International Journal of Advanced Technology and Engineering Exploration, Vol 10(104)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2022.10100468

Research Article

X-rays imaging analysis for early diagnosis of thoracic disorders using capsule neural network: a deep learning approach

Himanshu Pant[1*], Manoj Chandra Lohani[2] and Ashutosh Kumar Bhatt[3]
Assistant Professor, School of Computing, Graphic Era Hill University, Bhimtal, Nainital, Uttarakhand, India[1]
Professor, School of Computing, Graphic Era Hill University, Bhimtal, Nainital, Uttarakhand, India[2]
Associate Professor, Department of Computer Science, Uttarakhand Open University, Haldwani, Nainital, Uttarakhand, India[3]

Received: 26-November-2022; Revised: 12-July-2023; Accepted: 15-July-2023
©2023 Himanshu Pant et al. This is an open access article distributed under the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract
The battle against thoracic illnesses has been a major focus in the fields of machine learning and computer vision. Radiology has made significant advancements in the early detection of thoracic disorders. Chest X-rays have become a commonly employed procedure for identifying and categorizing such diseases. However, the shortage of qualified radiologists presents a challenge to accurate diagnosis. Recently, there has been growing interest in the utilization of convolutional neural network (CNN) models for classifying thoracic diseases. Nevertheless, CNNs have limitations in handling translation and rotation in input data, primarily due to the need for extensive training datasets. To address these limitations, capsule networks (CapsNN) have emerged as a novel architecture for automatic learning. CapsNNs are particularly well-suited for managing complex translation and rotation tasks. In this study, a thoracic-cap modelling framework based on CapsNNs was proposed as an alternative approach capable of working with small datasets. Experimental results utilizing a collection of X-ray images demonstrate that the proposed thoracic-caps model outperforms previous CNN-based models, achieving an accuracy of 96.46%. A total of 3043 images of healthy individuals and patients with various thoracic problems were utilized. These findings underscore the effectiveness of CapsNN, a deep fusion neural network, in identifying different thoracic diseases using chest X-ray radiography.

Keywords
Capsule neural network, Deep learning, Thoracic disorders, Chest X-ray, Early diagnosis, Medical imaging.

1.Introduction
The thorax is commonly referred to as the chest. The region is located on the body between the neck and the belly. It is surrounded and protected by the ribcage. Among its essential organs are the heart, lungs, oesophagus, and principal blood vessels. The thoracic region also includes the muscles, bones, ribs, sternum (breastbone), and thoracic vertebrae (spine) [1]. The thorax serves as a protective housing for the organs within it and plays a crucial role in respiration, as it contains the lungs and facilitates the movement of air in and out of the respiratory system. It is an important area for medical diagnosis and treatment as various disorders and conditions can affect the organs and structures within the thorax, including respiratory diseases, cardiovascular conditions, chest wall abnormalities, and tumors [2].

Chest X-ray (CHXR) imaging is a widely used diagnostic tool for evaluating thoracic disorders.

Its popularity stems from its accessibility and its ability to provide valuable insights into the condition of the thoracic region. A little muscle called the diaphragm separates the thorax and abdomen. Numerous conditions that may impede the chest from functioning properly are referred to as thoracic illnesses [3]. These conditions may also impair pulmonary, respiratory, and capacity to breathe functions. Bacterial, viral, or fungal infections link environmental variables to thoracic disorders.

Thoracic disorders pose a significant global health challenge. Early and accurate diagnosis of these disorders is crucial for timely intervention and improved patient outcomes. CHXR imaging is one of the most commonly used diagnostic modalities for thoracic disorders due to its cost-effectiveness and widespread availability. However, interpreting

*Author for correspondence

CHXR images can be complex and challenging, so we may require the expertise of skilled radiologists [4].

In recent times, the field of medical image analysis using the CHXR has witnessed significant advancements through convolutional neural networks (CNNs). These methods have demonstrated remarkable success in this domain. CNN models can automatically learn features from large datasets to extract relevant information from medical images and make accurate predictions [5]. However, CNN models have limitations in handling translation and rotation invariance, which can be critical in thoracic imaging, where the position of lesions or abnormalities can vary.

To address these limitations, capsule neural networks (CapsNN) have emerged as a promising deep learning approach. CapsNN uses "capsules" as basic building blocks instead of traditional neurons. Capsules are groups of neurons that encode not only the presence of features in an image but also their spatial relationship, orientation, and other attributes, making CapsNN inherently capable of handling complex spatial relationships and variations in image data [6]. CapsNN is designed to learn the hierarchical representation of features in images, which can potentially enable it to better capture the complex anatomical structures and abnormalities in CHXR images, leading to improved accuracy in thoracic disorder diagnosis [7].

Accurate interpretation of the CHXR images can be challenging and subjective, as it requires extensive experience and expertise to detect subtle abnormalities. Moreover, the increasing workload on radiologists and the potential for human error may lead to wrong interpretations [8].

Motivated by the potential advantages of CapsNN in medical imaging analysis, this study introduces a deep learning approach that relies on CapsNNs. The proposed method focuses on the early diagnosis of thoracic disorders by analyzing the CHXR images. The goal of this study is to develop a more accurate and efficient method for thoracic disorder diagnosis that can aid radiologists in making timely and informed decisions, ultimately improving patient care and outcomes.

The fast spread of these thoracic ailments has overburdened healthcare infrastructures, necessitating a more effective allocation of medical supplies and equipment. Early detection of these illnesses will hasten the separation of thoracic disorders from non-thoracic disorders, which will aid healthcare authorities in making the most efficient use of resource allocation strategies and early disease prevention [9]. Impaired tissues in the chest are the primary cause of thoracic diseases. It impacts the different physiological systems in a predictable order, including the veins, heart, lungs, mediastinum, oesophagus, chest wall, and diaphragm. Chest pain, hypoxia, paralysis, and a host of other symptoms are all signs of thoracic diseases.

One of the most innovative developments in medical diagnostics in recent decades is medical imaging. Medical imaging assists radiologists in establishing accurate diagnoses by providing a visual representation of the human body's interior. Traditional machine learning methods, such as support vector machines, have been employed for medical image categorization [10]. The performance of these approaches is much below the practical norm, and their recent development has been relatively gradual.

CHXR deformity evaluation is a laborious process for radiologists, and several researchers have created several techniques to assist them in doing it more successfully [11]. For approximately a decade, deep learning models have given rise to sophisticated computerized image applications that are far faster at spotting infections by utilizing image categorization and discovery methods [12].

The main contributions of this paper are as under:

Development of a novel deep learning framework: In the study, a deep learning method was introduced that is intended to evaluate X-ray images in the context of thoracic illnesses. The proposed framework utilizes a CapsNN, which offers a unique architecture and capabilities compared to traditional CNN.

Automation and efficiency: The paper aims to automate the process of thoracic disorder diagnosis using X-ray images by employing the deep learning approach. This automation can significantly reduce the time and effort required for manual interpretation by radiologists. It contributes to making the diagnosis process more efficient, enabling faster results, and potentially improving patient outcomes.

Early detection of thoracic disorders: Early diagnosis of thoracic disorders is crucial for timely intervention and improved treatment outcomes. The paper focuses on leveraging deep learning techniques

to identify potential abnormalities or signs of thoracic disorders in their early stages. By doing so, it contributes to the goal of improving early detection and facilitating prompt medical intervention.

The purpose of this paper is to study and put forward a deep learning-oriented methodology for the early detection and diagnosis of thoracic disorders through X-ray imaging. Specifically, the paper aims to utilize a neural network architecture called the CapsNN to analyze X-ray images and identify potential abnormalities or abnormalities associated with thoracic disorders at an early stage. The paper intends to address the limitations of traditional methods used in thoracic disorder diagnosis, which can be time-consuming and prone to human error. By leveraging deep learning techniques, specifically CapsNN, the authors aim to develop an automated and accurate system that can assist medical professionals in identifying and diagnosing thoracic disorders in their early stages. The block diagram of this approach is described in *Figure 1*.

