Information Sciences Letters

Volume 11
Issue 6 *Nov. 2022*

Article 34

2022

Diagnosis of COVID-19 from X-rays Using Recurrent Neural Network

S. Hussein
College of Computer Science and Informatics, Amman Arab University, Amman, Jordan, sabahussien7@gmail.com

M. Otair
College of Computer Science and Informatics, Amman Arab University, Amman, Jordan, otair@aau.edu.jo

S. Alzoubi
College of Computer Science and Informatics, Amman Arab University, Amman, Jordan, sabahussien7@gmail.com

A. Al-Sayyed
School of Dentistry, The University of Jordan, Amman, Jordan, sabahussien7@gmail.com

Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/110634

Diagnosis of COVID-19 from X-rays Using Recurrent Neural Network

S. Hussein[1,], M. Otair[1], S. Alzoubi[1], A. Al-Sayyed[2], S. Numan[1], A. Numan[1], S. Jamal[1] and R. Al-Sayyed[3]*

[1]College of Computer Science and Informatics, Amman Arab University, Amman, Jordan
[2]School of Dentistry, The University of Jordan, Amman, Jordan
[3]King Abdullah II School for Information Technology, The University of Jordan, Amman, Jordan

Received: 1 May 2022, Revised: 15 May 2022, Accepted: 23 May 2022.
Published online: 1 Nov. 2022.

Abstract: Nearly two years ago, the COVID-19 pandemic caused by the SARS-CoV-2 virus has caused drastic changes in many aspects of life at many levels in the world, and this has affected people's lifestyles. This impact was particularly significant and impactful on the health sectors, among many others. The COVID-19 virus has essentially increased the demand for treatment, diagnosis and testing. The definitive test for diagnosing COVID-19 is reverse transcriptase polymerase chain reaction (RT-PCR); nevertheless, chest x-ray is a quick, effective and inexpensive diagnosis to detect possible pneumonia associated with COVID-19. In this study, the feasibility of using a deep learning-based Recurrent Neural Network (RNN) classifier to detect COVID-19 from CXR images is investigated. The proposed classifier consists of an RNN, trained by a deep learning model. The RNN identifies abnormal images that contain signs of COVID-19. The experiment used in the study employed 286 COVID-19 samples from the Kaggle Repository. The proposed technique is compared with the decision tree algorithm in order to prove the efficiency of the proposed one. The results revealed that the accuracy of the RNN was 97.90%, with a low data loss rate of 2.10%, while the decision tree accuracy was 75.8741%, and a relatively high data loss rate of 24.1259%. These results support the usefulness of the proposed deep learning-based RNN classifier in pre-screening patients for triage and decision-making before RT-PCR data are available.

Keywords: COVID-19, chest x-ray, Recurrent Neural Network, RT-PCR, Kaggle Repository.

1 Introduction

In big data era, machine learning has become an increasingly popular approach for data processing. Data could be in various forms, such as text, images, audios, videos and signals [1]. The essence of machine learning is to learn any patterns from features of data [2]. In the above types of data, the number of features is massively high, which could result in the presence of a large number of irrelevant features. Most machine learning algorithms are sensitive to irrelevant features, so effective evaluation and selection of features in machine learning tasks are highly important [3][4]. Also, effective evaluation of features can help identify which features are necessary to be extracted from unstructured data. In this paper, we proposed Recurrent Neural Network (RNN) to classify patients' Xray images into either normal (not infected with COVID-19) or abnormal (infected with COVID-19) ones [5].
In this paper, an RNN class of Artificial Neural Networks (ANN) was proposed to identify coronavirus cases from chest X-rays. We used RNN to find out the best architecture within the limitations (including imbalanced training data to inadequate benchmarks) of the datasets [6]. When predicting images, datasets are split into either training or testing. Final step to apply RNN classifier is to predict whether the image is normal or COVID 19 positive [7], Fig.1 in section 3 illustrates the proposed model.
In this examination, we arrange datasets of Chest Xray images. The dataset contains of 286 number of Covid-19 positive X-ray chest images. This dataset is gathered from Kaggle repository [8][9]. The dataset is analyzed independently in the proposed model. We utilize this dataset for profound component extraction dependent on deep learning strategy. The contributions of this paper can be summarized as follows:

i) To distinguish and classify coronavirus from other infections, an RNN is developed and trained.

ii) Kaggle Repository is utilized to build a new dataset collected from many X-ray patients' images in order to classify them to either COVID-19 positive or negative.

