COMPUTOLOGY: Journal of Applied Computer Science and Intelligent Technologies
Volume 2, Issue 2, Jul-Dec 2022, pp. 14-22
DOI: 10.17492/computology.v2i2.2202
https://www.journalpressindia.com/cjacsit

Lung Infection Detection using Contemporary Techniques of Artificial Intelligence

Shiva Prasad Koyyada, Ajay Rawat** and Thipendra P. Singh****

ABSTRACT

Lung diseases are increasing day by day when compared to other diseases. The world has been experiencing COVID-19 since 2020 and one has no answer how to cure it in the initial days. Near future majority of the humanity will suffer from lung diseases based on studies. In this context, we are exploring contemporary artificial techniques (AI) and explaining how these will help in detecting the disease. Among Convolutional neural nets (CNNs), Regional Convolutional neural nets (RCNNs), and Vision Transformers (VT) state-of-the-art methods. The techniques are very diversified from the concept point of view. Each carries out the process in an expert wise. One who wants to automate the process can make an ensemble of these models.

Keywords: Convolutional neural nets; Vision transformers; Regional Convolutional neural nets, Lung disease.

1.0 Introduction

Lung diseases are one of the diseases that where majority of humanity is suffering from. Lung diseases include pneumonia, lung nodules, pneumothorax, and other lung infections. In 2020 a lung disease called COVID-19 sprung in Wuhan, a city in China. This disease has spread worldwide and everyone has experienced the pain of it. Medical research is not ready to face the number of cases increasing. RT-PCR is the method used to determine such disease however these kits are very limited. It is the responsibility of humans to determine under these circumstances. In this context, technology will be helpful in determining these diseases. One has to make use of the same. Chest X-ray and Computed tomography (CT) scans are the other modality methods to determine lung infections. We will be explaining further how state-of-the-art methods could be used to determine lung disease and how each is different from other.

2.0 Related work and Methodology

2.1 Related work

To stop the spread of COVID-19 among people, an automatic detection system could be used as a quick alternative diagnosis option [1].

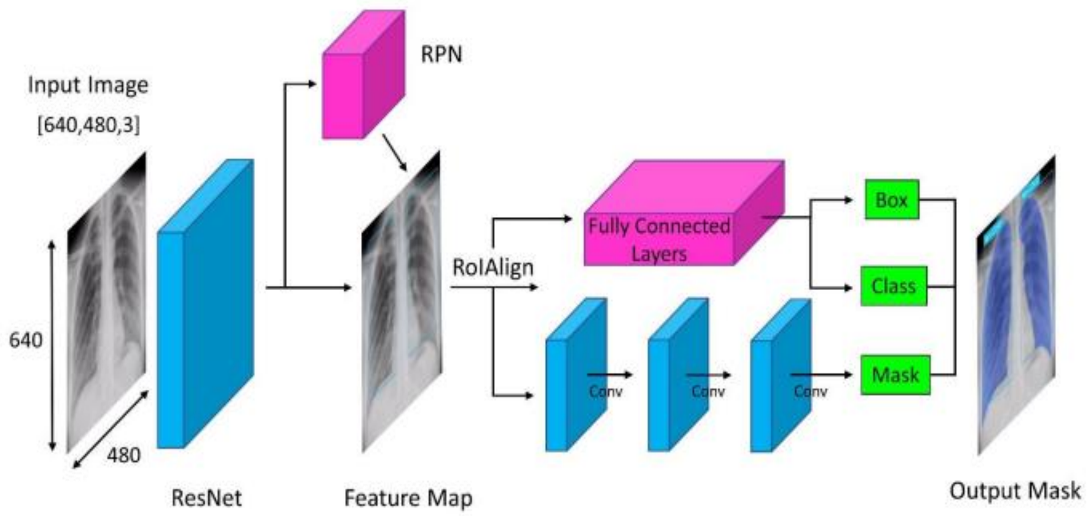
**Corresponding author; School of Computer Science, UPES University, Bidholi, Dehradun, Uttarakhand, India (E-mail: 500073479@stu.upes.ac.in)*
***Science and Technology Facilities Council, UK Research and Innovation (UKRI), Daresbury Warrington, United Kingdom (E-mail: ajay.rawat@stfc.ac.uk)*
****School of Computer Science, UPES University, Bidholi, Dehradun, Uttarakhand, India (E-mail: tpsingh@ddn.upes.ac.in)*

This study aims to use pre-trained transfer learning models such as the ResNet family, and Inception family in order to predict COVID-19. With no false negatives, we suggested a thin, shallow architecture for Convolutional Neural Networks (CNNs) that can automatically identify COVID-19-positive situations [2]. In comparison to other deep learning models, the shallow CNN-tailored architecture was created with fewer parameters. 321 CXRs that were COVID-19 positive were used to validate the shallow CNNtailored architecture.