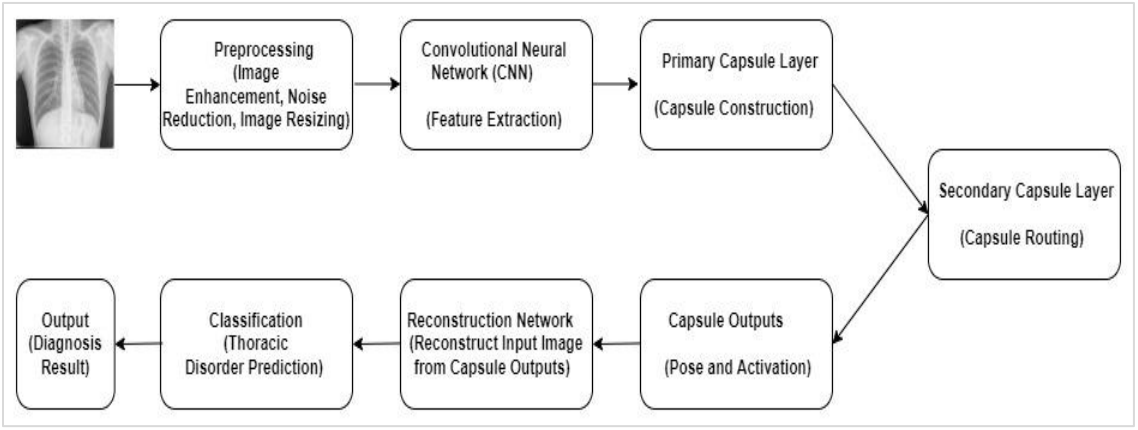

Figure 1 Block diagram of the approach

In Section 1 of this research study, we provide relevant theories and related efforts for the proposed thoracic disease screening model. Section 2 discusses related work and presents a literature review, exploring existing research and studies relevant to the topic. Section 3 focuses on the category of materials and processes, delving into various materials and processes pertinent to the study. Section 4 describes the acquired datasets and the data pre-processing with a result-oriented experimental setup. In Section 5, we present the performance and discussion of the screening model. Finally, the study is concluded in Section 6.

2.Literature review

CNN, which is based on deep learning, is chosen to perform diagnoses of these disorders. It has been used in several projects involving the categorization of restorative photographs. It has always achieved high levels of correctness, so it is expected that this arrangement would function admirably.

In a preceding study, researchers presented an exceptionally effective deep network for rapid tuberculosis (TB) identification and visualization in CHXR images. Their efficient CNN outperformed previous models in terms of speed and accuracy. They explored gradient-weighted class activation maps (Grad-CAMs) and saliency maps as visualization techniques, achieving high precision in TB detection. An unexpected finding was the reliable determination of cardiomegaly location using the heart and its surrounding area. Conventional models rely on the heart-to-lung ratio, which showcases an advantage of the method [13].

The current state of research in thoracic disorder detection using deep learning primarily relies on CNNs. However, CNNs have limitations in capturing spatial relationships and recognizing transformed images. A potential alternative is the capsule network, which comprises related neurons representing specific attributes of an entity, enabling a comprehensive capture of image features and spatial relationships. This network reduces reliance on large datasets, has shown promise as an alternative to CNNs, and also offers advantages in feature extraction. Future research can explore the potential of capsule networks in thoracic disorder detection [14]. A dynamic routing algorithm is proposed that enables capsules to communicate with each other during the forward pass, allowing them to reach a

consensus on the presence of specific features or objects in the input data. This routing mechanism helps to improve the efficiency of information flow in the network and allows capsules to collectively represent higher-level features [15].

The review provides an overview of deep learning methods for thoracic pathology detection and classification in CHXR images. The discussion revolves around the utilization of CNNs, recurrent neural networks (RNNs), and hybrid models trained on extensive datasets to acquire meaningful features. The result of this model shows high accuracy in identifying abnormalities. Deep learning offers advantages in handling complex datasets, achieving accuracy, and detecting multiple pathologies. The limitations of this research include the need for labelled data, potential bias, interpretability challenges, and generalization to rare cases. Standardized protocols, quality assurance, and further research are emphasized to address limitations and ensure reliability in clinical practice [16].

Another study introduces an artificial intelligence-assisted TB detection method using a deep learning normalization-free network model for CHXR images [17]. This model achieves effective TB detection and classification without the need for image normalization. It offers advantages such as accurate detection, the elimination of the normalization step, and the utilization of a large dataset. However, limitations include the reliance on labelled datasets, limited generalization to other thoracic abnormalities, potential interpretability challenges, and the need for further validation on diverse populations and datasets. Addressing these limitations through additional research and validation would enhance the reliability and applicability of the proposed method [18].

The study focuses on deep learning and transfer learning-based models for COVID-19 detection using radiography images. These models harness the capabilities of deep learning algorithms and transfer learning techniques to identify COVID-19 from chest radiographs. The advantages of these approaches include their ability to analyze large datasets, extract relevant features, and achieve high accuracy in COVID-19 detection. Transfer learning allows the models to leverage pre-trained networks and adapt them to the specific task of COVID-19 classification. However, limitations such as the requirement for labelled datasets, potential bias in training data, and challenges in generalizing to unseen or rare cases are present. Further research and validation are crucial to address these limitations and ensure the effectiveness and reliability of these models in real-world clinical practice [19].

The comparison study evaluates deep-learning architectures, including CNNs, RNNs, hybrid models, and transfer learning approaches, for classifying thoracic pathology. The result of this model shows accurate classification with advantages in capturing complex patterns and features. CNNs excel in spatial extraction, RNNs in temporal dependencies, and hybrid models perform well. Limitations include labelled datasets, overfitting, interpretability, and generalization challenges [20].

Another comprehensive review of capsule networks in X-ray Imaging provides an overview of the use of capsule networks in the analysis of X-ray images for the early detection of thoracic disorders. The review discusses the principles and advantages of capsule networks compared to traditional CNNs. It highlights the potential of capsule networks for capturing spatial relationships and hierarchical structures in X-ray images. The paper also discusses the challenges and limitations associated with capsule networks in this context [21].

In medical imaging with deep learning, a CapsNN for the automatic classification of CHXR diseases is proposed [22]. The proposed CapsNN architecture includes multiple convolutional layers followed by capsule layers and uses dynamic routing to facilitate information flow between capsules. The authors evaluate the performance of their proposed CapsNN on a publicly available CHXR dataset and compare it with other state-of-the-art methods. The results demonstrate that the proposed CapsNN achieves competitive performance in terms of accuracy and area under the receiver operating characteristic (AUC-ROC), showing its potential for automatic classification of CHXR diseases [23].

A capsule with CNN (Caps-CNN) for the classification of CHXR images to enable early detection of lung diseases is also proposed [24]. Multiple convolutional layers are followed by capsule layers in the proposed Caps-CNN, which uses dynamic routing to facilitate information flow between capsules. The Caps-CNN is trained on a dataset of CHXR images to learn the representations of different lung diseases. The authors assess the effectiveness of their proposed Caps-CNN by conducting performance evaluations on a publicly

accessible CHXR dataset. They also conduct a comparative analysis with other state-of-the-art methods. The results demonstrate that the proposed Caps-CNN achieves promising outcomes in terms of accuracy, sensitivity, specificity, and F1 score.

