PeerJ Computer Science

AI-driven deep CNN approach for multi-label pathology classification using chest X-Rays

Saleh Albahli[1], Hafiz Tayyab Rauf[2], Abdulelah Algosaibi[3] and Valentina Emilia Balas[4]

[1] Department of Information Technology, College of Computer Science, Qassim University, Buraydah, Saudi Arabia
[2] Centre for Smart Systems, AI and Cybersecurity, Staffordshire University, stoke on Trent, United Kingdom
[3] Department of Computer Science, King Faisal University, Hofuf, Saudi Arabia
[4] Department of Automation and Applied Informatics, Aurel Vlaicu University of Arad, Arad, Romania

ABSTRACT

Artificial intelligence (AI) has played a significant role in image analysis and feature extraction, applied to detect and diagnose a wide range of chest-related diseases. Although several researchers have used current state-of-the-art approaches and have produced impressive chest-related clinical outcomes, specific techniques may not contribute many advantages if one type of disease is detected without the rest being identified. Those who tried to identify multiple chest-related diseases were ineffective due to insufficient data and the available data not being balanced. This research provides a significant contribution to the healthcare industry and the research community by proposing a synthetic data augmentation in three deep Convolutional Neural Networks (CNNs) architectures for the detection of 14 chest-related diseases. The employed models are DenseNet121, InceptionResNetV2, and ResNet152V2; after training and validation, an average ROC-AUC score of 0.80 was obtained competitive as compared to the previous models that were trained for multi-class classification to detect anomalies in x-ray images. This research illustrates how the proposed model practices state-of-the-art deep neural networks to classify 14 chest-related diseases with better accuracy.

Subjects Artificial Intelligence, Computer Vision
Keywords Image classification, Chest diseases, InceptionResNetV2, Pathology

Submitted 17 February 2021
Accepted 27 March 2021
Published 20 April 2021

Corresponding authors
Hafiz Tayyab Rauf,
hafiztayyabrauf093@gmail.com
Valentina Emilia Balas,
balas@drbalas.ro,
valentina.balas@uav.ro

Academic editor
Robertas Damaševičius

Additional Information and Declarations can be found on page 13

DOI **10.7717/peerj-cs.495**

INTRODUCTION

Medical X-rays, short for X-radiation, are a way to look to visible light in classical physics but with higher energy hits the body (*Panwar et al., 2020*). X-ray is employed to generate images of tissues and structures inside the body; these include bones, chest, teeth, and so on *Rajaraman & Antani (2020)*. X-rays are handy diagnostic tools used for several decades by specialists to detect fractures, certain tumors, pneumonia, dental problems, and others. In advanced cases, CT (Computed Tomography) can produce a series of body images that are later assembled into a three-dimensional X-ray image processed by computer. However, the standard X-ray is faster, easier, cheaper, and less harmful than the CT scan (*Rajaraman & Antani, 2020*).

For most cases, X-rays are not enough for the radiologist to carry out their diagnosis; they usually require CT scans to confirm the diseases (*Hashmi et al., 2020*). Diseases with quick progression might even require multiple CT scans, which is very expensive, time-consuming, and even does the patient harm since both of them work with radiation (*Chatterjee et al., 2017*). Hence, an urgent need for an artificial intelligence algorithm that could accurately detect chest-related diseases (*Islam et al., 2020*; *Salehinejad et al., 2018c*; *Salehinejad et al., 2018a*).

Research done involved using deep learning to make predictions of medical images by extracting features from the images, including the shape and spatial rotation. Convolutional neural networks (CNNs) have played a very vital role in feature extraction and learning patterns that enabled prediction (*Demir, Sengur & Bajaj, 2020*; *Rustam et al., 2020*; *Toğaçar, Ergen & Cömert, 2020*; *Militante & Sibbaluca, 2020*). For example, a CNN was used to improve extraction high-speed video endoscopy (HSV) when the training data is very little (*Wang et al., 2017*).

AI is now mature to be embedded with the state of the art machine learning, and deep learning algorithm in different domains such as medicine (*Meraj et al., 2019*), agriculture (*Rauf et al., 2019*), biometrics (*Jin, Cruz & Goncalves, 2020*), renewable energy (*Gao, Wang & Shen, 2020b*), and cloud computing (*Gao, Wang & Shen, 2020c*; *Gao, Wang & Shen, 2020a*). AI is gradually changing medicine as we know it, as more and more ways are being researched to automate medical practices and support the doctors and specialists (*Brag, 2020*; *Rauf et al., 2020*; *Kulkarni et al., 2020*; *Briganti & Le Moine, 2020*). Many benefits are attached to artificial intelligence in medicine (*Jain et al., 2020*). For example, rural areas and third-world countries usually have a scarcity of specialist doctors who could give them the care they need; artificial intelligence tools help them get the required medical assistance. Also, researches like this will help significantly in the health-care industry as specialists can use it to confirm their diagnosis; it can also be used by people who have little to no medical background. Even the patients could use this tool (*Mongan, Moy & Kahn, 2020*; *Ali et al., 2020*; *Tabares-Soto et al., 2020*). All this is aimed at lessening the specialist's burden since there is just a hand full of them, so they are focused only mainly on the X-rays that are flagged as suspicious by the tools used for the prediction. They also reduce doctors' subjective opinions drastically, increase the speed of diagnosis, and accurately detect features that the human eye may ignore (*Zhang et al., 2020*).

In this research, we employed deep neural networks to predict a wide range of chest-related diseases. Previous research used state-of-the-art techniques and got highly accurate results, but these were only able to detect one disease and were virtually useless for other diseases. The proposed methodology's reliability and efficiency are slightly lower due to lack of data augmentation, highly normalized data, and ineffective hyperparameters. There is no practical way of accurately identifying multiple chest-related diseases from X-ray images with reliable results to the best of our knowledge.

The proposed deep learning models are employed to automatically detect the following chest-related diseases: Hernia, Pleural, Fibrosis, Emphysema, Pneumothorax, Pneumonia Atelectasis, Edema, Nodule, Mass, Infiltration, Effusion, Cardiomegaly and Consolidation and from the frontal view of chest X-ray images. Three different deep neural network

architectures with large input types which have been pre-trained on extensive datasets are employed.