The rest of this paper is organized as follows; a brief review of related works is presented in Section 2. Experimental Dataset is described in Section 3, Learning with RNN in Section 4, while Section 5 include result analysis with relative discussions. Finally, the conclusion of the paper is in Section 6.

*Corresponding author e-mail: sabahussien7@gmail.com

2 Related Work

Because of the COVID-19 pandemic, many efforts have been explored to develop a system for the diagnosis of COVID-19 using artificial intelligence techniques such as machine learning [12], deep learning [13]. In this section, a detailed description is provided about the recently developed systems to diagnose COVID-19 cases.

Luz et al. [14] introduced an extended Efficient Net model based on convolutional network architecture to analyze the Lung condition using X-ray images. The model used 183 samples of COVID-19 and achieved 93.9% accuracy and 80% sensitivity for coronavirus classification. Rahimzadeh and Attar [15] presented a concatenated Xception and ResNet50V2 network to find out the infected region of COVID-19 patients from chest X-rays. The network trained on eight phases and used 633 samples in each phase including 180 samples of COVID-19. The network obtained 97.56% accuracy and 80.53% recall to detect coronavirus infection.

Minaee et al. [16] illustrated a deep transfer learning architecture utilizing 71 COVID-19 samples to identify infected parts from other lung diseases. The architecture obtained an overall 97.5% sensitivity and 90% specificity to differentiate coronavirus cases. Punn and Agarwal [17] demonstrated a deep neural network to identify coronavirus symptoms. The scheme used 108 COVID-19 cases and obtained an average 97% accuracy.

Khan et al. [18] introduced a deep RNN to diagnose coronavirus disease from 284 COVID-19 samples. The proposed framework found an accuracy of 89.5%, and a precision of 97% to detect coronavirus. Wang and Wong [19] presented COVID Net to distinguish COVID-19 cases from others using chest X-ray samples. The system achieved 92.4% accuracy by analyzing 76 samples of COVID-19.

Narin et al. [20] proposed a deep transfer learning with three CNN architectures and used a small dataset including 50 chest X-rays for both COVID-19 and normal cases to detect coronavirus infection. The ResNet50 showed high performance with 98.6% accuracy, 96% recall, and 100% specificity among other networks. Table 1 shows the mostly related studies of the current study. Note that the achieved accuracy depends on the used dataset in each study.

Table 1: Summary of Mostly Related Studies

Author	Methods	Samples	Accuracy
Luz et al. [14]	CNN	COVID-19 samples	93.90%
Rahimzadeh and Attar [15]	Xception & ResNet50V2	Chest X-rays	97.56%
Minaee et al. [16]	Deep transfer learning	COVID-19 samples	97.50%
Punn and Agarwal [17]	NN	COVID-19	97%
Khan et al. [18]	RNN	COVID-19 samples	97%
Wang and Wong [19]	COVID Net	X-ray samples	92.40%
Narin et al. [20]	Transfer learning with three CNN	X-rays	98.60%

3 Experimental Dataset

In this paper, X-ray samples of COVID-19 were obtained from Kaggle repository. The train set has 15264 (512x512) chest X-ray pictures, whereas the test set contains 400 images. Positive and negative classes are included in the dataset to identify positive and negative COVID-19 cases. A sufficient number of samples (COVID-19) cases were used in the experiments for the case. CSV (Comma Separated Value) format used in this study consisting of two columns The first column shows the name of the picture in the train set, while the second column shows the image label. Either 1 or 0 are the labels. Positive samples are given a 1 and negative ones are given a 0. It's worth noting that the dataset is uneven, with only around 10% of the samples being positive, and you'll have to deal with unbalanced data throughout this task. These files converted into ARFF format in order to accepted by WEKA [4] [5]. Figure 1 shows the framework of the proposed model which contains several steps starting from collecting and preparing the dataset and ending with evaluation steps in terms of the used evaluation metrics.

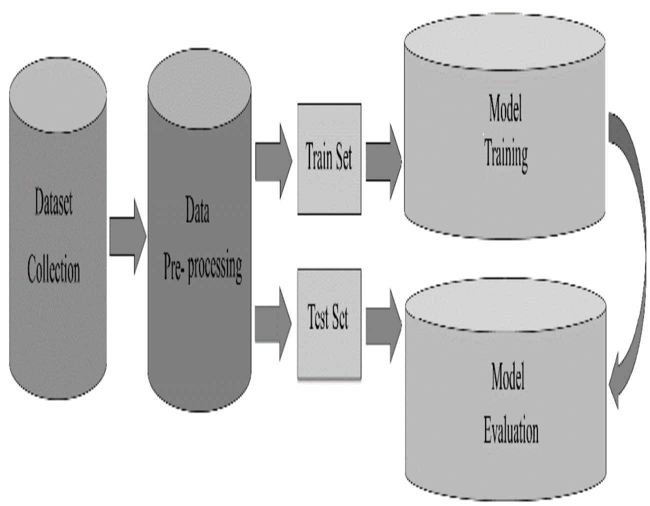

Fig.1: The Proposed Model.

