COVID-19 detection using hybrid deep learning model in chest x-rays images

Shubham Mahajan1 | Akshay Raina2 | Xiao-Zhi Gao3 | Amit Kant Pandit1

1School of Electronics & Communication Engineering, Shri Mata Vaishno Devi University, Katra, J&K, India
2School of Electrical Engineering, Shri Mata Vaishno Devi University, Katra, J&K, India
3School of Computing, University of Eastern Finland, Kuopio, Finland

Abstract
The novel-corona-virus is presently accountable for 547,782 deaths worldwide. It was first observed in China in late 2019 and, the increase in number of its affected cases seriously disturbed almost every nation in terms of its economical, structural, educational growth. Furthermore, with the advancement of data-analytics and machine learning towards enhanced diagnostic tools for the infection, the growth rate in the affected patients has reduced considerably, thereby making it critical for AI researchers and experts from medical radiology to put more efforts in this side. In this regard, we present a controlled study which provides analysis of various potential possibilities in terms of detection models/algorithms for COVID-19 detection from radiology-based images like chest x-rays. We provide a rigorous comparison between the VGG16, VGG19, Residual Network, Dark-Net as the foundational network with the Single Shot MultiBox Detector (SSD) for predictions. With some preprocessing techniques specific to the task like CLAHE, this study shows the potential of the methodology relative to the existing techniques. The highest of all precision and recall were achieved with DenseNet201 + SSD512 as 93.01 and 94.98 respectively.

KEYWORDS
COVID-19, deep learning, medical images, object detection

1 | INTRODUCTION

As of October 25, 2021, a total of 2,194,566,780 cases of COVID-19 and 4,547,782 deaths have occurred because of the novel-SARS corona virus infection. It is also an un-deniable fact that these numbers are of critical concern to almost every nation. The origin of outbreak is though officially unknown but its symptoms were first medically observed in China in December, last year. People are reporting to be affected from this infection even without having direct contact with a COVID-19 patient or traveling to an infection-prone area. This means it is also spreading via means of indirect media, thereby making it an urgent matter to be solved by joint efforts of researchers, professionals in medical, radiology, or AI community around the globe. The United States of America stands on top of the table with the most confirmed affected cases of 4,850,114 and death count of 159,128. It is followed by other largely populated nations like Brazil, India, and Russia. It is worth mentioning that many firms, business-cells and start-ups around the globe have been affected thus resulting in a great loss on development of humankind.

The World Health Organization (WHO) confirms that early symptoms of COVID-19 involve difficulty in breathing, speech/movement loss and chest pain/pressure, while later symptoms like fever, dry cough, tiredness and aches/pains, diarrhea, sore throat, conjunctivitis, skin rash, or tooth/toe discoloration. There have also been a number of studies showing potential hallmarks in radiology-based scan-images which may aid in an effective diagnostic tool for COVID-19.
1.1 Detection methods adopted and chest imaging findings

Presently, one of the approved drugs-based detection methods is reverse transcription polymerase chain reaction (RT-PCR) it mostly requires few hours. But another method, the real time PCR (qPCR) has become the preferred method as it provides advantages including automation, more reliable instrumentation and higher-throughput. Samples may also be obtained using a variety of methods, including nasopharyngeal swab, sputum, throat swabs, deep airway material collected by suction catheter or saliva. Though countries are working consistently to meet with the increasing demand of these methods, undoubtedly issues with access to RT-PCR and related wait times still exist. It can be inferred from this, that an alternative approach to detection, which is swift, conveniently available and with competitive specificity and sensitivity is an urgent requirement. Also, over the past few months several major US radiology societies have come out with statements making it clear that CT should be used sparingly in COVID-19 detection and only when, it will impact management. Thus, is becomes essential that healthcare providers and AI researchers be familiar with the imaging features of the infection.

