Research article
DOI: https://doi.org/10.18721/JCSTCS.17102
UDC 004

A STUDY OF HYPERPARAMETERS EFFECT ON CNN PERFORMANCE FOR CHEST X-RAY BASED COVID-19 DETECTION

A. Saaudi[1], *R. Mansoor*[2] ✉

[1] Al Muthanna University College of Engineering, Iraq;
[2] Electronics and Communication Engineering, Al Muthanna University, Iraq

✉ riyadhdmu@mu.edu.iq

Abstract. COVID-19 disease has been spreading around the world for the last four years. Different generations of corona viruses appeared: Alpha-, Beta-, Gamma-, and Delta variants. Thus, COVID-19 changed human lifestyle and affected economic development of many countries. According to clinical studies, most of the positive cases of COVID-19 patients suffer from lung infection. For this, a lot of efforts were aimed at developing fast and accurate detection methods. Thanks to the Deep Learning techniques that facilitate the process of identifying COVID-19 based on the chest images of the patients. X-ray and CT scan images are commonly used to evaluate corona virus lung infection. X-ray images are adopted by many researchers since they place less financial burden on the patient. In this work, we used chest X-ray images to develop eight CNN-based detection models. Three sets of images, i.e., COVID-19, pneumonia and normal cases were used for the training and testing. The performance of each model was optimized based on different hyperparameters to come up with the best results in terms of high detection accuracy, recall, precision and f1 score. These hyperparameters include Number of CNN layers, filters, dense layers, and number of nodes per dense layer. Our findings show that increasing both the CNN layers and number of filters result in high precision and f1 score of the positive samples, while increasing the number of dense layers leads to low precision recall and f1 score.

Keywords: Artificial intelligence, COVID-19, CNN, Deep learning, Chest X-ray

Citation: Saaudi A., Mansoor R. A study of hyperparameters effect on CNN performance for chest X-ray based COVID-19 detection. Computing, Telecommunications and Control, 2024, Vol. 17, No. 1, Pp. 20–32. DOI: 10.18721/JCSTCS.17102

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.17102
УДК 004

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ГИПЕРПАРАМЕТРОВ НА ЭФФЕКТИВНОСТЬ CNN ДЛЯ ВЫЯВЛЕНИЯ COVID-19 НА ОСНОВЕ РЕНТГЕНОГРАФИИ ГРУДНОЙ КЛЕТКИ

А. Сауди[1], *Р. Мансур*[2] ✉

[1] Инженерный колледж университета Аль-Мутанна, Ирак;
[2] Университет Аль-Мутанна, Самава, Ирак

✉ riyadhdmu@mu.edu.iq

Аннотация. В работе предложен метод для разработки быстрого и точного выявления развития осложнений после Covid-19, посредством метода глубокого обучения, который облегчает процесс идентификации Covid-19 на основе изображений грудной клетки пациентов. Рентгеновские снимки и компьютерная томография обычно используются для оценки легочной инфекции, вызванной коронавирусом. Рентгеновские снимки используются многими исследователями, поскольку они несут меньшую финансовую нагрузку на пациента. В этой работе мы использовали рентгеновские снимки грудной клетки для разработки восьми моделей обнаружения на основе CNN. Для обучения и тестирования используются три набора изображений: COVID-19, пневмония и обычные случаи. Производительность каждой модели оптимизирована на основе различных гиперпараметров для достижения наилучших результатов с точки зрения высокой точности обнаружения, отзыва, прецизионности и оценки f1. Эти гиперпараметры включают количество слоев CNN, фильтров, плотных слоев и количество узлов на плотный слой. Наши результаты показывают, что увеличение как количества слоев CNN, так и количества фильтров приводит к высокой точности и показателю f1 положительных образцов. В то время как увеличение количества плотных слоев приводит к низкой точности воспроизведения и оценке f1.