4 Deep Transfer Learning with RNN

Recurrent neural network [9] is an extended feedforward neural network with one or more feedback loops designed for processing sequential data. RNN is used whenever the input-output relationship is found based on time and capacity to handle long term dependencies [3][7]. The strategy of modeling sequence is to feed the input sequence to a fixed-sized vector using an RNN, and then to map the vector to a SoftMax layer. Unfortunately, a problem occurs in RNN when the gradient vector is increasing and decreasing exponentially for long sequences. This vanishing gradient and exploding problem [6][11] create difficulties to learn long-range relationships from the sequences of the RNN architecture. However, the Long Short-Term Memory (LSTM) [8] is capable to solve such a long-distance dependencies problem successfully. The main difference from RNN is that LSTM added a separate memory cell state to store long term states and updates or exposes them whenever necessary [10] [21]. The LSTM consists of three gates: input gate, forget gate, and output gate.

The paper's major goal is to build a model that uses the RNN approach as a prediction model to determine if an X-ray is normal or abnormal based on numerous factors linked with medical data. After pre-processing the dataset to (CSV format), the RNN algorithm is applied to the dataset using WEKA, and the data is classed as normal or abnormal based on the RNN's results.

5 Results Analysis

The main aim of this research is to analyze the classification algorithms performance for COVID-19 data (output) based on the numerous input parameters. They are analyzed using Recurrent Neural Network (RNN) and compared with Decision Tree algorithm. WEKA application is used as development environment for the performance evaluation. Each classifier is applied on the dataset contains of 286 number of Covid-19 positive X-ray chest images. This dataset is gathered from Kaggle repository. The screen shot of the WEKA preprocessing stage is shown in Figure 2. The figure show distribution of COVID-19 sorted by age and gender (The blue color represents female while the red for male).

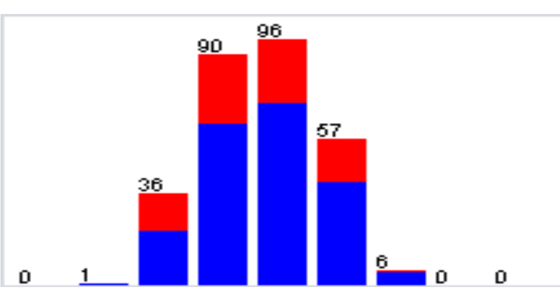

Fig. 2: Data distribution in the preprocessing stage.

In this phase, the classification of pre-processing is carried out based on all the values of taken nine attributes. A comparative study of classification accuracy in RNN and Decision Tree algorithm is carried out in this work. The various formal as used for the calculation of different evaluation measures are as follows.

$$Precision\ P = \frac{TP}{TP+FP} \quad (1)$$

Where TP is True Positive Rate and FP is a False Positive Rate.

Recall or Sensitivity means the proportion of positive cases that were correctly identified. It will be computed as:

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

When FN=False Negative Rate

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

The above formula will calculate the accuracy (the proportion of the total number of predictions that were correct) with TN = True Negative

F-Measure can be computed as some average of the information retrieval precision and recall metrics.

$$F \frac{2*Recall*precision}{precision+Recall} \quad (4)$$

The following are the summary statistics, detailed accuracy by class, and confusion matrix findings from the RNN classifier model assessment.

Total Number of Instances	286	
Correctly Classified Instances	280	97.90%
Incorrectly Classified Instances	6	2.10%
Kappa statistic	0.9498	
Mean absolute error	0.1407	
Root mean squared error	0.1869	
Relative absolute error	33.64%	
Root relative squared error	40.89%	

Accordingly, 280 (97.90%) of the occasions are effectively ordered by the model while 6 (2.10%) of the cases are mistakenly arranged. The Kappa statistic value (0.9498) is larger than 0, indicating that the classifier outperforms well. The general arrangement precision measure is given by the level of effectively ordered cases. The aftereffect of the itemized precision by class is the accompanying:

Table 2: Detailed results by RNN algorithm.

