Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 28 No. 2 (2025)

Deep Learning Methods for Detecting Chest X-Ray and CT Scan Abnormalities

M.Sangeetha[1], Shruti Sharma[2*], Jyoti Kanjalkar[3], Dr K Sreeramamurthy[4], D. Mangaiyarkarasi[5]

[1]Lecturer, Computer Engineering Department, University of Technology and Applied Sciences, Nizwa, Sultanate of Oman, sangeetha.mani@utas.edu.om

[2]Teaching Assistant, Dept. of AI & ML, COER University, Roorkee, Uttarakhand, cuteshruti323@gmail.com

[3]Assistant Professor, Vishwakarma Institute of Technology, Pune, jyoti.kanjalkar@vit.edu

[4]Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Bowrampet, Hyderabad-500043, Telangana, India, sreeram1203@gmail.com

[5]Assistant Professor, Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Avadi, Tamilnadu, India, dmangaibe@gmail.com

Article History:

Received: *22-07-2024*

Revised: *04-09-2024*

Accepted: *13-09-2024*

Abstract

Deep learning's quick development has created new opportunities to improve medical image analysis, especially in the identification of anomalies in chest CT and X-ray scans. This work investigates several deep learning techniques designed for this particular purpose to enhance the efficiency and accuracy of diagnosis in medical settings. We explore the use of 3D CNNs, transfer learning, and convolutional neural networks, or CNNs, for the analysis of volumetric CT scan information as well as 2D chest X-ray pictures. Comparative analyses show the benefits and drawbacks of various deep learning architectures for identifying a variety of anomalies, including tumors, tumors in the lungs, pneumonia, and other diseases. We also go over the significance of preprocessing methods, assessment metrics specifically designed for medical picture analysis, and dataset preparation. The results highlight how deep learning has the potential to revolutionize chest imaging diagnostics by facilitating the quicker and more accurate identification of anomalies, which will enhance patient outcomes and the effectiveness of healthcare delivery. To spur additional developments in the deep learning-powered analysis of medical images for chest problems, future research topics, and obstacles in this area are also covered.

Keywords: chest imaging diagnostics, convolutional neural networks (CNNs), transfer learning, 3D CNNs, detection of abnormalities, dataset, pre-processing.

1. Introduction

Several disorders can be diagnosed and treated with the aid of medical imaging, with CT scans and chest X-rays being two of the most commonly utilized modalities. However, radiologists may find it difficult and time-consuming to interpret these images, which could result in errors and interruptions in patient care. Multiple approaches to reducing these problems would be taken by a computer-assisted triage system. For starters, radiologists might then promptly concentrate their efforts on cases that pose a greater risk. More information would be available to radiologists, enabling them to

correct any potential misdiagnoses. [1] In summary, the doctor who treats the patient would have instant access to information regarding the patient's condition and risk level, enabling them to promptly order additional diagnostic tests and ask the interpreting radiologist pertinent and well-informed questions. The Tiff CXR images with a normal/abnormal designation will be the input for our algorithm. Next, a CNN will be used to classify each image as normal or abnormal. The issue will be purely a binary classification because we lack knowledge about the levels of abnormalities. Massive variations in CXR pictures and a dearth of labeled data have impeded the creation of such systems thus far. With a large enough dataset, a technique known as deep learning, which has experienced significant growth in recent years, can classify extremely heterogeneous photos. Prioritizing turnaround time can lead to subpar reports, misunderstandings, incorrect diagnoses, and communication breakdowns with primary care physicians, particularly in remote places where healthcare professionals rely on telecommuting for their interpretation of CXRs. They all have a detrimental effect on patient care and have the potential to change patients' lives.

Pneumonia is a periodic viral lung disease that, if not identified and treated promptly, can have potentially fatal consequences for young children (under five years old) and the elderly (over 60 years old). Chest X-rays, CT scans, and MRIs are among the imaging modalities used in clinics to diagnose pneumonia. Although chest X-ray radiography is the most economical diagnostic method for detecting pneumonia, diagnosing the illness from these pictures requires highly qualified radiologists since they frequently overlap with other abnormal lung illnesses. The time-consuming and frequently subjective nature of manual pneumonia detection might cause delays in diagnosis and treatment. Furthermore, X-ray pictures might not show how severe the pneumonia infection is. Utilizing machine learning (ML) and deep learning (DL) to identify pneumonia automatically, computer-aided diagnosis (CAD) practice solves these issues. Due to its broad relevance to issues involving automatic extraction of features and categorization, deep learning has been the subject of extensive research in the recent past. In object recognition and picture classification, neural network convolution (CNN)-based evaluations are commonly employed. CNNs use spatial filters that collect structural information from images automatically. As there is no requirement for image pre-processing subroutines, CNNs are pixel-based and operate directly on images, in contrast to traditional image classification techniques used in machine learning.

Cancer is the second leading cause of mortality worldwide, according to the World Health Organisation (WHO), with 9.5 million deaths expected from the disease in 2017.There are 2.08 million cases of cancer in the lungs alone. Additionally, according to the WHO, early detection and well-coordinated therapy can lower the death rate from cancer. [2] An operation radiation therapy, and chemotherapeutic are examples of treatment methods that can be used to lower the possibility of death if the tumor's benign or malignant categorization is correctly identified. Many imaging modalities can be used to identify lung cancer, but Computed Tomography (CT) represents one of the most affordable and useful modalities used in clinics to assess lung health. It takes a while to manually examine lung abnormalities based on CT pictures, so a pulmonologist with experience is needed. To aid in decision-making and treatment planning, if a lung abnormality is first evaluated using a computer-assisted method, the pulmonologist can receive the analytic report. This paper's main goal is to investigate the use of deep learning techniques for finding abnormalities in chest CT and X-ray images.

