7-1-2023

Transfer Learning For Covid-19 Image Classification Using Lightweight Architectures: A Systematic Review

Rajesh Godasu

David Zeng

Kruttika Sutrave

TRANSFER LEARNING FOR COVID-19 IMAGE CLASSIFICATION USING LIGHTWEIGHT ARCHITECTURES: A SYSTEMATIC REVIEW

Rajesh Godasu
Dakota State University
Rajesh.godasu@trojans.dsu.edu

David Zeng
Dakota State University
David.zeng@dsu.edu

Kruttika Sutrave
Dakota State University
Kruttika.sutrave@trojans.dsu.edu

ABSTRACT

Covid-19 has had a far-reaching influence on almost every aspect of everyday life. Deep learning techniques utilizing Chest CT images are currently proving to be significant resources in detecting Covid-19 cases. While standard CNN models have been utilized by various studies, a research stream focusing on Lightweight architectures is emerging. These compact models facilitate bringing the Covid-19 classification on Tablets and mobile phones, supporting healthcare professionals to rapidly identify the Covid-19 positive cases. We identified that the most used LWAs are MobileNetV2 and MobileNet for Covid-19 MIC. Both feature extraction and fine-tuning are extensively studied TL types. While dataset scarcity still exists X-ray images are mostly preferred due to cost-effectiveness and portability over CT images. Finally, we propose a truncated version of an ensemble model to be the future direction of our study.

Keywords

Transfer Learning, Lightweight architectures, Covid-19, medical image classification, Ensembled models

INTRODUCTION

Covid-19 disease was declared a pandemic by WHO. The rampant and uncontrollable nature of this disease caught the immediate attention of researchers worldwide to slow its transmission (Noteboom et al., 2021). Application of Deep Learning (DL) models on chest X-rays is emerging as a cost-efficient method to quickly detect Covid-19 (S. Tewari et al., 2020a). Researchers are extensively employing Convolutional Neural Networks (CNNs), along with Transfer Learning (TL) techniques, to classify Covid-19 positive images (Sufian et al., 2020). CNNs and TL are proving very effective in Medical Image Classification (MIC) tasks (Litjens et al., 2017), especially in data-scarce situations.

While classic architectures such as AlexNet have been performing exceptionally well in MIC, implementing them on resource constraint devices such as mobile phones and tablets is not practical as they contain millions of parameters and require GPUs to train (Erickson et al., 2018). These deep and complex networks that require high energy consumption are being overused by researchers making them unaligned with the sustainable AI goals (Vinuesa et al., 2020) to reduce global warming and carbon footprint (Haque et al., 2020). These networks have disadvantages, such as training difficulties and poor applicability (Yu et al., 2020). Architectural improvements by simplifying CNNs to reduce computations, thereby reducing the high energy consumption and maintaining high accuracies of CNN models could help achieve sustainable AI goals (Haque et al., 2020).

The quest to improve CNN models' efficiency and reduce computations led to developing resource-constraint or lightweight and compact CNN models (Khan et al., 2019), including but not limited to MobileNet and ShuffleNet. Lightweight Architectures (LWA) are highly efficient neural networks trying to mimic the accuracies of traditional deep CNNs with lesser parameters (Iandola et al., 2016). Architectures with fewer parameters require less time to train and can achieve higher accuracy with more scalable training when compared to complex CNN models in distributed training cases (Iandola et al., 2016). These models can be easily embedded in mobile devices as edge AI applications and can perform computations offline easily (Godasu et al., 2020). It is widely known that using GPUs is not cost-effective at this time; using an LWA CNN mobile application that could provide competitive performance can help optimal healthcare to reach the underprivileged countries (*WHO | World Health Organization*, 2020.).

LWA CNNs are widely used in MIC tasks. However, little attention is paid to synthesizing the existing literature on LWAs in Covid-19 diagnosis. Many current reviews have focused on standard CNNs and their applications in Covid-19 classification

(F. Shi et al., 2020) (Bhattacharya et al., 2021). These studies either focus on the standard DL techniques or recommend more research on using TL techniques. With this inspiration, our study solely focuses on systematically reviewing the works of LWAs using TL techniques for Covid-19 detection. To the best of our knowledge, this is the first literature review focusing on LWAs in the Covid-19 response contributing to sustainable AI. The main research objectives of this study are 1. What prevalent Lightweight CNNs are used in Transfer Learning for Covid-19 Classification? 2. What are the challenges in current research publishing LWA research in Covid-19 image classification? 3. To identify LWA research themes in TL approaches, classification type, dataset sizes, and data modalities utilized in Covid-19 and 4. Suggest actionable future research directions to improve Lightweight models in Covid-19 classification and contribute to sustainable AI goals.

