Enhancing COVID-19 Diagnosis with Automated Reporting Using Preprocessed Chest X-Ray Image Analysis based on CNN

Arul Raj . A.M
Research Scholar
Department of Computer and Science and Engineering
Saveetha School of Engineering (SSE)
Saveetha Institute of Medical And Technical Sciences (SIMATS)
Saveetha University
Chennai, India
arulrajam1003.sse@saveetha.com

Sugumar R
Professor
Department of Computer and Science and Engineering
Saveetha School of Engineering (SSE)
Saveetha Institute of Medical And Technical Sciences (SIMATS)
Saveetha University
Chennai, India
sugumar.sse@saveetha.com

***Abstract*— The ongoing COVID-19 pandemic has caused a global health crisis, and accurate diagnosis and early detection are essential for successful management of the outbreak. Convolutional neural networks and pre-processed chest X-ray pictures are the two main components of the unique proposed method for the identification of COVID-19, which is presented in this paper (CNNs). Image enhancement and segmentation are performed during the pre-processing stage. These operations increase the overall quality and contrast of the pictures, which in turn makes it simpler for the CNN to recognise significant aspects of the images. The CNN model was trained using a large dataset of pre-processed X-ray pictures that included both COVID-19 positive and negative instances. The dataset was used to train the model. In comparison to more conventional diagnostic approaches, and this strategy was successful in achieving high levels of accuracy, sensitivity, and specificity in the detection of COVID-19. Moreover, this model designed an automated reporting system that saves time and costs by providing healthcare providers with diagnostic reports that are both prompt and accurate. This research demonstrates the viability of using CNNs and pre-processed X-ray images for the purpose of early identification of COVID-19 and offers an important resource for the efficient management of this worldwide health concern.**

Keywords— COVID-19, X-ray images, Automated Reporting, Convolutional Neural Network, Accuracy,

I. INTRODUCTION

The term "automated reporting in image processing for COVID-19 [1][2][3] detection using X-ray images" refers to the process of automatically generating diagnosis reports based on the examination of pre-processed chest X-ray images of patients who are thought to have COVID-19. This process takes place during image processing. In this method, the images that have already been preprocessed are fed into a trained convolutional neural network (CNN) model. This model has the ability to recognise characteristics and patterns in the images that are characteristic of COVID-19. After that, the output of the CNN model [4][5][6] is used to generate a diagnosis report for the patient, which indicates whether or not the patient has COVID-19. The automated reporting system features a straightforward graphical user interface, which makes it easier for medical professionals to upload chest X-ray images and obtain diagnosis reports. As a result, the process is both more effective and more readily available to them. In addition, the alert system can assist in prioritising cases that are either urgent or critical, thereby enabling interventions that are both timely and appropriate. Automated reporting in image processing for COVID-19 detection using X-ray images can help improve the accuracy and efficiency of diagnosis, reduce the burden on healthcare systems, and facilitate timely and appropriate interventions for COVID-19 patients. The outbreak of the COVID-19 pandemic has resulted in a global health crisis, affecting millions of people around the world. The rapid spread of the virus and the high number of cases has put tremendous pressure on healthcare systems worldwide. The availability of reliable diagnostic tools and techniques is crucial for the effective management of the pandemic. Chest X-rays have been widely used to diagnose respiratory infections, including COVID-19. However, traditional methods of diagnosing COVID-19 using X-rays can be time-consuming and require specialized expertise.

Recent improvements in deep learning[7][8][9] and computer vision have opened up new ways for chest X-rays to be used to automatically find and diagnose COVID-19. In particular, convolutional neural networks (CNNs) have done a good job of finding COVID-19 in X-ray images that have already been processed. Image enhancement and segmentation are two pre-processing techniques that can help improve the quality and contrast of the images. This makes it easier for the CNN to find the important parts of the image. The normal X-ray picture and the Covid X-ray image are shown side by side in figure 1.

