity	F1-score
Internal	COVID-19+	I	0.9801	0.9643	0.9969	0.98
	COVID-19+	II	0.9932	0.9959	0.9932	0.99
External	COVID-19+	III – Dataset I Weight	0.9726	0.9621	0.9835	0.97
	COVID-19+	III – Dataset II Weight	0.9821	0.9911	0.9735	0.98
Consolidated averages			0.9819	0.9783	0.9867	0.98

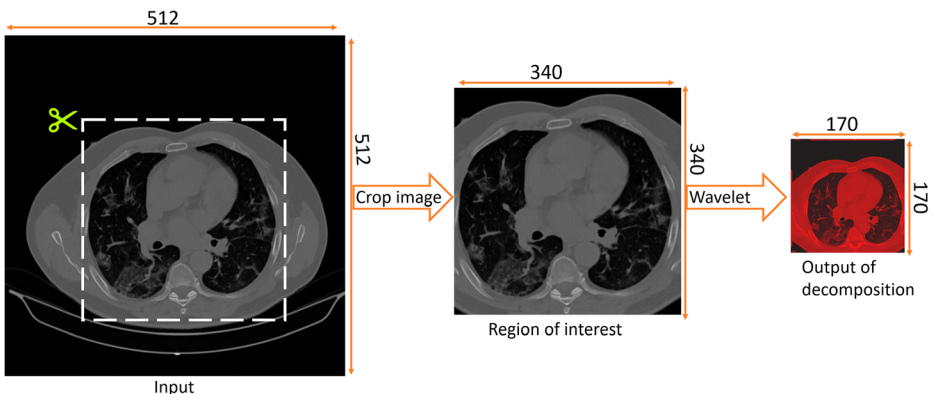
**Table 14** WCNN model versus state of art models

Research	Type of CNN	Public/ Private	Patient/ Image	Data augmentation	Image Resizing	ACC	Sen	Sp
[6]	Trans-CNN	Public	Image	True	True	0.9673	0.9776	0.9602
[41]	COVID-CT-Mask-Net	Public	Image	False	False	0.9166	0.9080	0.9162
[35]	Hybrid Model (squeezeNet e ShuffleNet)	Public	Image	True	True	0.9781	0.9615	0.9608
[32]	FusionModel (VGG-16, GoogleNet e ResNet-50)	Public	Image	False	False	0.9827	0.9893	0.9760
[18]	VGG16, DeseNet121, MobileNet, NASNet, Xception, and EfficientNet	Public	Image	True	True	0.9768	0.9579	0.9971
[3]	UNet++	Private	Image	False	False	0.9885	0.9434	0.9916
[4]	COVID-19-CNN	Public	Image	False	True	0.9670	0.9780	0.9556
Our model	<b>WCNN</b>	Both	Patient	False	False	0.9819	0.9783	0.9867

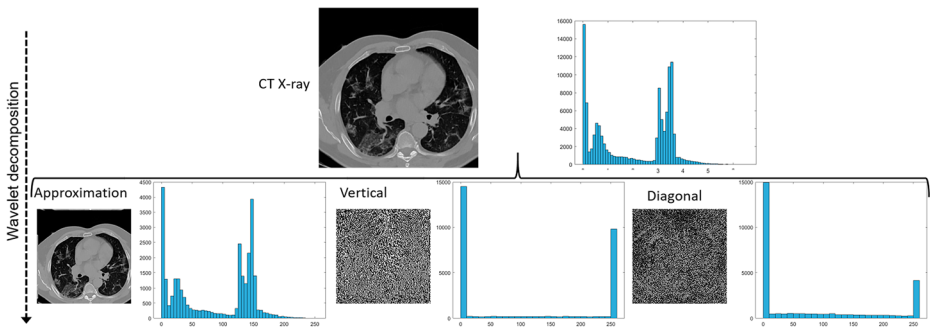
with this benefit, the wave layer processes the image through wavelet transform decomposition, in a single decomposition level, using the Coiflets 5 mother wavelet. In view of this, we considered that the use of WCNN in the wavelet domain can promote training procedures.

This impulse occurs, firstly, because the image generated by the wavelet transform has half the spatial resolution of the original image (Fig. 14), that is, it goes from a spatial resolution of  $340 \times 340$  to a spatial resolution of  $170 \times 170$ , that is, the spatial size of the output feature map is also reduced by half.