METHODOLOGY

Search Strategy and Data Sources

This study uses a systematic literature review approach proposed by (Liberati et al., 2009) to achieve the above-mentioned research goals. We performed a search from four digital databases Arxiv, IEEE explore, Web of Science, and PubMed. The search was restricted to articles from January 2020 to August 2022 as most of the Covid-19 research occurred during this period. The search query consisted of four groups. The first group of words represents the concept of TL, the second represents the medical images tasks, the third group includes Covid-19 synonyms and the fourth group is to capture the LWA models. Following are the inclusion criteria: 1. Articles that focused on only Lightweight CNN models and TL techniques for Covid-19 MIC tasks. 2. Comparative studies that use both LWAs and standard CNNs for their study. 3. Articles using Lightweight CNNs as backbone models that have a parameter size of fewer than 6 million parameters including Truncated standard models. 4. Articles in English only are considered.

Study Selection

A total of 280 records have been identified using PRISMA standards. Next, 85 duplicate articles were removed. The articles were then evaluated for eligibility using the criteria. 101 papers are selected for full-text evaluation after abstract and title screening. The final review of this study included 39 papers after 62 articles were excluded during the full-text review. The survey table and descriptive findings of this paper are illustrated using the final papers included in the research. To avoid any conflicts, two distinct researchers conducted the whole review procedure.

RESULTS AND DISCUSSION

The authors investigated the final set of papers based on three different categories - LWA models, TL types, and dataset characteristics. Broadly, three different research themes on LWA have been observed in Covid-19 image classification literature when TL is applied, and they are 1. Papers working on developing a single LWA, 2. Developing the ensemble models (LWA ensembled with standard CNNs), 3. Comparison of benchmark studies including both LWA and standard CNNs. Initially, access to Covid-19 imaging was difficult for the researchers as there were not enough publicly available datasets, however, as the availability of datasets increased, the publications grew rapidly in 2021 and 2022 as presented in Figure 1.a.

LWA Backbone Model:

Most of these studies utilized MobileNetV2 (n=15) as the backbone model. Besides MobileNetV2, the most prevalent model is its previous version MobileNet (n=10) which contains slightly more parameters. For example, a study on the Covid-19 CT image classification (Zhang et al., 2022) utilized mobilnetV2 and then applied Bayesian optimization to adjust the hyperparameters of the model to design a lightweight model that can be easily embedded in mobile devices. The authors performed the efficacy test on their model and found that the average time required to classify one CT image is just 1.06 s. Ten benchmark studies compared MobileNet with other standard CNNs and 5 (Ahsan, Ahad, et al., 2021; Ahsan, Nazim, et al., 2021; Huang & Liao, 2022; Ragab et al., 2022; Zhang et al., 2022) of them reported that it either competitively performed or outperformed the other studies with a much lesser model size. EfficientNetB0 (n=7), NasNetMobile (n=6), and SqueezeNet (n=6) are the next commonly used LWAs. Finally, a few studies (Das et al., 2020; F. Montalbo, 2021; F. J. Montalbo, 2022) utilized truncated versions of the standard and Lightweight CNNs, namely DenseNet (n=2), InceptionV3 (n=2), Xception, Resnet50V2, InceptionResnetV2, and EfficientnetB0 with n=1 study using each model.

Transfer Learning Types:

