The 4th International Conference on Emerging Data and Industry 4.0 (EDI40)
March 23-26, 2021, Warsaw, Poland

Using Deep Learning Model for Adapting and Managing COVID-19 Pandemic Crisis

Dr. Mohammad Alodat*

Sur University College,440,Sur 411,Sultanate of Oman

Abstract

The purpose of current paper is to create a smart and effective tool for telemedicine to early detect and diagnose COVID-19 disease and therefore help to manage Pandemic Crisis (MCPC) in Sultanate of Oman, as a tool for future pandemic containment. In this paper, we used tools to create robust models in real-time to support Telemedicine, it is Machine Learning (ML), Deep Learning (DL), Convolutional Neural Networks using Tensorflow (CNN-TF), and CNN Deployment. These models will assist telemedicine, 1) developing Automated Medical Immediate Diagnosis service (AMID). 2) Analysis of Chest X-rays image (CXRs). 3) Simplifying Classification of confirmed cases according to its severity. 4) Overcoming the lack of experience, by improving the performance of medical diagnostics and providing recommendations to the medical staff. The results show that the best Regression among the five Regression models is Random Forest Regression. while the best classification among the eight classification models and Recurrent Neural Network using Tensorflow (RNNTF) is Random Forest classification, and the best Clustering model among two Clustering models is K-Means++. Furthermore, CNN-TF model was able to discriminate between those with positive cases Covid-19 and those with negative cases.

Keywords: COVID-19;Sultanate of Oman;Machine Learning;Deep Learning;Convolutional Neural Networks;Tensorflow

1. Introduction

Coronavirus is continuing its spread across the world. Coronavirus family is a group of microorganisms which infect the respiratory tract and digestive system [1]. It has many sub types including, but not limited to, SARS-CoV,

* Corresponding author. Tel.: +96825565590, +96825565064; fax: +96825565555.
E-mail address: dr.maalodat@suc.edu.om, dr.maalodat@gmail.com

SARS-CoV-2, MERS-CoV, covid-19. Covid-19 outbreak spread quickly across the globe after the first case reported at the end of December 2019 and declared a global pandemic by the World Health Organization (WHO) on 11 march 2020 [2][3]. The outbreak of COVID-19 virus has affected human's health as well as global economy.

The aim of this research is to pandemic state in the future on specific area and its severity on individual level. The predictions will be in terms of the rate of viral spread in cites according to previous pandemic state. Anticipate who is at risk of death and equip to save their lives. Predict and Categorize patients according to the degree of severity of symptoms of the infection that affects the Omani society, from minor to serious. Helping to develop a medical diagnosis through two angles: the first angle in case a person is already infected with the virus, the model will predict if the patient will suffer minor or severe illness then recommend how the case will be treated, either stay at home, or being admitted to a health care facility. The second angle helps in improving the analysis of chest X-rays and articulate an accurate diagnosis of the disease, as it is difficult to differentiate between types of respiratory tract infection diseases, beside that manual diagnosis is complex, time-consuming and inaccurate.

2. Related Work

Deep Learning began in the early 1960s and has affected nonlinearity, processing capacity and enhanced accuracy, debriefing, disease classification and analysis of medical-related images [4]. Regression models are the individual values and mathematical collections of processes used to estimate or forecast the target based on independent and non-independent variables [5]. It was used to predict the number of confirmed cases through future days. Classification model is an example of pattern recognition belongs to the category of supervised machine learning algorithms. K-means++ was proposed in 2007 by David Arthur and Sergei Vassilvitskii, purpose to find cluster centers that minimize the intra-class variance. The goal of clustering is to group a collection of objects in such a way that objects in the same group (cluster) are more similar in nature, characteristics and need to be put together, than objects in other groups (clusters). K-Means++ Clustering, where k<n and k is the number of groups (clusters), n is the data points (objects) and attributes into k partition and be within a homogeneous and heterogeneous. Centroids is cluster centers through process re-calculated until there is no change to the centroids becoming centered cluster (considered settled) [6]. Tensor Flow and Keras are artificial neural network and ways to program deep learning models framework for high performance numerical computation. Keras was originally created and developed by Francois Chollet [7]. Tensor Flow developed by Google Brain Team in 2015 for internal use to google. Google announced TensorFlow 2.0 in June 2019, they declared that Keras is high-level API of TensorFlow [8]. Convolution Neural Networks (ConvNets or CNNs) are a category of deep learning techniques have the strong ability of nonlinear modeling. CNN is considered opaque meaning that, it is there is no clear and specific reason why a decision or prediction was made and it have internal structures and multistage operation [9].