Furthermore, the use of wavelet coefficients stimulates activation dispersion in hidden layers and in the output layer. The wavelet coefficients become sparser and therefore it is easier for the network to learn sparse maps rather than dense maps. The histograms in Fig. 15 illustrate the sparse distribution of the vertical, diagonal and approximation coefficients. The important level of sparsity further reduces the training time required for the network to locate the global minimum [8].

**Fig. 14** Region of Interest (ROI) Selection





**Fig. 15** Histogram sample from original image and corresponding diagonal, vertical, and approximation coefficients

After that, we calculated the average of the metric values of the internal and external datasets, thus obtaining an average accuracy of 0.9819, sensitivity of 0.9783 and specificity of 0.9867. When comparing the result of our model with the state of the art's, we found that the WCNN was among the top three.

## 12 Conclusion

In view of the above, we can conclude that the WCNN model has the following advantages compared to the previous related works: 1) Inclusion and exclusion criteria were adopted to form public and private databases with the help of a medical specialist, thus eliminating image duplicates, child patients, patient images with different spatial resolutions, allocation other than 16 bits; 2) Our study did not use data augmentation or image resizing, thus avoiding loss of relevant information [21]; and, 3) the WCNN model is based on a deep neural network using wavelet transform to extract features to classify images of patients with COVID-19, who already present lung changes.

Furthermore, the creation of the new Wave input layer, which replaces the Input layer from the Keras library, selects the region of interest, normalizes the region through its mean and standard deviation, and forms a new image through the decomposition of the wavelet transform, using the Coiflet family 5. Selecting the region of interest eliminates the image background; the normalization eliminates the variations caused by different equipment and the decomposition of the wavelet transform results in an image with a spatial resolution of  $170 \times 170$ , which retains essential information for the classification of the disease, in addition to accelerating the network training process. The WCNN Model is limited by the size of the input image ( $512 \times 512$  pixels), which precludes other sizes of spatial resolution of images, but instigates future assignments.

The results obtained indicate that the investment of time, human, financial and computational resources, in the creation of the WCNN, is a promising approach to assist professionals in the prognosis of the new coronavirus through chest computed tomography images.

**Acknowledgments** This research was financed, in part, by the “Coordination for the Improvement of Higher Education Personnel – Brazil” (CAPES) – Finance Code 001.

**Data availability statements** The datasets analysed in the current study are both public, and private (not public) being:

- 1) The BIMCV (Valencian Region Medical ImageBank) dataset analysed during the current study are PUBLICLY available in the [BIMCV-COVID-19] repository at <https://doi.org/10.48550/arXiv.2006.01174>
- 2) Both HSL-PUCRS (Hospital São Lucas of the Pontifical Catholic University of Rio Grande do Sul, and HC-UFU (Hospital de Clínicas of the Federal University of Uberlândia) datasets analysed during the current study are NOT PUBLICLY available due to concerns about patients data privacy, but are available from the corresponding author on reasonable request and data management privacy policies be discussed and committed by the author and the requester.

## Declarations

**Conflict of interest and ethical standards** All authors declare no conflict of interest, and this article does not contain studies with human or animal participants performed by any of the authors.