Figure 1.b. demonstrates several types of TL applied with LWA backbones in the literature. Fine-tuning (n=13) is the most popular TL approach followed by Feature Extraction (n=10), Feature Extraction+ML (n=5), and Feature Extraction +Ensemble (n=3). While Feature Extraction can be especially beneficial in the case of developing LWAs as all the pre-trained weights are frozen and reduce the computation costs, the Fine-tuning approach is still popular because of the ability to identify the optimal performance of the model by retraining few or all the layers. Seven of the thirteen fine-tuning studies reported that LWA either performed competitively or outperformed the standard CNNs. Overall, Feature Extraction when compared to Fine-tuning with

a combination of either hybrid or ensemble techniques is widely researched (n=19) and achieved success. Not only did LWA Feature extractors use ML models as a classifier, but also experimented with DL models as classifiers. The various types of TL are illustrated in figure 3.a. For instance, a study (Tiwari & Jain, 2022) utilized MobileNet and three other standard CNNs as a Feature extractor on CT images and then input the features into Capsule net (Sabour et al., 2017) classifier, a DL model. It is worth noting that the authors suggest the MobileCapsNet as the preferred network over the DenseNet, VGG16, and ResNet50 used in this study due to its compact size and high performance. Another highly observed approach to utilizing TL is to employ Feature extraction or Fine-tuning on LWA and Standard CNNs and then perform Feature Fusion. EfficientNetB0 and MobileNetV2 are two prominent LWAs utilized in this way.(Ahmad et al., 2021; S. Tewari et al., 2020b; Sharma et al., 2022) While this method can become computationally expensive, there are opportunities to explore ensembled truncated models.

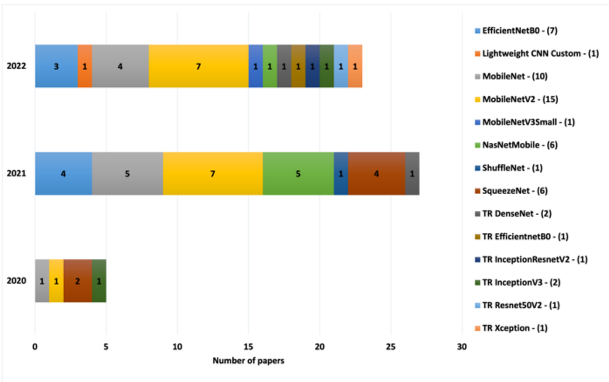

Figure with number of Lightweight Architecture Backbone models utilized across 39 papers during 2020 January -2021 August

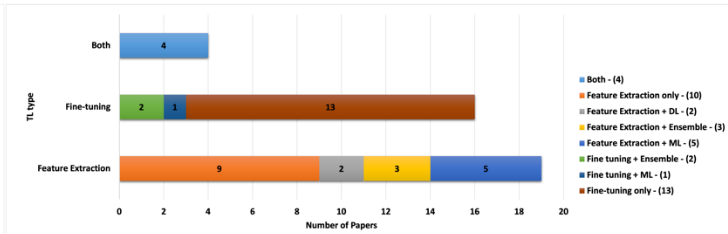

Figure with types of Transfer Learning observed in 39 included articles

Figure 1.a. LWA backbone models and 1.b. Types of TL

Data Characteristics:

Coming to imaging modality, the majority of the papers utilize Chest X-rays (n=26) instead of CT scans (n=8) as depicted in figure 2.a. due to the cost efficiency and portability (Chakraborty et al., 2021). Several papers have utilized multiple datasets of different sizes in their studies, but we found out that data availability is still a big problem as most of the available dataset range is concentrated below 1000 images (n=22). The dataset sizes are represented in figure 2. b. It is known that CNNs are data-hungry models but only a few papers (n=14) utilized the datasets containing more than 5000 images. To counter this issue, researchers are resorting to using data augmentation and it is proving to be an important tool to improve the accuracy of the model. In one such study (Rangarajan & Ramachandran, 2021), authors used GANs for the augmentation of Chest X-rays and developed a mobile-based application using MobileNetV2. Another study extensively investigated several traditional augmentation methods (resize value, resize method, rotate, zoom, warp, light, flip and normalize) on X-ray images and concluded that warping performed better than cropping to preserve image pixels (Monshi et al., 2021). Not only data quality and quantity but also data privacy protection is explored in one study (Heidari et al., 2022), they used blockchain technology to fine-tune a custom pre-trained lightweight model for TL on CT scans. This Edge AI model facilitates the protection of institutions by a Blockchain verification system and trains the LWA globally. Incorporating privacy protection elements into LWAs is a great way to develop Edge AI mobile-based applications that keep the data confidential.

