IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 18 July 2023, accepted 31 July 2023, date of publication 4 August 2023, date of current version 10 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3302180

RESEARCH ARTICLE

Classification and Localization of Multi-Type Abnormalities on Chest X-Rays Images

ABDUSSALAM ELHANASHI[1], (Member, IEEE), SERGIO SAPONARA[1], AND QINGHE ZHENG[2], (Member, IEEE)

[1]Dipartimento di Ingegneria dell'Informazione, University of Pisa, 56126 Pisa, Italy
[2]School of Intelligent Engineering, Shandong Management University, Jinan, Shandong 250357, China

Corresponding author: Abdussalam Elhanashi (a.elhanashi@studenti.unipi.it)

This work was supported by the Restart Toscana Project with the University of Pisa.

ABSTRACT Chest X-ray images are among the most common diagnostic tools for detecting and managing bronchopneumonia and lung abnormalities, such as those caused by COVID-19. However, interpreting these images requires significant expertise, and misinterpretations can result in false negatives or positives. Deep learning techniques have recently been highly effective in analyzing medical images, including chest X-rays. In this study, we propose two deep learning approaches to classify and localize different abnormalities, including COVID-19, on chest X-rays, which include multi-classification and object detection models that can identify and localize the presence of disease as other common abnormalities. The proposed models are trained on a large dataset of chest X-ray images from sick people (including COVID-19 patients) and validated on an independent test set. Compared to single object models, this paper presents an ensemble of models by combining multiple object detection models to detect multiple abnormalities in the chest X-ray images. Our results demonstrate that the proposed method achieved promising results in both multi-classification and localization of abnormalities, including COVID-19, compared to the state-of-the-art methodologies. The proposed methods have the potential to assist radiologists in the diagnosis of the abnormalities on chest X-ray images and provide a more accurate and efficient interpretation, thereby improving patient outcomes and reducing the burden on healthcare systems.

INDEX TERMS Deep learning, multi-classification, localization, ensemble model, bronchopneumonia/lung abnormalities.

I. INTRODUCTION

Chest X-ray imaging is one of the primary diagnostic tools used to evaluate the presence of abnormalities in the lungs, which may include many different types of sickness to be detected, such as lung opacity, viral pneumonia, and recently, effects due to COVID-19 pandemic which, since December 2019, has presented a global challenge [1]. The analysis of chest X-ray images can provide valuable information on the severity and progression of COVID-19 pneumonia, making it an essential tool in managing the disease. However, analyzing chest X-ray images can be challenging, particularly in cases with multiple abnormalities. One of the difficulties encountered with chest X-ray images involves the potential confusion when diagnosing pneumonia, along with other types of abnormalities such as COVID-19, viral pneumonia, and lung opacity. This becomes particularly challenging when the radiologist or diagnosing doctor lacks experience or when the patient's medical history is unavailable. Previous studies have predominantly focused on binary or single-class classification tasks, where abnormalities are simply identified as present or absent. However, simultaneously classifying multiple types of abnormalities on chest X-rays presents additional challenges. Additionally, there is a need to improve the accuracy of localizing abnormalities on chest X-ray images, despite the advancements made in object detection models for localization. There is still room for improvement in this area. Future research should aim to develop robust algorithms capable of distinguishing between various disease categories and accurately pinpointing the location of abnormalities

The associate editor coordinating the review of this manuscript and approving it for publication was Yongming Li.

within the image, thus providing detailed information about their spatial extent. Accurately identifying the presence and location of COVID-19 abnormalities on chest X-ray images can significantly improve diagnostic accuracy and facilitate prompt treatment. A chest X-ray image is a susceptible technique that can detect COVID-19 and other chest-related issues. It plays a crucial part in identifying and treating the disease at an early stage. Manual interpretation of these images is time-consuming and error-prone and can be particularly challenging when multiple abnormalities are present. Therefore, algorithms for fast (real-time or near real-time) detection and classification of abnormalities in chest X-ray images are needed. Research has demonstrated that chest X-ray (CXR) images exhibit distinct variations in the imaging features of abnormalities and typical pneumonia [2], [3], [4]. Deep learning (DL) algorithms have shown promising results in different medical imaging applications, which include chest X-ray image analysis. Deep learning algorithms can potentially analyze large datasets, extract complex features, and provide accurate predictions, making them suitable for multi-classification and localization tasks. Various issues, such as image classification, object detection, medical imaging, and drug interaction, are addressed using deep neural networks [5], [6]. In image processing, Convolutional Neural Networks (CNN) have demonstrated significant potential, and multiple research studies have shown the efficacy of these techniques in image segmentation [7], [8]. Binary classification is a deep learning approach that classifies data into two categories. Chest X-ray images have diseases classified as either showing the presence or absence of abnormalities. Training deep learning models achieves this classification on a large dataset of X-ray images labeled as negative or positive for diseases. However, there are several drawbacks to using binary classification on these X-ray images for disease diagnosis. Chest X-ray images are not specific to a particular abnormality such as COVID-19 and can be positive for other respiratory conditions such as pneumonia or tuberculosis. This can result in false prediction and misidentification of coronavirus [9], [10], [11], [12]. Additionally, chest X-ray images are less sensitive than other diagnostic methods, such as computed tomography (CT) scans. This means there may be cases of such disease that are missed or misclassified as negative. Therefore, there is a need to develop efficient and accurate approaches for localization and multi-classification of multiple abnormalities on X-ray images of the lungs. We will utilize deep learning models that will CNN pre-trained architectures [13], [14] and object detection models [15], [16], which have been demonstrated to be effective in classifying and localizing the targeted objects in the images. The proposed frameworks will be trained and validated on vast chest X-ray images containing different abnormalities. We will evaluate its performance using standard metrics such as recall, accuracy, precision, receiver operating characteristic curve for multi-classification, and mean average precision for object detection. We aim to develop an efficient and accurate system that automatically detects the presence and location of coronavirus abnormalities, viral pneumonia, and other opacities. It addresses the critical issue of accurate and efficient multi-classification and localization of coronavirus abnormalities on X-ray images. The objective of the object detection is to develop a robust algorithm that can identify and locate the abnormalities with bounding boxes on chest x ray images. The model is able to categorize the radiographs into different classes, including negative for pneumonia, typical, indeterminate, or atypical for COVID-19. The proposed system utilized imaging data and annotations provided by a group of radiologists. The goal is to improve the mean average precision and efficiency of abnormalities detection using chest radiographs, which can aid in the diagnosis and management of the disease. The proposed framework can potentially improve diagnostic accuracy and facilitate prompt treatment of coronavirus patients, ultimately mitigating the ongoing pandemic. The summarized contributions of this research are presented as the following:

1) This paper proposes two deep learning techniques, which include multi-classification, and object detection of abnormalities in the chest X-ray images by utilizing two different datasets that are dedicated to each method. The models proposed in these approaches employ Convolutional Neural Networks (CNNs) that can automatically learn and extract relevant features, enabling accurate classification of different abnormalities and localization of the objects from complex patterns.

2) The multi-classification approach is employed to classify chest X-ray images into one of four categories, which include COVID-19, normal, lung opacity, and viral pneumonia. This enables clinicians to distinguish between multiple respiratory diseases and healthy lungs.

3) A novel ensemble model approach has been implemented to combine the results of three object detection models (EfficientNet+YOLOv7+Faster R-CNN) by applying weighted box fusion techniques. The proposed method improved the mean average precision and reduced the overfitting risk.

4) We employ advanced image pre-processing techniques, including augmentation, cropping, rotation, and normalization, to improve the quality and consistency of the chest X-ray images.

The motivation of this research is to exploit two different approaches, which include multi-classification, and localization to enhance medical diagnosis and patient care. By developing accurate algorithms, healthcare professionals can efficiently classify and localize various abnormalities, such as pneumonia, negative, and lung diseases, in chest X-rays. Early diagnosis and precise localization can lead to timely treatments, better outcomes, and reduced healthcare costs. This research aims to leverage deep learning approaches to improve the efficiency and accuracy of diagnosing chest-related medical conditions.

All of our codes for the proposed two approaches are available on the GitHub link at

https://github.com/ahanashi/Classification-and-localization-abnormalities-on-Chest-x-ray-images

The paper is structured as the following:- Section I describes the introduction. The related works are presented in Section II. Section III describes the overview of CNN for feature extraction in medical images. Section IV presents the experimental methodology of this research. Section V describes the obtained results and discussion. Finally, the conclusions and future work are drawn in Section VI.

