COVID-19 diagnosis from chest CT scan images using deep learning

Raghad ALASSIRI[1], Felwa ABUKHODAIR[2], Manal KALKATAWI[2], Khalid KHASHOGGI[3], Reem ALOTAIBI[2]

[1] King Abdulaziz and His Companions foundation for Giftedness and Creativity (Mawhiba), Saudi Arabia

[2] Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

[3] Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

raghad.alassiri@hotmail.com, fabukhodair@kau.edu.sa, mkalkatawi@kau.edu.sa, kkhashoggi@kau.edu.sa ralotibi@kau.edu.sa

Abstract: Coronavirus disease 2019 (COVID-19) has caused nearly 600 million individual infections worldwide and more than 6 million deaths were reported. With recent advancements in deep learning techniques, there have been significant efforts to detect and diagnose COVID-19 from computerized tomography (CT) scan medical images using deep learning. A retrospective study to detect COVID-19 using deep learning algorithms is conducted in this paper. It aims to improve training results of pre-trained models using transfer learning and data augmentation The performance of different models was measured and the difference in performance with and without using data augmentation was computed. Also, a Convolutional Neural Network (CNN) model was proposed and data augmentation was used to achieve high accuracy ratios. Finally, designed a website that uses the trained models where doctors can upload CT scan images and get COVID-19 classification (https://covid-e46e8.web.app/) was designed. The highest results from pre-trained models without using data augmentation were for DenseNet121, which was equal to 81.4%, and the highest accuracy after using the data augmentation was for MobileNet, which was equal to 83.4%. The rate of accuracy improvement percentage after using data augmentation was about 3%. The conclusion was that data augmentation could improve the accuracy of COVID-19 detection models as it increases the number of samples used to train these models.

Keywords: COVID-19, deep learning models, CT scan, data augmentation, transfer learning.

1. Introduction

Coronavirus is a disease that is rapidly spreading. Since its start, the virus has infected nearly 600 million individuals around the world and more than 6 million deaths were reported (COVID live, 2022).

There are two tests to diagnose COVID-19 infection that are currently approved. The first one is PCR. This molecular test is a diagnostic test that uses a laboratory technique called polymerase chain reaction (PCR) to detect the genetic material of the COVID-19 virus. In order to perform the test, patient's fluids are collected from a swab of the nose or throat, or from saliva. It takes 3 to 6 hours for the results to appear, and in some cases, it may take up to 48 hours. Molecular tests are considered very accurate when performed correctly, yet it appears that some infections are missed.

The second test is called Antigen Test. The Antigen Test quickly detects some of the virus proteins, using a swab from the nose or throat to obtain a sample of fluid. Antigen testing is considered a more practical solution given the low cost and rapid results. A positive antigen test result is regarded to be highly accurate, however there is an increased percentage of a false negative result. Compared to the molecular test, the antigen test is less sensitive (Mak et al., 2020; Robert, 2020).

Several studies have shown that COVID-19 infection can be detected through CT scan images. CT scans are considered to be the primary imaging modality for investigating cases with suspected COVID-19 infection (Mungmunpuntipantip & Wiwanitkit, 2020; Arnoldi et al., 2009). CT scan images are abundant in most hospitals and are more widely used compared to X-rays for diagnosing COVID-19. CT images with COVID-19 have commonly known manifestations such as infiltrates, ground-glass opacities (GGO), and sub-segmental consolidation (Yurdaisik, 2020).

https://doi.org/10.33436/v32i3y202205

Given the large number of cases, CT image analysis and reporting by a radiologist may take a long time. In seeking faster and more accurate results, there have been many efforts to implement AI techniques to automatically diagnose COVID-19. Such AI tools can be useful to help expedite the reporting process of a large number of cases by suggesting a fast and accurate diagnosis. Nevertheless, the limited availability of public dataset for CT images remains one of the biggest challenges for building accurate AI tools.

The contribution of this work is presented in two folds. The first one is the comparative study between prominent CNN pre-trained models like MobileNet, ResNet50, DenseNet121 and DenseNet169. These models are pre-trained on more than 1 million images from different open-source databases and can classify more than 1000 different objects such as animals, cars, nature, etc. these models were not trained on COVID-19 images. Transfer learning was used to train these models on COVID-19.