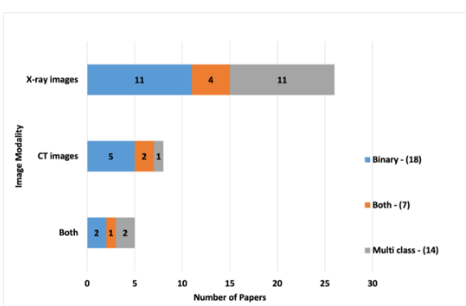

Figure with Covid-19 Imaging Modalities observed in the Literature

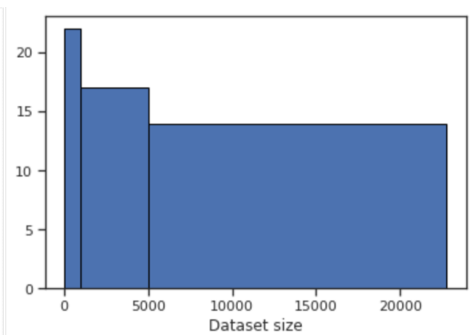

Figure 2.a. Imaging Modalities used for Covid-19 detection and 2.b. Dataset size range of images

CONCLUSION AND FUTURE DIRECTIONS

We identified that the most used LWAs are MobileNetV2 and MobileNet, due to their compact nature and highly efficient use of computations, they can be easily selected for Edge AI mobile based applications. We also identified that standard CNNs such as InceptionV3, Xception, Resnet50V2, InceptionResnetV2, and Densenet121 can be truncated to minimize their size by preserving the performance with the help of TL, few studies (F. Montalbo, 2021; F. J. Montalbo, 2022) implemented this idea. After analyzing all the LWA models. Additionally, we found that the smallest LWA used contains only 0.2 M parameters which is a truncated version of EfficientNetB0. This provides future opportunities to explore more truncated models in Covid-19 image analysis. Also, Feature extraction and Fine-tuning are both extensively applied TL techniques. Moreover, ensembled models with the help of feature stacking and feature fusion are observed in the literature, however, none of the studies implemented truncated models with ensemble techniques.

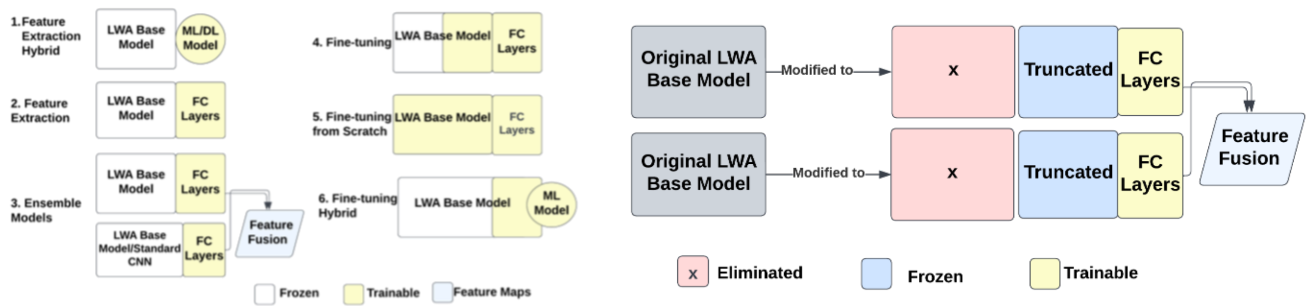

Figure 3.a. Types of TL and 3.b. proposed truncated ensemble model with LWA

In our future study, we intend to explore this area by truncating two different LWAs and then performing feature fusion as depicted in Figure 3.b. Several studies performed pre-processing methods such as augmentation and fuzzy color techniques to improve the data quality and quantity, however, only one study implemented Blockchain to protect data privacy, this calls for the need for more research in this area to combine Blockchain and Federated Learning techniques with LWA for Edge AI models. This study is not without limitations, we included articles from Arxiv that are not peer-reviewed. Also, our study only considers research with TL, but there are articles proposing LWAs without the use of TL (Chakraborty et al., 2021; Wang et al., 2022). Including these types of studies in future research can be beneficial for architectural improvements of LWAs.