II. RELATED WORKS

Chest X-ray and computed tomography (CT) images for patients are a relatively simple and inexpensive procedure, making their use in identifying coronavirus feasible in many countries. In a published article by [25], researchers utilized three distinct CNN models, ResNet50, InceptionV3, and Inception-ResNetV2, to identify coronavirus pneumonia in patients afflicted with the disease using X-ray radiographs [26]. Chest X-ray (CXR) is a prevalent medical tool that has the potential to offer an efficient diagnosis in medical applications. In order to facilitate the successful screening of individuals afflicted with abnormalities, numerous research studies utilizing algorithms of deep learning to identify such diseases have been carried out [27], [28], [29], [30], [31], [32]. Medical images effectively detect abnormalities, which can control the spread of the disease. Among the imaging techniques used for diagnosing coronavirus, X-ray images, CXR, and CT are crucial for diagnosing the disease. The concept of image diagnostic systems has been extensively studied using various approaches, which include feature learning. Convolutional Neural Networks (CNNs) are a prevalent and practical approach for detecting disease from digital images. Recent research has been carried out to highlight the contributions of convolutional neural networks for detecting coronavirus [33], [34]. For instance, Inception network-based CNN was utilized in [23] for detecting COVID-19 abnormalities within CT. At the same time, a modified ResNet-50 model was applied [20] to identify COVID-19 presence and bacterial pneumonia in the CT images. In [35], CXR was utilized, and a convolutional neural network was implemented based on different ImageNet architectures deep neural network (DNN) to extract features from the X-ray images. The extracted features were fed into a Support Vector Machine (SVM) to identify COVID-19 cases. The authors [36] utilized five AI-based models to detect coronavirus based on blood cell count. They applied a Bayesian optimization technique to tune the hyperparameters of the algorithms and transparently presented the results to allow for straightforward interpretation by healthcare professionals. Karim et al. [37] used a DNN model for coronavirus detection in chest X-ray images. They highlighted the class-distinguishing regions in the images using class activation maps of gradient and relevance propagation. They also provided an illustratable approach for description to help clinicians understand the reasoning behind the AI model's decision. In [38], the authors introduced an approach that automatically detects coronavirus in CT images utilizing DCNN models. The transfer learning approach uses a specific input to obtain the best accuracy. To help visualize the predicted models, they used visualization methods to illustrate better and explain the model's predictions. The authors [39] introduced a technique that involved utilizing deep learning to analyze the X-ray images of the lungs. They incorporated pre-trained models with domain extension transfer learning (DETL) to develop their system. Their focus was explicitly on identifying instances of coronavirus in the X-ray images. Author [40] implemented a convolutional neural network to detect COVID-19 disease. Although the employed models were validated with a limited set of chest X-ray images, specifically those showing positive cases in coronavirus patients, earlier studies faced a problem with imbalance for the classes. Additionally, these studies exclusively employed CNN as their pre-trained models for COVID-19 classification and segmentation [41]. It takes only a few minutes to complete the examination of a chest X-ray image. Although CT analysis provides more detailed information, it is considerably more costly and less accessible in low-income countries. On the other hand, diagnosing lung diseases through X-ray analysis is much more attainable, particularly in low-income nations. Nevertheless, in the initial stages of the illness, chest X-ray image analysis is not as informative as CT data analysis, as mentioned in the references [42]. Therefore, Jaiswal et al. employed deep transfer learning based on DenseNet201 to categorize CT images of patients who had contracted COVID-19 [43]. Their findings showed that CNN had high efficacy and achieved promising performance, allowing them to identify COVID-19 with this approach.

Mask R-CNN [44], [45] is a more advanced model for detecting objects than other models. It is an essential artificial intelligence method used for various purposes, such as identifying and separating nodules of lungs [46], segmentation of the liver [47], segmentation of multiorgan [48], face detection, hand segmentation [49], segmentation of gastric cancer [50], [51], and classification the breast tumor [52]. It is an improved version of object detector and has shown superior performance in object detection and segmentation compared to other models. Table 1 presents an overview of the key features of previous research studies mentioned in the recent literature for utilizing object detection and classification methods with DL techniques to automatically detect coronavirus abnormalities on chest X-ray images.

III. CNNS FOR FEATURE EXTRACTION IN MEDICAL IMAGES

The field of computer vision has been transformed by Neural Networks architectures, which are now the leading method for numerous visual recognition tasks, including feature extraction from medical images [53]. CNNs have been employed in medical imaging analysis, which has recently

TABLE 1. A summary of research endeavors employing deep learning methods with classification and object detection to identify COVID-19 cases, including their methodologies and performance.

Research article	Method	Model	Performance
Sitaula et al [17]	Classification	VGG16	87.49% Acc
Tiwari et al. [18]	Object detection	YOLOv5, RetinaNet-101	0.553 mAP
Marusani et al [19]	Object detection	YOLOv5	0.57 mAP
Song et al. [20]	Classification	ResNet50	86% Acc
Stephen et al. [21]	Classification	CNN	93% Acc
Wang et al. [22]	Classification	Tailored	87% Acc
Kang et al [23]	Classification	CNN	82.9% Acc
Sangeetha et al. [24]	Object detection	YOLOv5	0.55 mAP
Sangeetha et al. [24]	Object detection	Faster R-CNN	0.50 mAP

increased due to their capability to extract intricate features from images with little to no manual intervention. Medical images such as X-rays and MRI scans contain informative features that can help doctors make accurate diagnoses and provide appropriate treatment. However, interpreting medical images can be challenging even for experienced physicians, as these images often contain complex structures and patterns that are difficult to detect and analyze. Previously, researchers have developed various methods for obtaining features from images with high and low levels of the characteristics. These features include corners, edges, groups of color intensity, and features that remain consistent across scales, such as Scale-Invariant Feature Transform (SIFT) and speeded up robust features SURF. [54], [55]. These techniques have gained interest among researchers as they are impervious to image rotation and scaling, previously considered significant hurdles in computer vision and medical imaging. These features then train deep learning models to complete a specific supervised classification task. One drawback of using this method is that the effectiveness of the machine learning model selected would be heavily dependent on extracted features quality from the images. Features quality can be affected by different conditions such as lighting, the object's position in the image, noise, and other variables. Therefore, implementing solutions that rely on manually crafted features may not be as successful when tested in uncontrolled environments, as indicated by the results obtained from controlled testing. CNNs are deep learning architectures consisting of multiple convolutional filters that extract features from images. These filters are learned through training on large datasets of images, and the CNN is optimized to reduce the error between the predicted and actual outputs. The out of every convolutional layer is fed through an activation function, such as the Rectified Linear Unit (ReLU), which brings non-linear elements into the network, enabling it to learn complex features in the input data. In addition, it uses local receptive fields to capture local patterns in the input data. This allows the network to detect small patterns, such as illuminations, corners, and textures, which can then be combined to create more complex features. The spatial dimensions of the feature maps are reduced by pooling layers while retaining essential

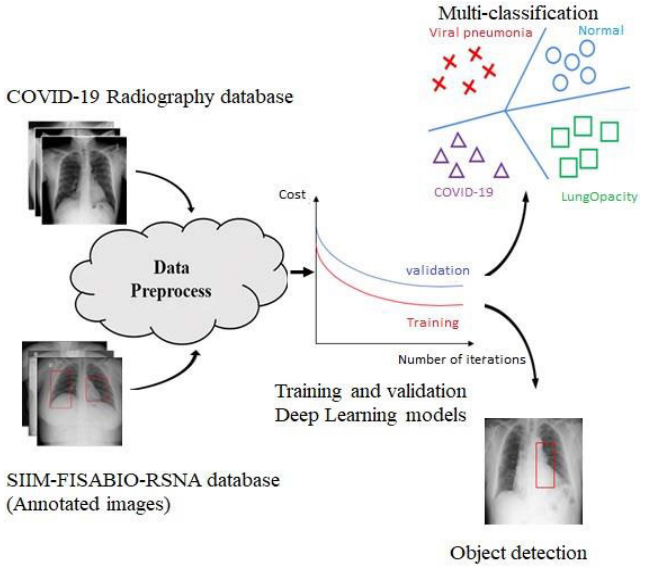


FIGURE 1. The proposed framework includes multi-classification and object detection of abnormalities on chest X-ray images. The workflow illustrates the utilization of two different databases that are preprocessed and then fed to the proposed deep-learning models for training and validation. Finally, we obtained the outcome results from each approach.

TABLE 2. The main description and parameters of COVID-19 Radiography, and SIIM-FISABIO-RSNA database.

Parameters	COVID-19 Radiography database	SIIM-FISABIO-RSNA
Image resolution	299 x 299	512 x 512
Image format	PNG	DICOM
Number of images	21165	6334
Annotation availability	No	Yes
Imaging modalities	X-ray images	X-ray and CT-scan
Publication year	2020	2021

features. This reduces the computational complexity of the network and helps to prevent overfitting. CNNs are used for different tasks in medical imaging applications, including segmentation, classification, and detection. CNNs are trained on vast datasets, which allow them to generalize effectively to new images and perform well on unseen data. Figure 1 shows the proposed two deep learning approaches, which include multi-classification, and object detection of abnormalities by utilizing two different databases of chest x-ray images.

IV. METHODOLOGY

A. DATASET COLLECTION

This research used two datasets to train and evaluate the proposed models. These datasets include the COVID-19 Radiography database, which targets multi-classification, and the SIIM-FISABIO-RSNA database, which is utilized for localization and object detection of abnormalities on X-ray images for lungs. In addition, these datasets are used to assess the robustness of the proposed models to multi-classify and locate the presence of abnormalities on chest X-ray images. Table 2 shows the parameters and description of each dataset.

These datasets were gathered from real cases of cases who have tested positive and negative for the virus. The two datasets are described as the following.

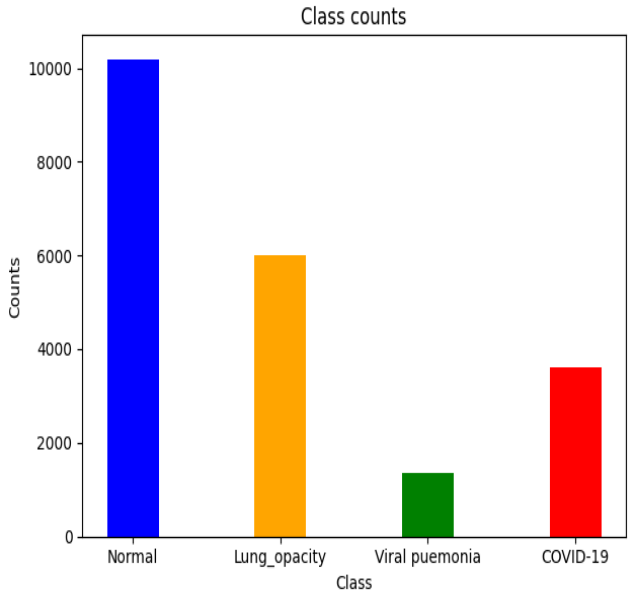


FIGURE 2. Distribution for different classes of COVID-19 Radiography dataset.