Convolutional Neural Networks (CNNs):

The increased performance of CNNs in image categorization has made them popular. Finding the temporal and spatial features in a picture is made easier by the network's convolutional layers working in tandem with filters. Another method to lessen computing requirements is the weight-sharing strategy among the layers. Simply put, CNNs are feedback artificial neural networks, or ANNs, with two limitations: to maintain spatial structure, neurons within the same filters are simply linked to small patches of the picture, and the weights they use are shared to minimise the overall number of model parameters. Three fundamental components make up a CNN: (i) a convolution layer, which learns features; (ii) a maximum pooling (subsampling) layer, which reduces the dimensionality of the image and hence the computational effort; and (iii) a layer that is fully connected, which gives the network classification capabilities. Figure 1.1 shows an overview of CNN's architecture.

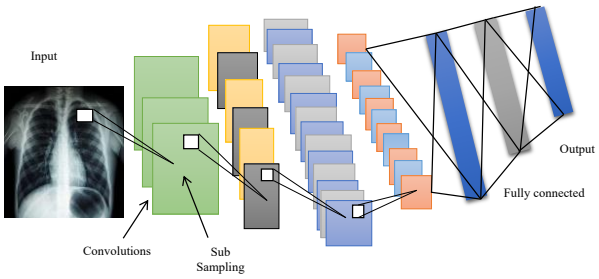

Fig. 1.1. CNN Architecture

Transfer Learning

Larger datasets tend to yield better results for CNNs than smaller ones. [3] In applications that use CNN when the dataset is small, transfer learning may be helpful. Figure 1.2 illustrates the idea of transfer learning, where a trained model from a larger dataset like ImageNet can be applied to a shorter dataset. Recent years have seen the successful implementation of transfer learning in a number of practical applications, including manufacturing, healthcare, and baggage security. This lowers the lengthy training period and eliminates the need for a huge dataset, which are prerequisite for the algorithm for deep learning when it is created from scratch.

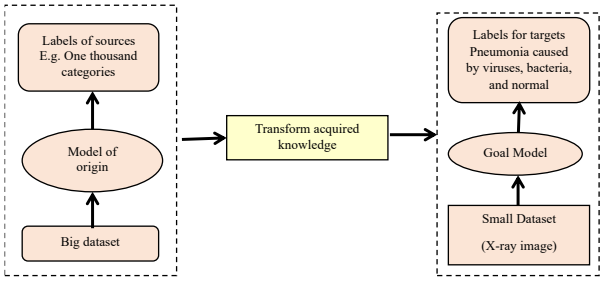

Fig. 1.2. Notion of Transfer Learning

3D Convolutional Neural Networks (3D CNNs)

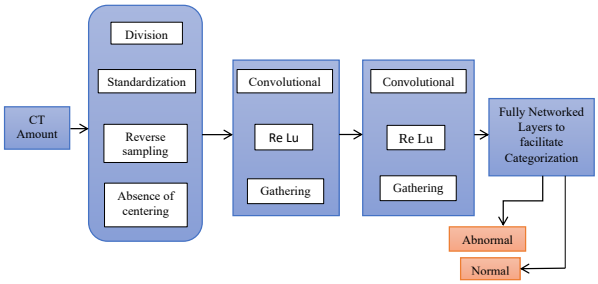

Fig. 1.3. Structure of three-dimensional convolutional neural networks

The first method was to just feed the 3D CT scans that had already been pre-processed into three-dimensional CNNs, but the outcomes were not good. [4] To feed the 3D CNNs with only areas of interest, more pre-processing was done. A U-Net received instruction for nodule applicant detection to pinpoint locations of interest. To finally classify the CT images as either negative or positive for cancer of the lung, input regions surrounding nodule candidates found by the U-Net were fed into 3D CNNs. Figure 2 depicts the general architecture; the next sections will go over each layer's specifics.

The essay's remaining sections are organized as follows: Section 2 presents the research on the pertinent prior work. Section 3 describes the features of the proposed system, including the proposed system architecture, implementation model, characteristics of the graph-based technique, and data analysis. The implementation environment is described, and the system's effectiveness is rated in Section 4. Section 5 provides the resolution.

2. Related Works

There are numerous methods for assessing, keeping track of, and controlling human behavior in the healthcare industry. These methods offer solutions from a variety of angles. An extensive literature review that describes the body of research on environmentally assisted living alternatives and facilitates comprehension of how it encourages and supports heart disease patients in taking care of themselves to lower mortality and morbidity associated with the condition. [5] The research is divided into four main themes: wearable technology, medical management systems, environmental assisted living for elderly people, self-monitoring, and deep learning-based cardiac disease diagnosis systems. The method offered a novel framework built on deep learning and cloud-oriented statistics for intelligent patient monitoring and recommendations. The unbalanced dataset collected on an individual with persistent blood pressure issues was the subject of a case study, and the patient's condition was noted. Based on the outcomes of the experiments, the research demonstrated the method's efficacy.

A method that uses pictures of blood cells to automatically identify and categorize parasites that cause malaria in blood. In addition to proposing a model for the identification and categorization of parasites from malaria in samples of blood obtained via light microscopy, they have experimented with a computerized diagnostic method for the quick and reliable diagnosis of problems in RBC. [6] To confirm the presence of the parasite that causes influenza inside thin blood smears, this study also used image classification. The features are produced based on color, texture, and cell and parasite

geometry. To differentiate between normal and infected blood cells, the study used a neural network classifier.

