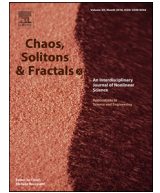




Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Comparison of deep learning approaches to predict COVID-19 infection

Talha Burak Alakus^{a,*}, Ibrahim Turkoglu^b^a Kırklareli University, Engineering Faculty, Department of Software Engineering, Kırklareli, 39000, Turkey^b Firat University, Technology Faculty, Department of Software Engineering, Elazığ, 23119, Turkey

ARTICLE INFO

Article history:

Received 22 June 2020

Accepted 10 July 2020

Available online 11 July 2020

Keywords:

SARS-CoV2

COVID-19

Coronavirus

Deep learning

Artificial intelligence

ABSTRACT

The SARS-CoV2 virus, which causes COVID-19 (coronavirus disease) has become a pandemic and has expanded all over the world. Because of increasing number of cases day by day, it takes time to interpret the laboratory findings thus the limitations in terms of both treatment and findings are emerged. Due to such limitations, the need for clinical decisions making system with predictive algorithms has arisen. Predictive algorithms could potentially ease the strain on healthcare systems by identifying the diseases. In this study, we perform clinical predictive models that estimate, using deep learning and laboratory data, which patients are likely to receive a COVID-19 disease. To evaluate the predictive performance of our models, precision, F1-score, recall, AUC, and accuracy scores calculated. Models were tested with 18 laboratory findings from 600 patients and validated with 10 fold cross-validation and train-test split approaches. The experimental results indicate that our predictive models identify patients that have COVID-19 disease at an accuracy of 86.66%, F1-score of 91.89%, precision of 86.75%, recall of 99.42%, and AUC of 62.50%. It is observed that predictive models trained on laboratory findings could be used to predict COVID-19 infection, and can be helpful for medical experts to prioritize the resources correctly. Our models (available at (<https://github.com/burakalakuss/COVID-19-Clinical>)) can be employed to assists medical experts in validating their initial laboratory findings, and can also be used for clinical prediction studies.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

On 31 December 2019, the virus SARS-CoV2, which causes coronavirus diseases (COVID-19) was detected in Wuhan, China and since December 2019, it has spread all over the world [1]. World Health Organization (WHO) declared the COVID-19 outbreak is now pandemic, it will be essential provide tools, mechanisms, and resources to quickly identify those at most risk of infirmity, and mortality. COVID-19 affects various people in different ways. Yet, over 80% infected people develop mild to moderate illness and recover without hospitalization [2,3]. Most common symptoms are fever, dry cough, and tiredness and in general, these symptoms begin as mild in all patients. However, severe symptoms such as chest pain or pressure, loss of speech or movement, and shortness of breath may be seen in a minority of patients [4,5]. Those who become more seriously ill are more likely to be older and male, with progressively more risk with each decade over the age of 50 [3]. In addition to these, the people with medical problems like diabetes, cancer, cardiovascular disease, and chronic respiratory disease are more likely to develop serious illness [2]. Although, there

are no specific treatments or vaccines for COVID-19, there are many ongoing clinical trials evaluating potential treatments. Despite the lack of vaccine or treatment, people can prevent the infection by washing hands, staying home, covering the mouth and nose when coughing or sneezing, refraining from smoking. These precautions are not for the treatment, yet they can protect people from the disease and slow the transmission of COVID-19.

Several studies have reported different laboratory findings at the beginning of the outbreak of COVID-19 [42,43]. Most of the cases are mild and clinical outcomes of patients have varied greatly [6–8]. Thus, it may be difficult to identify risk groups by using some features such as gender, age alone. In addition to these, it is essential to predict which patients will more likely to develop severe illness and will face a greater risk including, death itself. These are the important factors when the clinical resources and tools (hospital beds, medical mask, respirator, capacity of the hospital, etc.) are limited, and health care providers are forced to make judgments about the patients without any past experience to guide them. Because of all of these limitations, an artificial intelligence (AI) aided system is required to make such decisions. AI is actively used in healthcare systems to provide clinical decision support [9–11]. Machine learning classifiers are effective to interpret the medical findings such as epilepsy [12,13], nerve and muscle diseases [14,15], heart rhythms [16,17]. Deep learning algorithms also ef-

* Corresponding author.

E-mail addresses: talhaburakalakus@klu.edu.tr (T.B. Alakus), iturkoglu@firat.edu.tr (I. Turkoglu).

fective to predict clinical findings from cancers [18], virus diseases [19], and biomedical studies [20,21]. Such techniques are efficient and they can be used to predict COVID-19 infection.

In this study, we provide a prediction system for detection of COVID-19 infection by developing and applying various deep learning application models. Six various deep learning application models are designed and used on laboratory findings of patients. Performance of the models are measured with accuracy, precision, recall, AUC, and F1-scores. The main objectives of this research can be summarized as follows;

- To provide a prediction study for COVID-19 disease with deep learning application models with laboratory findings rather than X-ray or CT images,
- To ensure the prediction model for this novel pneumonia.

To the best of our knowledge there is no study to use deep learning models to predict COVID-19 infection with laboratory findings. This study may encourage the researches to validate the models by applying different laboratory data.

The paper is organized as follows. Section 2 describes the laboratory findings of the data set and deep learning models. The parameters and necessary information about the developed deep learning application models are given. Section 3 provides the experimental results of deep learning classifiers and the evaluation criteria including, accuracy, recall, precision, AUC, and F1-scores. Finally, Section 4 presents conclusion and provides potential future researches.