The objectives of the proposed research are as follow:

- To identify multiple chest-related diseases using deep CNNs and to tackle insufficient and unbalanced data through transfer learning.
- to employ synthetic data augmentation using deep CNNs for the detection of 14 chest-related diseases.
- To evaluate DenseNet121, InceptionResNetV2, and ResNet152V2 on multi-class identification problem.

RELATED WORK

Research done in medicine uses artificial intelligence to ease diagnosis, and some of them got beneficial and accurate results. Here we will discuss techniques used by previous researchers that used artificial intelligence to tackle chest-related diseases, using deep neural networks automatically. High-end medical infrastructures and qualified specialists are scarce, especially in rural areas and lower-middle-income countries (LMIC). Hence the medical teams available in these regions rely solely on CXRs for the detection of chest-related diseases. However, these CXRs are mainly used to diagnose pneumonia, which has to be done by a highly experienced radiologist as it is not an easy task (*Franquet, 2018*).

One-class detection of cardiothoracic diseases using deep neural networks

Sivasamy & Subashini (2020) used a Keras method to identify CXRs to predict lung diseases, and the model was 86.14 percent accurate, and it was then noted that the model's accuracy increased as the number of training epochs raised.

Zhang et al. (2018) utilized pixel wisely transcribed DRRs data to learn a supervised multi-organ segmentation model in X-ray images the X-ray image structure, the gap in Nodule Annotations is challenging and time-consuming.

Rajpurkar et al. (2020) proposed a binary X-rays pneumonia detection classifier and obtained 0.435 f1 score. *Salehinejad et al. (2018b)* adopted a custom DCGAN model for X-ray image processing, equipped to produce artificial chest X-ray photographs. Since trained on an expanded dataset with DCGAN synthesized CXRs for real dataset balance (d3), the model has been incredibly accurate.

To classify pneumonia diseases, *Chandra & Verma (2019)* used five different models and got 95,631 percent as the highest accuracy. The model was restricted only to observation on automatically driven non-rigid deformable registration in lung regions and segmented feature extraction confined to ROI.

Multi-class detection of cardiothoracic diseases using deep neural networks

Meta research was conducted in a few trials, which used state-of-the-art methods and had efficient results for one or two cardiothoracic diseases, but often they had the issue of

misclassification. There are very few strategies that have targeted the 14 types of diseases associated with the chest.

To detect degenerated lung tissues in X-ray images, the authors used image descriptors (*Ke et al., 2019*) dependent on Brightness, Hue's spatial distribution values, and saturation in X-ray images, as well as a deep neural network in conjunction with Ant Lion and Moth-Flame algorithm. Based on the proposed fitness feature, the neural network analyses the object; however, if the probability of lung infections is identified, the hybrid method determines the morphed tissues in the X-ray image in detail.

Wang et al. (2019) provided the largest publicly available dataset on X-ray images used by the research community for experiments and training of models, and most of the research has been done showed promising results using deep convolutional neural networks. They further claimed that in order to add more disease names, the dataset could be expanded. A deep learning technique to classify fourteen underlying chest-related diseases was proposed by *Rajpurkar et al. (2018)*. The proposed model was trained using a single X-ray image of the chest and generate the likelihood of each of the fourteen observations. They trained multiple models to see which one had the best performance, and in the end, DenseNet121 reached the best accuracy and was ultimately used for the testing, but it was limited to the CheXpert database and prone to over-fitting.

The authors suggested a hybrid approach (*Sahlol et al., 2020*) for classifying chest X-ray images efficiently. MobileNet, in combination with transfer learning based on a CNN model initially trained on the ImageNet dataset, is used to extract features from chest X-ray images. They used the AEO meta-heuristic algorithm as a component predictor to decide whichever of these attributes seem to be the most important. The proposed approach is tested against two publicly available benchmark datasets, Dataset 2 and Shenzhen. It enables them to obtain significant efficiency while cutting down on computing time.

Ho & Gwak (2019) used a pre-trained DenseNet121 model with feature extraction techniques to identify fourteen chest-related diseases accurately. Description of some related work indicating pros and cons of existing approaches are given in Table 1.

MATERIEL AND METHODS

ResNet (Residual Neural Network)

The ResNet model uses residual sharing, and the training process becomes more tedious and complicated as the depth of the network increases, the convergence time also increases significantly. The deep neural networks are exposed to degradation when the network starts the convergence process (*Kao et al., 2011*). For example, a deep neural network's accuracy starts to degrade rapidly when the network becomes stagnant due to saturation. This research aims at making use of residual mapping, which will fit the layers of the neural network, and this makes ResNet a perfect choice of a model for trial as it has sufficient accuracy and the depth of the model be adjusted.

DenseNet (Densely Connected Convolutional Networks)

DenseNet is a neural network used for visual object recognition, and it is quite similar to ResNet with a few fundamental differences. The traditional convolutional neural network

Table 1 Description of some related work indicating pros and cons of existing approaches.

Ref	Data set	Method	pros	cons
Pereira et al. (2020)	RYDLS-20	Early and late fusion techniques	Multi-class classification+ hierarchical classification	Data sparsity and feature optimization missing
Apostolopoulos, Aznaouridis & Tzani (2020)	Custom	Mobile Net	Considered feature extraction approaches	6-class problem
Oh, Park & Ye (2020)	Custom	Patch-based CNN	Employed Potential imaging biomarkers for validation	5-class problem
Apostolopoulos & Mpesiana (2020)	Custom	Transfer learning based CNN	Considered bacterial and viral pneumonia target class	5-class problem
Yoo et al. (2020)	Custom	Decision-tree classifier	Merging deep learning in tree classifier	3-class problem

architecture with L layers has L connections, and these connections are feed-forward, which carry information from one layer to the contagious layer. However, in *Wang et al. (2019)*, the authors suggest that there are direct (L(L+1)/2) links to a deep convolutional neural network, implying that previous layer maps for a coevolutionary neural network serve as input to the current layer. DenseNet provides densely connected front and back layers that let the context spread throughout the workout process, the main difference between Resnet and DenseNet. DenseNet also allows the reuse of features that channel connexions.

