Multi-Class Chest X-Ray Classification Using CheXNet and ResNeXt

Hariprasaath Velampalayam Veerakkumaar

A20528175

Bala Yaswanta Sai Chowdary Gotti

A20526959

Jaichiddharth Rangasamy Karthigeyan

A20527281

Abstract

Chest X-ray analysis stands at the forefront of modern medical diagnostics, offering invaluable insights into a myriad of respiratory conditions. As the volume of medical imaging data continues to grow, the demand for efficient and accurate diagnostic tools becomes paramount. Today, we delve into the realm of deep learning, aiming to elevate chest X-ray analysis beyond binary classifications. Our focus is on multi-class classification, addressing a spectrum of chest disease.

1 Introduction

A vital part of contemporary medical diagnostics is the study of chest X-rays, which offer important insights into a wide range of conditions affecting the heart, lungs, and chest wall. Lung cancer, pleural effusion, pneumonia, TB, and other illnesses are among them. Chest X-ray analysis principally can be used for finding an illness, track how they grow, and assess how well therapies are working. The conventional techniques of chest X-ray analysis, which depend on radiologists' manual interpretation, are challenged by the growing volume of medical imaging data. This procedure can be laborious, prone to mistakes, and uneven, particularly in places with few resources and a dearth of radiologists with the necessary training. Consequently, reliable and efficient diagnostic systems that can automate chest X-ray analysis and boost the standard of care must be developed.

Medical imaging is one of the industries that deep learning has completely revolutionized. Artificial neural networks are used in deep learning, a branch of machine learning, to learn from massive quantities of data and carry out challenging tasks like speech synthesis, image recognition, and natural language processing. In a number of image recognition benchmarks, deep learning systems have demonstrated outstanding performance, sometimes even surpassing human accuracy levels. Given that they are able to identify complex and subtle patterns in the pictures that might be symptomatic of various illnesses, this suggests that they have the potential to be extremely useful tools for chest X-ray analysis.

The traditional method of analyzing chest X-rays is based on binary classification, which is the process of figuring out whether or not an image has a particular condition. For example, by learning to differentiate between normal and unhealthy images, a binary classifier may be taught to identify pneumonia from chest X-rays. This method does not take into consideration the complexity and diversity of chest diseases, which presents a number of difficulties. For example, a patient might be suffering from multiple diseases concurrently, like pneumonia and TB, or separate illnesses could seem identical on an X-ray, such as lung cancer and tuberculosis. A binary classifier, which can only produce a yes or no response, would disregard these cases. Furthermore, as a binary classifier can only handle one condition at a time, several classifiers would be required to cover all potential disorders, raising the system's computing complexity and expense.

We suggest using multi-class classification for chest X-ray analysis, which identifies several illnesses from a single picture, to get over these restrictions. Since this method can identify the co-occurrence and overlap of several diseases, it can offer a more thorough and precise diagnosis of the patient's condition. Because a single multi-class classifier can handle several illnesses at once, it can also minimize the number of classifiers required, simplifying the system and cutting down on computing costs.

In addition to identifying pneumonia, CheXNet can identify more diseases in the dataset such as cardiomegaly, consolidation, and effusion. Another robust image recognition model that can be customized for a number of uses, such as the interpretation of chest X-rays, is called ResNeXt. The foundation of ResNeXt is a unique architecture known as ResNeXt, which adds a new dimension known as cardinality. ResNeXt's resilience and speed may be enhanced by learning more complex and diverse characteristics from the data. We compare the capabilities of CheXNet and ResNeXt for multi-class chest X-ray classification with the aim of discovering a more effective system to aid radiologists in diagnosis.

2 Problem Description

One of the most popular and significant diagnostic techniques in medicine is the study of chest X-rays. It facilitates the detection and diagnosis of a range of respiratory conditions affecting the chest region by physicians and other health care providers. Chest X-ray analysis is a useful tool for obtaining information on the health and condition of internal organs and tissues, such as the heart, lungs, and ribs, by producing pictures of these structures with a little amount of radiation. This makes it an extremely practical and non-invasive technique for treating, diagnosing, and monitoring a wide range of diseases associated to the chest.

However, there are several shortcomings and restrictions with the models that are currently being used to evaluate these images, including CheXNet. The primary purpose of these models is to perform binary classifications, meaning that they are trained to provide yes/no answers on the existence or non-existence of a disease in a chest X-ray image. For instance, a binary classification model would be able to identify the presence or absence of pneumonia in a patient, but not its kind or severity. This method may be effective in certain straightforward situations, but it falls short in capturing the variety and complexity of illnesses that can impact the chest.

