Academic Science Journal

The Use of Convolution Neural Networks to Classify Viral Pneumonia and COVID-19 by Using Chest X-ray Images

Mayssam Alwan Hasson, Taha Mohammed Hassan, Alaa Jalal Abdullah and Sarah Mohemmed Fawzi Hussein

Department of Computer Science- College of Science- University of Diyala,

Scicompms2127@uodiyala.edu.iq

Received: 22 December 2022 **Accepted: 12 February 2023**

DOI: https://doi.org/10.24237/ASJ.02.01.706D

Abstract

The novel coronavirus outbreak reached pandemic status in March-2020. Since then, many countries have collaborated in the fight against COVID-19. The main objective of these governments is the rapid and effective identification of COVID-19-positive patients. While many molecular tests currently exist, not all hospitals have immediate access to these. However, X-rays, which are easily accessible in the majority of hospitals, provide different ways to detect COVID-19. This article discusses the use of neural networks for the classification of radiographic images of patients with pneumonia and COVID-19. Precision, Recall, and F1-score were used to select the best resizing parameters and adaptive equalization of the brightness histogram of images, as well as the optimal architecture of the neural network and its hyperparameters. The high values of these classification quality metrics obtained accuracy (97%) for patients with COVID-19, and accuracy (99%) for patients with pneumonia. These results strongly indicate a reliable differentiation of radiographic images. This opens up the possibility of creating a model with good predictive ability without involving ready-made complex models and preliminary training on third-party data.

Keywords: Convolution Neural Networks, Chest CT scan image, COVID-19 Diagnosis using Deep Learning, Pneumonia, Deep Learning.

Academic Science Journal

استخدام الشبكات العصبية التلافيفية لتصنيف الالتهاب الرئوي الفيروسي وكوفيد-19 باستخدام صور الأشعة السينية للصدر

ميسم علوان حسون و طه احمد حسن والاء جلال عبد الله و ساره محمد فوزي حسين

قسم الحاسوب – كلية العلوم – جامعة ديالى

الخلاصة

اصبح فايروس كورونا المستجد وباءاً في مارس /2020 ، ومنذ ذلك الحين تعاونت الدول لمحاربة ذلك الوباء. وقد كان الهدف الاساسي لتلك الحكومات هو تشخيص المصابين ب كوفيد -19 بسرعة وفعالية . وبينما قد تكون بعض الفحوصات غير موجودة والمستشفى ليس لديها وصول مباشر لها , فعلى العكس من ذلك ان الشعة السينية متوفرة في اغلب المستشفيات. وتوَفر طريقة اضافية لتشخيص فايروس كورونا . هذا البحث يناقش استخدام الشبكات العصبية الاصطناعية لتصنيف الصور الطبية للمرضى المصابين بالالتهاب الرئوي وكورونا فايروس . معايير مهمة استخدمت و أفضل معلمات تغيير الحجم والمعادلة التكيفية للرسم البياني لسطوع الصور ، بالإضافة إلى البنية المثلى للشبكة العصبية ومعلماتها الفائقة.ان القيم الكبيرة لمعايير التصنيف ادت الى الحصول على دقة (97%) لمرضى كوفيد 19 ودقة (99%) لمرضى الالتهاب الرئوي . تشير هذه النتائج بقوة إلى تمايز موثوق للصور الشعاعية. هذا يفتح إمكانية إنشاء نموذج بقدرة تنبؤية جيدة دون تضمين نماذج معقدة جاهزة وبدون تدريب أولي على بيانات الطرف الثالث.

الكلمات المفتاحية: الشبكات العصبية الاصطناعية ، الأشعة السينية للصدر، تشخيص كوفيد – 19باستخدام التعلم العميق، التعلم العميق.

Introduction

The novel SARS-CoV-2 virus, responsible for the dangerous pneumonia-type disease, COVID-19, has undoubtedly changed the world by killing at least 3,900,000 people as of June 2021 and compromising the health of millions across the globe [1].

The rapid worldwide spread of the SARS-CoV-2 coronavirus and the associated disease COVID-19 is largely due to the lack of operational diagnostics [2]. The health facilities of many developed countries have been exhausted due to the rapid increase in corona-affected persons. These countries are facing a shortage of corona testing kits and ventilators. Due to this, some countries announced the lockdown to break the chain of coronavirus and save their population.