Due to the low availability of COVID-19 CT images, the improvement of training results for pre-trained models using data augmentation and the measurement of accuracy with and without using data augmentation are intended. The second contribution is the proposal of a standard CNN model that was trained from scratch using COVID-19 dataset only with data augmentation techniques again to increase the number of images in the training set to improve the model's performance.

To contribute to the field of diagnosing COVID-19 from CT scan images using deep learning, some related work that will be presented was studied.

He et al. (2020) suggests the Self-Trans approach, which combines both contrastive self-supervised learning and transfer learning approaches to learn robust and unbiased representations of features to reduce the risk of overfitting. The proposed approach achieves F1-score of 85% and an Area Under the Curve (AUC) of 94% in diagnosing COVID-19 from CT scans.

Pathak et al. (2020) used transfer learning to tune the initial set of deep learning parameters. The ResNet-50 network was adopted by them. The proposed approach achieved training classification accuracy up to 96.22% and testing classification accuracy up to 93.01%.

To address the lack of publicly available dataset, Maghdid et al. (2021) assembled a comprehensive dataset of both x-ray and CT scan images. Moreover, a convolution neural network (CNN) was proposed and AlexNet, a pre-trained network was modified.

The experiments' results showed that the adopted models achieved 98% accuracy using a pre-trained network and 94.1% accuracy using the modified CNN model.

The previous reviews found that using deep learning models to diagnose people with COVID-19 using CT and X-rays may achieve significant results, and hence the motivation to contribute to this challenging study.

In Section II of this paper the materials and methods are discussed while results are presented in detail in Section III. A discussion about the study is presented in Section IV. Finally, the whole study is concluded in Section V.

2. Materials and methods

In this section a retrospective deep learning study to detect COVID-19 from CT scan images is presented. In this study, the role of deep learning, CNN and data augmentation in automating the detection of COVID-19 from CT-scan images are explored.

Firstly, the dataset that is used in the experiments is presented; secondly, the experiment settings is discussed, followed by the evaluation and results.

2.1. COVID-19 CT scan images dataset

The images are downloaded from GitHub2 which were originally assembled from COVID19-study papers from various sources such as medRxiv, bioRxiv, NEJM, JAMA, Lancet, etc. All copyrights are reserved to the original authors and publishers. All CT scan images were approved and consented by patients. Further information related to patients' details can be found in CVID-CT-MetaInfo.xlsx.

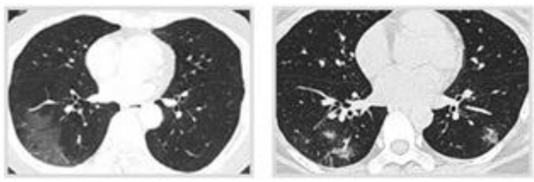

Figure 1. CT scan of two lungs infected with COVID-19 from the dataset

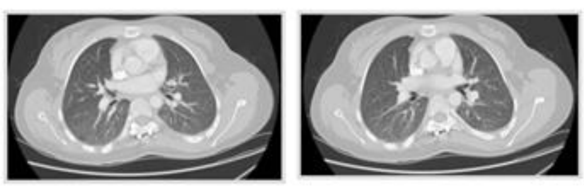

Figure 2. CT scan of two non-COVID-19 lungs from the dataset

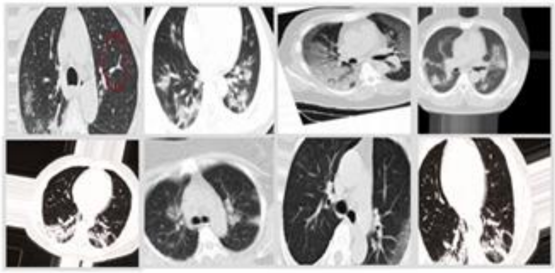

Figure 3. CT scans of people infected with COVID-19 from the dataset after using data augmentation

Figure 4. CT scans of people not infected with COVID-19 from the dataset after using data augmentation

The COVID-19 CT scan images dataset is an open access database that contains 349 COVID-19 CT scans images, displayed in Figure 1 and 397 CT scan images of non-COVID-19, shown in Figure 2. The dataset was collected between January and April 2020 The CT scan images were resized to 224x224. Then the dataset was divided into 80% for training, of which 10% was used for validation purposes. The remaining 20% of the dataset was used for testing. It is worth mentioning here that the non-COVID data were not images of normal lungs. On the contrary, it contains images of lungs with other chest diseases.

