The 1st International and 3rd National Conference on Biomathematics
19-21 January 2022, Damghan University, Iran.

A CNN model for the detection of Covid-19 using lung X-ray images

Amin Alidoost Jehezdani[1]*, Seyed Hashem Tabasi [1]

[1] Faculty of Mathematics and Computer Science, Damghan University, Damghan, Iran
Aminalidoost.ai@gmail.com

***Abstract*— Covid 19 disease causes the most damage to the lungs and it can be diagnosed by X-ray imaging of the lungs. In the period of the Covid-19 epidemic, the speed with which medical images are examined is important. By building high-precision computer models instead of using the human eyes, check speed is increased and the error rate in diagnosing the can be minimized. In this work, medical images were collected with different percentages of infection and a model is constructed using convolutional neural network.**

Keywords*— *classification, covid-19, convolutional, neural network, model, x-ray

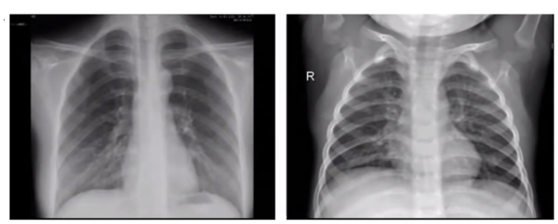
I. Introduction

In December 2019, the new SARS-CoV 2 virus-induced coronavirus was first observed in Wuhan and spread rapidly worldwide. The massive outbreak of the disease resulted in millions of people being infected. By the end of October 2021, more than 5 million people had been killed worldwide. Due to the high similarity of Covid 19 with the flu and cold, a correct and timely diagnosis of the disease can prevent the death of the patient. In case of mutation of the virus and increase in the number of patients, due to the lack of medical staff in front of a large number of patients, the possibility of misdiagnosis increases.

Today, computers are increasingly used in the medical profession. The main applications of computers in medicine include hospital information system, data analysis in medicine, medical imaging laboratory calculations, computer-aided medical decision-making, and care for critically ill patients, Computer-assisted therapy and so on [1].

COVID-19 is a respiratory disease. Doctors can use chest imaging tests to diagnose covid-19 when they are waiting for RT-PCR[1] test results or when the RT-PCR test is negative and the patient is experiencing covid-19 symptoms.

Chest imaging includes the use of x-rays, nuclear scanning, magnetic resonance imaging (MRI), computed tomography (CT) scanning, including positron emission tomography (PET) scanning, and ultrasonography [2]. Chest X-rays provide images of structures inside and around the chest and are very useful for identifying abnormalities of the heart, lung parenchyma, chest wall, and diaphragm. They are usually the initial test to assess the lungs [3]. In some cases, such as Fig.1, it may be difficult to diagnose Covid 19 disease. In such situations, having high-precision computer technology can be helpful.



a b

Fig. 1. Picture of Covid 19 patient's lung (a) and picture of healthy lung

Artificial neural networks are new computational methods for machine learning. The main task of the artificial neural network is to model the processing properties of the human brain to approximate conventional computational methods with biological

[1] reverse transcription polymerase chain reaction

processing methods. In other words, an artificial neural network is a method that, through training, learns the knowledge of the connection between several data sets and stores them for use in similar cases.

convolutional neural network (CNN) is a class of deep neural networks commonly used for visual or verbal analysis in machine learning [4]. In this CNN model, 320 images related to sick and healthy people have been used.

II. LITERATURE REVIEW

A. Artificial Neural Networks

The human brain is the most complex system ever observed and studied. The complexity of this unique system goes back to the many connections between its components. The excellent performance of the brain in solving all kinds of problems and its high efficiency has made brain simulation the most important goal of scientists.[5] In the field of artificial neural networks, several mathematical and software models inspired by the human brain have been proposed, which are used to solve a wide range of scientific, engineering and applied problems in various fields.

The human brain is made up of small units called neurons. In natural neurons, information is received from dendrites. All inputs are collected in the kernel and processed there. Eventually, the processed information leaves the neuron through axons and synapses.