3. Methodology

The following is used in this paper, Python libraries in Jupyter Notebook and Spyder as package NumPy and Pandas, because using them takes a few lines of codes and very accessible [10]. The dataset used in this research was collected from different sources as follows: 1) Wikipedia for analysis of COVID-19 cases. The included cases were during time from 24th February 2020 to 18th June 2020. 2) Questionnaire On 11.08.2020 was distributed on social media in the Sultanate of Oman, The number of participants reached 315, the questionnaire contains Symptoms and chronic diseases. 3) We got an open source dataset from Kaggle website composed of 3375 images of non-infected patients, 70 image of patients with Covid-19 virus, 2772 image for patients with pneumonia bacterial, and 1493 image of patients with pneumonia virus to train and test. 4) Eleven participants were randomly chosen from different areas in the sultanate of Oman who has respiratory symptoms with different medical diagnosis. Consent form taken and confidentiality of data was protected. Eligible participants have made Chest X-rays image (CXRs) to analyze and predict COVID-19.

In order to assist Manage COVID-19 Pandemic Crisis (MCPC) making medical correct and accurate decisions and facilitate dealing with the pandemic in the country, the workflow with the models used will be as follows: 1) entering the personal details, acute and chronic diseases: a) Regression Model used to predict the expected

increment of COVID-19 cases in the number of cases in the future in a specific city. b) Classification Model that predict if the patient will live or die and accordingly the decision of admission will be taken. c) Clustering Model used to categorize patients according to symptoms severity of COVID-19 cases, from Normal to Dangerous situation, to help health authorities to take suitable decision of home quarantine or admission to hospitals. 2) Chest X-rays will be analyzed through two methods: Convolution Neural Networks with Tensorflow (CNN-TF), and Recurrent Neural Network (RCNN).

4. Machine Learning

4.1. Regression models

Regression models are one of the most popular machine learning algorithms, it is fast, simple, primary and easy to use [11]. In this research regression models were used to predict COVID-19 cases after 18th June 2020 until 31th Dec 2020, which means it predicts through 196 days. The regression models used in this paper were: Multiple Linear Regression (MLR), Polynomial Regression (PR), Support Vector Regression (SVR), Decision Tree Regression (DTR) and Random Forest Regression (RFR) [12].

4.2. Classification Model

Classification is used to discover the category and performance comparison of linear and Gaussian kernel implementations unlike regression models where you predict a continuous number [13]. In this research Classification Models were used to diagnose the patient's condition before being infected with COVID-19, by predict death or survival based on the symptoms. Choosing the best classifier by evaluating the accuracy of the classification among the following: Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (K-NN), Kernel SVM, Naive Bayes (NB), Decision Tree Classification (DTC) and Random Forest Classification (RFC).

4.3. Clustering Model

Clustering is one of the approaches of unsupervised learning, by clustering events into groups, unexpected things can suddenly pop up when used for statistical data mining and analysis [14] [15]. In this research using Clustering Models to categorize patients on groups according to the severity of COVID-19 symptoms. By the choosing the most efficient model among two clustering models as follows: K-Means++ and Hierarchical Clustering. Clustering analysis ideal way is to determine the number of clusters (K) when observe the knee of a curve or elbow method in point graph by calculate the Within-Cluster-Sum-of-Squares (WCSS).