Additional non-COVID-19 5856 instances (publicly accessible, source: Kaggle) that included cases of normal, viral, and bacterial pneumonia were taken into consideration in addition to COVID-19-positive patients. The authors of [3] come up with a model for diagnosing COVID-19 that uses a deep CNN approach which employs VGG family, ResNet family, DenseNet, and InceptionV3. In comparison to the other the model VGG16 has performed well.

In the same vein, [4] designed a Mask RCNN to predict COVID-19. The architecture has shown in figure 1. Pneumonia detection through Fast RCNN [5] where CXR pictures were used to detect and locate pneumonia using DeepConv-DilatedNet. Faster R-CNN with two stages is used as the network's structure. Feature Pyramid Network (FPN) is inserted into the residual neural network of a dilated bottleneck to focus on the deep features and positional information of the object.

Figure 1: The Architecture of Mask R-CNN for COVID-19 Image Segmentation



Recent radiological research suggests that patients' infection status can be determined by the specific distribution of ground-glass opacities (GGOs), which are present on specific lung tissue. Computed tomography (CT) can be utilized to diagnose COVID-19 in addition to RT-PCR. For the purpose of detecting ground glass opacities (GGOs) in chest CT scans of COVID-19-infected individuals, the authors of this paper [6] have presented a Mask R-CNN (region-based convolution neural network) technique. Deep learning pipeline for COVID-19 identification from chest X-ray-based images using Vision Transformer [7].

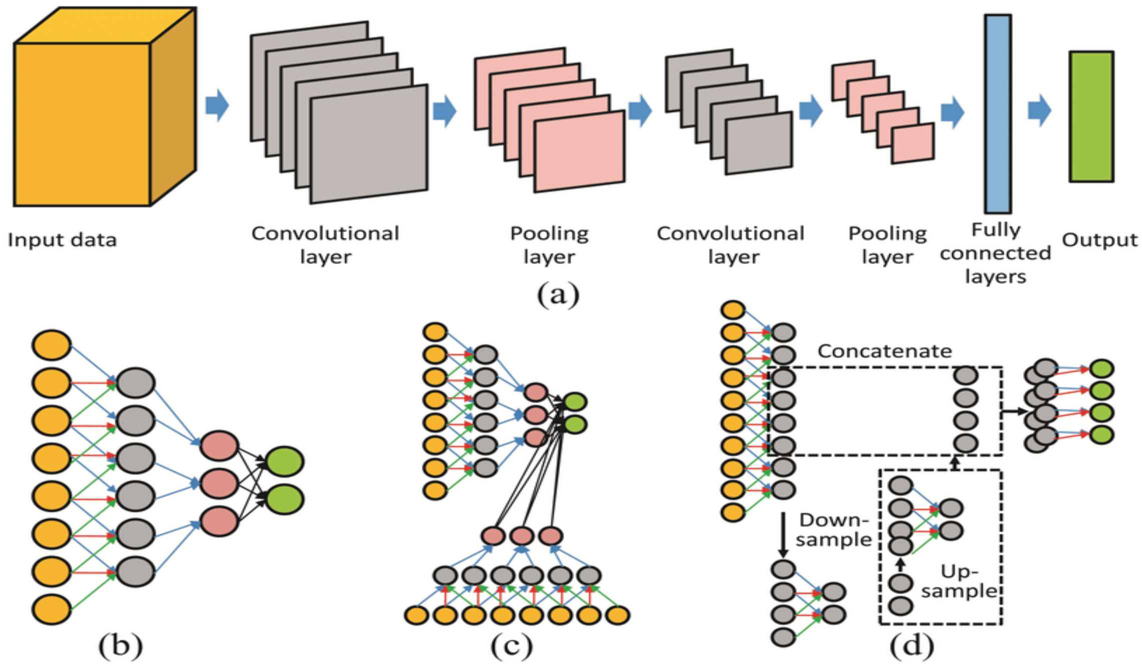
2.2 Methodology

CNNs, RCNNs, and Vision Transformers are used to detect lung diseases. Now we see how these methods are fundamentally different and how these can be used further.