REFERENCES

1. Ahmad, F., Farooq, A., & Ghani, M. (2021). Deep Ensemble Model for Classification of Novel Coronavirus in Chest X-Ray Images. COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE, 2021. https://doi.org/10.1155/2021/8890226
2. Ahsan, M., Ahad, M., Soma, F., Paul, S., Chowdhury, A., Luna, S., Yazdan, M., Rahman, A., Siddique, Z., & Huebner, P. (2021). Detecting SARS-CoV-2 From Chest X-Ray Using Artificial Intelligence. IEEE ACCESS, 9, 35501–35513. https://doi.org/10.1109/ACCESS.2021.3061621
3. Ahsan, M., Nazim, R., Siddique, Z., & Huebner, P. (2021). Detection of COVID-19 Patients from CT Scan and Chest X-ray Data Using Modified MobileNetV2 and LIME. HEALTHCARE, 9(9). https://doi.org/10.3390/healthcare9091099
4. Bhattacharya, S., Reddy Maddikunta, P. K., Pham, Q.-V., Gadekallu, T. R., Krishnan S, S. R., Chowdhary, C. L., Alazab, M., & Jalil Piran, Md. (2021). Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey. Sustainable Cities and Society, 65, 102589. https://doi.org/10.1016/j.scs.2020.102589
5. Chakraborty, M., Dhavale, S. V., & Ingole, J. (2021). Corona-Nidaan: Lightweight deep convolutional neural network for chest X-Ray based COVID-19 infection detection. Applied Intelligence (Dordrecht, Netherlands), 51(5), 3026–3043. https://doi.org/10.1007/s10489-020-01978-9
6. Das, D., Santosh, K. C., & Pal, U. (2020). Truncated inception net: COVID-19 outbreak screening using chest X-rays. Physical and Engineering Sciences in Medicine, 43(3), 915–925. https://doi.org/10.1007/s13246-020-00888-x
7. Erickson, B. J., Korfiatis, P., Kline, T. L., Akkus, Z., Philbrick, K., & Weston, A. D. (2018). Deep Learning in Radiology: Does One Size Fit All? Journal of the American College of Radiology, 15(3), 521–526. https://doi.org/10.1016/j.jacr.2017.12.027

8. F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi, & D. Shen. (2020). Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation and Diagnosis for COVID-19. IEEE Reviews in Biomedical Engineering, 1–1. https://doi.org/10.1109/RBME.2020.2987975
9. Godasu, R., Zeng, D., & Sutrave, K. (2020). Transfer Learning in Medical Image Classification: Challenges and Opportunities. 7.
10. Haque, W., Arefin, S., Shihavuddin, A., & Hasan, M. (2020). DeepThin: A novel lightweight CNN architecture for traffic sign recognition without GPU requirements. Expert Systems with Applications, 168. https://doi.org/10.1016/j.eswa.2020.114481
11. Heidari, A., Toumaj, S., Navimipour, N. J., & Unal, M. (2022). A privacy-aware method for COVID-19 detection in chest CT images using lightweight deep conventional neural network and blockchain. Computers in Biology and Medicine, 145, 105461. https://doi.org/10.1016/j.compbiomed.2022.105461
12. Huang, M., & Liao, Y. (2022). A lightweight CNN-based network on COVID-19 detection using X-ray and CT images. COMPUTERS IN BIOLOGY AND MEDICINE, 146. https://doi.org/10.1016/j.compbiomed.2022.105604
13. Iandola, F. N., Ashraf, K., Moskewicz, M. W., & Keutzer, K. (2016). FireCaffe: Near-linear acceleration of deep neural network training on compute clusters. ArXiv:1511.00175 [Cs]. http://arxiv.org/abs/1511.00175
14. Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. PLoS Medicine, 6(7), e1000100.
15. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A Survey on Deep Learning in Medical Image Analysis. Medical Image Analysis, 42, 60–88. https://doi.org/10.1016/j.media.2017.07.005
16. Monshi, M. M. A., Poon, J., Chung, V., & Monshi, F. M. (2021). CovidXrayNet: Optimizing data augmentation and CNN hyperparameters for improved COVID-19 detection from CXR. Computers in Biology and Medicine, 133, 104375. https://doi.org/10.1016/j.compbiomed.2021.104375
17. Montalbo, F. (2021). Diagnosing Covid-19 chest x-rays with a lightweight truncated DenseNet with partial layer freezing and feature fusion. BIOMEDICAL SIGNAL PROCESSING AND CONTROL, 68. https://doi.org/10.1016/j.bspc.2021.102583
18. Montalbo, F. J. (2022). Truncating fined-tuned vision-based models to lightweight deployable diagnostic tools for SARS-CoV-2 infected chest X-rays and CT-scans. Multimedia Tools and Applications, 81(12), 16411–16439. https://doi.org/10.1007/s11042-022-12484-0
19. Noteboom, C., Zeng, D., Godasu, R., & Sutrave, K. (2021). Applications of Deep Learning Augmented Systems for Covid-19 Predictions- A Literature Review. AMCIS 2021 Proceedings. https://aisel.aisnet.org/amcis2021/adv_info_systems_general_track/adv_info_systems_general_track/31
20. Ragab, M., Alshehri, S., Azim, G. A., Aldawsari, H. M., Noor, A., Alyami, J., & Abdel-Khalek, S. (2022). COVID-19 Identification System Using Transfer Learning Technique With Mobile-NetV2 and Chest X-Ray Images. Frontiers in Public Health, 10, 819156. https://doi.org/10.3389/fpubh.2022.819156
21. Rangarajan, A., & Ramachandran, H. (2021). A preliminary analysis of AI based smartphone application for diagnosis of COVID-19 using chest X-ray images. EXPERT SYSTEMS WITH APPLICATIONS, 183. https://doi.org/10.1016/j.eswa.2021.115401
22. S. Tewari, U. Agrawal, S. Verma, S. Kumar, & S. Jeevaraj. (2020a). Ensemble Model for COVID-19 detection from chest X-ray Scans using Image Segmentation, Fuzzy Color and Stacking Approaches. 2020 IEEE 4th Conference on Information & Communication Technology (CICT), 1–6. https://doi.org/10.1109/CICT51604.2020.9312076
23. S. Tewari, U. Agrawal, S. Verma, S. Kumar, & S. Jeevaraj. (2020b). Ensemble Model for COVID-19 detection from chest X-ray Scans using Image Segmentation, Fuzzy Color and Stacking Approaches. 2020 IEEE 4th Conference on Information & Communication Technology (CICT), 1–6. https://doi.org/10.1109/CICT51604.2020.9312076
24. Sabour, S., Frosst, N., & Hinton, G. E. (2017). Dynamic Routing Between Capsules. Advances in Neural Information Processing Systems, 30. https://proceedings.neurips.cc/paper/2017/hash/2cad8fa47bbef282badbb8de5374b894-Abstract.html