There are numerous studies on COVID-19 like which clearly state that the hallmarks of this infection are distribution of conspicuous ground glass opacities (GGO) and patchy shadows in peripheral region of lungs. A hazy shade of gray appears in case of GGO which appears like frosted glass of windows in winters. In severe cases, another visual, namely, the solid white consolidation (SWC) appear in combination of GGO or as sole marking. There have also been seen, the crazy paving patterns (CPP) contrary to the hazy GGO background. The first sign usually is GGO which is then followed by one or both of the others. Detection via chest CT is very sensitive for COVID-19 but non-acceptable specificity due to similarities in these visuals with chest scan images of other viral pneumonia cases like influenza and also in certain non-infectious lung disease.

2 LITERATURE REVIEW

For past few years, AI based detection models have been used extensively in medical and other applications. Massive screening programs, which is a requirement in present scenario, will be supported greatly if we slowly transition to AI based models. There have been many studies around the globe in past year making an attempt to address the task effectively. And most of these studies rely on use of deep-learning based techniques, particularly the use of convolutional neural networks (CNNs)-based classification and localization models.

For detection of COVID-19, one such work is done by Bo Kang et al. which uses the Inception-migration-learning model over 217 computed tomography scans. They achieved specificity of 0.805 and a sensitivity of 0.84 for validation with an accuracy of 83% after following random selection of the regions of interest (ROIs). Another renowned study among deep-learning researchers at this task, by Wang et al. is the proposal of the COVID-Net CNN architecture. This is also one of first open-source network for COVID-19 detection and the dataset they used, the COVIDx is also open-sourced. They managed to achieve comparable values of sensibility of 0.80, specificity of 0.889 and accuracy of 92.4%. In the same context, Showman et al. achieved F1 scores of 0.91 and 0.89 for COVID-19 and normal cases detection respectively in their comparison study, COVIDx-Net. This gives a controlled comparison of seven major architectures of deep CNNs namely, the VGG19, ResNet-V2, DenseNet121, Inception-V3, Xception, InceptionResNet-V2, and MobileNet-V2. Experimentally, using the metrics of precision, recall and so forth, they claimed that the VGG19 and the DenseNet201 outperforms others. Saiz and Inigo used the VGG16 with the SSD300 along with some pre-processing of the images and got to a specificity of 0.92 and a sensibility of 0.9492. Singh et al. proposed a deep learning pipeline utilizing chest x-ray images within which certain relevant features were extracted, selected and fed into a Hybrid Social Group Optimization algorithm. They managed to achieve a high classification accuracy of 99.65% using the support vector machine for classification. Makris et al. used a dataset comprising a mixture of images from COVID-19 infected, bacterial pneumonia and healthy individuals. They employed numerous classic CNN based architectures like MobileNetV2, Inception Net and so forth, along with transfer learning with the VGG16 and VGG19 obtaining an overall classification accuracy of 95%.

We also proposed a novel methodology on the task previously using the deconvolutional SSD with some improvisations and achieved competitive specificity of 0.9474 and an accuracy of 0.9597. This is considerable now that deep learning methods can get us good results, aid the presently-unbalancing management systems.

3 METHODOLOGY

First, in Section 3.1, we will establish the overview of the study, followed by discussions on various networks implemented as base networks for detection heads in Section 3.2–3.4. Afterwards, we will head to discussion on the SSD and the potential changes and pre-processing we implemented in Section 3.5 and 3.6. The dataset used; its split structure has been discussed in Section 4. Later, in Section 5, the analysis on the comparison of all models with the discussion on hyperparameter selection and dataset used will be done.
3.1 Detection methods adopted and chest imaging findings

There is a need of modeling of a base-network which is to be truncated with the detection heads like the Single Shot MultiBox Detector which is used in this study. The COVIDx-Net research compares seven separate classification networks for detection of COVID-19 utilizing chest-radiology images. Each deep neural network model is capable of analyzing normalized concentrations of x-ray image in order to identify patient status as a negative or a positive COVID-19 scenario. This study reported the comparative performance of various deep learning techniques with observations to give most accurate classification results of COVID-19 utilizing a small x-ray image dataset. They used one-hot encoding on the labels of the dataset and uniformly resized as pre-processing of the dataset. Then, the dataset was split as 0.8:0.2 as per Pareto Principle for training and testing respectively. They utilized the publicly accessible dataset of 50 x-ray images with 25 having each label namely, COVID-19 and NORMAL. Without implementing any data augmentation, they used the stochastic gradient descent (SGD) with learning rate of $e^{-3}$ and a batch size and number of epochs as 7 and 50 respectively.