2. Related work

It is important to predict clinical tasks for health base systems. Computer aided clinical predictive models have been used in various areas including risk of heart failure [29], mortality in pneumonia [30,31], mortality risk in critical care [32–34]. With these systems medical experts are enable to comprehend and assess clinical findings better. In this study, we build on recent methodological advances to provide clinical predictive model for COVID-19. Similar studies about clinical prediction for COVID-19 are limited in the literature. Authors in [26], used machine learning techniques to predict the clinical severity of coronavirus. Data was obtained from Wenzhou Central Hospital and Cangnan People's Hospital in Wenzhou, China and cannot be accessible since the data is private. Eleven clinical features were considered and six different - Logistic regression, k nearest neighborhood (KNN), 2 different decision trees, random forests and support vector machines (SVM) -classifiers were applied. The performance of the classifiers was evaluated with only accuracy values. Best accuracy was obtained with SVM classifier with 80%. In the another study [27], authors applied machine learning classifiers to predict COVID-19 diagnosis. Clinical data was obtained from Hospital Israelita Albert Einstein at Sao Paulo Brazil. 18 clinical findings were considered in the study and classifiers were evaluated with AUC, sensitivity, specificity, F1-score, Brier score, positive predictive value, and negative predictive value. Only five different classifiers were applied including, SVM, random forests, neural networks, logistic regression, and gradient boosted trees. The best AUC scores were obtained with both SVM, and random forest classifiers with 0.847. In the study of [28], clinical predictive model for COVID-19 was proposed. In the study, data was collected from Hospital Israelita Albert Einstein at Sao Paulo, Brazil like in this study and [27]. Authors applied various machine learning applications including RF, NN (Neural Network), LR, SVM, XFB (Gradient Boosting) and determined the performance of classifiers by calculating sensitivity, specificity, and AUC scores. The best performance was obtained with XGB with 66% AUC score.

Table 1

18 Laboratory findings of the patients in the dataset.

Laboratory Findings	Hematocrit, hemoglobin, platelets, red blood cells, lymphocytes, leukocytes, basophils, eosinophils, monocytes, serum glucose, neutrophils, urea, C reactive protein, creatinine, potassium, sodium, alanine transaminase, aspartate transaminase
---------------------	---

3. Methods and data

3.1. Data description

Dataset includes the laboratory findings of the patients seen at the Hospital Israelita Albert Einstein at Sao Paulo Brazil and can be accessed through [28]. Samples were collected from patients to detect SARS-CoV2 in the early months of 2020. Dataset contains 111 laboratory findings from 5644 various patients. In the dataset, the rate of positive patients was around 10% of which around 6.5% and 2.5% required hospitalization and critical care. In the dataset, there is no gender information. According to the study of [26–28], 18 laboratory findings have a vital role on COVID-19 disease. Thus, we wiped away remaining laboratory features to balance the dataset and to perform COVID-19 detection. After the balancing process, dataset includes 18 laboratory findings from 600 patients, since some of the 18 laboratory findings are unknown to some patients, the number of patients decreased from 5644 to 600. In the balanced dataset, we have 520 no findings and 80 COVID-19 patients. Table 1 shows the laboratory findings. Researchers can access the balanced dataset via <https://github.com/burakalakuss/COVID-19-Clinical>.

3.2. Deep learning application models

AI based algorithms learn from the historical data to provide predictions for the future outcomes. Machine learning (ML) and deep learning (DL) algorithms can be considered as a subsets of the AI. It is an area that is based on learning and improving on its own by analyzing computer algorithms. There are certain differences between machine learning and deep learning. Until recently, DL algorithms were limited by computing power and complexity. Yet, developments in big data have allowed larger and deeper networks, providing computers to learn, observe and react to complex situations faster than humans. In general DL is used for image classification [22], speech recognition [23], bioinformatics [24], etc.

In this study, we develop and evaluate clinical predictive models to determine the COVID-19 infection with laboratory findings. To evaluate the study, we trained six different model types: Artificial Neural Network (ANN), Convolutional Neural Networks (CNN), Long-Short Term Memory (LSTM), Recurrent Neural Networks (RNN), CNNLSTM, and CNNRNN. ANN is an information processing approach that is inspired by the biological nervous system of human brain. It is composed of neurons, activation functions, input, output, and hidden layers. CNN is one of the variants of neural networks and is highly used in image classification studies. It includes convolutional layers, pooling layers, fully-connected layers, and a classification layer. Convolution layers are responsible for feature extraction. Unlike machine learning, CNN obtains features by itself. In the pooling layer, the dimension of the inputs is reduced. RNN is a kind of feedforward neural network which has an internal memory. It uses the same function for every input while the output of the current input depends on the past one computation. RNN uses its internal memory to process the inputs. LSTM is the modified version of the RNN. In the LSTM, it is easier to remember the past data in the memory. The vanishing gradient problem of RNN is resolved in the LSTM networks. Alongside all off CNN, RNN, LSTM, and ANN deep learning models, we de-

Table 2
Parameters of each DL classifier.