Another distinction is that ResNet uses the additive approach (+) that combines the previous layer with the next layer, while DenseNet concatenates the previous layer's output layer with the next layer. It was discovered that DenseNet could successfully reduce the number of training parameters according to the experiments in *Wang et al. (2019)*. It also reduces the problem of vanishing gradient, strengthening the feature propagation and effective reuse of features.

InceptionResNetV2

InceptionResNetV2 is a new and more accurate model based on CNN architecture validated accuracy on the ImageNet database with more than a million images. It is a variation of the InceptionV3 (*Szegedy et al., 2016*) and has 164 layers that use residual connections in the Inception architecture.

Model architecture

A convolutional neural network is a series of deep networks where consecutive layers receive input from the layers before them and pass output to the next layer; this continues until the final layer. The amount of deep networks is proportional to the prediction's strength and efficiency, meaning models with deeper networks tend to perform better than models with external networks (*Szegedy et al., 2016*). The reliability of the model also depends on the degree of approximation, where deeper abstractions would describe the various oriented boundaries of the image, medium abstraction levels may describe the portions of an image, and larger object parts and the object as a whole are represented in

higher levels. In this research, we experiment on deep neural networks' capabilities in the automatic detection of cases in chest pathology.

For this research, three deep convolutional neural networks are experimented on. Figure 1 depicts our proposed framework, representing the conceptual process of different deep learning model architectures used in this research. TensorFlow, which Google develops, is the leading deep learning framework used throughout the development process to create, train, validate, fine-tune and test different architectures. Apart from TensorFlow, Pandas has been used for data analysis and matplotlib for creating a visualization. The training data is first loaded into the workspace and the corresponding labels using panda's data frame. There are a total of 14 pathological disease labels that the entire dataset has been based on.

Figure 2 shows the gender and age-wise distribution of diseases. It shows several patients of each gender concerning age. It can be observed that Cardiomegaly affects people starting around the age of 10 upwards but most common between the age of 35 to 60, with the median age around 50. It can also be seen that the male gender is more affected than the female by most of the diseases.

EXPERIMENTAL STUDY AND ANALYSIS

Dataset preparation and preprocessing

This analysis uses a detailed state-of-the-art datasetcollected from *Wang et al. (2019)*. A total of 112,120 frontal-view chest X-rays (CXRs) images collected from 30,805 unique patients are included in the dataset. All the images were labeled with one of the 14 pathological diseases and classified as "No Findings" for those with no disease. The Hernia class eventually withdrew from the label, leaving 13 main classes because the images were smaller than the others. Some examples and their corresponding annotations in Chest X-ray14 (*Wang et al., 2019*) are shown in Fig. 3.

One of the significant problems faced with this research is that the classes' images are not enough for proper training of our models; hence, to get effective results with high accuracy, we generalized the data by considering data augmentation. Image augmentation provides training images by alternate processing or assembly methods of many systems, such as random rotation, turns, shear, flips, zooming, etc.

Next, as part of the data pre-processing, a folder was created for each of the class labels, and each folder was further subdivided into training and validation sets. The training and validation folders contained the class label as a sub-folder, enabling the model to read the data accurately. Before the model's training began, the training and validation datasets were reshaped to (150, 150, 3), and they were normalized, then image augmentation was applied with a certain degree of rotation on the training dataset. Finally, using one-hot encoding, the class labels were encoded, and the array was reshaped to (128, 128, 3), which were stored as pickle objects. It was not easy to recognize any patterns that divided the images into their different groups when these images were observed.

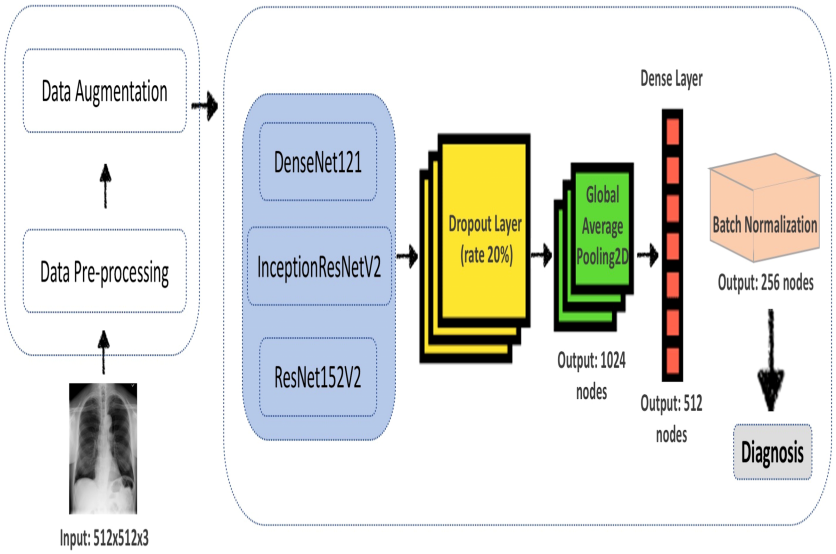

Figure 1 **Overview of the proposed work framework with three deep convolutional neural networks models.**