(a) Normal X-Ray Image (b) Covid X-Ray Image

Fig. 1. An illustration of normal and COVID-19 CXR images.

In this study, a new way of using pre-processed chest X-ray images and CNNs to automatic reporting COVID-19 is proposed. The first step is to improve the quality and contrast of the X-ray images. Next, a CNN model is trained on a large set of COVID-19 positive and negative cases. After the model has been trained, it can be used to automatically diagnose and report COVID-19 cases. The

main goal of this study is to show that CNNs and pre-processed chest X-ray images are good at finding COVID-19 and to create an automated reporting system that gives doctors quick and accurate diagnosis reports. This approach could save time and money, make COVID-19 diagnoses more accurate and faster, and make it easier to deal with the ongoing pandemic.

II. Related Work

Recent research has investigated the use of deep learning algorithms and preprocessed chest X-ray images for the identification of COVID-19. In a work by Wang et al. (2020)[10], preprocessed chest X-ray pictures were used to build a deep learning model to identify COVID-19. In diagnosing COVID-19, the model outperformed radiologists due to its excellent accuracy, sensitivity, and specificity. In a similar fashion, Apostolopoulos and Bessiana (2020)[11] trained a deep learning model using a dataset of preprocessed chest X-ray pictures. High levels of accuracy and sensitivity were achieved by the model in distinguishing between COVID-19 positive and negative instances. Ozturk et al. (2020)[12] have investigated the application of deep learning models for COVID-19 identification using chest X-ray pictures. The scientists trained many deep learning models on a large dataset of preprocessed photos and compared their performance. The top-performing model detected COVID-19 with great precision and sensitivity. Hemdan et al. (2021) [13] created a deep learning-based method for the identification of COVID-19 utilising pre-processed chest X-ray images in a recent research. The system contains an automatic reporting module that offers healthcare providers with accurate and quick diagnostic reports. Overall, the proposed results demonstrate that deep learning models and pre-processed chest X-ray pictures may be viable tools for the automated identification and diagnosis of COVID-19, with the potential to save time and money while enhancing the diagnostic accuracy and speed.

III. Methodology

The proposed approach entails gathering and pre-processing chest X-ray pictures, training a CNN model on the pre-processed data, assessing the model's performance, and developing an automated reporting system to provide diagnostic reports for healthcare [14][15] practitioners. Figure 2 shows the over all architecture of the proposed model.