Methods based on the polymerase chain reaction have the highest reliability, but the result, as a rule, is ready in a few hours, while ensuring rather high requirements for compliance with the analysis technology and f the professionalism of the medical staff. Therefore, developing rapid and reliable detection systems for COVID-19 is essential to prevent a pandemic of this disease. The lungs are most severely affected by COVID-19 [3]. To recognize pneumonia caused by various viruses, computed tomography with analysis of the got images using deep learning methods is widely used [4]. Real-time reverse transcription-polymerase chain reaction (RT-PCR) is performed on the clinical samples of patients and results are obtained within a few hours to two days. Due to the low sensitivity of RT-PCR and the unavailability of kits, radiography imaging techniques can be utilized to detect COVID-19 in patients. It has been observed from research articles that chest scans may be suitable for COVID-19 detection [5]. X-ray diagnostics, although having a lower resolution, are more common because of the greater availability of x-ray machines and has less radiation exposure to patients. Currently, there are some works that propose the use of neural networks (NN) to detect COVID-19 based on the analysis of radiographic images (RI) [6]. A deep learning framework to detect Covid-19 disease (AI) involving machine learning (ML) and deep learning (DL) has a good track record in the medical image understanding scope due to its high strength of classification and feature extraction [7]. Convolutional neural network (CNN) has been widely applied to detect viral pneumonia and differentiate bacterial in chest radiographs [8].

Related Work

In LindaWang, et al [9], a convolutional neural network was first trained on an external ImageNet dataset, and then further trained using a dataset used to train and evaluate the proposed COVID-Net, which we will refer to as COVIDx, is comprised of a total of 13,975 CXR images across 13,870 patient cases, of which 183 were with COVID-19. The proposed method got positive prediction as normal 90.5, pneumonia 91.3 covid 98.9. In Ozturk T, et al [10], to detect COVID-19 diseases, a NN was created by reducing the layers and filters of the already existing DarkNet, and 127 radiographs with a diagnosis of COVID-19 were used to obtain a model. The model produced a classification accuracy of 98.08% for binary classes and

87.02% for multi-class cases. In Mohamed Loey, et al [11] the possibility of using ready-made NNs Alexnet, Googlenet, Resnet18 for diagnosing COVID-19 using only 307 RIs is shown, and the advantage of Googlenet in some scenarios is noted. These studies show the promise of this area and emphasize that its development is hindered by an extremely small number of publicly available RI with COVID-19. The main model as it achieves 85.2 in 3 class testing accuracy.

In Ioannis D. Apostolopoulos & Tzani A. Mpesiana [12], using pre-trained models, it was shown that from VGG19, ResNetv2 gives the best results MobileNet v2, with it is noted that the classification accuracy strongly depends on the number of samples presented in each class. The datasets utilized in this experiment are two. Firstly, a collection of 1427 X-ray images including 224 images with confirmed Covid-19 disease, 700 images with confirmed common bacterial pneumonia, and 504 images of normal conditions. Secondly, a dataset including 224 images with confirmed Covid-19 disease, 714 images with confirmed bacterial and viral pneumonia, and 504 images of normal conditions. The data was collected from the available X-ray images in public medical repositories. The best accuracy obtained is 96.78%.

In Turker Tuncer [13], special attention is paid to data preprocessing to improve classification accuracy. The used dataset represents a total of 234 healthy X-ray chest images obtained from the Kaggle site. The proposed feature generation method is called Residual Exemplar Local Binary Pattern (ResExLBP). In the feature selection phase, a novel iterative ReliefF (IRF) based feature selection is used. Decision tree (DT), linear discriminant (LD), support vector machine (SVM), k nearest neighborhood (kNN), and subspace discriminant (SD) methods are chosen as classifiers in the classification phase. Leave-one-out cross-validation (LOOCV), 10-fold cross-validation, and holdout validation are used for training and testing. In this work, the SVM classifier achieved 100.0% classification accuracy by using 10-fold cross-validation.

The purpose of this study is to develop a RI classifier for patients with viral pneumonia and COVID-19 based on the simplest NS, without using learning ready-made models of other developers and without preliminary training on third-party data, having only a small set of RI.