## References

1. Aggarwal CC et al (2018) Neural networks and deep learning. Springer. <https://doi.org/10.1007/978-3-319-94463-0>
2. Balas VE, Roy SS, Sharma D, Samui P (2019) Handbook of deep learning applications, vol 136. Springer. <https://doi.org/10.1007/978-3-030-11479-4>
3. Chen J, Wu L, Zhang J, Zhang L, Gong D, Zhao Y, Chen Q, Huang S, Yang M, Yang X, Hu S, Wang Y, Hu X, Zheng B, Zhang K, Wu H, Dong Z, Xu Y, Zhu Y, ... Yu H (2020) Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. *Sci Rep* 10:1–11. <https://doi.org/10.1038/s41598-020-76282-0>
4. Chen YM, Chen YJ, Ho WH, Tsai JT (2021) Classifying chest CT images as COVID-19 positive/negative using a convolutional neural network ensemble model and uniform experimental design method. *BMC Bioinformatics* 22:1–19. <https://doi.org/10.1186/s12859-021-04083-x>
5. Dai W, Zhang H, Yu J, Xu H, Chen H, Luo S, Zhang H, Liang L, Wu X, Lei Y, Lin F (2020) CT imaging and differential diagnosis of COVID-19. *Can Assoc Radiol J* 71:195–200. <https://doi.org/10.1177/0846537120913033>
6. Fan X, Feng X, Dong Y, Hou H (2022) COVID-19 CT image recognition algorithm based on transformer and CNN. *Displays* 72:102150. <https://doi.org/10.1016/j.displa.2022.102150>
7. Gonzalez RC, Woods RE (2006) Digital image processing. Prentice Hall, New Jersey
8. Guo T, Mousavi HS, Vu TH, Monga V (2017) Deep wavelet prediction for image super-resolution. Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 104–113
9. Huang Y, Wang Q, Jia W, Lu Y, He X (2021) See more than once: kernel-sharing atrous convolution for semantic segmentation. *Neurocomputing* 443:26–34. <https://doi.org/10.1016/j.neucom.2021.02.091>
10. Ishitaki T, Oda T, Barolli L (2016) A neural network-based user identification for Tor networks: data analysis using Friedman test. In 2016 30th international conference on advanced information networking and applications workshops, pp. 7–13. <https://doi.org/10.1109/waina.2016.143>
11. Jansen M (2012) Noise reduction by wavelet thresholding, vol 161. Springer Science & Business Media. <https://doi.org/10.1007/978-1-4613-0145-5>
12. Junior CADC (2019) Proposta de uma Metodologia para Suavização de Ruído em Imagens Mamográficas de Mamas Densas." Dissertação (Mestrado em Engenharia Biomédica) – Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia, Dissertação (Mestrado em Engenharia Biomédica) – Faculdade de Engenharia Elétrica. <https://doi.org/10.14393/ufu.di.2019.2036>
13. Junior CADC, Patrocínio AC (2019) Performance evaluation of Denoising techniques applied to mammograms of dense breasts. XXVI Brazilian congress on biomedical engineering, pp. 369–374. [https://doi.org/10.1007/978-981-13-2517-5\\_56](https://doi.org/10.1007/978-981-13-2517-5_56)
14. Kasongo SM, Sun Y (2020) A deep learning method with wrapper based feature extraction for wireless intrusion detection system. *Comput Secur* 92:101752. <https://doi.org/10.1016/j.cose.2020.101752>