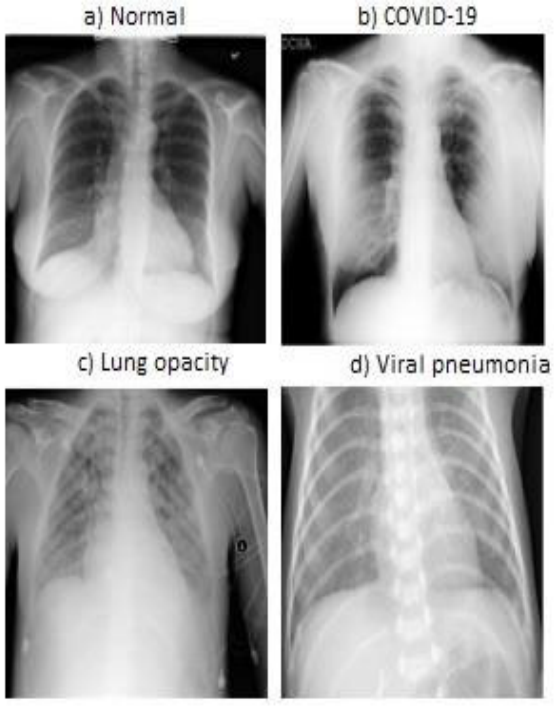


FIGURE 3. Four classes for radiography database, which include a) Normal, b) COVID-19, c) Lung opacity, and d) Viral pneumonia.

1) COVID-19 RADIOGRAPHY DATABASE

COVID-19 Radiography database, available on Kaggle [56], [57]. It is a dataset created by medical doctors, researchers, and scientists from Qatar University and Dhaka. It includes X-ray images of lungs that can be used for COVID-19 detection, which include categories, see Figure 3. The images were collected from various sources, including public repositories, scientific articles, and COVID-19 image databases. This Database contains 21165 chest X-ray images, with 3616 COVID-19, 6012 lung opacity, 10192 normal, and 1345 viral pneumonia cases. Figure 2 shows the distribution for different classes of X-ray images. The image format is in PNG and has a 299 × 299 pixels resolution. Each image is labeled with its corresponding case, which can be used to train and validate deep learning models. The significant advantages of this dataset are its size, which makes it suitable for building and evaluating deep learning models. The dataset also includes images with different severities of coronavirus cases, which can develop models distinguishing mild and severe cases. Another advantage of this dataset is that the pre-trained CNNs can be trained on this Database. Pre-trained architectures include the ResNet50, VGG19, VGG16, and Xception, trained on chest X-ray images. COVID-19 Radiography Database has limitations, which do not include patient information, such as age, sex, or medical history, which can be crucial in abnormalities diagnosis.

2) SIIM-FISABIO-RSNA COVID-19 DATASET

It is a vast available dataset of X-ray and chest CT images, along with their corresponding clinical data, intended for developing and testing deep learning (DL) algorithms for detecting and diagnosing diseases. The Society created the dataset for Imaging Informatics in Medicine (SIIM) [58]. This Database includes 6334 in Digital Imaging and Communications in Medicine (DICOM) format anonymized images from different ages of patients, including individuals with COVID-19 and those with other respiratory diagnoses. In addition, a panel of expert radiologists has labeled these images for further appearances of illness. The availability of such a large and diverse dataset is expected to accelerate the improvements of deep learning algorithms for abnormalities detection and analysis, which could ultimately help enhance the outcomes of the patients and minimize the time-consumption on healthcare systems. In addition, the dataset has been categorized and labeled with four different classes, typical, atypical, indeterminate, and negative for pneumonia, as shown in Figures 4 and 5. These labels correspond to various patterns observed in the chest images and can help doctors diagnose and manage coronavirus and other lung abnormalities.

B. THE PROPOSED METHODS

Medical imaging has led to an urgent need for efficient disease diagnosis. Deep learning models have significant potential to detect abnormalities from medical images. In this research, we will explore using multi-classification and object detection algorithms to classify and detect coronavirus diseases the other abnormalities on the chest on X-ray images. Multi-classification models are employed to classify the images into multiple categories. Multi-classification models can be trained to classify X-ray images such as COVID-19, pneumonia, or normal. This approach efficiently classifies coronavirus cases while identifying other chest abnormalities

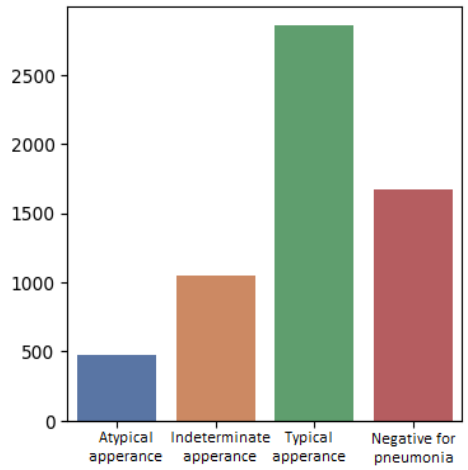


FIGURE 4. Distribution for different classes for the dataset, and as shown, typical appearance has the most images versus the other classes in this dataset.

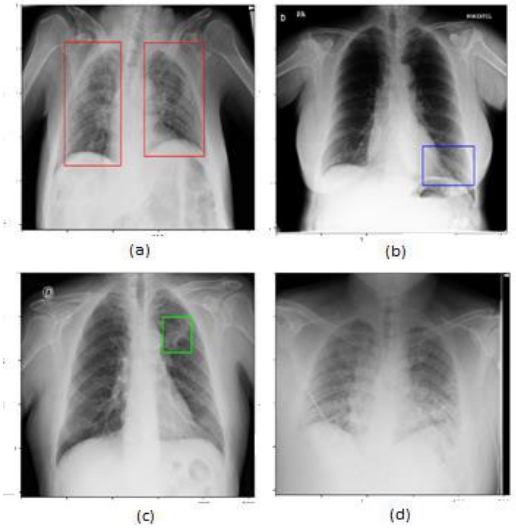


FIGURE 5. Four classes for the proposed dataset to detect and localize abnormalities on the X-ray images.

on the X-ray images. Object detection models, however, are designed to detect abnormalities within an image. Therefore, an object detection model can be trained to detect special patterns along with these abnormalities and highlight them in the image. This approach provides a more precise diagnosis by identifying specific chest areas affected by COVID-19.

1) MULTI-CLASSIFICATION OF COVID-19 ON CHEST X-RAY IMAGES

The multi-classification technique is illustrated in Figure 6. It includes several analysis blocks that involve image pre-processing, exploiting the proposed CNN models, model training, and fine-tuning parameter setting. This study aims to explore different convolutional neural network architectures employed on medical images of chest X-rays to identify COVID-19. The research utilized four highly efficient pre-trained CNN models, which include VGG16 [59], VGG19 [60], ResNet50 [61], and Xception [62], to accomplish the multi-classification task of detecting diseases in the images. Table 3 shows four neural networks' characteristics, which relate to their size, the number of parameters, and the size of an input image. The description of these convolutional neural network models is presented in the following:

TABLE 3. The main characteristics of four pre-trained neural network models.

Network	Size (MB)	Parameters (Millions)	Input image
VGG16	515	138	224x224
VGG19	535	144	224x224
ResNet50	96	25.6	224x224
Xception	85	22.9	299x299

VGG16 is a deep learning model consisting of 16 convolutional layers, which a research group at the University of Oxford has implemented. This model is utilized for object recognition and classification tasks on images. VGG16 architecture uses multiple 3 × 3 convolutional and max-pooling filters to learn more complex features from the input images. VGG16 model has been validated on the bench of datasets such as ImageNet Large Scale Visual Recognition Challenge (ILSVRC) and achieved promising performance in comparison to the other deep learning pre-trained models.

VGG19 is an upgraded version of VGG16. It has 19 layers. It uses small 3 × 3 convolutional filters for feature extraction from the images. It increases the channel numbers and reduces the spatial size of the feature maps. VGG19 has been utilized in different applications, including image recognition, object detection, and style transfer. The architecture can extract and classify the features in images, making it popular in computer vision research and applications. VGG19 has been used in medical imaging to diagnose diseases and in artistic style transfer to create unique visual effects.

ResNet50 is a pre-trained model that Microsoft Research Group has developed. This architecture is utilized for image classification and recognition tasks. ResNet50 has 50 layers, and it utilizes a technique called residual learning. This technique allows the architecture to learn effectively by using skip connections to bypass specific layers. This architecture has produced highly accurate results in various image-related tasks, including object detection and facial recognition. As a result, it has been widely adopted in both research and industry for multiple applications in computer vision.

XCEPTION was introduced by Franç ois Chollet in 2016 and is based on the Inception architecture. This model utilizes depth-wise separable convolutions to enhance efficiency while maintaining accuracy. To achieve this, the convolution operation is divided into depth-wise and point-wise convolution, reducing the number of parameters. This results in a network with fewer parameters and a lower computational cost while maintaining the performance on image classification and recognition tasks. Xception has been utilized in various applications, including object detection, segmentation, and image captioning.

