COVID-19 Classification in X-ray Chest images using a New Convolutional Neural Network: CNN-COVID

Pedro Moises de Sousa (profpedromoises@gmail.com)
Federal University of Uberlândia https://orcid.org/0000-0003-4563-0033

Marianne Modesto
Federal University of Uberlândia https://orcid.org/0000-0002-1645-1112

Gabrielle Macedo Pereira
Federal University of Uberlândia https://orcid.org/0000-0002-1482-4946

Carlos Alberto da Costa Junior
Federal University of Uberlândia https://orcid.org/0000-0003-1915-9137

Luis Vinicius de Moura
Pontifical Catholic University of Rio Grande do Sul https://orcid.org/0000-0003-3429-3289

Christian Mattjie
Pontifical Catholic University of Rio Grande do Sul https://orcid.org/0000-0002-3745-8686

Pedro Cunha Carneiro
Federal University of Uberlândia https://orcid.org/0000-0002-0120-5273

Ana Maria Marques da Silva
Pontifical Catholic University of Rio Grande do Sul https://orcid.org/0000-0002-5924-6852

Ana Claudia Patrocinio
Federal University of Uberlândia https://orcid.org/0000-0001-9376-7689

Research Article

Keywords: chest x-ray images, coronavirus, deep learning, convolutional neural network, CNN, COVID-19, CNN-COVID

Posted Date: July 14th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-40913/v1

COVID-19 Classification in X-ray Chest images using a New Convolutional Neural Network: CNN-COVID

Pedro Moisés de Sousa[1], Marianne Modesto[2], Gabrielle Macedo Pereira[3], Carlos Alberto da Costa Junior[4], Luis Vinicius de Moura[5], Christian Mattjie[6], Pedro Cunha Carneiro[7], Ana Maria Marques da Silva[8], Ana Claudia Patrocinio[9].

[1] Biomedical Lab, Faculty of Electrical Engineering, Federal University of Uberlândia, Campus Sta Mônica, Av. João Naves de Ávila, 2121, Bloco 1E, CEP 38400-000, Uberlândia, MG, Brazil. E-mail: profpedromoises@gmail.com. ORCID: 0000-0003-4563-033.

[2] Biomedical Lab, Faculty of Electrical Engineering, Federal University of Uberlândia, Campus Sta Mônica, Av. João Naves de Ávila, 2121, Bloco 1E, CEP 38400-000, Uberlândia, MG, Brazil. E-mail: modestomariane@gmail.com. ORCID: 0000-0002-1645-1112.

[3] Biomedical Lab, Faculty of Electrical Engineering, Federal University of Uberlândia, Campus Sta Mônica, Av. João Naves de Ávila, 2121, Bloco 1E, CEP 38400-000, Uberlândia, MG, Brazil. E-mail: gabriellemp@ufu.br. ORCID: 0000-0002-1482-4946.

[4] Biomedical Lab, Faculty of Electrical Engineering, Federal University of Uberlândia, Campus Sta Mônica, Av. João Naves de Ávila, 2121, Bloco 1E, CEP 38400-000, Uberlândia, MG, Brazil. E-mail: carlosjunior@ufu.br. ORCID: 0000-0003-1915-9137.

[5] Medical Image Computing Laboratory, Pontifical Catholic University of Rio Grande do Sul, Av. Ipiranga, 6681 Partenon, CEP 90619-900, Porto Alegre, RS, Brazil. E-mail: luis.moura@edu.pucrs.br. ORCID: 0000-0003-3429-3289.

[6] Medical Image Computing Laboratory, Pontifical Catholic University of Rio Grande do Sul, Av. Ipiranga, 6681 Partenon, CEP 90619-900, Porto Alegre, RS, Brazil. E-mail: christian.mattjie@gmail.com. ORCID: 0000-0002-3745-8686.

[7] Biomedical Lab, Faculty of Electrical Engineering, Federal University of Uberlândia, Campus Sta Mônica, Av. João Naves de Ávila, 2121, Bloco 1E, CEP 38400-000, Uberlândia, MG, Brazil. E-mail: pedrocarneiro@ufu.br. ORCID: 0000-0002-0120-5273.

[8] Medical Image Computing Laboratory, Pontifical Catholic University of Rio Grande do Sul, Av. Ipiranga, 6681 Partenon, CEP 90619-900, Porto Alegre, RS, Brazil. E-mail: ana.marques@pucrs.br. ORCID: 0000-0002-5924-6852.

[9] Biomedical Lab, Faculty of Electrical Engineering, Federal University of Uberlândia, Campus Sta Mônica, Av. João Naves de Ávila, 2121, Bloco 1E, CEP 38400-000, Uberlândia, MG, Brazil. E-mail: ana.patrocinio@ufu.br. ORCID: 0000-0001-9376-7689.

Abstract

Purpose

As COVID-19 causes lung inflamation and lesions, several scientists have worked on seeking a computer model able to identify medical images of patients with this disease and improve triage. Chest x-ray and computed tomography images are remarkably like images from patients with other lung diseases, which makes it hard to diagnose, and that is why there was an urge to seek a computer model. Thus, this paper proposes a computer model able to classify x-ray images of patients as with the new coronavirus. Chest x-ray exams were chosen for this study over computed tomography scans because they are low cost, results are obtained quickly, and x-ray equipment availability is higher in regions impacted with the disease.

Methods

A new CNN network, CNN-COVID, was developed to classify chest images as with and without COVID. This article collected images of patients with and without COVID-19 from Covid-19 image data collection and ChestXray14 banks. To assess the network's accuracy, 10 training sessions and tests were done using CNN-COVID. A confusion matrix was generated to assess the performance of the model and calculate the following metrics: sensitivity (Se), specificity (Sp), and F1 score. Besides, the Receiver Operating Characteristic (ROC) curves and the Areas Under the Curve (AUCs) were used for assessment.

Results

The following values were obtained: AUC = 0.9720, Se = 98%, and Sp = 96% for the validation set. A total of 10 tests were executed and the average was 0.9787, the lowest result being 0.9740 and the highest, 0.9870.

Conclusions

The results proved that the CNN-COVID model is a promising tool to help physicians classify

chest images with pneumonia, considering pneumonia caused by COVID-19 and pneumonia due to other causes.

Keywords: chest x-ray images, coronavirus, deep learning, convolutional neural network, CNN, COVID-19, CNN-COVID

Introduction

In December 2019, a group of patients with atypical pneumonia, cause unknown, was associated to the consumption of bat meat bought at an exotic animal meat market in Wuhan, Hubei, China. The disease has quickly spread to other corners of the world and, on March 11, 2020, the World Health Organization declared the COVID-19 global pandemic outbreak, which is still ongoing (Zhu et al. 2019; Yang et al. 2020). By using impartial sequencing of samples from patients, it was possible to identify a new type of beta coronavirus. This novel coronavirus, named 2019-nCoV, was compared to Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS), presented a lower mortality rate and a higher transmission capacity (Zhu et al. 2019;Yang et al. 2020).

Coronaviruses are RNA viruses with a lipid envelope able to cause respiratory, enteric, liver and neurological diseases in domestic animals and human beings (Zhu et al. 2019; Weiss and Leibowitz 2011). The existence of wild viruses in nature is well-known by researchers. They are highly prevalent, broadly distributed, genetically diverse, and frequently genetically recombined. There is a higher probability of natural hosts disseminating these viruses to human beings as the man-animal interface activity is increasing.

Both SARS and MERS were associated with a zoonotic origin and transmitted by civets and camels, respectively. The appearance of diseases such as SARS and MERS, supposedly caused by the new coronavirus, tends to become more frequent events when no barriers are imposed between human society and wild nature (Zhu et al. 2019; Cui et al. 2019).