Ключевые слова: Искусственный интеллект, COVID-19, CNN, Глубокое обучение, рентген грудной клетки

Для цитирования: Saaudi A., Mansoor R. A study of hyperparameters effect on CNN performance for chest X-ray based COVID-19 detection // Computing, Telecommunications and Control. 2024. Т. 17, № 1. С. 20–32. DOI: 10.18721/JCSTCS.17102

Introduction

Artificial intelligence (AI) has witnessed remarkable development in the past few years, and has become a spearhead in facing the challenges we face on the planet; the latest of which is the Coronavirus, a viral-based disease, which has become the main concern of the world [1]. This disease affects people's lives in the East and West countries. Governments tried hard to defeat this virus. Economically, billions of dollars have been spent to develop a cure or to develop a diagnosis system. However, identifying COVID-19 in its early stages is not an easy task. Scientists employ machine learning methodologies to deal with various healthcare challenges. Machine learning is one of the capabilities of artificial intelligence, and this ability enables the identification of complex patterns in large sets of data, whether text or images [2]. And if used correctly, AI can surpass humans, not only in speed but also in accuracy when identifying patterns in data that humans might ignore. Since AI requires large amounts of data, the challenge with the coronavirus is to provide reliable and quality data. Fortunately, there are two available solutions to deal with the lack of data samples problem. First, medical data is now available as some countries are starting to understand the problem and work to tackle the emerging virus. Second, Machine learning provides augmentation algorithms that generate data samples synthetically.

Many treatments have been developed for COVID-19 disease [3]. However, there is no reliable treatment for the virus mainly with advanced lung infections. Most efforts focus on developing lung scanning-based methods to identify COVID-19 cases. For instance, X-ray and CT scans are the most common lung imaging techniques used today to diagnose COVID-19 infections [4]. In comparison to CT, people can afford X-ray imaging because of its low cost. In addition, lung X-rays are well-known and used commonly in most countries. Traditionally, specialists are responsible for evaluating lung X-ray samples. This process could be affected by human vulnerabilities such as overwhelming, tiredness, social effects, and personal emotions.

In this work, we present a comprehensive study on using deep learning techniques with chest X-ray data samples. Mainly convolution neural network CNN-based systems are developed to classify COVID-19 cases. For this several CNN- based structures are presented and evaluated to reach the optimum performance. Four metrics are adopted to evaluate the model performance i.e. precession, recall, accuracy, and f1 score. Moreover, this work relies on studying several deep learning model components. These components include the number of CNN layers, number of filters per layer, number of dense layers, and number of nodes per dense layer. A baseline model is built first by using one CNN layer of 32 filters and one dense layer of 32 nodes. The obtained results of this model show low performance in terms of the four aforementioned metrics. Then, seven different model component combinations are tested to achieve the best hyperparameter configuration that leads to high performance compared to the baseline model. In addition to the proposed model, transferred-learning models are adopted to develop three COVID-19 detection models: Resnet50, Exception, and VGG16.

Dataset and Proposed Model

Dataset

The proposed model relies on the chest x-ray data sets, e.g., COVID-19, Normal, and Pneumonia. These data sets are used to train and test the models. We use the chest x-ray data set from Kaggle [5] which provides three types of samples COVID-19, Normal, and Pneumonia. A subset of 780 samples is divided into training, testing, and validation groups. For this, the training data set has 600 samples evenly distributed among three classes COVID-19, Normal, and Pneumonia. The testing data set consists of 120 data samples, 40 samples from each class. Finally, the validation data set involves 60 samples to validate the training process.

The reason for selecting this number of data sets is to overcome the unbalanced issue that comes from the fact that the Kaggle data set in its actual distribution results in overfitting/underfitting of the model [6]. The size of the selected subset is not large enough compared to the actual size of the original data set from Kaggle [4]. To overcome the unbalanced issue, we partition from the same distribution. So, the model learning is based on the same number of samples from each class. Moreover, we use several augmentation techniques to avoid overfitting [7]. These techniques include rotation, zoom, and sharing of images.

Base-line Model

In this section, we use a convolutional neural network to develop a COVID-19 detection system [8]. The proposed model consists of one convolutional layer, one max pooling layer, one dense layer, one flattened layer, and one input and output layer as shown in Fig. 1.