In real life, there are multiple kinds of chest conditions, and some of them share symptoms and looks on X-rays. For example, unusual shadows or opacities on the lungs can be caused by lung diseases such as TB, lung cancer, and pulmonary edema, but they differ in their origins, therapies, and prognoses. Additionally, a patient may be suffering from many diseases concurrently, which can make diagnosis and therapy more difficult. By concentrating solely on binary classifications, current models fail to account for this complexity and diversity, which can result in oversimplified or incorrect diagnoses.

Furthermore, as medical imaging technology develops, the shortcomings of these binary classification approaches become increasingly apparent. There is a need for increasingly complicated and advanced diagnostic models that can manage illnesses across various groups and categories due to the growing volume and complexity of medical imaging data. This calls for a change from binary classifications to multi-class classifications models that can recognize and characterize the entire spectrum of potential chest disorders.

The goal of raising the standard of patient care and diagnostic precision is what drives the development of such all-encompassing solutions. Because multi-class classification algorithms allow for the early finding and treating of several diseases from one single X-ray image, they could improve patient outcomes. A multi-class classification model, for instance, could be able to recognize co-occurring conditions like heart failure or effusion in addition to pneumonia, as well as the kind, severity, and location of the illness. This might make it easier for medical professionals to give their patients more individualized and efficient treatment programs.

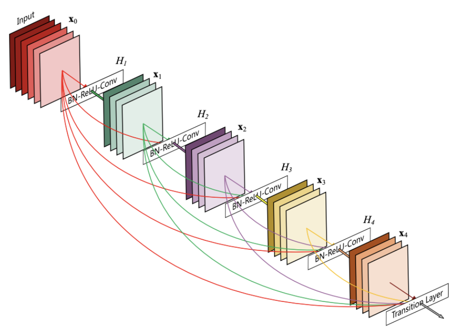

Figure 1: A 5-layer dense block[2]

3 CheXNet

Deep learning algorithms like CheXNet[1] have completely changed the area of medical imaging, especially when it comes to analyzing chest X-rays. It is a model built on top of a 121-layer DenseNet[2], a kind of Convolutional Neural Network (CNN).

DenseNet, an acronym for Densely connected Convolutional Networks, is a network design in which all layers are feed-forwardly linked to all other layers. This structure of dense connection improves gradients and information flow within the network, making extremely deep network optimization possible.

The ChestX-ray14, dataset served as the training set for CheXNet. This dataset consists of several X-ray pictures of the chest, each tagged with up to 14 distinct diseases. CheXNet is able to learn a broad variety of characteristics and patterns related with different chest ailments because of the utilization of such an extensive dataset.

CheXNet's ability to identify pneumonia, a prevalent and potentially dangerous lung illness, is one of its main advantages. Indeed, studies have shown that CheXNet can diagnose pneumonia more accurately than radiologists[1], demonstrating the strength and promise of this deep learning model.

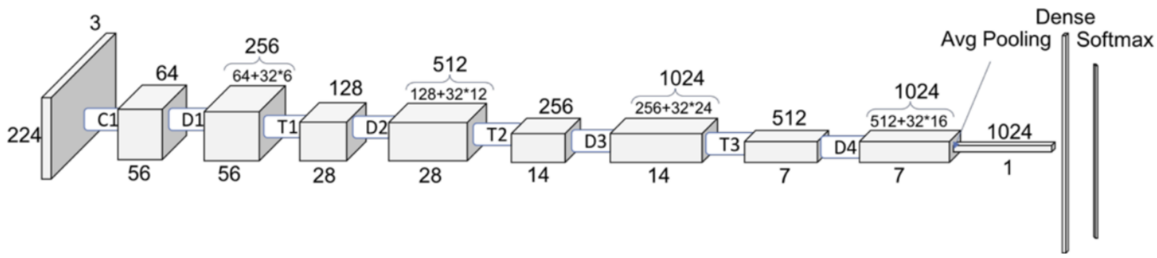

Figure 2: Simple architecture of DenseNet-121 [2]

The measurements under each volume in Figure 2 show the width and depth measurements. The feature map dimension is shown by the numbers at the top.