As there is no specific drug or vaccine to treat the new coronavirus, early detection is crucial so that the patient is isolated from the healthy population as soon as possible (Ai, 2020). Thus, the research for methods of early detection has become essential to fight this pandemic outbreak.

Currently, the gold standard of the COVID-19 diagnosis is made by viral nucleic acid detection using reverse transcription-polymerase chain reaction (RT-PCR) in real-time, although its effective

accuracy is 30 to 50%(Ai 2020; Zhang et al. 2020; Ozturk et al. 2020).

An underlying problem with this method is how unavailable it is in several regions and countries impacted; this generates logistic and political issues for providing enough test kits for the increasing number of patients suspected of having the disease (Zhang et al. 2020). Moreover, the delay in processing and getting results and the significant amout of false negatives urged researchers from all over the world to try and find a solution for this problem in various areas of knowledge (Ozturk et al. 2020).

Medical image processing is one of the areas which has been contributing to promising studies. Research in this area is being done to clinically aid to diagnose the disease for patients who develop lung the atypical pneumonia, using chest x-ray images or computed tomography (CT) scans (Ozturk et al. 2020). As several patients with COVID-19 develop lung infection, CT scans has had very useful to detect lung impairment, as well as classify its progression (Zhang et al. 2020; Dai et al. 2020).

Radiological images from COVID-19 patients may present similarities with those from patients with bacterial or viral pneumonia, specifically the ones caused by SARS and MERS. Thus, the ability to accurately differentiate diseases by analyzing medical images has become a vital challenge and overcoming it means helping healthcare professionals with detecting the disease early and isolating affected patients as soon as possible (Ozturk et al. 2020; Chung et al. 2020).

Medical image processing area has several studies aimed at developing machine learning methods able to help to diagnose COVID-19, either using CT scan images or chest x-ray ones (Ozturk et al. 2020). Problems using CT scans rather than chest x-rays includes low availability of equipment, radiologists, and physicians; higher cost; and longer time to obtain images (Zhang et al. 2020).

A machine learning technique used in research is called deep learning. It allows computer models with several layers of processing to learn how to represent data in various levels of abstraction (Zhang et al. 2020; Lecun et al. 2015; Martin et al. 2020). This technique allows design of applications which can to perform recognition, such as speech recognition, visual recognition and object detection.

Medical images of patients with COVID-19 presents common features that might show a pattern. Deep learning is an effective technique used by researchers to help healthcare professionals to analyze vast volumes of data generated by chest x-ray images, for instance (Zhang et al. 2020; Martin et al. 2020).

When it comes to applying artificial intelligence and machine learning to the medical field, convolutional neural networks (CNNs) stand out. CNNs are deep artificial neural networks which can be used to classify images, group them by similarity and run object recognition. These networks are inspired in the human visual cortex processing and used for medical images where irregularities in tissue morphology may be used to classify tumors. CNNs can detect patterns which are hard to be found by human specialists, for instance, initial stages of disease in tissue samples (Balas et al., 2020).

This paper proposes a new Convolutional Neural Network model called CNN-COVID, to classify images of COVID-19 patients and differentiate them from those who do not have COVID-19, with a focus in analysis, classification, and high accuracy. The CNN-COVID model and its related work are presented in the following sections.

Related work

In the study from Zhang et al. (2020) a model to detect anomalies was developed by using deep learning. The goal is to do a quick and trustworthy triage of patients with COVID-19. One hundred chest x-ray images from 70 patients with COVID-19 were sourced from a GitHub repository to assess the performance of the model; and 1431 chest x-ray images from 1008 patients with other types of pneumonia were sourced from ChestX-ray14, a public data pool. The Zhang model in (Zhang et al. 2020) is composed of three components, namely, a backbone network, a classification head, and an anomaly detection head. Thus, the first component extracts the high-level features from a chest x-ray which will be the input data for the classification head, and then, the input data for the anomaly

detection head. So, with this paper, it resulted 90.00% sensitivity, 87.84% specificity (when parameter T from the study was 0.25) or 96.00% sensitivity and 70.65% specificity (when parameter T was 0.15). Nonetheless, this model presented some limitations, such as missing 4% of COVID-19 cases and having an approximately 30% false positives.

In a similar study, Sethy and Behe (2020) suggested a model to detect COVID-19 from chest x-rays using deep learning. Support vector machine (SVM) classified images of patients suffering from this disease differentiating from images of patients who suffer from other diseases. A subset of 25 COVID-19 images were sourced from the GitHub repository, and a subset of 25 pneumonia images was sourced from the Kaggle repository. The ResNet50 deep neural network model with SVM classification has proven to be the best approach to detect COVID-19, with 95.38% accuracy, 97.2% sensitivity, and 93.4% specificity.

In Wang and Wong (2020) a CNN was adapted to detect COVID-19 cases. It was called COVID-Net, and it used chest x-rays images from open source code and available to the general audience. Thus, COVIDX was created, a database sourcing samples from five different databases (Wang and Wong 2020): 100 samples of chest x-rays of healthy patients, 100 samples of pneumonia patients, and 100 images of COVID-19 patients. COVID-NET presented the following results: 93.3% accuracy, 91.0% sensitivity, and 99.9% specificity when detecting COVID-19 x-ray images out of healthy cases and severe acute respiratory syndrome cases.

The study from Abbas et al. (2020) uses the transfer learning technique, an effective mechanism to provide a promising solution by transferring knowledge from generic tasks for object recognition to domain-specific tasks. Therefore, DeTraC, a deep CNN was adapted with functions to decompose, transfer, and compose samples to classify chest x-ray images as COVID-19. They sourced 80 chest x-ray images (4020 x 4892 pixels) of healthy patients from the Japanese Society of Radiological Technology (JSRT), 105 images (4248 × 3480 pixels) of COVID-19 patients and 11 images (4248 × 3480 pixels) of SARS patients from GitHub. DeTraC presented the following results:

95.12% accuracy, 97.91% sensitivity, and 91.87% specificity when detecting COVID-19 x-ray images out of healthy cases and severe acute respiratory syndrome cases.

Dataset and Data Augmentation

Dataset

This research used two distinct databases. The first database has 217 chest x-ray images from the COVID-19 image data collection (Cohen and Dao 2020). In this database, the images are from 141 patients who tested positive for COVID-19. From the second database, ChestXray14 (Wang et al. 2019), 1126 images were used. All these images correspond to chest x-ray images labeled for the presence of 14 common chest radiographic observations which, in this paper, were labeled as NON-COVID.

A subset of 166 images, out of the 217 images from the COVID-19 set were selected randomly (Cohen and Dao 2020). Out of these 166 selected images, around 75% were for the training phase and 25% were for the testing phase. From the NON-COVID set, 1000 images were selected randomly, 80% were for the training phase and 20% were for the testing phase.

For the validation phase, 126 images were selected from the COVID-19 set (out of the 217 total images) and 126 images from the NON-COVID set (out of the 1126 total images), as shown in Table 1.

Table 1: Database used for CNN-COVID. Created by the author

Database	Database total	Training	Test	Validation
COVID-19	217	126	40	126
NON-COVID	1126	800	200	126

Data Augmentation

One of the challenges of this study was having a limited number of x-rays from COVID-19 patients for the training of deep learning models. We used an *ImageDataGenerator* class to overcome

this issue, providing for new images to be generated for the training. The new images were generated from digital processing, using geometric transformations of the original images.