5. Deep Learning

5.1. Recurrent Neural Network (RNN)

We had built a Recurrent Neural Network using Tensorflow (RNNTF), consisting of: 1) the input layer and independent variables: The Age stage, clinical symptoms, rare symptoms and chronic diseases. 2) Output layer and dependent variable the patient who will survive or die. 3) Activation function used sigmoid in output layer and rectifier (ReLU) function in two hidden layers. 4) Epochs equal 100, which represent the number of iterations that will be used in the training process. And 5) selecting the optimizer called Adam optimization algorithm [16].

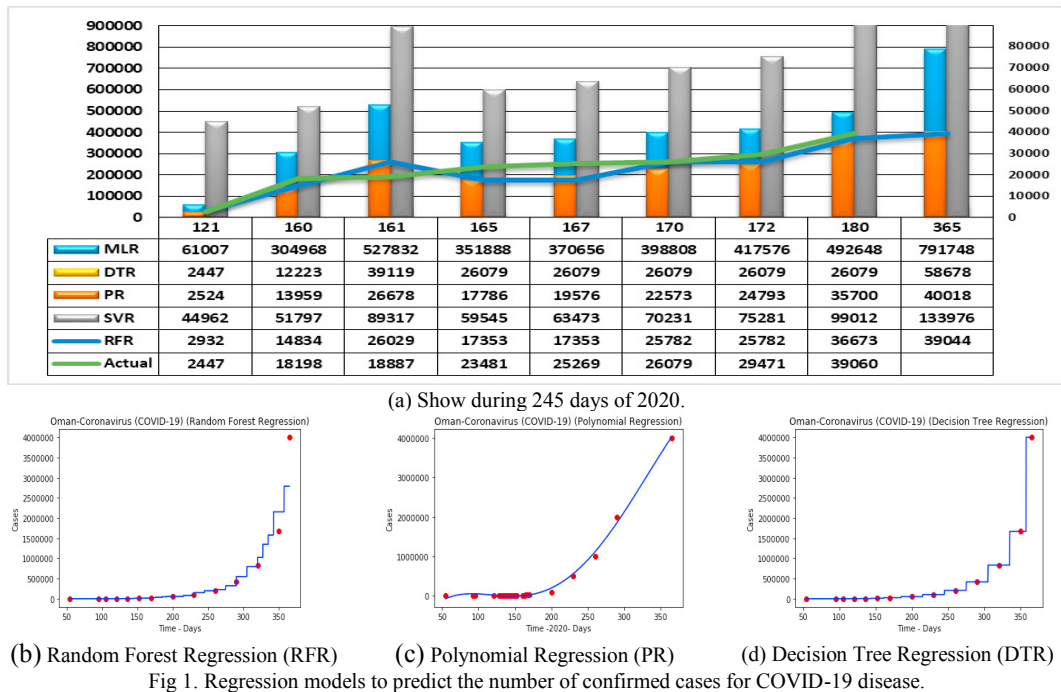### 5.2. Convolutional Neural Networks (Cnns)

Manual diagnosis of chest X-rays (CXRs) images is a complex and time-consuming work and lead to differences in diagnosis because pneumonia diseases have similar appearance in CXR images [17]. In order to overcome these problems, we used in this research Convolution Neural Networks (CNN) techniques that combine medical image

processing with computer vision. We had built two CNN using Recurrent Neural Network (RCNN) and Tensorflow (CNN-TF). These models were used to classify images to either negative case or positive cases, by passing images across successive filters. The Last layer of the model has Soft-Max classifier and four neurons, where one neuron for negative cases (normal), and three neurons for positive cases (pneumonia viral, pneumonia bacterial and Covid-19) [18].

6. Result and Discussion

6.1. Regression Models

Regression models were used to predict the number of confirmed cases after 18.6.2020, about 196 days of the year 2020. The most accurate predictive model is the Random Forest Regression (RFR) is (91%). Figure one shows two curves, one of them related to Random Forest Regression (RFR) model, and the other represents actual positive cases and both curves coincides to a large degree, see Figure 1 (a). For example the model predicted the number of confirmed cases in the 170th day to be 25,782 patients, and the actual positive cases reported for that day was 26079 which is so close to each other. The figure shows the distribution of samples on the different models, as shown in Figure 1 (b, c, d).