The authors highlight the limitations of traditional CNNs in handling translation and rotation invariance, which can be crucial in accurately diagnosing thoracic diseases. They introduce capsule networks as a potential solution to overcome these limitations. Capsule networks utilize a hierarchical structure of capsules that can capture the spatial relationships between objects or features in the input data [25]. The paper presents a framework for training capsule networks for thoracic disease classification, which involves capsule-based convolution and routing mechanisms. The proposed approach is evaluated on a large dataset of CHXR images, and its performance is compared with traditional CNN-based approaches. The results show that the capsule network-based approach outperforms the CNN-based approaches in terms of accuracy, sensitivity, specificity, and F1-score for thoracic disease classification. The authors conclude that capsule networks, with their ability to capture spatial relationships and handle translation and rotation invariance, can be an effective and promising approach for the automatic diagnosis of thoracic diseases from CHXR images.

CapsNN is a type of neural network that can capture hierarchical relationships between features to improve the accuracy of thoracic disease diagnosis. The paper presents a detailed description of the proposed CapsNN-based model, including its architecture and training process. The authors conducted experiments on a dataset of CHXR images to evaluate the performance of their CapsNN model for thoracic disease diagnosis [26]. They compare their CapsNN model with other CNN models commonly used in thoracic imaging analysis. The experimental results show that the CapsNN model outperforms the traditional CNN models in terms of accuracy, sensitivity, specificity, and area under the curve (AUC) for thoracic disease diagnosis. The authors also discuss the interpretability and computational complexity of CapsNN as potential limitations of their approach.

A novel approach for thoracic disease classification from the CHXR images using CapsNN with an attention mechanism is proposed to enhance the discriminative capability of the CapsNN model by selectively focusing on relevant regions in the input images [27]. The proposed CapsNN with attention mechanism achieves improved accuracy, sensitivity, and specificity in the AUC-ROC compared to traditional CNN models. The experimental results demonstrate the potential of the proposed approach for accurate and effective thoracic disease classification from CHXR images, which suggests its potential utility in clinical practice for early diagnosis of thoracic disorders.

An efficient approach for the diagnosis of thoracic disorders from CHXR images using CapsNN is also proposed. In this approach, the author addresses the limitations of traditional CNNs in handling translation and rotation invariance, which can affect the accuracy of thoracic disease diagnosis [28]. They proposed a CapsNN-based model that incorporates capsule layers with dynamic routing by agreement to capture local and global contextual information from the input images. The proposed approach achieves high accuracy in the diagnosis of thoracic disorders while also demonstrating efficiency in terms of computational resources and training time. The experimental results indicate that the CapsNN-based approach outperforms traditional CNNs, highlighting the potential of CapsNN for efficient and accurate diagnosis of thoracic disorders from CHXR images.

A deep learning approach for thoracic disease classification from the CHXR images using CapsNN is also presented. The authors propose a deep CapsNN model with multiple capsule layers to capture complex spatial relationships and hierarchical features in the CHXR images. They also introduce a dynamic routing mechanism to enable effective information flow between capsules. The proposed approach achieves high accuracy in thoracic disease classification, outperforming traditional CNN in terms of accuracy and robustness [29]. The experimental results demonstrate the potential of CapsNN in capturing fine-grained features and improving the accuracy of thoracic disease classification from the CHXR images.

To surpass the performance of radiologists, an algorithm is developed specifically for diagnosing pneumonia from frontal-view CHXR images. The algorithm is trained using the ChestX-ray14 dataset with the 121-layer dense convolutional network (DenseNet) [30]. To account for label dependencies, a technique based on the potential of training deep convolutional neural networks (DCNNs) for computer-aided diagnosis systems (CADs) is employed. In a separate study, multiple CNN models

were employed to classify chest radiographs into TB-positive and TB-negative groups. The system's performance is evaluated using two publicly available datasets: the Shenzhen CHXR set and the Montgomery County CHXR set. More than 80% accuracy is achieved by the suggested computer-aided demonstration framework for TB screening, which is equivalent to radiologists' performance.

Several pre-processing techniques were used to train the model on CHXR pictures. The classification approach determined whether or not pneumonia was present in the CHXR pictures using CNNs and residual network (ResNet) architecture [31]. They show that their method outperforms the comparable reference standard in both classification and localization assignments and can successfully employ both class data and explanations of limited areas.

In this paper, we provide a system for categorizing different thoracic illnesses using X-ray images. The literature study claims that the CNNs utilized a lot of data, although there isn't much data accessible for medical imaging. Additionally, CNN performance is decreased by not taking into account the spatial orientation of the input picture. So, for the objective of classifying thoracic illnesses, we emphasize the CapsNN.

The proposed architecture, which utilizes a CapsNN for X-ray imaging analysis of thoracic disorders, is designed for better representation of spatial relationships, robustness to image transformations, reconstruction-based regularization, and early diagnosis and intervention of the disorders.

3.Methodology

3.1Dataset

Accurate data collection for research purposes is vital and energizing in a DCNN. Authorized open-source datasets National Institute of Health (NIH CHXR dataset and COVID-19 CHXR Dataset) that are publicly available for the disease detection competition were used to get the thoracic sickness dataset for this research. THE NIH CHXR dataset is a subset of the ChestX-ray14 dataset that includes 14,863 labelled CHXR images from the NIH Clinical Center [32]. It is commonly used for training and evaluating deep-learning models for thoracic disease classification. The dataset includes X-ray images of patients with various thoracic diseases, such as pneumonia, TB, Lung nodules, and other abnormalities [33]. The images are annotated with binary labels indicating the presence or absence of each of the 14 thoracic diseases, allowing for supervised training of deep learning models. The COVID-19 CHXR Dataset is a specialized dataset that focuses on the CHXR images of patients with confirmed or suspected COVID-19 infections. It is designed specifically for research related to COVID-19 diagnosis using X-ray imaging. The dataset typically includes X-ray images of patients' chests, annotated with labels indicating the presence or absence of COVID-19 infection [34]. We categorized the dataset into four groups using the concepts of the capsule network: normal, pneumonia, TB, and COVID-19. The total number of images in training and testing datasets is 2291 and 752 respectively in *Table 1*. The four labels on this package stand for the four distinct disease categories that were used to classify and forecast thoracic diseases. 3043 CHXR images in total were compiled as a dataset for this research.

Table 1 Train and test data set split

S. No.	Classes	Training images	Testing images
1	Pneumonia	996	349
2	Covid	161	56
3	TB	99	39
4	Healthy	1035	306
	Total	**2291**	**752**

The ratio of the training dataset to the testing dataset for the collected photos is 80:20. Every collection of photographs is in the uncompressed portable network graphics (PNG) format and is a series of grey-scale pictures. We prepared a thorough description of the various case features after collecting the data consistently. *Figure 2* represents the sample images of all four disease categories. Both lung textures thickened in *Figure 2(a).* The left lung's lower lobe has consolidated. That means bacterial pneumonia should be determined to be the cause of the patient's symptoms.

The frosted glass-like structure with hazy borders that can be seen at the edge of the pulmonary arteries in *Figure 2(b)* is readily apparent. There is widespread air turbidity and uneven patchiness in lesions. In conclusion, a patient with COVID-19 has to be identified. White arrowheads in *Figure 2(c)* indicate an infection in both lungs, whereas black arrows indicate the development of a cavity. TB is seen on this CHXR. In *Figure 2(d),* there were no parenchymal lesions and clean veins in both lungs [35]. Both lungs exhibited a small hilum, and there were no nodules or bump shadows to be detected.