Despite the introduction of the first computer-aided design system in the late 1980s to identify nodules in the lungs or impacted lung cells, it was not enough. This is due to the fact that there were many computers at the time that were not powerful enough to use advanced image processing techniques. Diagnosing lung disease with basic image processing techniques takes longer. [7] The performance of decision-making systems and CAD has significantly improved as a result of the creation of CNN and GPUs. A wide range of deep learning models have been proposed by numerous studies for the diagnosis of lung diseases, including lung cancer. It is recommended to use a 3D deep CNN to locate lung nodules utilising segmented images and multi-scale forecasting methods.

The number of tests carried out is extremely low about the pace of spread because many testing kits are not readily available. To solve this problem, a new system that can accurately diagnose the existence of a viral infection in the human body must be created. [8] This can be accomplished by examining the COVID-19 patient's X-ray or computed tomography (CT) pictures. When paired with other AI algorithms, image analysis may be the most effective way to give radiologists a second opinion when determining whether or not this virus is present. Recently, several detection technologies have been created to identify influenza in the pulmonary system and other tissues. CT and X-ray scans are the imaging modalities taken into consideration.

Deep convolutional neural networks (CNNs) can identify lymph nodes in clinical diagnostic tasks and produce dramatic outcomes, even when surrounding structures from computer tomography have low contrast. [9] A different study used deep CNN to solve the issues of interstitial lung disease classification and thoracoabdominal lymph detection. With three positives that were false per patient and 85% sensitivity, they built various CNN designs and had encouraging results. Ronneburger et al. used data augmentation to construct a CNN technique. According to their suggestion, the constructed model achieved great accuracy even when educated on tiny amounts of data from images obtained using light that traveled microscope.

By looking into the application of DL models to analyze CXR pictures having SARS-CoV infections caused by viruses, 34 publications were analyzed. Most of the research (71%) used publicly accessible CNN structures built using the [10] ImageNet data set to implement transfer learning. The public can access these architectures together with their variables and hyperparameter settings. Yet, 29% of the research used unique architectures instead of using commercial tools. The primary methodologies and datasets employed in the research projects this survey assessed are briefly described in the subsections that follow.

The use of AI and deep learning in the diagnosis of medical pictures, such as CT (computed tomography) scans, has been the subject of numerous studies and research projects. To evaluate 78 brain CT scans, the DenseNet design and recurrent neural network layer are included. At the CT level, RADnet exhibits an accuracy of 82.80% in hemorrhage prediction. [11] For the categorization of lung cancer, three different types of deep neural networks—CNN, DNN, and SAE—were created. It was discovered that the accuracy of the model used by CNN was superior to that of the other models. Deep learning, more especially the use of convolutional neural network evaluations, has the

potential to identify and classify chronic obstructive pulmonary disease as well as forecast smokers' mortality and episodes of acute respiratory illnesses.

3. Methods and Materials

Applications of AI are being used to address a wide range of biological issues and complications, including the detection of brain tumors, lung diseases, tumours in the breast, and various other oncological emergencies. Convolutional neural networks, or CNNs, are a type of deep learning technique that have been shown to be effective in revealing picture features that aren't obvious in the original image. The level of detail of the imaging determines how accurate the deep learning system is. CNN can enhance the quality of images obtained from a fast-paced video endoscopy in low light, detect pediatric influenza from CXR images, find pulmonary nodules from CT images, automatically label polyps from a colonoscopy, and analyze cystoscopy images from videos. Therefore, photos of solely confirmed probable COVID-19 patients were chosen. When physical testing, especially RT-PCR, is few, deep learning algorithms and models might be built as tools for triage situations. The United States College of Imaging, or ACR, only recommended using mobile devices CXR in a functioning care facility when needed to inform choices about a presumed COVID-19 patient, which includes whether to perform an RT-PCR test, recognise the individual, use additional medications, and hinder the patient from receiving treatment in isolation or being put with others. They also strongly discouraged the use of CT. On the other hand, deep learning algorithms and models need to forecast patient outcomes and enable the doctor to expeditiously assist with treatment and management.

LIDC-IDRI Dataset

We assess the framework's performance using the LUNA16 task. The dataset is derived from the LIDC-IDRI, the biggest publicly accessible standard database for lung nodules, which comprises 1017 CT scans in total. The LUNA16 dataset only contains detection annotations, but LIDC-IDRI has almost every relevant detail for low-dose lung CT scans, with comments from numerous clinicians on nodule dimensions, locations, diagnostic results, nodular appearance, nodular margin, and other information. CTs with identified nodules below three mm, uneven or missing cut spacing in the LIDC-IDRI dataset, and slice depths greater than 3 mm are all removed from the LUNA16 dataset. The goal of splitting the remaining 887 scans into 10-folds is to do cross-validation on them.

There are now three candidate detection techniques available for computing the candidate locations. Lesions that have a diameter ≤ 6 mm are combined because numerous candidates can identify the lesion. 745,967 candidates in total are included in LUNA16, and each candidate has the associated label of class. Remember that any nodule may have more than one contender. [13]A significant disparity between the erroneous positive options and the genuine nodules presents a problem for the false positive reduction step of the dataset. Data augmentation is the most popular and straightforward way to address the imbalance. Using picture data, we perform horizontal reflection and image transformations that are comparable. To address the imbalance of classes, a pre-screening procedure is used. By applying down sampling at random to the negative nodular class, the total amount of negative samples is brought closer to that of the positive samples. Then, we use the AlexNet model in the experiment; they combine to train a small CNN. The model that is trained only retains the incorrectly classified nodules of samples that are negative for the subsequent training

cycle. The ratio of negative to positive samples is 100:2 following the initial screening of negative samples. To address the extreme class imbalance and pick more accurate and discriminative training and testing samples, only the misclassified nodules need to be kept. This will improve the model's robustness.