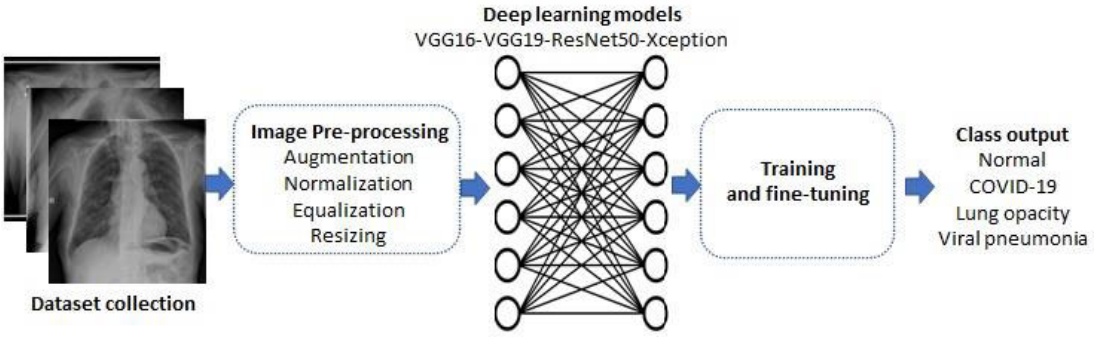


FIGURE 6. The proposed technique for multi-classification, which involve dataset collection, image pre-processing, and training deep learning models.

We started the experiments with the pre-processing activities. Augmentation is an approach, which improves the performance of deep learning models. It enhances the model's ability to generalize by applying different transformations to the original data. The following steps have been performed as part of pre-processing the images for the COVID-19 Radiography database:-

- The images have been rotated by ±10%, scaling them by ±10%, and translating them by ±10% in both the horizontal and vertical directions. In addition to that zooming has been applied by ±15% in both the horizontal and vertical directions.
- The images have been converted into pixel arrays and performed data normalization. This process involves scaling the pixel values to fit within the interval of [0,1].
- Adaptive histogram equalization was utilized to minimize distortion and improve the brightness of the images.
- Finally, the normalized images have been resized to a standardized size of 224 × 224 pixels. This is to meet the requirements of the input layer of each pre-trained Convolutional Neural Network (CNN) model used in the image pre-processing pipeline for the proposed Database.

80% of the dataset has been used for training and 20% for validation. Table 4 illustrates the training and testing split for the proposed dataset. The pre-trained model's Stochastic gradient descent with momentum optimizer was utilized for training, with a value of momentum at 0.9 and a gradient threshold method of L2 norm. The minimum batch size was 32, while the maximum number of epochs was 50. The learning rate was selected at 0.0003 and remained constant throughout training. A factor for L2 regularization (weight decay) was also introduced at 0.0001.

2) LOCALIZATION OF THE ABNORMALITIES OF COVID-19 ON CHEST X-RAY IMAGES

The proposed models for object detection have been trained and evaluated on the SIIM-FISABIO-RSNA dataset.

Object detection has widely emerged in video and imaging applications. There are two popular approaches for this task,

TABLE 4. The COVID-19 radiography database used in our research.

Category	Train	Test
COVID-19	2893	723
lung_opacity	4810	1202
normal	8154	2038
viral_pnuemonia	1076	269
Total	16933	4232

one-stage, and two-stage detectors. One-stage detectors, such as YOLO and single shot detector (SSD), detect objects in a single pass through the network, making them faster but less accurate than two-stage detectors. On the other hand, two-stage detectors, such as Faster R-CNN, employ a region proposal network to produce feasible object locations. Subsequently, these locations are refined by a secondary network. This approach is slower but more accurate than one-stage detectors. The primary difference between the two is the balance between speed and accuracy. One-stage detectors prioritize speed over accuracy, while two-stage detectors prioritize accuracy over speed. The decision about which approach to use depends on the specific requirements and limitations of the application. This study examined and combination of multiple object detectors to boost the predictions. The models chosen were based on Faster R-CNN [63], YOLOv7 [64], YOLOv8, and EfficientNet [65], selected for their performance and hardware requirements. The subsequent subsections contain a concise explanation of these architectures and their configurations:

FASTER R-CNN architecture takes an input image and applies the convolutional neural network to extract a set of feature maps from the images. The object detector is composed of two main elements: the first is a region proposal network (RPN), while the second is a Faster R-CNN detector. The RPN creates region proposals, locating the objects in the image. Fast R-CNN detector classifies these proposals and improves their performance for prediction. This two-stage approach enables Faster R-CNN to achieve high accuracy while maintaining fast processing speeds. Faster R-CNN is employed in different computer vision applications, such as autonomous driving, pedestrians, and face detection.

YOLO (You Only Look Once) is an object detection target for real-time applications. It identifies and localizes objects

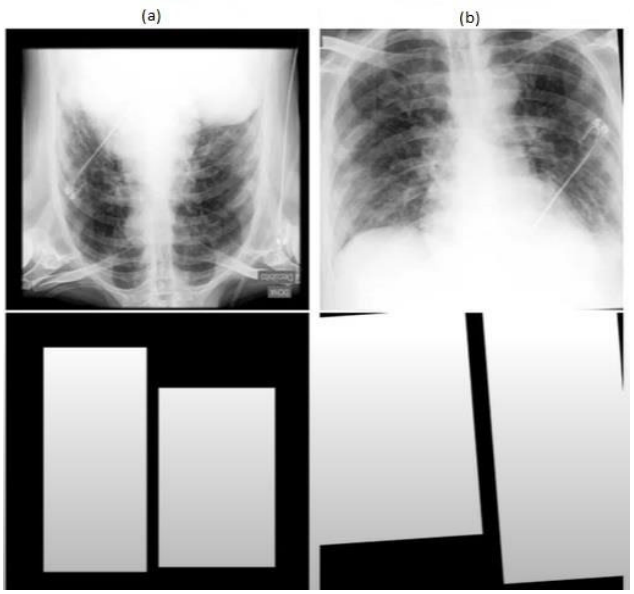


FIGURE 7. Augmentation technique for SIIM-FISABIO-RSNA COVID-19 dataset a) H-Flip & V-FLIP b) Rotation & cropping the images.

within an image or video. YOLO architecture divides an image into multiple cells and makes predictions for each cell, including the bounding boxes, scores, and probabilities for different classes. It is popular for autonomous driving, surveillance, and robotics applications. YOLO has been developed and improved over the years, with the latest version being YOLOv8, which incorporates advancements in deep learning techniques and architecture to improve accuracy and speed.

EFFICIENTNET is a model that aims to achieve superior performance in image classification tasks while utilizing fewer computational resources and parameters than the other models, such as ResNet and Inception models. The design of EfficientNet is based on a unique scaling technique that evenly optimizes the depth and width of the neural network architecture. This approach allows EfficientNet to perform highly in various benchmark datasets, from essential object recognition to more intricate tasks such as fine-grained image classification.

As part of the initial processing, the images were then converted from DICOM to NumPy arrays using the *dicom2array* function and encoded with 8 bits. Augmentation techniques were utilized to establish new training samples from the original data, see Figure 7. The following steps were employed for data pre-processing in the x-ray image SIIM-FISABIO-RSNA dataset:-

- The images have been converted from DICOM to PNG.
- We rotate the images by a maximum of limit 10^2, and this transformation was applied with a probability of 0.20.
- The flipping process has been performed to the images vertically and horizontally with a given probability of 0.25. This is to introduce additional variations to the data that can improve the performance of the proposed models.
- The contrast of the images underwent randomly with a probability of 0.25. This is to the visibility of subtle features, making it easier for the deep learning model to detect and analyze patterns in the image.
- A cropping technique has been performed to eliminate unwanted artifacts, and noise present in the original image.

TABLE 5. The SIIM-FISABIO-RSNA COVID-19 was used in our research.

Category	Train	Test
Atypical	379	95
Typical	2275	569
Indeterminate	836	209
normal	1337	334
Total	4827	1207

The dataset has been divided into 80% for training and 20% for validation. Table 5 shows the summary of the dataset (training and testing) after the split. A number of experiments were performed to ascertain the epochs number requirement for the model to converge, and the models have been trained and tuned with 80 epochs. Stochastic Gradient Descent has been selected for all the proposed models, and momentum has been adjusted at 0.9 to accelerate the neural network. We set the mini-batch at 16. The learning rate was optimized at 0.0001 to control the model error response.