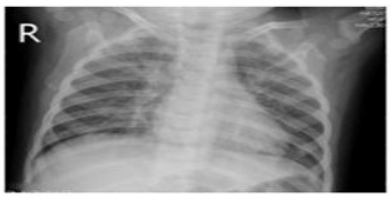

Figure 2(a) Pneumonia

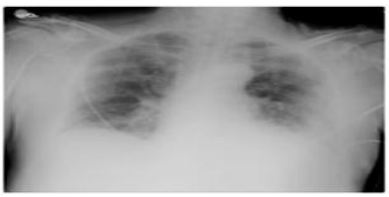

Figure 2(b) Covid

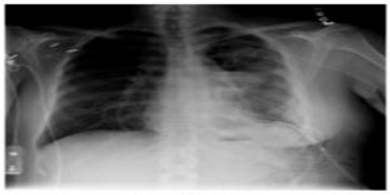

Figure 2(c) Tuberculosis

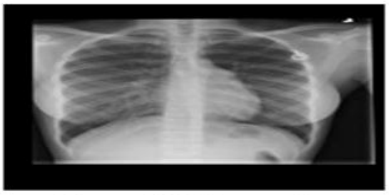

Figure 2(d) Normal

Figure 2 Classified X-ray images

3.2Data preprocessing

When employing various scanning tools and procedures, there will undoubtedly be some variation in terms of challenges related to data resolution, image size, image definition, and others [36]. We suggest some pre-processing steps for the CHXR, which are detailed in *Figure 3*. Pre-processing reduces the impact of data heterogeneity on the effectiveness of networks. The original CHXR image is first converted to a grey scale, and to boost local contrast and gain additional image details, histogram equalization is used. Data augmentation technique will be carried out after the CHXR image processing for image improvement. Data augmentation techniques ought to be as arbitrary as feasible to produce more useful training data [37].

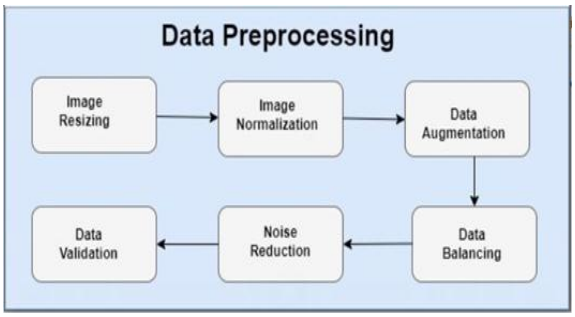

Figure 3 Data preprocessing

The initial step involves acquiring the CHXR report images that are unmarked and unannotated. These images are then associated with four separate directories, each representing a distinct subtype of thoracic diseases [38]. To facilitate categorization the authors labelled and annotated all the images in their respective categories. The directory names employed by the authors include terms such as "pneumonia," "Covid," "TB," and "normal" to indicate the different types of images contained within them. The obtained CHXR images exhibit varying sizes and grayscale intensities. To overcome this challenge, image processing techniques are employed to standardize the size, shape, and dimensions of all the CHXR images. The images are processed using the Tensor Flow API and the Python Keras API, ensuring they are uniform in size, measuring 256 x 256 pixels, and displayed as grayscale images. Additionally, all the X-ray images are annotated to facilitate classification and detection tasks. This dataset contains some noises that are removed by using various steps, as shown in *Figure 4.*

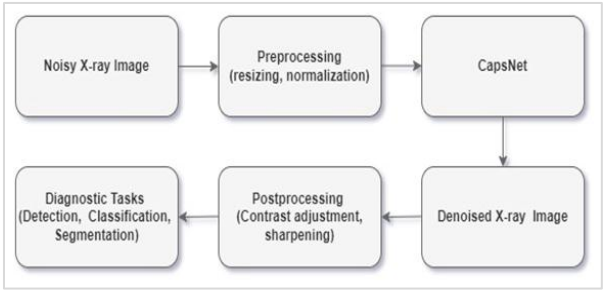

Figure 4 Steps for noise reduction

The input X-ray image may contain various types of noise and artefacts. Preprocessing steps, such as resizing and normalization, are applied to prepare the image for further processing. The CapsNN which consists of multiple capsule layers is the core component of architecture. It extracts and encodes hierarchical visual features from the input image. The CapsNN outputs a de-noised X-ray image, which has undergone noise removal based on the learned

representations. Post-processing steps, such as contrast adjustment or sharpening, can be applied to further refine the de-noised image if necessary. The de-noised X-ray image can be used for various diagnostic tasks, including disease detection, classification, or segmentation.

A histogram is a graphical representation of the distribution of pixel intensity values in an image. In the context of X-ray images, a histogram can provide valuable insights into the overall intensity distribution, which can be useful for image analysis and processing tasks [39]. Python and OpenCV are used to implement the unique implementation method with the help of Google Colab. Histograms are used to compute the pixel intensity values using a histogram function or method provided by the image processing library. The histogram function will return the frequency of occurrence of each intensity value in the image. To create a single-channel picture, grey-scale processing is first applied to the original three-channel CHXR images [40]. After pretreatment, the details on the CHXR image were more evident, the histogram was more balanced, and the lung shape was more prominent. A selection of typical fixed-pixel-sized illustrations of pictures from all four categories is shown in *Figure 5*.

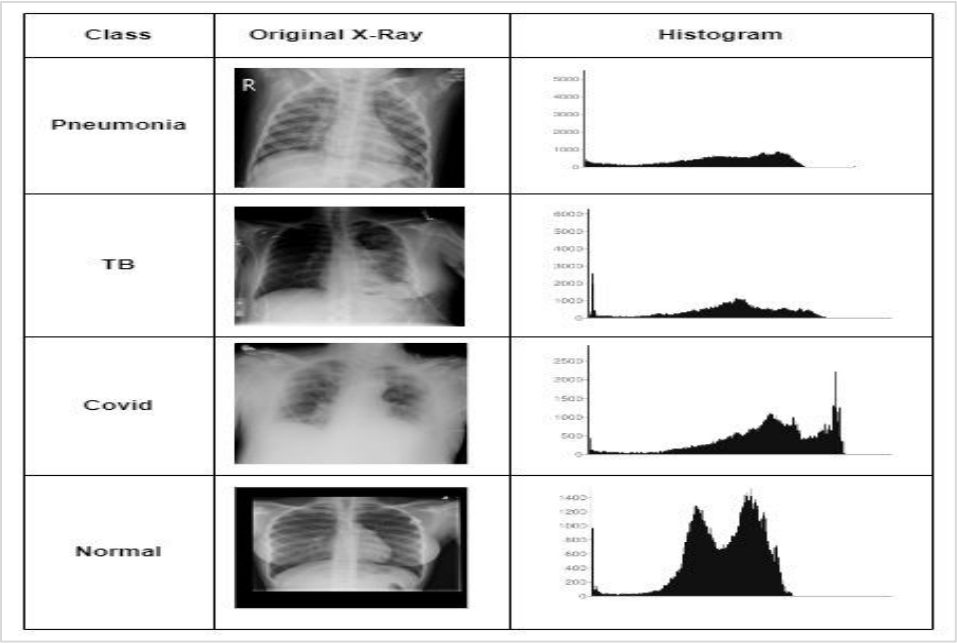

Figure 5 Original CHXR and its histogram

3.3Proposed framework

The novelty of X-ray imaging analysis for early diagnosis of thoracic disorders using a CapsNN in a deep learning approach can be attributed to some potential elements. *Figure 6* depicts the entire framework's suggested structure. First, radiographs are pre-processed to achieve data normalization which addresses the issue of image heterogeneity between various dataset resources [41]. The segmentation network then receives the pre-processed data and extracts the thorax portion of the CHXR images. To determine the type of sickness depicted in the X-ray image, the classification network's DenseNet component collects specific information from the thorax region and provides the features to the capsule network [42]. The proposed architecture for X-ray imaging analysis using a CapsNN is introduced to address the need for early diagnosis of thoracic disorders. The goal is to leverage deep learning techniques to accurately classify thoracic diseases at an early stage, enabling timely intervention and improving patient outcomes. The novelty of the proposed approach lies in the utilization of a CapsNN for thoracic disease classification. The key benefits of using this model architecture for thoracic disease analysis are hierarchical feature learning, better generalization, reduced sensitivity to image variations, and early diagnosis and timely intervention. The justification for using deep learning, specifically CapsNN, for thoracic disease classification is supported by the increasing success of deep learning in medical

imaging analysis. Deep learning techniques have demonstrated remarkable capabilities in image recognition, feature extraction, and classification tasks. CapsNN, with its unique architecture and capabilities, holds promise for improving the accuracy and efficiency of thoracic disease classification compared to traditional CNN-based approaches.