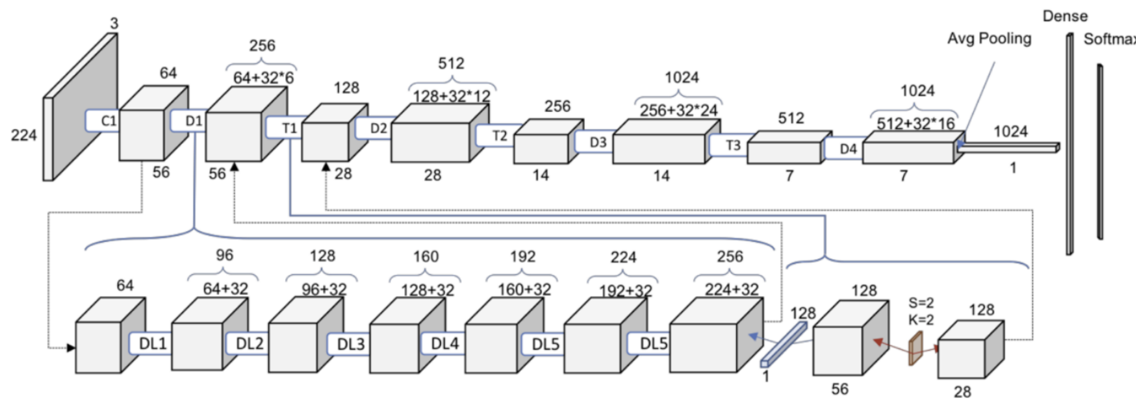

Figure 3: View of DenseNet-121 from a deeper perspective. Transition Block and Dense Block [2]

Each layer adds these 32 fresh feature maps to the preceding volume. 256 layers later following 64 layers. Additionally, Transition Block uses 128 filters to operate as a 1x1 convolution. Subsequently, a 2x2 pooling with a stride of 2 was performed, which led to a 50% reduction in both the volume and feature map count. This demonstrates in detail how:

- The volume inside a Dense Block never changes and
- Following each Transition Block, both the volume and the feature maps are cut in half.

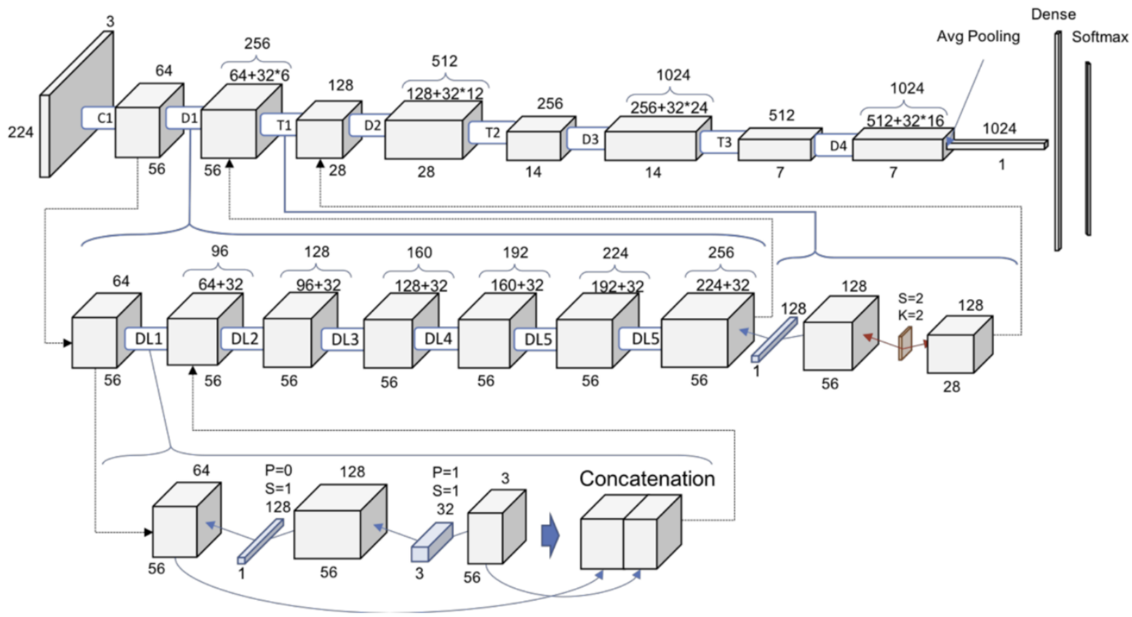

Figure 4: Two levels deeper. Complete schematic diagram of DenseNet-121 [2]

How the action of adding 32 times the total number of layers is really achieved is shown in the first Dense Layer that comprises the first Dense Block. In order to decrease the size of the feature maps, we execute a 1x1 convolution on 128 filters. After that, we perform a more costly 3x3 convolution with the selected 32 feature maps of growth rate. In order to contribute fresh data to the network's collective knowledge, the input volume as well as the outcomes of the two procedures are concatenated.