Parameters	ANN	CNN	LSTM	RNN	CNNLSTM	CNNRNN
Number of units	32,16,8	512,256	-	-	512,256	512,256
Number of layers	1,2,3	1,2	1	1	1,2	1,2
Activation function	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU
Learning rate	1e-3	1e-3	1e-3	1e-3	1e-3	1e-3
Loss function	Binary crossentropy	Binary crossentropy	Binary crossentropy	Binary crossentropy	Binary crossentropy	Binary crossentropy
Number of epoch	250	250	250	250	250	250
Optimizer	SGD	SGD	SGD	SGD	SGD	SGD
Decay	1e-5	1e-5	1e-5	1e-5	1e-5	1e-5
Momentum	0.3	0.3	0.3	0.3	0.3	0.3
Number of fully connected units	-	2048,1024	2048,1024	2048,1024	2048,1024	2048,1024
Number of fully connected layers	-	1,2	1,2	1,2	1,2	1,2
Number of LSTM units	-	-	512	-	512	512
Number of RNN units	-	-	-	512	-	512
Dropout	-	-	-	0.25	0.15	0.15

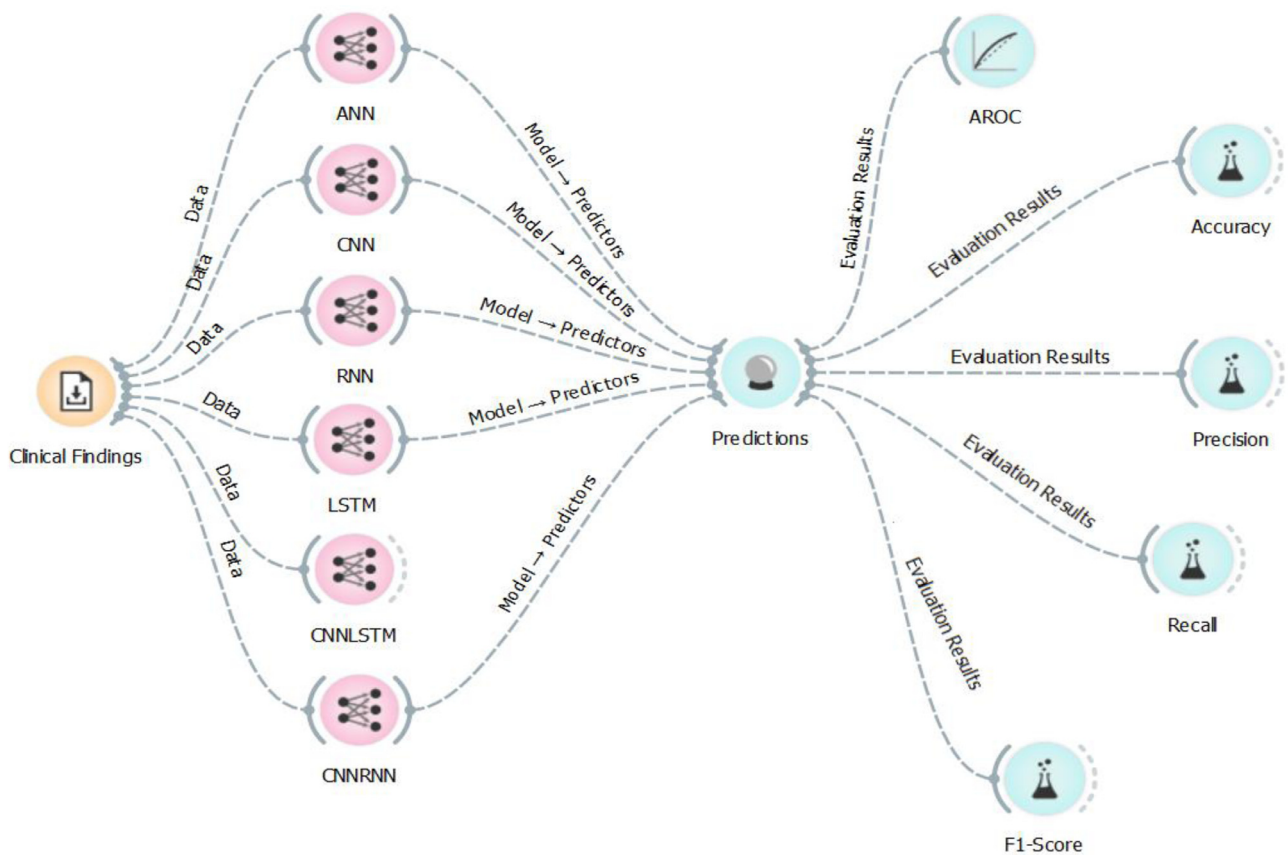


Fig. 1. Flowchart of this study. The orange icon indicates the dataset, which is laboratory findings in this study. The pink ones represent the deep learning models including, ANN, CNN, RNN, LSTM, CNNLSTM, and CNNRNN. All of these models were used to predict the No findings and COVID-19 patients. AUC, Accuracy, Precision, Recall, and F1-Scores were applied to evaluate the results. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

veloped two hybrid models including CNNLSTM, and CNNRNN. We followed a trial and error approach to set the parameters for each DL models. Table 2 emphasizes the parameters of each classifier.

To assess the predictive performance of each of the developed predictive models, we calculated their performance in terms of accuracy, f1-score, precision, recall, and area under roc curve (AUC). To validate the data, we both used 10-fold cross validation and 80–20 train-test split approach. Fig. 1 images the flowchart of the predictive model.