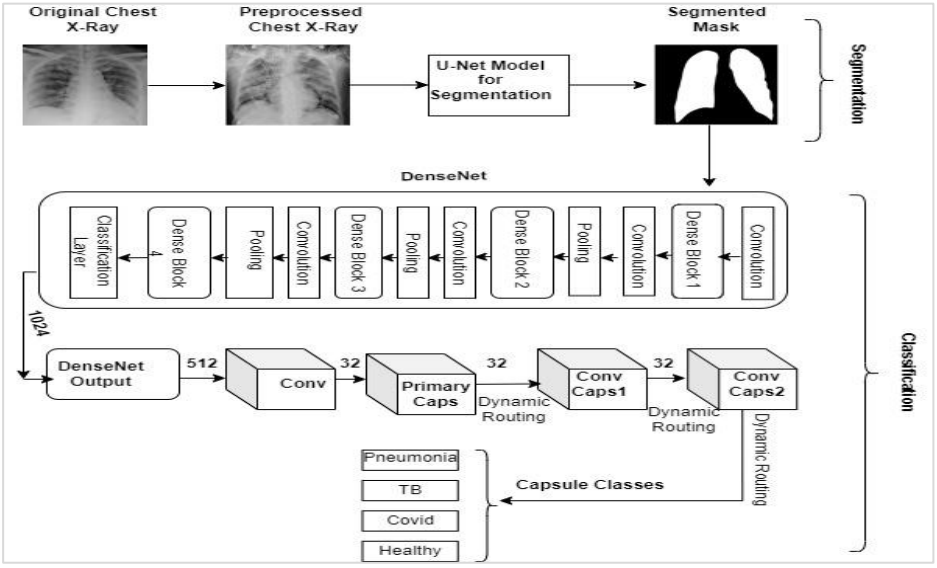

Figure 6 Model architecture

3.4Image segmentation

In this study, the lung contour must be extracted from the CHXR pictures using a segmented network. Finally, for the semantic segmentation component, the recommended technique is the U-Net model. The U-Net model generates an output quickly and precisely using pre-trained ResNet weights. This method helps medical professionals make a rapid diagnosis of several thoracic disorders. Encoder and decoder sets make up the U-Net model [43]. By focusing on the first half of the U-Net model, the encoder condenses the information in the picture into a latent space. The second part of the U-Net model contains a decoder that receives data from latent space and generates a mask-like representation of an image. The typical U-net model is shown in *Figure 7*. The created mask pinpoints the location of thoracic disorders in the CHXR picture. Computer vision and picture pre-processing are essential in medical imaging because the size and quality of the entire image's shot may vary.

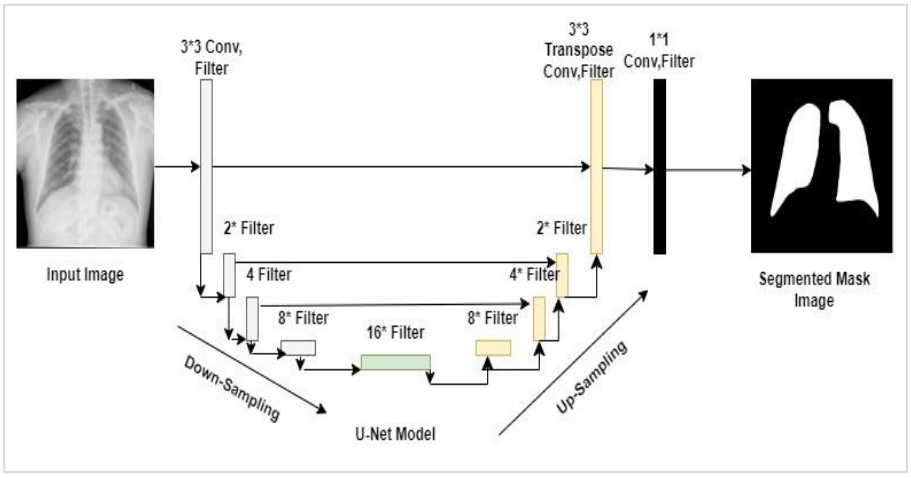

Figure 7 Classical U-net model

3.5Image classification

3.5.1Feature extraction by convolution neural network

Feature extraction is a crucial step in X-ray imaging analysis for early diagnosis of thoracic disorders using CapsNNs. CNNs are commonly used for feature extraction from medical images, including X-ray images, due to their ability to automatically learn relevant features from raw pixel data. In the proposed deep learning approach using a CapsNN, the CNNs can be used as the feature extraction component. The CNNs consist of multiple layers, including convolutional layers, pooling layers, and activation functions, which work together to extract relevant features from the input X-ray images [44].

The convolutional layers in the CNNs apply convolution operations to the input images, which involve applying a set of filters (also known as kernels) to the image pixels to extract local patterns, edges, and textures. This helps reduce the computational complexity and extract relevant features at different scales. The activation functions in the CNNs introduce non-linearity into the model, allowing the CNNs to learn complex patterns and representations from the input images. The extracted features from the CNNs are then fed into CapsNN, which is a type of neural network that is capable of capturing rich spatial relationships between features. Capsules are groups of neurons that represent specific features and their poses (i.e., orientation, size, etc.), and they are trained to learn the probability of the existence of a specific feature in the input image. This allows the CapsNN to capture more holistic and hierarchical features from the input X-ray images, potentially leading to improved accuracy and robustness in the diagnosis of thoracic disorders.

3.5.2Feature extraction by capsule neural network

CapsNN is a network that can extract geographical information as well as other critical attributes, preventing the information loss witnessed during pooling operations. Capsule gives us a vector with a direction as an output [45]. To preserve the locations of items and their attributes inside the image as well as simulate their hierarchical relationships, capsule networks have been created.

The pooling layer in CNN brings important input from the background to the fore. The network may not be able to learn tiny features since data is pooled and sent to the subsequent layer. The neural output of CNN also includes a scalar value. Because capsules hold several neurons, they offer vectorial output with the same size but different routings. The parameters of the images are represented by a vector's routings. Scalar input activation techniques like Tangent, Sigmoid, and rectified linear unit (ReLU) are used by CNNs. On the other hand, squashing, which is defined in Equation 1, is a vectorial activation function used in capsule networks.

$$v_j = \left(\|s_j\|^2 / \left(1 + \|s_j\|^2\right)\right) \times \left(s_j / \|s_j\|\right) \qquad (1)$$

Where, v_j is the output of the j[th] capsule, which is a vector representing the instantiation parameters (e.g., pose, orientation) of the corresponding entity in the input image. s_j is the input to the j[th] capsule, which is the weighted sum of the predictions from the lower-level capsules, usually obtained through a matrix multiplication between the input capsules and the weight matrices. $\|s_j\|$ represents the Euclidean norm (magnitude) of s_j, calculated as the square root of the sum of the squares of its elements.

The weighted sum of the prediction vectors (Uj|i) in the capsules placed in the lower layers, except the first layer of capsule networks, is used to determine the total input value of capsule Sj in Equation 2. The weight matrix (W) and a capsule's output (Oi) from the lowest layer are multiplied to create the prediction vector (Uj|i) (Wij).

$$s_j = W_j \times u_j \qquad (2)$$

Where s_j is the input value to the j[th] capsule. W_j is the weight matrix associated with the j[th] capsule, which determines the strength of the connections between the input capsules and the j[th] capsule. u_j is the output vector of the lower-level capsules that connect to the j[th] capsule, typically obtained from the output of the previous layer or capsule.