	121	160	161	165	167	170	172	180	365
MLR	61007	304968	527832	351888	370656	398808	417576	492648	791748
DTR	2447	12223	39119	26079	26079	26079	26079	26079	58678
PR	2524	13959	26678	17786	19576	22573	24793	35700	40018
SVR	44962	51797	89317	59545	63473	70231	75281	99012	133976
RFR	2932	14834	26029	17353	17353	25782	25782	36673	39044
Actual	2447	18198	18887	23481	25269	26079	29471	39060	

(a) Show during 245 days of 2020.

(b) Random Forest Regression (RFR)   (c) Polynomial Regression (PR)   (d) Decision Tree Regression (DTR)

Fig 1. Regression models to predict the number of confirmed cases for COVID-19 disease.

6.2. Classification and RNNTF Models

The best performer was proved to be Random Forest classification, it has the highest accuracy from among the Classification Models, scored 90% correct predictions, while Receiver Operating Characteristic Curve (ROC) achieved 64%, Precision scored 91%, Recall scored 99%, and finally the F1 score is equals 94%, as shown in Table 1.

Table 1. Assess performances RNNTF Models and Classification Models.

Models	Accuracy	Precision	Recall	MSE	F1 score	ROC
SVM	0.89	0.89	0.99	0.11	0.94	0.59
Kernel SVM	0.87	0.88	0.99	0.13	0.93	0.54
K-NN	0.86	0.87	0.99	0.14	0.93	0.49
LR	0.89	0.89	0.99	0.11	0.94	0.59
DTC	0.86	0.94	0.94	0.14	0.92	0.62
NB	0.86	0.87	0.99	0.14	0.93	0.49
RFC	0.90	0.91	0.99	0.10	0.94	0.64
RNNTF	0.79	0.87	0.89	0.21	0.88	0.56

6.3. Clustering Models

After using K-Means++ Clustering , the results of the analysis revealed that COVID-19 positive cases are divided into four groups according to the severity of disease: No symptoms, stable Condition, unstable Condition and Dangerous condition, as shown in Figure 2.
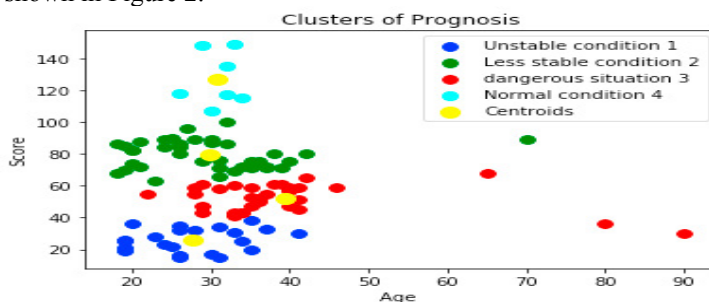

Fig 2. K-Means++; Distribute patients according to the severity of disease.

6.4. Convolutional Neural Networks (CNN)

The number of volunteer patients who sent their chest X-rays (CXRs) reached 11, as shown in Figure 3, the best model was CNN-TF. It was able to assess the following: a) Model Accuracy over the time increased, (accuracy= 0.93625 and validation= 0.782. b) Model Loss will be decreased, (loss accuracy =0.0659, loss validation= 0.9243, and x-axis the epoch= 100).
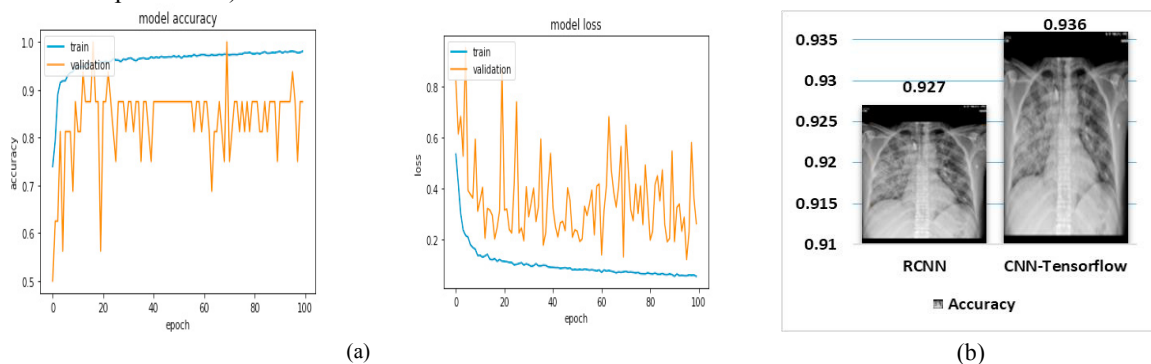