2.2. Experimental settings

In the first half of the study pre-trained CNN models, such as MobileNet, ResNet50, DenseNet121, and DenseNet169 were used. These models have been previously trained on different daily objects such as dog, cat, chair, keyboard, mouse, car, table and so on. These models are adopted and transfer learning is used to re-train the models using COVID-19 images to classify CT medical images into COVID-19 and non-COVID-19.

So, firstly various CNN pre-trained models using transfer learning to train the models on COVID-19 images were experimented with. After a few rounds of experiments with different parameters and different models, the ones that achieved higher results for classifying COVID-19 images were chosen and the ones with poor results were eliminated.

In order to increase the performance, data augmentation was used. This is a known technique that is used to increase the diversity of available data for training models and to eliminate the need for collecting and annotating new data. Examples of data augmentation methods include cropping, padding, and horizontal flipping. These methods are most commonly used to train large deep neural networks (Shorten and Khoshgoftaar, 2019; Seita, 2022). Figure 3 and Figure 4 show samples of data generated by using data augmentation and after being resized to 224×224 for infected and non-infected patients, respectively.

Then the results were compared between pre- and post-augmentation. Secondly, a standard CNN model architecture that was trained from scratch using COVID-19 datasets was proposed and the model was trained to classify COVID/ non-COVID images.

After preparing the dataset to train the proposed computational models, building and training those deep learning-based models were started.

Proposed CNN Model: At first, a deep learning model based on standard CNN was proposed and trained on the previously explained dataset; but that did not work well because of the small training data. Thus, the data augmentation technique was applied, and the augmented data was used to train the CNN model.

The proposed architecture of the CNN model consists of four convolutional layers in which each layer is followed by a max-pooling layer After that, a dense layer was added, then a flatten layer, and finally two more dense layers for COVID-19 classification.

Pre-trained Models (Transfer Learning): The experiments were conducted on different predefined models' architecture, namely: ResNet-50, MobileNet, DenseNet-121, and DenseNet-169. A set of parameters of the pre-trained models were configured, such as batch size, the number of training epochs, and the learning rate, which are tuned to 32, 100, and 3e-4, respectively.

2.3. Evaluations

To evaluate our proposed models' efficacy, three widely used measures were used for binary classification problems, accuracy, sensitivity, and specificity. The accuracy is used to measure the number of correctly predicted data points out of all the data points 6 (Iwendi et al., 2020). Sensitivity is used to measures the model's ability to predict true positive cases among all available class. Whereas specificity is the metric that computes the model's ability to detect true negatives of each available class (Kovács et al., 2021; Alsharif et al., 2020; Alsharif et al., 2020). Those three measurements can be calculated by Equations 1, 2 and 3 as shown below.

$$Accuracy = (True\ predictions)/(All\ predictions) = (Tp+Tn)/(Tp+Tn+Fp+Tn) \quad (1)$$

$$Sensitivity = (True\ positives)/(Actual\ positives) = Tp/(Tp+Fn) \quad (2)$$

$$Specificity = (True\ negatives)/(Actual\ negatives) = Tn/(Tn+Fp) \quad (3)$$

where Tp, Tn, Fp and Fn are the number of true positives, true negatives, false positives, and false negatives, respectively.

3. Experiments' results

The results of different models are shown in Table 1 and Table 2 with and without using data augmentation, respectively. DenseNet121 has the highest accuracy score in Table 1, which was 81.46%, while MobileNet achieved the highest accuracy score in Table 2, 83.44%.

Table 1. Models' Performance without Using Data Augmentation

	Accuracy	Sensitivity	Specificity
MobileNet	79.50%	79.01%	80.00%
ResNet50	79.58%	78.87%	80.28%
DenseNet121	**81.46%**	73.24%	88.75%
DeneseNet169	78.26%	81.48%	75.00%

Table 2. Models' Performance with using data augmentation

	Accuracy	Sensitivity	Specificity
MobileNet	**83.44%**	85.92%	81.25%
ResNet50	80.79%	81.69%	80.00%
DenseNet121	82.78%	87.32%	78.75%
DeneseNet169	78.15%	84.51%	72.50%

Figure 5 shows the difference between accuracy results with and without data augmentation. MobileNet has made the highest improvement in accuracy with data augmentation by about 4%. Also, there was a decrease of 0.1% in DeneseNet169. For the rest of the models, there was an increase of approximately 1%.