2.2.1. Convolutional neural networks functionality

Convolutional Neural Networks (CNNs) play a crucial role in lung infection detection using chest X-ray (CXR) images. The basic architecture of CNN has shown in the figure 2. Feature Extraction: CNNs are skilled at automatically identifying pertinent elements in images. Patterns, forms, and textures that are suggestive of lung infections may be present in CXR images as these features. In order to find these features, CNNs employ a sequence of convolutional layers, gradually learning more complicated representations as they progress deeper into the network. Pattern Recognition: Lung infections frequently appear as distinct patterns on CXR images. For instance, patches of opacity in the lung fields may be a sign of pneumonia. CNNs can distinguish between healthy and sick areas by identifying these patterns and capturing them as features.

Figure 2: CNNs architectures commonly used in medical imaging.(a) Basic CNN with a sequence of convolution and max-pooling (b) classic 2D CNN(c) Multi-stream CNN (d) U-NET with only one down-sampling stage [8]



Classification: After CXR image features have been retrieved, CNNs are applied to classification tasks. They got the expertise to identify whether or not an image shows signs of a lung infection. Usually, this is done by categorizing illnesses into multiple categories using multi-class classification or binary classification (as infected or noninfected).

Localization: Localizing the regions of interest in CXR images can also be done using CNNs. In order to detect lung infections, it may be necessary to pinpoint the precise locations of the illnesses. This can help doctors and radiologists determine the scope and severity of the infection.

Real-time Decision Support: CNNs can be embedded in healthcare systems to offer radiologists real-time decision help. They can help to speed up the process of CXR image interpretation and possibly increase diagnostic precision by assisting in the identification of probable infection-related abnormalities. CNNs are useful tools for large-scale screening programs, such as those employed during disease outbreaks, since they can process enormous numbers of CXR images quickly. They can swiftly spot cases that warrant additional investigation by medical experts. Integration into Telemedicine:CNN-based lung infection detection models can be integrated into telemedicine platforms, allowing for remote screening and diagnosis, especially in areas with limited access to healthcare facilities.

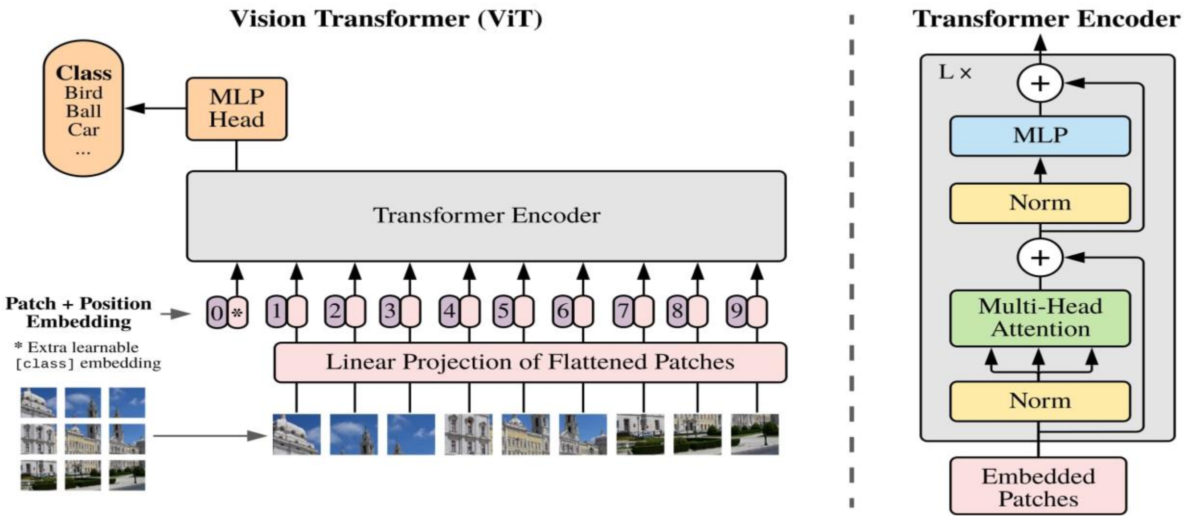
It's important to note that while CNNs are powerful tools for lung infection detection in CXR images [9],[10] they are typically used as aids to healthcare professionals rather than as standalone diagnostic tools. Their predictions should be interpreted by trained radiologists and clinicians, and any definitive diagnosis and treatment decisions should be made by healthcare experts. Additionally, the accuracy and reliability of CNN-based models depend on the quality and diversity of the training data and the specific architecture and parameters of the network.