These geometric transformations, such as be translation, rotation, patch extraction and reflection, do not make changes to the image object properties, making "data augmentation" possible. The positive side of this technique is that it increases the ability to generalize CNN-COVID when trained with an augmented dataset (Aggarwal 2018; Chollet 2016). Thus, overfitting, which is when the network is no longer able to generalize when presented with new data, can be reduced.

The following common methods were used for dataset augmentation:

- range rotation
- width shift range
- height shift range
- zoom range
- horizontal flip
- vertical flip

After these changes, it was possible to balance the dataset from both COVID and NON-COVID classifications on the testing and training sets. This database augmentation happens during run time, when chest x-ray images are presented as an input for CNN-COVID.

CNN-COVID Creation

A CNN is composed of two stages: feature extraction stage and classification stage. In the CNN, the pooling and convolution layers act as a stage of feature extraction. In contrast the classification stage is made of one or more fully connected layers followed by a sigmoid function

layer (Wani et al. 2020), which are presented below. Python programming language (using Keras library) (Chollet, 2016) was used to create and train CNN-COVID. The work was developed using a I7-8750H Intel processor, 2.21GHz CPU, 16.0 GB RAM and a GeForce GTX 1060 graphic card with Max-Q Design.

Convolution layer

A new convolution operation is 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 kernel in each possible 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 a kernel is applied - in the image. The convolution operation is executed in every region the image (or hidden layer) to define the next layer (in which activations keep their spatial relations in the previous layer) (Ponti and Da Costa, 2018; Aggarwal 2018; Lecun et al. 2015).

There may be several kernels in the convolutional layer. Every kernel has a feature, such as an edge or a corner. During the forward pass, every kernel is slid to the image width and height (or hidden layer), thus, generating the feature map (Ponti and Da Costa, 2018; Aggarwal 2018; Lecun et al. 2015; Balas et al. 2020).

Adaptive Moment Estimation (ADAM)

CNN-COVID uses Adaptive Moment Estimation (ADAM), an adaptive optimization technique which saves an exponentially decaying average of previously squared gradientsv_t. Besides, ADAM also computes the average of the second moments of the gradients m_t (Wani et al. 2020; Kingma and Ba 2014).

Average and non-centered variance values $m_t$are presented in Equations 1 and 2, respectively:

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

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g t^2 \qquad (2)$$

ADAM updates exponential moving averages of the gradient and the squared gradient where the hyperparameters β1, β2 ∈ [0, 1] control the decay rates of these moving averages (Equations 3 and 4):

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \qquad (3)$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \qquad (4)$$

The final formula for the update is presented in Equation (5):

$$w_{t+1} = w_t - \frac{\alpha.\hat{m}_t}{\sqrt{\hat{v}_t}+\epsilon} \qquad (5)$$

Where α is the learning rate and ϵ is a constant added to the denominator for quick conversion methods in order to avoid the division by 0 (Wani et al. 2020; Kingma and Ba 2014).

Dropout technique

CNN-COVID 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 add 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 (Wani et al. 2020; Balas et al. 2020).

Activation functions

An activation function feeds the output of every convolutional layer. The activation function layer consists of an activation function which uses the resource 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 (Wani et al. 2020). The functions used were the following:

- Sigmoid / Logistic activation function: The sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$ is a curve shaped like an S (Ponti and Da Costa, 2018).
- The activation function $f(x) = max(0, x)$ is called Rectified Linear Unit – ReLU (Ponti and Da Costa, 2018) and generates a non-linear activation map.

Pooling Layer

The pooling layer, or down sampling layer, is used to reduce the receptive field spatial size, thus, reducing the number of network parameters. The pooling layer selects each convolutional layer feature map and creates a reduced sample. Max-pooling was the technique used for this work. It generates the maximum value in the receptive field. The receptive field is 2x2, therefore, max-pooling will issue the maximum of the four input values (Wani et al. 2020).

Fully Connected Layer

After the convolution and pooling processes, the next step is to decide based on the features detected. This is done by adding one or more fully connected layers to the end. In the fully connected layer, each neuron from the previous layer is connected to each neuron from the following one. All values contribute to predict how strong a value correlates to a given class (Wani et al. 2020). The fully connected layers may be layered on top of one another to learn even more sophisticated features combinations. The output of the last fully connected layer is fed by an activation function which generates the class scores. The sigmoid activation function is the one used for CNN-COVID. It produces class scores, and the class with the highest score is treated as the correct one (Wani et al.

2020).

CNN-COVID Structure

Convolutional Neural Networks (CNNs) were proposed to assess image data. The name comes from the convolution operator, a simple way of doing complex operations using the convolution kernel (Ravi et al. 2017).

Many variations of CNN were proposed, such as AlexNet (Krizhevsky et al. 2012), Clarifai (Zeiler and Fergus 2014), GoogleNet (Szegedy et al. 2015). CNN-COVID structure is also a variation of CNN with the following architecture: an input layer, a convolutional layer, a dense layer, and an output layer, as per Figure 1.

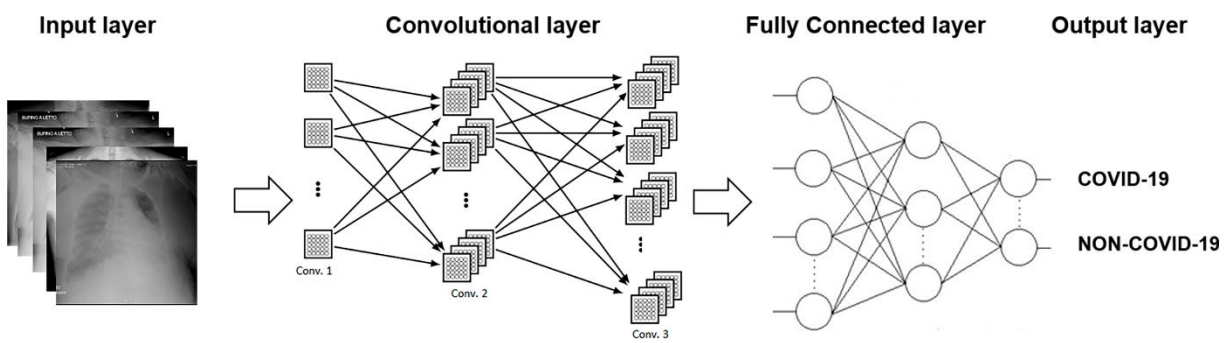

Fig. 1: Deep neural network classification scheme. Created by the author

CNN-COVID detailed architecture is illustrated in Table 2. The network consists of conventional layers, including input, convolution layer, max-pooling layer and fully-connected layers.

Besides, 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).

Table 2: CNN-COVID architecture. The network contains the input (I), the convolution (C), the max-pooling (M) layers and the fully connected network (F). Created by the author

Layer	**CNN-COVID**		
Filter	**Dimensions**	**Input/Output**	**Dimensions**
0	I		300x300
1	C	5x5x256	296x296x256
2	M	2x2	148x148x256
3	C	3x3x128	146x146x128
4	M	2x2	73x73x128
5	C	3x3x64	71x71x64
6	M	2x2	35x35x64
7	C	3x3x32	33x33x32x
8	M	2x2	16x16x32
9	F	16x16x32x256	1x256
10	F	1x1x256x128	1x128
11	F	1x1x128x64	1x64
12	F	1x1x64x32	1x32
13	F	1x132x2	1x2

CNN-COVID Training

In the training phase, weights are initialized randomly. The network was trained as per the ADAM model (Wani et al. 2020). Standard parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$ were used (Kingma and Ba, 2014), as well as the initial learning rate $\alpha = 0.001$ reduced by a factor of 10.