After rigorous discussion on comparison of all the models using metrics of F1 score, precision, recall, and accuracy, they recommended to use the VGG19 and the DenseNet201 to identify health status of patients potentially being cases of COVID-19.

Therefore, in our study, we decided to compare the classifier networks namely, the VGG16, VGG19, ResNet101, and the DarkNet201 truncated to the SSD-layers. These base networks are allegedly one of the classic CNN-based deep neural networks which have proven their robustness for long. Studying all of them with another classic network for detection, the SSD is potentially highly contributing research to the AI-community.

3.2 The Visual Geometry Group (VGG16, VGG19)

The VGG16 was introduced by Karen and Andrew at the Oxford University Visual Geometry Group in 2014 and won first runner-up place in the ILSVRC 2014 challenge. It is 16 layers deep with the image input size of [224,224] RGB and classifies the images into either of 1000 object categories. It performed well on the ImageNet dataset. This network used smaller filter sizes of [3 × 3] unlike its predecessor, the Alex Net which used [11 × 11] filter. The VGG19 has 19 layers and hence is deeper than the VGG16. Also, it has a greater number of parameters to learn, thereby making it a bit expensive computationally.

3.3 The Residual Network (ResNet101)

Developed by K. He, X. Zhang, S. Ren and J. Sun, the deep convolutional residual network won ILSVRC 2015 image classification challenge as well as the Microsoft COCO 2015 detection/segmentation challenge. There are three realizations of ResNet namely, the ResNet50, ResNet101, and the ResNet152. A core thought in this study is the use of residual learning or implementation of “skip-connections” between convolutional blocks as depicted in Figure 1 as well. These skip connections greatly support the gradient flow by helping diminish it is vanish. This allows much deeper training, that is, use of much deeper networks. In image classification tasks, the residual networks hence have proven to be better than the VGG. The ResNet particularly, accomplishes 76.4% top-1 and 92.9% top-5 accuracy in 1-crop validation.

3.4 The Densely Connected Convolutional Network (DenseNet201)

Proposed by G. Huang, the Densely Connected Convolutional Network (DenseNet201) is 201 layers deep and is an improved model in terms of information flow between layers with a different connectivity pattern. The skip connections added by ResNet bypass the non-linear

---

**FIGURE 1** Shows the residual learning via skip connections
transformations with an identity function and the gradient directly flows through this function to earlier layers from later ones. The information flow is impeded by the straight-summation of identity function with the output. This issue has been tried to be solved in the DenseNet architecture. This study also encourages feature reuse and the parameters’ count has been reduced substantially.

3.5 | The Single Shot MultiBox Detector (SSD512)

The SSD is built on top of a base network and is truncated with some convolutional layers to the end and a series of progressively smaller convolution layers to allow predictions of detections at multiple scales, is added as in blue in Figure 2. For every feature map cell, a number of priors or default boxes of variable scales and aspect ratios are generated by regular tiling of the input map. These boxes are equivalent to the anchors in the study of Faster R-CNN. These priors are matched with the ground truth boxes using a threshold of Jaccard index and classified as negative or positive samples. To detect objects at different scales, this model combines multi-scale feature maps and default boundary boxes. At each feature map cell, offsets relative to default box shapes in cell, as well as the per-class scores are predicted, that indicate presence of class instance in each of those boxes. The non-maximum suppression (NMS) is used to post-process the predictions, and the final detection results are obtained. It is represented in blocks in Figure 2.

3.6 | Pre-processing and modifications

As explained in the article, some liberated electrons due to thermionic emission, get electrically attracted towards the anode. This collision onto the target (tungsten) results into an emission of photons in x-ray spectrum thereby forming the basis of x-ray image formation. But an important point to note here is that the filament gets heated (resulting into thermionic emission) as a result of flow of current. That means the visual measures of an x-ray image are directly affected by the voltage spikes and hence, all of the x-ray scans around the world cannot be synchronized in terms of these parameters. Another factor affecting the contrast of these images is the exposure time which refers to time interval during which the x-rays are formed.