The input layer takes a chest X-ray sample of size 300X300. Then, convolutional operations are applied using 32 filters of size 3X3. Each CNN filter iterates over each input sample, with one step stride, applying dot product to extract a feature map. Thus, the result of this layer is 32 feature maps, each of 298X298 dimensions. The dot product operation is performed using Relu as an activation function. The Relu function evaluates the summation of the weighted input signals to eliminate the weak signals and pass the strong ones [9]. Max-pooling layer is added to the convolution layer to reduce the dimensions and collect the most important features [10]. The Max-pooling layer iterates over each feature map with a window of size 2x2, four values. The window takes the maximum value under window reign, which in

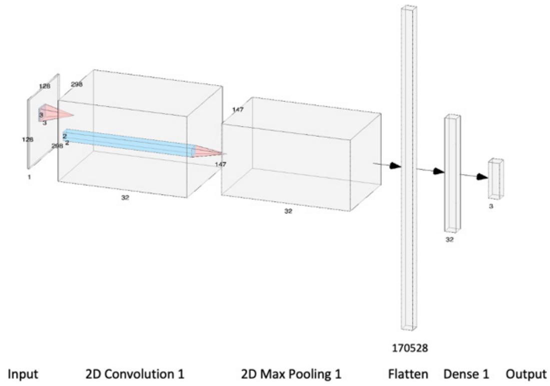

Fig. 1. The structure of the Base-line model

our work has four values. The result of the first convolutional and max-pooling layers is 32 feature maps of size 149x149 each.

To feed the 32 feature maps to the next layer, a flattened layer is added to form a vector of 170528 nodes. Then, one dense layer of 32 nodes in length is added. At the end, a three-node output layer is added to evaluate the input sample to one of three classes COVID-19, Normal, or Pneumonia. A sigmoid activation faction is used with the output layer to assess the summation of the weighted signals [11].

Hyperparameters and evaluation metrics

We consider four hyperparameters for our study: the number of filters, CNN layers, Dense layer, and the number of nodes of dense layers. For the baseline model, the hyperparameters are set as follows, one CNN layer of 32 filters, a Max-pooling layer, one dense layer of 32 nodes, one input layer receiving data example of size 300x300, and one output layer of 3 nodes.

The performance of the baseline model is evaluated using four metrics which are accuracy, precession, recall, and f1 score [14, 15]. The accuracy metric evaluates the overall performance of the model by dividing the number of accurately predicted samples by the total number of testing samples. It indicates how close the predicted labels are to their actual labels, see equation 1. Moreover, Precision matric refers to how predicted samples are close to each other. It illustrates the relationship between the predicted sample and the class it belongs to, see equation 2. Precision is not related to accuracy. In other words, the predicted labels could be very precise but not accurate, or they could be accurate but not precise. In addition, Recall matric examines the model by showing the relation between the true positive and the false negative samples in such a way that completes the Precision task. The recall is used in extreme cases such as cancer, when any false negative leads to high-risk consequences. Recall is calculated by dividing the number of true positive samples by the summation number of true positive and false negative samples, see equation 3 [12]. Finally, to test the balance between Precision and Recall, the F1 score is calculated, see equation 4 [13].

$$Accurasy = \frac{True\ Positives + True\ Negatives}{Total\ Number\ of\ Testing\ Samples}; \tag{1}$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}; \tag{2}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}; \tag{3}$$

$$F1\ Score = 2x\frac{Precision\ x\ Recall}{Precision + Recall}. \tag{4}$$

Result and discussion of the baseline model

This section presents the analysis of the results. The results of the baseline model are listed in Table 1. The behavior is evaluated with the help of using four metrics: accuracy, Precision, Recall, and F1 score. The proposed model is tested with 120 chest X-ray samples. These samples belong to three classes COVID-19, Normal, and Pneumonia. The baseline model shows a low average accuracy of 0.33 approximately. This model misidentifies all of the normal and Pneumonia samples. Whereas 27 positive cases are recognized falsely, see Table 1.

Table 1

Evaluation metrics of base-line Model

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		0.33	1	0.5
Normal		0	0	0
Pneumonia		0	0	0
Average/Total	0.33	0.11	0.33	0.17

In the next section, the effect of the hyperparameters will be examined with the aim of obtaining the best performance in terms of evaluation metrics.

Optimization Plan

This section examines seven deep-learning models to detect COVID-19. The proposed models are developed with four hyperparameters: number of CNN layers, number of filters, number of dense layers, and number of nodes per dense layer. An analysis study is conducted to evaluate each model regarding the baseline model. For this, four metrics are used: Accuracy, Precision, Recall, and F1 score, as follows:

Model I

In this model, two CNN layers are used compared to the baseline model. The number of filters is set to 32 in each layer. While the dense layer is kept to be 32 nodes as shown in Fig. 2. Additionally, Table 2 summarizes the evaluation metrics of this model.