2-X-ray images dataset

X-ray images were obtained from the source [14], Figure1, where they were in the public domain as an archive containing 219 PNG RIs of patients with COVID-19 and 1341 RIs of patients with viral pneumonia, with a size of 1024 × 1024 pixels. The accompanying file indicates the origin of the RI: "We have developed the database of COVID-19 x-ray images from the Italian Society of Medical and Interventional Radiology (SIRM) COVID-19 DATABASE, the Novel Corona Virus 2019 Dataset developed by Joseph Paul Cohen, Paul Morrison and Lan Dao in GitHub and images extracted from 43 different publications." All initial RI at the first stage was divided into two groups according to the classification criteria and placed in the appropriate folders.

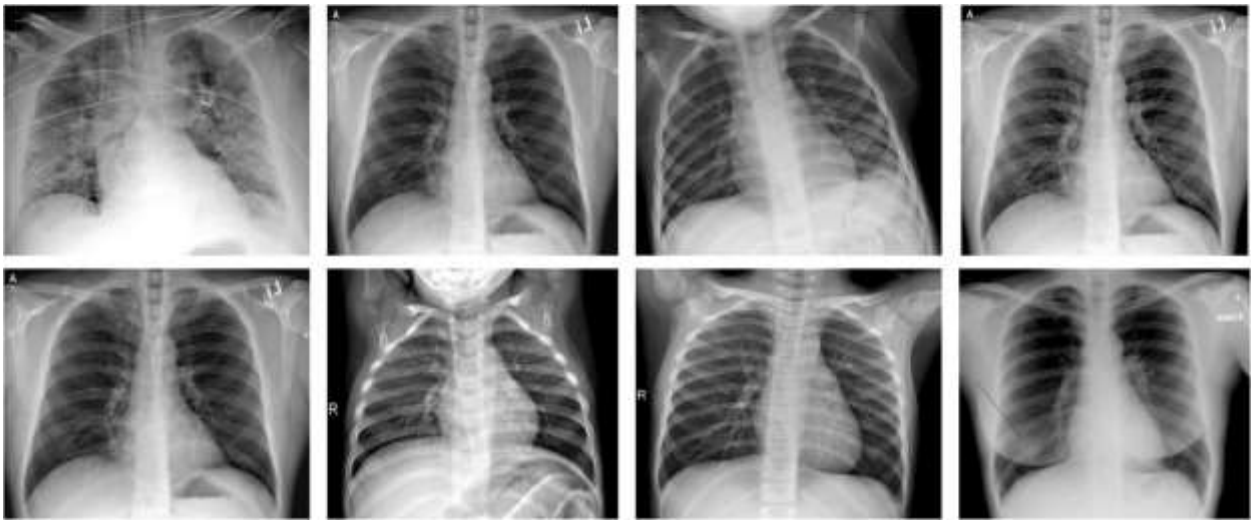

Figure 1: Samples of X-ray images for binary classification

Image pre-processing

The main stages of data preparation and the movement of information flows are shown in Figure. 2.

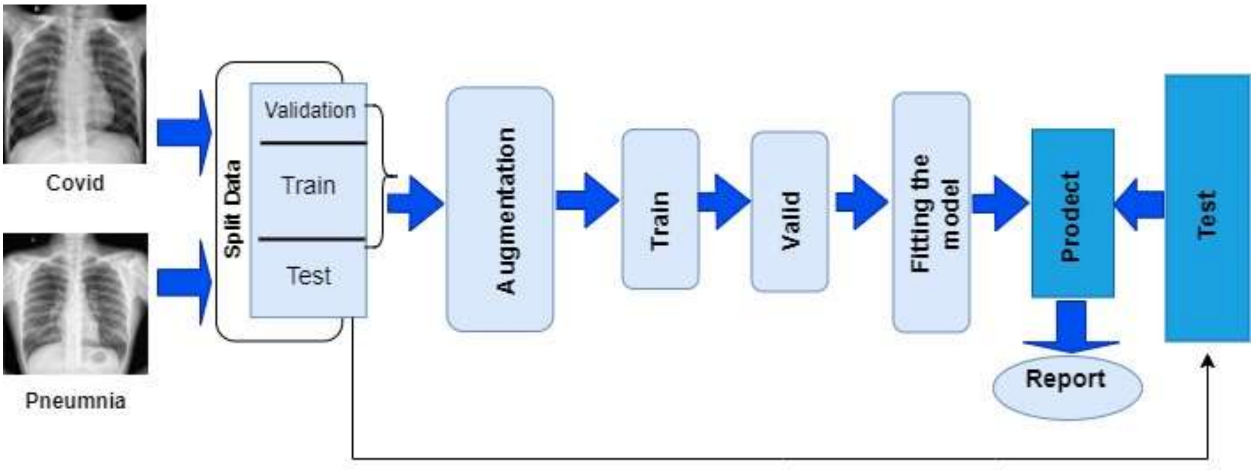

Figure 2: Scheme of image preprocessing and movement of information flows

A preliminary analysis of the RI showed that at the edges of some images there are service symbols that, in an encoded form, can carry information about the image belonging to a particular class. In order for the National Assembly to look for informational signs precisely on the images of the lungs, through the operation