2.3 Vision transformers

An increasingly popular neural network architecture called Vision Transformers (ViTs) is used for computer vision tasks including medical image processing. The Transformer architecture, which was created originally for natural language processing, is used by ViTs to process image data. Below is an outline of how ViTs might be utilized to detect lung infections in chest X-ray (CXR) images based on the broad principles of ViT utilization in computer vision.[11].

Its architecture as Data Preparation: Data Collection and Annotation: It is crucial to have a sizable collection of CXR images with annotations indicating the presence or absence of lung illnesses (such as pneumonia and tuberculosis). Preprocessing: In order to improve the network's capacity to learn features, preprocessing operations may include scaling, normalization, and augmentation for CXR images because their sizes often fluctuate. Patch embedings: PViTs divide the input image into a grid of nonoverlapping patches. Each patch is treated as a sequence of embeddings, just like words in natural language processing tasks.

Figure 3: Vision Transformer Architecture



Positional Encodings: To capture spatial information, positional encodings are added to the patch embeddings. These encodings provide the network with information about the patch's position in the image. Multi-Head Self-Attention: ViTs utilize multi-head self-attention mechanisms to allow patches to attend to each other. This enables the model to learn dependencies and relationships between different regions of the CXR image.

Transformer Blocks: These blocks consist of multi-head self-attention layers followed by feedforward neural networks. They are stacked to process the patch embeddings hierarchically.

Classification Head: After processing the patch embeddings through the Transformer layers, a classification head is added to make predictions. For lung infection detection, this head typically consists of fully connected layers, which output the probability of infection or non-infection.

Supervised Learning: ViTs are trained using supervised learning, where the network learns to predict infection labels based on the annotated dataset.

Loss Function: A common choice for binary classification tasks like this is binary cross-entropy loss.

Optimization: Optimizers like Adam or SGD are used to update the network's parameters during training.

Evaluation: The trained ViT model is evaluated on a separate test dataset to assess its performance. Evaluation metrics such as accuracy, precision, recall, F1-score, and AUC-ROC may be used to quantify its ability to detect lung infections. Post-processing and Interpretability: A probability threshold may be applied to the model's output to classify an image as infected or not infected. Interpretability: techniques such as attention maps can be used to understand which regions of the CXR image were most informative in making the prediction. This can help in providing insights into the decision-making process of the model.

It's crucial to keep in mind that while ViTs have demonstrated promise in a number of computer vision applications, their application to medical image analysis, especially the detection of lung infections [12],[13], [14] may call for modifications to account for the unique qualities and difficulties of medical images. Additionally, for assuring the validity and generalizability of ViT-based models in healthcare applications, extensive and diverse datasets, thorough model evaluation, and model validation are essential. 2.4. Regional Convolutional Neural networks(RCNNs) R-CNN [15] includes variants like Fast R-CNN and Faster R-CNN, can be used in the detection of lung infections in medical images [16],[17],[18],[19] including chest X-rays (CXR) and computed tomography (CT) scans. Here's an overview of how R-CNNs can be applied to this task.

2.4.1 Data preparation

Dataset Collection and Annotation: A dataset of CXR or CT scan images is collected and annotated to mark regions of interest (ROIs) within the images where infections or abnormalities may be present. Annotations should include the type and location of infections. Preprocessing: Images are preprocessed, which may involve resizing, normalization, and augmentation to enhance the network's ability to detect infections.