V. EXPERIMENT RESULTS

All the experiments have been conducted on Kaggle with NVIDIA Tesla P100 GPU to train & validate the proposed architectures with their requisites and libraries. In this section, we will explore the obtained results from both localization and multi-classification of the abnormalities on the X-ray images. Furthermore, this section provides a comprehensive analysis of the experimental results from both tasks and their implications for developing accurate and reliable object detection models for abnormalities diagnosis and detection using X-ray images. We used confusion matrix criteria, and ROC curves, which include the metric measurement such as Accuracy (ACC), Precision (P), Recall (R), and F1-score (F1) to evaluate the proposed pre-trained models, see Eq [1,2,3,4]. In this research, we used the mean average precision for evaluating object detection performance on chest X-ray images. Mean average precision considers the impact of varying confidence thresholds on detection performance. Object detectors assign confidence scores to each detected object, indicating the algorithm's confidence in its accuracy. By varying the threshold for considering an object as detected, mAP captures the trade-off between precision and recall, offering insights into the algorithm's performance at different operating points. see Eq [5,6], where k is queries number and AP_i is the average precision (AP) for a given query (i)

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

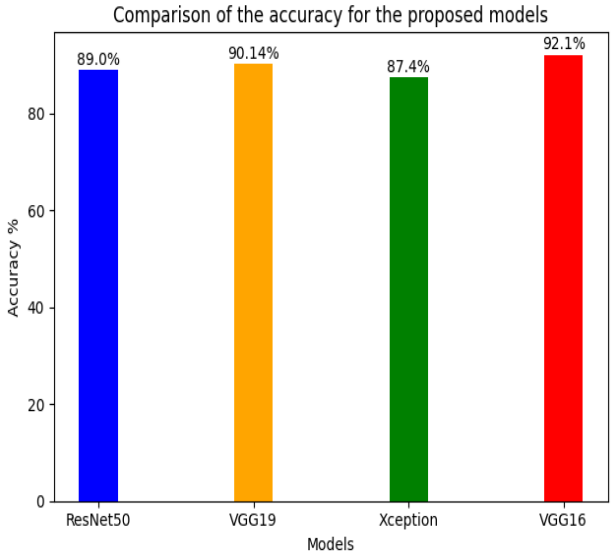


FIGURE 8. Comparison of the accuracies of the proposed models.

$$P = \frac{TP}{TP + FP} \tag{2}$$

$$R = \frac{TP}{TP + FN} \tag{3}$$

$$F1 = 2 \times \frac{(P \times R)}{(P + R)} \tag{4}$$

$$\text{Average Precision (AP)} = \int_{r=0}^{1} p(r)dr \tag{5}$$

$$mAP = \frac{1}{k} \sum_{i}^{k} AP_i \tag{6}$$

A. MULTI-CLASSIFICATION RESULTS OF PRE-TRAINED MODELS

Tables [6-9] summarize the results for multi-classification of the abnormalities on chest X-ray images, including comparing the models based on various confusion metrics. The area under the curve (AUC) is shown in Figure 9. The results showed that the proposed models could classify the four patient statues (Normal, Lung opacity, viral pneumonia, and COVID-19) starting with normal images, with the highest number of images followed by lung opacity, COVID-19, and viral pneumonia. ResNet50 performs reasonably well, with F1 scores ranging from 0.80 to 0.91 across the four classes. The model achieves the highest F1-score for the COVID-19 class (0.91), indicating that it is better at identifying this class than the others. The precision and recall values are also relatively high for all classes, indicating that the model is able to both correctly identify positive cases and avoid false positives. VGG19 also performs reasonably well, with F1 scores ranging from 0.85 to 0.91 across the four classes. As with the ResNet50 model, the COVID-19 class has the highest F1-score (0.91), indicating that the model is best at identifying this class. The precision and recall values for the COVID-19 and viral_pneumonia classes are exceptionally high, indicating that the model is able to identify these positive cases accurately. Xception architecture performs somewhat worse than the other models, with F1 scores ranging from 0.74 to 0.87 across the four classes. The normal class has the lowest F1-score (0.74), indicating that the model struggles to identify this class accurately. The precision and recall values for the normal class are particularly low, indicating that the model is not very good at identifying true negative cases. The VGG16 model performs similarly to the VGG19 model, with F1 scores ranging from 0.86 to 0.92 across the four classes. The COVID-19 class again has the highest F1-score (0.92), indicating that the model is best at identifying this class. The precision and recall values for the COVID-19 and viral_pneumonia classes are once again high, indicating that the model is able to identify these positive cases accurately. All four models perform reasonably well, with F1 scores ranging from 0.74 to 0.92 across the four classes. However, there are some differences in performance between the models, with the VGG19 and VGG16 models appearing to perform slightly better than the ResNet50 and Xception models. The COVID-19 and viral_pneumonia classes appear to be the easiest to identify for all four models, while the normal class is the most difficult for the Xception model. It is noted that VGG16 model achieved the highest accuracy with 92.1% and performed better than the other three models for multi-classification of COVID-19, as evidenced by Figure 8. VGG16 performs better than the other models for X-ray image classification is its simpler architecture. VGG16 has 16 layers, making it easier to train than VGG19 (which has 19 layers), ResNet50, and Xception (which have even more layers). The simpler architecture of VGG16 reduces the risk of overfitting, which is especially important when working with limited data. VGG16 uses 3×3 filters, while ResNet50 and Xception use larger 5×5 filters. Smaller filters allow the network to learn more detailed features, which is crucial when working with chest X-ray images. Furthermore, VGG16 uses max pooling after every two convolution layers, which helps reduce the feature maps' spatial dimensionality. In addition, VGG16 overcomes the state-of-the-art methodologies, as reported in Table 1.

TABLE 6. Classification report of VGG19 model.

Classes	Precision	Recall	F1-score
COVID-19	0.95	0.87	0.91
lung opacity	0.85	0.85	0.85
normal	0.90	0.80	0.85
viral_pneumonia	0.89	0.93	0.91

TABLE 7. Classification report of ResNet50 model.

Classes	Precision	Recall	F1-score
COVID-19	0.90	0.88	0.91
lung opacity	0.80	0.83	0.82
normal	0.84	0.76	0.80
viral_pneumonia	0.88	0.89	0.88

B. LOCALIZATION RESULTS OF THE PROPOSED OBJECT DETECTION MODELS

Each model underwent different experiments using the proposed pre-processing technique, which involves cropping

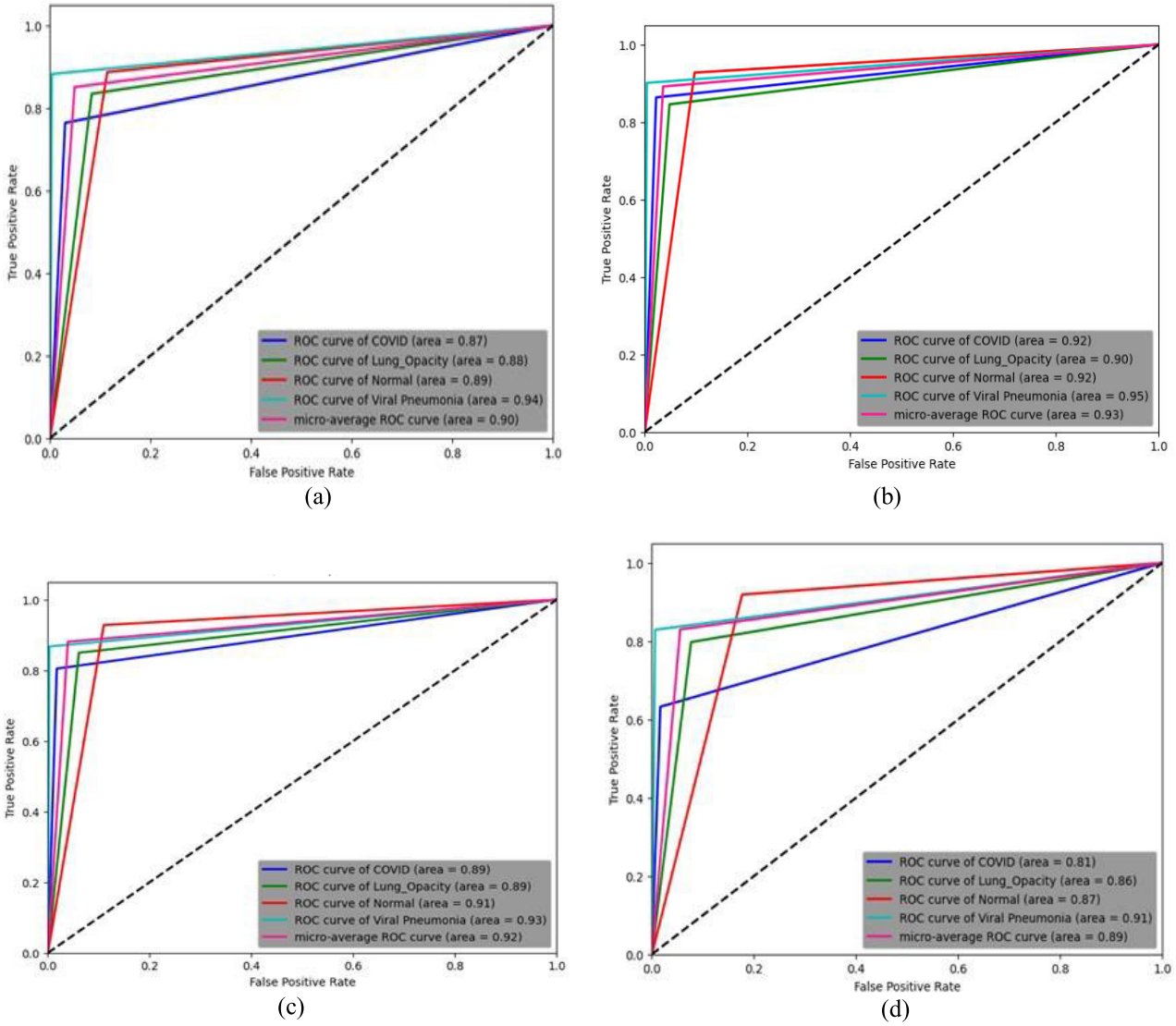


FIGURE 9. ROC curves that include a) ResNet50 and b) VGG16. c) VGG19 d) Xception architectures for multi-classification of abnormalities on X-ray images.

TABLE 8. Classification report of VGG16 model.

Classes	Precision	Recall	F1-score
COVID-19	0.94	0.90	0.92
lung opacity	0.87	0.85	0.86
normal	0.89	0.86	0.88
viral_pneumonia	0.90	0.93	0.91

TABLE 9. Classification report of Xception model.