A public database, comprising thoracic computed tomography (CT) scans of 1011 patients with lung cancer and annotations (nodule outlines) from a maximum of four radiologists, is available due to the LIDC. Although annotations are available for more than a thousand patients, only 155 of those individuals have diagnostic data that includes nodule ratings. The scores were acquired through biopsy, surgical excision, development, or radiographic image assessment to demonstrate the nodule condition during two years at two levels: the patient's level initially, and the nodule level second. [14] Over an extended period using a variety of scanners, the LIDC database containing thoracic CT examinations for 1010 participants was gathered. A distinct quantity of slices may be connected to certain nodule research since there can be several nodules connected to a single study. Additionally, each slice may include annotation data for a tumor from a maximum of four radiologists. These annotations are only accessible for nodules larger than 3 mm in size in the LIDC dataset. We opted to utilize the evaluations derived from diagnostic data as the true basis to teach the system of classification and assessing the outcomes rather than utilizing the radiologists' ratings in the dataset because the diagnostic information is the sole way to determine the certainty of malignancy.

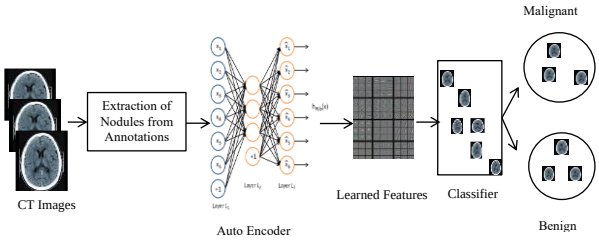

Fig. 3.1. An illustration of the various CAD system components that are suggested.

NIH dataset

The National Institutes of Health (NIH) provided us with datasets of chest X-rays for this investigation. The NIH dataset is made up of 120,000 1024 × 1024 photos annotated with 13 different diseases. The images were obtained from 30,805 different patients. Thirteen diseases are taken into consideration. To enhance the image, 2D chest X-rays are processed as input. Next, we feed the image into a neural network that has been trained on a variety of illnesses. Next, the model's capacity to recognise disease is assessed using an X-ray of the chest from the learning dataset, and the output is displayed. The test and training folders contain the dataset. To assess the effectiveness of our model, we used CX-Ultranet's weights that were previously trained from the NIH Dataset in the test folder. [15] There are 234 normal chest X-rays and 390 chest X-rays showing pneumonia in the test folder.

Using pictures from the NIH ChestX-ray14 dataset, we use the bone extractor U-Net to investigate the practicality of our technique to real X-rays. Perceptual loss was used throughout the training process to get the desired results. The outcomes of the augmentation on a real X-ray are displayed in

Figure 3.2. It is evident that the excised bone layer no longer contains the heart or any other soft components.