Full-size DOI: 10.7717/peerjcs.495/fig-1

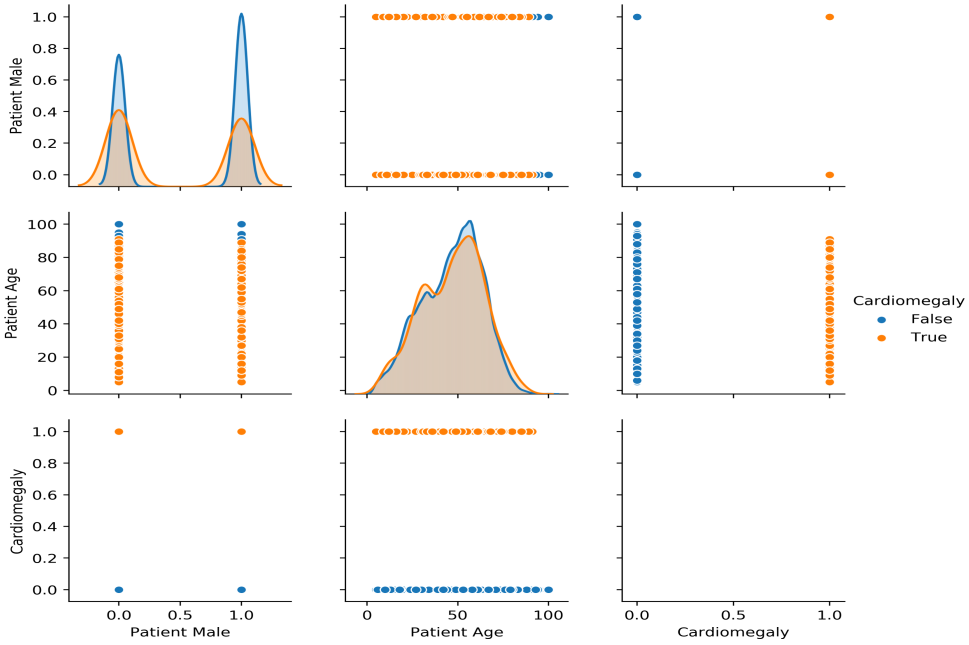

Figure 2 **Gender and Age wise distribution of chest related diseases.**

Full-size DOI: 10.7717/peerjcs.495/fig-2

Training methodology

This section explains the training procedure undertaken to train different deep learning model architectures. TensorFlow is the leading deep learning framework used throughout the development process to create, train, validate, fine-tune and test different architectures.

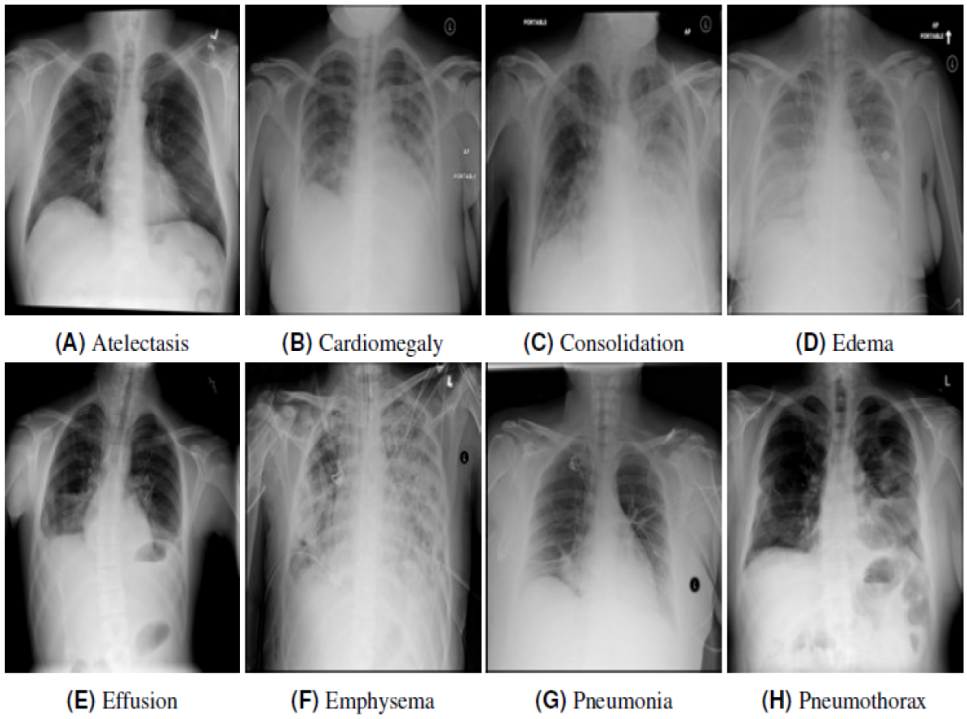

Figure 3 Chest X-rays categories pathology with related labels in Chest X-ray14. Each image is labelled with one pathology. (A) Atelectasis, (B) cardiomegaly, (C) consolidation, (D) edema, (E) effusion, (F) emphysema, (G) pneumonia, (H) phenothorax.

Full-size DOI: 10.7717/peerjcs.495/fig-3

Apart from TensorFlow, Pandas has been used for data analysis and matplotlib for creating a visualization. The training data is first loaded into the workspace and the corresponding labels using panda's data frame. There are a total of 14 pathological disease labels that the entire dataset has been based on.

Training process

1. The model of choice, as seen in Fig. 1, is first loaded into a workspace with a pre-trained weight set to the 'ImageNet' dataset. To add the classifying layer to this model's head, a GlobalAveragePooling2D layer is added at the end. Then the output of this layer is flattened, and added to it is another dense layer of 512 activation units using 'ReLU' as the activation function, finally. An output layer of 14 activation units is added to complete the model corresponding to the 14 labels with activation function as 'sigmoid.'
2. To add image augmentation to the dataset, the following *ImagDataGenerator* has been added:

```
core_idg = ImageDataGenerator(rescale=1 / 255,
 samplewise_center=True,
 samplewise_std_normalization=True,
 horizontal_flip=True,
 vertical_flip=False,
```

height_shift_range=0.05,
width_shift_range=0.1,
rotation_range=10,
shear_range=0.1,
fill_mode='nearest',
zoom_range=0.15)

This mainly rescales the images and horizontal flips the images and few other functionalities like shift range, zooming, rotation, etc.

3. To compile the model, '*binary_crossentropy*' has been used as a loss function and 'Adam' as the main optimizing algorithm with 'accuracy' as a metric.
4. The training starts with a learning rate of 0.001 but is eventually lowered using *ReduceLROnPlateau* scheduler till it reaches 0.00001.
5. The model is initially trained with all the layers of the pre-trained model set to freeze for 30-40 epochs, then to further fine-tune the model, a second training phase is initiated with the same number of epochs, but this time is training all the layers of the model including the pre-trained weights from *ImageNet'*. This further helps to improve our accuracy and transforms our model to generalize better on our X-ray dataset.