The crop on all original images was removed by 10% of the peripheral part of the image. Since the image is mathematically a matrix, the cut image is obtained by copying the matrix, skipping 10% of the first and last rows, as well as skipping 10% of the left and right columns of the images. In Python, this mathematical operation is formalized as follows:

im_crop = im [102:922,102:922], where im_crop is the matrix of the cropped image.

Further visual analysis of the RR showed that they are extremely inhomogeneous in brightness. In informationally important areas of the lungs, either excessive darkening or lightening was observed, i.e. the pixel values were limited to a small range or, accordingly, there was a narrow peak in the brightness histogram. Simply stretching the histogram for the entire image can lead to unwanted changes in brightness. Therefore, for the better manifestation of information features, the operation of adaptive histogram equalization was applied to the RI using the CLAHE function of the OpenCV computer vision library. In this algorithm, images are divided into rectangular sections and for each, the pixel brightness distribution function is calculated as a function of the relative number of pixels from specific brightness values. The transformed image is obtained by changing the brightness of each pixel R j of the input image into the brightness value Sk of the corresponding pixel of the output image, calculated by the formula(1) [15]:

$$S_j = \sum_{j=0}^{k} \frac{n_j}{n} \qquad (1)$$

k– Pixel number, varies from 0 to L-one, L is the maximum number of gradations of brightness, and j changes from 0 to k; nj number of pixels with brightness Rj, n is the total number of image pixels. Then, to remove artifacts at the boundaries of the sections in the form of a checkerboard effect, bilinear interpolation is applied, i.e. calculating the weighted average of the brightness of the pixels by linear interpolation first in one diagonal direction and then in the other.

To get images of the required size, ra, according to the recommendation of the developers of the OpenCV library, the cv2.imresize() function was used with the cv2.INTER_AREA interpolation algorithm, due to using this algorithm, we eliminate the appearance of moiré-type distortions. Further, for each class (SPLIT block), samples of training (train), verification (val) and testing (test) of the NN were formed. To do this, all RI was randomly shuffled using the Python3 random.shuffle() function. In this case, the random module generates random numbers, and the shuffle() extension provides random shuffling of images.

Then, in each class, 10% of RI were also randomly selected, and they were placed in test samples (blocks TEST_COV and TEST_VPN).

We randomly divided the remaining data into two independent sets in a ratio of 90% and 10% for training (TRAIN_COV, TRAIN_VPN blocks) and validation (VAL_COV and VAL_VPN blocks). After that, all data were normalized. In order for each image to fall into the test sample, all calculations and, accordingly, random mixing were performed 50 times. To reduce overfitting and partially compensate for the imbalance, the number of images in each class for the training sample was increased to 10000 by adding transformed images (AUGM block). Transformations were changes in random order of the following parameters: rotation in the range from +15 to -15 degrees, shift to the left and right in the range from 1 to 15 pixels, change in scale and brightness in the range from 0 to 10%, rotation around the horizontal axis in the range from 1 to 180 degrees, which was provided by the ImageDataGenerator function, which is part of the Keras library. Further, by merging the corresponding samples of the two classes, complete training, verification, and testing samples were formed (TRAIN, VAL, and TEST blocks).To train the neural network (FIT block), the full training and test samples, as well as class labels of these samples (LABEL_TRAIN and LABEL_VAL blocks), are fed to its input. The learning result is the NN model.

The proposed Convolutional Neural Network Structure

To develop the RI classifier, a convolutional NN was used, for the input of which RIs are fed. The result of the work of the trained NN is a label that refers to the image of a particular class. Cross-validation and lattice search methods have shown that the best results are obtained by the NN, which comprises five successive convolutional layers (Figure 3). These layers had 32, 64, 128, 192, and 256 filters for the first and subsequent convolutional layers and the same kernel sizes (3 × 3).

