IMPROVING CHEST PATHOLOGIES DETECTION FROM CHEST X-RAY WITH DEEP

LEARNING USING TRANSFER LEARNING AND IMAGE ENHANCEMENT

TANABUT TAKSINAVONGSKUL

Graduate School Srinakharinwirot University

2022

IMPROVING CHEST PATHOLOGIES DETECTION FROM CHEST X-RAY WITH

DEEP LEARNING USING TRANSFER LEARNING AND IMAGE

ENHANCEMENT

ธนบุตร ทักษิณาวงศ์สกุล

IMPROVING CHEST PATHOLOGIES DETECTION FROM CHEST X-RAY WITH DEEP

LEARNING USING TRANSFER LEARNING AND IMAGE ENHANCEMENT

TANABUT TAKSINAVONGSKUL

A Master's Project Submitted in Partial Fulfillment of the Requirements

for the Degree of MASTER OF SCIENCE

(Data Science)

Faculty of Science, Srinakharinwirot University

2022

THE MASTER'S PROJECT TITLED

IMPROVING CHEST PATHOLOGIES DETECTION FROM CHEST X-RAY WITH DEEP

LEARNING USING TRANSFER LEARNING AND IMAGE ENHANCEMENT

BY

TANABUT TAKSINAVONGSKUL

HAS BEEN APPROVED BY THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE MASTER OF SCIENCE

IN DATA SCIENCE AT SRINAKHARINWIROT UNIVERSITY

....

(Assoc. Prof. Dr. Chatchai Ekpanyaskul, MD.)

Dean of Graduate School

....

ORAL DEFENSE COMMITTEE

.... Major-advisor

(Dr.Sophon Mongkolluksamee)

.... Chair

(Dr.Aimaschana Niruntasukrat)

.... Co-advisor

(Assoc. Prof. Dr.Worapan Kusakunniran)

.... Committee

(Asst. Prof. Dr.Sirisup Laohakiat)

Title	IMPROVING CHEST PATHOLOGIES DETECTION FROM CHEST X-RAY WITH DEEP LEARNING USING TRANSFER LEARNING AND IMAGE ENHANCEMENT
Author	TANABUT TAKSINAVONGSKUL
Degree	MASTER OF SCIENCE
Academic Year	2022
Thesis Advisor	Dr. Sophon Mongkolluksamee
Co Advisor	Associate Professor Dr. Worapan Kusakunniran

This research is concerned with chest radiography, which is essential for doctors to determine and follow up on lung disease. However, practicing radiologists have an insufficient ability to identify diseases in chest x-ray images. Therefore, the researchers developed deep-learning models to mitigate this problem, and CheXNet is one of the state-of-the-art models that can detect 14 lung pathologies. This research applied six image enhancement techniques to the x-ray images before using ChexNet to improve detection performance. The six techniques consisted of Gamma, Complement, HE, CLAHE, BCET, and MMCS. In addition, we studied the effectiveness of using a single enhancement technique (single channel) and a combination of them to the original image (multi-channel). Gamma gave the highest and most stable detection improvement using a single enhancement technique at 0.628% AUCROC in 14 diseases. Combining the original image, Gamma-enhanced image, and CLAHE-enhanced image shows 0.7% AUCROC improvement for 14 diseases. Moreover, this combination offers outstanding Pneumonia detection, which is 2% more than CheXNet.

Keyword : CheXNet Chest x-ray image enhancement multichannel input image DenseNet

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

Page

LIST OF TABLES

К

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

1.1 Introduction

Diagnostic imaging is a vital tool in today's medicine. Medical imaging techniques such as computed radiography, computed tomography, digital mammography, and magnetic resonance imaging, among others, help map a subject's anatomy. They are essential to diagnosing and therapy planning because they show normal and abnormal anatomy. Knowledge has considerably increased as a result of these technologies.

Medical imaging aids technology can support radiologists in making quicker and even more accurate diagnoses by providing a visual image of the inside of the human body. As a result, the doctor can treat diseases more effectively, resulting in better patient care. Medical imaging has progressed in measuring speed, spatial resolution, and contrast. Having this helpful tool necessitates having enough capacity to have qualified radiologists evaluate the required data.

Medical X-rays are images that diagnose several of the most sensitive human body organs, such as the bones, chest, teeth, and head. For generations, medical professionals have utilized this approach to investigate and visualize fractures or anomalies in specific body areas. Since X-rays are excellent diagnostic instruments for chest illnesses, they are non-invasive and cost-efficient. X-rays can reveal pathological changes, cavitations, consolidations, infiltrates, blunted costophrenic angles, and small, widely scattered nodules can all be seen on CXR images. Pleurisy, effusion, pneumonia, bronchitis, infiltration, nodule, atelectasis, pericarditis, Cardiomegaly, Pneumothorax, fractures, and many more disorders and diseases can be diagnosed with a chest X-ray[1].
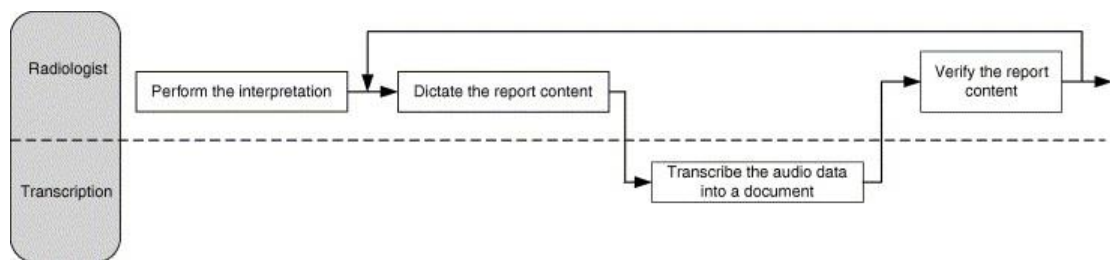
Figure 1: Common interpretation workflow.

Source: [2]

Figure 1 shows the typical process of how the radiologist interpreted the radiograph. Radiologists face a complex problem in classifying abnormalities on chest x-rays. As a result, computer-aided diagnostic (CAD) systems have been created in recent decades to extract meaningful information from X-rays to assist doctors in gaining a quantitative understanding of an X-ray. However, such CAD systems have not yet reached a level of significance that allows them to judge the types of diseases shown in X-rays. [3]. Thus, the role of CAD was left as visualization functionality that helps doctors in making decisions.

The field of medical image analysis is now intensely focused on deep learning. In 2012, Krizhevsky et al. presented AlexNet[4]—a convolutional neural network—for image classification in computer vision and won the ImageNet challenge by a large margin. The increased computer capacity (i.e., parallel computing of graphical processing units (GPUs)) and the large amount of data available made this possible. Such success reintroduced neural networks as a machine learning technique. Deep learning has already proven its capacity to interpret medical with excellent accuracy[5].

1.2 Problem Statement

Plain radiography is the most common imaging modality in radiology departments, and chest X-rays are the most frequent examination type [Bundesamt für Strahlenschutz, 2020; NHS England, 2020]. The limitation to getting all chest X-ray images evaluated by radiologists is inadequate capacity [Care Quality Commission, 2017; Royal College of Radiologists, 2018]. As the amount of data produced from various medical imaging methods rises[6] and the growing world population [United Nations DESA, 2019], The demand for expert reading capacity is likely to rise soon.

Chest pathologies interpreted from chest X-rays require an expert radiologist. The improvement of time consumption and accuracy of interpreting is required to mitigate the shortage of expert radiologists. Recently, deep learning (with Convolution Neuron Network: CNN) has been successful in medical mage interpreting. Implementing Transfer learning and Image enhancement techniques would improve the performance deep learning model.

The existing deep learning model, CheXNet[7], for diagnosing 14 chest pathologies performs well on most of them. However, some pathology still required improvement, e.g., Infiltration, Nodule, Pneumonia, and Consolidation. This experiment will focus on improving the model's performance using three image enhancement techniques. Then, Transfer learning and finetune the model with a new form of chest X-ray from Image enhancement techniques.

1.3 Objective

- To improve the performance of deep learning CNN, CheXNet[7], classification model on Infiltration, Nodule, Pneumonia, and Consolidation.
- To investigate and study the suitability of 3 different Image enhancement techniques with Transfer learning and fine-tuning that affect the performance of deep learning model.

O To construct the method used for Image enhancement technique to select the best detection performance on the specific pathologies, Infiltration, Nodule, Pneumonia, and Consolidation.

1.4 Scope and Limitation

1. The pathologies detection performance interest is:
    a. Infiltration
    b. Nodule
    c. Pneumonia
    d. Consolidation
2. Preprocessing for image enhancement technique used:
    a. Gamma correction
    b. Contrast limited adaptive histogram equalization (CLAHE)
    c. Balance Contrast Enhancement Technique (BCET)
3. Deep learning CNN technique is:
    a. Transfer Learning with the CheXNet model
    b. Fine-tune with the CheXNet model

1.5 Thesis Structure

The following paragraphs outline the structure of this thesis and provide an overview of each chapter and its contributions. Chapter 2 summarize the background information and essential literature.

Chapter one is a general introduction to the thesis, where the aims and significance of the thesis.

Chapter two is a general overview of the radiology X-rays and types of chest diseases that may be found in a radiograph and a brief on Deep convolutional

Neuron Networks in medical radiography. The image enhancement technique has been mentioned as well.

Chapter three is a dataset describing and flow of work in this study. In addition, Current progress is included the preliminary result, e.g., image post enhancement on gamma technique example.

Chapter four are results from experiments separated into two sections; The Single Channel Image Enhancement Results and the Multi-Channel Image Enhancement Results. Then, the table comparison of overall pathology detection performance and specific pathology detection performance is shown in this chapter.

Chapter five is the conclusion and discussion of the results.
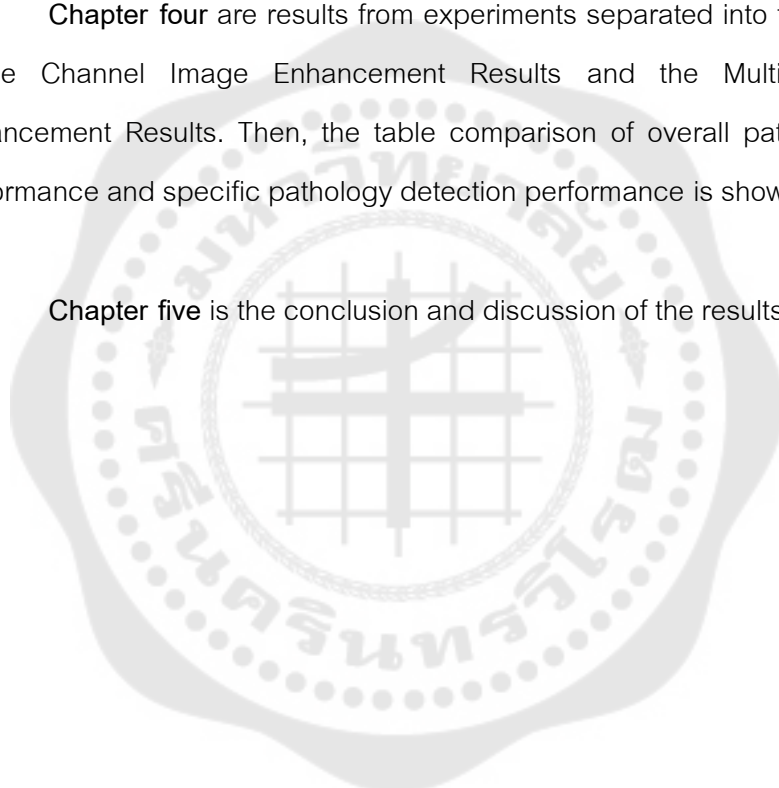
CHAPTER 2
LITERATURE REVIEW

2.1 Conventional radiography imaging

Wilhelm Röntgen, who discovered X-rays in 1895 and was the first to take a two-dimensional X-ray image of a human body part (see Figure 2 (b)), was the first to capture a two-dimensional X-ray image of a human body part. This discovery ushered in a new age in medical imaging, which has since grown in popularity to become the most common examination type. A two-dimensional projection imaging technique that includes projecting an object onto a detector is known as conventional radiography. X-radiation is generated by the X-ray tube and travels through things. Depending on the varying densities and attenuation coefficients of materials, the intensity of X-radiation is dispersed or muted (i.e., bones, tissues, and fluids)

2.2 Digital Radiography

Digitalization can convert results into a digital image or planar radiograph. Planar radiographs can be thought of as two-dimensional (2D) arrays of gray values. Each array element or pixel (picture element) represents precisely one image point of the detector. The gray level encodes the optical density related to the amount of the transmitted energy imparted at the corresponding pixel area. [8]
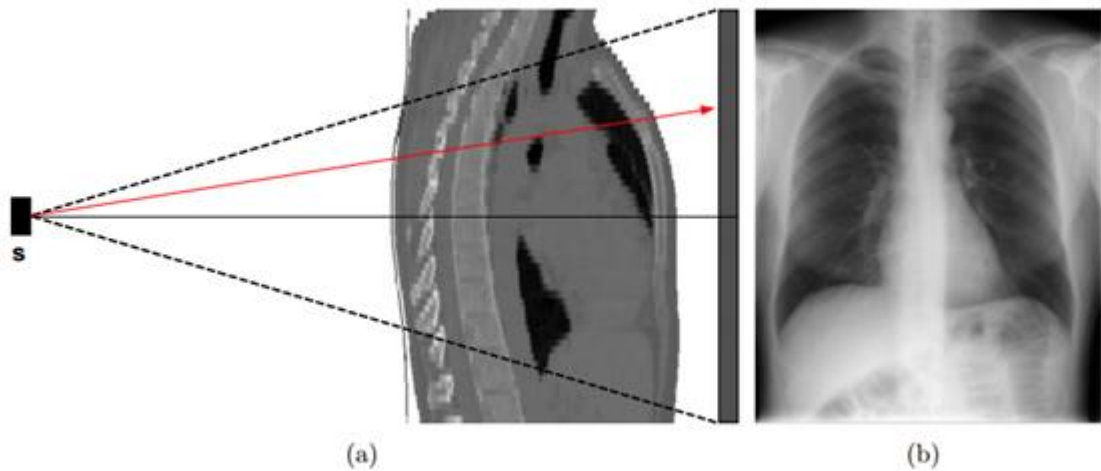
Figure 2: (a) Schematic representation of the X-ray imaging system for Digital PA chest radiography. (b) The digital image (planar radiograph)

Source: [9] Medical X-ray Images of the Human Thorax

A typical digital PA chest radiograph is shown in Figure 2(a); PA stands for posterior-anterior, meaning that the patient faces the observer (the radiation passes through the patient from back to front). By convention, the brightness indicates absorbed radiation.

2.3 Chest X-ray Abnormalities

The findings of the frequency study are shown in Table 1. There are two issues with this for image processing. First, the wide range of findings makes it difficult to create an automatic picture analysis that categorizes the majority of discoveries based on hand-made criteria. This challenge explains why, when utilizing hand-crafted feature extraction techniques, researchers frequently solely focus on aberrant individual discoveries. Because feature engineering is no longer necessary thanks to deep learning, researchers no longer need to concentrate on specific results.

Frequency ranking	Finding	Count	% of all images (N = 1089)	% of all abnormal (N = 877)
1.	Pulmonary infiltrates	482	44%	55%
2.	I. V. catheters	291	27%	33%
3.	Heart size/contour	239	22%	27%
4.	Endotracheal/tracheostomy tubes	193	18%	22%
5.	Pleural effusions	130	12%	12%
6.	Linear atelectasis/scar	86	8%	10%
7.	Drainage catheters and tubes	78	7%	9%
8.	Pulmonary vascularity	77	7%	9%
9.	Pleural scarring	69	6%	8%
10.	Rib lesions	65	6%	7%
11.	Mediastinal masses	56	5%	6%
12.	Diaphragm	44	4%	5%
13.	Calcified granulomas	43	4%	5%
14.	Pneumothorax	42	4%	5%
15.	Lung nodules	40	4%	5%
16.	Extrathoracic abnormalities	36	3%	4%
17.	Lung masses	17	2%	2%
18.	Calcified nodes	13	1%	1%
19.	Mediastinal shift/contour	13	1%	1%
20.	Cardiac pacemakers	12	1%	1%

Table 1: Abnormal finding distribution in chest X-rays.