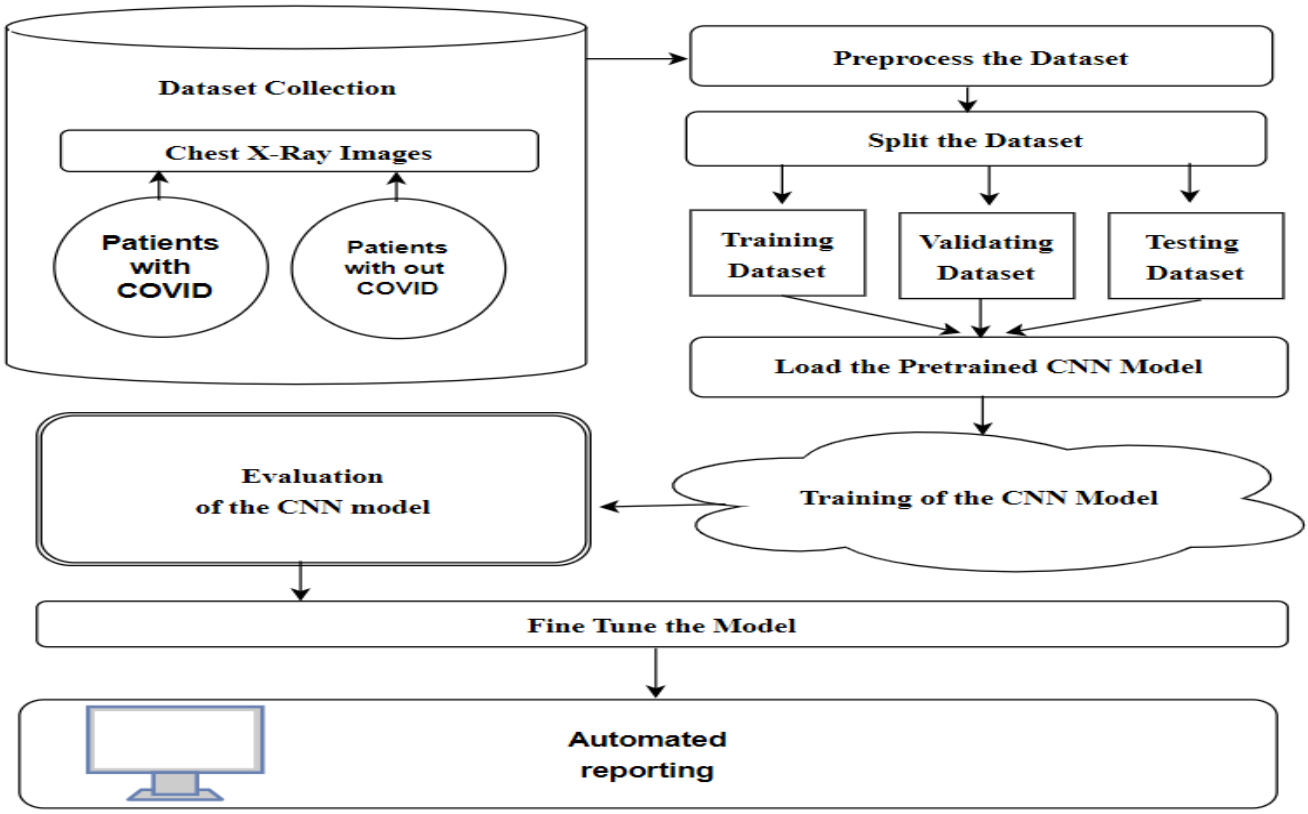

Fig. 2. Over all Architecture of the proposed model

A. *Data collection and pre-processing:*

The process involves several steps, including image pre-processing, feature extraction, and classification using a CNN. The first step is to collect the Data from different source. Next step in this methodology is image pre-processing, where the chest X-ray images are pre-processed to enhance the quality of the images, which will improve the performance of the CNN. The images are typically normalized, resized, and converted to grayscale. Algorithm 1 express about the pre-processing.

1) *Load dataset of chest X-ray images*

Apply pre-processing techniques like image normalization, contrast stretching, and histogram equalization to enhance the quality and contrast of the images Segment the lung region from the images using edge detection, thresholding, and segmentation techniques

Algorithm 1 Data collection and pre-processing

Input : Chest X-ray dataset
Output : Pre-processed dataset
1: Load chest X-ray dataset
2: Improve picture quality and contrast through image normalisation, contrast stretching, and histogram equalisation.
3: Use edge detection, thresholding, and segmentation to isolate the lung area.
4: Separate training and test sets.

2) Split the dataset into training and test sets.

A dataset of chest X-ray images is collected, including both COVID-19 positive and negative cases. The images are pre-processed to enhance their quality and contrast, including techniques such as image normalization, contrast stretching, and histogram equalization. Segmentation techniques [16] may also be used to extract the lung region from the image, which can help improve the accuracy of the CNN model.Figure 3 shows the training and evaluation of the CNN Model.