CheXNet's architecture is built to manage the complexity and variability of chest X-ray pictures. The network's 121 layers enable it to learn a hierarchy of characteristics, ranging from basic textures and edges to more intricate patterns and structures. Because of its depth and DenseNet's dense connection, CheXNet is able to pick up on details and patterns that other networks with shallower or less dense connections could miss.

4 ResNeXt

Image recognition has advanced significantly with the help of ResNeXt, an efficient deep learning model. It is an expansion of the ResNet model, which was presented by Xie et al. and is well-known for its residual connections that aid in the training of deeper networks[3].

The primary innovation of ResNeXt is the addition of a new dimension to the traditional dimensions. The new dimension "cardinality," or the size of the set of transformations[3]. The performance of the model is greatly influenced by this cardinality dimension, which enables more intricate and varied feature learning.

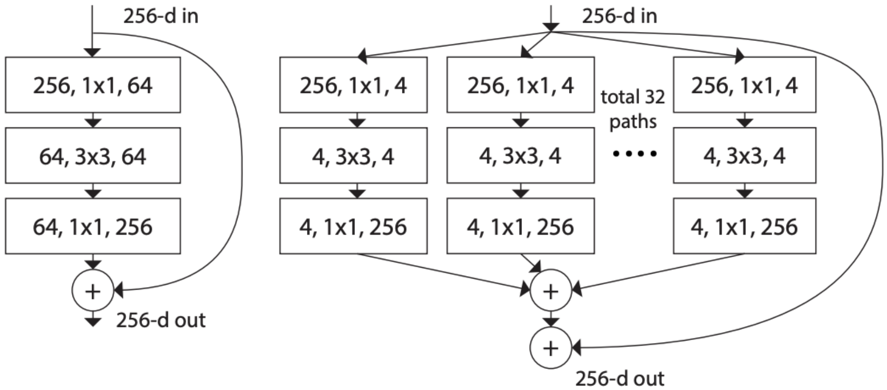

Figure 5: Left: A ResNet block. Right: A ResNeXt block of cardinality:32 [3]

ResNeXt's design is homogenous, which makes it simpler to implement and modify because it is made up of blocks of the same kind. It simplifies the design process by reducing the number of hyperparameters needed for traditional ResNet.

ResNeXt is a powerful and flexible model that increases ResNet's functionality. It enables more sophisticated feature learning without significantly raising the model's complexity by adding the concept of cardinality. Because of this, ResNeXt is a useful tool for jobs like image recognition, where learning a large variety of characteristics is important.

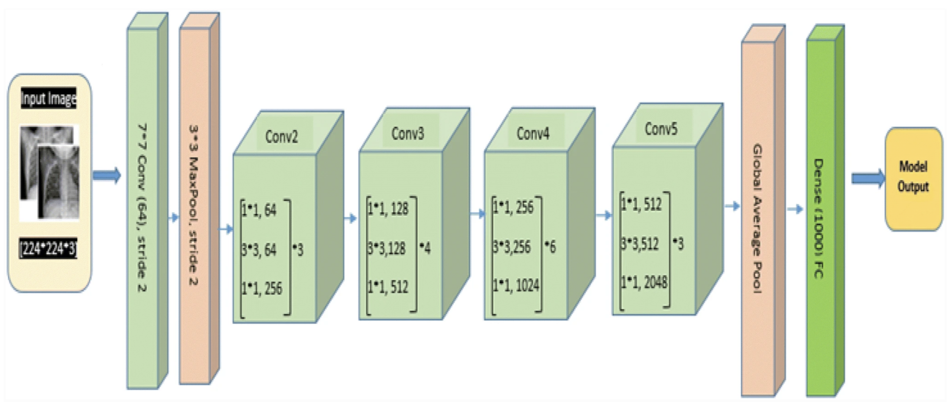

Figure 6: ResNeXt Architectture for Chest X-ray [4]

4.1 Cardinality Block: A Deep Dive

4.1.1 Design Philosophy

ResNeXt's cardinality block, which has 32 parallel pathways in one block, embodies a new design concept. Three convolutional layers with different kernel sizes are stacked in the cardinality block, in contrast to conventional convolutional blocks. Because the model may investigate several receptive fields at once, this approach encourages feature diversity.