Source: [10] Comparison of imaging properties of a computed radiography system and screen–film systems.

Most exams performed in radiology departments are chest X-rays [Bundesamt für Strahlenschutz, 2020; NHS England, 2020]. Therefore, software support is required, given the radiology field's increasing workload and declining profitability. There are numerous additional clinical uses outside completely automated chest X-ray processing, where a radiologist merely has to cross-check the data.

The research for detecting all regular chest X-rays (i.e., no abnormal findings on the chest X-ray) can significantly reduce the workload in a radiology department.

2.4 Deep Convolutional Neural Networks

Due to their better performance compared to other machine learning paradigms, deep CNNs have been extensively applied in image classification. The network structure automatically extracts the spatial and temporal features of an image. Many applications have successfully implemented the transfer learning method, particularly those where finding an extensive dataset might be challenging.

Information is extracted hierarchically in a convolutional neural network [11]. The initial layers extract basic information like edges and color blobs. Deeper layers collect feature combinations from prior layers by linearly combining previously extracted features. High-level convolutional layers are used in the final convolutional layers; high-level features are extracted from the image. Figure 3 demonstrates a hierarchical feature extraction. The top row shows a multi-layered convolutional neural network. Color blobs and edges are extracted in the first layers, while circle combinations are extracted in the middle layers. Then, specific items that should be linearly separable by a classifier (i.e., the final fully-connected layer) are retrieved.
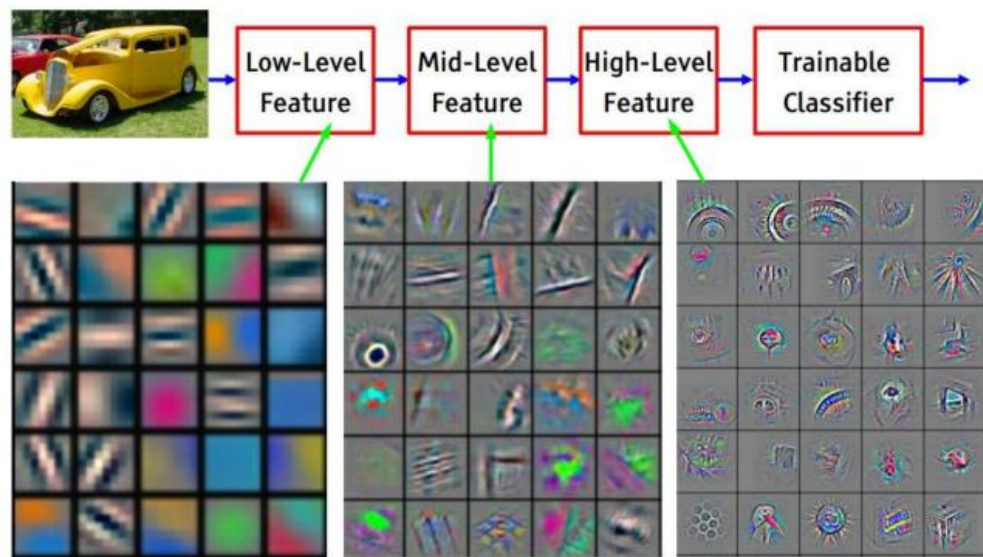
Figure 3: Hierarchical feature extraction of a convolutional neural network. The top row illustrates the layers of a convolutional neural network. The bottom row presents the feature visualization of a convolutional network trained on the ImageNet dataset[12].

source: [11, 13]

2.4.1 DenseNet121, CheXNet

CheXNet is the model based on the architecture of DenseNet121, as shown in the architecture in Figure 4, with pretrained weight from the research[7]. With an increasing number of parameters, DenseNets tend to provide constant accuracy improvements without any indications of performance deterioration or overfitting. It produced cutting-edge findings in various scenarios on several highly competitive datasets. Additionally, DenseNets require far fewer parameters and processing to execute at the highest level. We expect that more accurate tuning of the hyperparameters and learning rate schedules might result in even more significant improvements in the accuracy of DenseNets. Because we used hyperparameter settings intended for residual networks in our study.[14]

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

Figure 4: DenseNets architectures for ImageNet. Note that each "conv" layer in the table corresponds to the sequence BN-ReLU-Conv.

Source: [14]

CheXNet is a 121-layer Dense Convolutional Network (DenseNets) trained on the Chest X-ray 14 dataset by Huang et al. (2016). DenseNets enhance the network's information flow and gradients, making very deep network optimization manageable. After switching out the final fully linked layer for one with a single output, we add a sigmoid nonlinearity.

The network weights are initialized from a model pretrained on ImageNet (Deng et al., 2009). The network is trained end-to-end using Adam with standard parameters ($\beta_1 = 0.9$ and $\beta_2 = 0.999$) (Kingma & Ba, 2014). We used minibatches of size 16 to train the model. The initial learning rate of 0.001 decayed by a factor of 10 each time the validation loss plateaus after an epoch, and pick the model with the lowest validation loss.[7]

2.4.2 Result from Image classification as multi-label

The classification problem in this study is multi-label classification. One CXR image could have more than one pathology or non-mutually exclusive, as shown in figure 5 and figure 6. The evaluation metric used for comparison is AUCROC on each pathology class.
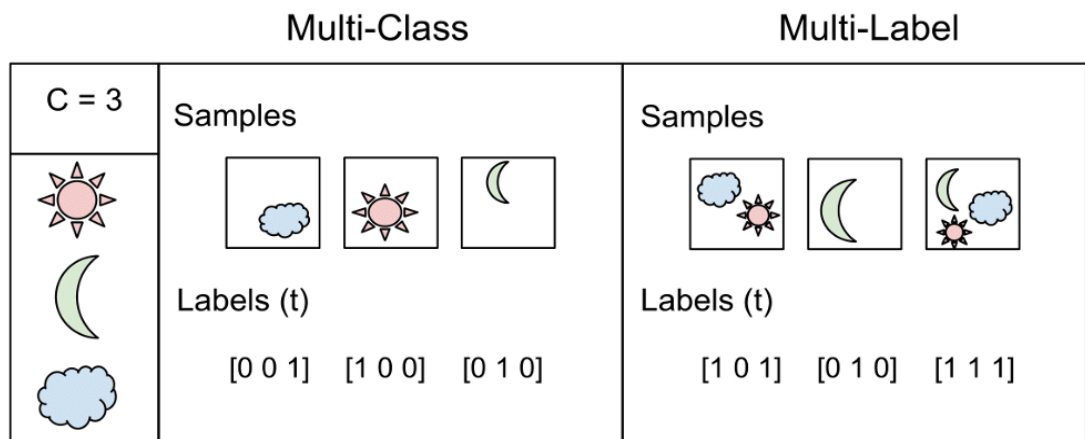
Figure 5: Multiclass problem and multi-label problem comparison

Source: https://prakhartechviz.blogspot.com/2019/02/multi-label-classification-python.html

(a) *alone, athletics, high-jump, women, sign-of-triumph, joy.* Photo by Cornelius Poppe / Scanpix
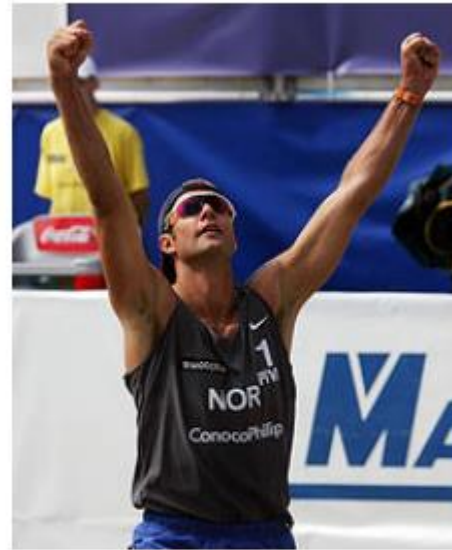
(b) *alone, beach-volleyball, action, joy.* Photo by Alf Ove Hansen / Scanpix

Figure 6: Example of two similar images but with different labels. These two images have "joy" and "alone" non-mutually exclusive.[15]

2.5 Image Enhancement

A pixel can be understood as the intensity value at a specific point in an image. It can be considered the visual perception of a collection of pixels. Typically, 2D descriptions of pixels, like f (x, y).

Picture processing is essential to computer vision because it allows for properly conditioning image data prior to machine learning. The number of gray levels utilized can affect the pixel values of an image. For a picture with a gray level of m, the pixel range can be written as 0 to 2m.

In medicine, image processing has been widely applied. In this industry, image enhancement is always the most frequently required process. A medial image

has numerous components and may be noisy. Finding the correct diagnosis is particularly difficult for doctors due to the noise and unclear medical image. Since all components of an image, including noise, differ in brightness and intensity, image processing can be a valuable tool for detecting and enhancing images. As a result, image processing technologies are employed in this work to improve chest X-ray images and eliminate any possible noise. Several methods are used in image processing to enhance images, including filtering, histogram equalization, and intensity correction.

Numerous filters can be applied when filtering data, including Gaussian, median, and mean filters. Images are screened for median filters because some of them have noise artifacts that need to be eliminated to improve the quality of the images. As it rejects the Salt and Pepper noise seen in some medical imaging, the median filter effectively reduces noise.

Moreover, image intensity adjustment can also be used to enhance the quality of images. This technique involves mapping the pixel's intensity distribution from one level to another. The intensities of pixels are increased by mapping them into other values to highlight the images more and more. The image ended up with brighter images where the cells are more apparent, including the cancerous cells.

Image enhancement is a crucial image-processing method that suppresses or eliminates some secondary information from images to increase the classification quality. The goal is to improve upon the original photos such that the objective images are more suited for a particular purpose. In this study, we use five different enhancing techniques. These image enhancement methods will be briefly explained in the section that follows:[7]

2.5.1 Histogram Equalization (HE)

The goal of the histogram equalization (HE) method is to distribute the grayscale values in an image evenly. As a result, the likelihood of each gray level is equal. To improve image quality, HE adjusts the brightness and contrast of dark and low-contrast images[60]. A dark image would cause the histogram to be skewed towards the lower end of the grayscale, and the image data would be packed into the dark end of the histogram. The grey levels can be re-distributed in a more evenly distributed histogram at the dark end, making the picture clear. The histogram of a digital image with intensity levels in the range [0, L-1] is a discrete function represented as follows:

$$h(r_k) = n_k \qquad (1)$$

Where, r_k is kth intensity value, n_k is the number of pixels in the image with intensity, r_k. Histograms are frequently normalized by the total number of pixels in the image. Assuming an M x N image, a normalized histogram is related to the probability of occurrence of r_k in the image, as shown in equation 2.

$$P(r_k) = \frac{r_k}{M * N} \qquad (2)$$

2.5.2 Contrast-limited adaptive histogram equalization (CLAHE)

An improved histogram equalization (HE) variant is called Adaptive Histogram Equalization (AHE). AHE increases the contrast of each region independently by applying histogram equalization over small regions (i.e., patches) in the image. As a result, rather than using the image's general information, it enhances local contrast and edges in each region according to the local distribution of pixel intensities. AHE, however, could exaggerate the image's noise component. [61]. Contrast-limited adaptive histogram equalization (CLAHE), on the other hand,

produces photos that look more naturally boosted than HE does. It was found that the HE approach can oversaturate some areas when used on the X-ray images. CLAHE adopts the same strategy as AHE to address this issue. A threshold parameter, however, limits the amount of contrast enhancement produced inside the chosen region. First, the original image is changed from RGB (red, green, and blue) to HSV (hue, saturation, and value) color space to create a form of color that is more closely related to how people see color. Second, CLAHE processes the value portion of HSV without modifying the hue or saturation. Each gray level is redistributed to the original histogram's cropped pixels once it has been cropped. Each pixel's value is decreased until it reaches a preset limit. The image that has undergone HSV processing is then changed to RGB color space.

2.5.3 Image Invert/ Complement

When black and white are reversed in a binary image using the image inversion or complement approach, zeros become ones, and ones become zeros. The original pixel value for an 8-bit grayscale image is subtracted from the greatest intensity value, 255; the result is the new image's pixel value. In x-ray photographs, the light spots get darker, and the dark spots get lighter than in the original images. The mathematical formulation is simple:

$$y = 255 - x \qquad (3)$$

Where x and y are the intensity values of the original and the transformed (new) images, this technique shows the lungs area (i.e., the region of interest) lighter and the bones are dark as It can be noted that the histogram for the complemented image is a flipped copy of the original image. As this is a standard procedure used widely by radiologists, it may help deep networks for better classification.

2.5.4 Gamma correction

Gamma correction carries out a nonlinear operation on the pixels of the source image. The projection relationship between the pixel value and the gamma value following the internal map, gamma correction changes the pixel value to improve the image. Image normalization often involves performing linear operations on each pixel, such as scalar multiplication, addition, and subtraction. If P represents the pixel value inside the [0,255] range, Ω represents the angle value, Γ is the symbol of the gamma value set, and x is the grayscale value of the pixel (x $\in$ P). Let xm be the range midpoint [0, 255]. The linear map from group P to group Ω is defined as:

$$\varphi\colon P \rightarrow \Omega, \Omega = \{\omega | \omega = \varphi(x)\}, \varphi(x) = \frac{\pi x}{2x_m} \qquad (4)$$

The mapping from Ω to Γ is defined as:

$$h\colon \Omega \rightarrow \Gamma, \Gamma = \{\gamma | \gamma = h(x)\} \qquad (5)$$

$$\begin{cases} h(x) = 1 + f_1(x) & (6) \\ f_1(x) = acos(\varphi(x)) & (7) \end{cases}$$

Where a $\in$ [0, 1] denotes a weighted factor.

Group P can be related to Γ group pixel values based on this map. The arbitrary pixel value is calculated with a given Gamma number. Let γ (x) = h(x), and the Gamma correction function is as follows

$$g(x) = 255 \left(\frac{x}{255}\right)^{\frac{1}{\gamma/x}} \qquad (8)$$

Where g(x) represents the output pixel correction value in grayscale.