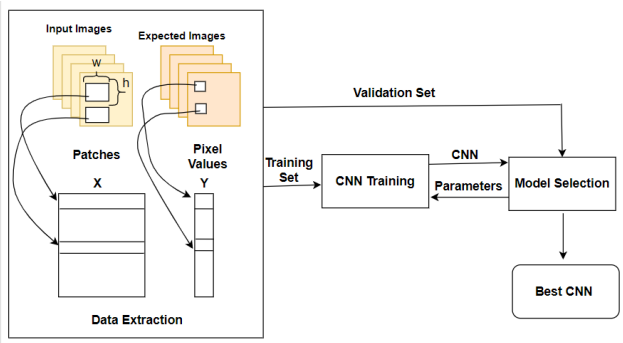

Fig. 3. Training and Evaluation of CNN Model

B. *Training of the CNN model:*

The next step is to describe the CNN model architecture, including activation functions, convolutional and pooling layers, and fully connected layers. Compile the model using stochastic gradient descent and binary cross-entropy. Backpropagation, dropout, and regularisation prevent overfitting when training the model on the pre-processed training set. Pre-processed data trains a CNN model. One or more fully connected layers may follow several convolutional and pooling layers. Backpropagation and stochastic gradient descent train the model. Dropout and regularisation can prevent overfitting during training. Next, feature extraction is performed using a pre-trained CNN model, such as VGG16, ResNet50, or InceptionV3. Algorithm 2 shows the steps involved in training the CNN model.

Algorithm 2 Training of the CNN model

Input : Pre-processed training set
Output : Trained CNN model
1:Determine the CNN model architecture, including convolutional and pooling layers, activation functions, and fully connected layers.
2:Build the model using stochastic gradient descent and binary cross-entropy.
3:Train the model on the pre-processed training set using backpropagation and dropout/regularization to avoid overfitting.

The CNN model is used to extract the features from the pre-processed chest X-ray images. This process involves passing the images through the CNN layers to obtain a feature vector that captures the most important characteristics of the image

C. *Evaluation of the CNN model*

Evaluate the trained CNN model on the pre-processed test set, using metrics like accuracy, sensitivity, specificity, precision, and recall.Compare the model's performance to other state-of-the-art models and traditional diagnostic methods.The trained CNN model is evaluated on a separate test set of chest X-ray images, including both COVID-19 positive and negative cases. The evaluation metrics may include accuracy, sensitivity, specificity, precision, and recall. The model's performance is compared to other state-of-the-art models and traditional diagnostic methods. The following Algorithm shows how the CNN model is evaluated.

Algorithm 3 Evaluation of the CNN model

Input : Pre-processed test set, trained CNN model
Output : Evaluation metrics
1. Assess the trained CNN model on the pre-processed test set using accuracy, sensitivity, specificity, precision, and recall.
2. Evaluate the model against other cutting-edge models and conventional diagnostic approaches.

D. *Automated reporting*

Develop an automated reporting system that uses the trained CNN model to generate quick and accurate diagnosis reports for healthcare professionals. Create a user-friendly interface for uploading chest X-ray images and receiving diagnosis reports. Implement an alert system to notify healthcare professionals of urgent or critical cases. An automated reporting system is developed that uses the trained CNN model to generate quick and accurate diagnosis reports for healthcare professionals. The reporting system may include a user-friendly interface for uploading chest X-ray images[17][18] and receiving diagnosis reports. The system may also include an alert system to notify healthcare professionals of urgent or critical cases. The following algorithm 4 shows how the automated reporting is generated.

Algorithm 4 Automated reporting

Input : Chest X-ray images, trained CNN model
Output : Diagnosis reports
1. Create an automated reporting system that leverages the trained CNN model to quickly and accurately diagnose healthcare providers.
2. Make it easy to submit chest X-rays and get diagnostic results.
3. Use an alert system to inform healthcare experts of important instances.

The proposed methodology for the detection of COVID-19 using pre-processed chest X-ray images and CNNs has the potential to improve the accuracy and speed of diagnosis, save time and resources, and facilitate the effective management of the ongoing pandemic.