4.1.2 Parallel Processing

The core of the cardinality block is its 32 identical elements, which allow it to process information in parallel. A series of 1x1, 3x3, and 1x1 convolutions are followed by each element. In feature extraction, the 3x3 convolution adds a spatial hierarchy, whereas the 1x1 convolutions are used for dimensionality reduction and expansion. The model's ability to recognise complex patterns and relationships in the input data is improved by this parallel processing technique.

4.1.3 Adaptive Feature Learning

In feature learning, the cardinality block demonstrates flexibility. The ability to aggregate data from several paths allows the model to dynamically adapt its representation to the intricacies found in the data. This flexibility is especially helpful in medical picture classification tasks because different conditions can have dramatically different visual patterns suggestive of those disorders.

4.2 Model Initialization and Training

4.2.1 Initial Layers and Feature Extraction

A 7x7 convolutional layer is used for model initialization, and it is followed by ReLU activation and batch normalization. For efficient feature extraction, the foundation is laid by this first layer. To ensure that the model can process input hierarchically, the spatial dimensions are downsampled using max-pooling.

4.2.2 Residual Blocks and Training Strategies

After the first levels, the architecture consists of residual blocks, each containing a cardinality block. The optimisation process is aided by these residual blocks, which make residual mapping learning easier. To improve the model's capacity to generalise to unknown variances in the input data, data augmentation techniques such random rotations and flips are used during training.

4.2.3 Hyperparameter Choices

Iterative optimisation of the model's parameters occurs during the training process. The learning rate and momentum are crucial hyperparameters that determine the model's convergence and generalisation capabilities. Strategic testing is required to fine-tune these settings and ensure optimal training performance.

5 Dataset

5.1 National Institutes of Health Chest X-Ray Dataset

The NIH Chest X-ray Dataset consists of 112,120 X-ray pictures from 30,805 distinct individuals that include illness classifications. The authors text-mined illness categories from the resembling radiological reports with help of NLP to get these labels. For weakly-supervised learning, the labels should be more than 90% correct. There is no public access to the original radiological reports.

- Dataset Used : NIH Chest X-Ray Dataset
- Source : Kaggle (https://www.kaggle.com/datasets/nih-chest-xrays/dat)
- Dataset Size : 112,120 total images with size 1024 x 1024

- No. of Classes : 15
- Classes Considered : 3 (Effusion, Consolidation, Cardiomegaly)
- Size of Dataset used : 1100 images per class

5.1.1 Why this Dataset?

- The NIH Chest X-ray dataset contains an extensive coverage of various chest diseases, providing a diverse and representative set of images.
- This dataset spans a spectrum of chest diseases (15 Classes), enabling a holistic approach to multi-class classification, a key requirement for our objectives.
- Ground Truth Annotations: The availability of annotated labels by NLP from reports provided by radiologists serves as invaluable ground truth for training and evaluating our models.

5.2 Classes Considered

Out of 15 available total classes including the normal class, 3 classes were chosen for this project. The three classes are Effusion, Consolidation and Cardiomegaly.

- **Effusion:** Happens when fluid accumulates in the area between the lungs and heart
- **Consolidation:** A fluid, solid, or other substance, such as pus, blood, or water, replaces the air in the lung passages.
- **Cardiomegaly:** An enlarged heart is one that is bigger than normal due to excessive stretching or abnormal thickening.