(a)   (b)

Fig 3. Evaluation Performance CNN-TF and RCNN Model, (a) Loss and Accuracy Curves for CNN-TF, (b) The comparison between RCNN and CNN-TF.

We were able to distinguish between 11 patients with pneumonia virus, pneumonia bacterial and Covid-19 and those without the disease. CNN-TF matched the actual diagnosis except for one patient who volunteered did not have Covid-19 disease with the actual diagnosis, but it was not-infection, as shown in Figure 4.
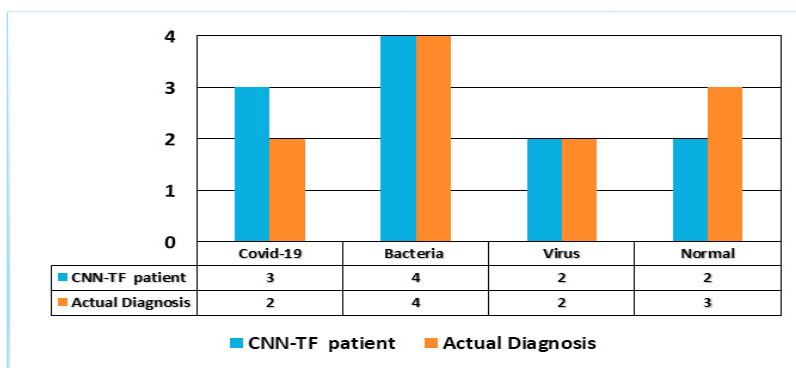
Fig 4. Distinguish patients by CXRs analysis using CNN-TF and Actual Diagnosis.

7. Conclusion

Accurate prediction in the event of a Covid-19 disease helped avoid confusion when diagnosis and gain time, and relieve stress on healthcare systems.

In this paper we presented future recommendations to Medical decision making for each stage of the Management of COVID-19 Pandemic Crisis (MCPC) in Sultanate of Oman.

By Random Forest Regression (RFR) model which predicts the number of future covid-19 case in a specific geographic area. b) Random Forest classification model was able to predict who will be at risk of death and to provide appropriate healthcare assistance, c) K-Means++ Model was the best in categorizing patients on groups according to the severity of symptoms, the categories were four groups from no symptoms to severe condition.

Convolution Neural Networks with Tensorflow (CNN-TF) and RCNN was used to analyze Chest X-rays images to provide differential diagnosis to respiratory diseases, of them the best model was CNN-TF that presented the most accurate medical diagnosis.

Our models helped in early detection of Covid-19, give a better insight of the hospital needs before catching the disease, reduce severity and complexity of medical treatment, improve medical diagnosis, and saved time. We were able to analyze chest X-rays image (CXRs), therefore differentiate between those infected with the pneumonia virus, pneumonia bacterial and Covid-19 and not infected.

In future studies, we will develop and improve the CNN architecture and to overcome its defects by using Deep transfer learning (DTL) from pre-trained network on ImageNet of convolutional neural networks (CNN) and TensorFlow 2.0 with tf.keras.

ACKNOWLEDGMENT

I would like to thank the management of Sur University College for the continued support and encouragement to conduct this research.