2.5.5 Balance Contrast Enhancement Technique (BCET)

By stretching or compressing the image's contrast without changing the histogram pattern of the image data, BCET is a method for enhancing balance contrast[16]. The parabolic function obtained from the picture data is the foundation for the solution. The general parabolic functional form is defined as

$$y = a(x - b)^2 + c \qquad (9)$$

The three coefficients, a, b and c, are determined from the following equations using the minimum, the maximum, and the mean of the input and output image values.

$$b = \frac{h^2(E - L) - s(H - L) + l^2(H - E)}{2[h(E - L) - e(H - l) + l(H - E)]} \qquad (10)$$

$$a = \frac{H - L}{(h - l)(h + l - 2b)} \qquad (11)$$

$$c = L - a(l - b)^2 \qquad (12)$$

Where 'l' represents the input image's minimum value of the input image. 'h' denotes the maximum value. 'e' denotes the mean value of the input image. 'L' is the minimum value of the output image, 'H' denotes the maximum value of the output image and 'E' denotes the mean value of the output image.
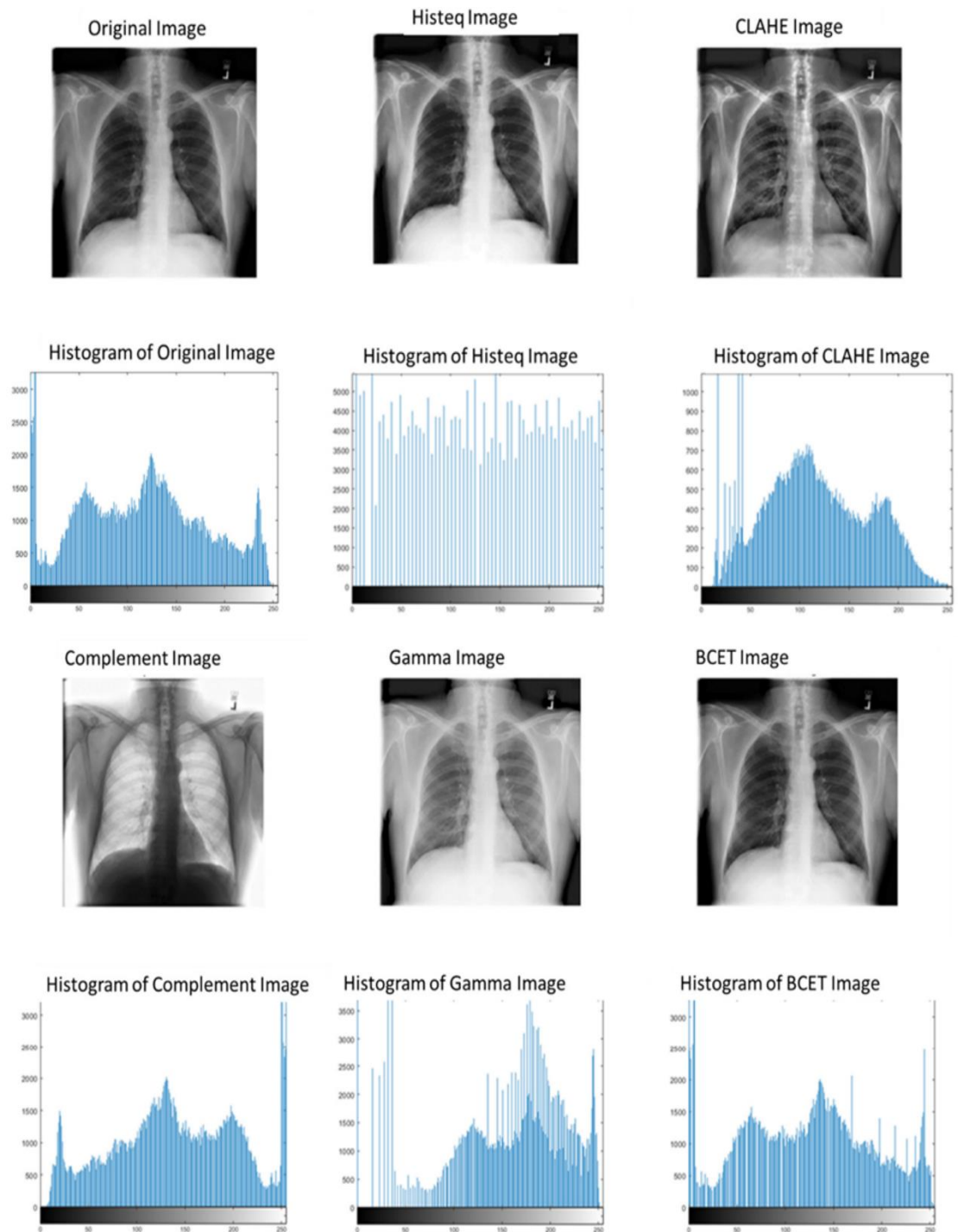
Figure 7: Histogram for original X‑ray image and images undergo different enhancement techniques[17].

2.5.6 Min Max Linear Contrast Stretching

Contrast stretching involves extending an image's contrast beyond its intensity values to a specified range of values. Another name for it is normalization. Minimum-Maximum, Percentage, and Piecewise Contrast Enhancement are a few methods of contrast stretching.

In Min-Max Contrast Stretching for each pixel:

$$pixel = (pixel - min)/(max - min)) * 255$$

Where min and max are the image's maximum and minimum pixel values.[18]

2.6 Related Works

The National Institutes of Health (NIH) CXR dataset [19] comprises 112,120 frontal CXRs, individually labeled to include up to 14 distinct pathologies. The authors employed Natural Language Processing to text-mine illness diagnoses with an estimated accuracy of more than 90% to create these labels from the related radiological reports.

The CheXNet deep CNN model, which uses this NIH CXR dataset, is stated to outperform the average radiologist on the pneumonia diagnosis task significantly. CheXNet image classification models accept input images with dimensions of 224x224. The CNN architecture is a 121-layer convolutional neural network trained using ChestX-ray14, the world's most extensive publicly available chest X-ray dataset, which contains over 100,000 frontal view X-ray pictures of 14 illnesses. In the following study, the weight from model training might be used for Transfer learning.[7]

The CheXNet is based on Dense Convolutional Network (DenseNet). It establishes direct links between any two layers with the same feature-map size. We

demonstrated that DenseNets grow quickly to hundreds of layers while posing no optimization difficulties. DenseNets yield constant improvements in accuracy as the number of parameters increases, with no indication of performance degradation or overfitting. It delivered cutting-edge outcomes across different datasets in a variety of settings. Furthermore, DenseNets require far fewer parameters and less computation to reach state-of-the-art performance.[14] The propose of DenseNet121 to be implemented came from the performance comparison of ResNet152, DenseNet121, InceptionV4, and SEResNeXt101 on CheXpert, finding that DenseNet121 performed best.[20]

On 12 image classification datasets, 16 convolutional neural networks (CNNs) have been tested using ImageNet Transfer. They discovered that applying these ImageNet pretrained structures to logistic regression as feature extractors or fine-tuning them on the target dataset produced Spearman ρ = 0.99 and ρ = 0.97 between ImageNet accuracy and transfer accuracy, respectively. Regularizes that improve ImageNet performance are highly detrimental to transfer learning performance based on penultimate layer features, and better ImageNet architectures obtain outstanding accuracy.[21]

This research investigates the impact of prominent image-enhancing techniques and reports on their impact on detection performance. Five image improvement techniques were used To increase COVID-19 detection accuracy: histogram equalization (HE), contrast limited adaptive histogram equalization (CLAHE), image complement, gamma correction, and Balance Contrast Enhancement Technique (BCET). The gamma correction approach outperforms other enhancing techniques in detecting COVID-19 from typical and segmented lung Chest X-ray images.[17]

CHAPTER 3

METHODOLOGY

3.1 Dataset

The dataset includes around 60% of all frontal chest x-rays taken at the hospital and was taken from the clinical PACS database at the National Institutes of Health Clinical Center. As a result, compared to earlier chest x-ray datasets, we anticipate that this dataset is much more representative of the actual patient population distributions and clinical diagnosis problems. Of course, the dataset size—the overall number of images and the frequencies of lung diseases—would improve the deep learning training efficiency.[22]. Refer to [19] for the details of how the dataset is extracted and image labels are mined through natural language processing (NLP).

The whole corpus of ChestX-ray14 is used to train and evaluate techniques for multilabel pathology classification. Figure 8 shows 8 chosen ChestX-ray14 samples. The collection includes 112,120 frontal chest X-rays from 30,805 patients.

In this research, they randomly split the dataset into training (28744 patients, 98637 images), validation (1672 patients, 6351 images), and test (389 patients, 420 images). There is no patient overlap between the sets.

The collection only includes preprocessed images and does not include the raw DICOM images. [19] used the encoded display settings to conduct a simple preprocessing while the pixel depth was decreased to 8 bits. Also, each image was resized to 1024 x1024 pixels without concern for the aspect ratio.
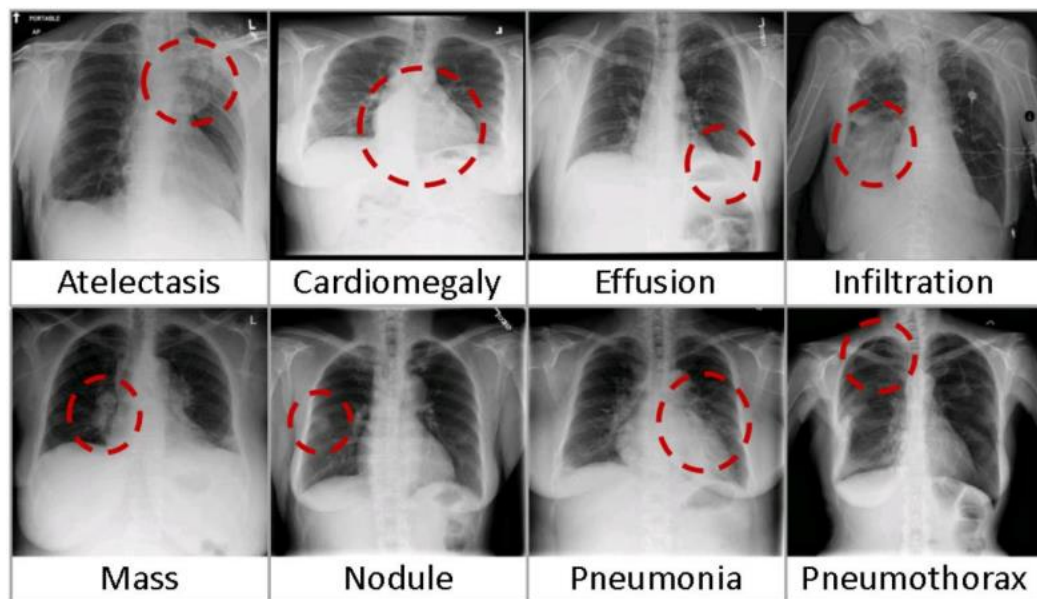
Figure 8: eight visual examples of common thorax diseases

Source: [19]

The distribution of each class and statistics for non-image data are provided in Tables 2 and 3, as well as Figure 9. The prevalence of each pathology was usually rare, with frequency ranging from 0.2 percent to 17.74 percent (see Table 2). The patient gender and view position distributions were relatively equal, with ratios of 1.3 and 1.5, respectively (see Table 3). The histogram in Figure 9 shows the distribution of patient age in ChestX-ray14. The average age of the patients was 46.87 years, with a standard deviation of 16.60 years.

Pathology	True	False	Prevalence [%] $N = 112,120$
Cardiomegaly	2,776	109,344	2.48
Emphysema	2,516	109,604	2.24
Edema	2,303	109,817	2.05
Hernia	227	111,893	0.20
Pneumothorax	5,302	106,818	4.73
Effusion	13,317	98,803	11.88
Mass	5,782	106,338	5.16
Fibrosis	1,686	110,434	1.50
Atelectasis	11,559	100,561	10.31
Consolidation	4,667	107,453	4.16
Pleural thickening	3,385	108,735	3.02
Nodule	6,331	105,789	5.65
Pneumonia	1,431	110,689	1.28
Infiltration	19,894	92,226	17.74
No findings	60,412	51,700	53.89

Table 2: Summary of disease distribution in the ChestX-ray14 dataset. For each disease, the total number of "true" and "false" (i.e., whether the disease is present or not) and their prevalence are given. The last row shows the number of "true" and "false" items for the implicit label "No Finding."

Prevalence is the value shown ratio of the positive number found in a sample of people studied and usually used in medical conditions for risk of disease. In this case, lung disease positive value by a total number of CXR images.

$$Prevalence = \frac{\# \ of \ people \ in \ sample \ with \ characteristic}{Total \ \# \ of \ people \ in \ sample}$$

Example on Nodule:

$$Prevalence\ on\ Nodule\ =\ \frac{6331}{(6331 + 105789)} * 100\ =\ 5.65\,\%$$

	Female	Male	Ratio
Patient gender	63, 340	48, 780	1.30
	PA	**AP**	**Ratio**
View position	67, 310	44, 810	1.50

Table 3: Distribution of patient gender and view position in the ChestX-ray14 dataset. For patient gender, the total count of female and male is shown, and for view position, the total count of posterior-anterior (PA) and anterior-posterior (AP) is given. In the third column, the ratio between the first and second columns
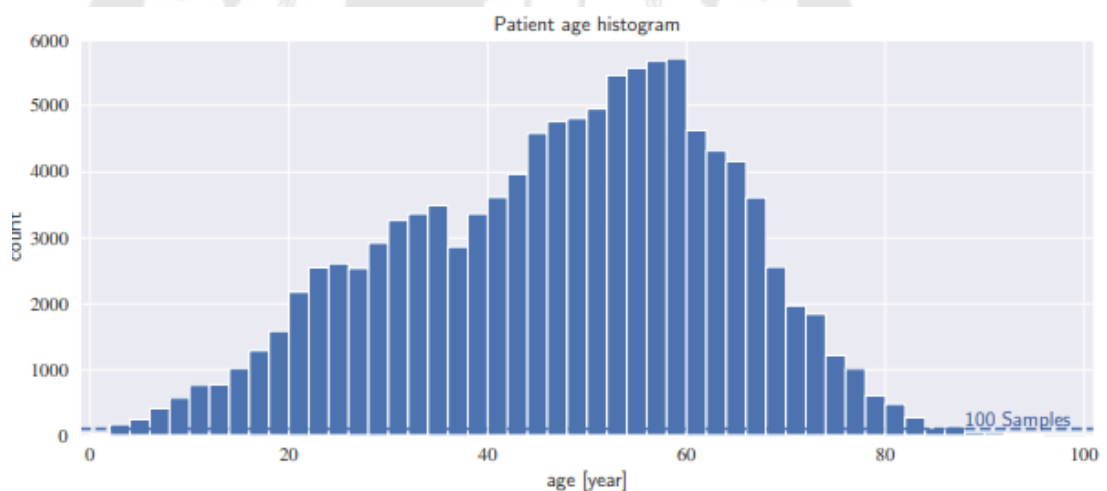
Figure 9: Distribution of patient age in the ChestX-ray14 dataset. Each bin covers a width of two years. The average patient age was 46.87 years, with a standard deviation of 16.60 years.

More exploration into the dataset. As per Figure 10, check each CXR image, how many pathologies will show per image, and visualize the bar chart.

Evident exist found some insight information on each disease that have some correlation.
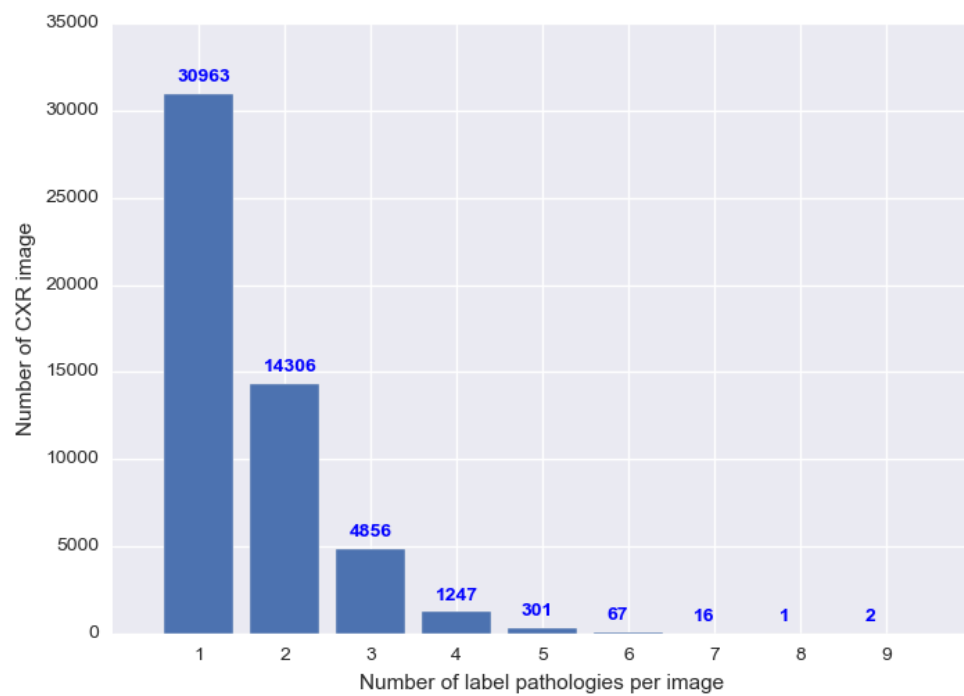
Figure 10: Explore the Number of labels per one image distribution without "No Finding" included

3.2 Method for study

This section exhibits experimental testing for chest X-ray pathology detection using deep learning. And background about the dataset and CheXNet deep learning model. There is a requirement to improve the effectiveness and performance of the image enhancement technique on the transfer learning CheXNet model. The Flow from the start to the end of the prediction of each deep model is as follows.