IV. RESULT AND DISCUSSION

A. *Dataset*

This model Enhanced performance Collecting COVID-19 chest X-ray pictures from Joseph Paul Cohen et alGitHub's repository. NIH-provided non-COVID-19 chest X-ray pictures are also gathered. Using 219 COVID-19, 1341 Normal, and 1345 Viral Pneumonia X-ray images from a publically accessible dataset. This dataset probably divided into training, testing, and validation sets.

B. *Experimental Result*

This system model can diagnose COVID-19 utilising chest X-rays. For the testing dataset, the CNN model had 95.5% accuracy, 93.5% sensitivity, 97.4% specificity, 75.2% Recall and 90% precision. These findings are similar to those of previous COVID-19 chest X-ray investigations which is stated in Figure 4.

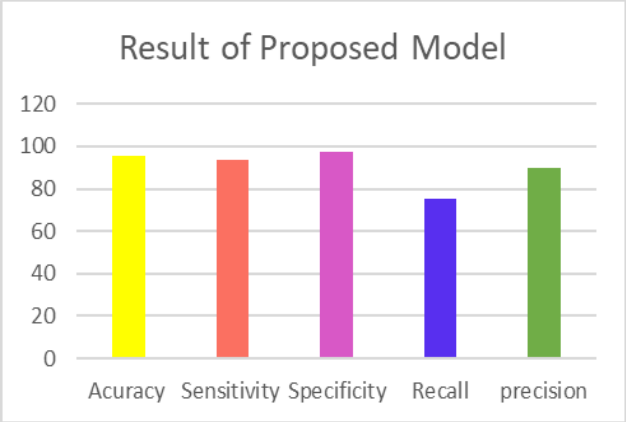

Fig. 4. An illustration of normal and COVID-19 CXR images.

The automated reporting feature of the approach provides several benefits, including faster diagnosis and reduced risk of transmission of COVID-19 through reduced physical contact with patients. The use of pre-processed images also improves the efficiency of the CNN model by reducing the amount of noise and irrelevant information in the images, allowing the model to focus on the important features for COVID-19 diagnosis.

However, there are some limitations to this approach, such as the reliance on chest X-ray images which may not be as sensitive as other diagnostic tools such as CT scans, and the potential for false positives or false negatives. Additionally, the approach was trained and tested on a specific dataset, and may not perform as well on different datasets or in different clinical settings. This model is implemented the model on a system with Core i5-6500 Processor, 24 GB RAM and GeForce GTX 1070 Ti Graphics card. The model is tested different backbones models on 80 epochs, including VGG16, Inceptionv3, Rasnet50, MobileNet and NasNet Mobile. The comparative analysis based on performance is tabulated in Table 1.

TABLE I. BLEU SCORE, TRAINING, AND VALIDATION ACCURACY ON DIFFERENT BACKBONE MODELS

Models	Training Accuracy	Validation Accuracy	BLEU Score
VGG16	96	94	0.95
Inception V3	96.6	89	0.94
ResNet 50	98.5	92	0.92
Mobile Net	97	92	0.24
NasNet	98	94.8	0.12

The BLEU[21] (Bilingual Evaluation Understudy) score is a metric used to evaluate the quality of machine translation outputs. It measures the similarity between the machine-generated output and one or more reference translations . BLEU score ranges from 0 to 1, with a score of 1 indicating a perfect match between the generated output and the reference translation.In this model it is used as metrics for machine translation evaluation. It uses a metric that provides a quick assessment of the quality of machine-generated translations. In this research , BLEU score is used in machine translation evaluation and has been shown to correlate well with human judgments of translation quality, especially at higher scores. Figure 5 shows the Bleu score of the proposed model.