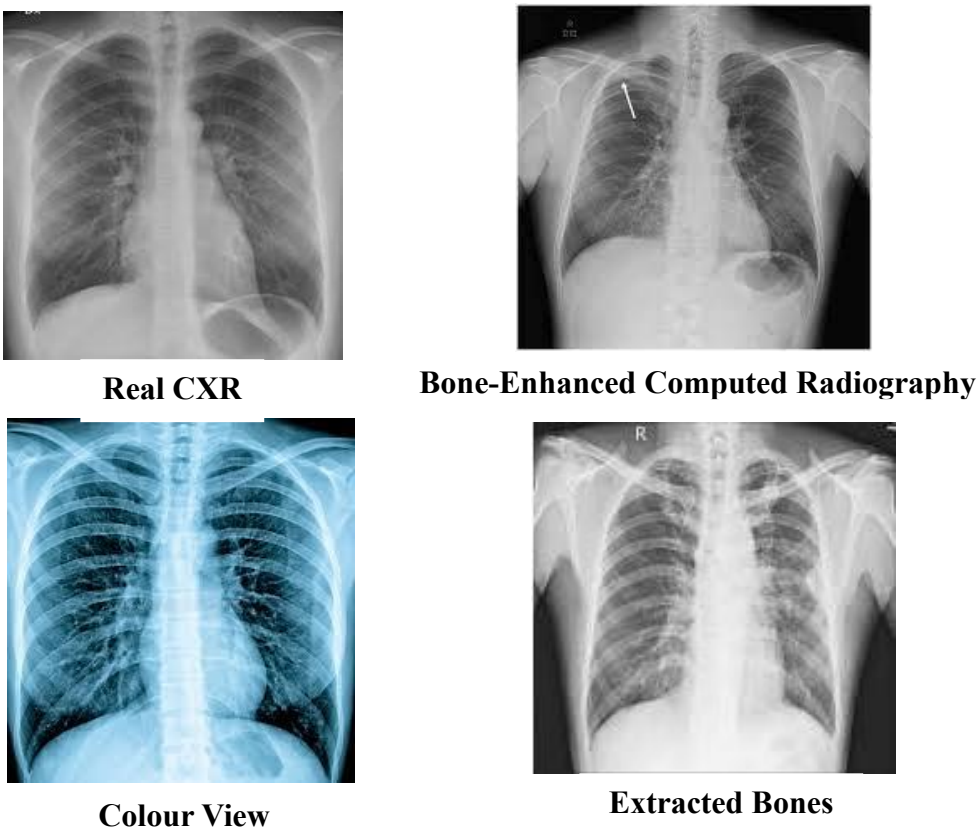

Fig. 3.2. Actual X-ray Findings - NIH X-ray14

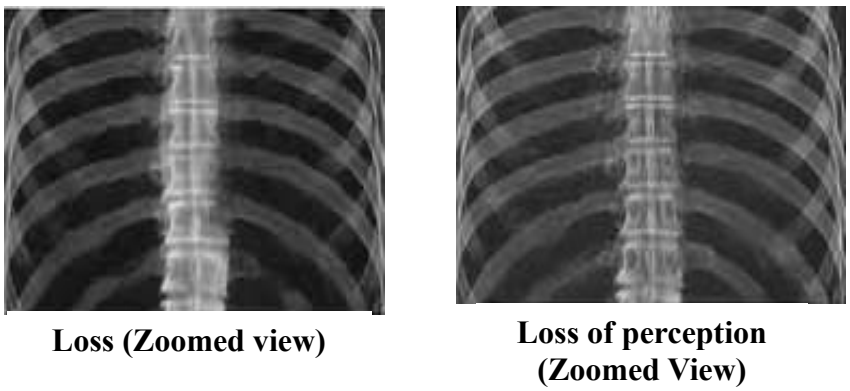

Fig. 3.3. Loss versus Loss of perception- NIH X-ray14

Two U-Nets one trained with L1 and the other with Perceptual Loss created the two bone images shown in Figure 3.3. The picture's sharpness is improved and we were able to capture more detail in the vertebrae by using Perceptual Loss.

Using local X-ray images of the chest and online datasets, the model was tested, validated, and trained. 11,716 labelled X-ray pictures were collected using the NIH datasets source. Further data was gathered from LIDC-IDRI, including 443 local X-ray pictures. An example of the kind and quantity of picture data gathered from NIH and LIDC-IDRI is provided Table 3.1.

Table 3.1. Quantity and types of image data gathered from LIDC-IDRI and the NIH database

Disease	Data from NIH	Data from LIDC-IDRI
Pneumonia	1813	15
Lung cancer	908	45
TB	1251	156
COPD	756	65
Normal	1552	89
Pneumothorax	1535	36
Total	7,815	406

Pre-processing:

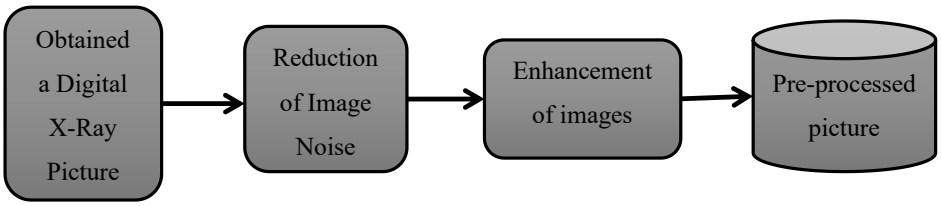

Fig. 3.4. Pre-processing steps for images include median filter reduction of noise and CLAHE enhancement

Pre-processing is mostly done to improve the quality of photographs by eliminating artifacts and emphasizing key elements. Images in X-ray radiography are typically grayscale, with individual pixels detecting different photons; certain pixels receive more photons from the X-ray and look darker, while other pixels receive less radiation and appear lighter. [16] Nevertheless, the random and shaded aspect of this pixel distribution leads to a pepper and salt distribution. X-ray pictures are typically primarily impacted by Poisson, sprinkle, pepper, and salt noises. While there are many different de-noising methods, the median filtering technique is the most useful one for removing speckle, Poisson, and pepper and salt noises from X-ray images. Every image in this has been given a median filter. The CLAHE method was applied to improve the image after the noise was removed. Histogram equalisation (HE), adaptive histogram equalising (AHE), and CLAHE were employed as a comparative tool. The enhanced adaptive histogram equalisation method, known as CLAHE, applies contrast to a confined area of a grayscale image. The purpose of CLAHE is to stop adaptive histogram equalization from causing other desired regions to be affected by noise and from amplifying noise overall. The pre-processing method used in this paper is shown in Figure 3.4.

Data Augmentation

Deep learning models for identifying images face significant challenges due to limited data. Unbalanced classes are frequently associated with problems; essential classes may have enough data, while under-sampled classes will have low class-specific precision. The issues brought about by incomplete data and unequal class sizes can be resolved in a variety of ways. Enhancing existing photographs is a helpful method to expand the training set's size without having to buy new ones. To address the class imbalance and enhance the quantity of data, data augmentation employing rotations of 48, 160, and 2750 degrees has been used in this work. The total number of images rose from 13,149 to 17,215 after image augmentation was applied to the pre-processed image. Furthermore, the

model required that the photos be scaled from their original 1066 × 1066 resolution to an identical dimension of 299 × 299.

Deep Learning Methods

Training a CNN from scratch is typically challenging since choosing the right model design for correct convergence necessitates a huge amount of training data and a high level of experience. Professional annotation is costly, and data in medical applications is usually limited. Deep CNN training takes a very long time since it needs a lot of memory and processing power. When there is insufficient data, transfer learning (TL) presents a viable option for fine-tuning a CNN that has already been pre-trained on a sizable collection of labeled images from a different category. This reduces computing complexity throughout training and accelerates convergence.