Table 2

Evaluation metrics of Model I

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		0.97	0.95	0.96
Normal		0.91	1	0.95
Pneumonia		1	0.93	0.96
Average/Total	0.95	0.95	0.95	0.95

Model II

In this model, the number of dense layers is increased by adding another layer of 64 nodes to model I. The sequence of dense layers becomes 64, and 32 nodes respectively instead of 32 nodes in the previous model, as shown in Fig. 3. The performance of Model II is evaluated using four metrics, see Table 3.

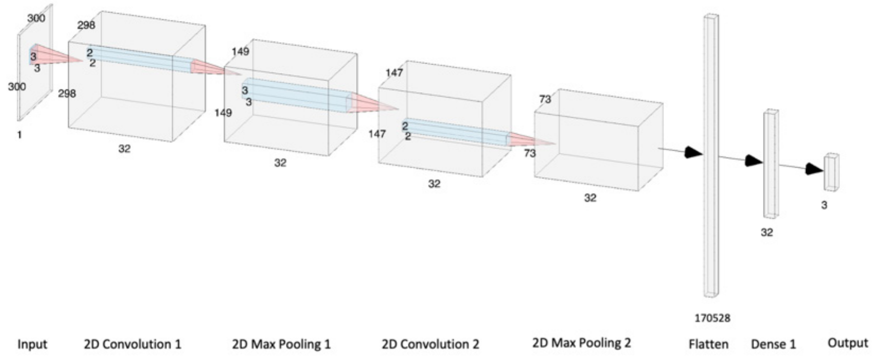

Fig. 2. The structure of Model I

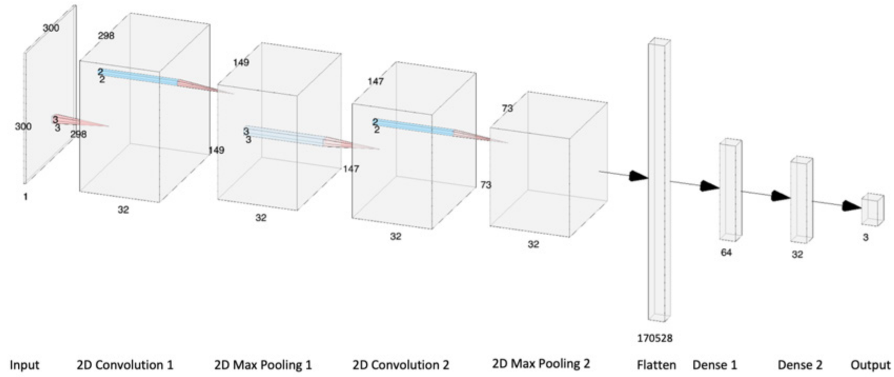

Fig. 3. The structure of Model I

Table 3

Evaluation metrics of Model II

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		0.97	0.95	0.96
Normal		0.95	0.95	0.95
Pneumonia		0.95	0.97	0.96
Average/Total	0.96	0.96	0.96	0.96

Model III

Here, the effect of the number of nodes per dense layer is examined. So, two dense layers each of 64 nodes, compared to mode II, is adopted as shown in Fig. 4. Along with the models' structure, table 4 illustrates the evaluation metrics of the model.

Model IV

Now, the effect of the filter number is tested. The structure of Model II is modified by changing the number of filters of CNN layers to 64, and 32 with the same dense layers, as explained in Fig. 5. To analyze the models' performance, Table 5 shows the evaluation metrics.