3.3. Application results

Totally 18 laboratory findings from 600 patients were considered for the prediction of COVID-19 infection. All of the samples

are the laboratory findings of the patients. Six different deep learning application models were developed and applied as classifiers. Later, predictions were performed and the performance of the deep learning applications models were evaluated. Table 3 shows the evaluation results of all deep learning application models with 10 fold cross-validation approach.

In terms of predictive performance, we observed that the overall best identified models by AUC score were 62.50 by LSTM for predicting COVID-19 disease. It is noticed that predicting COVID-19 disease from laboratory findings was a challenging task, since collecting the samples need a time and complex procedures. Nevertheless, the best clinical prediction results achieved a respectable accuracy of 86.66%, f1-score of 91.89%, and recall of 99.42%, respectively with LSTM. It is not a surprising result, since LSTM is good

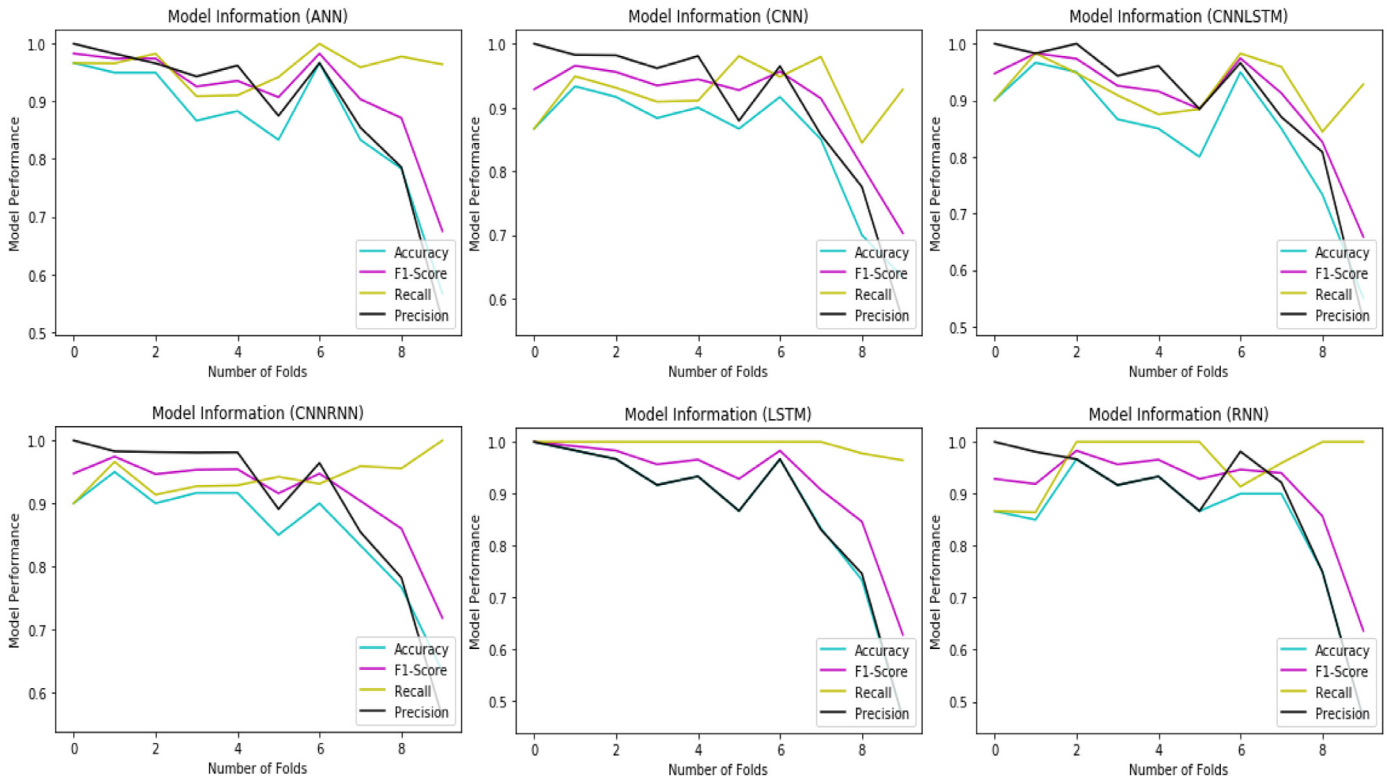


Fig. 2. Evaluation results of all deep learning models with 10 fold cross-validation approach.

Table 3

Evaluation results of all deep learning application models with 10 fold cross-validation approach.

	Accuracy	F1-Score	Precision	Recall	AUC
ANN	0.8600	0.9134	0.8855	0.9578	0.5615
CNN	0.8800	0.9038	0.8948	0.9248	0.6149
CNNLSTM	0.8416	0.9001	0.8926	0.9214	0.5889
CNNRNN	0.8566	0.9120	0.8977	0.9423	0.6408
LSTM	0.8666	0.9189	0.8675	0.9942	0.6250
RNN	0.8416	0.9061	0.8783	0.9604	0.5245

Table 4

Evaluation results of all deep learning application models with train-test split approach.