3.2.1 Flow of work

3.2.1.1 Image Enhancement and preparation

In This study, six types of Image enhancement are implemented from the flowchart. To experiment with preprocessing on each Chest X-ray image of each 14 diseases.

Figure 11 shows frontal-view chest X-ray images are put through the image enhancement process. Prepared datasets were separated on each processed image from each technique to prepare for training the Transfer Learning CheXNet model.

In the case of the Gamma correction technique, The initial brightness of the Chest X-ray image is required to be the base starting point of adjustment of the gamma. After obtaining the initial brightness of each Chest X-ray image, adjust of gamma value on each Chest X-ray image in the condition range. Then gamma corrected CXR image can feed to the Transfer Learning CheXNet model.
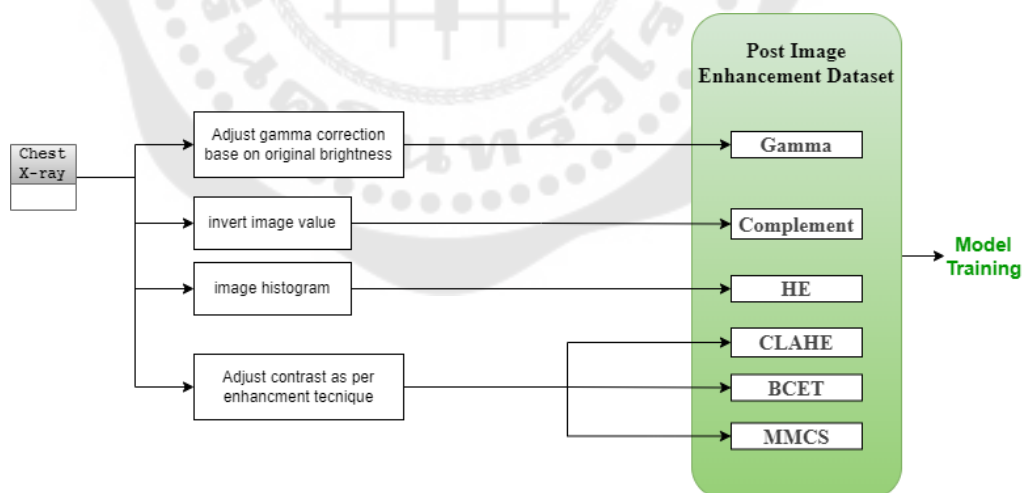
Figure 11: Flowchart of a subprocess for post Image Enhancement dataset

From the image enhancement preparation process, Gamma Correction must adjust the gamma parameter based on the original image's brightness value. Gamma adjustment will not be too high or too low, which can

affect to model training result. For example, from Figure 12, the Gamma correction process has been implemented with the heuristic process by selecting the start point of brightness and the gamma setting. As per Figure 13, the brightness of this image (00002364_001.png) is 42.98%. Then, the Gamma setting is chosen as 0.6. From the exploration of dataset brightness, the distribution is shown in Figure 13.
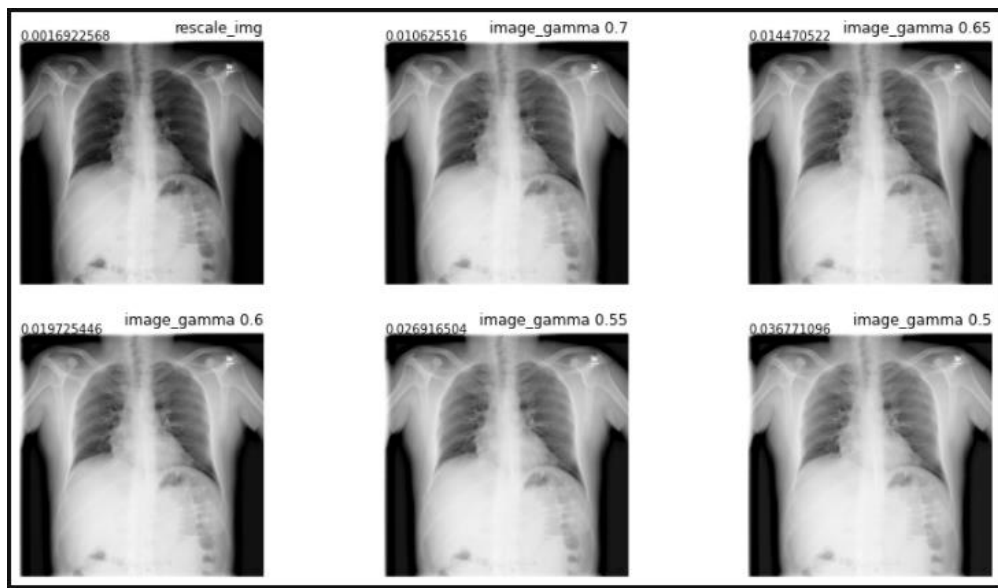
Figure 12: Post-gamma correction variation
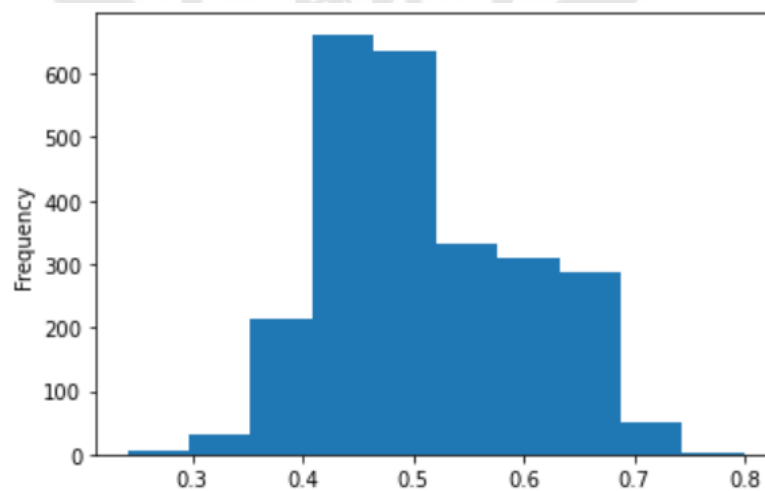
Figure 13: datasets brightness distribution

3.2.1.2 Single Channel Image Enhancement process

The proposed work is shown In Figure 14. Initially, the test dataset of chest X-ray images will go through the Deep CNN pertained model, the "CheXNet" model. The result is used for the baseline in performance compared with the result from different image enhancement techniques on CheXNet fine-tune model.
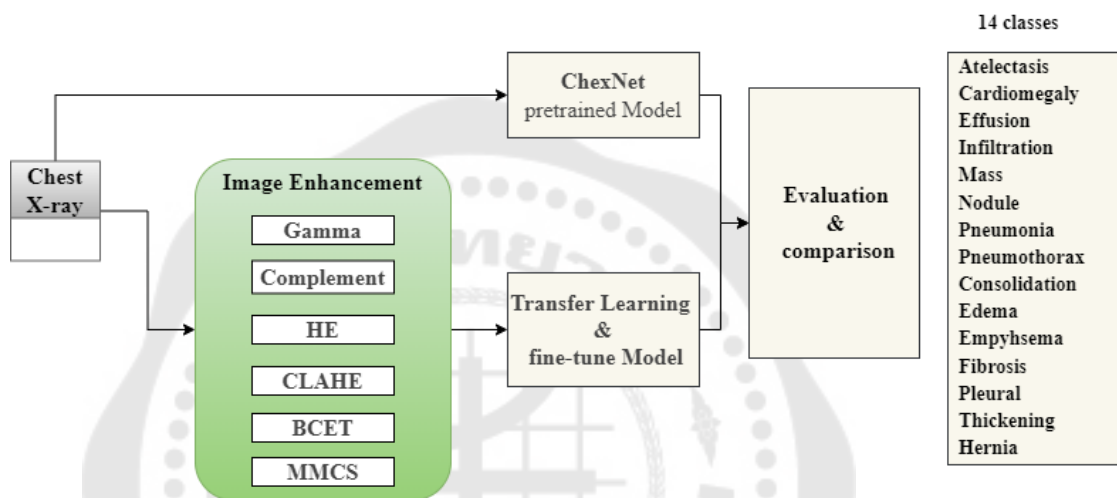
Figure 14: Flow of the Methodology for Single Channel Image Enhancement

3.2.1.3 Multi-Channels Image Enhancement process

The proposed work is shown In Figure 15—post enhancements chest. X-ray image kept in source storage. The next step is to get the array of each combination Image enhancement technique. The first stack of an array is the original image. The second and Third stacks are the combination of the experiment shown in the flow of work below. For example, in the Gamma+CLAHE combination, the first array stack is the original image. The second is the Post Gamma correction image, and the third stack is the Post CLAHE image. Afterward, the image stack with the three channels of mention here was input into the model to train the loop from all data sources of this combination. The model training use CheXNet transfer learning and fine-tuning. Then, the evaluation process proceeds accordingly.
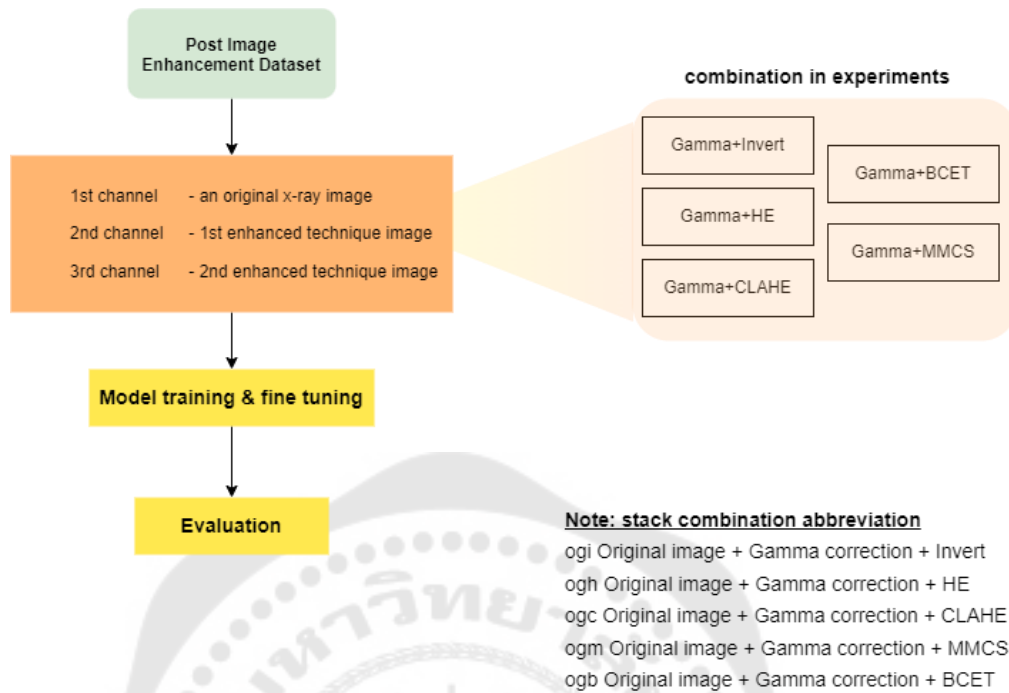
Figure 15: Flow of the Methodology for Multi-Channels Image Enhancement

3.2.2 Training Procedure

We use a 121-layers Dense Convolutional Network (DenseNets) training on Chest X-ray 14 dataset. The architecture of the network is shown in Figure 16. The network takes as input an image of a chest X-ray and outputs a class prediction. The image used for the training model is from different image preprocessing techniques according to the preparation step above.

```
Layer (type)                   Output Shape        Param #    Connected to
==================================================================================================
input_1 (InputLayer)           [(None, 224, 224, 3) 0

zero_padding2d (ZeroPadding2D) (None, 230, 230, 3)  0          input_1[0][0]

conv1/conv (Conv2D)            (None, 112, 112, 64) 9408       zero_padding2d[0][0]

conv1/bn (BatchNormalization)  (None, 112, 112, 64) 256        conv1/conv[0][0]

conv1/relu (Activation)        (None, 112, 112, 64) 0          conv1/bn[0][0]

zero_padding2d_1 (ZeroPadding2D (None, 114, 114, 64) 0         conv1/relu[0][0]

pool1 (MaxPooling2D)           (None, 56, 56, 64)   0          zero_padding2d_1[0][0]

conv2_block1_0_bn (BatchNormali (None, 56, 56, 64)  256        pool1[0][0]

conv2_block1_0_relu (Activation (None, 56, 56, 64)  0          conv2_block1_0_bn[0][0]

conv2_block1_1_conv (Conv2D)   (None, 56, 56, 128)  8192       conv2_block1_0_relu[0][0]

conv2_block1_1_bn (BatchNormali (None, 56, 56, 128) 512        conv2_block1_1_conv[0][0]

conv2_block1_1_relu (Activation (None, 56, 56, 128) 0          conv2_block1_1_bn[0][0]

conv2_block1_2_conv (Conv2D)   (None, 56, 56, 32)   36864      conv2_block1_1_relu[0][0]

conv2_block1_concat (Concatenat (None, 56, 56, 96)  0          pool1[0][0]
                                                               conv2_block1_2_conv[0][0]

conv2_block2_0_bn (BatchNormali (None, 56, 56, 96)  384        conv2_block1_concat[0][0]

conv2_block2_0_relu (Activation (None, 56, 56, 96)  0          conv2_block2_0_bn[0][0]

conv2_block2_1_conv (Conv2D)   (None, 56, 56, 128)  12288      conv2_block2_0_relu[0][0]

conv2_block2_1_bn (BatchNormali (None, 56, 56, 128) 512        conv2_block2_1_conv[0][0]

conv2_block2_1_relu (Activation (None, 56, 56, 128) 0          conv2_block2_1_bn[0][0]

conv2_block2_2_conv (Conv2D)   (None, 56, 56, 32)   36864      conv2_block2_1_relu[0][0]

conv2_block2_concat (Concatenat (None, 56, 56, 128) 0          conv2_block1_concat[0][0]
                                                               conv2_block2_2_conv[0][0]

conv5_block16_0_bn (BatchNormal (None, 7, 7, 992)   3968       conv5_block15_concat[0][0]

conv5_block16_0_relu (Activatio (None, 7, 7, 992)   0          conv5_block16_0_bn[0][0]

conv5_block16_1_conv (Conv2D)  (None, 7, 7, 128)    126976     conv5_block16_0_relu[0][0]

conv5_block16_1_bn (BatchNormal (None, 7, 7, 128)   512        conv5_block16_1_conv[0][0]

conv5_block16_1_relu (Activatio (None, 7, 7, 128)   0          conv5_block16_1_bn[0][0]

conv5_block16_2_conv (Conv2D)  (None, 7, 7, 32)     36864      conv5_block16_1_relu[0][0]

conv5_block16_concat (Concatena (None, 7, 7, 1024)  0          conv5_block15_concat[0][0]
                                                               conv5_block16_2_conv[0][0]

bn (BatchNormalization)        (None, 7, 7, 1024)   4096       conv5_block16_concat[0][0]

relu (Activation)              (None, 7, 7, 1024)   0          bn[0][0]

avg_pool (GlobalAveragePooling2 (None, 1024)        0          relu[0][0]

predictions (Dense)            (None, 14)           14350      avg_pool[0][0]
==================================================================================================
Total params: 7,051,854
Trainable params: 6,968,206
Non-trainable params: 83,648
```

Figure 16: a partial example from the block of DenseNet121 architecture top and

bottom part

We train chest X-ray classification models with pretrained ImageNet and use the weight obtained from our training condition without CheXNet weight.

The task of interest is to predict the probability of different pathologies from chest X-rays. We use the 112,120 frontal chest X-rays from 30,805 patients labeled for the presence or absence of 14 radiological observations, split the dataset into training (28744 patients, 98637 images), validation (1672 patients, 6351 images), and test (389 patients, 420 images). There is no patient overlap between the sets. The base model used is DensNet121, ImageNet weight loaded, batch size 32, and the initial learning rate is 0.001. The image dimension input 224x224 from Chest X-ray image 1024x1024, used callback "ReduceLROnPlateau" to decay the learning rate each epoch, and the minimum learning rate is 1e-8, optimizer "Adam" with standard parameters (β_1 = 0.9 and β_2 = 0.999), loss function "binary_crossentropy" and train 50 epochs. Imbalance data optimization by class weighting. An activation function is Sigmoid. Data Augmentation was implemented by flipping the image horizontally only to alter between the PA and AP types of the Chest X-ray image. The best weight from training with an original image of the Chest X-ray 14 dataset is CheXNet weight. Moreover, the models were trained on Nvidia GeForce RTX 3080 GPUs with 23 GB of memory.