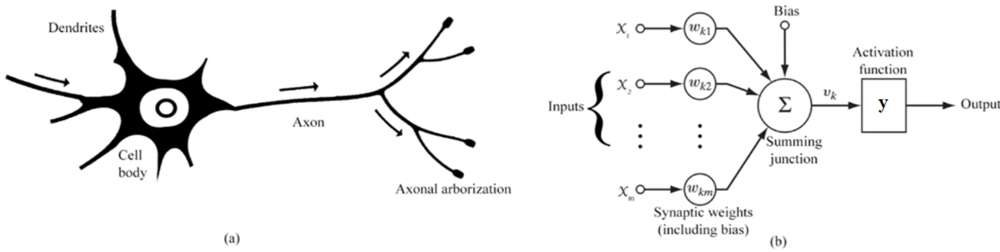


Fig. 2. Natural neuron (a), Artificial neuron (b)

In artificial neurons, the x_1 to x_m inputs act as dendrites. The thickness of the dendrites is directly related to the importance of the information in that branch. This is done in artificial neurons through weights. In fact, weights w_{k2} to $w_{km}$are equivalent to the same dendritic thickness. In an artificial neuron, the inputs are multiplied by a weight and then sum together. In artificial neurons, after the sum operation, the neuron information will be output through y. Axons and synapses are equivalent to y in artificial neurons.

The activation function adds a nonlinear property to the neuron. Linear, sigmoid, tangent and ReLU are some of the activation functions. In addition to the activation function, another parameter is added to the artificial neuron and it is called bias. The bias is a fixed number that is entered directly into the activation function. The first feature of the bias is that it makes it difficult to output zero. The second property of bias is to give activation flexibility, which with the addition of bias, the activation function can be shifted and help to learn better.

$$\boldsymbol{y = f(z) = f(W^T X + x_0)} \qquad (1)$$

Equation (**1**) is a mathematical representation of neurons in which W represents the weight vector, X represents the input vector, and W_0 represents the bias.[6]

B. CNN Model creating

There are two approaches to machine learning: supervised and unsupervised. In a supervised model, a training dataset is fed into the classification algorithm. That lets the model know what is, for example, “positive” transactions. Then the test data sample is compared with that to determine if there is a “negative” transaction. This type of learning falls under “Classification”. Classification is a science that is one of the subfields of machine learning and data mining and creates a model for predicting new data labels based on previously labeled data.

Clustering is framed in unsupervised learning, for this type of algorithm we only have one set of input data (not labelled), about which we must obtain information, without previously knowing what the output will be.

Convolution neural network, like other neural networks, consists of layers with weights and bias with the ability to learn. The purpose of building a convolutional neural network is to build many layers of feature detectors to calculate the spatial arrangement of pixels in an image. CNN differs from some neural networks, such as MLP (multilayer perceptron), in input.[7]

CNN is designed to work well for inputs with matrix structure (two-dimensional and three-dimensional). The MLP network changes the structure of the input data and converts a two-dimensional 100 × 100 matrix into a 10,000-dimensional vector. But CNN does not change the structure of the input and cares about the connection between neighboring pixels.

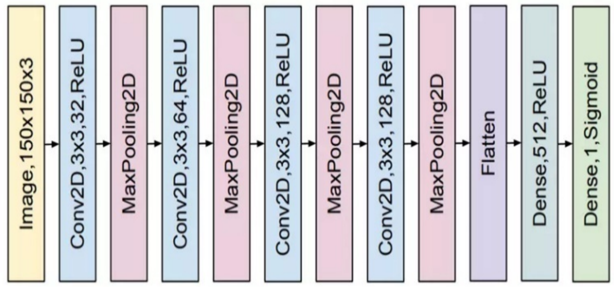


Fig. 3. A simple architecture of convolutional neural network

Fig. 3. shows the yellow block of the input layer, blue blocks are convolutional layers and nonlinear activation function that usually show in one layer. Pooling layers are the red blocks and the green block is for fully connected layer. The core of CNN is the convolution layer, which accounts for the vast majority of convolution neural network calculations. In the convolution operator, according to Fig. 4, there are four important components, which are: input matrix or image, convolution filter, operator * and Output. The convolution operator (*) picks up the convolution kernel or filter and slides it over the image or input matrix.