Region Proposal Network (RPN): In Faster R-CNN, a Region Proposal Network is used to generate a set of candidate regions (ROIs) likely to contain infections. These ROIs are generated based on anchor boxes and their likelihood of containing objects of interest.

Feature Extraction Backbone: The CNN component of the R-CNN (e.g., ResNet, VGG, etc.) extracts features from the entire image, including both the ROIs and the background. This feature extraction backbone is pretrained on a large dataset (e.g., ImageNet) to capture generic image features.

ROI Pooling: ROIs generated by the RPN are individually cropped from the feature maps produced by the backbone network and resized to a fixed size. This creates a consistent input size for the subsequent layers.

Classification Head: The feature representations of the ROIs are passed through a classification head, which classifies each ROI as either infected or not infected.

Regression Head: A regression head is used to refine the bounding box coordinates of the ROIs, improving the localization accuracy of infections.

Supervised Learning: The R-CNN is trained in a supervised manner using the annotated dataset. It learns to classify ROIs as infected or non-infected and refines bounding boxes to better align with the infections' actual locations.

Loss Functions: Typical loss functions used include binary cross-entropy for classification and smooth L1 loss for bounding box regression.

Evaluation: The trained R-CNN model is evaluated on a separate test dataset to assess its performance in detecting lung infections. Metrics like precision, recall, F1-score, and IoU (Intersection over Union) can be used to measure its accuracy and localization ability. Post-processing,

Visualization and Reporting: Post-processing steps may involve applying a confidence threshold to the classification scores and filtering out detections with low confidence to reduce false positives. The R-CNN can provide visualizations of the detected infections overlaid on the original images, aiding radiologists in their assessment.

The fact that R-CNNs, particularly Faster R-CNNs, are effective for object identification tasks and can be modified for lung infection diagnosis must be emphasized. However, the dataset's size and quality, the backbone network of choice, and the meticulous hyperparameter tweaking all have a role in how effective they are. Additionally, rather than serving as a sole diagnostic tool, they are frequently employed as a support for radiologists, with medical professionals making the final determinations of diagnosis and course of treatment.

3.0 Experiment and Result

In order to compare the performance of different methods, we have started experimenting with vision transformers. The reason for choosing vision transformers is that most of the researchers have worked with CNNs. Our experimentation as shown in the figure 4 has been used to 3500 images from kaggle. Each image will be divided into 64 patches where each patch of size 8*8. Each patch will be converted into a vector. Each vector will be passed to the transformer for further training. The model was trained on google colab for 100 epochs. Adam optimizer was used to optimize the weights. Figure 5 shows the error curve where the blue color indicates on the train and the other indicates on validation data. The experiment has shown an accuracy of 99.15% on test data when compared to the existing CNN performance [20].