	Accuracy	F1-Score	Precision	Recall	AUC
ANN	0.8690	0.8713	0.8713	0.8713	0.85
CNN	0.8735	0.8856	0.8847	0.8867	0.80
CNNLSTM	0.9230	0.9300	0.9235	0.9368	0.90
CNNRNN	0.8624	0.8755	0.8755	0.8755	0.69
LSTM	0.9034	0.8997	0.8997	0.8998	0.83
RNN	0.8400	0.8427	0.8428	0.8427	0.83

for such sequences which have long term dependencies in it and is powerful when the data contains time series. Fig. 2 shows the model evaluation results.

In addition to these, we tested the performance of the algorithms using 80–20 train-test split approach. Although k fold cross-validation approach is frequently used in artificial intelligence in health studies especially in cases of relatively small samples, it generates less clearly the results in clinical applications [27]. The clinical predictive performance of all algorithms was better in comparison with 10 fold cross-validation strategy with an AUC of 0.90, accuracy of 0.9230, f1-score of 0.93, precision of 0.9235, and recall of 0.9368 for the best-performing algorithm, which was CNNLSTM hybrid model. Table 4 shows the evaluation results of all deep learning models with train-test split approach.

As can be seen in Table 4, the accuracy results of all deep learning models were reached at least 84.00% and above. The best evaluation performance was obtained with CNNLSTM hybrid model with 92.30%. LSTM was observed as the second best model. Although LSTM is powerful and performs well in time series, it did not surpass the hybrid model CNNLSTM. The main reason for this result is, CNNLSTM is a model which is both spatially and temporally deep, and has the flexibility to be applied to a variety of tasks involving sequential inputs and outputs [35,36]. In addition

to these, CNN performs as an encoder and feature extractor, while LSTM is responsible for decode. This provides an advantage to CNNLSTM model [36].

All F1-score, precision, and recall results were observed above 84.00%. Precision can be defined as the ration of correctly predicted positive observations to the total predicted positive observations. In information retrieval studies, a perfect precision should be 1. In this research, the best precision score was obtained with CNNLSTM with 0.9235. Recall is the ratio of correctly predicted positive observations to the all observations. Like precision, a recall score must reach to the 1 for the perfect classification process. The best recall value was obtained from CNNLSTM deep learning application model with 0.9368. F1 score is the weighted average of precision and recall values. This evaluation criterion takes both false positives and false negatives. A good F1-score means that classifier has low false positives and low false negatives. In this case, classifier identify the real threats and not disturbed by false alarms. An F1-score is considered perfect when the value is 1. Like any other evaluation criteria, the best F1-score obtained with CNNLSTM with 0.9300. AUC is used in the classification analysis to determine which of the used models predicts the classes best. In general, an AUC score of 0.5 means that there is no discrimination, a score between 0.6 and 0.8 is considered acceptable, a score be-