RESULTS

In the present study, we investigate three deep convolutional neural network models (DenseNet121, InceptionResNetV2, and ResNet152V2) to design the predictive model. Later, we compare the first two models to show competitiveness compared to another model called ResNet152V2. Thus, we later focused our analysis based on DenseNet121 and InceptionResNetV2 models designed and tasked with classifying X-ray images of chest-related diseases for the final prediction. Every neural network has a $150 \times 150 \times 3$ input structure. All three networks are built either with or without available data for each of the simulations' models. We hardly have used orientation parameters to increase data, which produces data through a rotation of 10-degree angles of the images. The data was distributed as 80% and 20% for training and testing, respectively.

The hyperparameters of the models are learning rate (0.001), batch size (32), Softmax and ReLU, respectively, and epochs (40) for DenseNet121 and epochs (30) for InceptionResNetV2.

The predictive model is trained by the state-of-the-art training datasets collected by *Wang et al. (2019)*. A summary of the comparison between actual and predicted classes is provided in Table 2. The accuracy between the actual and predicted classes shown in Table 2 presents the relevance in evaluating classification performance. It shows the result's high accuracy and minimizes the errors in between predictive and actual data sets.

The loss is a quantitative measure of how far from the actual output the prediction made; this means that models with little losses make more accurate predictions than the models with large losses. The loss is calculated for both the training and validation to see how well the model is working and ensure no problems of under-fitting or over-fitting.

An epoch is a term in deep learning to show the number of times the whole training dataset is utilized once to update the weights (*Albahli, 2020*). It is chosen by the network

Table 2 A comparison of actual and predicted accuracies of DenseNet121 and InceptionResNetV2.

Pathology	Model	Actual	Predicted
Atelectasis	DenseNet121	20.21%	22.10%
	InceptionResNetV2	20.51%	21.57%
Cardiomegaly	DenseNet121	4.79%	5.71%
	InceptionResNetV2	6.15%	5.63%
Consolidation	DenseNet121	9.86%	8.56%
	InceptionResNetV2	8.30%	8.88%
Edema	DenseNet121	4.49%	4.39%
	InceptionResNetV2	4.10%	4.67%
Effusion	DenseNet121	25.88%	28.07%
	InceptionResNetV2	26.07%	25.39%
Emphysema	DenseNet121	4.79%	3.59%
	InceptionResNetV2	4.49%	4.04%
Fibrosis	DenseNet121	3.03%	2.56%
	InceptionResNetV2	2.44%	3.26%
Infiltration	DenseNet121	40.14%	37.92%
	InceptionResNetV2	40.53%	39.67%
Mass	DenseNet121	12.30%	11.13%
	InceptionResNetV2	13.28%	11.30%
Nodule	DenseNet121	11.91%	11.28%
	InceptionResNetV2	12.70%	12.21%
Pleural_Thickening	DenseNet121	7.23%	6.34%
	InceptionResNetV2	7.03%	6.20%
Pneumonia	DenseNet121	3.22%	2.56%
	InceptionResNetV2	2.83%	2.57%
Pneumothorax	DenseNet121	8.89%	8.84%
	InceptionResNetV2	9.77%	9.83%

Table 3 The proposed models and comparisons of their results.

Models	Train_loss	Train_accuracy	Val_loss	Val_accuracy	ROC-AUC score
DenseNet121	0.2596	0.4238	0.2645	0.4043	0.793
InceptionResNetV2	0.2446	0.4547	0.2641	0.4102	0.801
ResNet152V2	0.2827	0.3686	0.2821	0.3760	0.751

designer so that the loss is minimal and does not increase, and since loss and accuracy are inversely proportional, at the point where loss is minimum, the accuracy will be at its highest point. We set our epochs to 40 because both models reached a stable low loss when the epochs were less than or equal to 40.

Table 3 shows the comparison of the training and validation for both loss and accuracy of all the three convolutional neural networks that were initially used in this research. InceptionResNetV2 model got an average Area Under the Receiver Operating Characteristic Curve (ROC-AUC) score of 0.801 when detecting the pathological diseases.

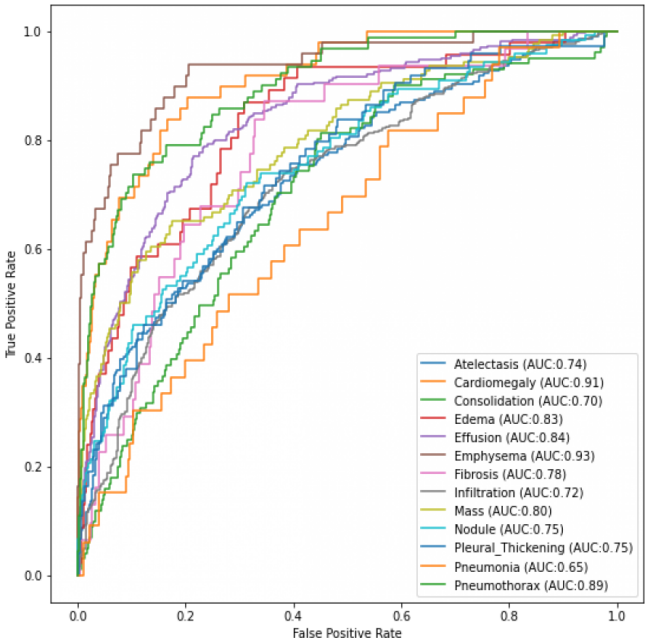

Figure 4 **ROC AUC scores for DenseNet121 of the 13th classes of chest-related diseases.**
Full-size DOI: 10.7717/peerjcs.495/fig-4

Figures 4 and 5 shows the ROC-AUC graphs for the DenseNet121 and Inception-ResNetV2 respectively. The curves have two parameters plotted: True Positive Rate (TPR) and False Positive Rate (FPR). When we compare the two models, it is seen that InceptionResNetV2 got a higher score for the number of classes predicted. This makes the InceptionResNetV2 better than the other two deep learning models used in this research, as seen in Table 3. It can also be seen that the ResNet152V2 had the lowest performance when compared to the other two in terms of training and validation accuracies when data augmentation was used. However, in terms of the loss, Table 3 shows the training and validation losses instead of the training and validation accuracy.