Figure 4: Experimental Setup Block Diagram

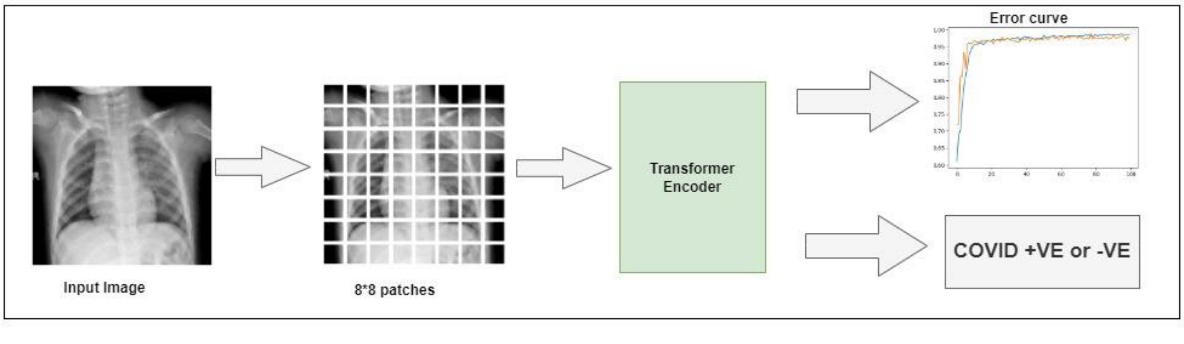
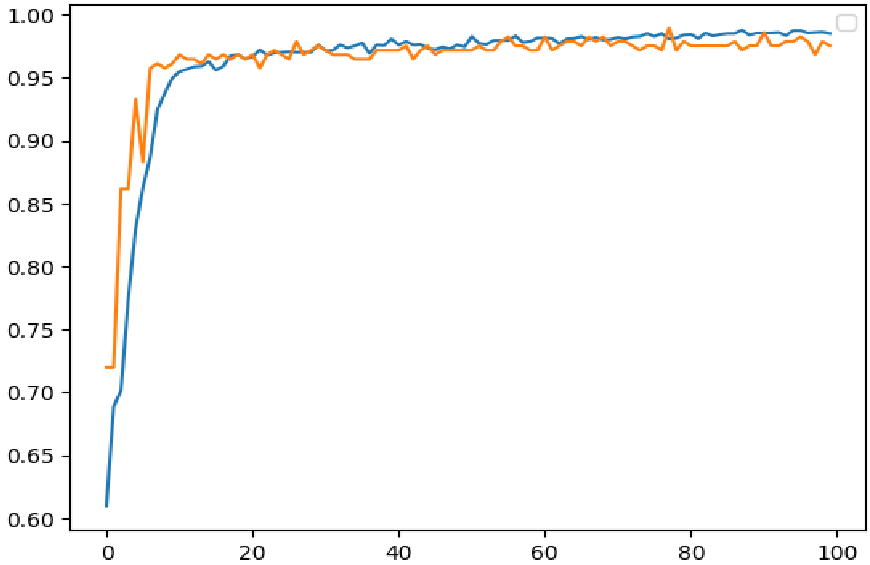


Figure 5: Error Curve Obtained during Training



4.0 Conclusion and Future Scope

As we observe all these methods have a data pre-processing stage that is different for each. CNNs are focusing on resizing of images and vision transformers are preparing data patch-wise as RCNN prepares data through annotations. Now all these focus on extracting features from the image through CNN based method. Localization is better performed through a vision transformer and RCNN instead of CNN. Instead, CNN focuses on the entire image to extract the features. As all the methods focus on various parts of the CXR, one can come up with a combination of these models to get better accuracy. As these models are diversified in nature, each comes up with its own prediction helps medical practitioners to make better decisions. for example, RCNN may find ground glass opacities by highlighting region of interest features. This way a robust AI model is possible. As research progresses more lung diseases can be included and the future is AI.

References

1. A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, Pattern Analysis and Applications 24 (3) (2021) 1207–1220. doi:10.1007/s10044-021-00984-y. URL https://doi.org/10.1007/s10044-021-00984-y
2. H. Mukherjee, S. Ghosh, A. Dhar, S. M. Obaidullah, K. C. Santosh, K. Roy, Shallow convolutional neural network for covid-19 outbreak screening using chest x-rays, Cognitive Computation (Feb 2021). doi:10.1007/s12559-020-09775-9. URL https://doi.org/10.1007/s12559-020-09775-9
3. K. Sahinbas, F. O. Catak, 24 - transfer learning-based convolutional neural network for covid-19 detection with x-ray images, in: U. Kose, D. Gupta, V. H. C. de Albuquerque, A. Khanna (Eds.), Data Science for COVID-19, Academic Press, 2021, pp. 451–466. doi:https://doi.org/10.1016/B978-0-12-824536-1.00003-4. URL https://www.sciencedirect.com/science/article/pii/B9780128245361000034
4. S. Podder, S. Bhattacharjee, A. Roy, An efficient method of detection of covid-19 using mask r-cnn on chest x-ray images, AIMS Biophys 8 (3) (2021) 281–290.