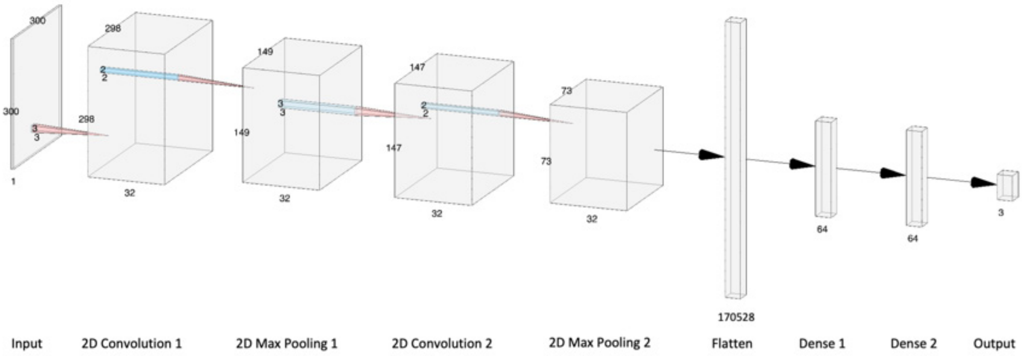

Fig. 4. The structure of Model III

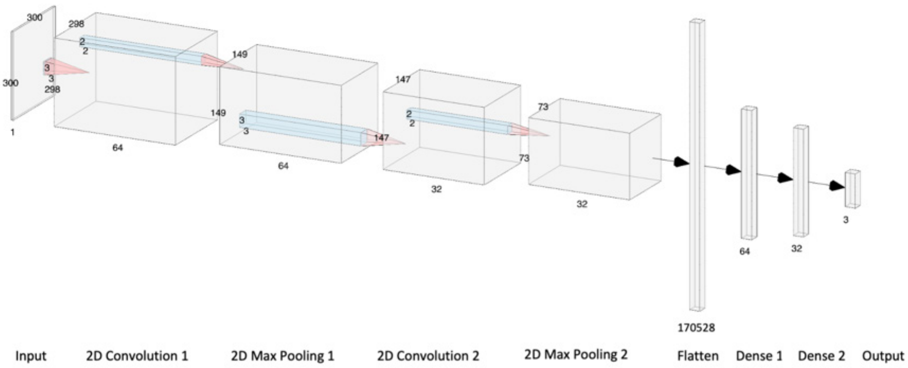

Fig. 5. The structure of Model IV

Table 4

Evaluation metrics of Model III

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		1	0.95	0.97
Normal		0.87	1	0.93
Pneumonia		1	0.9	0.95
Average/Total	0.95	0.96	0.95	0.95

Model V

Model V is composed of three CNN layers (32, 32, 32) with two dense layers (64, 32) as shown in Fig. 6. The effect of adding a third CNN layer of 32 filters is studied compared to model II. Moreover, Table 6 presents the evaluation metrics to study the performance of the optimized model.

Model VI

Model VI consists of three CNN layers with two dense layers (64, 32) as shown in Fig. 7. The effect of tuning the number of filters to be (64, 32, 32), compared to model V, is examined. The tuning of the filters' number improves the model performance slightly as shown in Table 7.

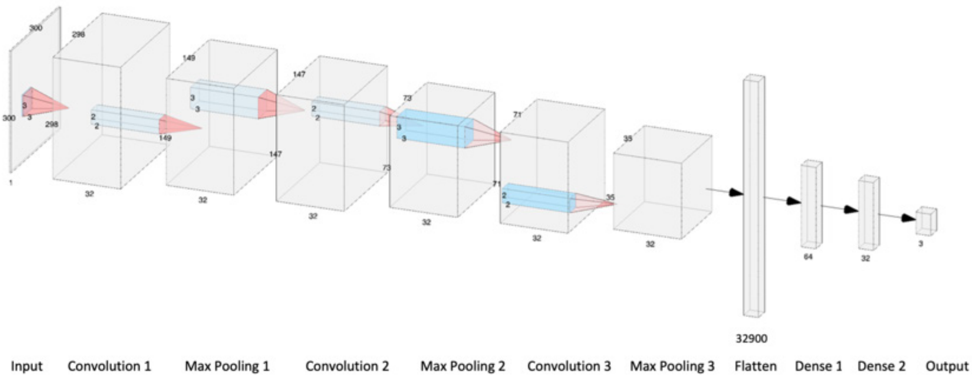

Fig. 6. The structure of Model V

Table 5

Evaluation metrics of Model IV

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		1	0.95	0.97
Normal		0.95	0.95	0.95
Pneumonia		0.95	1	0.98
Average/Total	0.97	0.97	0.97	0.97

Table 6

Evaluation metrics of Model V

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		0.97	0.95	0.96
Normal		0.95	0.95	0.95
Pneumonia		0.95	0.97	0.96
Average/Total	0.96	0.96	0.96	0.96

Table 7

Evaluation metrics of Model VI

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		1	0.95	0.97
Normal		0.95	0.95	0.95
Pneumonia		0.95	1	0.98
Average/Total	0.97	0.97	0.97	0.97

Model VII

In Fig. 8, Extra tuning for the number of filters is performed in this model. The filters per CNN layers are set to be (64, 32, 16) compared to model V. In addition to the model structure, Table 8 shows the models' performance metrics.