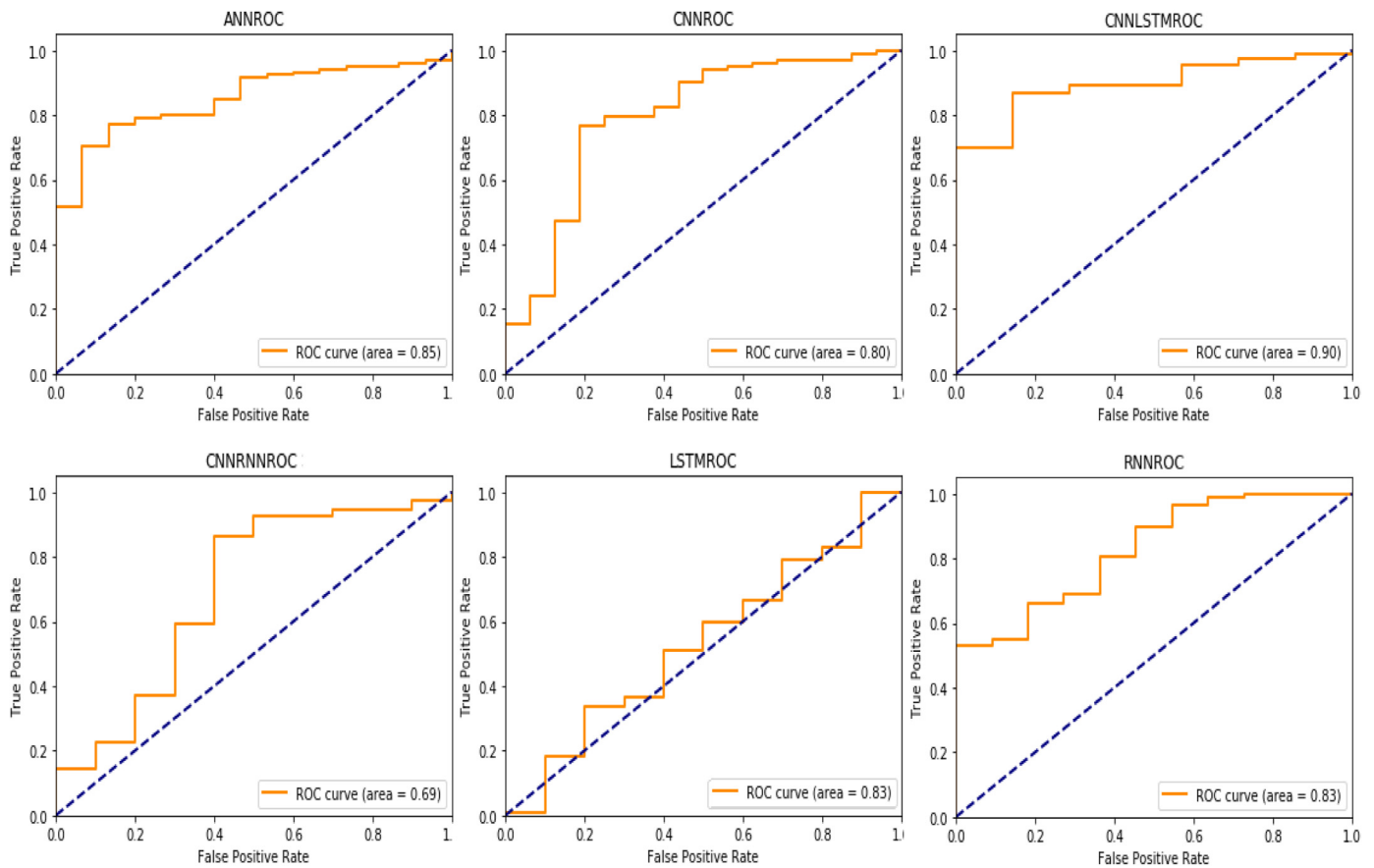


Fig. 3. AUC values of all deep learning application models with train-test split approach.

Table 5
Comparison of evaluation results.

Study	Dataset Location	AI Technique	Classifier	Accuracy	AUC	F1-Score
[26]	Wenzhou Central Hospital and Cangnan People's Hospital in Wenzhu, China	Machine learning	SVM	80.00%	-	-
[27]	Hospital Israelita Albert Einstein at Sao Paulo, Brazil	Machine learning	SVM, RF	-	0.87	0.72
[28]	Hospital Israelita Albert Einstein at Sao Paulo, Brazil	Machine learning	XGB	-	0.66	-
This work	Hospital Israelita Albert Einstein at Sao Paulo, Brazil	Deep learning	CNNLSTM	92.30%	0.90	0.93

tween 0.8 and 0.9 is considered excellent, and more than 0.9 is considered outstanding [25]. The AUC score of CNNRNN is considered acceptable since the results ranges between 0.6 and 0.8. The AUC scores of the remaining ones were excellent since all of the results were higher than 0.8. According to the AUC scores, all deep learning models may be used for clinical prediction of COVID-19. In critical medical and clinical studies, it is essential to obtain true positive rates since recall represents the percentage of actual positives are detected [37]. In this study, recall is important evaluation criteria since it is computed by taking the ratio of correctly identified COVID-19 patients to the total number of COVID-19 diseased patients. In addition to these, AUC score has a vital role on medical researches, since it has a meaningful interpretation for disease classification from healthy subjects [38,39]. Accuracy is a research characteristic, which provides a way to know how close are the sample parameters to population characteristics [40]. By measuring the accuracy of the models, the researcher can prove that the research is generalizable, reliable, and valid [41]. Thus in this study, only these three evaluation metrics were considered. Remaining ones were calculated to compare the results with [26–28]. In Fig. 3, we provided the AUC scores of deep learning models with train-test split approach. Fig. 3 shows the AUC scores of all deep learning models.

Table 5 lists the comparison result of classifiers between this research and other studies.

In the study of [26–28], authors used machine learning techniques. As seen in Table 5, best classification was obtained with SVM and XGB classifiers in these studies. Yet, in this research, we did not use machine learning. We developed six different deep learning application models, and reached better accuracy, and AUC scores than machine learning classifiers. It showed that, DL approaches can be more powerful than ML approaches even the data is small.

4. Conclusion and discussion

In this study, the prediction of COVID-19 outbreak was carried out with deep learning models based on laboratory findings. Various laboratory data were analyzed with 6 different deep learning models. In the first stage of the study, the data were standardized and then used as inputs for the deep learning models. Later, classification was carried out and the performances of the models were measured with precision, recall, accuracy, AUC, and F1-scores. To validate the models, we applied 10 fold cross-validation and train-test split approaches. In 10 fold cross-validation strategy, best meaningful results observed from LSTM deep learning model

with accuracy of 86.66%, recall of 99.42%, and AUC score of 62.50%. Although, this validation is popular, it did not yield the best validation result. The best accuracy, recall and AUC values were obtained with CNNLSTM model as 92.3%, 93.68%, and 90.00%, respectively in train-test split approach. All deep learning models developed in the study showed an accuracy of over 84%. Similar inferences can be made for precision and recall values.

The major limitation in this study is the size of the data. Data of 600 patients were used, and some laboratory findings could not be measured for some patients. However, in a measurable population range, the prediction took place between 84%, and 93%. In addition to these, the data was imbalanced, thus we balanced the data by deleting some materials. The performance of these models can be enhanced with a larger data set.

Further studies need to be carried out with other laboratory findings obtained from other locations to validate these results. We only analyzed the samples from Hospital Israelita Albert Einstein. In addition to these different stages of the disease may affect the predictive performance of the models.

Moreover, in this study, it has been observed that decision-making mechanisms can distinguish between patient and non-patient, and the values such as fever, and lymphopenia are not very essential for the prediction process. In future studies, with the use of artificial intelligence techniques and the increase in the number of data, early diagnosis of COVID-19 diseases and early treatment opportunities can be provided.