5. S. Yao, Y. Chen, X. Tian, R. Jiang, Pneumonia detection using an improved algorithm based on faster r-cnn, Computational and Mathematical Methods in Medicine 2021 (2021) 8854892. doi:10.1155/2021/8854892. URL https://doi.org/10.1155/2021/8854892
6. A. Kundu, C. Mishra, S. Bilgaiyan, Covid-segnet: Diagnosis of covid-19 cases on radiological images using mask r-cnn, in: 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII), 2021, pp. 1–5. doi:10.1109/ICBSII51839.2021.9445190.
7. D. Shome, T. Kar, S. N. Mohanty, P. Tiwari, K. Muhammad, A. Al- Tameem, Y. Zhang, A. K. J. Saudagar, Covid-transformer: Interpretable covid-19 detection using vision transformer for healthcare, International Journal of Environmental Research and Public Health 18 (21) (2021). doi:10.3390/ijerph182111086. URL https://www.mdpi.com/1660-4601/18/21/11086
8. P. Monkam, S. Qi, H. Ma, W. Gao, Y. Yao, W. Qian, Detection and classification of pulmonary nodules using convolutional neural networks: A survey, IEEE Access PP (2019) 1–1. doi:10.1109/ACCESS.2019.2920980.
9. I. Ahmed, A. Ahmad, G. Jeon, An iot-based deep learning framework for early assessment of covid-19, IEEE Internet of Things Journal 8 (21) (2021) 15855–15862. doi:10.1109/JIOT.2020.3034074.
10. R. C. Joshi, S. Yadav, V. K. Pathak, H. S. Malhotra, H. V. S. Khokhar, A. Parihar, N. Kohli, D. Himanshu, R. K. Garg, M. L. B. Bhatt, R. Kumar, N. P. Singh, V. Sardana, R. Burget, C. Alippi, C. M. Travieso-Gonzalez, M. K. Dutta, A deep learning-based covid-19 automatic diagnostic framework using chest x-ray images, Biocybernetics and Biomedical Engineering 41 (1) (2021) 239–254. doi:https://doi.org/10.1016/j.bbe.2021.01.002. URL https://www.sciencedirect.com/science/article/pii/S0208521621000036
11. A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby, An image is worth 16x16 words: Transformers for image recognition at scale (2021). arXiv:2010.11929.
12. K. S. Krishnan, K. S. Krishnan, Vision transformer based covid-19 detection using chest x-rays, in: 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), 2021, pp. 644–648. doi:10.1109/ISPCC53510.2021.9609375.
13. K. Tyagi, G. Pathak, R. Nijhawan, A. Mittal, Detecting pneumonia using vision transformer and comparing with other techniques, in: 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2021, pp. 12–16. doi:10.1109/ICECA52323.2021.9676146.
14. L. T. Duong, N. H. Le, T. B. Tran, V. M. Ngo, P. T. Nguyen, Detection of tuberculosis from chest x-ray images: Boosting the performance with vision transformer and transfer learning, Expert Systems with Applications 184 (2021) 115519. doi:https://doi.org/10.1016/j.eswa.2021.115519. URL https://www.sciencedirect.com/science/article/pii/S0957417421009295
15. S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks (2016). arXiv:1506.01497.
16. K. Long, L. Tang, X. Pu, Y. Ren, M. Zheng, L. Gao, C. Song, S. Han, M. Zhou, F. Deng, Probability-based mask r-cnn for pulmonary embolism detection, Neurocomputing 422 (2021) 345–353. doi:https://doi.org/10.1016/j.neucom.2020.10.022. URL https://www.sciencedirect.com/science/article/pii/S0925231220315186
17. L. J. Muhammad, J. M. Amshi, S. S. Usman, I. A. Badi, I. Mohammed, O. S. Dada, A. A. Haruna, Deep learning models for classification and localization of covid-19 abnormalities on chest radiographs, in: 2021 4th International Conference on Computing Information Sciences (ICCIS), 2021, pp. 1–6. doi:10.1109/ICCIS54243.2021.9676401.

18. S. R. Jena, S. T. George, D. N. Ponraj, Lung cancer detection and classification with dgmm-rbcnn technique, Neural Computing and Applications 33 (22) (2021) 15601–15617. doi:10.1007/s00521-021-06182-5. URL https://doi.org/10.1007/s00521-021-06182-5
19. D. Saikrishna, M. M. Rao, B. S. Dhanush, S. Harshavardhan, B. Prudhvi, P. Rana, U. Mittal, Pneumonia detection using deep learning algorithms, in: 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), 2021, pp. 282–287. doi:10.1109/ICIEM51511.2021.9445310.
20. E. E.-D. Hemdan, M. A. Shouman, M. E. Karar, Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images (2020). URL http://arxiv.org/abs/2003.11055