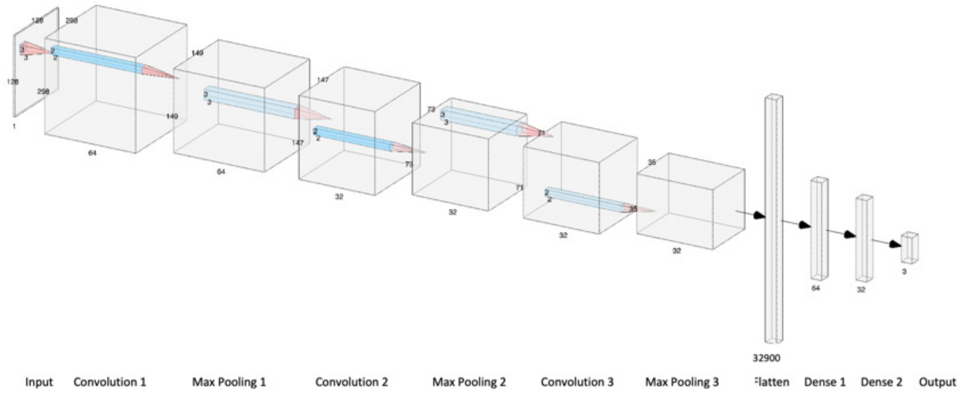

Fig. 7. The structure of Model VI

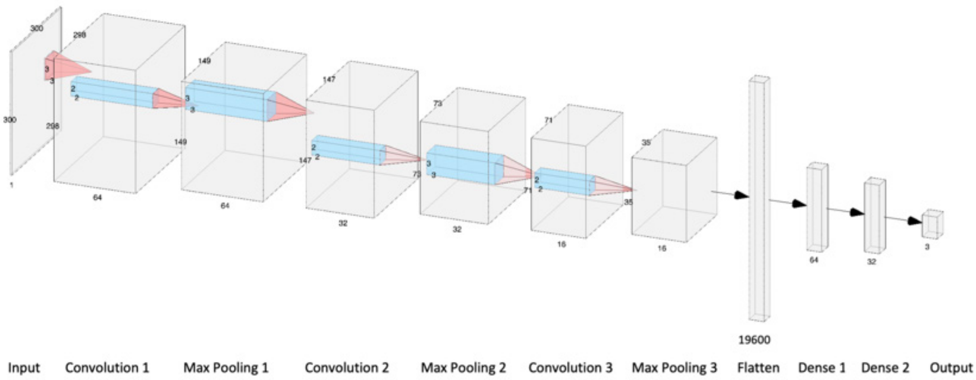

Fig. 8. The structure of Model VII

Table 8

Evaluation metrics of Model VII

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		1	0.95	0.97
Normal		0.95	1	0.98
Pneumonia		1	1	1
Average/Total	0.98	0.98	0.98	0.98

Transfer-Learning Models

This section presents the use of three of the well-known transferred learning models (Resnet50, Xception, and VGG16) models. In this work, the trained weights of these three models are adopted and customized to fit the purpose of the COVID-19 identification application.

The weights of the Resnet50, Xception, and VGG16 are adopted without the input and output layers. The latter are customized by setting the input layer dimensions as the one used with the proposed models (300x300) pixels. Moreover, two dense layers are added after the flattened layer with the sizes of 32 nodes, and 16 nodes respectively. Finally, an output layer with the size of three nodes is added to classify an input sample into one of three classes COVID-19, Pneumonia, or Normal class.

The following subsections illustrate the results of the experiments, and the evaluation of the findings using accuracy, precision, recall, and f1-score metrices, in addition to the use of the confusion matrices.

The Results and The Evaluation Metrics of transfer-learning Models

Resnet50

The resnet50 presents a good performance in detecting COVID-19 cases, see Table 9 and Table 10. The precession, recall, and f1-score have the same score value of 0.97, where the model identifies 39 of COVID-19 cases correctly. However, the behavior is different from the rest of the classes. The precision with the pneumonia class is 0.8 since the model considers 10 different classes as pneumonia. The recall with the normal class is 0.75 because ten normal samples are evaluated as not normal samples.