In their study on improving contrast for images, Reza stated that, in x-ray imaging, a low-level exposure is maintained until the scanning process for the region/s of interest (ROI) is finished. Hence, the images so-obtained are often with low signal-to-noise ratio. Hence, it is evidently clear now that, all the x-ray images need to be unified in contrast. Therefore, the Contrast Limited Adaptive Histogram Equalization (CLAHE) is implemented before training, onto the input images’ dataset. In CLAHE, the contrast amplification is limited as a modification to the adaptive histogram equalization. This is helpful in reducing the above-mentioned issue of amplification of noise. It has proven to provide good results on medical images. The target image is divided into equally sized regions (non-overlapping). This algorithm applied to an x-ray image is shown in Figure 3.

It has been observed that applying CLAHE to images before training improves the model-feature measurements (accuracy, sensibility, specificity) by considerable amount.

Extensive data augmentation techniques like randomly flipping, cropping, and photometric distortion, also need to be used.
Particularly for the experiment with ResNet101, we also used the so-called prediction-modules. This is because the primary structure that is, SSD + ResNet101 is not by itself a big improvement over other detection network. Cheng-Yang Fu et al. in their study on DSSD\textsuperscript{19} have shown that adding a Prediction Module (PM) increases the performance. The feature extraction layers in SSD have to learn to generate maps representing spatial, semantic information as well as the right transformations as per the optimal convergence condition of the setup. Also, it has to undo previous transformations before selecting the best for a scale. But adding PMs to the network would now require for feature extraction layers to just learn representing information from image and PMs are now able to learn the transformations. Following this and as used in Reference 19, we use one residual block as mentioned in Figure 4 for individual prediction layer.

4 DATA AND TRAINING

It has been well established in Reference 19 that the visual findings in chest x-ray Images of cases of COVID-19 and other viral pneumonia like influenza are very similar. Therefore, to make the model more robust and better specificity we used an image dataset constituting images of both the classes. This way the model will result into less false positives while inference. Same has also been applied to the work in Reference 7. Specifically, we used the COVIDx Dataset\textsuperscript{20} which was proposed by Wang et al.\textsuperscript{7} as the direct source. Although, there are five different open-source chest radiography datasets as the constituents of this source. There are 473+ C: x-ray images of COVID-19 cases and number for images of pneumonia cases is higher than “required.” We merged this with Reference 21 for enlarging the dataset.

The split of the dataset was done as follows-

- Randomly select images from the class “pneumonia.”
- Balancing the number of images in both classes with a difference in several image instances is roughly not more than 10%.
• For the training and validation dataset, roughly $(78 \pm 2)\%$ of available COVID-19 x-ray images and $(73 \pm 2)\%$ of selected pneumonia x-ray images to be used.

• For testing dataset, roughly $20\%$ of available COVID-19 x-ray images and $(25 \pm 1.5)\%$ of available pneumonia x-ray images to be used.

This should be certainly observable that there are more images of class pneumonia in the test dataset. This is for getting better insights about the performance of the model or its performance metric measures on specificity and number of false positives it results to.

**FIGURE 5** Shows trade-off between the mean IoU and the number of anchors

**FIGURE 6** Shows the implementation of the study as a flow-diagram
FIGURE 7  Shows training loss and accuracy versus epochs for first 25 epochs in all four models individually. (A) Shows plot for DenseNet201+SSD512; (B) shows plot for ResNet101+SSD512; (C) shows plot for VGG16+SSD512; (D) shows plot for VGG19+SSD512
The training steps followed are:

- Implement the conventional SSD training approach with each of the base networks one after other, that is, pairing set of anchors to target ground truth boxes using Jaccard index, selecting non-matched samples such that their ration with matched is 3:1, and finally minimizing the joint localization and confidence losses.
- The batch size is 16/32 preferably.
- All of the trainings are done in two steps, with a step-decrease in learning rate.
- Use SGD with momentum.
- Give less weightage to the localization loss as a higher priority for such a model should be given to address categorization of a chest x-ray Image properly.
- Choose the anchor boxes based on the specific training data as discussed below.