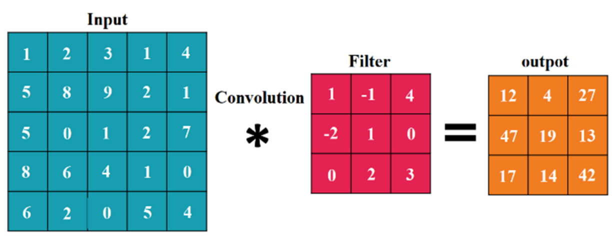


Fig. 4. Four main components in convolution

In convolutional neural networks, the input is usually a color image (three-dimensional matrix). In convolutional neural networks, the input is usually a color image (three-dimensional matrix). Therefore, the filter must have three channels. This means that if a 3 × 3 filter was used for convolution in a 2D image, now a 3 × 3 × 3 filter is needed, because each page of the color image has a separate filter for itself. [8]

To build a model, the data is divided into three parts: training data, validation data, and test data. The model is trained using training data, and the parameters are adjusted using a set of validations that the model did not see at the time of training, and then the performance of the model is tested on test data. After building the model, new data can be used to make predictions, then a confusion matrix is created to show how good the built model was. The confusion matrix has 4 parameters: true positive, true negative, false positive and false negative. Accurate and good model has higher true positive and higher true negative. Finally, the results are evaluated. Some evaluation criteria are: accuracy, precision, recall, F1 score.

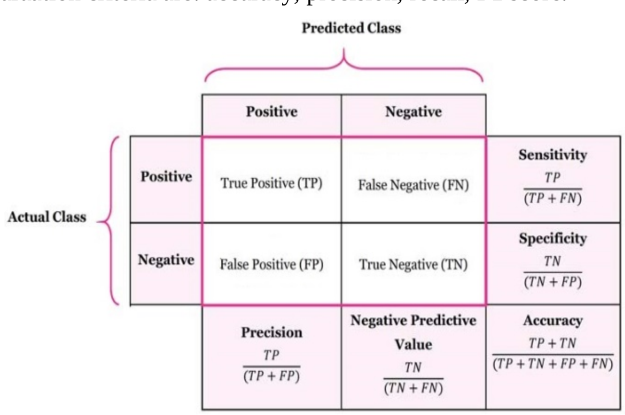


Fig. 5. Confusion-matrix-with-advanced-classification-metrics

III. METHODOLOGY

A. Dataset

The image datasets used to build the model are downloaded from sources GitHub chest x-ray dataset [9] and Radiopaedia COVID-19 database [10]. The images were divided into two categories: normal and covid. 320 images were used for training data and 60 images were used for validation data. The pictures were for people aged 18 to 94.

B. Method

Initially, the preprocessing operation is applied to the images in the dataset. In the first step, images that were taken from the wrong side should be removed. Fig. 6 shows two types of images in the dataset: front view and side view. In this work, images such as Fig. 6-b have been used.

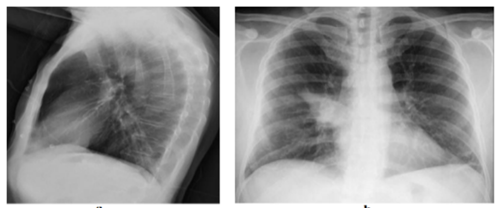


Fig. 6. Side view(a) and front view(b) of chest radiograph

Images are used in the standard size of 224 × 224. Much smaller images cause lost features and information, and with large images, model training becomes difficult. Normalization of images is done in the model training stage. In this model, 4 layers of convolution 32, 64, 64 and 128 have been used. The main libraries were Tensorflow, keras and numpy.

The activation function ReLU is used in the model. In ReLU, the negative components are set to 0. finally, because there was an output neuron, the value of Dense was set to 1, and the use of sigmoid was binary. In compiler section, “Adam” optimizer is used and in part metrics, “accuracy” is used. Throughout the learning process, Adam uses the average of the second gradient torque for the learning rate [11]. In training model section steps per epoch was10 and epochs was 9 which resulted in an accuracy of about 97%. In each epoch, the entire set passes through the neural network once forward and once backward. This will update the weights of each pass [12].