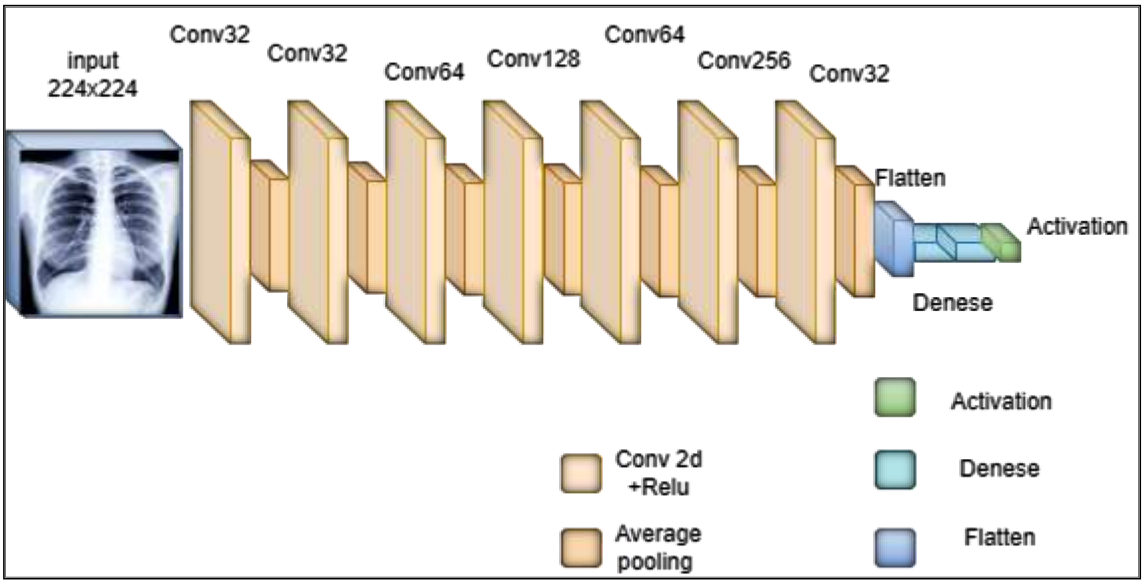

Figure 3: Structure and main parameters CNNs

When training, the Adam optimizer was used (lr = 0.001, beta_1 = 0.9, beta_2 = 0.999). In our case, we obtained the best results when categorical rather than binary cross-entropy was used as the loss function. The training was performed at 350 epochs. Regularization was used to prevent over-fitting.

When choosing the architecture and hyperparameters of the NN, we use the metric:

Accuracy = (TP + TN) / (TP + TN + FP + FN), reflecting the number of correctly classified results divided by their total number [17].

To test this model, it goes to the PREDICT block, which also receives the full test sample (TEST) and the corresponding set of class labels (LABEL_TEST block). The REPORT block shows the result of the model on test data. For each class, information is displayed on the main classification criteria: precision, recall, and F1-score. The number of RIs in each class is also shown.

Precision = TP / (TP + FP) indicates how many of the predicted positive outcomes.

Recall = TP / (TP + FN) denotes how much of the total number of actual positives was predicted to be positive.

F1-score is the harmonic mean of precision and recall and is used to calculate the balanced average result.

F1-score = 2*Recall*Precision / (Recall+Precision)

where TP, TN, FP, and FN are true positive, true negative, false positive, and false negative. [16, 17].

The classification report function of the scikit-learn machine learning library was used to generate the report. The nature of the change in the values of accuracy and loss of the function for the Train and Val samples Figure.3 shows the absence of retraining of the NN model. The quality of the classifier was evaluated on a test set using the most commonly used metrics Precision, Recall, and F1-score.