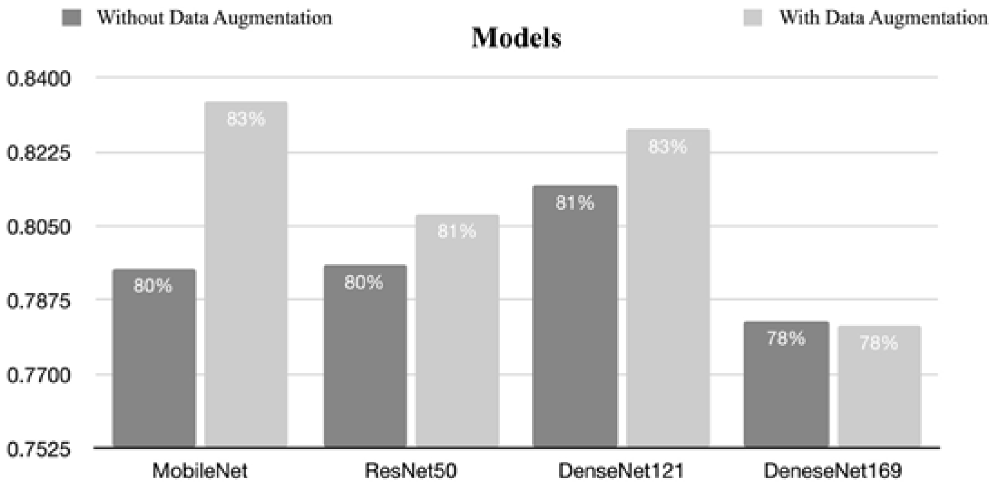

Figure 5. The differences between accuracy results with and without data augmentation

In our proposed CNN model that was built from scratch, the model achieved an accuracy of 94.2%, which is much better than the accuracy of the pre-trained models used in transfer leaning with an improvement that approximately equals 10% better classification results.

4. Discussion

As demonstrated by the results in the preceding section, utilizing data augmentation on a small number of samples would produce superior outcomes when training deep learning models than using the same samples without any data augmentation. The highest results that were achieved were 83.44% which is about 4% more than the results without using the data augmentation; that is clear proof of the hypothesis at the beginning of this research which is that the augmentation can be used in COVID-19 classification tasks to increase the estimation accuracy of the model.

Also, training a model from scratch using domain specific samples, even with the standard architecture, would gain much more accurate results than using transfer learning. In the proposed model, the trained parameters worked only on medical images while the pre-trained models were trained on diverse types of images.

This research helps develop an alternative test to the current tests to diagnose Coronavirus and speed up the diagnostic process. An effective website was further designed for COVID-19 classification. The website trained models where doctors can upload a CT scan of a patient's chest and classify his/her condition, whether she/he has COVID-19 as shown in Figure 6. (https://covid-e46e8.web.app/).