The majority of vision tasks are typically handled by CNNs with their early layers learning high-level image properties. Conversely, higher-level, more application-specific traits are learned by the subsequent levels. Because of this, transfer learning typically just requires a little tweaking of the final few layers. Typically, the final fully-connected layer of the pre-trained CNN is replaced with a new completely connected plane with precisely the same number of cells as the subclasses in the newly generated target application. The remaining layers hold the remaining weights of the pre-trained network. The characteristics produced in the layer before would be used to train a linear classifier in this manner. [17] Using models that have already been trained on massive datasets such as ImageNet7 through training through transfer learning (TL) for a new job that involves an additional set with fewer data points is a popular practice in the literature. However, it has recently been shown through experimentation that TL has very little use if the source and targeted domains are quite different, such as images from hospitals and photographs taken in the natural world, since the networks may learn entirely different high-level features in the two cases.

Our method is to use Data-A to train a CNN model from scratch such that it can distinguish between normal and sick X-ray pictures. Next, we add a new fully-connected phase with just as many cells as the number of classes (in our example, four: normal, other disease, pneumonia, and tuberculosis) to replace the final fully connected layer of the pre-trained CNN. The remaining weights in the pre-trained network's remaining layers are kept. Using the characteristics produced in the previous layer to train a linear classifier is the equivalent of this. Below is an outline of the adopted procedure.

- Pre-trained weights for the model trained on Data-A should be loaded into the convolutional base upon instantiation.
- Take off the pre-trained CNN's final fully-connected layer and add a fresh one in its place.
- Model layers should be frozen up to the final convolutional block.
- Finally, use the Stochastic Gradient Descent (SGD) optimization approach with a very slow learning rate to retrain the final Fourier block and the fully linked layers.

TB (Tuberculosis)

Pneumonia

Fig. 3.5. Extracted images of Tuberculosis and Pneumonia

Three widely-used CNN models AlexNet (eight layers), VGGNet (sixteen layers), and ResNet (fifty layers) with progressively more layers were employed. Because global average pooling is used instead of fully connected layers, ResNet's model size is significantly smaller even if it is deeper than VGGNet's. We used smaller three-by-three kernels instead of huge ones like in AlexNet (eleven × eleven) VGGNet, and ResNet. Because there are fewer trainable weights for smaller kernels, it can be observed that smaller kernels result in better regularisation, offering the opportunity to build larger networks without losing excessive data in the layers. At first, all models were trained using Data-A from scratch, enabling them to differentiate between normal and sick X-ray pictures. The models were then refined on Data-B using the previously mentioned TL technique to teach them to distinguish between the four classes: normal, other disease, pneumonia, and tuberculosis.

4. Result Analysis and Discussion

Dataset:

A collection of chest X-ray pictures and associated information made accessible for use in research by the National Institutes of Health Clinical Center is known as the NIH Lungs dataset. More than 100,000 anonymous chest X-ray pictures and related metadata, including patient gender, age, and diagnosis, make up the dataset. The goal of the NIH Lungs database is to support researchers studying image analysis and machine learning methods for recognizing and making a diagnosis of lung disorders such as COPD, respiratory infections, and lung cancer. The dataset is meant to be used for machine learning algorithm development and evaluation. It is separated into training and

testing sets. The NIH Clinical Centre's open data project includes several datasets, including the NIH Lungs dataset.

The National Institutes of Health chest X-ray dataset which has pictures and information from a wider range of illnesses, and the National Institutes of Health database, which contains pictures of anomalies and lesions in other body parts, are two further datasets made available by this program. Computerized tomography, or CT, images of the chest are included in the CT chest LIDC-IDRI dataset. For research purposes, it was made publicly available after being gathered by the LIDC-IDRI. Many CT scans and their accompanying annotations showing lung abnormalities, including lung nodules and pneumonia caused by tuberculosis, are included in the dataset.

Metrics of measurement:

A deep learning model's performance for computed tomography of the chest and X-ray images can be assessed using metrics such as recall, sensibility, adequacy, specificity, and precision.

- Accuracy is defined as the proportion of all samples that can be correctly identified. This widely used metric is straightforward, but it may be deceptive if there is an imbalance in the distribution of classes.
- The fraction of real negatives that are accurately classified as negatives is known as specificity.
- The proportion of true positives that are accurately classified as positive, relative to all real positives, is known as sensitivity. Other names for specificity and sensitivity are the real positive rate and real negative rate, respectively.
- The percentage of genuine positives among every sample that was anticipated to be positive is known as precision.
- The percentage of genuine positives among all actual samples that are positive is known as recall (sensitivity).

Table 4.1. Different evaluation metrics results on both training datasets (LIDC-IDRI CT and NIH) in terms of accuracy, sensitivity, specificity, and precision.

Model	Accuracy	Precision	Sensitivity	Specificity	F1-Score
Model A	82.45	83.02	85.57	86.54	87.86
Model B	86.75	84.36	83.75	87.46	86.48
Model C	83.67	88.07	82.46	84.67	84.36
Model D	80.64	86.76	89.46	86.55	82.67
Transfer Learning	**93.32**	**95.53**	**94.86**	**92.56**	**91.46**

Because they offer a reasonable balance between positive outcomes and false negatives—a critical aspect in the healthcare sector—the F1 score and accuracy are often utilised metrics for evaluation in medical imaging analysis. They do, however, come with drawbacks. [18] One disadvantage of the F1 ranking and accuracy is that it does not take into consideration the genuine negatives, which in the case of medical image analysis could make up a significant amount of the data set. Thus, the accuracy of the model as a whole could not be sufficiently reflected in the data. Additionally, in cases when the class dispersion is significantly unbalanced, these measurements may not be accurate enough to detect subtle variations in system performance. Several other assessment metrics may be

more appropriate for medical picture analysis, depending on the specific use case. For example, both the specificity and the sensitivity measurements can be used to find the fraction of real positives and true negatives, respectively. The ROC curves and the value of the area under the curve (AUC) can also be used to evaluate the efficacy of binary classifiers. Moreover, segmentation models can be evaluated using a variety of metrics, such as the Dice frequency and the Jacquard index. The specific work at hand and the investigation's goals ultimately determine the assessment metric to be used. Before to choosing the best metric for the given situation, it is necessary to carefully consider the advantages and disadvantages of each one. Therefore, we assess our suggested technique using additional measurement metrics, including ROC and sensitivity.