A margin loss function is described in Equation 3 and used to train the capsules by encouraging them to produce output vectors that meet certain criteria of magnitude or parameters.

$$L_k = T_k \times max(0, m + \|v_k\| - T_k\|)^2 + \lambda \times (1 - T_k) \times max(0, \|v_k\| - m)^2 \qquad (3)$$

Where, L_k is the margin loss for the k[th] capsule. T_k is the target or reference value for the k[th] capsule, which represents the presence or absence of the predicted entity in the input image. v_k is the output vector of the k[th] capsule, which represents the instantiation parameters (e.g., pose, orientation) of the corresponding entity in the input image. $\|v_k\|$ is the magnitude of the output vector v_k. m is a margin hyperparameter that determines the desired distance between the output vector and the target or reference

vector. λ is a regularization hyperparameter that controls the trade-off between margin-based penalties and regularization. The max () function represents the element-wise maximum function, which ensures that the margin loss is only computed for capsules that produce output vectors outside the desired margin.

4.Results

4.1Performance metrics for classification

The classification phase can recognize the values of true positive (TP), true negative (TN), false negative (FN), and false positive (FP), according to the confusion matrix, Addition of the aforementioned four variables will yield accuracy, precision, recall, specificity, and the F1 score [46].

4.1.1Sensitivity (Recall, true positive rate)

Sensitivity, also known as recall or TP rate, measures the proportion of TP cases (i.e., correctly predicted cases of the thoracic disease) out of all actual positive cases. It reflects the model's ability to accurately identify cases of thoracic disease (Equation 4).

$$\text{Sensitivity} = TP/(TP + FN) \quad (4)$$

4.1.2Specificity (True negative rate)

Specificity, also known as true negative rate, measures the proportion of TN cases (i.e., correctly predicted cases without the thoracic disease) out of all actual negative cases. It reflects the model's ability to accurately identify cases without thoracic disease (Equation 5).

$$\text{Specificity} = TN/(TN + FP) \quad (5)$$

4.1.3Precision (Positive predictive value)

Precision, also known as a positive predictive value, measures the proportion of TP cases out of all positive predictions made by the model. It reflects the model's accuracy in predicting positive cases of thoracic disease (Equation 6).

$$\text{Precision} = TP/(TP + FP) \quad (6)$$

4.1.4Negative predictive value

The negative predictive value measures the proportion of true negative cases out of all negative predictions made by the model. It reflects the model's accuracy in predicting negative cases without thoracic disease (Equation 7).

$$\text{Negative Predictive Value} = TN/(TN + FN) \quad (7)$$

4.1.5F1-Score

The F1-Score is the harmonic mean of precision and sensitivity and provides a balanced measure of both precision and sensitivity. It is a single value that combines both precision and sensitivity, with higher values indicating better performance (Equation 8).

$$\text{F1} - \text{Score} = 2 \times (\text{Precision} \times \text{Sensitivity})/(\text{Precision} + \text{Sensitivity}) \quad (8)$$

To construct the disease detection screening system, an experimental configuration was established utilizing 3×3-sized convolutional filters. These filters are utilized to identify distinctive features and components within an image. Subsequently, a max pooling layer is applied after the convolutional filter. Each convolutional block consists of one convolutional filter and one max pooling layer. The CNN kernel processes the sum of pixels in an input image utilizing the padding approach. Neural network architecture consists of multiple layers, each with a specific number of neurons or filters.

A summary of a sequential neural network model is described in *Table 2*. Each line describes a layer in the network along with its output shape and the number of parameters it contains. Here's a breakdown of the information provided:

- Dense layer 1: This layer has an output shape of (None, 256) and contains 1,048,832 parameters.
- Activation layer 1: Applies an activation function to the output of dense layer 1. The shape remains the same.
- Dropout layer 1: Regularizes the output of Activation layer 1 by randomly dropping out a fraction of elements. It has no parameters.
- Dense layer 2: This layer has an output shape of (None, 128) and contains 32,896 parameters.
- Activation layer 2: Applies an activation function to the output of dense layer 2. The shape remains the same.
- Dropout layer 2: Regularizes the output of Activation layer 2 by randomly dropping out a fraction of elements. It has no parameters.
- Dense layer 3: This layer has an output shape of (None, 4) and contains 516 parameters.
- Activation layer 3: Applies an activation function to the output of dense layer 3. The shape remains the same.

Table 2 Model parameters and layers

Layer(type)	Output shape	Parameters
Dense	(None,256)	1048832
Activation	(None,256)	0
Dropout	(None,256)	0
Dense	(None,128)	32896
Activation	(None,128)	0
Dropout	(None,128)	0
Dense	(None,4)	516
Activation	(None,4)	0
Total Parameters:	**1082244**	
Trainable Parameters	**1082244**	
Non-Trainable Parameters	**0**	

To calculate the number of training steps per epoch utilized in the model, divide the sum of training and testing thoracic illness pictures over all four categories by the given batch size (Equation 9).

$$\text{Total Steps} = \frac{\text{Total Train Xray+Total Test Xray}}{\text{Batch Size}} \quad (9)$$

Ten components of the data set were divided, nine of which were used for training and the tenth for testing. This process was applied to all parts, and the efficiency of the method was assessed by averaging all the parts. Since the data sets include photos from many sources, each image was first scaled down to 128×128 pixels. Because of the great quality of the photos, it takes a powerful machine to analyze the images in their original size using capsule networks. Processing high-quality photos takes time and money, as it does in all traditional deep-learning models. As a result, images have been scaled down to 128×128 pixels [47]. When dealing with more than two classes in X-ray imaging analysis for early diagnosis of thoracic disorders using CapsNN, a multiclass confusion matrix is employed as a table to assess the performance of the classification model. It shows the number of correct and incorrect predictions made by the model for each class, as shown in *Table 3*. Performance is shown as the number of rows and columns in the classification matrix [48]. To achieve optimal performance in terms of accuracy, recall, and F1 score, the classification report for the convolutional model can be derived from the confusion matrix. *Table 4* showcases the classification report, demonstrating the identification and classification of thoracic illnesses through this screening approach.

Table 3 Confusion matrix

		Actual			
		Covid	**TB**	**Normal**	**Pneumonia**
Predicted	**Covid**	169	7	11	13
	TB	2	165	26	7
	Normal	1	34	156	9
	Pneumonia	6	2	3	189

Table 4 Classification report

Thoracic class	Precision	Recall	F1 Score
Covid	0.9458	0.821	0.88
TB	0.787	0813	0.8
Normal	0.797	0774	0.785
Pneumonia	0.878	0892	0.885

With careful tuning of hyperparameters, the three-block sequential model demonstrates exceptional performance. Hyperparameter optimization is essential for maximizing the model's effectiveness. Hyperparameters, which are set before training, play a crucial role in influencing the model's performance. In *Figure 8*, the model achieves a validation accuracy of 96.469% with 50 epochs and 44 steps, along with a training accuracy of 97.29%.