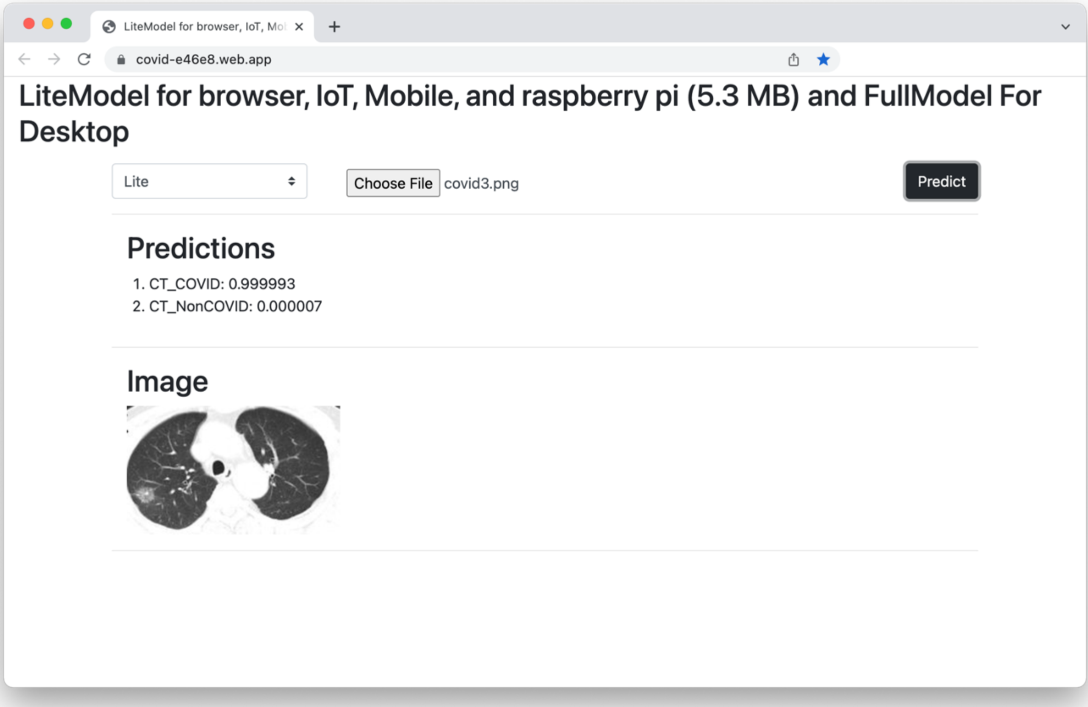

Figure 6. The main screen of the proposed application

This study, in conjunction with the results previously presented by Pathak et al. (2022) and Yang et al. (2020) indicates that the proposed CNN model that was built from scratch was able to reach an accuracy level of 94.2%. Pathak et al. (2022) used ResNet-50 network and transfer learning and achieved 93.01% accuracy. Yang et al. (2020) used data augmentation with CNNs to improve the classification results and achieved an accuracy of 84.7%, which is lower than the proposed model.

The main limitation of this study is that the models were trained on a small number of CT scan images. Another limitation is that clinical data was not considered during the training of the model to avoid misclassifying other conditions such as pneumonia.

5. Conclusion

In this research, several pre-trained models were trained to detect COVID-19 from CT scan images. Percentage variance was measured using data augmentation and without using data augmentation. For most of the models, an increase between 1% - 4% of accuracy after using data augmentation was achieved.

In addition, a CNN model was built from scratch and data augmentation was used to increase the number of data, and it gained an accuracy of 94.2%, which is much better than the pre-trained models due to the dedication of this model to be trained on medical images.

As for future studies the training of the model on larger datasets is considered, which could help increase the confidence about the results and make sure the proposed model was not overfitted. Also, as for future work, multiclassification might be considered, which is a more complicated problem as distinguishing between COVID-19 and pneumonia is hard to accomplish.

Future research could also consider the different types of machines that produce CT scan images. Improving the performance by categorizing the images in advance might be considered.

Funding

This research was partially supported by King Abdulaziz and His Companions Foundation for Giftedness and Creativity (Mawhiba).