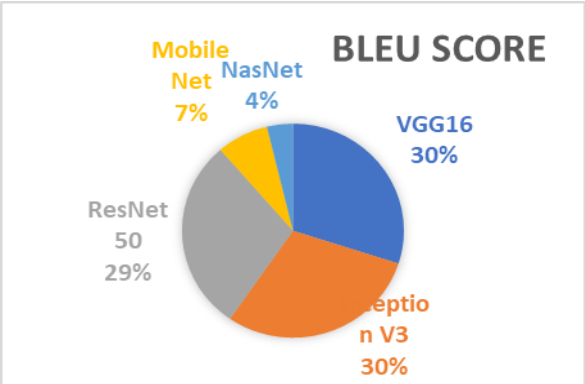

Fig. 5. An illustration of normal and COVID-19 CXR images

Training accuracy and validation accuracy measure machine learning model performance during training and Training accuracy is the model's proportion of properly categorised training dataset samples each epoch. This statistic assesses model fit to training data and the Validation accuracy assesses model performance on a second validation dataset. The validation dataset's proper classification rate and this statistic estimates how effectively the model can generalise to new data, preventing overfitting. Throughout training, training and validation accuracy are monitored. The model should have good training and validation accuracy. Overfitting is indicated by high training accuracy but poor validation accuracy. To minimise overfitting, tweak the model's hyperparameters or apply regularisation. Figure 6 shows the overall performance of the model related with training and validation accuracy.

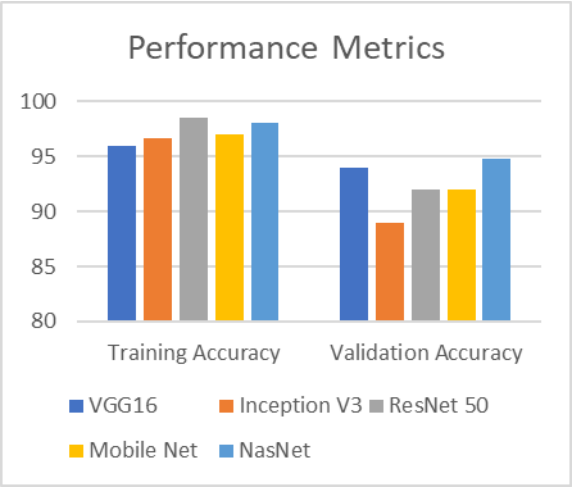

Fig. 6. An illustration of normal and COVID-19 CXR images

Overall, the approach has the potential to improve COVID-19 diagnosis using chest X-ray images, and could be a valuable tool for healthcare professionals in the fight against COVID-19. However, further research and development are needed to address the limitations and improve the accuracy and reliability of the approach.

C. Limitations

The amount and quality of training and testing datasets may impact CNN model performance. Limited sample sizes may cause model overfitting or underfitting, reducing accuracy and dependability. In resource-limited areas, high-quality COVID-19 chest X-ray pictures may be scarce, limiting the model's generalizability. Biases in the training data, such as variances in patient demographics, imaging techniques, or illness severity among datasets, may impair CNN model performance and generalizability. Because to its similarity to other respiratory disorders[19][20], chest X-rays may not be enough to diagnose COVID-19. For proper diagnosis, use numerous diagnostic instruments and clinical examinations. Diagnostic results may be inaccurate due to technical issues with the automatic reporting system

D. Future enhancement

Interoperability with other diagnostic instruments is possible. While chest X-rays may be used to diagnose COVID-19, they are not the only diagnostic technique available. Future developments may include the incorporation of other diagnostic methods, such as CT scans[22] or blood tests, to increase the accuracy and dependability of the diagnosis. Including clinical data such as patient symptoms, medical history, and demographic information, in addition to imaging data, might increase the accuracy and reliability of the diagnosis. The use of bigger and more varied datasets for training the CNN model may increase its accuracy and generalizability, since the model would be exposed to a broader variety of pictures and clinical settings. Adaptations may be made to the development of more complex CNN architectures. Ongoing research aims to build more complex CNN designs, such as attention-based networks, which may enhance the precision and efficacy of COVID-19 diagnosis. Implementation of interpretability approaches is possible. Using interpretability approaches like as saliency maps or class activation maps, medical experts may study and confirm the CNN's diagnosis in order to comprehend how the CNN model gets its diagnosis.