We estimated the anchor boxes from the training data using the IoU distance metric and the number of anchor boxes were chosen empirically that is, using the measure of mean IoU of the boxes in each cluster via k-means clustering with IoU. The trade-off between the mean IoU and the number of anchors is shown in Figure 5. Empirically, we determined the optimal number of anchor boxes to be nine.

The flow of the implementation of this study can be interpreted from Figure 6 below.

The flow of loss and accuracy simultaneously versus number of epochs while training for first 25 epochs for each of the classifier networks along with SSD512 is shown in Figure 7. It can be noticed clearly that DenseNet201 converges better than other models while the ResNet101 is a close competitor.

For better comparative analysis, a plot for all of these four models can be studied from the plot in Figure 8. Second, Table 2 below provides an overlook into the hyperparameters selected within the methodology (Table 1).

![Training Loss vs Accuracy for first 25 epochs](image)

**FIGURE 8** Shows training loss versus accuracy for first 25 epochs in all four models

| Hyper parameter          | Value                                                                 |
|--------------------------|----------------------------------------------------------------------|
| Hyper parameter          | Value                                                                 |
| Data split               | (78% + 73%)/(20% + 25%) (train: COVID+pneumonia/test: COVID+pneumonia) |
| Batch size               | 16/32                                                                |
| Optimizer                | Stochastic gradient descent with momentum                            |
| Non-matched: matched     | 3:1                                                                  |
| Number of anchor boxes   | 9                                                                    |
TABLE 2 The confusion matrix

| Predicted positive | Predicted negative |
|--------------------|--------------------|
| True positive (TP) | False negative (TN) |
| False positive (FP) | True negative (FN) |

5 EVALUATION AND COMPARISON

For evaluating the performance of each deep-learning model in this study, we analyzed different metrics for comparing the robustness and potential of each. A confusion matrix, in Table 2 resulted as for use of cross validation estimator. The four possible outcomes namely, true positive (TP), which is a measure of the outcomes when the true class is positive and the model has correctly predicted; true negative (TN), which is the measure of the outcome when model correctly predicts a negative class; false positive (FP) which represents the outcomes when the true class is positive and the model fails to predict correctly; false negative (FN) refers to an outcome when the model fails to predict a negative class.

Values for all four types of outcomes was first achieved and different performance parameters as discussed below were used.

**Precision** represents the proportion of correctly predicted positives and total correct predictions.

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

**Recall** or **sensitivity** measures the proportion of correctly identified true positives and summation of incorrectly predicted negative values and true positives.

\[
\text{Recall} = \frac{TP}{TP + FP}
\]

TABLE 3 Comparative analysis

| Model                  | Precision | Recall     | F1 score  |
|------------------------|-----------|------------|-----------|
| VGG16 + SSD512         | 0.9243    | 0.949292   | 0.936629  |
| VGG19 + SSD512         | 0.9241    | 0.949331   | 0.936545  |
| ResNet101 + SSD512     | 0.9245    | 0.949485   | 0.936825  |
| DenseNet201 + SSD512   | 0.9301    | 0.949801   | 0.939847  |

FIGURE 9 Shows plot for recall versus false negative rate
FIGURE 10  Shows training RMSE versus loss for first 25 epochs in all four models individually. (A) Shows plot for DenseNet201+SSD512; (B) shows plot for ResNet101+SSD512; (C) shows plot for VGG16+SSD512; (D) shows plot for VGG19+SSD512
The values in Table 3 provide comparative analysis of these metrics for all of the models implemented in this study. Since these metrics are statistical measures for binary classifications only, the pneumonia and normal classes should be considered as negative.

6 | CONCLUSION

The infectious and swiftly spreading COVID-19 has shaken the world in terms of management, economic, and educational stability threatening lives of billions of people. In this work, we try to provide a comparative analysis of four different networks, namely, the VGG16, VGG19, ResNet101, and the DenseNet201 as base with SSD as the detection sub-structure. The results of our experimental study suggest the use of DenseNet as the best classifier to be used as base network with the SSD512 among others particularly for task of COVID-19 infection detection in chest x-ray Images. This demonstrates use of deep-learning techniques in domain of computer vision/medical imaging and hence claims as an appeal to all AI researchers around the globe to research, make modifications and take this to a “practical” phase for the swift stabilizing of the increasing cases of this highly spreading pandemic.