Classes	Precision	Recall	F1-score
COVID-19	0.90	0.83	0.86
lung opacity	0.80	0.80	0.80
normal	0.89	0.63	0.74
viral_pneumonia	0.83	0.92	0.87

and rotating the images. The results show that the mean average precision cannot exceed 0.49 for Faster R-CNN, 0.47 for YOLOv7, and 0.46 for the YOLOv8x model. Chest X-rays are complex images with overlapping organs and structures and subtle differences in disease presentation. These complexities can make it difficult for the exploited object detection models to identify the presence and location of abnormalities accurately. SIIM-FISABIO-RSNA COVID-19 has a class imbalance, where the number of annotated labels differs from one class to another. This can cause the model to perform poorly in the minority class and achieve lower mean average precision. In order to boost the mean average performance, the weighted box fusion (WBF) technique by the ensemble and fusing the predicted confidence scores and bounding boxes of multiple object detection models. The weighted box fusion approach utilizes the confidence scores of all proposed bounding boxes to establish average boxes. WBF first calculates the Intersection over Union (IoU) between all pairs of bounding boxes. IoU measures the overlap between two bounding boxes as the ratio of the area of their intersection to the area of their union, considering two bounding boxes $b_1 \& b_2$ The formula is presented in Eq [7].

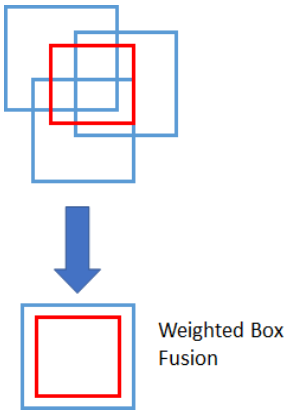


FIGURE 10. Schematic weighted box fusion (WBF) diagram for ensemble multiple bounding boxes.

Then, it assigns a weight to each bounding box based on its confidence score and IoU with other boxes. Finally, it calculates a weighted average of the bounding boxes to obtain the final prediction, as shown in Figure 10.

$$\text{IoU}\,(b_1, b_2) = \frac{\text{area}\,(b_1 \cap b_2)}{\text{area}\,(b_1 \cup b_2)} \tag{7}$$

Assuming you have N object detectors in your ensemble, and each detector produces a set of bounding boxes with their corresponding confidence scores, the weighted box fusion Eq [8] is:

$$B^* = \sum_{i=1}^{N} w_i B_i \tag{8}$$

where B^* is the fused set of bounding boxes, B_i Is the set of bounding boxes produced by detector i. w_i the weight assigned to detector i is based on its performance and reliability. The weight w_i can be computed in several ways, such as using the precision and recall of each detector on a validation set or the mean average precision (mAP) score of each detector on a test set. First, the weights should be normalized so that their sum equals 1. After computing the weights, the final set of bounding boxes B^* is obtained by linearly combining the bounding boxes of each detector, weighted by their corresponding weights. This research has developed an ensemble of multiple object detectors for COVID-19 detection. Each model has been trained and evaluated by using SIIM-FISABIO-RSNA COVID-19 dataset. First, we have combined YOLOv7+EfficientNet, and the mean average precision has been improved to 0.58 mAP@0.5. After that, we developed the proposed system to add Faster R-CNN to combine EfficientNet+YOLOv7+Faster R-CNN as shown in Figure 11, and mean average precision has been increased up to 0.612 mAP@0.5. See Table 10.

Based on the table, we can observe that the ensemble models, EfficientNet+YOLOv7 and EfficientNet+YOLOv7+ Faster R-CNN, are giving better mean average precision (mAP) than the other single object detection models. Ensemble models combine multiple models to improve the overall performance, leveraging the strengths of each model

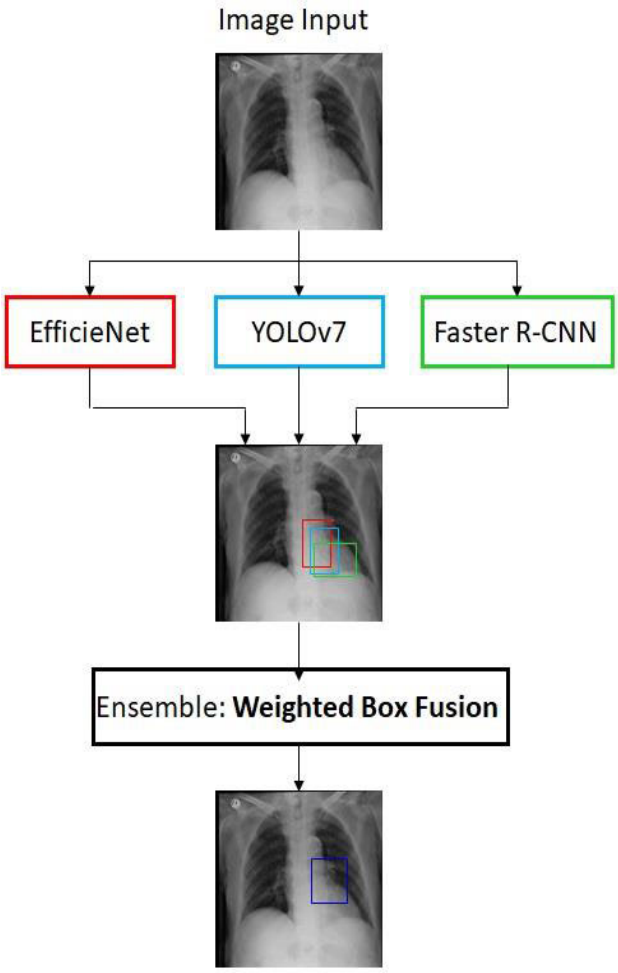


FIGURE 11. The final ensemble models combined with EfficientNet+YOLOv7+Faster R-CNN for localization and object detection on chest X-ray images.

to overcome their weaknesses. In this case, the EfficientNet model, a state-of-the-art convolutional neural network for image classification, is combined with object detection models YOLOv7 and Faster R-CNN to improve their performance in detecting abnormalities. The YOLOv7 and Faster R-CNN models are popular object detection models with different strengths and weaknesses. YOLOv7 is known for its speed and real-time performance, while Faster R-CNN is known for its high accuracy but is slower. By combining these models with EfficientNet, the resulting ensemble models can benefit from both the speed of YOLOv7 and the accuracy of Faster R-CNN. The mAP metric measures the average precision of a model at different levels of recall, which measures how well the model can correctly identify objects of interest. The higher the mAP, the better the model's performance. We can see that the EfficientNet+YOLOv7 ensemble model achieves an mAP of 0.58, which is higher than any of the single object detection models. The EfficientNet+YOLOv7+Faster R-CNN model achieves an even higher mAP of 0.612, indicating that the combination of all three models produces the best

TABLE 10. The performance of object detection and ensemble models for abnormalities detection.

Model	m AP@0.5
YOLOv7	0.47
YOLOv7-e6e	0.4
YOLOv7-w6	0.41
YOLOv7x	0.46
YOLOv8n	0.44
YOLOv8x	0.46
Faster R-CNN	0.49
EfficientNet+YOLOv7	0.58
EfficientNet+YOLOv7+Faster R-CNN	**0.612**

performance. The ensemble models EfficientNet+YOLOv7 and EfficientNet+YOLOv7+Faster R-CNN give better mean average precision than the other single object detection models because they leverage the strengths of multiple models to improve overall performance. By combining the speed of YOLOv7, the accuracy of Faster R-CNN, and the image classification capabilities of EfficientNet, the ensemble models can detect abnormalities more accurately and efficiently.

C. DISCUSSION & COMPARATIVE ANALYSIS

In this work, two different deep learning approaches for multi-classification and object detection of abnormalities on chest X-ray images are proposed. The multi-classification approach is utilized to classify different types of diseases, which include normal, negative, COVID-19, and viral pneumonia. The obtained results showed that VGG16 achieved the highest performance. The proposed approach obtained better accuracy than the work [17], [20], although the authors utilized ResNet50 and VGG16 to classify abnormalities on chest X-ray images. The results of the method [21] showed better accuracy than the proposed model. However, the authors of this approach utilized only 5856 images and conducted only binary classification with positive or negative for the presence of COVID-19 on chest X-ray images. The proposed object detection models for localizing abnormalities in chest X-ray images indicated that the ensemble of these 3 models achieved better results in comparison to other systems in the literature [18], [19], [24], which utilized SIIM-FISABIO-RSNA COVID-19 dataset. as reported in Table 1. This system outperformed all previous implementations. ensemble models improved the results of localization on X-ray images by combining the predictions of multiple models. Individual models have different biases, which can create overfitting to specific patterns in the data. Overfitting occurs when a model learns to perform well on the training data but fails to generalize to new, unseen data. These issues can lead to poor predictive performance and generalization on X-ray images. With the ensemble method, bias, and variance are reduced. The proposed ensemble model works by aggregating the predictions of individual models, typically through voting or averaging. Each model's prediction contributes to the final decision, and the ensemble model's output is determined based on the consensus and weighted average of these predictions.