From the results obtained, it can be seen that InceptionResNetV2 and DenseNet121 are more efficient for the classification task using data augmentation. It can also be seen that InceptionResetV2 and DenseNet121 have an 80% improvement in the AUC score over the ResNet152V2 model. There is also a small gap between the training and validation results regarding accuracy and loss. As more data is gotten, there is a very high possibility of the current performing better, and different models could also give highly effective results.

To sum up, Table 4 shows a comparison between our DenseNet121 and InceptionResNetV2 model with similar researches done in the field of chest pathology using deep learning. It can be seen that of the 14 cardiothoracic diseases that were tested for, our models got a much higher accuracy in 13 of them, except for Consolidation. It should be noted that *Gundel et al. (2019)* trained their model with 180,000 images from the PLCO dataset (*Ho & Gwak, 2019*) as extra training data.

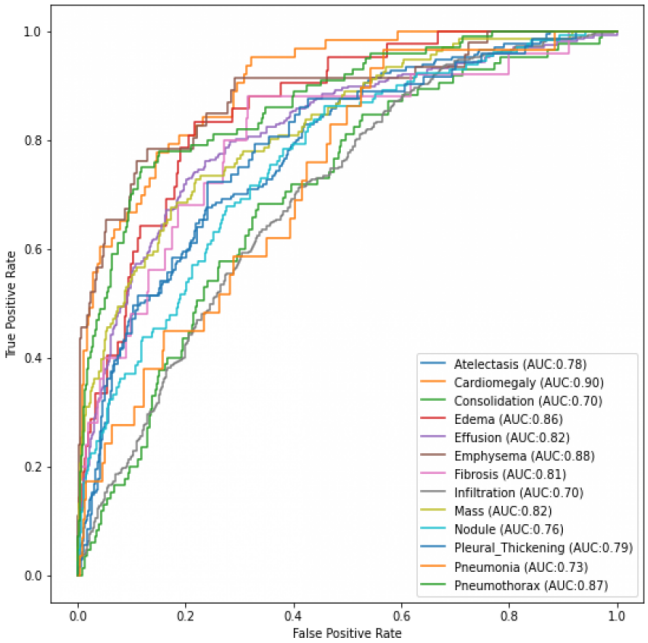

Figure 5 ROC AUC scores for InceptionResNetV2 of the 13th classes of chest-related diseases.
Full-size DOI: 10.7717/peerjcs.495/fig-5

Table 4 AUC scores comparisons of the 13th chest-related diseases.

Pathology	DensNet	InceptionResNetV2	Wang et al.	*Yao et al. (2017)*	*Wang & Xia (2018)*	Gundel et al.
Atelectasis	0.74	0.78	0.71	0.77	0.74	0.76
Cardiomegaly	0.91	0.90	0.80	0.90	0.87	0.88
Consolidation	0.70	0.70	0.70	0.78	0.72	0.74
Edema	0.83	0.86	0.83	0.88	0.83	0.83
Effusion	0.84	0.82	0.78	0.85	0.81	0.82
Emphysema	0.93	0.88	0.81	0.82	0.82	0.89
Fibrosis	0.78	0.81	0.76	0.76	0.80	0.80
Infiltration	0.72	0.70	0.60	0.69	0.67	0.70
Mass	0.80	0.82	0.70	0.79	0.78	0.82
Nodule	0.75	0.76	0.67	0.71	0.69	0.75
Pleural_Thickening	0.75	0.79	0.70	0.76	0.75	0.76
Pneumonia	0.65	0.73	0.63	0.71	0.69	0.73
Pneumothorax	0.89	0.87	0.80	0.84	0.81	0.84

Limitation of proposed framework

In DenseNet121, we utilize L-1 total layers for replacing the convolution operation with a CNN layer and a transformation up average pooling operation. An extrapolated convolution layer uses the previous function to maps in the next pooling layer. To form a new dense block's input, the input images feature maps are concatenated with the skip relation ones. Since the fully connected layers raise the dynamic range of the feature maps, exponential growth in the range of features would have been too storage-intensive,

particularly for the pre-softmax layer's full resolution features. The proposed study's order limitation is the lack of training parameters in InceptionResNetV2, which is tackle by adopting transfer learning.

CONCLUSION

Diagnosis of chest-related diseases from X-ray images is a challenging task that requires special attention from the research community because of the rising rate at which the diseases are contracted. However, there are a few types of research done related to chest-related diseases. All the previous research has been done in chest-related diseases centered around only one class disease, but research is multi-class and accurately investigated 14 classes of chest-related diseases with reliable results. This research used three deep neural network architectures; DenseNet121, InceptionResNetV2, and ResNet152V2, and we were able to get a beneficial result in detecting chest-related diseases from X-ray images. InceptionResNetV2 got the highest ROC-AUC score in the three models we used for this research. The major limitation of this research is that all the images used for both training and validation are images of the chest's frontal view. However, specialists say literal view yields more accurate results. Hence, for future works, researches will be done using literal views of the chest X-rays, and the results will be compared.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The Deanship of Scientific Research, Qassim University funded the publication of this project. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
The Deanship of Scientific Research, Qassim University.

Competing Interests

Valentina Emilia Balas is an Academic Editor for PeerJ.

Author Contributions

- Saleh Albahli conceived and designed the experiments, performed the experiments, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Hafiz Tayyab Rauf conceived and designed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Abdulelah Algosaibi performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Valentina Emilia Balas analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

Data is publicly available at Kaggle https://www.kaggle.com/nih-chest-xrays/data.