Table 4.2 presents the evaluation's results, which are obtained through the use of k-fold cross-validation. Under fold-5, the model demonstrates the best scores of 96.88% specificity and 97.31% sensitivity, indicating that it has good specificity and sensitivity across the board. Further evidence is provided that the fused images preserve the initial X-ray image properties.

Table 4.2. Calculating the F1-Score for accuracy, sensitivity, specificity, and precision on two separate datasets (NIH and LIDC-IDRI CT).

Model	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**F1-Score**
Model A	92.45	93.02	95.57	96.54	97.86
Model B	96.75	94.36	93.75	97.46	96.48
Model C	93.67	98.07	92.46	94.67	94.36
Model D	90.64	96.76	99.46	96.55	92.67
Model E	93.32	95.53	94.86	92.56	91.46
Average	**95.57**	**95.23**	**95.46**	**95.67**	**95.64**

Performance metrics, including accuracy, specificity, recall, precision, and F1 score were taken into account when assessing the suggested model. The measures' formulas are outlined in Eqs. (1) to (4). Precision is a fidelity metric and a useful tool for identifying false positives or chest x-ray images that do not show viral or bacterial pneumonia. A high accuracy rating suggests fewer false positives. [19] As a measure of thoroughness, recall is frequently used to identify pneumonia in many applications. High detection rate and recall value are correlated. For verification or classification problems, accuracy is the most useful metric. The accuracy is tested using the F1 score. The score is calculated using both recall as well as accuracy.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \times 100\% \quad (1)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \times 100\% \quad (2)$$

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Negative + True\ Negative + False\ Positive} \times 100\% \quad (3)$$

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Real positives, true negatives, false positives, and false negatives are denoted by the symbols TP, TN, FP, and FN in Eqs. (1) Through (4), respectively. The pixel intensity histogram, graphical feature maps, practical evaluation of the three models, and comparison with the reference models are covered in the ensuing sections.

The CNN architecture's performance is shown in this part, along with a comparison of the model with related studies. Using various hyperparameter variations, we conducted experiments on the datasets with the suggested CNN model. For example, we showed the output of the model and examined how well it performs when the SGD and Adam optimization techniques are used. [20] Additionally, using the model we suggested as a test, we investigated the impact of two alternative L2 regularisation approach values. The Adam optimization algorithm and a decaying weight (L2) with a value of 0.0002 were employed in the first set of tests. Figure 4.1 depicts the model's performance as measured by the loss function, and Figure 4.2 displays the accuracy trajectory for both the training and validation examples. Take note that the following is the Adam optimizer configuration: epsilon = 1e-8, learning rate = 0.001, which is beta1 = 0.9 and beta2 = 0.999.

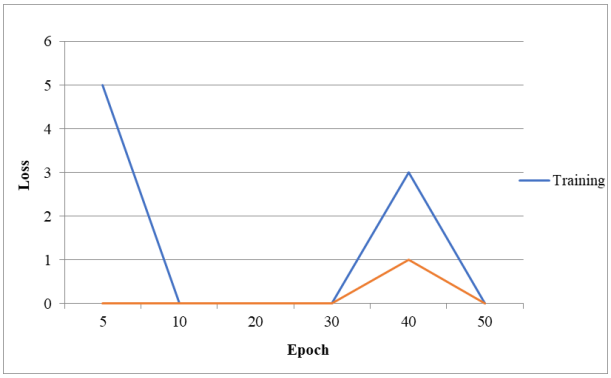

Fig. 4.1. Pattern of variation in the training & verification loss functions for the combined dataset

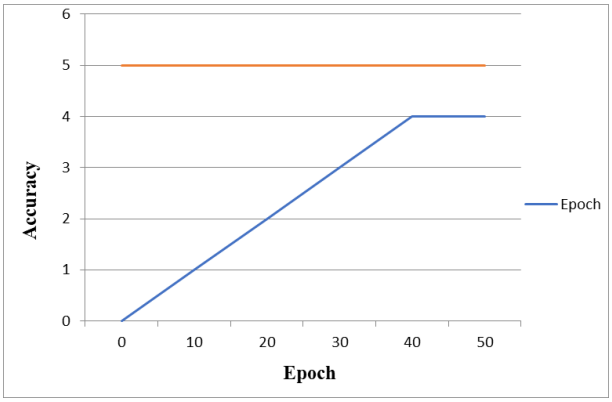

Fig. 4.2. Pattern of variation in the training & verification accuracy functions for the combined dataset

5. Conclusion

An automatic method for detecting pneumonia and its classifications using deep-CNN-based transfer learning techniques is presented in this work. Using chest X-rays and CT images, four distinct widely used CNN-based deep learning methods were developed and evaluated to distinguishing patients with pneumonia and those without. It was found that DenseNet201 performs better than the three other deep CNN networks. To attain an excellent F1 score and precision, the technique combines a distributed CNN framework (Hybrid Feature Fusion, or HFF) for several modalities in conjunction with a class balance algorithm and Support Vector Machine (SVM) for image classification in clustering. The method has strong generalization and accuracy, making it a potential tool for doctors and medical professionals. This paper proposes a method that detects lung illness

symptoms by using a class-balancing algorithm and deep learning methods to train and analyze chest X-rays and CT pictures.