Table 9

Evaluation metrics of Resnet50 Model

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		0.97	0.97	0.97
Pneumonia		0.8	1	0.89
Normal		1	0.75	0.86
Average/Total	0.91	0.93	0.91	0.91

Table 10

Confusion matrices of the transfer-learning models with (C: COVID-19, P: Pneumonia, and N: Normal) cases

		Resnet50				Xception				VGG16			
		Predicted				Predicted				Predicted			
			C	**P**	**N**		**C**	**P**	**N**		**C**	**P**	**N**
Testing	**Actual**	**C**	39	1	0	**C**	38	0	2	**C**	40	0	0
		P	0	40	0	**P**	6	0	34	**P**	0	40	0
		N	1	9	30	**N**	5	0	35	**N**	0	5	35

Xception

The Xception model illustrates the worst performance of detecting the input samples, see Table 10 and Table 11. The model does not detect any Pneumonia cases and miss identifies two COVID-19 cases and five normal cases.

Table 11

Evaluation metrics of Xception Model

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		0.78	0.95	0.85
Pneumonia		0	0	0
Normal		0.49	0.88	0.63
Average/Total	0.68	0.42	0.61	0.49

VGG16

This model presents the best performance of identifying the input data samples, see Table 11 and Table 12. The model identifies all COVID-19, and pneumonia classes correctly, and misidentifies five samples of normal cases.

Table 12

Evaluation metrics of VGG16 Model

Metrics	Accuracy	Precision	Recall	F1 score
COVID-19		1	1	1
Pneumonia		0.89	1	0.94
Normal		1	0.88	0.93
Average/Total	0.96	0.96	0.96	0.96

Discussion

This study presents an optimization map to find the best hyperparameter configurations for a CNN-based detection system. Moreover, the work includes the use of transfer-learning models to develop COVID-19 identification systems using pre-trained weights. The work started with a baseline model of one CNN and one dense layer, section *2. Base-line model*. The model is trained for 100 epochs with samples of three chest x-ray data sets, COVID-19, Pneumonia, and Normal sets. The evaluation metrics show low performance in recognizing positive cases of COVID-19. For instance, the average weights of evaluation metrics obtained are as follows: 33% accuracy, 0.11 precision, 0.33 recall, and 0.17 f1 score. An optimization plan is conducted with the aim of getting better performance in terms of Identification accuracy. For this, seven models are developed. The effect of four hyperparameters (CNN layers, Number of Filters, Dense layers, and the number of nodes per dense layer) on model performance is considered. The results show that the increasing number of CNN layers presents a major effect on model performance. For example, the average weight accuracy of Model I, two CNN layers, is 0.95 compared to the baseline model which is 0.33. Moreover, tuning of filter numbers provides an extra enhancement for system identification ability. For instance, increasing the number of filters of CNN layers in model IV improves the performance compared to model II. In addition to that, adding an extra dense layer helps in identifying the right cases as well, see Model I and Model II. Table 9 summarizes the structure and accuracy performance of each model. Regarding the transfer-learning models, VGG16 present the best performance in term of accuracy, precision, recall, and f1-score.

Conclusion

In this work, we used chest X-ray images to develop eight CNN-based detection models. Three sets of images, i.e. COVID-19, pneumonia, and normal cases were used for the training and testing. The performance of each model was optimized based on different hyperparameters to come up with the best results in terms of high detection accuracy, recall, precession, and f1 score. These hyperparameters included the number of CNN layers, filters, dense layers, and the number of nodes per dense layer. A baseline model of one CNN and one dense layer was developed first. The number of filters and nodes were selected to be 32, 32 respectively. The result shows a low level of accuracy (33 %). However, we ran optimization in different scenarios. First, the effect of increasing the number of CNN layers was examined by adding another CNN layer of 32 filters. The accuracy was highly improved compared to the baseline model. Then, the optimization process was expanded to include different combinations of CNN layers and the number of filters per layer. Moreover, the number of dense layers and nodes per dense layer was also tested to examine their effect on system performance. This work concluded with a

Table 13

Evaluation Metrics of the Proposed Models

ID	Model	Description	Accuracy
0	Base Model	One CNN 32 and one dense 32	0.33
1	Model I	Two CNN 32, 32 and one dense 32	0.95
2	Model II	Two CNN 32,32 and two dense 64 and 32	0.96
3	Model III	Two CNN layers 32, 32 and two dense 64, 64	0.95
4	Model IV	Two CNN 64,32, and two dense 64 and 32	0.97
5	Model V	Three CNN 32, 32, 32, and two dense 64, 32	0.96
6	Model VI	Three CNN 64, 32, 32, and two dense 64, 32	0.97
7	Model VII	Three CNN 64, 32, 16 and two dense 64, 32	0.98

model design of three CNN layers of (64, 32, 16) filters and two dense layers of (64, 32) nodes that show the highest accuracy score of 98%. The adopted transfer-learning models show irregular performance in terms of evaluation metrics, Exception model presented the worst behavior while VGG16 presented the best performance.