The ADAM training model (Wani et al. 2020) has a better performance compared with other adaptive techniques; it has a quick conversion rate, thus, it reduces the chances of error and increases accuracy. It also overcomes problems faced by other optimization techniques, such as decaying learning rate, high variance in updates, and slow convergence (Wani et al. 2020).

CNN-COVID input parameters

Several options were tested when choosing the input parameter values and the CNN-COVID batch size, considering the performance capacity of the available hardware. For input sizes 200x200 and 220x220, the batch size was 20. For larger input sizes, the batch size was 10. The input parameter

tests were run in 500 epochs. The test with better accuracy results was the one with input size 300x300 and batch size 10, as per Figure 2.

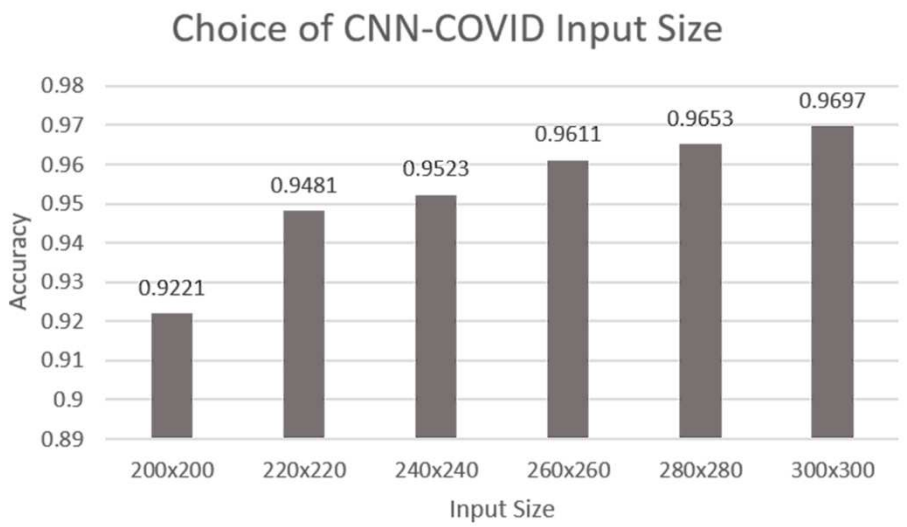

Fig. 2: Choice of CNN-COVID Input Size (200x200 300x300). Created by the author

CNN-COVID training, testing, and validation

For the training and testing phases, the network was tested 10 times varying the parameters used by the ImageDataGenerator class to generate new samples on the image databases for both COVID-19 and NON-COVID classes. For the validation phase, the network with the best accuracy in the test phase (10 tests) was considered.

Performance metrics

The following metrics were used to validate the CNN-COVID system:

- Accuracy (ACC): accurate classification rate as per the total number of elements.
- Recall/Sensitivity (Se): true positive rate.
- Specificity (Sp): true negative rate.

- F1-Score: relationship between precision and recall/sensitivity.

They are commonly used to assess the performance of classification algorithms (Ruuska et al. 2018; Skansi 2018; Khatami et al. 2017). 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. For a classification of two classes, the confusion matrix is presented on Table 3.

Table 3: Example of a confusion matrix. Adapted from (Skansi 2018)

	Classifier says YES	**Classifier says NO**
In reality YES	True positives	False positives
In reality NO	False negatives	True negatives

The confusion matrix allows us to determine the following metrics (Narin et al. 2020; Ruuska et al. 2018; Skansi 2018; Khatami et al. 2017):

- Accuracy: $ACC = \frac{\mathrm{TN+TP}}{\mathrm{TN+TP+FN+FP}}$
- Recall/Sensitivity: $Se = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$
- Specificity: $Sp = \frac{\mathrm{TN}}{\mathrm{TN+FP}}$
- F1 Score: $F1 - Score = \frac{2\times(Precision\times Recall)}{(Precision+Recall)}$

It is also possible to get the Receiver Operating Characteristic Curve (ROC curve). ROC analysis is often called the ROC accuracy ratio, a common technique for judging the accuracy of default probability models (Shirazi et al. 2018).

I. Results

A total of 10 trainings and tests were done with CNN-COVID. A total of 2000 epochs were applied to each training and test. The results are presented in Figure 3. The lowest accuracy was 0.9740 and the highest, 0.987. The test overall average for CNN-COVID showed 0.9787 accuracy.

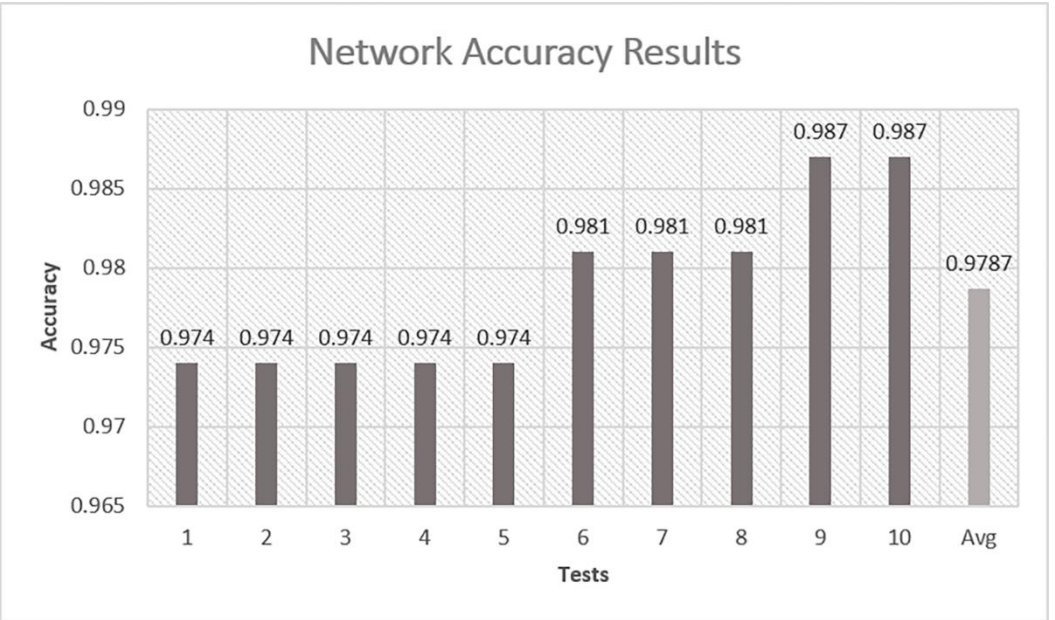

Fig. 3: CNN-COVID results: results for each test and overall average (avg) of all 10 tests. Source: Created by the author

After analyzing the results, the CNN-COVID test presenting the best accuracy amongst the 10 tests was chosen for the validation phase. Thus, the confusion matrix was generated with 126 COVID-19 images and 126 NON-COVID images, without using data augmentation to verify whether the network was avoiding overfitting. It resulted in a total of 252 images for validating CNN-COVID. Therefore, as per the confusion matrix from Figure 4, we see the following: true positives = 124, true negatives = 121, false positives = 5 and false negatives = 2.