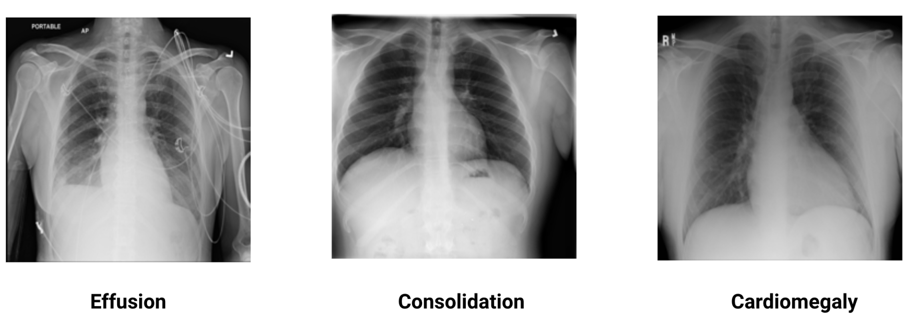

Figure 7: Sample Images

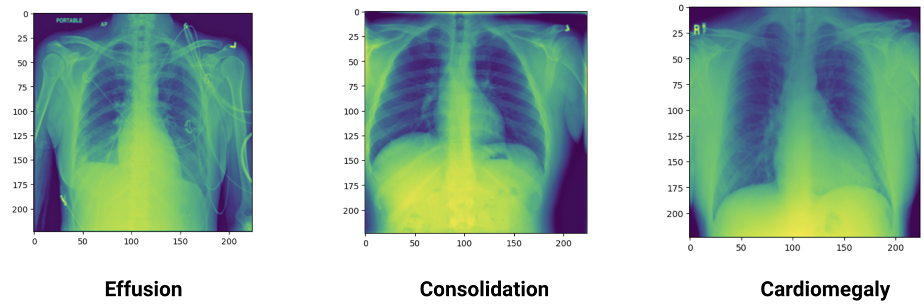

Figure 8: Sample Images after Transforming

Same images displayed using imshow() after transformation 8.

6 Results

Following the transformation of images, as well as modifying the networks to our requirements, the next step is to train and evaluate how efficient and adaptable our CheXNet and ResNeXt model would perform on the chest X-ray for classification.

6.1 CheXNet

The CheXNet model was trained for 3 class classification with the hyperparameters,

- Learning Rate: 0.0001
- Betas: (0.9,0.999)
- Weight Decay: 0.01
- Epoch: 100

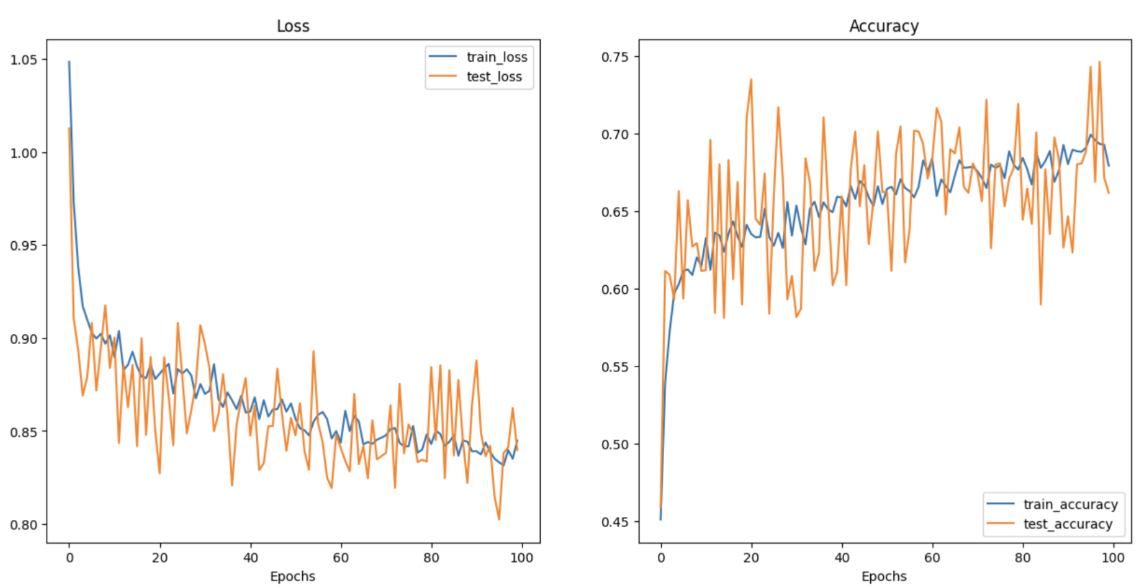

Figure 9: CheXNet Training Loss and Accuracy Graph

Number of trainable parameters: 6956931

Overall Accuracy : 73%

	Precision	Recall	F1 Score
Effusion	0.78	0.59	0.67
Consolidation	0.62	0.73	0.67
Cardiomegaly	0.80	0.88	0.83

Table 1: CheXNet Test Data Metrics

6.2 ResNeXt

The ResNeXt model was trained for 3 class classification with the hyperparameters,