Class	TP Rate	FP Rate	Precision	Recall	F-Measure
no-recurrence-events	0.985	0.035	0.985	0.985	0.985
recurrence-events	0.965	0.015	0.965	0.965	0.965
Weighted Avg.	0.979	0.029	0.979	0.979	0.979

The confusion matrix has additionally displayed in the accompanying:

Table 3: Confusion Matrix of RNN algorithm.

a	b	classified as
198	3	a = no-recurrence-events
3	82	b = recurrence-events

Assuming the class is 'no-recurrence-events' (non-Covid19 patient), there are 198 right and 3 erroneous forecasts while on the off chance that the class is 'recurrence-events' (Covid19 patient), there are 82 right and 3 wrong predictions.

We obtained the following summary information for the Decision Table method after testing the classifier model.

Total Number of Instances	286	
Correctly Classified Instances	217	75.8741 %
Incorrectly Classified Instances	69	24.1259 %
Kappa statistic	0.2899	
Mean absolute error	0.3658	
Root mean squared error	0.4269	
Relative absolute error	87.4491 %	
Root relative squared error	93.4017 %	

Accordingly, 217 (75.8741%) of the occasions are effectively ordered by the model while 69 (24.1259%) of the cases are mistakenly arranged. The general arrangement precision measure is given by the level of effectively ordered cases. The aftereffect of the itemized precision by class is the accompanying:

Table 4 shows the details of the evaluation metrics depending on the measures used to measure the classification accuracy for Decision Tree.

Assuming the class is 'no-recurrence-events' (non-Covid19 patient), there are 194 right and 7 erroneous forecasts while on the off chance that the class is 'recurrence-events' (Covid19 patient), there are 23 right and 62 wrong predictions.

Table 4: Detailed results by Decision Tree algorithm.

Class	TP Rate	FP Rate	Precision	Recall	F-Measure
no-recurrence-events	0.965	0.729	0.758	0.965	0.849
recurrence-events	0.271	0.035	0.767	0.271	0.400
Weighted Avg.	0.759	0.523	0.760	0.759	0.716

Table 5: Confusion Matrix of Decision Tree algorithm.

a	b	classified as
194	7	a = no-recurrence-events
62	23	b = recurrence-events

The accuracy and loss in the training and validation phases are shown in figures 3 and 4. For RNN architecture, the highest training and validation accuracy is observed 97.90% and loss is 2.10% at iterations 100. On the contrary, the lowest training and validation accuracy is obtained 75.8741% and loss is 24.1259% at iterations 100 for the Decision Tree network.

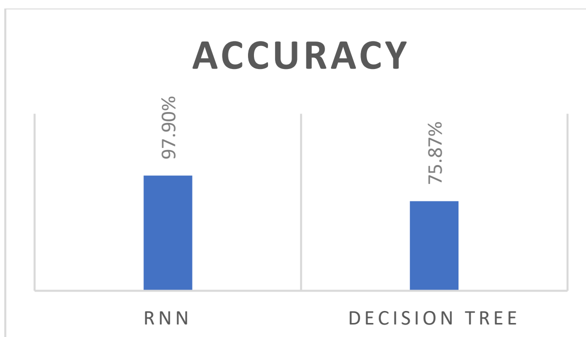

Fig.3: accuracy of algorithms.

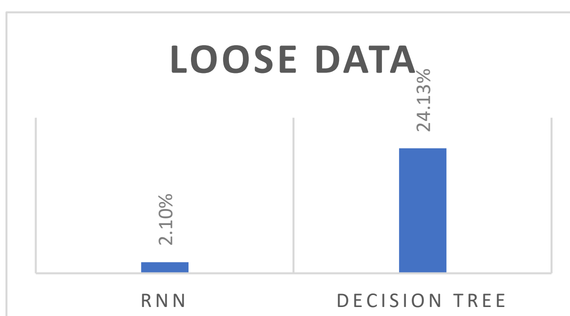

Fig.4: Loose data of algorithms.

6 Conclusions

Herein, a deep learning-based RNN classifier for detecting COVID-19, has been presented and validated using patient data. For classification , an RNN was used with trained data. This research work evaluates the performances in terms of classification accuracy of RNN algorithm using various accuracy measures like FP rate, TP rate, Recall, Precision, and F-measure. Accuracy of 97.90%, were achieved for RNN with loss 2.10% . We believe that this study will have significant clinical applications, allowing fast follow-up decision making and pre-screening in suspected COVID-19 cases prior to the availability of RT-PCR results.