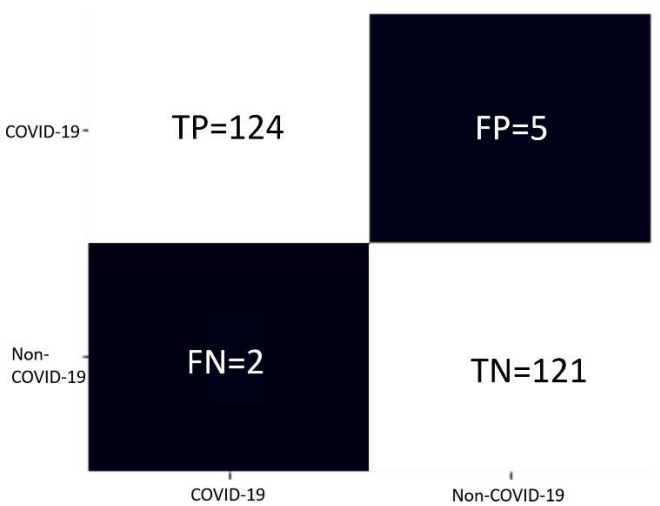

Fig. 4: CNN-COVID confusion matrix. Source: Created by the author

Using the TP, TN, FP and FN parameters, the following metrics were calculated: accuracy,

sensitivity, specificity and F1 Score, as per Table 4.

Table 4: Metrics results. Source: Created by the author

Class	Accuracy	Recall/	Class	Accuracy
COVID-19	0.9722	0.98	0.96	0.97

As per the values presented in Table 4, the ROC curve was calculated as (1-Sp) = 0.04 and Se = 0.98, as x and y, respectively. As per the ROC assessment, the area (AUC) was 9.972, as per Figure 5.

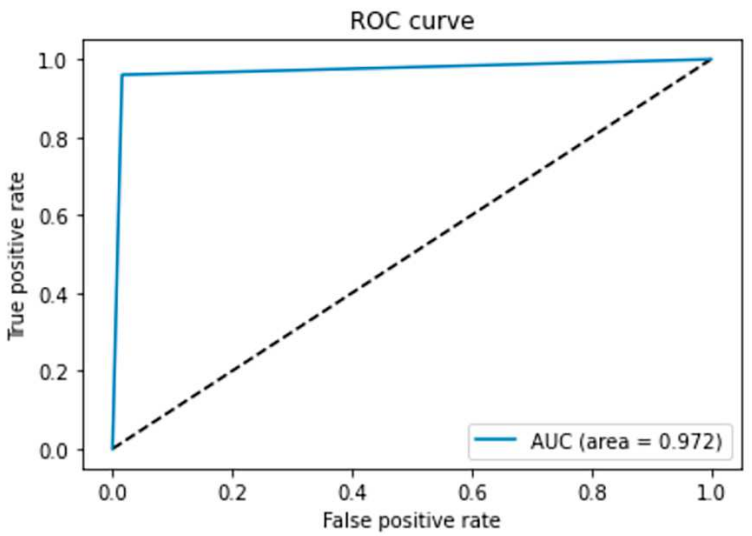

Fig. 5: CNN-COVID ROC curve/AUC = 0.972. Source: Created by the author

Thus, the CNN-COVID validation results were compared with state-of-the-art methods used in this paper and the best results were found in the following metrics: 98% Se and 97.2% ACC, as per Table 5. Besides, the methodology used for CNN-COVID is superior to the other papers mentioned, for it uses an average of 10 tests, that is, the average accuracy from 10 tests run with CNN-COVID was used, in which the test accuracy minimum value was 0.9740 and the maximum, 0.987.

Table 5: comparison of CNN-COVID with state-of-the-art methods. Created by the author

Class	Accuracy	Recall/	Class
(Zhang et al. 2020)	0.9518	0.9600	0.7060
(Sethy and Behera 2020)-ResNet50	0.9538	0.9727	0.9347
(Sethy and Behera 2020) -VGG16	0.9276	0.9747	0.8805
(Wang and Wong 2020)	0.9333	0.9100	**0.9900**
(Abbas et al. 2020)	0.9512	0.9791	0.9187

Proposed Method: CNN-COVID	**0.9720**	**0.9800**	0.9600

Discussion

For the chest x-ray images have various sizes over 1000x1000 pixels, a study was needed to decide which image sizes should be remodeled to have the best input size for CNN-COVID. The initial input size was 200x200, and then, it was increased by 20 until it reached 300x300. CNN-COVID processed every input size in 500 epochs. This study showed that the higher the input size, the more accurate the network was. As per Figure 2, when the input size was 300x300, the accuracy was 0.9697.

The algorithm for gradient-based optimization of stochastic objective functions chosen was the Adaptive Moment Estimation (ADAM) . It is an adaptive optimization technique which leverages both AdaGrad (the ability to deal with sparse gradients) and RMSProp (the ability deal with non-stationary objectives) (Wani et al. 2020). Besides, the method is straightforward to implement and requires little memory.

As there was a reduced number of COVID-19 images, augmenting the database was needed, to increase the CNN-COVID generalization. The database augmentation was done during run time. In the training phase, 126 COVID-19 images generated 252,000 new images, and 800 NON-COVID images generated 1,600,000 new images bringing it all to total of 1,852,000 images. In the test phase, 40 COVID-19 images generated 80,000 new images, and 200 NON-COVID images generated 400,000 new images bringing it all to total of 480,000 images. In the validation phase, the data generator was not used.

In this paper, we considered testing CNN-COVID 10 times to have a more trustworthy ACC. Therefore, it was possible to get an average accuracy of 0.9787 for 10 tests. For the validation phase, ACC was 0.9722. With these results, it was demonstrated that the validation accuracy was close to the average accuracy of tests, 0.9787, with less than 1% error, which proves that the average accuracy

is, indeed, the most trustworthy. Thus, these results indicate that the investment in time, human resources, money, as well as the financial and computational investments, to create and enhance techniques based on machine learning is a promising approach to help professionals diagnose COVID-19 though chest x-ray images.

Conclusion and future work

This article proposed a deep neural network, CNN-COVID, to aid diagnose COVID-19. A real dataset was applied using both COVID-19 image data collection (Cohen and Dao 2020) and ChestXray14 dataset (Wang et al. 2019). A total of 10 tests were run varying the database using data augmentation and the test average was 0.9787 accurate.

The results proved that the CNN-COVID model is a promising tool to help physicians classify chest images with pneumonia, considering atypical pneumonia caused by COVID-19 and pneumonia due to other causes. We hope this technology enhances the provision of healthcare services, contributing to the disease prognosis through straightforward exams, such as chest x-ray, and broadening the access to information through tools that help with using images for diagnosing.

For future work, we plan to improve the CNN-COVID accuracy as new COVID-19 data is collected.

Conflict of Interest and Ethical Standards

Conflict of Interest: Author Pedro Moises de Sousa declares that he has no conflict of interest. Author Mariane Modesto Oliveira declares that she has no conflict of interest. Author Gabrielle Macedo Pereira declares that she has no conflict of interest. Author Carlos Alberto da Costa Junior declares that he has no conflict of interest. Author Luis Vinicius de Moura declares that he has no conflict of interest. Author Christian Mattjie declares that he has no conflict of interest. Author Pedro Cunha

Carneiro declares that he has no conflict of interest. Author Ana Maria Marques da Silva declares that she has no conflict of interest. Author Ana Cláudia Patrocínio declares that she has no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

References

Abbas, A.; Abdelsamea , M. M.; Gaber, M. M. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. medRxiv,2020; https://doi.org/10.1101/2020.03.30.20047456

Aggarwal, Charu C. Neural networks and deep learning. Springer, v. 10, p. 978-3, 2018. https://doi.org/10.1007/978-3-319-94463-0

Ai, Tao et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology, p. 200642, 2020; https://doi.org/10.1148/radiol.2020200642

Balas, Valentina Emilia et al. (Ed.). Handbook of Deep Learning Applications. Springer, 2019; https://doi.org/10.1007/978-3-030-11479-4

Chollet, Francois. Building powerful image classification models using very little data. Keras Blog, 2016. https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html. Acessado em 01 de junho de 2020.