VI. CONCLUSION & FUTURE WORK

This research aimed to implement deep learning architectures for automatically classifying and localizing abnormalities on chest X-rays, including those due to COVID-19. Two approaches were proposed: a multi-classification model and an object detection model that can identify and localize disease and other pulmonary abnormalities. We trained the proposed models on a large dataset of chest X-ray images for different classes and validated them on an independent test set. The pre-processing technique has been implemented in both tasks, which minimizes the search space, reduces the artifacts, and removes any irrelevant details that could confuse the proposed models, such as patient information and recording procedure data. Our results showed that the ensemble approach outperformed the single-object models in detecting multiple abnormalities in X-ray images. This research encountered limitations and challenges, which is related to the dataset. Medical imaging datasets often involve a class imbalance problem where one class has significantly fewer examples. This class imbalance can cause issues for deep learning models trained on the datasets, leading to biased predictions and reduced performance. Medical imaging datasets are often limited in size due to the high cost of acquiring medical images and the ethical issues surrounding the use of patient data. Overall, the results are systematically presented, making it suitable to compare the various methods and ensure a fair assessment of the developed approaches. Additionally, this approach establishes a framework that can be applied to other architectures, ensembles, and different diseases on X-ray and CT images for the same purposes. In this work, it is observed that ensemble models have certain drawbacks that make them computationally expensive and time-consuming. This is due to the need for training and storing multiple models, as well as combining their outputs. Consequently, this complexity results in increased memory requirements for the system. Further to our exploration, there is an ongoing effort to boost the accuracy of the proposed models by improving the pre-processing techniques, balancing the number of classes of the images, and using other deep learning architectures in both multi-classification and object detection.

ACKNOWLEDGMENT

The authors would like to thank the Re-start Toscana COVID-19 project and the Testarossa EuroHPC project for their support.

REFERENCES

[1] H. N. Monday, J. Li, G. U. Nneji, S. Nahar, M. A. Hossin, J. Jackson, and C. J. Ejiyi, "COVID-19 diagnosis from chest X-ray images using a robust multi-resolution analysis Siamese neural network with super-resolution convolutional neural network," *Diagnostics*, vol. 12, no. 3, p. 741, Mar. 2022.

[2] S. H. Khan, A. Sohail, A. Khan, and Y.-S. Lee, "COVID-19 detection in chest X-ray images using a new channel boosted CNN," *Diagnostics*, vol. 12, no. 2, p. 267, Jan. 2022, doi: 10.3390/diagnostics12020267.

[3] S. H. S. Tali, J. J. LeBlanc, Z. Sadiq, O. D. Oyewunmi, C. Camargo, B. Nikpour, N. Armanfard, S. M. Sagan, and S. Jahanshahi-Anbuhi, "Tools and techniques for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)/COVID-19 detection," *Clin. Microbiol. Rev.*, vol. 34, no. 3, pp.'1–63, Jun. 2021, doi: 10.1128/CMR.00228-20.

[4] O. Filchakova, D. Dossym, A. Ilyas, T. Kuanysheva, A. Abdizhamil, and R. Bukasov, "Review of COVID-19 testing and diagnostic methods," *Talanta*, vol. 244, Jul. 2022, Art. no. 123409, doi: 10.1016/j.talanta.2022.123409.

[5] N. M. Dipu, S. A. Shohan, and K. M. A. Salam, "Deep learning based brain tumor detection and classification," in *Proc. Int. Conf. Intell. Technol. (CONIT)*, Hubli, India, Jun. 2021, pp. 1–6, doi: 10.1109/CONIT51480.2021.9498384.

[6] M. A. Azeem, M. I. Khan, and S. A. Khan, "COVID-19 detection via image classification using deep learning on chest X-ray," in *Proc. Ethics Explainability Responsible Data Science (EE-RDS)*, Johannesburg, South Africa, 2021, pp. 1–4, doi: 10.1109/EE-RDS53766.2021.9708588.

[7] S. Sudha, K. B. Jayanthi, C. Rajasekaran, and T. Sunder, "Segmentation of RoI in medical images using CNN—A comparative study," in *Proc. IEEE Region Conf. (TENCON)*, Kochi, India, Oct. 2019, pp. 767–771, doi: 10.1109/TENCON.2019.8929648.

[8] R. Gu, G. Wang, T. Song, R. Huang, M. Aertsen, J. Deprest, S. Ourselin, T. Vercauteren, and S. Zhang, "CA-Net: Comprehensive attention convolutional neural networks for explainable medical image segmentation," *IEEE Trans. Med. Imag.*, vol. 40, no. 2, pp. 699–711, Feb. 2021, doi: 10.1109/TMI.2020.3035253.

[9] H. C. Reis and V. Turk, "COVID-DSNet: A novel deep convolutional neural network for detection of coronavirus (SARS-CoV-2) cases from CT and chest X-ray images," *Artif. Intell. Med.*, vol. 134, Dec. 2022, Art. no. 102427, doi: 10.1016/j.artmed.2022.102427.

[10] A. Elhanashi, D. Lowe, S. Saponara, and Y. Moshfeghi, "Deep learning techniques to identify and classify COVID-19 abnormalities on chest X-ray images," *Proc. SPIE*, vol. 12102, May 2022, Art. no. 1210204, doi: 10.1117/12.2618762.

[11] R. Kumar, R. Arora, V. Bansal, V. J. Sahayasheela, H. Buckchash, J. Imran, N. Narayanan, G. N. Pandian, and B. Raman, "Classification of COVID-19 from chest X-ray images using deep features and correlation coefficient," *Multimedia Tools Appl.*, vol. 81, no. 19, pp. 27631–27655, Aug. 2022, doi: 10.1007/s11042-022-12500-3.

[12] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network," 2020, *arXiv:2003.13815*.

[13] Y. Xie, F. Zaccagna, L. Rundo, C. Testa, R. Agati, R. Lodi, D. N. Manners, and C. Tonon, "Convolutional neural network techniques for brain tumor classification (from 2015 to 2022): Review, challenges, and future perspectives," *Diagnostics*, vol. 12, no. 8, p. 1850, Jul. 2022, doi: 10.3390/diagnostics12081850.

[14] D. Albashish, "Ensemble of adapted convolutional neural networks (CNN) methods for classifying colon histopathological images," *PeerJ Comput. Sci.*, vol. 8, p. e1031, Jul. 2022, doi: 10.7717/peerj-cs.1031.

[15] M. I. Daoud, A. Al-Ali, R. Alazrai, M. S. Al-Najar, B. A. Alsaify, M. Z. Ali, and S. Alouneh, "An edge-based selection method for improving regions-of-interest localizations obtained using multiple deep learning object-detection models in breast ultrasound images," *Sensors*, vol. 22, no. 18, p. 6721, Sep. 2022, doi: 10.3390/s22186721.

[16] F. Hardalaç, F. Uysal, O. Peker, M. Çiçeklidag, T. Tolunay, N. Tokgöz, U. Kutbay, B. Demirciler, and F. Mert, "Fracture detection in wrist X-ray images using deep learning-based object detection models," *Sensors*, vol. 22, no. 3, p. 1285, Feb. 2022, doi: 10.3390/s22031285.

[17] C. Sitaula and M. B. Hossain, "Attention-based VGG-16 model for COVID-19 chest X-ray image classification," *Appl. Intell.*, vol. 51, pp. 2850–2863, Nov. 2021, doi: 10.1007/s10489-020-02055-x.

[18] V. Tiwari, A. Singhal, and N. Dhankhar, "Detecting COVID-19 opacity in X-ray images using YOLO and RetinaNet ensemble," in *Proc. IEEE Delhi Sect. Conf. (DELCON)*, New Delhi, India, Feb. 2022, pp. 1–5.

[19] J. Marusani, B. G. Sudha, and N. Darapaneni, "Small-scale CNN-N model for COVID-19 anomaly detection and localization from chest X-rays," in *Proc. 1st Int. Conf. Artif. Intell. Trends Pattern Recognit. (ICAITPR)*, Hyderabad, India, Mar. 2022, pp. 1–6.

[20] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, H. Zhao, Y. Jie, R. Wang, H. Zhao, Y. Chong, J. Shen, Y. Zha, and Y. Yang, "Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images," *medRxiv*, vol. 18, no. 6, pp. 2775–2780, Nov./Dec. 2021.

[21] O. Stephen, M. Sain, U. J. Maduh, and D.-U. Jeong, "An efficient deep learning approach to pneumonia classification in healthcare," *J. Healthcare Eng.*, vol. 2019, pp. 1–7, Mar. 2019.

[22] S. Wang, Y. Zha, W. Li, Q. Wu, X. Li, M. Niu, M. Wang, X. Qiu, H. Li, H. Yu, W. Gong, Y. Bai, L. Li, Y. Zhu, L. Wang, and J. Tian, "A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis," *Eur. Respiratory J.*, no. 2, Aug. 2020.

[23] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, and B. Xu, "A deep learning algorithm using CT images to screen for corona virus disease (COVID-19)," *medRxiv*, vol. 31, pp. 6096–6104, 2020.

[24] B. Sangeetha, A. A. Joshuaa, A. Roshini, P. Sujaybharath, and K. Sutharsan, "Identification and localization of COVID-19Abnormalities on chest radiographs using computer vision," in *Proc. Int. Conf. Intell. Syst. Commun., IoT Secur. (ICISCoIS)*, Coimbatore, India, Feb. 2023, pp. 591–595, doi: 10.1109/ICISCoIS56541.2023.10100499.

[25] J. Hou and T. Gao, "Explainable DCNN based chest X-ray image analysis and classification for COVID-19 pneumonia detection," *Sci. Rep.*, vol. 11, no. 1, pp. 1–15, Aug. 2021.

[26] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," 2020, *arXiv:2003.10849*.

[27] M. F. Aslan, K. Sabanci, A. Durdu, and M. F. Unlersen, "COVID-19 diagnosis using state-of-the-art CNN architecture features and Bayesian optimization," *Comput. Biol. Med.*, vol. 142, Mar. 2022, Art. no. 105244, doi: 10.1016/j.compbiomed.2022.105244.

[28] S. Tang, C. Wang, J. Nie, N. Kumar, Y. Zhang, Z. Xiong, and A. Barnawi, "EDL-COVID: Ensemble deep learning for COVID-19 case detection from chest X-ray images," *IEEE Trans. Ind. Informat.*, vol. 17, no. 9, pp. 6539–6549, Sep. 2021.