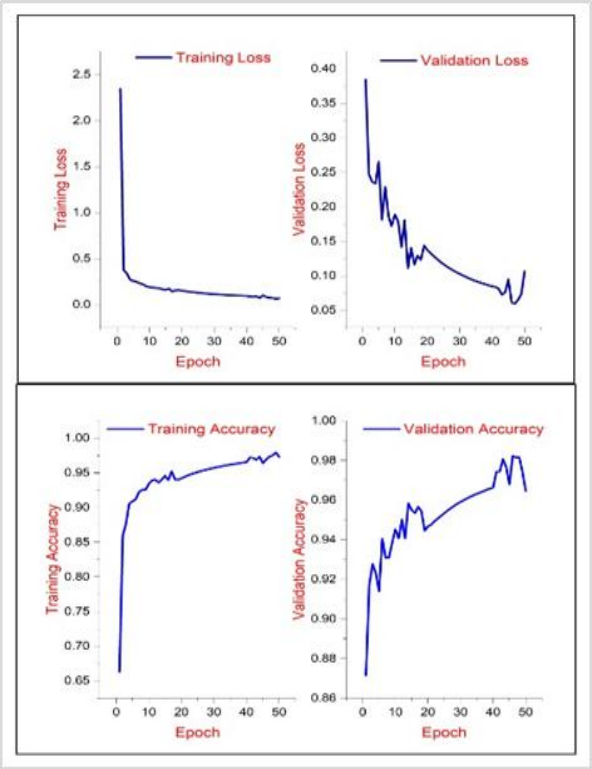

Figure 8 Model performance and cross entropy loss

5.Discussion

A high validation accuracy of 96.469% and a training accuracy of 97.29% have been achieved by the suggested technique for diagnosing thoracic illnesses using X-ray images. These high accuracies show that the model performs well in diagnosing thoracic disorders based on the supplied information. The high accuracy of the model points to the fact that it has discovered important patterns and characteristics from the input images, allowing it to make precise predictions. Evaluation of the model's performance for various disorder classes has also benefited from the use of a multiclass confusion matrix and classification report. With this method, it is possible to fully comprehend the model's advantages and disadvantages when it comes to categorizing various thoracic disorders. Healthcare workers might spot regions where the model could have trouble correctly predicting certain groups of diseases by looking at the confusion matrix. This data can be useful for enhancing the model's performance in difficult circumstances or for creating tailored solutions. Due

to computational constraints, the study used a special dataset that included images that were scaled down. The model is generalizable to additional datasets and full-resolution pictures. It is advised to validate the suggested strategy on a variety of bigger datasets to evaluate how well it performs in various demographics and circumstances. Additionally, the model's sensitivity, specificity, and resilience to changes in data quality and imaging methods should be taken into account. Hyper parameter adjustment was important in further optimizing the suggested approach. The performance of the model is heavily influenced by hyper parameters, which are chosen before training. To maximize the model's performance, careful parameter optimization is advised. To identify the ideal configuration that produces the best results, this procedure comprises a methodical examination and change of hyper parameters, including learning rate, batch size, and network design.

Limitations

The study on X-ray imaging analysis for early diagnosis of thoracic disorders using a CapsNN approach has several limitations that can impact the generalizability, reliability, and applicability of the findings. Some considerable limitations of the model include dataset biases, limited sample size, data imbalance, variability in imaging quality, and limited clinical validation.

A complete list of abbreviations is shown in *Appendix I.*

6.Conclusion and future work

In this study, a neural network-based model combining CNN and CapsNN was developed to detect thoracic diseases. The optimized CapsNN model outperformed other scenarios, achieving an accuracy rate of 96.58% and an F1-score of 97.08%. Future work involves the incorporation of explainable artificial intelligence (XAI) to improve outcome comprehension. The exploration of X-Ray and lung sounds as markers for thoracic ailments is planned. The integration of comprehensive X-ray datasets and established image classification models like VGG19 and AlexNet will enhance validation accuracy. Leveraging CHXR images demonstrates the potential for predicting and detecting various thoracic disorders. Additionally, there is an envisioned development of a user-friendly Android-based mobile application for medical illness detection and self-diagnosis.

Acknowledgment

None.

Conflicts of interest

The authors have no conflicts of interest to declare.

Author's contribution statement

Himanshu Pant: Data collection, data preprocessing, drafting, review and editing. **Manoj Chandra Lohani:** Analysis and interpretation of results. **Ashutosh Kumar Bhatt:** Model design, supervision, investigation on challenges.

References

[1] Pant H, Lohani MC, Bhatt AK, Pant J, Sharma RK. Thoracic disease detection using deep learning. In 5th international conference on computing methodologies and communication 2021 (pp. 1197-203). IEEE.

[2] Palani S, Kulkarni A, Kochara A, Kiruthika M. Detection of thoracic diseases using deep learning. In ITM web of conferences 2020 (pp. 1-7). EDP Sciences.

[3] Fating S, Kotambkar DM. Characterization of common thoracic diseases from chest x-ray images using CNN. In advanced machine intelligence and signal processing 2022(pp. 665-77). Singapore: Springer Nature Singapore.

[4] Heidarian S, Afshar P, Enshaei N, Naderkhani F, Rafiee MJ, Babaki FF, et al. Covid-fact: a fully-automated capsule network-based framework for identification of covid-19 cases from chest CT scans. Frontiers in Artificial Intelligence. 2021; 4:1-13.

[5] Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. Evolutionary Intelligence. 2022; 15(1):1-22.

[6] Sabour S, Frosst N, Hinton GE. Dynamic routing between capsules. In conference on neural information processing systems (pp. 1-12). 2017.

[7] Yamashita R, Nishio M, Do RK, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights into Imaging. 2018; 9:611-29.

[8] Yadav SS, Jadhav SM. Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big Data. 2019; 6(1):1-8.

[9] Li Q, Cai W, Wang X, Zhou Y, Feng DD, Chen M. Medical image classification with convolutional neural network. In 13th international conference on control automation robotics & vision 2014 (pp. 844-8). IEEE.

[10] Jiménez-sánchez A, Albarqouni S, Mateus D. Capsule networks against medical imaging data challenges. In intravascular imaging and computer assisted stenting and large-scale annotation of biomedical data and expert label synthesis: 7th joint international workshop, CVII-STENT and third international workshop, LABELS, held in conjunction with MICCAI, Granada, Spain 2018 (pp. 150-60). Springer International Publishing.