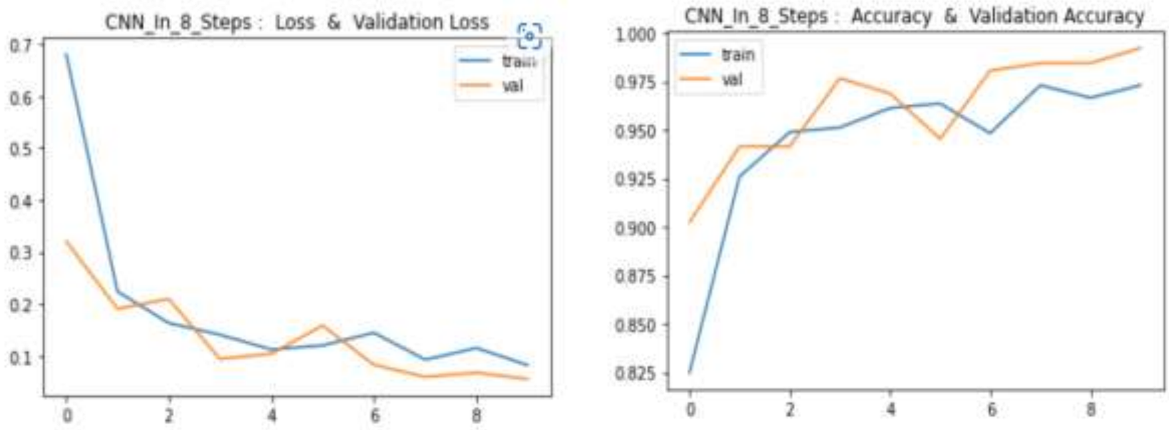

Figure 4: The accuracy and loss functions in the process of CNN training and validation

The choice of the best architecture and hyperparameters of the NN was carried out according to the maximum values of these metrics and according to their uniformity for each class.

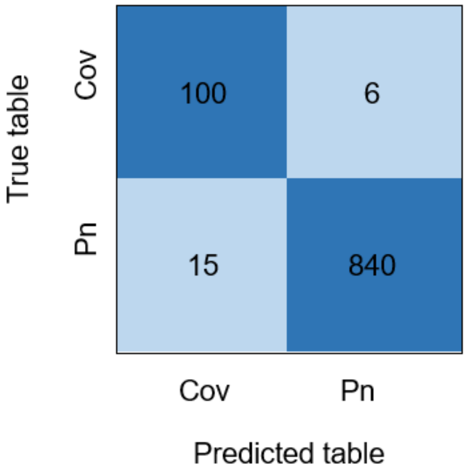

Figure 5: confusion matrix for binary classification (Covid-19)

A confusion matrix was obtained for the proposed model. When the confusion matrix includes covid-19 and pneumonia classes, the model classified COVID-19 with 100 true positives and six non-false negative classes, while the pneumonia class was classified by the model with 840 as the true case and fifteen false negatives in the total normal class.

Analysis of the results

To assess the predictive ability of the resulting classifier, images with the following sizes were selected: 768 × 768, 512 × 512, 224 × 224, 112 × 112, and 64 × 64 pixels. The best results were obtained on images with a size of 224 × 224, and the results of processing these images are given below. In Table. 1 for the test sample shows the values of the quality metrics of this classifier for patients with pneumonia and COVID-19.

Table 1: Classifier accuracy metrics

	PRECISION	RECALL	F1-SCORE	ACCURACY
Pneumonia	0.99	0.99	0.99	0.99
COVID-19	0.91	0.95	0.91	0.97

The high values of the classification criteria (> 0.91) convincingly indicates a reliable classification, i.e. the reliable difference between RI in patients with pneumonia and RI in patients with COVID-19.

The high values of precision given in Table. 1 (0.99 for pneumonia and 0.91 for COVID-19) indicates a few false positive detections of these diseases, and high recall values (0.99 for pneumonia and 0.95 for COVID-19) - testify that the classifier gives an incorrect result in only 5% of cases with COVID-19. Higher metrics for pneumonia are possible but are associated with a significantly higher number of RIs with this disease.

All indicators are given the in the Table. 1 indicates the high quality of the classification, which clearly indicates the possibility of using NN to detect cases of COVID-19 by chest X-ray.

This work's software and hardware specifications are as follows:

Software: Windows 10x64 Pro, Python 3.9, Numpy 1.16.5, OpenCV 4.2.0, Matplotlib 3.1.1, TensorFlow 2.1.0, Scikit-Learn 0.22 with coLab environment.

Hardware: Intel Core i5- 6200U CPU @ 2.30GHz 2.40, Lenovo X1 Carbon 512GB SSD, NVIDIA GeForce GTX1080.

Comparison with Previous Studies

The following Table 3. is a comparison of the proposed system with the previous works, presented in summary form for comparison in terms of accuracy.