- Learning Rate: 0.0001
- Betas: (0.9,0.999)
- Weight Decay: 0.01
- Epoch: 100

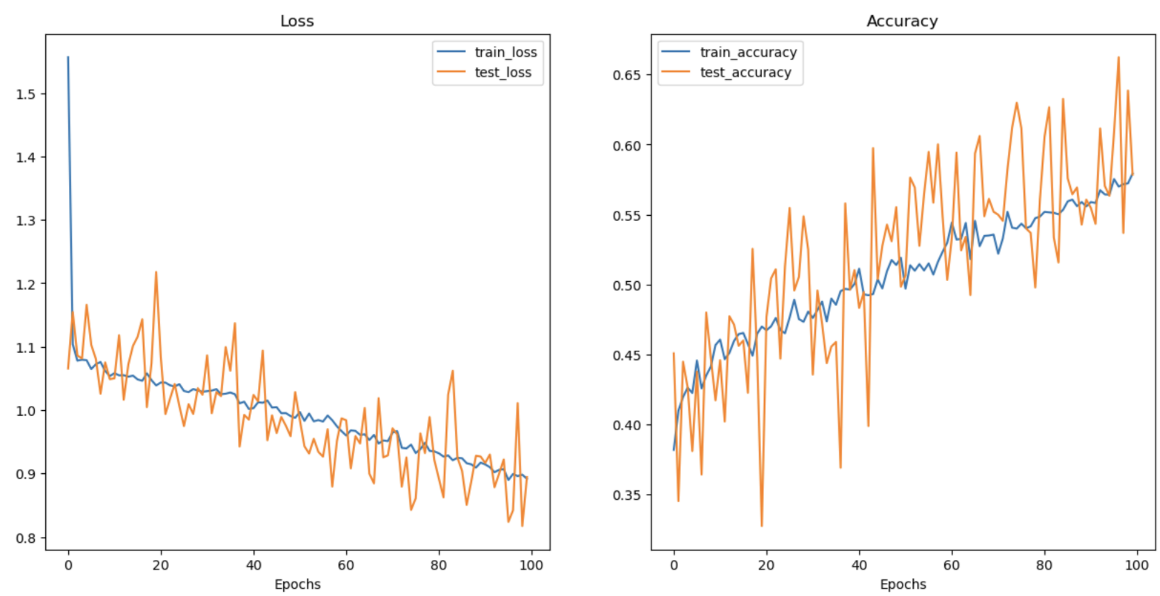

Figure 10: ResNeXt Training Loss and Accuracy Graph

Number of trainable parameters: 5166472

Overall Accuracy : 62%

	Precision	Recall	F1 Score
Effusion	0.79	0.48	0.60
Consolidation	0.47	0.83	0.60
Cardiomegaly	0.85	0.57	0.68

Table 2: ResNeXt Test Data Metrics

6.3 Performance

Both CheXNet and ResNeXt models were trained with same hyperparameters, optimizer (ADAM) and each trained for 100 epochs. The overall accuracy on the test set for CheXNet was 73% and for ResNeXt was 62%.

It can be inferred from 1 and 2 that the per class F1-Score of the classes Effusion and Consolidation was similar with both the models and for the class Cardiomegaly CheXNet model performed well with the score 83.

7 Conclusions

To sum up, the NIH dataset was used in this study to apply the ChexNet and ResNext models for multiclass picture classification. The effectiveness of these models was carefully assessed using critical metrics like precision, recall, and F1 score. The results demonstrated how well both algorithms classified various medical photos correctly across a number of classes. Notably, careful examination turned up subtle differences in how well they performed in each class. To astonishingly, the ChexNet model outperformed ResNext in every measurement criterion, including overall accuracy. Furthermore demonstrating the resilience and effectiveness of the ChexNet architecture in this medical imaging environment are random factors, such as an 11% increase in accuracy.

7.1 Future Work

A combinatorial use of transfer learning and Class Activation Maps (CAM) can be helpful in improving the accuracy of multi-class picture classification. To start, pre-train a neural network on a sizable dataset to take advantage of transfer learning and extract general properties that are applicable to various applications. Fine-tuning this pre-trained models with various hyperparameters using parameter grid search for the target multi-class classification problem, allowing it to respond to domain-specific complexities. Concurrently, integrate CAM to give interpretability by visualizing the model's attention on critical image regions. This allows ResNext and ChexNet models to better understand complex class boundaries for targeted refining during fine-tuning.

References

[1] Rajpurkar, Pranav, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding et al. "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." arXiv preprint arXiv:1711.05225 (2017).