The CT scan and chest X-ray pictures from the NIH chest-X-ray dataset and the LIDC-IDRI were used. To improve the quality and remove noise distortion, these photos are first pre-processed using CLAHE. We use the Honey Badger Algorithm to segment the lung pictures. In the end, a deep pyramid residual neural network is used to classify these segmented images as cancerous or non-cancerous. The results obtained from the LIDC-IDRI database indicate that the accuracy levels are 98.5 % for healthy lungs 97.44 % for abnormal lungs, 98.14 % for normal lungs 94.66% for abnormal lungs, respectively, based on these databases.

References

[1] Tataru, C., Yi, D., Shenoyas, A., & Ma, A. (2017, June). Deep Learning for abnormality detection in Chest X-Ray images. In *IEEE conference on deep learning*. Adelaide: IEEE.

[2] Bhandary, A., Prabhu, G. A., Rajinikanth, V., Thanaraj, K. P., Satapathy, S. C., Robbins, D. E., ... & Raja, N. S. M. (2020). Deep-learning framework to detect lung abnormality–A study with chest X-Ray and lung CT scan images. *Pattern Recognition Letters*, *129*, 271-278.

[3] Rahman, T., Chowdhury, M. E., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., ... & Kashem, S. (2020). Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. *Applied Sciences*, *10*(9), 3233.

[4] Alakwaa, W., Nassef, M., & Badr, A. (2017). Lung cancer detection and classification with 3D convolutional neural network (3D-CNN). *International Journal of Advanced Computer Science and Applications*, *8*(8).

[5] Wang, Y., Nazir, S., & Shafiq, M. (2021). An overview on analyzing deep learning and transfer learning approaches for health monitoring. *Computational and Mathematical Methods in Medicine*, *2021*, 1-10.

[6] Gourisaria, M. K., Das, S., Sharma, R., Rautaray, S. S., & Pandey, M. (2020). A deep learning model for malaria disease detection and analysis using deep convolutional neural networks. *International Journal of Emerging Technologies*, *11*(2), 699-704.

[7] Bharati, S., Podder, P., & Mondal, M. R. H. (2020). Hybrid deep learning for detecting lung diseases from X-ray images. *Informatics in Medicine Unlocked*, *20*, 100391.

[8] Nigam, B., Nigam, A., Jain, R., Dodia, S., Arora, N., & Annappa, B. (2021). COVID-19: Automatic detection from X-ray images by utilizing deep learning methods. *Expert Systems with Applications*, *176*, 114883.

[9] Jaiswal, A. K., Tiwari, P., Kumar, S., Gupta, D., Khanna, A., & Rodrigues, J. J. (2019). Identifying pneumonia in chest X-rays: A deep learning approach. *Measurement*, *145*, 511-518.

[10] Alghamdi, H. S., Amoudi, G., Elhag, S., Saeedi, K., & Nasser, J. (2021). Deep learning approaches for detecting COVID-19 from chest X-ray images: A survey. *Ieee Access*, *9*, 20235-20254.

[11] Shah, V., Keniya, R., Shridharani, A., Punjabi, M., Shah, J., & Mehendale, N. (2021). Diagnosis of COVID-19 using CT scan images and deep learning techniques. Emergency radiology, 28, 497-505.

[12] Serena Low, W. C., Chuah, J. H., Tee, C. A. T., Anis, S., Shoaib, M. A., Faisal, A., ... & Lai, K. W. (2021). An overview of deep learning techniques on chest X-ray and CT scan identification of COVID-19. Computational and Mathematical Methods in Medicine, 2021, 1-17.

[13] Xie, H., Yang, D., Sun, N., Chen, Z., & Zhang, Y. (2019). Automated pulmonary nodule detection in CT images using deep convolutional neural networks. Pattern recognition, 85, 109-119.

[14] Kumar, D., Wong, A., & Clausi, D. A. (2015, June). Lung nodule classification using deep features in CT images. In 2015 12th conference on computer and robot vision (pp. 133-138). IEEE.

[15] Gozes, O., & Greenspan, H. (2020, April). Bone structures extraction and enhancement in chest radiographs via CNN trained on synthetic data. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (pp. 858-861). IEEE.

[16] Yimer, F., Tessema, A. W., & Simegn, G. L. (2021). Multiple lung diseases classification from chest X-ray images using deep learning approach. Int. J, 10, 2936-2946.

[17] Basu, S., Mitra, S., & Saha, N. (2020, December). Deep learning for screening covid-19 using chest x-ray images. In 2020 IEEE symposium series on computational intelligence (SSCI) (pp. 2521-2527). IEEE.
[18] Iqbal, S., Qureshi, A. N., Li, J., Choudhry, I. A., & Mahmood, T. (2023). Dynamic learning for imbalanced data in learning chest X-ray and CT images. Heliyon, 9(6).
[19] Manickam, A., Jiang, J., Zhou, Y., Sagar, A., Soundrapandiyan, R., & Samuel, R. D. J. (2021). Automated pneumonia detection on chest X-ray images: A deep learning approach with different optimizers and transfer learning architectures. Measurement, 184, 109953.
[20] Oyelade, O. N., Ezugwu, A. E. S., & Chiroma, H. (2021). CovFrameNet: An enhanced deep learning framework for COVID-19 detection. Ieee Access, 9, 77905-77919.