REFERENCES

Albahli S. 2020. A deep neural network to distinguish COVID-19 from other chest diseases using X-ray images. *Current Medical Imaging Formerly Current Medical Imaging Reviews* **17**:109–119 DOI 10.2174/1573405616666200604163954.

Ali A-R, Li J, Yang G, OShea SJ. 2020. A machine learning approach to automatic detection of irregularity in skin lesion border using dermoscopic images. *PeerJ Computer Science* **6**:e268 DOI 10.7717/peerj-cs.268.

Apostolopoulos ID, Aznaouridis SI, Tzani MA. 2020. Extracting possibly representative COVID-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases. *Journal of Medical and Biological Engineering* **40**:462–469 DOI 10.1007/s40846-020-00529-4.

Apostolopoulos ID, Mpesiana TA. 2020. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine* **43(2)**:635–640 DOI 10.1007/s13246-020-00865-4.

Brag J. 2020. Artificial intelligence in medical imaging. In: *Healthcare and artificial intelligence*. New York City: Springer International Publishing, 93–103 DOI 10.1007/978-3-030-32161-1_14.

Briganti G, Le Moine O. 2020. Artificial intelligence in medicine: today and tomorrow. *Frontiers in Medicine* **7**:27 DOI 10.3389/fmed.2020.00027.

Chandra TB, Verma K. 2019. Pneumonia detection on chest X-Ray using machine learning paradigm. In: *Proceedings of 3rd international conference on computer vision and image processing*. Springer Singapore, 21–33 DOI 10.1007/978-981-32-9088-4_3.

Chatterjee S, Dzitac S, Sen S, Rohatinovici NC, Dey N, Ashour AS, Balas VE. 2017. Hybrid modified Cuckoo Search-Neural Network in chronic kidney disease classification. In: *2017 14th international conference on engineering of modern electric systems (EMES)*. Piscataway: IEEE, 164–167.

Demir F, Sengur A, Bajaj V. 2020. Convolutional neural networks based efficient approach for classification of lung diseases. *Health Information Science and Systems* **8(1)**:1–8.

Franquet T. 2018. Imaging of community-acquired pneumonia. *Journal of Thoracic Imaging* **33(5)**:282–294 DOI 10.1097/rti.0000000000000347.

Gao J, Wang H, Shen H. 2020a. Machine learning based workload prediction in cloud computing. In: *2020 29th international conference on computer communications and networks (ICCCN)*. Piscataway: IEEE, 1–9.

Gao J, Wang H, Shen H. 2020b. Smartly handling renewable energy instability in supporting a cloud datacenter. In: *2020 IEEE international parallel and distributed processing symposium (IPDPS)*. Piscataway: IEEE, 769–778.

Gao J, Wang H, Shen H. 2020c. Task failure prediction in cloud data centers using deep learning. In: *IEEE Transactions on Services Computing*. Piscataway: IEEE.

Gundel S, Grbic S, Georgescu B, Liu S, Maier A, Comaniciu D. 2019. Learning to recognize abnormalities in chest X-Rays with location-aware dense networks. In: *Progress in pattern recognition, image analysis, computer vision, and applications*. New York City: Springer International Publishing, 757–765 DOI 10.1007/978-3-030-13469-3_88.

Hashmi MF, Katiyar S, Keskar AG, Bokde ND, Geem ZW. 2020. Efficient pneumonia detection in chest Xray images using deep transfer learning. *Diagnostics* **10(6)**:1–23 DOI 10.3390/diagnostics10060417.

Ho TKK, Gwak J. 2019. Multiple feature integration for classification of thoracic disease in chest radiography. *Applied Sciences* **9(19)**:1–15 DOI 10.3390/app9194130.

Islam K, Wijewickrema S, Collins A, O'Leary S. 2020. A deep transfer learning framework for pneumonia detection from chest X-ray images. In: *Proceedings of the 15th international joint conference on computer vision, imaging and computer graphics theory and applications*. SCITEPRESS - Science and Technology Publications DOI 10.5220/0008927002860293.

Jain R, Nagrath P, Kataria G, Kaushik VS, Hemanth DJ. 2020. Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning. *Measurement* **165**:108046 DOI 10.1016/j.measurement.2020.108046.

Jin B, Cruz L, Goncalves N. 2020. Deep facial diagnosis: deep transfer learning from face recognition to facial diagnosis. *IEEE Access* **8**:123649–123661 DOI 10.1109/ACCESS.2020.3005687.

Kao E-F, Lin W-C, Hsu J-S, Chou M-C, Jaw T-S, Liu G-C. 2011. A computerized method for automated identification of erect posteroanterior and supine anteroposterior chest radiographs. *Physics in Medicine and Biology* **56(24)**:7737–7753 DOI 10.1088/0031-9155/56/24/004.

Ke Q, Zhang J, Wei W, Polap D, Wozniak M, Kosmider L, Damavsevuicius R. 2019. A neuro-heuristic approach for recognition of lung diseases from X-ray images. *Expert Systems with Applications* **126**:218–232 DOI 10.1016/j.eswa.2019.01.060.

Kulkarni S, Seneviratne N, Baig MS, Khan AHA. 2020. Artificial intelligence in medicine: where are we now? *Academic Radiology* **27(1)**:62–70 DOI 10.1016/j.acra.2019.10.001.

Meraj T, Rauf HT, Zahoor S, Hassan A, Lali MI, Ali L, Bukhari SAC, Shoaib U. 2019. Lung nodules detection using semantic segmentation and classification with optimal features. *Neural Computing and Applications*. Epub ahead of print 2020 11 May.

Militante SV, Sibbaluca BG. 2020. Pneumonia detection using convolutional neural networks. *International Journal of Scientific & Technology Research* **9(04)**:1332–1337.

Mongan J, Moy L, Kahn CE. 2020. Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers. *Radiology: Artificial Intelligence* **2(2)**:e200029 DOI 10.1148/ryai.2020200029.

Oh Y, Park S, Ye JC. 2020. Deep learning covid-19 features on cxr using limited training data sets. *IEEE Transactions on Medical Imaging* **39(8)**:2688–2700 DOI 10.1109/TMI.2020.2993291.