Further study starts from the CheXNet weight. Then it uses this weight to fine-tune the deep learning model with the same parameter above but alter the input source from 6 different image enhancement techniques prepared earlier. The model architecture is still the same, except we unfreeze the layer from the start to adjust the weight of each neural network on each layer.

3.2.3 Evaluation and Comparison

We evaluate models using the average of their AUROC metrics (AUC) on the 14 radiological observations (Atelectasis, Cardiomegaly, Effusion, Infiltration, Mass, Nodule, Pneumonia, Pneumothorax, Consolidation, Edema, Emphysema, Fibrosis, Pleural Thickening, Hernia) and Comparison between image enhancement

technique. The ROC curve is plotted with TPR against the FPR where TPR is on the

y-axis and FPR is on the x-axis.
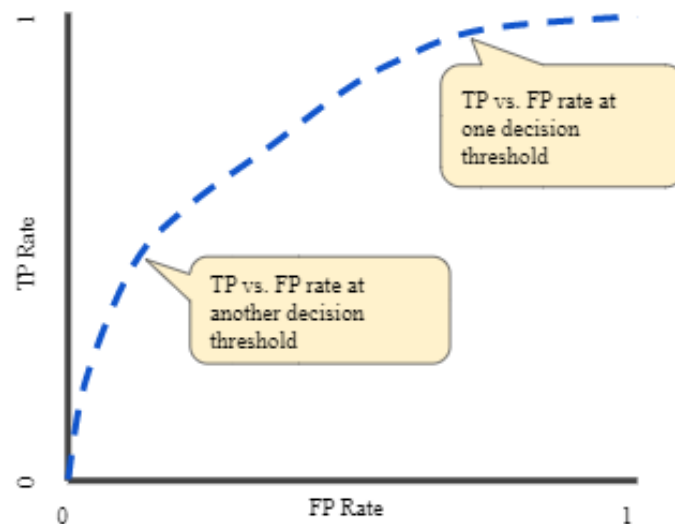
Figure 17: AUC - ROC Curve

source: https://developers.google.com/machine-learning/crash-

course/classification/roc-and-auc

Defining terms used in AUC and ROC Curve

$$TPR \text{ or } recall \text{ or } Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$FDR = 1 - Specificity = \frac{FP}{TN + FP}$$

AUC measures the volume that the ROC curve is generating by
computing the sensitivity and 1-specificity by evaluating all possible threshold
values. The greater this area, the better the algorithm tends to be. The axis of a ROC

plot consists of the false positive rate (1- specificity, FPR) against the true positive rate (sensitivity, TPR). An excellent model has an AUC near one, meaning it has a good separability measure. A poor model has an AUC near 0, meaning it has the worst separability measure. It means it is reciprocating the result. It predicts 0s as 1s and 1s as 0s.

AUC - ROC curve is a performance measurement for classification problems at various threshold settings. ROC is a probability curve, and AUC represents the degree or measure of separability. It tells how much the model is capable of distinguishing between classes. The higher the AUC, the better the model predicts 0 classes as 0 and 1 classes as 1. By analogy, the Higher the AUC, the better the model is at distinguishing between patients with the disease and no disease. Furthermore, when AUC is 0.5, the model has no class separation capacity.

3.3 Experiment baseline (from paper and our result)

Pretrained with CheXNet Table 4 below shows the result of CheXNet Pretrained from paper and "Our_weight" training from scratch on DenseNets with ImageNet weight use for the based line on this study. The result is slightly different from the CheXNet paper due to the python dependencies environment and configurations that affect the training process.

Pathology	cheXNet	Our_weight
Atelectasis	0.8094	0.8107
Cardiomegaly	0.9248	0.8914
Effusion	0.8638	0.8776
Infiltration	0.7345	0.7152
Mass	0.8676	0.8455
Nodule	0.7802	0.7145
Pneumonia	0.7680	0.7721
Pneumothorax	0.8870	0.8761
Consolidation	0.7901	0.7986
Edema	0.8878	0.8878
Emphysema	0.9371	0.8967
Fibrosis	0.8047	0.7571
Pleural_Thickening	0.8062	0.7835
Hernia	0.9164	0.8727
mean auroc	0.8413	0.8214

Table 4: AUROC from CheXNet result with the test set to be Based line for comparison.

CHAPTER 4
RESULTS

After training on each source image from 6 different image enhancement techniques, in this chapter, we present the results of our performance of each image enhancement technique through fine-tuning the CheXNet model with Chest X-ray 14 dataset. We separately show the AUROC result on each image enhancement technique and the relative percentage difference.

4.1 Single Channel Image Enhancement Result

In the single channel image enhancement experiment, each image enhancement technique was fed to train five times to confirm the trend and performance is not occasionally result. The detail on the input image is 224x224 pixels from 1024x1024 pixels, and three channel is the layer of the processed image from the image enhancement on that technique. The performance metric use AUROC to compare with the performance from previous paper research, our weight reimplements, and the detail above in chapter 3. The result table below compares AUROC from each image enhancement technique and the AUROC from the original image model with the weight train by ourselves with reference AUROC from the previous research paper of CheXNet[7].

4.1.1 Gamma correction

Table 5 shows the AUROC performance result from gamma correction. The AUROC improve clearly on Nodule, Pneumonia, Plural Thickening, and Hernia. However, the performance in diagnosing Mass, Pneumothorax, Edema, and Fibrosis decreased. 5 repeated experiments confirmed that the result was in the same direction. The mean AUROC on the 5th experiment indicates the best overall performance of these studies on gamma correction—the performance reduction in other pathologies is nearly unchanged.

Pathology	Paper	myW	gam1	gam2	gam3	gam4	gam5	Percent increment				
								%gam1	%gam2	%gam3	%gam4	%gam5
Atelectasis	0.8094	0.810745	0.8144	0.8137	0.8111	0.8123	0.8103	0.4469	0.3706	0.0444	0.1962	-0.0545
Cardiomeg	0.9248	0.891407	0.8985	0.8958	0.8911	0.8938	0.8904	0.7921	0.4872	-0.0398	0.2665	-0.1129
Effusion	0.8638	0.877551	0.8790	0.8797	0.8779	0.8790	0.8801	0.1605	0.2490	0.0360	0.1694	0.2881
Infiltration	0.7345	0.715202	0.7177	0.7187	0.7151	0.7178	0.7138	0.3533	0.4893	-0.0167	0.3567	-0.1956
Mass	0.8676	0.845511	0.8435	0.8395	0.8456	0.8462	0.8479	-0.2363	0.7166	0.0116	0.0868	0.2829
Nodule	0.7802	0.714531	0.7333	0.7379	0.7392	0.7331	0.7361	2.6273	3.2684	3.4464	2.6043	3.0228
Pneumonia	0.768	0.772134	0.7817	0.7817	0.7739	0.7831	0.7877	1.2411	1.2411	0.2273	1.4238	2.0143
Pneumotho	0.887	0.87609	0.8733	0.8760	0.8703	0.8741	0.8793	-0.3239	-0.0142	0.6553	-0.2233	0.3691
Consolidati	0.7901	0.798645	0.7992	0.7991	0.7960	0.8008	0.8016	0.0648	0.0538	-0.3296	0.2649	0.3745
Edema	0.8878	0.887843	0.8852	0.8882	0.8860	0.8878	0.8877	-0.2976	0.0450	-0.2056	-0.0038	-0.0124
Emphysem	0.9371	0.896654	0.8971	0.9044	0.9011	0.9052	0.9012	0.0480	0.8619	0.4956	0.9530	0.5066
Fibrosis	0.8047	0.757111	0.7539	0.7456	0.7552	0.7502	0.7634	-0.4301	-1.5217	-0.2530	-0.9129	0.8337
Pleural_Thi	0.8062	0.78352	0.7918	0.7893	0.7893	0.7885	0.7912	1.0588	0.7393	0.7416	0.6394	0.9825
Hernia	0.9164	0.872689	0.8996	0.8770	0.8765	0.8819	0.8810	3.0782	0.4955	0.4423	1.0587	0.9535
mean auro	0.8413	0.8214	0.8263	0.8248	0.8235	0.8253	0.8266	0.5943	0.4082	0.2494	0.4726	0.6280

Table 5: AUROC result from Transfer Learning and fine-tuning with gamma correction compare with the AUROC from the original image model

4.1.2 Contrast-limited adaptive histogram equalization (CLAHE)

Table 6 shows the AUROC performance result from Contrast limited adaptive histogram equalization. The AUROC improves clearly on nodule, Pneumonia, Fibrosis, Plural Thickening, and Hernia. However, the performance in diagnosing Mass, Pneumothorax, and Edema decreased. 5 repeated experiments confirmed that the result was in the same direction. The mean AUROC in the 3rd experiment indicates the best overall performance of these studies on CLAHE. The reduction performance occurred only on Mass, Edema, and Pneumothorax.

Patho	Paper	myW	clahe1	clahe2	clahe3	clahe4	clahe5	Percent increment				
								%clahe1	%clahe2	%clahe3	%clahe4	%clahe5
Atelect	0.8094	0.810745	0.8135	0.8118	0.8100	0.8139	0.8123	0.3446	0.1280	0.0884	0.3874	0.1960
Cardio	0.9248	0.891407	0.8967	0.8916	0.8966	0.8909	0.8920	0.5894	0.0219	0.5820	0.0619	0.0689
Effusio	0.8638	0.877551	0.8787	0.8796	0.8787	0.8774	0.8796	0.1314	0.2327	0.1366	0.0148	0.2372
Infiltra	0.7345	0.715202	0.7117	0.7145	0.7152	0.7125	0.7180	0.4943	0.0917	0.0019	0.3747	0.3851
Mass	0.8676	0.845511	0.8388	0.8350	0.8393	0.8379	0.8365	0.7972	1.2391	0.7397	0.8993	1.0644
Nodule	0.7802	0.714531	0.7468	0.7455	0.7482	0.7497	0.7451	4.5151	4.3300	4.7177	4.9245	4.2725
Pneum	0.768	0.772134	0.7806	0.7844	0.7791	0.7845	0.7811	1.0976	1.5833	0.9039	1.6006	1.1644
Pneum	0.887	0.87609	0.8686	0.8735	0.8742	0.8742	0.8707	0.8585	0.2925	0.2141	0.2103	0.6191
Consol	0.7901	0.798645	0.8005	0.7993	0.7983	0.8033	0.8032	0.2277	0.0768	0.0417	0.5822	0.5696
Edema	0.8878	0.887843	0.8772	0.8828	0.8806	0.8820	0.8816	1.2039	0.5697	0.8214	0.6609	0.7065
Emphy	0.9371	0.896654	0.8985	0.9005	0.9068	0.8953	0.9038	0.2067	0.4303	1.1344	0.1529	0.7953
Fibrosi	0.8047	0.757111	0.7655	0.7675	0.7616	0.7632	0.7662	1.1080	1.3711	0.5929	0.8018	1.2061
Pleural	0.8062	0.78352	0.7925	0.7969	0.7861	0.7913	0.7921	1.1514	1.7018	0.3350	0.9893	1.0950
Hernia	0.9164	0.872689	0.8806	0.8665	0.8796	0.8715	0.8618	0.9088	0.7130	0.7928	0.1305	1.2448
mean	0.8413	0.8214	0.8250	0.8249	0.8253	0.8248	0.8246	0.4388	0.4317	0.4767	0.4172	0.3860

Table 6: AUROC result from Transfer Learning and fine-tuning with Contrast limited adaptive histogram equalization compared with the AUROC from the original image model

52

4.1.3 Histogram Equalization

Table 7 shows the AUROC performance result from Histogram Equalization. The AUROC improves clearly on Cardiomegaly, Mass, Nodule, and Plural Thickening. However, the performance in diagnosing other pathology decreases, as the red color shows in the table. Five repeated experiments confirmed that the result was in the same direction. The mean AUROC in the 1st experiment indicates the best overall performance of these studies on Histogram Equalization. The performance reduction on the overall pathology improved, but the Nodule detection performance slightly decreased from the four remaining experiments.

Pathol	Paper	myW	he1	he2	he3	he4	he5	%he1	%he2	%he3	%he4	%he5
Atelecta	0.8094	0.810745	0.8090	0.8050	0.8108	0.8086	0.8121	-0.2112	-0.7072	0.0114	-0.2688	0.1669
Cardiom	0.9248	0.891407	0.8975	0.9014	0.9019	0.9027	0.9025	0.6875	1.1173	1.1758	1.2655	1.2478
Effusion	0.8638	0.877551	0.8759	0.8788	0.8756	0.8769	0.8775	-0.1851	0.1417	-0.2202	-0.0688	-0.0047
Infiltrati	0.7345	0.715202	0.7148	0.7148	0.7147	0.7159	0.7157	-0.0581	-0.0532	-0.0724	0.0986	0.0749
Mass	0.8676	0.845511	0.8494	0.8496	0.8488	0.8524	0.8492	0.4600	0.4836	0.3847	0.8098	0.4306
Nodule	0.7802	0.714531	0.7349	0.7410	0.7343	0.7422	0.7381	2.8451	3.6986	2.7645	3.8681	3.3038
Pneumo	0.768	0.772134	0.7730	0.7673	0.7710	0.7790	0.7665	0.1176	-0.5288	-0.1433	0.8921	-0.7286
Pneumo	0.887	0.87609	0.8739	0.8669	0.8717	0.8671	0.8711	-0.2514	-1.0517	-0.5030	-1.0220	-0.5718
Consolic	0.7901	0.798645	0.7981	0.7969	0.7937	0.7980	0.7951	-0.0667	-0.2194	-0.6153	-0.0852	-0.4481
Edema	0.8878	0.887843	0.8832	0.8807	0.8840	0.8813	0.8819	-0.5280	-0.8100	-0.4381	-0.7400	-0.6660
Emphys	0.9371	0.896654	0.8931	0.8911	0.8986	0.8938	0.8951	-0.3941	-0.6150	0.2212	-0.3228	-0.1718
Fibrosis	0.8047	0.757111	0.7589	0.7490	0.7467	0.7461	0.7427	0.2302	-1.0691	-1.3740	-1.4498	-1.8978
Pleural_	0.8062	0.78352	0.7928	0.7912	0.7920	0.7885	0.7902	1.1825	0.9802	1.0792	0.6396	0.8546
Hernia	0.9164	0.872689	0.8695	0.8510	0.8721	0.8657	0.8449	-0.3686	-2.4848	-0.0689	-0.8001	-3.1894
mean a	0.8413	0.8214	0.8231	0.8203	0.8226	0.8227	0.8202	0.2116	-0.1306	0.1413	0.1611	-0.1474

Table 7: AUROC result from Transfer Learning and fine-tuning with Histogram Equalization compare with the AUROC from the original image model

4.1.4 Balance Contrast Enhancement Technique (BCET)

Table 8 shows the AUROC performance result from Balance Contrast Enhancement Technique. The AUROC improve clearly on Nodule, Plural Thickening, and Hernia. However, the performance in diagnosing the rest of the pathology decreases, as the red color shows in the table. Five repeated experiments confirmed that the result was in the same direction. The mean AUROC on the 5th experiment indicate the best overall performance of these study on BCET even if it has no improvement compared with the performance without using the image enhancement technique. The most improvement from BCET pathology is Hernia when comparing all six image Enhancement Techniques.