Table I. model summary

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 222, 222, 32)	896
conv2d_1 (Conv2D)	(None, 220, 220, 64)	18496
max_pooling2d(MaxPooling2D)	(None, 110, 110, 64)	0
dropout (Dropout)	(None, 110, 110, 64)	0
conv2d_2 (Conv2D)	(None, 108, 108, 64)	36928
max_pooling2d_1(MaxPooling2D)	(None, 54, 54, 64)	0
dropout_1 (Dropout)	(None, 54, 54, 64)	0
conv2d_3 (Conv2D)	(None, 52, 52, 128)	73856
max_pooling2d_2(MaxPooling2D)	(None, 26, 26, 128)	0
dropout_2 (Dropout)	(None, 26, 26, 128)	0
flatten (Flatten)	(None, 86528)	0
dense (Dense)	(None, 64)	5537856
dropout_3 (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 1)	65
Total params: 5,668,097 Trainable params: 5,668,097 Non-trainable params: 0		

Table I. shows the details of the model layers.

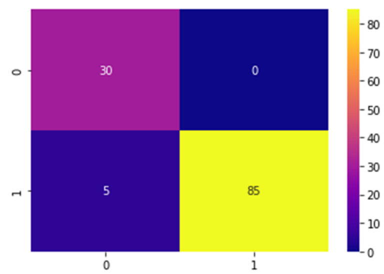
```
Epoch 1/9
10/10 [==============================] - 226s 24s/step - loss: 1.0580 - accuracy: 0.5375 - val_loss: 0.6840 - val_accuracy: 0.9000
Epoch 2/9
10/10 [==============================] - 91s 9s/step - loss: 0.6111 - accuracy: 0.6781 - val_loss: 0.4905 - val_accuracy: 0.9500
Epoch 3/9
10/10 [==============================] - 90s 9s/step - loss: 0.3909 - accuracy: 0.8250 - val_loss: 0.2933 - val_accuracy: 0.9667
Epoch 4/9
10/10 [==============================] - 90s 9s/step - loss: 0.3727 - accuracy: 0.8562 - val_loss: 0.2297 - val_accuracy: 0.9833
Epoch 5/9
10/10 [==============================] - 91s 9s/step - loss: 0.2987 - accuracy: 0.8781 - val_loss: 0.1052 - val_accuracy: 0.9667
Epoch 6/9
10/10 [==============================] - 90s 9s/step - loss: 0.3516 - accuracy: 0.8781 - val_loss: 0.2217 - val_accuracy: 0.9833
Epoch 7/9
10/10 [==============================] - 90s 9s/step - loss: 0.2450 - accuracy: 0.9031 - val_loss: 0.1854 - val_accuracy: 0.9500
Epoch 8/9
10/10 [==============================] - 90s 9s/step - loss: 0.2243 - accuracy: 0.9125 - val_loss: 0.1219 - val_accuracy: 0.9833
Epoch 9/9
10/10 [==============================] - 90s 9s/step - loss: 0.2110 - accuracy: 0.9312 - val_loss: 0.0755 - val_accuracy: 0.9667
```

Fig. 7. model training epochs

Fig. 7. shows the upward trend in the accuracy and the validation during the training.

C. Results

The results obtained from the model show its appropriate accuracy. The performance evaluation of the model was examined with some metrics as shown in Fig. 8.



a

```
classification report:
              precision    recall  f1-score   support

       Covid       0.86      1.00      0.92        30
      Normal       1.00      0.94      0.97        90

    accuracy                           0.96       120
   macro avg       0.93      0.97      0.95       120
weighted avg       0.96      0.96      0.96       120

>>>
```

b

Fig. 8. Confusion matrix(a) and classification full report(b)

In the confusion matrix, it is an attempt to maximize the TP, TN, which are the correct performance of the model, and closing to zero the FP and PN, which are the incorrect performance of the model.