Panwar H, Gupta P, Siddiqui MK, Morales-Menendez R, Singh V. 2020. Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. *Chaos, Solitons & Fractals* **138**:109944 DOI 10.1016/j.chaos.2020.109944.

Pereira RM, Bertolini D, Teixeira LO, Silla CN, Costa YM. 2020. COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. *Computer Methods and Programs in Biomedicine* **194**:105532 DOI 10.1016/j.cmpb.2020.105532.

Rajaraman S, Antani SK. 2020. Training deep learning algorithms with weakly labeled pneumonia chest X-ray data for COVID-19 detection. *medrXiv* DOI 10.1101/2020.05.04.20090803.

Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz CP, Patel BN, Yeom KW, Shpanskaya K, Blankenberg FG, Seekins J, Amrhein TJ, Mong DA, Halabi SS, Zucker EJ, Ng AY, Lungren MP. 2018. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLOS Medicine* **15(11)**:e1002686 DOI 10.1371/journal.pmed.1002686.

Rajpurkar P, Park A, Irvin J, Chute C, Bereket M, Mastrodicasa D, Langlotz CP, Lungren MP, Ng AY, Patel BN. 2020. AppendiXNet: deep learning for diagnosis of appendicitis from a small dataset of CT exams using video pretraining. *Scientific Reports* **10**:3958 DOI 10.1038/s41598-020-61055-6.

Rauf HT, Malik S, Shoaib U, Irfan MN, Lali MI. 2020. Adaptive inertia weight Bat algorithm with Sugeno-Function fuzzy search. *Applied Soft Computing* **90**:106159 DOI 10.1016/j.asoc.2020.106159.

Rauf HT, Saleem BA, Lali MIU, Khan MA, Sharif M, Bukhari SAC. 2019. A citrus fruits and leaves dataset for detection and classification of citrus diseases through machine learning. *Data in Brief* **26**:104340 DOI 10.1016/j.dib.2019.104340.

Rustam Z, Yuda RP, Alatas H, Aroef C. 2020. Pulmonary rontgen classification to detect pneumonia disease using convolutional neural networks. *Telkomnika* **18(3)**:1522–1528 DOI 10.12928/telkomnika.v18i3.14839.

Sahlol AT, Abd Elaziz M, Tariq Jamal A, Damavsevuicius R, Farouk Hassan O. 2020. A novel method for detection of tuberculosis in chest radiographs using artificial ecosystem-based optimisation of deep neural network features. *Symmetry* **12(7)**:1146 DOI 10.3390/sym12071146.

Salehinejad H, Colak E, Dowdell T, Barfett J, Valaee S. 2018a. Synthesizing chest x-ray pathology for training deep convolutional neural networks. *IEEE Transactions on Medical Imaging* **38(5)**:1197–1206.

Salehinejad H, Valaee S, Dowdell T, Colak E, Barfett J. 2018b. Generalization of deep neural networks for chest pathology classification in X-Rays using generative adversarial networks. In: *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. Piscataway: IEEE, DOI 10.1109/icassp.2018.8461430.

Salehinejad H, Valaee S, Dowdell T, Colak E, Barfett J. 2018c. Generalization of deep neural networks for chest pathology classification in x-rays using generative adversarial networks. In: *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. 990–994.

Sivasamy J, Subashini T. 2020. Classification and predictions of lung diseases from Chest X-rays using MobileNet. *International Journal of Analytical and Experimental Modal Analysis* **12(3)**:665–672.

Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. 2016. Rethinking the inception architecture for computer vision. In: *2016 IEEE conference on computer vision and pattern recognition (CVPR)*. IEEE DOI 10.1109/cvpr.2016.308.

Tabares-Soto R, Orozco-Arias S, Romero-Cano V, Bucheli VS, Rodríguez-Sotelo JL, Jiménez-Varón CF. 2020. A comparative study of machine learning and deep learning algorithms to classify cancer types based on microarray gene expression data. *PeerJ Computer Science* **6**:e270 DOI 10.7717/peerj-cs.270.

Toğaçar M, Ergen B, Cömert Z. 2020. Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks. *Biocybernetics and Biomedical Engineering* **40(1)**:23–39 DOI 10.1016/j.bbe.2019.11.004.

Wang H, Xia Y. 2018. Chestnet: a deep neural network for classification of thoracic diseases on chest radiography. ArXiv preprint. arXiv:1807.03058.

Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. 2017. ChestX-Ray8: hospital-scale chest X-Ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: *2017 IEEE conference on computer vision and pattern recognition (CVPR)*. IEEE, 1 DOI 10.1109/cvpr.2017.369.

Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. 2019. ChestX-ray: hospital-scale Chest X-ray database and benchmarks on weakly supervised classification and localization of common thorax diseases. In: *Deep learning and convolutional neural networks for medical imaging and clinical informatics*. Hawaii: Springer International Publishing, 369–392 DOI 10.1007/978-3-030-13969-8_18.

Yao L, Poblenz E, Dagunts D, Covington B, Bernard D, Lyman K. 2017. Learning to diagnose from scratch by exploiting dependencies among labels. ArXiv preprint. arXiv:1710.10501.

Yoo SH, Geng H, Chiu TL, Yu SK, Cho DC, Heo J, Choi MS, Choi IH, Cung Van C, Nhung NV, Byung JM, Ho L. 2020. Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest X-ray imaging. *Frontiers in Medicine* **7**:427 DOI 10.3389/fmed.2020.00427.

Zhang J, Xie Y, Li Y, Shen C, Xia Y. 2020. Covid-19 screening on chest x-ray images using deep learning based anomaly detection. ArXiv preprint. arXiv:2003.12338.

Zhang Y, Miao S, Mansi T, Liao R. 2018. Task driven generative modeling for unsupervised domain adaptation: application to x-ray image segmentation. In: *Medical image computing and computer assisted intervention –MICCAI 2018*. Granada: Springer International Publishing, 599–607 DOI 10.1007/978-3-030-00934-2_67.