REFERENCES

1. Alsharif, M. H., Alsharif, Y. H., Chaudhry, S. A., et al. (2020). Artificial intelligence technology for diagnosing COVID-19 cases: a review of substantial issues. *European review for medical and pharmacological sciences*, 24(17), 9226–9233. doi: 10.26355/eurrev_202009_22875.
2. Alsharif, M. H., Alsharif, Y. H., Yahya, K., et al. (2020). Deep learning applications to combat the dissemination of COVID-19 disease: a review. *European Review for medical and pharmacological sciences*, 24(21), 11455–11460. doi: 10.26355/eurrev_202011_23640.
3. Arnoldi, E. et al. (2009). Letters to the editor. *Tumori*, *95*(1), 131–131. doi: 10.1177/030089160909500127.
4. COVID live - *Coronavirus statistics - Worldometer (no date) Worldometers.info*. Available at: https://www.worldometers.info/coronavirus (Accessed: June 20, 2022).
5. He, X. et al. (2020). Sample-efficient deep learning for COVID-19 diagnosis based on CT scans. *bioRxiv*. doi: 10.1101/2020.04.13.20063941.
6. Iwendi, C. et al. (2020). COVID-19 patient health prediction using boosted Random Forest algorithm. *Frontiers in Public Health,* 8, p. 357. doi: 10.3389/fpubh.2020.00357.
7. Kovács, A. et al. (2021). The sensitivity and specificity of chest CT in the diagnosis of COVID-19. *European Radiology*, 31(5), 2819–2824. doi: 10.1007/s00330-020-07347-x.
8. Maghdid, H.S. et al. (2021). Diagnosing COVID-19 pneumonia from x-ray and CT images using deep learning and transfer learning algorithms. In Agaian, S. S. et al. (eds.) *Multimodal Image Exploitation and Learning 2021*. SPIE.
9. Mak, G. C. et al. (2020). Evaluation of rapid antigen test for detection of SARS-CoV-2 virus. *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology*, 129(104500), p. 104500. doi: 10.1016/j.jcv.2020.104500.
10. Mungmunpuntipantip, R. & Wiwanitkit, V. (2020). Chest computed tomography in children with COVID-19. *Pediatric radiology*, 50(7), p. 1018. doi: 10.1007/s00247-020-04676-3.
11. Pathak, Y. et al. (2022). Deep transfer learning based classification model for COVID-19 disease. *IRBM,* 43(2), 87–92. doi: 10.1016/j.irbm.2020.05.003.
12. Robert, S. (2020). Coronavirus antigen tests: quick and cheap, but too often wrong?. *Science*, New York, N.Y.. doi: 10.1126/science.abc9586.
13. Seita, D. (2022). *1000x faster data augmentation*. The Berkeley Artificial Intelligence Research Blog. Available at: https://bair.berkeley.edu/blog/2019/06/07/data_aug (Accessed: June 20, 2022).
14. Shorten, C. & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1). doi: 10.1186/s40537-019-0197-0.
15. Yang, X. et al. (2020). *COVID-CT-Dataset: A CT Scan Dataset about COVID-19*. arXiv [cs.LG]. Available at: http://arxiv.org/abs/2003.13865.
16. Yurdaisik, I. (2020). Effectiveness of computed tomography in the diagnosis of novel Coronavirus-2019. *Cureus*, 12(5), p. e8134. doi: 10.7759/cureus.8134.

***

Raghad ALASSIRI is a high school graduate and a student of the King Abdulaziz & His Companions Foundation for Giftedness & Creativity MAWHIBA, interested in computer science, Artificial Intelligence, and scientific research.

***

Felwa ABUKHODAIR is an Assistant Professor at the Information Technology Department of the Faculty of Computing and Information Technology of King Abdulaziz University, Jeddah, Saudi Arabia. Dr. Abukhodair received her PhD from Simon Fraser University, Vancouver, Canada. She is the director of the Center of Excellence in Smart Environment Research (CESER). Her research interests include Visual Analytics, Visualization, Machine Learning and Health Informatics.

***

Manal KALKATAWI is an Assistant Professor at King Abdulaziz University (KAU), Faculty of Computing and Information Technology, Information Technology Department. She completed her M.A. degree in 2011 and her Ph.D. degree in 2017 at King Abdullah University of Science and Technology (KAUST), in Computer Science focusing on Bioinformatics, Data mining and Machine/Deep learning. She has extensive experience in genome analysis, genomic signals recognition and genomic sequences data extraction and processing. She has worked on applying deep learning in Arabic sentiment analysis and recognition of breast cancer images.

***

Khalid KHASHOGGI is an Associate Professor at the King Abdulaziz University Hospital. He is also the former President of the Radiological Society of Saudi Arabia. He earned his Bachelor's degree in Medicine and Surgery (MBBS) from King Abdulaziz University. He completed his internship at the Royal Columbian Hospital in New Westminster, followed by a Diagnostic Radiology residency at the University of British Columbia. He subsequently completed his fellowship training in Musculoskeletal Radiology and Pediatric radiology at the University of British Columbia as well. Dr. Khashoggi has maintained a keen interest and involvement in all aspects of diagnostic imaging.

***

Reem ALOTAIBI is an Associate Professor at the Faculty of Computing and Information Technology of King Abdulaziz University, Jeddah, Saudi Arabia. Currently, she is the supervisor of the Information Technology Department. Dr. Alotaibi received her Ph.D. in Computer Science from University of Bristol, Bristol, U.K., in 2017. During 2017-2018 she was a Visiting Lecturer at the Intelligent Systems Laboratory of the University of Bristol. Her research interests include Artificial Intelligence, Machine learning, Data mining and Crowd management. Dr. Alotaibi's research has been funded by several sources in Saudi Arabia, including Deputyship for Research & Innovation, Ministry of Education, King Abdulaziz City for Science and Technology (KACST) and Deanship of Scientific Research (DSR), King Abdulaziz University.