The accuracy of a model on a test data set is the percentage of samples that are properly labeled. Recall and precision are widely used in classification. Recall measure is a measure of integrity, that is, the percentage of samples that are positively labeled and their class is really positive. Another way to use recall and precision is to combine them together and turn them into another simple metric called f1-score.

IV. Conclusion

Every day, computer science is more connected to other sciences than before, and medical science is no exception. It is observed that to diagnose the disease using artificial neural networks, high-precision models can be built to prevent human eyes error. It has been observed that to diagnose the disease using artificial neural networks, high-precision models can be developed to prevent human eye error. Progress in this area requires more attention to computer medicine and the provision of appropriate facilities in all hospitals and the use of software engineers as part of the patient treatment process. In future work, a model for diagnosing more diseases with different types of medical imaging and simulating the progression or improvement of the disease with different drugs will be examined.

References

[1] Perera Molligoda Arachchige, Arosh S.; Svet, Afanasy (2021-09-10). "Integrating artificial intelligence into radiology practice: undergraduate students' perspective". European Journal of Nuclear Medicine and Molecular Imaging. doi:10.1007/s00259-021-05558-y. ISSN 1619-7089.

[2] Dezube, Rebecca (2021), MD, MHS, Johns Hopkins University, Chest Imaging, https://www.msdmanuals.com/professional/pulmonary-disorders/diagnostic-and-therapeutic-pulmonary-procedures/chest-imaging.

[3] Brant WE, Helms CA (2007). "Diagnostic Imaging Methods". Fundamentals of Diagnostic Radiology (3rd ed.). Philadelphia: Lippincott Williams & Wilkins. p. 3. ISBN 9780781761352.

[4] LeCun, Yann. "LeNet-5, convolutional neural networks". Retrieved 16 November 2013.

[5] Russell, Ingrid. "Neural Networks Module". Archived from the original on 29 May 2014. Retrieved 2012.

[6] Howsam, Tim (2020), The structure of artificial neurons, https://howsam.org/artificial-neuron.

[7] Y.LeCun, L. Bemgio, and P.Haffner. Gradient-based Learning Applied to Document Recognition. Proceeding of the IEEE, 86(11):2278-2324,1998.

[8] P. Y. Simard, D. Steinkraus, and J. C. Platt. Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis. IEEE, 2003, p.958.

[9] Joseph Paul Cohen and Paul Morrison and Lan Dao and Karsten Roth and Tim Q Duong and Marzyeh Ghassemi. cohen2020covidProspective. “COVID-19 Image Data Collection: Prospective Predictions Are the Future”. arXiv 2006.11988. https://github.com/ieee8023/covid-chestxray-dataset.(2020)

[10] Bell, D., Hacking, C. COVID-19. Reference article, Radiopaedia.org. https://doi.org/10.53347/rID-73913

[11] Kingma, Diederik; Ba, Jimmy (2014). "Adam: A Method for Stochastic Optimization". arXiv:1412.6980

[12] Bertsekas, D. P. (2011). Incremental gradient, subgradient, and proximal methods for convex optimization: a survey. Optimization for Machine Learning, 85.

یک مدل شبکه عصبی کانولوشن برای تشخیص بیماری کووید - 19 با استفاده از تصاویر اشعه X ریه

امین علی دوست جهزدانی[1]*، سید هاشم طبسی[1]

[1] دانشکده ریاضی و علوم کامپیوتر، دانشگاه دامغان، دامغان، ایران
Aminalidoost.ai@gmail.com

چکیده: بیماری کووید 19 بیشترین آسیب را به ریه ها وارد می کند و با تصویربرداری اشعه ایکس از ریه ها می توان آن را تشخیص داد. در دوره اپیدمی کووید-19، سرعت بررسی تصاویر پزشکی مهم است. با ساخت مدل های کامپیوتری با دقت بالا به جای استفاده از چشم انسان، سرعت بررسی افزایش می یابد و می توان میزان خطا در تشخیص را به حداقل رساند. در این کار، تصاویر پزشکی با درصدهای مختلف آلودگی جمع‌آوری شد و مدلی با استفاده از شبکه عصبی کانولوشن ساخته شده است.

کلید واژگان: اشعه x، شبکه عصبی، کانولوشن، کووید19، مدل.