Compliance with Ethical Standards

The author declares that there is No conflict of interest and no fund was obtained. Institution permission and Institution Review Board (IRB) was taken from Sur University College. Written informed consent was obtained from all individual participants included in the Study. The researcher explained the purpose, and the possible outcomes of the research. Participation was completely voluntary and participants were assured that they have rights to withdraw at any time throughout the study and non-participation would not have any detrimental effects in terms

of the essential or regular professional issues or any penalty. Also participants were assured that their responses will be treated confidentially.

References

[1] Mavrikou, S.; Moschopoulou, G.; Tsekouras, V.; Kintzios, S. Development of a Portable, Ultra-Rapid and Ultra-Sensitive Cell-Based Biosensor for the Direct Detection of the SARS-CoV-2 S1 Spike Protein Antigen. Sensors 2020, 20, 3121.

[2] Chatterjee, A.; Gerdes, M.W.; Martinez, S.G. Statistical Explorations and Univariate Timeseries Analysis on COVID-19 Datasets to Understand the Trend of Disease Spreading and Death. Sensors 2020, 20, 3089.

[3] Lvov, D. K., et al. "Etiology of epidemic outbreaks COVID-19 in Wuhan, Hubei province, Chinese People Republic associated with 2019-nCoV (Nidovirales, Coronaviridae, Coronavirinae, Betacoronavirus, Subgenus Sarbecovirus): lessons of SARS-CoV outbreak." https://virusjour. elpub. ru/jour/article/download/261/157 (2020).

[4] Militante, Sammy V., Nanette V. Dionisio, and Brandon G. Sibbaluca. "Pneumonia Detection through Adaptive Deep Learning Models of Convolutional Neural Networks." 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC). IEEE, 2020.

[5] Pandey, Gaurav, et al. "SEIR and Regression Model based COVID-19 outbreak predictions in India." arXiv preprint arXiv:2004.00958 (2020).

[6] Arthur, David, and Sergei Vassilvitskii. k-means++: The advantages of careful seeding. Stanford, 2006.

[7] Chollet, Francois. "Building autoencoders in keras." The Keras Blog (2016).

[8] Nguyen, Giang, et al. "Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey." Artificial Intelligence Review 52.1 (2019): 77-124.

[9] Liu, Tianyi, et al. "Implementation of training convolutional neural networks." arXiv preprint arXiv:1506.01195 (2015).

[10] Pereira, Rodolfo M., et al. "COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios." Computer Methods and Programs in Biomedicine (2020): 105532.

[11] Alodat, Mohammad, and Iyas Abdullah. "Surveillance Rapid Detection of Signs of Traffic Services in Real Time." Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 10.2-4 (2018): 193-196.

[12] Li, Na, and Xiang Yu. "Outbreak and Regression of COVID-19 Epidemic among Chinese Medical Staff." Risk Management and Healthcare Policy 13 (2020): 1095.

[13] Jaiswal, Aayush, et al. "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning." Journal of Biomolecular Structure and Dynamics (2020): 1-8.

[14] Carrillo-Larco, Rodrigo M., and Manuel Castillo-Cara. "Using country-level variables to classify countries according to the number of confirmed COVID-19 cases: An unsupervised machine learning approach." Wellcome Open Research 5.56 (2020): 56.

[15] Alodat, Mohammad. "Predicting Student Final Score Using Deep Learning." Advances in Computer, Communication and Computational Sciences. Springer, Singapore, 2020. 429-436.

[16] Zaremba, Wojciech, Ilya Sutskever, and Oriol Vinyals. "Recurrent neural network regularization." arXiv preprint arXiv:1409.2329 (2014).

[17] Jha, Sujeet, et al. "Prevalence of flu-like symptoms and COVID-19 in healthcare workers from India." J Assoc Physicians India 68.7 (2020): 27-9.

[18] Sim, Yongsik, et al. "Deep convolutional neural network–based software improves radiologist detection of malignant lung nodules on chest radiographs." Radiology 294.1 (2020): 199-209.