Chung, Michael et al. CT imaging features of 2019 novel coronavirus (2019-nCoV). Radiology, v. 295, n. 1, p. 202-207, 2020. https://doi.org/10.1148/radiol.2020200230

Cui, Jie; Li, Fang; Shi, Zheng-Li. Origin and evolution of pathogenic coronaviruses. Nature reviews Microbiology, v. 17, n. 3, p. 181-192, 2019; https://doi.org/10.1038/s41579-018-0118-9

Cohen, J.P., Morrison, P., Dao, L., Covid-19 image data collection. arXiv 2003.11597, 2020. URL:https://github.com/ieee8023/covid-chestxray-dataset. Acessado em 1 de junho de 2020.

Dai, Wei-cai et al. CT imaging and differential diagnosis of COVID-19. Canadian Association of Radiologists Journal, v. 71, n. 2, p. 195-200, 2020; https://doi.org/10.1177/0846537120913033

Khatami, Amin et al. Medical image analysis using wavelet transform and deep belief networks. Expert Systems with Applications, v. 86, p. 190-198, 2017;

https://doi.org/10.1016/j.eswa.2017.05.073

Kingma, Diederik and Ba, Jimmy. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.

Krizhevsky, Alex; Sutskever, Ilya; Hinton, Geoffrey E. Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. 2012. p. 1097-1105.

Lecun, Yann; Bengio, Yoshua; Hinton, Geoffrey. Deep learning. Nature, v. 521, n. 7553, p. 436-444, 2015; https://doi.org/10.1038/nature14539

Martin, David R. et al. A deep learning convolutional neural network can recognize common patterns of injury in gastric pathology. Archives of pathology & laboratory medicine, v. 144, n. 3, p. 370-378, 2020; https://doi.org/10.5858/arpa.2019-0004-OA

Narin, Ali; Kaya, Ceren; Pamuk, Ziynet. Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. arXiv, p. arXiv: 2003.10849, 2020.

Ozturk, Saban; Ozkya, Umut; Barstugan, Mucahid. Classification of Coronavirus Images using Shrunken Features. medRxiv, p. 2020.04.03.20048868, 2020; https://doi.org/10.1101/2020.04.03.20048868

Ponti, Moacir Antonelli; Da Costa, Gabriel B. Paranhos. Como funciona o deep learning. arXiv preprint arXiv:1806.07908, 2018.

Ravi, D. et al. Deep Learning for Health Informatics. IEEE journal of biomedical and health informatics, v. 21, n. 1, p. 4-21, 2017; https://doi.org/10.1109/JBHI.2016.2636665

Ruuska, Salla et al. Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. Behavioural processes, v. 148, p. 56-62, 2018; https://doi.org/10.1016/j.beproc.2018.01.004

Sethy, Prabira Kumar; Behera, Santi Kumari. Detection of coronavirus disease (COVID-19) based on deep features. Preprints, v. 2020030300, p. 2020, 2020; https://doi.org/10.20944/preprints202003.0300.v1

Shirazi, Amin Zadeh; Chabok, Seyyed Javad Seyyed Mahdavi; Mohammadi, Zahra. A novel and reliable computational intelligence system for breast cancer detection. Medical & biological engineering & computing, v. 56, n. 5, p. 721-732, 2018. https://doi.org/10.1007/s11517-017-1721-z

Skansi, Sandro. Introduction to Deep Learning: from logical calculus to artificial intelligence. Springer, 2018. https://doi.org/10.1007/978-3-319-73004-2

Szegedy, Christian et al. Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. p. 1-9, 2015; https://doi.org/10.1109/CVPR.2015.7298594

Wang, Linda; Wong, Alexander. COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images. arXiv preprint arXiv:2003.09871, 2020

Wang, Xiaosong et al. ChestX-ray: Hospital-Scale Chest X-ray Database and Benchmarks on

Weakly. Deep Learning and Convolutional Neural Networks for Medical Imaging and Clinical Informatics, p. 369, 2019; https://doi.org/10.1007/978-3-030-13969-8_18

Wani, M. Arif et al. Advances in Deep Learning. Springer, 2020; https://doi.org/10.1007/978-981-13-6794-6

Weiss, Susan R.; Leibowitz, Julian L. Coronavirus pathogenesis. In: Advances in virus research. Academic Press, v.81, n. 1, p. 85-164, 2011; https://doi.org/10.1016/B978-0-12-385885-6.00009-2

Yang, Wenjie et al. Clinical characteristics and imaging manifestations of the 2019 novel coronavirus disease (COVID-19): A multi-center study in Wenzhou city, Zhejiang, China. Journal of Infection, 2020; https://doi.org/10.1016/j.jinf.2020.02.016

Zeiler, Matthew D.; Fergus, Rob. Visualizing and understanding convolutional networks. In: European conference on computer vision. Springer, Cham, 2014. p. 818-833; https://doi.org/10.1007/978-3-319-10590-1_53

Zhang, Jianpeng et al. Covid-19 screening on chest x-ray images using deep learning based anomaly detection. arXiv preprint arXiv:2003.12338, 2020.

Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. New England Journal of Medicine, v. 382, n. 8, p. 727–733, 2020; https://doi.org/10.1056/NEJMoa2001017