Patho	Paper	myW	bcet1	bcet2	bcet3	bcet4	bcet5	%bcet1	%bcet2	%bcet3	%bcet4	%bcet5
Atelect	0.8094	0.810745	0.8068	0.8000	0.8024	0.8014	0.8045	-0.4895	-1.3204	-1.0251	-1.1553	-0.7758
Cardio	0.9248	0.891407	0.8879	0.8918	0.8863	0.8873	0.8828	-0.3941	0.0392	-0.5727	-0.4657	-0.9663
Effusio	0.8638	0.877551	0.8726	0.8733	0.8719	0.8733	0.8723	-0.5665	-0.4873	-0.6384	-0.4842	-0.6001
Infiltra	0.7345	0.715202	0.7096	0.7112	0.7126	0.7089	0.7097	-0.7825	-0.5575	-0.3614	-0.8770	-0.7627
Mass	0.8676	0.845511	0.8325	0.8347	0.8352	0.8347	0.8349	-1.5423	-1.2829	-1.2228	-1.2767	-1.2547
Nodule	0.7802	0.714531	0.7275	0.7307	0.7398	0.7355	0.7379	1.8134	2.2640	3.5428	2.9384	3.2673
Pneum	0.768	0.772134	0.7661	0.7676	0.7712	0.7735	0.7826	-0.7753	-0.5874	-0.1182	0.1714	1.3501
Pneum	0.887	0.87609	0.8589	0.8681	0.8658	0.8611	0.8604	-1.9640	-0.9107	-1.1705	-1.7090	-1.7951
Consol	0.7901	0.798645	0.7931	0.7930	0.7958	0.7972	0.7940	-0.6915	-0.7029	-0.3622	-0.1805	-0.5775
Edema	0.8878	0.887843	0.8793	0.8746	0.8806	0.8792	0.8787	-0.9579	-1.4898	-0.8136	-0.9774	-1.0295
Emphy	0.9371	0.896654	0.8790	0.8846	0.8849	0.8815	0.8878	-1.9742	-1.3455	-1.3054	-1.6928	-0.9902
Fibrosi	0.8047	0.757111	0.7448	0.7461	0.7512	0.7541	0.7600	-1.6315	-1.4605	-0.7754	-0.4037	0.3847
Pleural	0.8062	0.78352	0.7864	0.7943	0.7880	0.7901	0.7896	0.3729	1.3735	0.5731	0.8389	0.7802
Hernia	0.9164	0.872689	0.9015	0.8897	0.8992	0.8987	0.9031	3.3005	1.9446	3.0419	2.9839	3.4842
mean	0.8413	0.8214	0.8176	0.8185	0.8204	0.8197	0.8213	-0.4668	-0.3481	-0.1256	-0.2020	-0.0121

Table 8: AUROC result from Transfer Learning and fine-tuning with Balance Contrast Enhancement Technique compare with the AUROC from the original image model

4.1.5 Min Max Linear Contrast Stretching (MMCS)

Table 9 shows the AUROC performance result from Min-Max Contrast Stretching. The AUROC improves clearly on nodules, Pneumonia, and Hernia. However, in some diagnoses, some pathologies were slightly reduced. 5 repeated experiments confirmed the result was in the same direction in most pathology. Nevertheless, pathology differences showed in the 3rd experiment; overall performance decreased, and the Emphysema and Hernia reduced in different directions compared with the remaining experiments. The mean AUROC on the 5th experiment indicates the best overall performance of these studies on MMCS. The performance reduction in other pathologies is close to unchanged.

Pathol	Paper	myW	mmcs1	mmcs2	mmcs3	mmcs4	mmcs5	%mmcs1	%mmcs2	%mmcs3	%mmcs4	%mmcs5
Atelecta	0.8094	0.810745	0.8141	0.8149	0.8116	0.8140	0.8121	0.4111	0.5110	0.1093	0.3975	0.1732
Cardiom	0.9248	0.891407	0.8965	0.8911	0.8973	0.9008	0.8987	0.5720	-0.0331	0.6582	1.0578	0.8174
Effusion	0.8638	0.877551	0.8796	0.8800	0.8780	0.8780	0.8786	0.2309	0.2814	0.0469	0.0542	0.1185
Infiltrati	0.7345	0.715202	0.7155	0.7142	0.7163	0.7171	0.7172	0.0410	-0.1411	0.1495	0.2668	0.2730
Mass	0.8676	0.845511	0.8425	0.8477	0.8483	0.8413	0.8459	0.3527	0.2569	0.3327	0.4958	0.0510
Nodule	0.7802	0.714531	0.7398	0.7379	0.7337	0.7381	0.7379	3.5432	3.2667	2.6795	3.3022	3.2709
Pneumc	0.768	0.772134	0.7780	0.7824	0.7740	0.7776	0.7809	0.7643	1.3323	0.2460	0.7082	1.1332
Pneumc	0.887	0.87609	0.8728	0.8724	0.8727	0.8710	0.8753	0.3737	0.4198	0.3865	0.5828	-0.0936
Consolic	0.7901	0.798645	0.7982	0.8006	0.7996	0.7966	0.8007	-0.0526	0.2435	0.1192	0.2515	0.2550
Edema	0.8878	0.887843	0.8862	0.8873	0.8852	0.8845	0.8853	-0.1850	-0.0560	-0.2947	-0.3776	-0.2839
Emphys	0.9371	0.896654	0.8999	0.9030	0.8985	0.8984	0.9047	0.3598	0.7102	0.2022	0.1999	0.8987
Fibrosis	0.8047	0.757111	0.7544	0.7639	0.7474	0.7602	0.7582	0.3553	0.8942	-1.2846	0.4100	0.1405
Pleural_	0.8062	0.78352	0.7865	0.7877	0.7878	0.7865	0.7823	0.3793	0.5398	0.5513	0.3791	-0.1495
Hernia	0.9164	0.872689	0.8783	0.8772	0.8637	0.8881	0.8959	0.6381	0.5160	1.0321	1.7657	2.6588
mean a	0.8413	0.8214	0.8245	0.8257	0.8224	0.8252	0.8267	0.3716	0.5283	0.1256	0.4584	0.6442

Table 9: AUROC result from Transfer Learning and fine-tuning with Min-Max Contrast Stretching compared with the AUROC from the original image model

4.1.6 Image Invert/ Complement

Table 10 shows the AUROC performance result from Complement. The AUROC improves clearly on Nodule, Cardiomegaly, Pneumonia, and Fibrosis. However, some pathology diagnoses decreased as red color showed in Pneumothorax, Consolidation, Edema, and Emphysema. 5 repeated experiments confirmed the result was in the same direction in most pathology. The mean AUROC on the 4th experiment indicates the best overall performance of these studies on Complement—the performance reduction in other pathologies is nearly unchanged.

Pathology	Paper	myW	invert1	invert2	invert3	invert4	invert5	Percent increment %invert1	%invert2	%invert3	%invert4	%invert5
Atelectasis	0.8094	0.810745	0.8091	0.8093	0.8114	0.8104	0.8073	0.1986	0.1784	0.0837	-0.0414	-0.4244
Cardiomeg	0.9248	0.891407	0.9023	0.8978	0.8997	0.8999	0.8960	1.2188	0.7167	0.9308	0.9558	0.5117
Effusion	0.8638	0.877551	0.8779	0.8793	0.8812	0.8810	0.8785	0.0350	0.2034	0.4200	0.3952	0.1084
Infiltration	0.7345	0.715202	0.7194	0.7132	0.7150	0.7148	0.7158	0.5915	0.2788	-0.0339	-0.0493	0.0829
Mass	0.8676	0.845511	0.8494	0.8368	0.8404	0.8494	0.8472	0.4643	1.0325	0.6069	0.4579	0.1971
Nodule	0.7802	0.714531	0.7359	0.7404	0.7353	0.7334	0.7380	2.9921	3.6258	2.9004	2.6356	3.2779
Pneumonia	0.768	0.772134	0.7810	0.7720	0.7876	0.7842	0.7734	1.1431	-0.0136	2.0010	1.5655	0.1679
Pneumotho	0.887	0.87609	0.8688	0.8732	0.8727	0.8755	0.8735	0.8336	0.3254	0.3919	-0.0715	-0.2924
Consolidati	0.7901	0.798645	0.7944	0.7980	0.7947	0.7954	0.7956	0.5285	0.0866	0.4953	-0.4001	-0.3819
Edema	0.8878	0.887843	0.8838	0.8850	0.8861	0.8881	0.8884	0.4534	0.3221	0.1998	0.0294	0.0673
Emphysem	0.9371	0.896654	0.8951	0.8945	0.8892	0.9024	0.8974	0.1777	0.2354	0.8292	0.6455	0.0809
Fibrosis	0.8047	0.757111	0.7673	0.7604	0.7669	0.7588	0.7556	1.3502	0.4335	1.2953	0.2205	-0.1984
Pleural_Thi	0.8062	0.78352	0.7831	0.7824	0.7867	0.7845	0.7913	-0.0537	0.1372	0.4107	0.1263	0.9987
Hernia	0.9164	0.872689	0.8715	0.8701	0.8742	0.8770	0.8792	0.1354	0.2997	0.1728	0.4931	0.7493
mean auro	0.8413	0.8214	0.8242	0.8223	0.8244	0.8254	0.8241	0.3426	0.1120	0.3600	0.4807	0.3272

Table 10: AUROC result from Transfer Learning and fine-tuning with Complement compare with the AUROC from the original image model

4.1.7 Single Channel image enhancement result comparison

We observe pathologies improve differently depending on each image enhancement technique. Based on the selection of the best mean AUROC from each Single Channel image enhancement model training, the comparison result is shown in table 11. The best overall can use mean AUROC to indicate from the table; MMCS and Gamma correction is the best overall performance on 14 pathologies from these single channel input training experiments. For the pathology example of a Nodule alone, the CLAHE is the best technique to improve disease detection by 3.37 percent. The second best for Nodule detection is the Min Max Contrast Stretching technique, as shown in Table 11.

Moreover, The BCET and Gamma correction improve Hernia detection. However, The Balance Contrast Enhancement Technique decreases the performance of the remaining pathology except for Nodule and Hernia. The best improvement in Pneumonia detection used Gamma correction.

Each pathology detection base on each image enhancement technique, from figure 18, shows the potential performance for the specific uses for one class of pathology detection. When observing different perspectives from all the experiments on image enhancement techniques specific to the disease, the disease which significantly improves is Cardiomegaly, Nodule, Pneumonia, Emphysema, Fibrosis, Pleural Thickening, and Hernia. Cardiomegaly specifically chooses the Histogram Equalization technique to maximize improvement from these experiments. Nodule Fibrosis Emphysema and Pleural Thickening chose the Contrast limited adaptive histogram equalization; Pneumonia used the Gamma correction technique, and Hernia chose the Balance Contrast Enhancement Technique. The best weight trained from the experiment is shown in figures 19-25, according to each pathology.

Pathology	Paper	myW	mmcs5	gam5	invert4	clahe3	he1	bcet5	Percent Increment					
									mmcs5	gam5	invert4	clahe3	he1	bcet5
Atelectasis	0.8094	0.810745	0.8121	0.8103	0.8104	0.8100	0.8090	0.8045	0.1732	-0.0545	-0.0414	-0.0884	-0.2112	-0.7758
Cardiomegaly	0.9248	0.891407	0.8987	0.8904	0.8999	0.8966	0.8975	0.8828	0.8174	-0.1129	0.9558	0.5820	0.6875	-0.9663
Effusion	0.8638	0.877551	0.8786	0.8801	0.8810	0.8787	0.8759	0.8723	0.1185	0.2881	0.3952	0.1366	-0.1851	-0.6001
Infiltration	0.7345	0.715202	0.7172	0.7138	0.7148	0.7152	0.7148	0.7097	0.2730	-0.1956	-0.0493	0.0019	-0.0581	-0.7627
Mass	0.8676	0.845511	0.8459	0.8479	0.8494	0.8393	0.8494	0.8349	0.0510	0.2829	0.4579	-0.7397	0.4600	-1.2547
Nodule	0.7802	0.714531	0.7379	0.7361	0.7334	0.7482	0.7349	0.7379	3.2709	3.0228	2.6356	4.7177	2.8451	3.2673
Pneumonia	0.768	0.772134	0.7809	0.7877	0.7842	0.7791	0.7730	0.7826	1.1332	2.0143	1.5655	0.9039	0.1176	1.3501
Pneumothorax	0.887	0.87609	0.8753	0.8793	0.8755	0.8742	0.8739	0.8604	-0.0936	0.3691	-0.0715	-0.2141	-0.2514	-1.7951
Consolidation	0.7901	0.798645	0.8007	0.8016	0.7954	0.7983	0.7981	0.7940	0.2550	0.3745	-0.4001	-0.0417	-0.0667	-0.5775
Edema	0.8878	0.887843	0.8853	0.8877	0.8881	0.8806	0.8832	0.8787	-0.2839	-0.0124	0.0294	-0.8214	-0.5280	-1.0295
Emphysema	0.9371	0.896654	0.9047	0.9012	0.9024	0.9068	0.8931	0.8878	0.8987	0.5066	0.6455	1.1344	-0.3941	-0.9902
Fibrosis	0.8047	0.757111	0.7582	0.7634	0.7588	0.7616	0.7589	0.7600	0.1405	0.8337	0.2205	0.5929	0.2302	0.3847
Pleural_Thickening	0.8062	0.78352	0.7823	0.7912	0.7845	0.7861	0.7928	0.7896	-0.1495	0.9825	0.1263	0.3350	1.1825	0.7802
Hernia	0.9164	0.872689	0.8959	0.8810	0.8770	0.8796	0.8695	0.9031	2.6588	0.9535	0.4931	0.7928	-0.3686	3.4842
mean auroc	0.8413	0.8214	0.8267	0.8266	0.8254	0.8253	0.8231	0.8213	0.6442	0.6280	0.4807	0.4767	0.2116	-0.0121