Globally, various real-time RT-PCR protocols have been proposed for the diagnosis of COVID-19 [44]. RT-PCR tests performance is impacted by several factors that are difficult to measure, such as low levels of shedding during incubation and early infection, variability in the site of specimen acquisition, and sufficiency of sample collected [45–47]. In the light of all these data, these modeling techniques reveal the importance for early detection of COVID-19 infection and to start treatment without delay.

In conclusion, we found evidence to suggest that deep learning application models can be applied to predict COVID-19 infection with laboratory findings. Our experimental results indicate that may be useful to help prioritize scarce healthcare resources by assigning personalized risk scores using laboratory and blood analysis data. In addition to these, our findings on the importance of laboratory measurements towards predicting COVID-19 infection for patients increase our understanding of the outcomes of COVID-19 disease. Based on our study's results, we conclude that healthcare systems should explore the use of predictive models that assess individual COVID-19 risk in order to improve healthcare resource prioritization and inform patient care.

Declaration of Competing Interest

There is no conflict of interest in this research.

Acknowledgment

We would like to thank Safak Ozer Balin, assistant professor at Firat University, Faculty of Medicine, Department of Infectious Disease, for her valuable and informative comments during the research.

References

- [1] World Health Organization, Report of the WHO-China joint mission on coronavirus disease (COVID-19). 2020 <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>.
- [2] World Health Organization, Health topics, coronavirus. 2020 https://www.who.int/health-topics/coronavirus#tab=tab_3.
- [3] National Institute of Infection Diseases, Field briefing: diamond princess COVID-19 cases. 2020 <https://www.niid.go.jp/niid/en/2019-ncov-e/9407-covid-dp-fe-01.html>.
- [4] Del Rio C, Malani PN. Novel coronavirus – important information for clinicians. *J Am Med Assoc* 2019;323(11):2020. doi:10.1001/jama.2020.1490.
- [5] Wang D, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *J Am Med Assoc* 2020;323(11) DOI: 1061-1069. doi:10.1001/jama.2020.1585.
- [6] Jiehao C, et al. A case series of children with 2019 novel coronavirus infection: clinical and epidemiological features. *Clin Infect Dis* 2020;ciaa198. doi:10.1093/cid/ciaa198.
- [7] Karm KQ, et al. A well infant with coronavirus diseases 2019 (COVID-19) with high viral load. *Clin Infect Dis* 2020;ciaa201. doi:10.1093/cid/ciaa201.
- [8] Bai Y, Yao L, Wei T, et al. Presumed asymptomatic carrier transmission of COVID-19. *J Am Med Assoc* 2020;323(14):1406–7. doi:10.1001/jama.2020.2565.
- [9] Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017;2(4). doi:10.1136/svn-2017-000101.
- [10] Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthcare J* 2019;6(2):92–8. doi:10.7861/futurehosp.6-2-94.
- [11] Reddy S, Fox J, Purohit MP. Artificial intelligence-enabled healthcare delivery. *J R Soc Med* 2019;112(1):22–8. doi:10.1177/014107681881551.
- [12] Alakus TB, Turkoglu I. Detection of pre-epileptic seizure by using wavelet packet decomposition and artificial neural networks. In: 10th International Conference on Electrical and Electronic Engineering; 2017. p. 511–15.
- [13] Memarian N, Kim S, Dewar S, Engel J, Staba RJ. Multimodal data and machine learning for surgery outcome prediction in complicated cases of mesial temporal lobe epilepsy. *Comput Biol Med* 2015;64(1):67–78. doi:10.1016/j.compbiomed.2015.06.008.
- [14] Yousefi J, Hamilton-Wright A. Characterizing EMG data using machine-learning tools. *Comput Biol Med* 2014;51:1–13. doi:10.1016/j.compbiomed.2014.04.018.
- [15] Karthick PA, Ghosh DM, Ramakrishnan S. Surface electromyography based muscle fatigue detection using high-resolution time-frequency methods and machine learning algorithms. *Comput Methods Programs Biomed* 2018;154:45–56. doi:10.1016/j.cmpb.2017.10.024.
- [16] Alfaras M, Soriano MC, Ortin S. A fast machine learning model for ECG-based heartbeat classification and arrhythmia detection. *Front Phys* 2019. doi:10.3389/fphys.2019.00103.
- [17] Ledezma CA, Zhou X, Rodriguez B, Tan PJ, Diaz-Zuccarini V. A modeling and machine learning approach to ECG feature engineering for the detection of ischemia using pseudo-ECG. *PLoS ONE* 2019;14(8) PMC690680. doi:10.1371/journal.pone.0220294.
- [18] Munir K, Elahi H, Ayub A, Frezza F, Rizzi A. Cancer diagnosis using deep learning: a bibliographic review. *Cancers (Basel)* 2019;11(9):E1235. doi:10.3390/cancers11091235.
- [19] Andriasyan, V., Yakimovich, Georgi, F. et al., Deep learning of virus infections reveals mechanics of lytic cells, bioRxiv, 2019. doi: <https://doi.org/10.1101/798074>.
- [20] Senior AW, Evans R, Jumper J, et al. Improved protein structure prediction using potentials from deep learning. *Nature* 2020;577:706–10. doi:10.1038/s41586-019-1923-7.
- [21] Bosco G, Gangi MA. Deep learning architectures for DNA sequence classification. *Lect Notes Comput Sci* 2017:162–71. doi:10.1007/978-3-319-52962-2_14.
- [22] Krishna MM, Neelima M, Harshali M, Rao MVG. Image classification using deep learning. *Int J Eng Technol* 2018;7(2.7):614–17. doi:10.14419/ijet.v7i2.7.10892.
- [23] Nassif AB, Shahin I, Attili I, Azzeh M, Shaalan K. Speech recognition using deep neural networks: a systematic review. *IEEE Access* 2019;7 19413–19165. doi:10.1109/ACCESS.2019.2896880.
- [24] Li, Y, Huang, C, Ding, L, Li, Z, Pan, Y, Gao, X. Deep learning in bioinformatics: introduction, application, and perspective in big data era, *arXiv*, 2019.
- [25] Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thoracic Oncol* 2010;5(9):1315–16. doi:10.1097/JTO.0b013e3181ec173d.
- [26] Jiang X, Coffee M, Bari A, Wang J, Jiang X, et al. Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. *Compu Mater Continua* 2020;63(1):537–51. doi:10.32604/cmc.2020.010691.
- [27] Batista, A.F., Miraglia, J.L., Donato, T.H.R., and Filho, A.D.P.C., COVID-19 diagnosis prediction in emergency care patients: a machine learning approach, *medRxiv*, 2020. doi: 10.1101/2020.04.04.20052092.
- [28] Schwab, P., Schütte, A.D., Dietz, B., and Bauer, S. “predCOVID-19: a systematic study of clinical predictive models for coronavirus disease 2019, arXiv:2005.08302, 2020.
- [29] Wu J, Roy J, Stewart WF. Prediction modeling using EHR data: challenges, strategies, and a comparison of machine learning approaches. *Med Care* 2010;48(6):106–13. doi:10.1097/MLR.0b013e3181de9e17.
- [30] Cooper GF, Aliferis CF, Ambrosino R, Aronis J, Buchanan BG, et al. An evaluation of machine-learning methods for predicting pneumonia mortality. *Artif Intell Med* 1997;9(2):107–38. doi:10.1016/S0933-3657(96)00367-3.
- [31] Wu C, Rosenfeld R, Clermont G. Using data-driven rules to predict mortality in severe community acquired pneumonia. *Plos ONE* 2014;9(4):e89053. doi:10.1371/journal.pone.0089053.
- [32] Clermont G, Angus DC, DiRusso SM, Griffin M, Linde-Zwirble WT. Predicting hospital mortality for patients in the intensive care unit: a comparison of artificial neural networks with logistic regression models. *Crit Care Med* 2001;29(2):291–6.
- [33] Ghassemi M, Naumann T, Doshi-Velez F, Brimmer N, Joshi R, Rumshisky A, Szolovits P. Unfolding physiological state: mortality modelling in intensive care units. *KDD* 2014:75–84. doi:10.1145/2623330.2623742.
- [34] Johnson AEW, Pollard TJ, Mark RG. Reproducibility in critical care: a mortality prediction case study. *Proc Mach Learn Res* 2017;68:361–76.