[29] P. Bhowal, S. Sen, J. H. Yoon, Z. W. Geem, and R. Sarkar, "Choquet integral and coalition game-based ensemble of deep learning models for COVID-19 screening from chest X-ray images," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 12, pp. 4328–4339, Dec. 2021.

[30] O. Attallah, "ECG-BiCoNet: An ECG-based pipeline for COVID-19 diagnosis using bi-layers of deep features integration," *Comput. Biol. Med.*, vol. 142, Mar. 2022, Art. no. 105210, doi: 10.1016/j.compbiomed.2022.105210.

[31] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. J. Soufi, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," *Med. Image Anal.*, vol. 65, Oct. 2020, Art. no. 101794.

[32] H. Chen, Y. Jiang, M. Loew, and H. Ko, "Unsupervised domain adaptation based COVID-19 CT infection segmentation network," *Int. J. Speech Technol.*, vol. 52, no. 6, pp. 6340–6353, Apr. 2022, doi: 10.1007/s10489-021-02691-x.

[33] D. Dong, Z. Tang, S. Wang, H. Hui, L. Gong, Y. Lu, Z. Xue, H. Liao, F. Chen, F. Yang, R. Jin, K. Wang, Z. Liu, J. Wei, W. Mu, H. Zhang, J. Jiang, J. Tian, and H. Li, "The role of imaging in the detection and management of COVID-19: A review," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 16–29, 2021.

[34] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi, and D. Shen, "Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 4–15, 2021.

[35] P. K. Sethy and S. K. Behera, "Detection of coronavirus disease (COVID-19) based on deep features," Tech. Rep., 2020.

[36] K. W. Reis and K. P. Oliveira-Esquerre, "Diagnosis of patients with blood cell count for COVID-19: An explainable artificial intelligence approach," *J. Health Informat.*, vol. 13, no. 2, pp. 49–56, 2021.

[37] M. R. Karim, T. Döhmen, D. Rebholz-Schuhmann, S. Decker, M. Cochez, and O. Beyan, "DeepCOVIDExplainer: Explainable COVID-19 diagnosis based on chest X-ray images," 2020, *arXiv:2004.04582*.

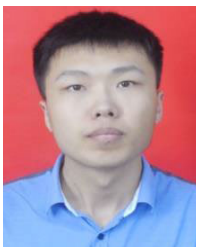
[38] R. Geetha, M. Balasubramanian, and K. R. Devi, "COVIDetection: Deep convolutional neural networks-based automatic detection of COVID-19 with chest X-ray images," *Res. Biomed. Eng.*, vol. 38, no. 3, pp. 955–964, Jun. 2022, doi: 10.1007/s42600-022-00230-2.

[39] S. Basu, S. Mitra, and N. Saha, "Deep learning for screening COVID-19 using chest X-ray images," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2020, pp. 2521–2527, doi: 10.1109/SSCI47803.2020.9308571.

[40] R. Sethi, M. Mehrotra, and D. Sethi, "Deep learning based diagnosis recommendation for COVID-19 using chest X-rays images," in *Proc. 2nd Int. Conf. Inventive Res. Comput. Appl. (ICIRCA)*, Coimbatore, India, Jul. 2020, pp. 1–4, doi: 10.1109/ICIRCA48905.2020.9183278.
[41] A. Elhanashi, D. Lowe, S. Saponara, and Y. Moshfeghi, "Deep learning techniques to identify and classify COVID-19 abnormalities on chest X-ray images," *Proc. SPIE*, vol. 12102, May 2022, Art. no. 1210204, doi: 10.1117/12.2618762.
[42] H. Jiang, S. Tang, W. Liu, and Y. Zhang, "Deep learning for COVID-19 chest CT (computed tomography) image analysis: A lesson from lung cancer," *Comput. Struct. Biotechnol. J.*, vol. 19, pp. 1391–1399, Jan. 2021, doi: 10.1016/j.csbj.2021.02.016.
[43] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *J. Biomolecular Struct. Dyn.*, vol. 39, no. 15, pp. 5682–5689, 2021.
[44] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988.
[45] K. Lin, H. Zhao, J. Lv, C. Li, X. Liu, R. Chen, and R. Zhao, "Face detection and segmentation based on improved mask R-CNN," *Discrete Dyn. Nature Soc.*, vol. 2020, May 2020, Art. no. 9242917.
[46] M. Liu, J. Dong, X. Dong, H. Yu, and L. Qi, "Segmentation of lung nodule in CT images based on mask R-CNN," in *Proc. 9th Int. Conf. Awareness Sci. Technol. (iCAST)*, Sep. 2018, pp. 1–6.
[47] S. Mulay, G. Deepika, S. Jeevakala, K. Ram, and M. Sivaprakasam, "Liver segmentation from multimodal images using HED-mask R-CNN," in *Proc. Int. Workshop Multiscale Multimodal Med. Imag.* Cham, Switzerland: Springer, 2019, pp. 68–75.
[48] J.-H. Shu, F.-D. Nian, M.-H. Yu, and X. Li, "An improved mask R-CNN model for multiorgan segmentation," *Math. Problems Eng.*, vol. 2020, pp. 1–11, Jul. 2020.
[49] D.-H. Nguyen, T.-H. Le, T.-H. Tran, H. Vu, T.-L. Le, and H.-G. Doan, "Hand segmentation under different viewpoints by combination of mask R-CNN with tracking," in *Proc. 5th Asian Conf. Defense Technol. (ACDT)*, 2018, pp. 14–20.
[50] T. Shibata, A. Teramoto, H. Yamada, N. Ohmiya, K. Saito, and H. Fujita, "Automated detection and segmentation of early gastric cancer from endoscopic images using mask R-CNN," *Appl. Sci.*, vol. 10, no. 11, p. 3842, May 2020.
[51] G. Cao, W. Song, and Z. Zhao, "Gastric cancer diagnosis with mask R-CNN," in *Proc. 11th Int. Conf. Intell. Hum.-Mach. Syst. Cybern. (IHMSC)*, vol. 1, Aug. 2019, pp. 60–63.
[52] J.-Y. Chiao, K.-Y. Chen, K. Y.-K. Liao, P.-H. Hsieh, G. Zhang, and T.-C. Huang, "Detection and classification the breast tumors using mask R-CNN on sonograms," *Medicine*, vol. 98, no. 19, May 2019, Art. no. e15200.
[53] D. R. Sarvamangala and R. V. Kulkarni, "Convolutional neural networks in medical image understanding: A survey," *Evol. Intell.*, vol. 15, no. 1, pp. 1–22, Mar. 2022, doi: 10.1007/s12065-020-00540-3.
[54] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[55] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Computer Vision—ECCV 2006*. Berlin, Germany: Springer, 2006, pp. 404–417.
[56] M. E. H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. A. Emadi, M. B. I. Reaz, and M. T. Islam, "Can AI help in screening viral and COVID-19 pneumonia?" *IEEE Access*, vol. 8, pp. 132665–132676, 2020.
[57] T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S. B. A. Kashem, M. T. Islam, S. Al Maadeed, S. M. Zughaier, M. S. Khan, and M. E. H. Chowdhury, "Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images," *Comput. Biol. Med.*, vol. 132, May 2021, Art. no. 104319.
[58] P. Lakhani et al., "The 2021 SIIM-FISABIO-RSNA machine learning COVID-19 challenge: Annotation and standard exam classification of COVID-19 chest radiographs," *J. Digit. Imag.*, vol. 36, no. 1, pp. 365–372, Sep. 2022, doi: 10.1007/s10278-022-00706-8.
[59] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–14.
[60] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
[61] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[62] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," 2016, *arXiv:1610.02357*.
[63] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," 2015, *arXiv:1506.01497*.
[64] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," 2022, *arXiv:2207.02696*.
[65] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and efficient object detection," 2019, *arXiv:1911.09070*.

ABDUSSALAM ELHANASHI (Member, IEEE) received the M.Sc. degree in electronic engineering from the University of Glasgow, Scotland, in 2018. He is currently a Researcher with the University of Pisa, Italy. He has authored and coauthored several scientific articles in international conferences and journals. His research interests include deep learning, imaging processing, medical images, embedded systems, power optimization management, and the IoT devices. He is a member of the Society of Imaging Informatics in Medicine.

SERGIO SAPONARA received the Ph.D. degree in electronics from the I-CAS Laboratory, University of Pisa. He was a Marie Curie Fellow with Imec, Belgium. He is currently a Full Professor of electronics and the Leader of the I-CAS Laboratory, University of Pisa, Italy. He has coauthored 340 scientific articles and 20 patents. He is a TPC Member of more than 100 international IEEE/SPIE conferences. He is an associate editor of several IEEE/Springer journals. In addition, he has been a IEEE Distinguished Lecturer.



QINGHE ZHENG (Member, IEEE) received the B.E. degree from the Xi'an University of Posts and Telecommunications, in 2014, and the M.E. and Ph.D. degrees from Shandong University, in 2018 and 2022, respectively. He is currently an Associate Professor with the Shandong Management University. He has coauthored about 50 peer-reviewed scientific journal articles and holds 11 patents. His research interests include image processing, pattern recognition, deep learning, and edge computing. He is a TPC Member of some international academic conferences, including ICAML 2020, ICBTA 2020, BDMIP 2020, IVPAI 2020, and ICCAES 2021. He is a Guest Editor of several journals, including *Sensors*, *Symmetry*, and *Electronics*.

Open Access funding provided by 'Università di Pisa' within the CRUI CARE Agreement