Table 3: comparison with previous studies

RF.	Dataset	Methods	Accuracy
[9]	COVIDx X-ray	COVID-Net	93.3%
[10]	Two X-ray Datasets	Dark-Net	98.08%
[11]	X-ray collection from different sites	Alexnet, Googlenet, Resnet18	85.5%
[12]	Two X-ray Datasets	VGG19, ResNetv2	96.78%
[13]	X-ray chest images from the Kaggle site	SVM classifier	100%
The proposed model	CoronaHack X-ray From Kaggle site	NN	99% for Pneumonia 97% for COVID-19

Conclusions

An assessment of the predictive ability of the presented NN showed that with optimal sample preparation and choice of architecture, as well as the correct selection of hyperparameters, it is possible to create a RI classifier for patients with viral pneumonia and COVID-19 without involving ready-made complex models and preliminary training on third-party data. An increase in the number of RIs for patients with COVID-19 will lead to an increase in the predictive ability of the classifier and the creation of sensitive and reliable express diagnostic methods, which will reduce the burden on the healthcare system. Recognizing the urgent need to create high-precision and practical solutions for the rapid detection of COVID-19, the authors are ready to provide all the information obtained in this work, and the source code to the National Assembly, the scientific community, and interested developers.

References

1. Wu F, Zhao S, Yu B, et al." A new Coronavirus associated with human respiratory disease in China". Nature 2020; 579(7798): 265.
2. World Health Organization. Pneumonia of unknown cause -China. Source:- https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/-
3. Veselova EI, Russkikh AE, Kaminskiy GD, Lovacheva OV,Samoylova AG, Vasilyeva IA." Novel coronavirus infection." Tuberculosis and Lung Diseases, 2020; 98(4): 6-14. DOI: 10.21292/2075-1230-2020-98-4-6-14.
4. Pashina TA, Gaidel AV, Zelter PM, Kapishnikov AV, Nikonorov AV." Automatic highlighting of the region of interest in computed tomography images of the lungs".Computer Optics 2020; 44(1): 74-81. DOI: 10.18287/2412-6179-CO-659.
5. Li L, et al. "Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the diagnostic accuracy". Radiology 2020; 296(2): E65-E71. DOI: 10.1148/radiol.2020200905.
6. Nath M, Choudhury C." Automatic detection of pneumonia from chest X-Rays using deep learning". In Book: Bhattacharjee A, Borgohain S, Soni B, Verma G, Gao XZ, eds. Machine learning, image processing, network security, and data sciences. Singapore: Springer; 2020:175-182. DOI: 10.1007/978-981-15-6315-7_14.
7. Okeke S, et al. "An efficient deep learning approach to pneumonia classification in healthcare". J Healthc Eng 2019; 2019: 4180949. DOI: 10.1155/2019/4180949.
8. Swapnarekha H, et al. "Role of intelligent computing in COVID-19 prognosis: A state-of-the-art review". Chaos Solitons Fractals 2020; 138:109947. DOI: 10.1016/j.chaos.2020.109947
9. Wang L, Wong A." COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-Ray images". 2020. Source:- https://arxiv.org/abs/2003.09871.
10. Ozturk T, et al." Automated detection of COVID-19 cases using deep neural networks with X-ray images". Comput Biol Med 2020; 121: 103792. DOI: 10.1016/ j.compbiomed.2020.103792.

11. Loey M, Smarandache F, Khalifa NEM. "Within the lack of chest COVID-19 X-Ray dataset: A novel detection model based on GAN and deep transfer learning". Symmetry 2020; 12:651. DOI: 10.3390/sym12040651.

12. Apostolopoulos ID, Mpesiana TA." COVID-19: automatic detection from X-Ray images utilizing transfer learning with convolutional neural networks". Phys Eng Sci Med 2020; 43(2): 635-640. DOI: 10.1007/s13246-020-00865-4.

13. Tuncer T, Dogan S, Ozyurt F." An automated residual exemplar local binary pattern and iterative reliefF based COVID-19 detection method using chest X-Ray image". Chemom Intelli Lab Syst 2020; 203: 104054. DOI: 10.1016/ j.chemolab.2020.104054.

14. CoronaHack –"Chest X-Ray-Dataset. Classify the X-Ray image which is having Corona." source:-https://www.kaggle.com/datasets/praveengovi/coronahack-chest-xraydataset

15. Gonzalez RC, Woods RE. "Digital image processing. 3rd ed. Pearson Education "Inc.; 2008.

16. Chollet F." Deep learning with Python. New York: Manning "Publications; 2017. ISBN 9781617294433 384 pages.

17. Müller AC, Guido S. "Introduction to machine learning with Python: A guide for data scientists". O'Reilly Media; 2016.