Table 11: AUROC comparison from Transfer Learning and fine-tuning with six image enhancement techniques and the AUROC from the original image base on the best overall AUROC (mean AUROC).
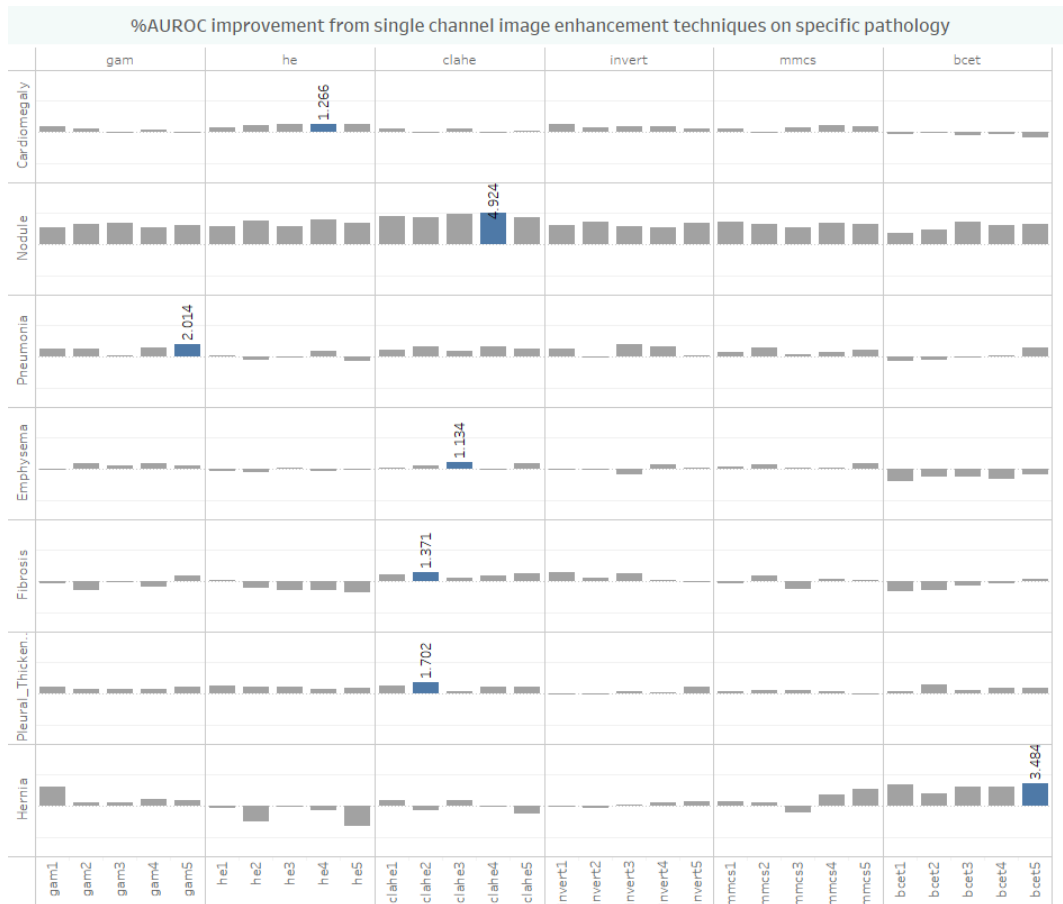
Figure 18: The percentage of AUROC improvement from image enhancement
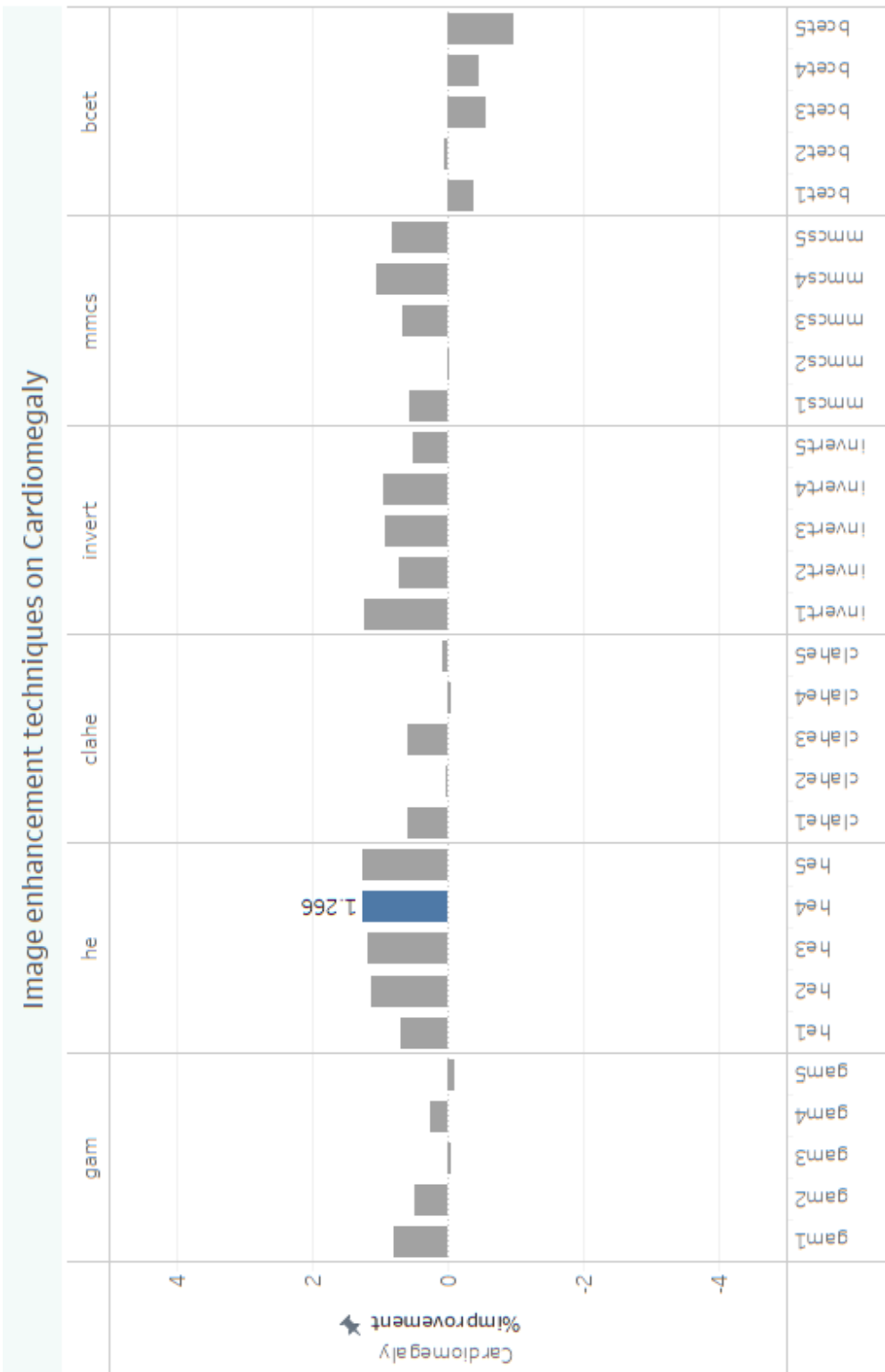techniques on the specific pathology

Figure 19: The percentage of AUROC improvement from image enhancement techniques on Cardiomegaly.
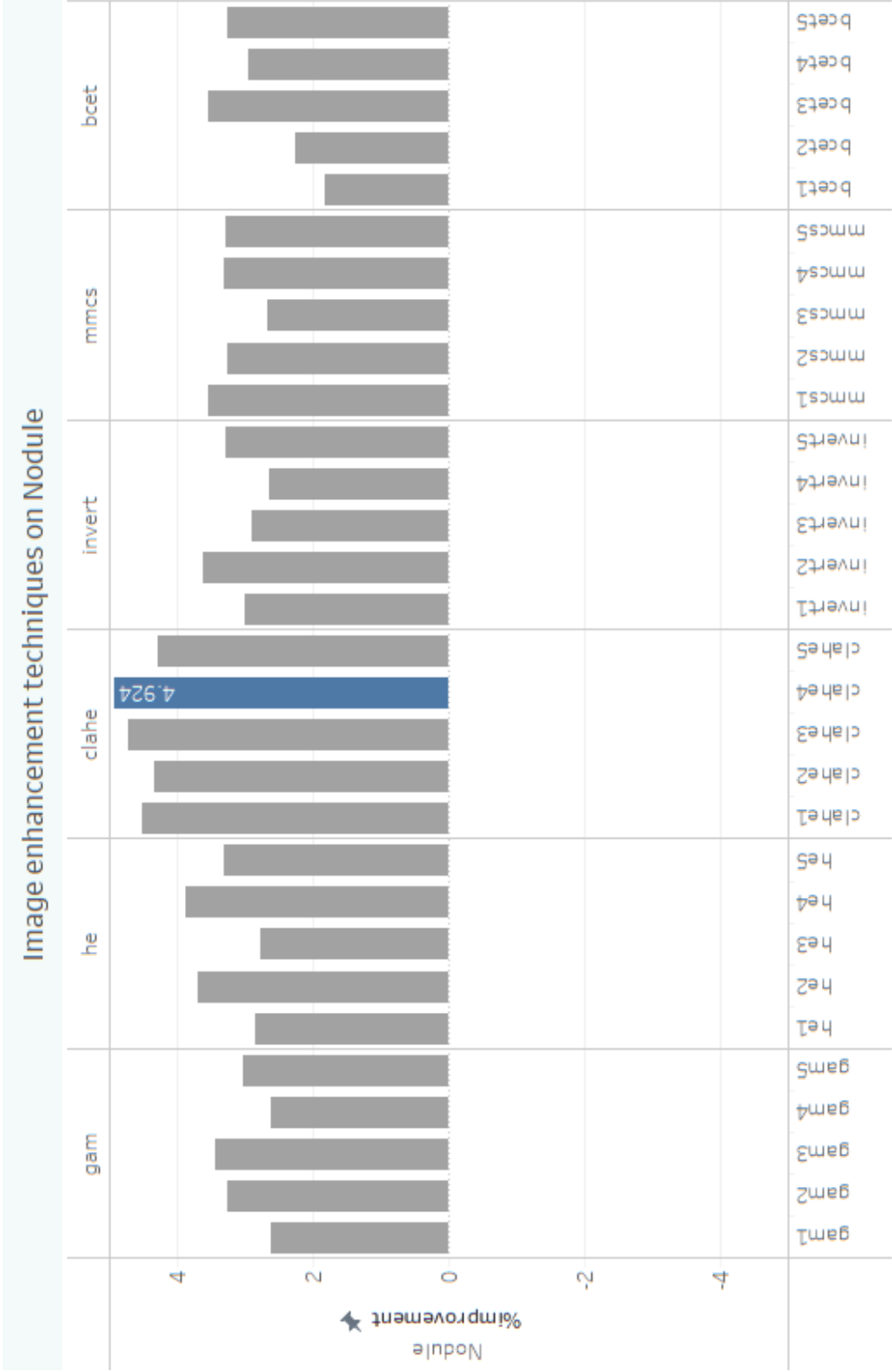
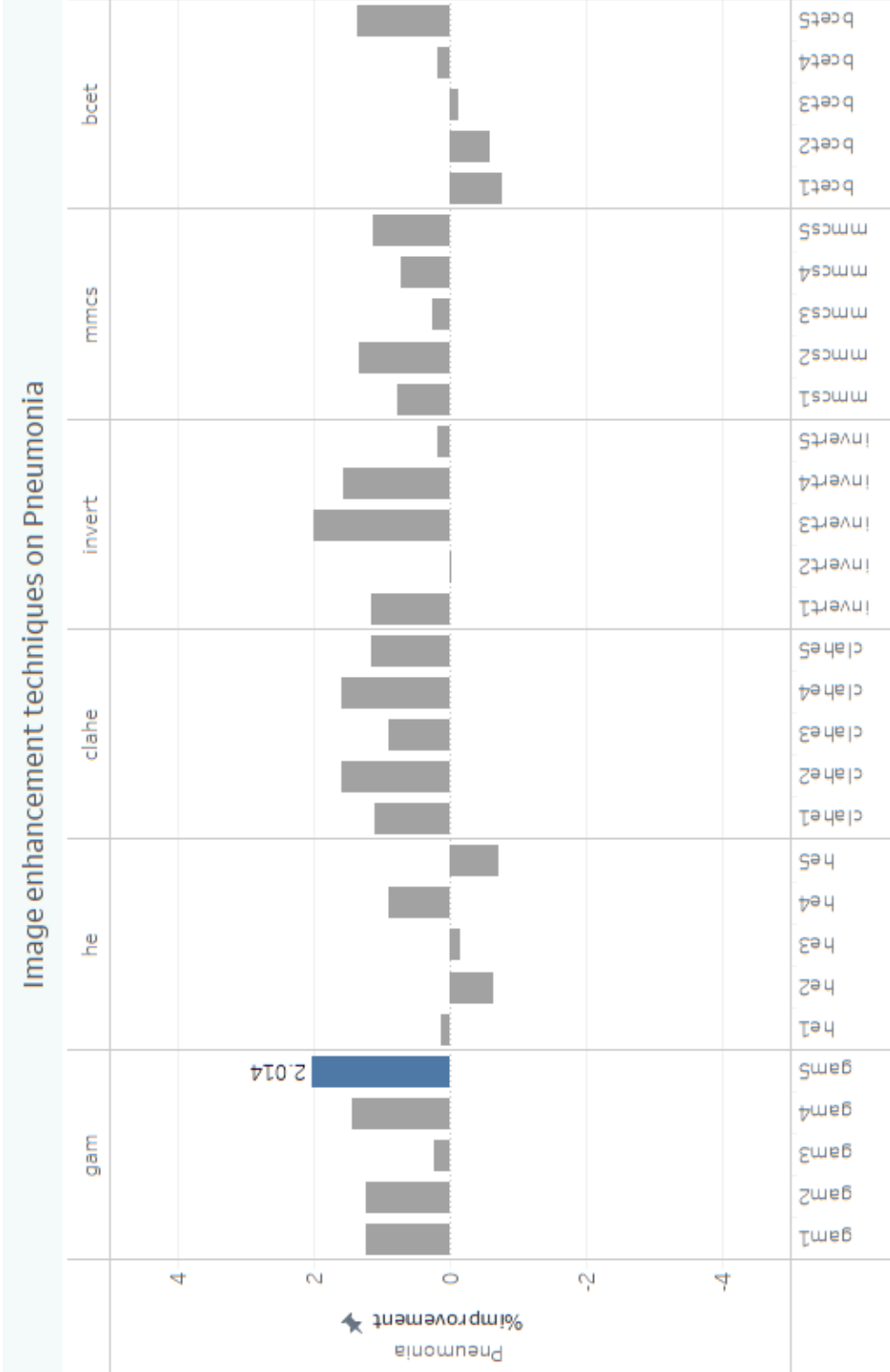Figure 20: The percentage of AUROC improvement from image enhancement techniques on Nodule.

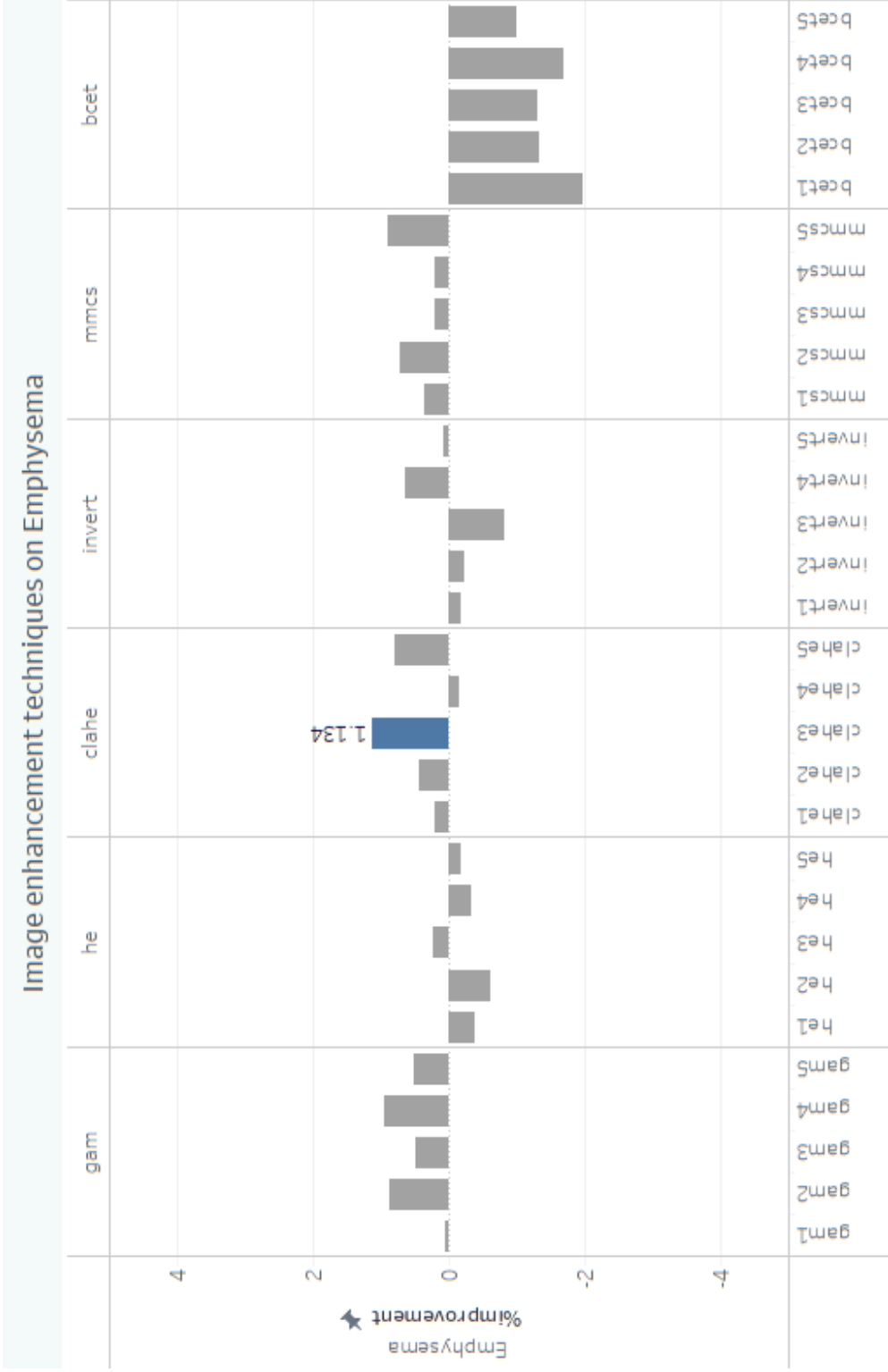Figure 21: The percentage of AUROC improvement from image enhancement techniques on Pneumonia.