- [35] Donahue J, Hendricks LA, Rohrbach M, Venugopalan S, Guaddarrama S, Saenko K, et al. Long-term recurrent convolutional networks for visual recognition and description. *IEEE Trans. Patt. Anal. Mach. Intelli.* 2017;39(4):677–91. doi:10.1109/TPAMI.2016.2599174.
- [36] Vinyals O, Toshev A, Bengio S, Erhan D. Show and tell: a neural image caption. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2015. doi:10.1109/CVPR.2015.7298935.
- [37] Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Improving palliative care with deep learning. *BMC Med Inform Decis Mak* 2018;18(4). doi:10.1186/s12911-018-0677-8.
- [38] Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med* 2013;4(2):627–35.
- [39] Kamarudin AN, Cox T, Kolamunnaage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med Res Methodol* 2017;17(53). doi:10.1186/s12874-017-0332-6.
- [40] Wynants L, Calster BV, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction model for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 2020;369. doi:10.1136/bmj.m1328.
- [41] Pierce R. Evaluating information: validity, reliability, accuracy, triangulation. *Res Methods Polit* 2008;79–99. doi:10.4135/9780857024589.
- [42] Li H, Li C, Liu HG. Clinical characteristics of novel coronavirus cases in tertiary hospitals in Hubei Province. *Chin Med J* 2020;133(9):1025–31. doi:10.1097/CM9.0000000000000744.
- [43] Huang C, Wang Y, Li X, Ren L, Zhao J, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395(10223):497–506. doi:10.1016/S0140-6736(20)30183-5.
- [44] World Health Organization. Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases: interim guidance. <https://www.who.int/publications-detail/laboratory-testing-for-2019-novel-coronavirus-in-suspected-human-cases-20200117>. (Updated on March 19, 2020).
- [45] Wölfel R, Corman VM, Guggemos W, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature* 2020;581:465–9. doi:10.1038/s41586-020-2196-x.
- [46] Wang, W., Xu, Y., Gao, R., and et al. "Detection of SARS-CoV-2 in different types of clinical specimens," *JAMA*, 323(18), 1843–4, 220. doi: 10.1001/jama.2020.3786.
- [47] Fang Y, Zhang H, Xie J, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology* 2020. doi:10.1148/radiol.2020200432.