15. Kassania SH, Kassanib PH, Wesolowski MJ, Schneidera KA, Detersa R (2021) Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: a machine learning based approach. *Biocybern Biomed Eng* 41:867–879. <https://doi.org/10.1016/j.bbe.2021.05.013>
16. Khatami A, Khosravi A, Nguyen T, Lim CP, Nahavandi S (2017) Medical image analysis using wavelet transform and deep belief networks. *Expert Syst Appl* 86:190–198. <https://doi.org/10.1016/j.eswa.2017.05.073>
17. Kingma DP, Ba JL (2014) Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980
18. Kogilavani SV, Prabhu J, Sandhiya R, Kumar MS, Subramaniam U, Karthick A, Muhibullah M, Imam SBS (2022) COVID-19 detection based on lung Ct scan using deep learning techniques. *Comput Math Methods Med* 2022:1–13. <https://doi.org/10.1155/2022/7672196>
19. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, pp 1097–1105.
20. Lecun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86:2278–2324. <https://doi.org/10.1109/5.726791>
21. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444. <https://doi.org/10.1038/nature14539>
22. Li X, Zhang W, Ding Q (2019) Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction. *Reliab Eng Syst Saf* 182:208–218. <https://doi.org/10.1016/j.res.2018.11.011>
23. Long W, Lu Z, Cui L (2019) Deep learning-based feature engineering for stock price movement prediction. *Knowl-Based Syst* 164:163–173. <https://doi.org/10.1016/j.knosys.2018.10.034>
24. Lu X, Wang W, Ma C, Shen J, Shao L, Porikli F (2019) See more, know more: unsupervised video object segmentation with co-attention siamese networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 3623–3632
25. Lu X, Ma C, Shen J, Yang X, Reid I, Yang MH (2020) Deep object tracking with shrinkage loss. *IEEE Trans Pattern Anal Mach Intell* 44:2386–2401. <https://doi.org/10.1109/TPAMI.2020.3041332>
26. Lu X, Wang W, Shen J, Crandall DJ, Gool LV (2021) Segmenting objects from relational visual data. *IEEE Trans Pattern Anal Mach Intell* 44:7885–7897. <https://doi.org/10.1109/TPAMI.2021.3115815>
27. Lu X, Wang W, Shen J, Crandall D, Luo J (2022) Zero-shot video object segmentation with co-attention Siamese networks. *IEEE Trans Pattern Anal Mach Intell* 44:2228–2242. <https://doi.org/10.1109/TPAMI.2020.3040258>
28. Martin DR, Hanson JA, Gullapalli RR, Schultz FA, Sethi A, Clark DP (2020) A deep learning convolutional neural network can recognize common patterns of injury in gastric pathology. *Arch Pathol Lab Med* 144:370–378. <https://doi.org/10.5858/arpa.2019-0004-OA>
29. Merry RJE (2005) Wavelet theory and applications: a literature study. DCT rapporten, vol. 2005
30. Meyers R, Lu M, Puisseau CWD, Meisen T (2019) Ablation studies in artificial neural networks. *arXiv preprint arXiv:1901.08644*, 2019. <https://doi.org/10.48550/arXiv.1901.08644>
31. Mittal M, Goyal LM, Kaur S, Kaur I, Verma A, Hemanth J (2019) Deep learning based enhanced tumor segmentation approach for MR brain images. *Appl Soft Comput* 78:346–354. <https://doi.org/10.1016/j.asoc.2019.02.036>
32. Özkaya U, Öztürk S, Barstugan M (2020) Coronavirus (COVID-19) classification using deep features fusion and ranking technique. *Big data analytics and artificial intelligence against COVID-19: Innovation Vision and Approach.*: Springer, 020, pp. 281–295. [https://doi.org/10.1007/978-3-030-55258-9\\_17](https://doi.org/10.1007/978-3-030-55258-9_17)
33. Ozturk S, Ozkaya U, Barstugan M (2020) Classification of coronavirus images using shrunken features. *medRxiv*. <https://doi.org/10.1101/2020.04.03.20048868>
34. Ponti MA, Costa GBP (2018) Como funciona o deep learning. arXiv preprint arXiv:1806.07908
35. Rangarajan AK, Ramachandran HK (2022) A fused lightweight CNN model for the diagnosis of COVID-19 using CT scan images. *Automatika* 63:171–184. <https://doi.org/10.1080/00051144.2021.2014037>
36. Ravi D, Wong C, Deligianni F, Berthelot M, Perez JA, Lo B, Yang GZ (2017) Deep learning for health informatics. *IEEE J Biomed Health Inform* 21:4–21. <https://doi.org/10.1109/JBHI.2016.2636665>
37. Ribeiro CDS, Roode MV, Haringhuizen GB, Koopmans MP, Claassen E, Burgwal LHMV (2018) How ownership rights over microorganisms affect infectious disease control and innovation: a root-cause analysis of barriers to data sharing as experienced by key stakeholders. *PLoS One* 13:e0195885. <https://doi.org/10.1371/journal.pone.0195885>
38. Ribeiro CDS, Koopmans MP, Haringhuizen GB (2018) Threats to timely sharing of pathogen sequence data. *Science* 362:404–406. <https://doi.org/10.1126/science.aau5229>
39. Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, Riveiro AIA, Etmann C, McCague C, Beer L, McCall JRW, Teng Z, Klotsas EG, Rudd JHF, Sala E (2021) Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell* 3:199–217. <https://doi.org/10.1038/s42256-021-00307-0>