References

[1] Islam, M., Karray, F., Alhajj, R., & Zeng, J. A Review on Deep Learning Techniques for the Diagnosis of Novel Coronavirus (COVID-19). *arXiv preprint arXiv*:2008.04815. (2020)

[2] Organization, World Health.. COVID-19 Vaccines Advice. COVID-19 Vaccines Advice;. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/covid-19-vaccines/adviceworld health organization,

[3] Everything about the Corona virus - Medicine and Health. https://medicine-and-mentalhealth.xyz/archieves/4510. (Accessed 06 June, 2020).

[4] T. Ai, Z. Yang, L. Xia, Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease, Radiology. 2019 (2020) 1–8. https://doi.org/10.14358/PERS.80.2.000.

[5] C. Long, H. Xu, Q. Shen, X. Zhang, B. Fan, C. Wang, B. Zeng, Z. Li, X. Li, H. Li, Diagnosis of the Coronavirus disease (COVID-19) *Eur. J. Radiol.* 126 (2020) 108961. https://doi.org/10.1016/j.ejrad.2020.108961.

[6] H. Shi, X. Han, N. Jiang, Y. Cao, O. Alwalid, J. Gu, Y. Fan, C. Zheng, Articles Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan , China : a descriptive study, *Lancet Infect. Dis.* 20 (2020) 425–434. https://doi.org/10.1016/S1473-3099(20)30086-4.

[7] Z.Y. Zu, M. Di Jiang, P.P. Xu, W. Chen, Q.Q. Ni, G.M. Lu, L.J. Zhang, Coronavirus Disease 2019 (COVID-19): A Perspective from China, Radiology. (2020) 200490. https://doi.org/10.1148/radiol.2020200490.

[8] I.D. Apostolopoulos, T.A. Mpesiana, Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks, *Phys. Eng. Sci. Med. (2020) 1–8*. https://doi.org/10.1007/s13246-020-00865-4.

[9] M.J. Horry, S. Chakraborty, M. Paul, A. Ulhaq, B. Pradhan, X-Ray Image based COVID-19 Detection using Pre-trained Deep Learning Models, (2020)

[10] G. Gaál, B. Maga, A. Lukács, Attention U-Net Based Adversarial Architectures for Chest X-ray Lung Segmentation, (2020) 1–7.

[11] Y. Celik, M. Talo, O. Yildirim, M. Karabatak, U.R. Acharya, Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images, *Pattern Recognit. Lett. 133 (2020) 232–239.* https://doi.org/10.1016/j.patrec.2020.03.011.

[12] L.J. Muhammad, M.M. Islam, S.S. Usman, S.I. Ayon, Predictive Data Mining Models for Novel Coronavirus (COVID-19) Infected Patients' Recovery, *SN Comput. Sci. 1 (2020) 206.* https://doi.org/10.1007/s42979-020-00216-w.

[13] T. Mahmud, M.A. Rahman, S.A. Fattah, CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization, *Comput. Biol. Med.* 122 (2020) 103869. https://doi.org/10.1016/j.compbiomed.2020.103869

[14] E. Luz, P.L. Silva, R. Silva, L. Silva, G. Moreira, D. Menotti, Towards an Effective and Efficient Deep Learning Model for COVID-19 Patterns Detection in X-ray Images, (2020) 1–10.

[15] M. Rahimzadeh, A. Attar, A New Modified Deep Convolutional Neural Network for Detecting COVID-19 from X-ray Images, (2020). http://arxiv.org/abs/2004.08052.

[16] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, G.J. Soufi, Deep-COVID: Predicting COVID-19 From Chest X-Ray Images Using Deep Transfer Learning, (2020). http://arxiv.org/abs/2004.09363.

[17] N.S. Punn, S. Agarwal, Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using fine-tuned deep neural networks, (2020). http://arxiv.org/abs/2004.11676.

[18] A.I. Khan, J.L. Shah, M. Bhat, CoroNet: A Deep Neural Network for Detection and Diagnosis of Covid-19 from Chest Xray Images, (2020). http://arxiv.org/abs/2004.04931.

[19] L. Wang, A. Wong, COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images, (2020). http://arxiv.org/abs/2003.09871.

[20] A.Narin, C. Kaya, Z. Pamuk, Department of Biomedical Engineering, Zonguldak Bulent Ecevit University, 67100, Zonguldak, Turkey., ArXiv Prepr. ArXiv2003.10849. (2020). https://arxiv.org/abs/2003.10849.

[21] Tarawneh, O., Otair, M., Husni, M., Abuaddous, H. Y., Tarawneh, M., & Almomani, M. A. (2022). Breast Cancer Classification using Decision Tree Algorithms. International Journal of Advanced Computer Science and Applications, 13(4).