25. Sharma, A., Singh, K., & Koundal, D. (2022). A novel fusion based convolutional neural network approach for classification of COVID-19 from chest X-ray images. BIOMEDICAL SIGNAL PROCESSING AND CONTROL, 77. https://doi.org/10.1016/j.bspc.2022.103778
26. Sufian, A., Ghosh, A., Sadiq, A. S., & Smarandache, F. (2020). A Survey on Deep Transfer Learning to Edge Computing for Mitigating the COVID-19 Pandemic. Journal of Systems Architecture, 108, 101830. https://doi.org/10.1016/j.sysarc.2020.101830
27. Tiwari, S., & Jain, A. (2022). A lightweight capsule network architecture for detection of COVID-19 from lung CT scans. INTERNATIONAL JOURNAL OF IMAGING SYSTEMS AND TECHNOLOGY, 32(2), 419–434. https://doi.org/10.1002/ima.22706
28. Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. Nature Communications, 11(1), Article 1. https://doi.org/10.1038/s41467-019-14108-y
29. Wang, W., Huang, W., Wang, X., Zhang, P., & Zhang, N. (2022). A COVID-19 CXR image recognition method based on MSA-DDCovidNet. IET Image Processing, 16(8), 2101–2113. https://doi.org/10.1049/ipr2.12474
30. WHO | World Health Organization. (n.d.). WHO; World Health Organization. Retrieved March 18, 2020, from http://www.who.int/gho/health_workforce/en/
31. Yu, D., Xu, Q., Guo, H., Zhao, C., Lin, Y., & Li, D. (2020). An Efficient and Lightweight Convolutional Neural Network for Remote Sensing Image Scene Classification. Sensors (Basel, Switzerland), 20(7). https://doi.org/10.3390/s20071999
32. Zhang, X., Zhou, J., Sun, W., & Jha, S. (2022). A Lightweight CNN Based on Transfer Learning for COVID-19 Diagnosis. CMC-COMPUTERS MATERIALS & CONTINUA, 72(1), 1123–1137. https://doi.org/10.32604/cmc.2022.024589