REFERENCES

1. **Albahri O.S., Zaidan A.A., Albahri A.S., Zaidan B.B., Abdulkareem K.H., Al-Qaysi Z.T., Rashid N.A.** Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: Taxonomy analysis, challenges, future solutions and methodological aspects. Journal of infection and public health, 13 (10) (2020) 1381–1396.

2. **Lalmuanawma Samuel, Jamal Hussain, Lalrinfela Chhakchhuak.** Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. Chaos, Solitons & Fractals, 139 (2020) 110059.

3. **Stasi Cristina, Silvia Fallani, Fabio Voller, Caterina Silvestri.** Treatment for COVID-19: An overview. European journal of pharmacology, 889 (2020) 173644.

4. **Saaudi A., Mansoor R., Abed A.K.** Clustering and Visualizing of Chest X-ray Images for Covid-19 Detection, 2021 2nd Information Technology To Enhance e-learning and Other Application (IT-ELA), Baghdad, Iraq, 2021, pp. 35–39. DOI: 10.1109/IT-ELA52201.2021.9773539

5. **Prashant Patel.** Chest X-ray (Covid-19 & Pneumonia), Available: https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia, (Accessed 4 June 2021).

6. **Jabbar H., Rafiqul Zaman Khan.** Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study). Computer Science, Communication and Instrumentation Devices, 70 (2015) 163–172.

7. **Shorten Connor, Taghi M.** Khoshgoftaar, A survey on image data augmentation for deep learning. Journal of big data, 6.1 (2019) 1–48.

8. **Bassi Pedro RAS, Romis Attux.** A deep convolutional neural network for COVID-19 detection using chest X-rays. Research on Biomedical Engineering, (2021) 1–10.

9. **Eckle Konstantin, Johannes Schmidt-Hieber.** A comparison of deep networks with ReLU activation function and linear spline-type methods. Neural Networks, 110 (2019) 232–242.

10. **Nagi J., Ducatelle F., Di Caro G.A., Cireşan D., Meier U., Giusti A., Gambardella L.M.** Max-pooling convolutional neural networks for vision-based hand gesture recognition. 2011, November. In 2011 IEEE international conference on signal and image processing applications (ICSIPA). 2011, Pp. 342–347.

11. **Wanto A., Windarto A.P., Hartama D., Parlina I.** Use of binary sigmoid function and linear identity in artificial neural networks for forecasting population density. IJISTECH (International Journal of Information System and Technology), 1 (1) (2017) 43–54.

12. **Qi Q., Luo Y., Xu Z., Ji S., Yang T.** Stochastic optimization of areas under precision-recall curves with provable convergence. Advances in Neural Information Processing Systems, 34 (2021) 1752–1765.

13. **Goutte Cyril, Eric Gaussier.** A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. Advances in Information Retrieval: 27th European Conference on IR Research, ECIR 2005, Santiago de Compostela, Spain, March 21–23, 2005. Proceedings 27. Springer Berlin Heidelberg, 2005.

14. **Saaudi Ahmed.** Insider's Misuse Detection: From Hidden Markov Model to Deep Learning. Diss. University of South Carolina, 2019.

15. **Hossin Mohammad, Md Nasir Sulaiman.** A review on evaluation metrics for data classification evaluations. International journal of data mining & knowledge management process, 5.2 (2015) 1. DOI: 10.5121/ijdkp.2015.5201

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Ahmed Saaudi
Ахмед Сауди
E-mail: ahmed.saauid@mu.edu.iq

Riyadh Mansoor
Рияд Мансур
E-mail: riyadhdmu@mu.edu.iq
ORCID: https://orcid.org/0000-0002-6542-0087

Submitted: 27.12.2023; Approved: 30.03.2024; Accepted: 12.04.2024.
Поступила: 27.12.2023; Одобрена: 30.03.2024; Принята: 12.04.2024.