Figures

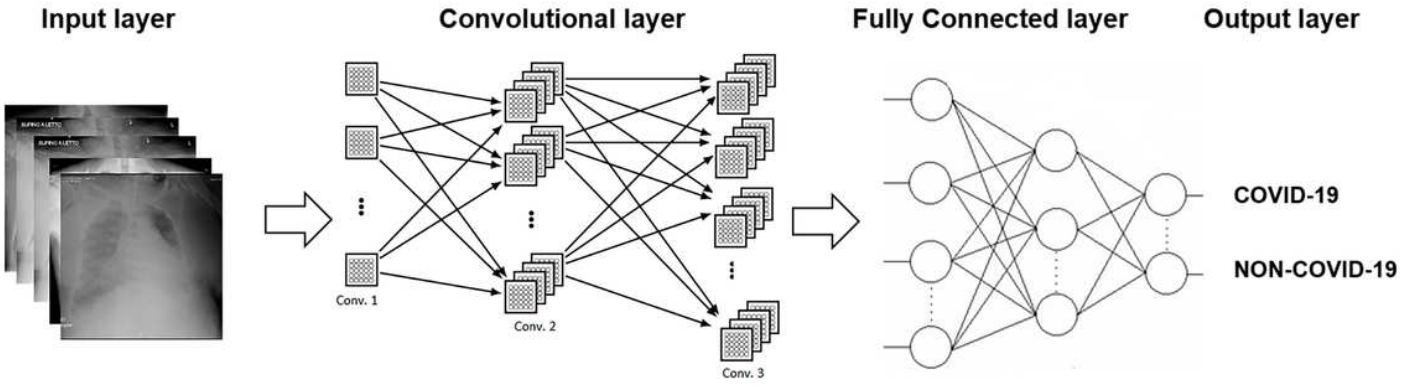

Figure 1

Deep neural network classification scheme. Created by the author

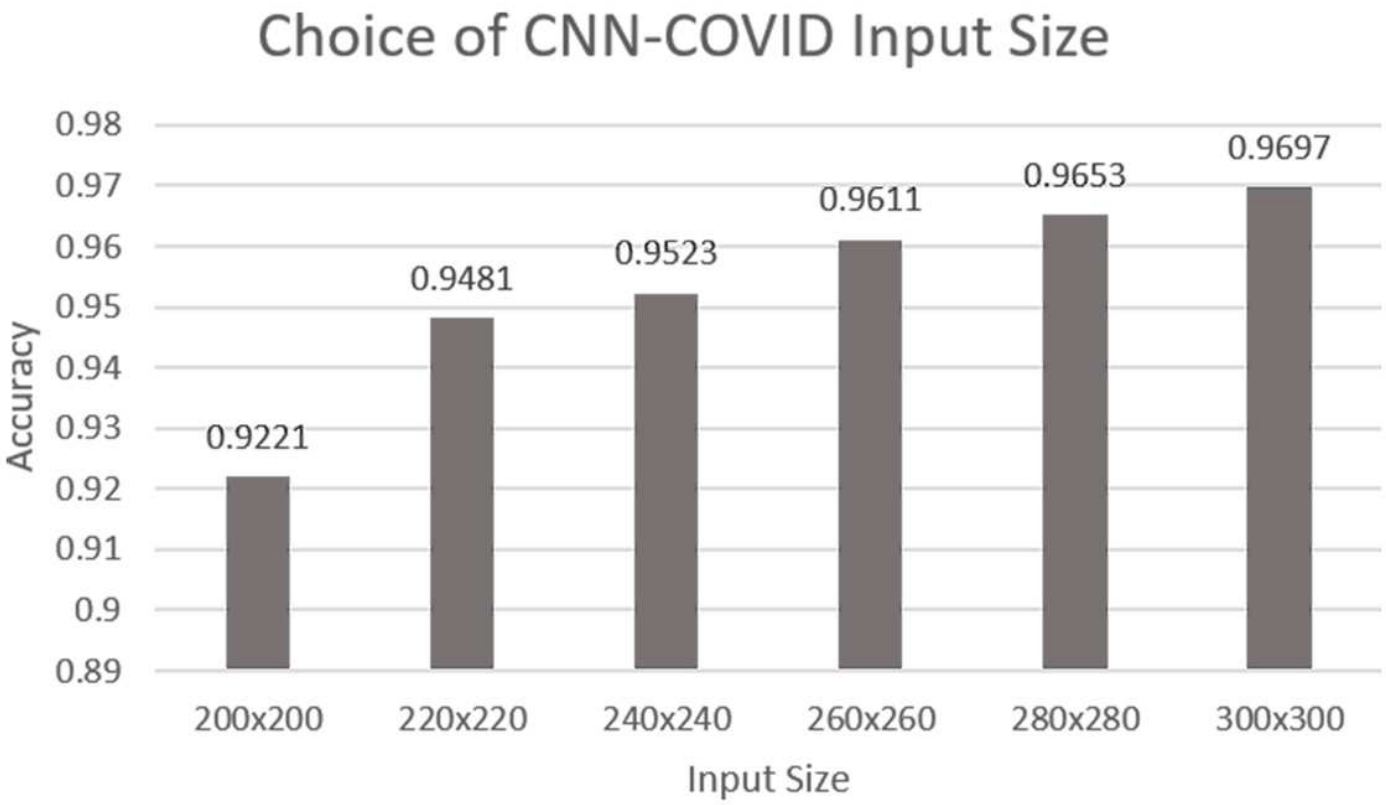

Figure 2

Choice of CNN-COVID Input Size (200x200 300x300). Created by the author

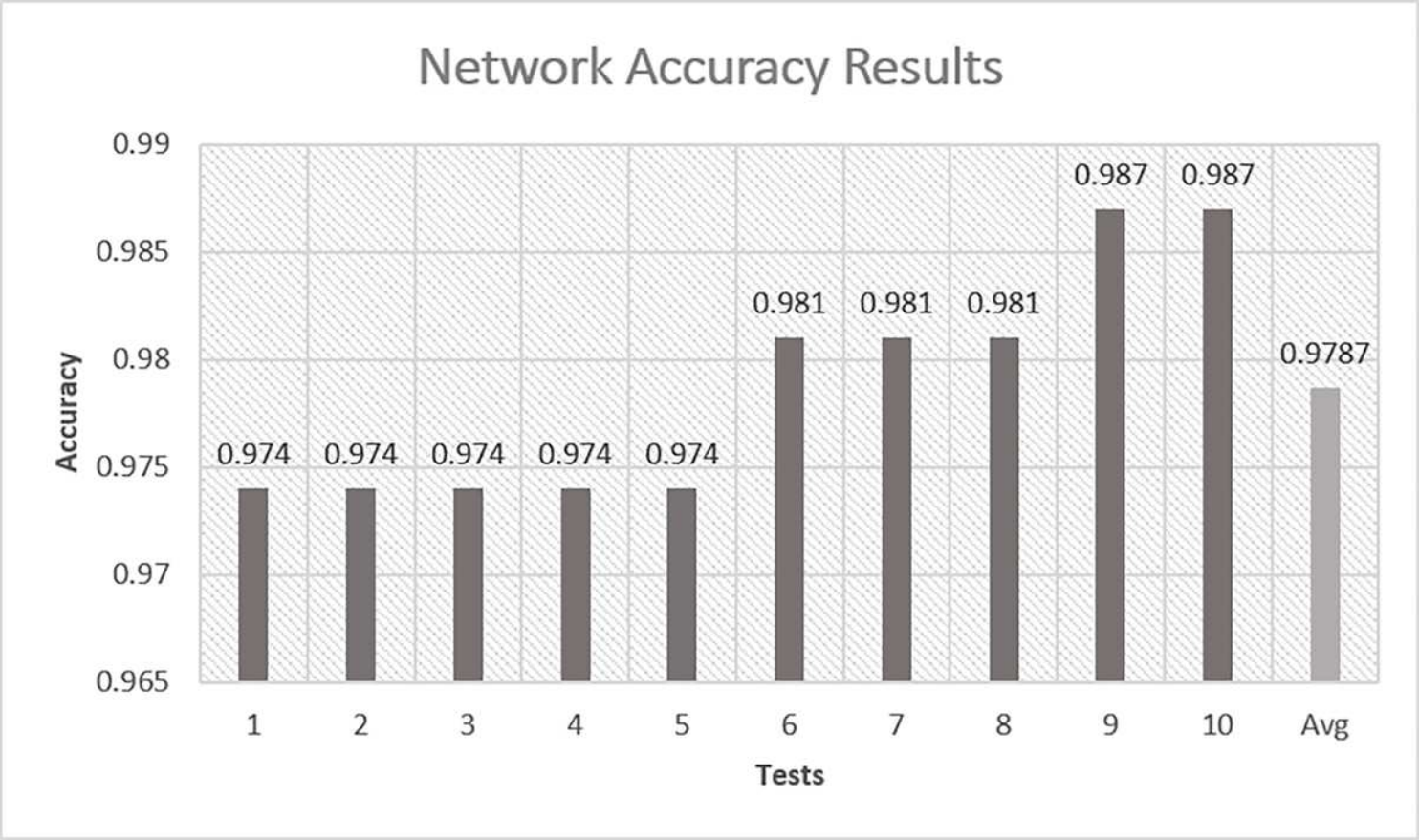

Figure 3

CNN-COVID results: results for each test and overall average (avg) of all 10 tests. Source: Created by the author