40. Ruuska S, Wilhemiina H, Sari K, Mikaela M, Pekka M, Jaakko M (2018) Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. *Behav Process* 148:56–62. <https://doi.org/10.1016/j.beproc.2018.01.004>
41. Sarkisov AT (2022) COVID-CT-mask-net: prediction of COVID-19 from CT scans using regional features. *Appl Intell* 52:9664–9675. <https://doi.org/10.1007/s10489-021-02731-6>
42. Shafiq M, Gu Z (2022) Deep residual learning for image recognition: a survey. *Appl Sci* 12:8972. <https://doi.org/10.3390/app12188972>
43. Sherry ST, Ward MH, Kholodv M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 29:308–311. <https://doi.org/10.1093/nar/29.1.308>
44. Shirazi AZ, Chabok SJS, Mohammadi Z (2018) A novel and reliable computational intelligence system for breast cancer detection. *Med Biol Eng Comput* 56:721–732. <https://doi.org/10.1007/s11517-017-1721-z>
45. Simon JHM, Claassen E, Correa CE, Osterhaus ADME (2005) Managing severe acute respiratory syndrome (SARS) intellectual property rights: the possible role of patent pooling. *Bull World Health Organ* 83:707–710
46. Simonyan K, Zisserman A (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. <https://doi.org/10.48550/arXiv.1409.1556>
47. Skansi S (2018) Introduction to deep learning: from logical calculus to artificial intelligence. Springer. <https://doi.org/10.1007/978-3-319-73004-2>
48. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition* pp. 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
49. Vayá MDLI et al (2020) BIMCV COVID-19+: a large annotated dataset of RX and CT images from COVID-19 patients. arXiv preprint arXiv:2006.01174, 2020. <https://doi.org/10.48550/arXiv.2006.01174>
50. Wang L, Lin ZQ, Wong A (2020) Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Sci Rep* 10:1–12. <https://doi.org/10.1038/s41598-020-76550-z>
51. Wani MA, Bhat FA, Afzal S, Khan AI (2020) *Advances in deep learning*, vol 57. Springer. <https://doi.org/10.1007/978-981-13-6794-6>
52. Weiss SR, Leibowitz JL (2011) Coronavirus pathogenesis. *Adv Virus Res Elsevier* 81:85–164. <https://doi.org/10.1016/B978-0-12-385885-6.00009-2>
53. Wu J (2020) Institute of Genomics, Chinese Academy of Science, China National Center for Bioinformatics & National Genomics Data Center. <https://bigd.big.ac.cn/ncov/?lang=en>. Accessed 01 June 2020
54. Yang W, Cao Q, Wang X, Cheng Z, Pan A, Dai J, Sun Q, Zhao F, Qu J, Yan F (2020) Clinical characteristics and imaging manifestations of the 2019 novel coronavirus disease (COVID-19): a multi-center study in Wenzhou city, Zhejiang, China. *J Infect* 80:388–393. <https://doi.org/10.1016/j.jinf.2020.02.016>
55. Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. *European conference on computer visio*, pp 818–833. [https://doi.org/10.1007/978-3-319-10590-1\\_53](https://doi.org/10.1007/978-3-319-10590-1_53)
56. Zhang J, Xie Y, Li Y, Shen C, Xia Y (2020) Covid-19 screening on chest x-ray images using deep learning-based anomaly detection. arXiv preprint arXiv:2003.12338
57. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W (2020) China novel coronavirus investigating and research team. A novel coronavirus from patients with pneumonia in China. *N Engl J Med* 382:727–733. <https://doi.org/10.1056/NEJMoa2001017>
58. Zimmerman DW, Zumbo BD (1993) Relative power of the Wilcoxon test, the Friedman test, and repeated-measures ANOVA on ranks. *J Exp Educ* 62:75–86. <https://doi.org/10.1080/00220973.1993.9943832>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Affiliations

**Pedro Moises de Sousa**<sup>1</sup>  • **Pedro Cunha Carneiro**<sup>1</sup> • **Gabrielle Macedo Pereira**<sup>1</sup> • **Mariane Modesto Oliveira**<sup>1</sup> • **Carlos Alberto da Costa Junior**<sup>1</sup> • **Luis Vinicius de Moura**<sup>2</sup> • **Christian Mattjie**<sup>2</sup> • **Ana Maria Marques da Silva**<sup>2</sup> • **Túlio Augusto Alves Macedo**<sup>3</sup> • **Ana Claudia Patrocinio**<sup>1</sup>

Pedro Cunha Carneiro  
pedrocarneiro@ufu.br

Gabrielle Macedo Pereira  
gabriellemp@ufu.br

Mariane Modesto Oliveira  
modestomariane@gmail.com

Carlos Alberto da Costa Junior  
carlosjunior@ufu.br

Luis Vinicius de Moura  
luis.moura@edu.pucrs.br

Christian Mattjie  
christian.mattjie@gmail.com

Ana Maria Marques da Silva  
ana.marques@pucrs.br

Túlio Augusto Alves Macedo  
tuliomacedo@ufu.br

Ana Claudia Patrocinio  
ana.patrocinio@ufu.br

- <sup>1</sup> Biomedical Lab, Faculty of Electrical Engineering, Federal University of Uberlândia, Campus Sta Mônica, Av. João Naves de Avila, 2121, Bloco 1E, Uberlândia, MG CEP 38400-000, Brazil
- <sup>2</sup> Medical Image Computing Laboratory, Pontifical Catholic University of Rio Grande do Sul, Av. Ipiranga, 6681 Partenon, Porto Alegre, RS CEP 90619-900, Brazil
- <sup>3</sup> Clinic Hospital of the Federal University, Campus Umuarama - Bloco UMU2H - Sala 01 Av. Pará - 1720 - Bairro Umuarama Uberlândia - MG - CEP, Uberlândia, MG 38405-320, Brazil