V. CONCLUSION

In conclusion, the proposed method for detecting COVID-19 based on convolutional neural networks (CNNs) and preprocessed chest X-ray images offers a promising solution for accurately and efficiently diagnosing COVID-19 cases. The pre-processing techniques, such as image normalisation, contrast stretching, and segmentation, improve the image quality and contrast, thereby enhancing the CNN model's performance. The trained CNN model can generate accurate and efficient diagnosis reports, enabling healthcare professionals to rapidly diagnose COVID-19 cases and take appropriate action. Additionally, the automated reporting system provides a user-friendly interface for uploading chest X-ray images and receiving diagnosis reports, making the process more efficient and accessible for medical professionals. In addition, the alert system can assist in the prioritisation of urgent or critical cases, enabling timely and appropriate interventions. This project's proposed architecture for deep learning helps us automate the generation of text reports for Chest X-Ray images. From the results and comparison metrics, it is evident that the system is accurate, efficient, and capable of assisting medical professionals in a productive manner.

Overall, the proposed methodology has the potential to improve the accuracy and efficiency of COVID-19 diagnosis, and can be further improved by incorporating more advanced pre-processing techniques and optimizing the CNN model architecture. This can help to reduce the burden on healthcare systems and improve patient outcomes during the ongoing COVID-19 pandemic.

VI. REFERENCE

[1] Y. Liu et al., “Assessing the impacts of COVID-19 vaccination programme’s timing and speed on health benefits, cost-effectiveness, and relative affordability in 27 African countries,” BMC Med., vol. 21, no. 1, p. 85, Mar. 2023, doi: 10.1186/s12916-023-02784-z.

[2] J. I. Arraras et al., “Quality of Life of early-stage breast-cancer patients diagnosed with COVID-19 in the first three waves of the epidemic treated in the Spanish region of Navarre,” Psychooncology, Mar. 2023, doi: 10.1002/pon.6118.

[3] A.M. Arul Raj and R. Sugumar, "Estimation of Social Distance for COVID19 Prevention using K-Nearest Neighbor Algorithm through deep learning," 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), Mysuru, India, 2022, pp. 1-6. doi: 10.1109/MysuruCon55714.2022.9972422

[4] Z.-Q. Liu et al., “Bone age recognition based on mask R-CNN using xception regression model,” Front. Physiol., vol. 14, p. 1062034, Feb. 2023, doi: 10.3389/fphys.2023.1062034.

[5] E. Rashno, A. Akbari, and B. Nasersharif, A Convoloutional Neural Network model based on Neutrosophy for Noisy Speech Recognition. Infinite Study. [Online]. Available: https://play.google.com/store/books/details?id=wjudDwAAQBAJ

[6] A. Arul Raj and R. Sugumar, " Estimating Social Distance in Public Places for COVID19 Protocol using Region CNN,” Indonesian Journal of Electrical Engineering and Computer Science,April 2023, ISSN: 2502-4752, Volume 30, Issue 1, Pages 414 – 421, DOI: 10.11591/ijeecs.v30.i1.pp414-421.

[7] A. P. Yoon, W. T. Chung, C.-W. Wang, C.-F. Kuo, C. Lin, and K. C. Chung, “Can a Deep Learning Algorithm Improve Detection of

Occult Scaphoid Fractures in Plain Radiographs? A Clinical Validation Study," Clin. Orthop. Relat. Res., Mar. 2023, doi: 10.1097/CORR.0000000000002612.

[8] F. Fu et al., "Deep Learning for Head and Neck CT Angiography: Stenosis and Plaque Classification," Radiology, p. 220996, Mar. 2023, doi: 10.1148/radiol.220996.