7 | FUTURE WORK

The infectious and swiftly spreading COVID-19 has shaken the world in terms of management, economic, and educational stability.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available on request from the corresponding author.

ORCID
Shubham Mahajan  https://orcid.org/0000-0003-0385-3933
REFERENCES
1. Use of chest imaging in COVID-19: a rapid advice guide. Geneva, World Health Organization; 2020. (WHO/2019-nCoV/Clinical/Radiology_imaging/2020.1). Licence: CC BY-NC-SA 3.0 IGO.
2. Song F, Shi N, Shan F, et al. Emerging 2019 novel coronavirus (2019-nCoV) pneumonia. Radiology. 2020;295:210-217.
3. Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan. China Jama. 2020;323:1061-1069.
4. Wang S, Bo Kang JM, Zeng X, Xiao M. A deep learning algorithm using CT images to screen for Corona virus disease COVID-19); 2020. medRxiv.
5. Wang L, Wong A. COVID-net a tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images; 2020. arXiv Preprint arXiv:2003.09871.
6. Hemdan EED, Shouman MA, Karar ME. COVIDX-net a framework of deep learning classifiers to diagnose COVID-19 in X-ray images; 2020. arXiv Preprint arXiv:2003.11055.
7. Saiz FA, Barandiarian I. COVID-19 detection in chest X-ray images using a deep learning approach. Int J Interact Multim Artif Intell. 2020:6. doi:10.9781/ijimai.2020.04.003(2):1–4.
8. Singh AK, Kumar A, Mahmud M, Kaiser MS, Kishore A. COVID-19 infection detection from chest X-ray images using hybrid social group optimization and support vector classifier. Cognit Comput. 2021 Mar;4:1-3.
9. Makris A, Kontopoulos I, Tserpes K. COVID-19 detection from chest X-ray images using deep learning and convolutional neural networks. Proceedings of the 11th Hellenic Conference on Artificial Intelligence; September 2, 2020:60-66.
10. Mahajan S, Raina A, Abouhawwash M, Gao X, Pandit AK. Covid-19 detection from chest X-Ray images using advanced deep learning techniques. Comput Mater Continua. 2021;70(1):1541-1556.
11. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition; 2014. arXiv Preprint arXiv:1409.1556.
12. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016.
13. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2017:4700-4708.
14. Liu W, Anguelov D, Erhan D, Christian S, Reed S, Fu CY, Berg AC. SSD: single shot multibox detector. ECCV, 2016.
15. Ren S, He K, Girshick R, Sun J. Faster r-cnn: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell; 2016:39(6):1137-1149.
16. https://www.radiologycafe.com/radiology-trainees/frcr-physics-notes/production-of-x-rays
17. Reza M. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. J VLSI Signal Process Syst Signal Image Video Technol. 2004;38:35-44.
18. Fu CY, Liu W, Ranga A, Tjagi A, Berg AC. Dssd: deconvolutional single shot detector. arXiv 2017; 2018. arXiv preprint arXiv:1701.06659.
19. Cellina M, Orsi M, Panzeri M, Byl G, Oliva G. Chest X-Ray findings and visual quantitative assessment of COVID-19 pneumonia; 2020. 10.21203/rs.3.rs-32480/v1.
20. https://github.com/lindawangg/COVID-Net/blob/master/docs/COVIDx.md
21. https://data.mendeley.com/datasets/2fxz4px6d8/4#folder-2eec69c7-9b6c-4864-b9d5-bd059e2dd24b
22. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Inf Process Manag. 2009;45(4):427-437.
23. Young GO. Synthetic structure of industrial plastics (Book style with paper title and editor). In: Peters J, ed. Plastics. Vol 3. 2nd ed. McGraw-Hill; 1964:15-64.

How to cite this article: Mahajan S, Raina A, Gao X-Z, Pandit AK. COVID-19 detection using hybrid deep learning model in chest x-rays images. Concurrency Computat Pract Exper. 2022;34(5):e6747. doi: 10.1002/cpe.6747