[2] Huang, Gao, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q. Weinberger. "Densely connected convolutional networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700-4708. 2017.

[3] Xie, Saining, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. "Aggregated residual transformations for deep neural networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1492-1500. 2017.

[4] Hira, S., Bai, A. Hira, S. An automatic approach based on CNN architecture to detect Covid-19 disease from chest X-ray images. Appl Intell 51, 2864–2889 (2021). https://doi.org/10.1007/s10489-020-02010-w

[5] Bell, Sean, C. Lawrence Zitnick, Kavita Bala, and Ross Girshick. "Inside-outside net: Detecting objects in context with skip pooling and recurrent neural networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2874-2883. 2016.

[6] Lakhani, Paras and Sundaram, Baskaran. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology, pp. 162326, 2017.

[7] Islam, Mohammad Tariqul, Aowal, Md Abdul, Minhaz, Ahmed Tahseen, and Ashraf, Khalid. Abnormality detection and localization in chest x-rays using deep convolutional neural networks. arXivpreprint arXiv:1705.09850, 2017.

[8] Cherian, Thomas, Mulholland, E Kim, Carlin, John B, Ostensen, Harald, Amin, Ruhul, Campo, Margaret de, Greenberg, David, Lagos, Rosanna, Lucero, Marilla, Madhi, Shabir A, et al. Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies. Bulletin of the World Health Organization, 83(5):353–359, 2005.

[9] Grewal, Monika, Srivastava, Muktabh Mayank, Kumar, Pulkit, and Varadarajan, Srikrishna. Radnet: Radiologist level accuracy using deep learning for hemorrhage detection in ct scans. arXiv preprint arXiv:1710.04934, 2017.

[10] Huang, Peng, Park, Seyoun, Yan, Rongkai, Lee, Junghoon, Chu, Linda C, Lin, Cheng T, Hussien, Amira, Rathmell, Joshua, Thomas, Brett, Chen, Chen, et al. Added value of computer aided ct image features for early lung cancer diagnosis with small pulmonary nodules: A matched case-control study. Radiology, pp. 162725, 2017.

[11] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. Imagenet large scale visual recognition challenge. IJCV.

[12] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell. Caffe: Convolutional architecture for fast feature embedding. arXiv:1408.5093, 2014.

[13] Khan, Ejaz, Muhammad Zia Ur Rehman, Fawad Ahmed, Faisal Abdulaziz Alfouzan, Nouf M. Alzahrani, and Jawad Ahmad. "Chest X-ray classification for the detection of COVID-19 using deep learning techniques." Sensors 22, no. 3 (2022): 1211.

[14] Jaiswal, Amit Kumar, Prayag Tiwari, Sachin Kumar, Deepak Gupta, Ashish Khanna, and Joel JPC Rodrigues. "Identifying pneumonia in chest X-rays: A deep learning approach." Measurement 145 (2019): 511-518.

[15] Avola, Danilo, Andrea Bacciu, Luigi Cinque, Alessio Fagioli, Marco Raoul Marini, and Riccardo Taiello. "Study on transfer learning capabilities for pneumonia classification in chest-x-rays images." Computer Methods and Programs in Biomedicine 221 (2022): 106833.

[16] Gabruseva, Tatiana, Dmytro Poplavskiy, and Alexandr Kalinin. "Deep learning for automatic pneumonia detection." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp. 350-351. 2020.

[17] Alshmrani, Goram Mufarah M., Qiang Ni, Richard Jiang, Haris Pervaiz, and Nada M. Elshennawy. "A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images." Alexandria Engineering Journal 64 (2023): 923-935.

[18] Fan, Zong, Shenghua He, Su Ruan, Xiaowei Wang, and Hua Li. "Deep learning-based multi-class COVID-19 classification with x-ray images." In Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling, vol. 11598, pp. 286-296. SPIE, 2021.

[19] Darici, Muazzez Buket, Zumray Dokur, and Tamer Olmez. "Pneumonia detection and classification using deep learning on chest x-ray images." International Journal of Intelligent Systems and Applications in Engineering 8, no. 4 (2020): 177-183.

[20] Rajaraman, Sivaramakrishnan, Ghada Zamzmi, and Sameer Antani. "Multi-loss ensemble deep learning for chest X-ray classification." arXiv preprint arXiv:2109.14433 (2021).

GitHub Repository Link: `https://github.com/hariprasathvv/Multi-Class-Chest-X-Ray-Classification-Usi`