[11] Wang D, Liu Q. An optimization view on dynamic routing between capsules. Workshop track - ICLR 2018 (pp. 1-4).
[12] Yan F, Huang X, Yao Y, Lu M, Li M. Combining LSTM and dense net for automatic annotation and classification of chest x-ray images. IEEE Access. 2019; 7:74181-9.
[13] Hertel R, Benlamri R. Deep learning techniques for COVID-19 diagnosis and prognosis based on radiological imaging. ACM Computing Surveys. 2023; 55(12):1-39.
[14] Cardoso MJ, Arbel T, Lee SL, Cheplygina V, Balocco S, Mateus D, et al. Intravascular imaging and computer assisted stenting, and large-scale annotation of biomedical data and expert label synthesis. In CVII-STENT and second international workshop, LABELS 2017. Springer.
[15] Gazda M, Plavka J, Gazda J, Drotar P. Self-supervised deep convolutional neural network for chest X-ray classification. IEEE Access. 2021; 9:151972-82.
[16] Bai HX, Hsieh B, Xiong Z, Halsey K, Choi JW, Tran TM, et al. Performance of radiologists in differentiating COVID-19 from non-COVID-19 viral pneumonia at chest CT. Radiology. 2020; 296(2):E46-54.
[17] Nijiati M, Ma J, Hu C, Tuersun A, Abulizi A, Kelimu A, et al. Artificial intelligence assisting the early detection of active pulmonary tuberculosis from chest X-rays: a population-based study. Frontiers in Molecular Biosciences. 2022; 9:874475.
[18] Bhattacharyya A, Bhaik D, Kumar S, Thakur P, Sharma R, Pachori RB. A deep learning based approach for automatic detection of COVID-19 cases using chest X-ray images. Biomedical Signal Processing and Control. 2022; 71:1-13.
[19] Hu S, Gao Y, Niu Z, Jiang Y, Li L, Xiao X, et al. Weakly supervised deep learning for covid-19 infection detection and classification from CT images. IEEE Access. 2020; 8:118869-83.
[20] Chouhan V, Singh SK, Khamparia A, Gupta D, Tiwari P, Moreira C, et al. A novel transfer learning based approach for pneumonia detection in chest X-ray images. Applied Sciences. 2020; 10(2):1-17.
[21] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet. 2020; 395(10223):497-506.
[22] Rajaraman S, Antani SK. Modality-specific deep learning model ensembles toward improving TB detection in chest radiographs. IEEE Access. 2020; 8:27318-26.
[23] Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Medicine. 2018; 15(11):e1002686.
[24] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems. 2012; 25:84-90.
[25] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In proceedings of the IEEE conference on computer vision and pattern recognition 2017 (pp. 2097-106). IEEE.
[26] Qin C, Yao D, Shi Y, Song Z. Computer-aided detection in chest radiography based on artificial intelligence: a survey. Biomedical Engineering Online. 2018; 17(1):1-23.
[27] Parveen S, Khan KB. Detection and classification of pneumonia in chest X-ray images by supervised learning. In 23rd international multitopic conference 2020 (pp. 1-5). IEEE.
[28] Mittal A, Kumar D, Mittal M, Saba T, Abunadi I, Rehman A, et al. Detecting pneumonia using convolutions and dynamic capsule routing for chest X-ray images. Sensors. 2020; 20(4):1-30.
[29] Huang G, Liu Z, Van DML, Weinberger KQ. Densely connected convolutional networks. In proceedings of the IEEE conference on computer vision and pattern recognition 2017 (pp. 4700-8). IEEE.
[30] Abedalla A, Abdullah M, Al-ayyoub M, Benkhelifa E. Chest X-ray pneumothorax segmentation using U-Net with EfficientNet and ResNet architectures. Peer Journal Computer Science. 2021; 7:1-36.
[31] Heidari M, Mirniaharikandehei S, Khuzani AZ, Danala G, Qiu Y, Zheng B. Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms. International Journal of Medical Informatics. 2020; 144:1-9.
[32] Afshar P, Heidarian S, Naderkhani F, Oikonomou A, Plataniotis KN, Mohammadi A. Covid-caps: a capsule network-based framework for identification of covid-19 cases from x-ray images. Pattern Recognition Letters. 2020; 138:638-43.
[33] Teto JK, Xie Y. Automatically Identifying of animals in the wilderness: comparative studies between CNN and C-Capsule network. In proceedings of the 3rd international conference on compute and data analysis 2019 (pp. 128-33). ACM.
[34] Rohilla A, Hooda R, Mittal A. TB detection in chest radiograph using deep learning architecture. International conference on emerging trends in engineering, technology, science and management (pp. 136-47). 2017.
[35] Nardelli P, Jimenez-Carretero D, Bermejo-Pelaez D, Washko GR, Rahaghi FN, Ledesma-Carbayo MJ, et al. Pulmonary artery–vein classification in CT images using deep learning. IEEE Transactions on Medical Imaging. 2018; 37(11):2428-40.
[36] Chen HJ, Ruan SJ, Huang SW, Peng YT. Lung x-ray segmentation using deep convolutional neural networks on contrast-enhanced binarized images. Mathematics. 2020; 8(4):1-12.
[37] Berman M, Blaschko MB. Optimization of the Jaccard index for image segmentation with the Lovász hinge. CoRR. 2017: 1-9.

[38] Li Z, Wang C, Han M, Xue Y, Wei W, Li LJ, et al. Thoracic disease identification and localization with limited supervision. In proceedings of the IEEE conference on computer vision and pattern recognition 2018 (pp. 8290-9). IEEE.
[39] Rios A, Kavuluru R. Convolutional neural networks for biomedical text classification: application in indexing biomedical articles. In proceedings of the 6th ACM conference on bioinformatics, computational biology and health informatics 2015 (pp. 258-67). ACM.
[40] Baltruschat IM, Nickisch H, Grass M, Knopp T, Saalbach A. Comparison of deep learning approaches for multi-label chest X-ray classification. Scientific Reports. 2019; 9(1):1-10.
[41] Alapat DJ, Menon MV, Ashok S. A review on detection of pneumonia in chest X-ray images using neural networks. Journal of Biomedical Physics & Engineering. 2022; 12(6):551- 8.
[42] Ananthakrishnan B, Shaik A, Kumar S, Narendran SO, Mattu K, Kavitha MS. Automated detection and classification of oral squamous cell carcinoma using deep neural networks. Diagnostics. 2023; 13(5):1-19.
[43] Yang J, Shi R, Wei D, Liu Z, Zhao L, Ke B, et al. MedMNIST v2-A large-scale lightweight benchmark for 2D and 3D biomedical image classification. Scientific Data. 2023; 10(1):1-10.
[44] https://www.sciencedaily.com/releases/2010/10/101022123749.htm. Accessed 22 April 2022.
[45] Nampally A, Koyyada S, Vaishnavi S, Shaik M. Diagnosis of COVID-19 cases on X-Ray images using CNN. International Journal of Science and Research Archive. 2022; 8(1):31-7.
[46] Ma Y, Lv W. Identification of pneumonia in chest X-ray image based on transformer. International Journal of Antennas and Propagation. 2022; 2022:1-8.
[47] Ar B, Rs VK, Ss K. LCD-capsule network for the detection and classification of lung cancer on computed tomography images. Multimedia Tools and Applications. 2023: 1-20.
[48] Abouel-magd LM, Darwish A, Snasel V, Hassanien AE. A pre-trained convolutional neural network with optimized capsule networks for chest X-rays COVID-19 diagnosis. Cluster Computing. 2023; 26(2):1389-403.

Himanshu Pant is currently pursuing a Doctorate from Graphic Era Hill University. He obtained his MCA from UPTU Lucknow. His areas of interest include Machine Learning, Soft Computing, and Data Mining. With more than 15 publications in reputed National and International SCOPUS, UGC Approved Journals, and conferences, he has made significant contributions to his field. Presently, he is working as an Assistant Professor at Graphic Era Hill University Bhimtal, with approximately 11 years of teaching experience.
Email: himpant7@gmail.com

Dr. Manoj Chandra Lohani earned his Doctorate in Computer Science from Kumaun University Nainital in 2009. With over 20 years of academic teaching experience and more than 20 publications in SCI and SCOPUS journals and conferences, he has made significant contributions to the field. His research area encompasses Network Security, Machine Learning, Soft Computing, and Data Science. Presently, he serves as the Director and Professor at Graphic Era Hill University Bhimtal, Uttarakhand. Additionally, he is a member of Computer Society of India (CSI).
Email: getmlohani@gmail.com

Dr Ashutosh Kumar Bhatt earned his Doctorate in Computer Science from Kumaun University Nainital in 2009. With over 19 years of academic teaching experience and more than 40 publications in SCI and SCOPUS journals and conferences, he has made significant contributions to the field. His research area encompasses Deep Learning, Machine Learning, Soft Computing, and Data Science. Currently, he holds the position of Associate Professor at Uttarakhand Open University, Haldwani, Uttarakhand. He is actively associated with numerous renowned National/International Journals as Lead Guest Editor, reviewer, and editorial board member. Additionally, he is a life member of CSI (Computer Society of India) and a senior member of IEEE. Presently, he is also serving as the District Coordinator of Uttarakhand Science Education & Research Centre (USERC). Furthermore, he had served as State Student Coordinator of Region I Uttarakhand of CSI for three years and also held the position of Secretary of Uttarakhand ACM Professional Chapter.
Email: ashutoshbhatt123@gmail.com

Appendix I

S. No.	Abbreviation	Description
1	AUC	Area Under the Curve
2	AUC-ROC	Area Under the Receiver Operating Characteristic
3	CADS	Computer-Aided Diagnosis Systems
4	Caps-CNN	Capsule with Convolutional Neural Network
5	CapsNN	Capsule Neural Network
6	CHXR	Chest X-Ray
7	CNN	Convolutional Neural Network
8	DCNN	Deep Convolutional Neural Network
9	DenseNet	Dense Convolutional Network
10	FN	False Negative
11	FP	False Positive
12	Grad-CAM	Gradient Weighted Class Activation Mapping
13	NIH	National Institute of Health
14	PNG	Portable Network Graphics
15	ResNet	Residual Network
16	ReLU	Rectified Linear Unit
17	RNN	Recurrent Neural Network
18	TB	Tuberculosis
19	TN	True Negative
20	TP	True Positive
21	XAI	Explainable Artificial Intelligence