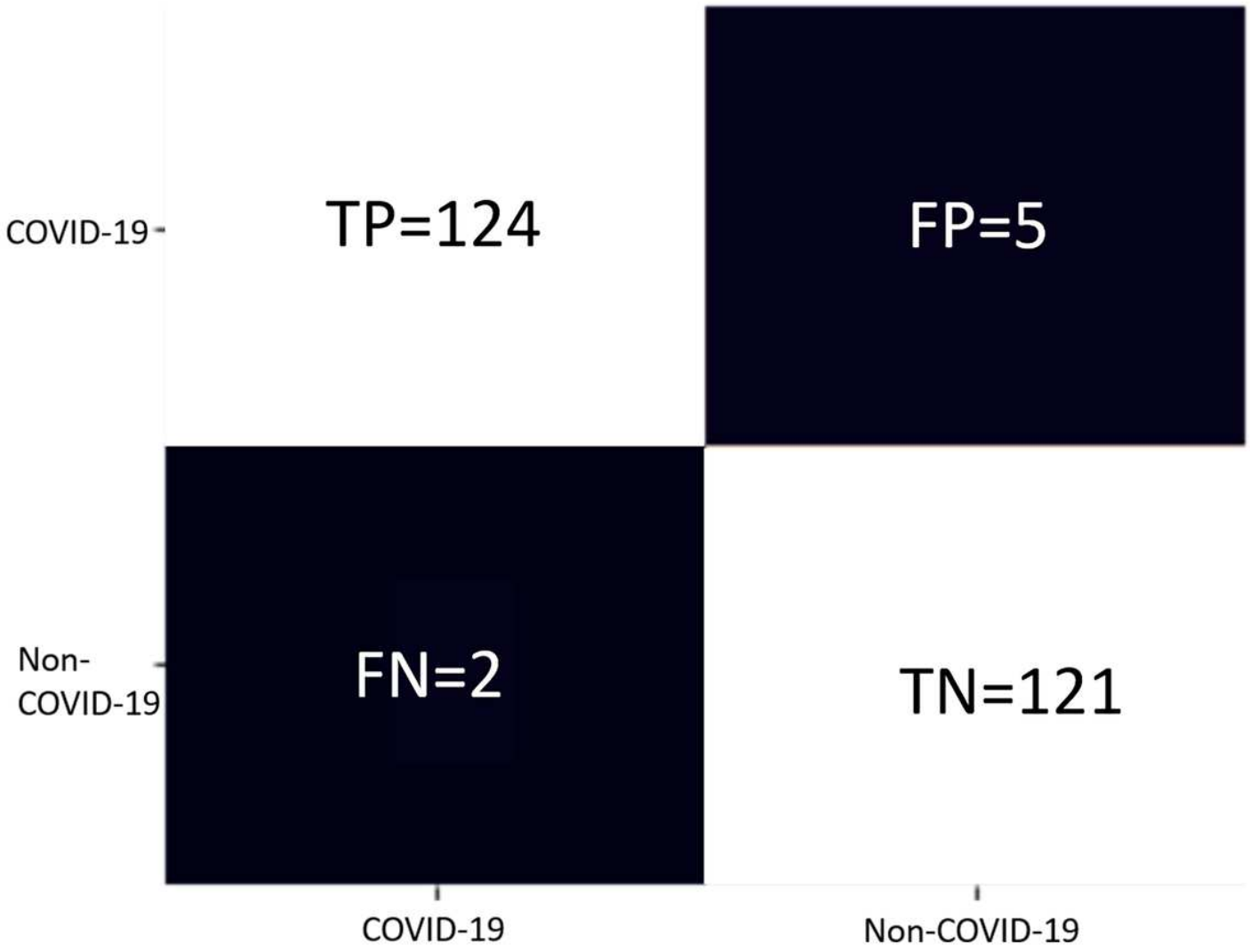

Figure 4

CNN-COVID confusion matrix. Source: Created by the author

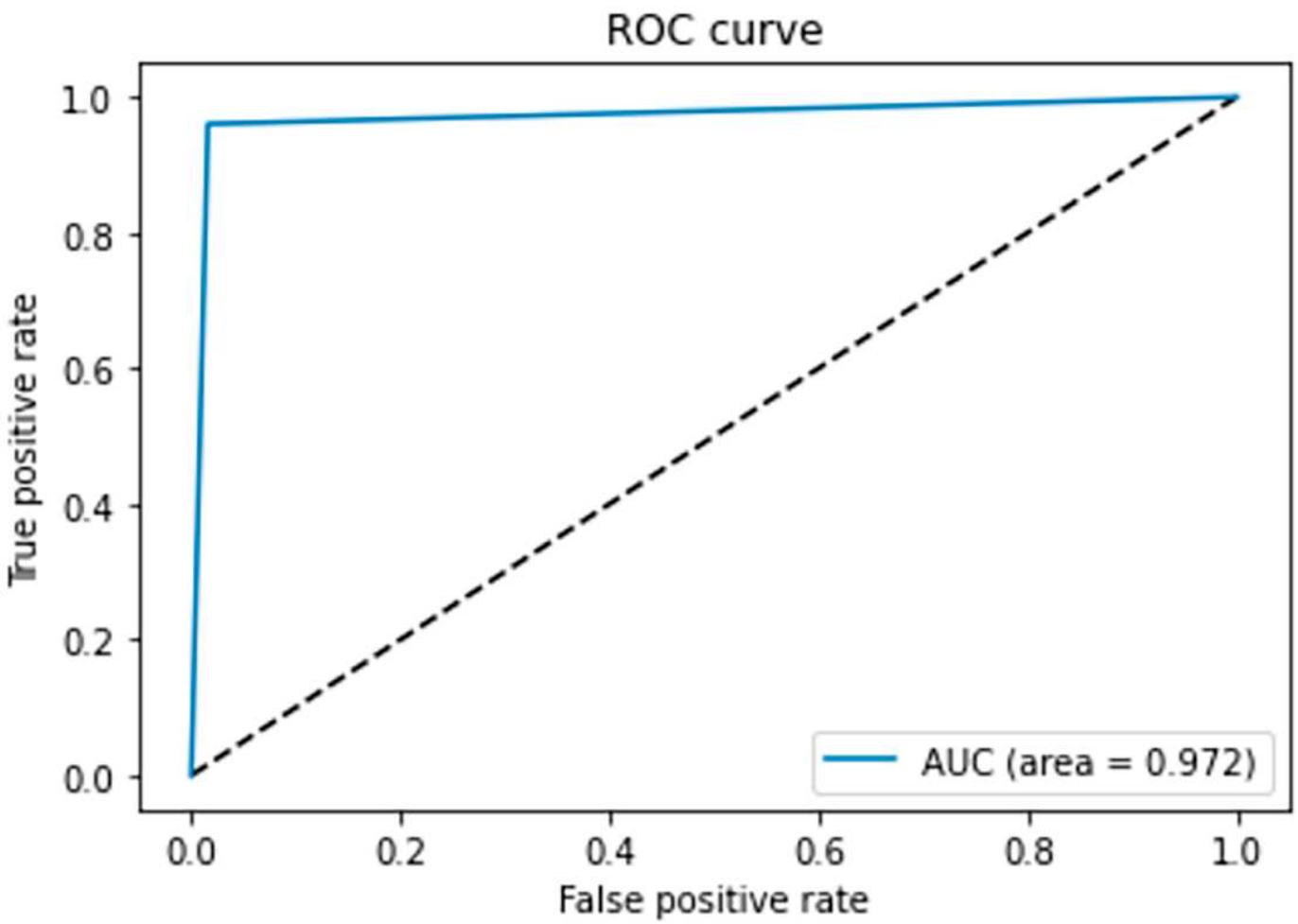

Figure 5

CNN-COVID ROC curve/AUC = 0.972. Source: Created by the author

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- databaseCovidPublica.txt
- databaseNCovidPublica.txt