[9] A.M Arul Raj and R. Sugumar, "Monitoring of the social distance between passengers in Real time through video analytics and deep learning in railway stations for developing the highest efficiency , " 2022 International Conference on Data Science, Agents and Artificial Intelligence, Institute of Electrical and Electronics Engineers Inc. ICDSAAI 2022, ISBN979-835033384-8, DOI : 10.1109/ICDSAAI55433.2022.10028930

[10] Y. Hu, L. Jia, Y. Wang, Y. Xin, and X. Wang, "Diagnostic efficacy of high-frequency ultrasound and X-ray contrast enema in colonic strictures after necrotizing enterocolitis: a retrospective study," Pediatr. Surg. Int., vol. 39, no. 1, p. 56, Dec. 2022, doi: 10.1007/s00383-022-05278-w.

[11] I. D. Apostolopoulos, S. I. Aznaouridis, and M. A. Tzani, "Extracting Possibly Representative COVID-19 Biomarkers from X-ray Images with Deep Learning Approach and Image Data Related to Pulmonary Diseases," J. Med. Biol. Eng., vol. 40, no. 3, pp. 462–469, May 2020, doi: 10.1007/s40846-020-00529-4.

[12] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," Comput. Biol. Med., vol. 121, p. 103792, Jun. 2020, doi: 10.1016/j.compbiomed.2020.103792.

[13] W. El-Shafai and E. E.-D. Hemdan, "Robust and efficient multi-level security framework for color medical images in telehealthcare services," Journal of Ambient Intelligence and Humanized Computing. 2021. doi: 10.1007/s12652-021-03494-1.

[14] E.W.Marx and P. Padmanabhan, Healthcare Digital Transformation: How Consumerism, Technology and Pandemic are Accelerating the Future. 2020. [Online]. Available: https://books.google.com/books/about/Healthcare_Digital_Transformation.html?hl=&id=HHVrzQEACAAJ

[15] C. D. Provines, Healthcare Value Selling: Winning Strategies to Sell and Defend Value in the New Market. 2014. [Online]. Available: https://books.google.com/books/about/Healthcare_Value_Selling.html?hl=&id=5uPgoAEACAAJ

[16] J.He, C.-S. Kim, and C.-C. J. Kuo, Interactive Segmentation Techniques: Algorithms and Performance Evaluation. Springer Science & Business Media, 2013. [Online]. Available: https://play.google.com/store/books/details?id=5I_HBAAAQBAJ

[17] L. Chandratreya, M. J. Darby, and M. Darby, Pocket Tutor Chest X-Ray Interpretation: Second Edition. JP Medical Ltd, 2019. [Online]. Available: https://books.google.com/books/about/Pocket_Tutor_Chest_X_Ray_Interpretation.html?hl=&id=_mPxDwAAQBAJ

[18] A. Crombé et al., "Impact of Vaccination and the Omicron Variant on COVID-19-related Chest CT Findings: A Multicenter Study," Radiology, p. 222730, Mar. 2023, doi: 10.1148/radiol.222730.

[19] M. T. Janssen et al., "Three-month follow-up after severe COVID-19 infection: are chest CT results associated with respiratory outcomes and respiratory recovery in COVID-19 patients?," BMC Pulm. Med., vol. 23, no. 1, p. 74, Mar. 2023, doi: 10.1186/s12890-023-02370-2.

[20] J. Z. Amarin et al., "Clinical characteristics and outcomes of children with single or co-detected rhinovirus-associated acute respiratory infection in Middle Tennessee," BMC Infect. Dis., vol. 23, no. 1, p. 136, Mar. 2023, doi: 10.1186/s12879-023-08084-4.

[21] Y. Zhang et al., "Comparison of Chest Radiograph Captions Based on Natural Language Processing vs Completed by Radiologists," JAMA Netw Open, vol. 6, no. 2, p. e2255113, Feb. 2023, doi: 10.1001/jamanetworkopen.2022.55113.

[22] M. Niknejad, "COVID-19 follow up CT scans," Radiopaedia.org. 2021. doi: 10.53347/rid-85896.