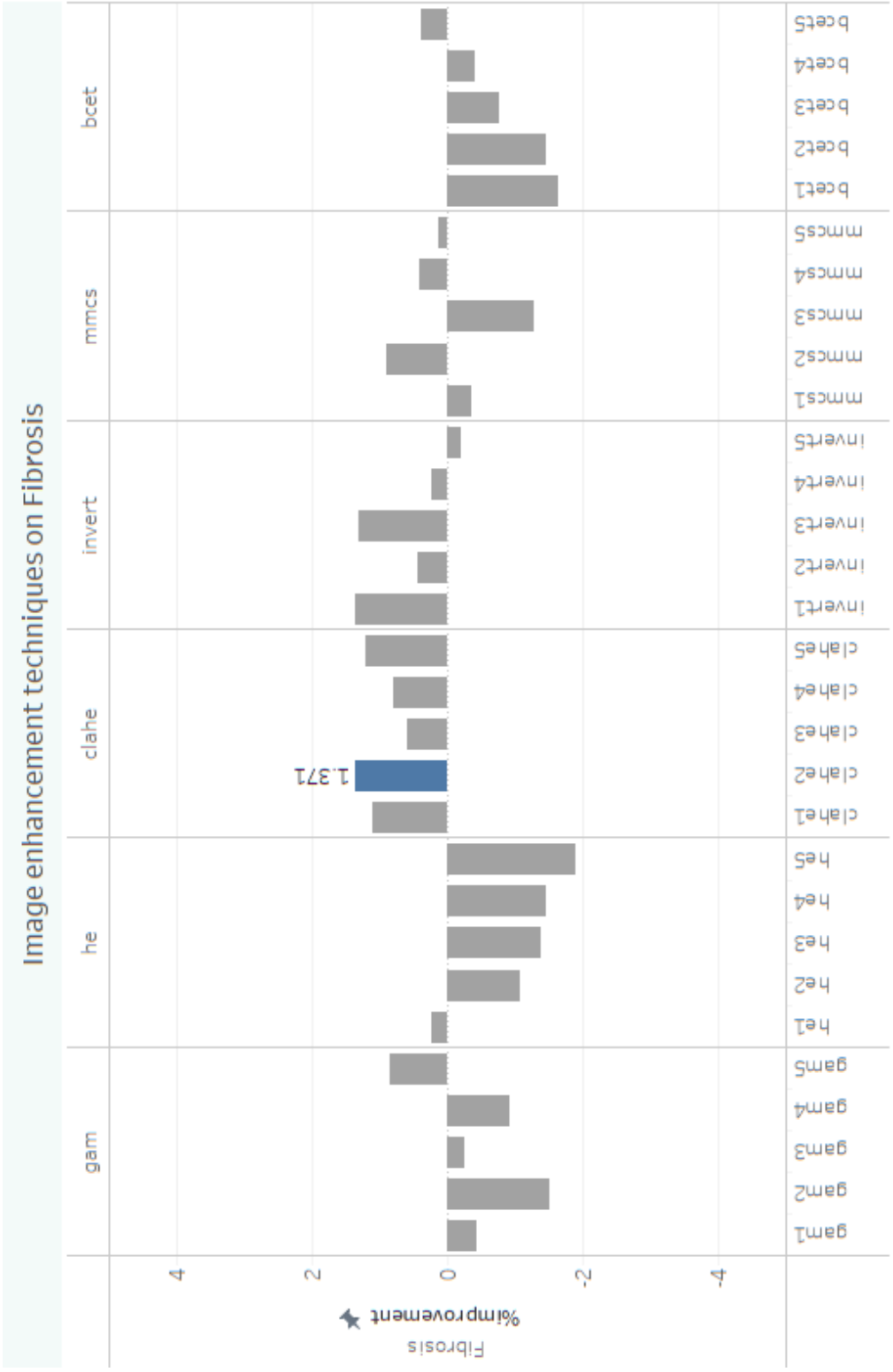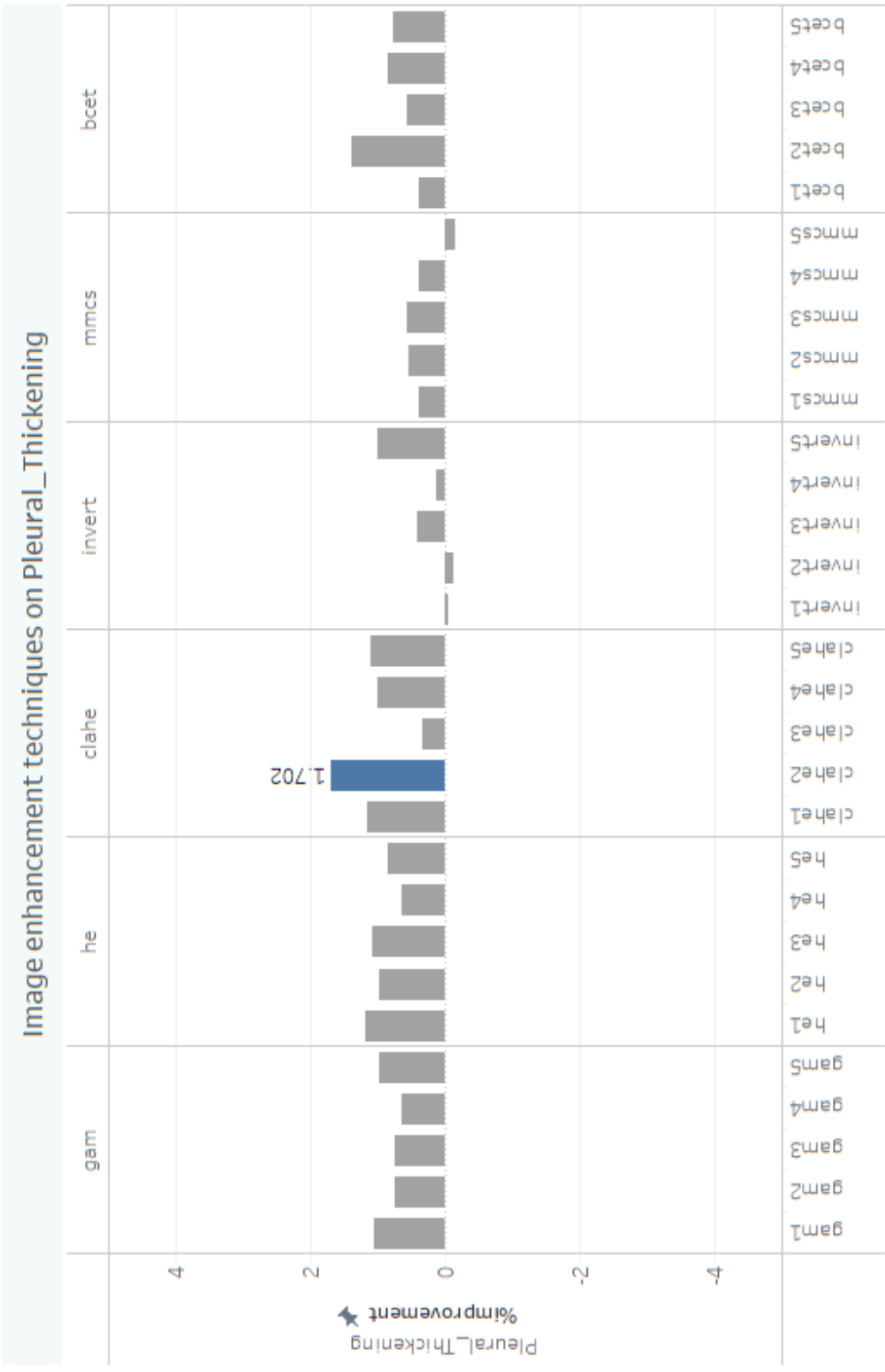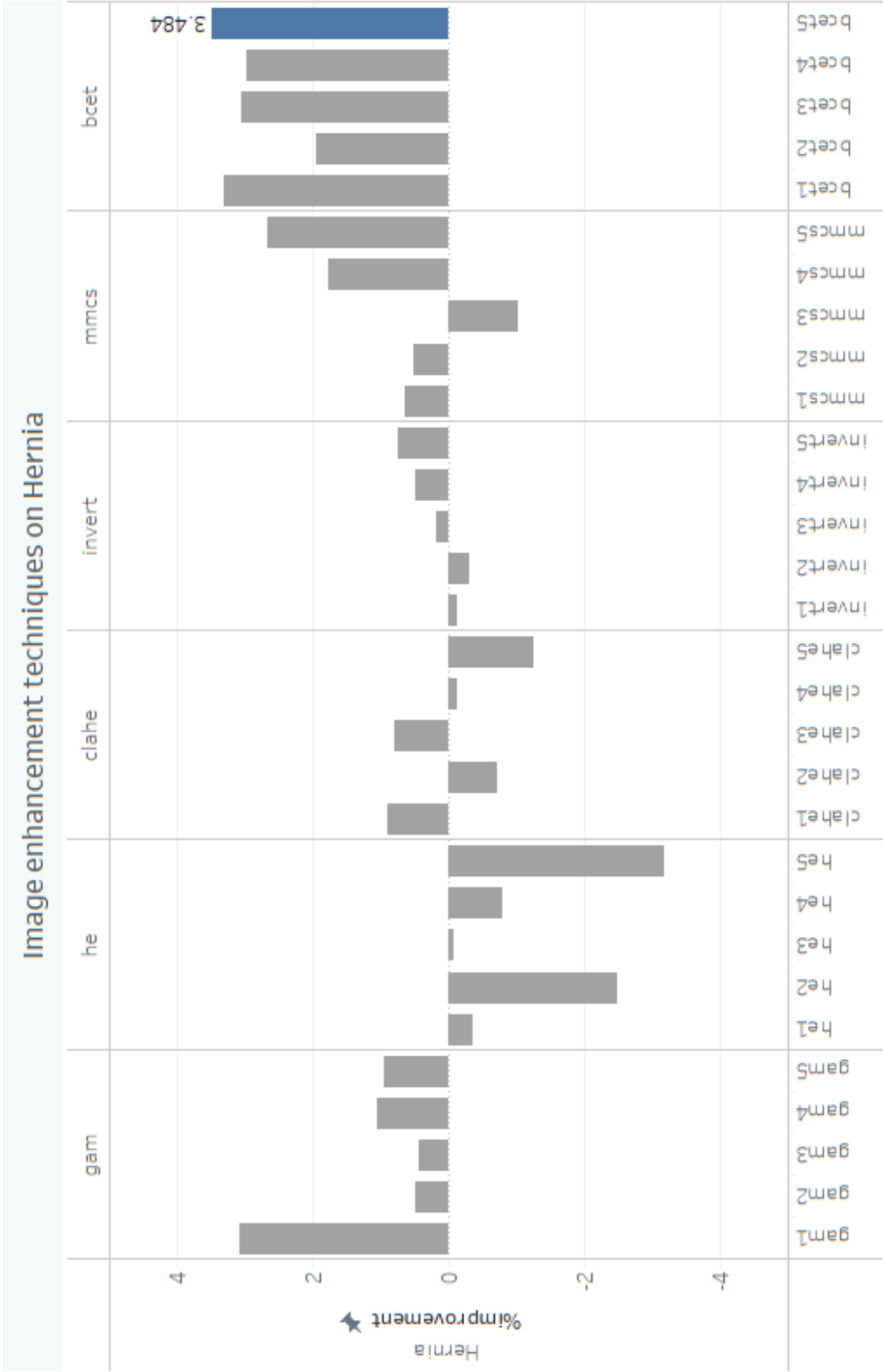Figure 22: The percentage of AUROC improvement from image enhancement techniques on Emphysema.

Figure 23: The percentage of AUROC improvement from image enhancement techniques on Fibrosis.

Figure 24: The percentage of AUROC improvement from image enhancement techniques on Pleural Thickening.

Figure 25: The percentage of AUROC improvement from image enhancement techniques on Hernia.

4.2 Multi-Channels Image Enhancement Result

The channel of each image enhancement technique will obtain and stack to a new image with three different channels. The three-channel consists of the original dimension, the first image enhancement technique, and the second. We found that the exclusion of the original image causes the model to be mistakenly realized into different images, and the performance results in a drastic decrease. Therefore, training on multi-channel input must include the original channel with two other combinations of image enhancement techniques.

4.2.1 Original image + Gamma correction + Invert

Table 12 shows the AUROC performance result from the combination. The AUROC improves clearly on nodules, Pneumonia, and Hernia. However, the performance in diagnosing the rest of the pathology decreases, as shown in the red color. The five repeated experiments confirm that the result was in the same direction. The mean AUROC on the 3rd experiment is 0.4243 percent increment, indicating the best overall performance on the combination.

The most improvement from this combination is Pneumonia. A percent increment of 1.924 is the best improvement on Pneumonia, specifically both Single Channel and Multi-Channel Image Enhancement.

	paper	myW	ogi1	ogi2	ogi3	ogi4	ogi5	Percent increment %ogi1	%ogi2	%ogi3	%ogi4	%ogi5
Atelectasis	0.8094	0.8107	0.8102	0.8089	0.8112	0.8111	0.8087	-0.0682	-0.2286	0.0604	0.0422	-0.2493
Cardiomegaly	0.9248	0.8914	0.8974	0.8939	0.8929	0.8935	0.8926	0.6739	0.2765	0.1639	0.2311	0.1368
Effusion	0.8638	0.8776	0.8791	0.8785	0.8775	0.8776	0.8803	0.1728	0.1116	-0.0099	0.0061	0.3135
Infiltration	0.7345	0.7152	0.7149	0.7148	0.7162	0.7197	0.7183	-0.0489	-0.0618	0.1332	0.6357	0.4323
Mass	0.8676	0.8455	0.8427	0.8456	0.8524	0.8473	0.8451	-0.3330	0.0143	0.8134	0.2127	-0.0513
Nodule	0.7802	0.7145	0.7324	0.7369	0.7340	0.7354	0.7337	2.5021	3.1298	2.7224	2.9196	2.6880
Pneumonia	0.7680	0.7721	0.7831	0.7874	0.7914	0.7839	0.7810	1.4214	1.9824	2.4920	1.5228	1.1450
Pneumothorax	0.8870	0.8761	0.8694	0.8709	0.8721	0.8685	0.8723	-0.7656	-0.5977	-0.4503	-0.8640	-0.4270
Consolidation	0.7901	0.7986	0.8004	0.8026	0.8005	0.8005	0.8015	0.2191	0.4930	0.2268	0.2347	0.3625
Edema	0.8878	0.8878	0.8895	0.8868	0.8883	0.8865	0.8875	0.1835	-0.1218	0.0532	-0.1532	-0.0431
Emphysema	0.9371	0.8967	0.9034	0.9010	0.8995	0.8955	0.8914	0.7495	0.4803	0.3228	-0.1261	-0.5830
Fibrosis	0.8047	0.7571	0.7570	0.7477	0.7517	0.7561	0.7539	-0.0118	-1.2367	-0.7154	-0.1286	-0.4275
Pleural_Thickening	0.8062	0.7835	0.7823	0.7798	0.7879	0.7863	0.7908	-0.1535	-0.4802	0.5633	0.3491	0.9227
Hernia	0.9164	0.8727	0.8809	0.8776	0.8835	0.8851	0.8800	0.9378	0.5571	1.2351	1.4200	0.8327
mean auroc	0.8413	0.8214	0.8245	0.8237	0.8256	0.8248	0.8241	0.3734	0.2835	0.5165	0.4121	0.3257

Table 12: AUROC result from Transfer Learning and fine-tuning with gamma correction and invert combination compare with the AUROC from the original image model

4.2.2 Original image + Gamma correction + HE

Table 13 shows the AUROC performance result from the combination. The AUROC improves clearly on Nodules, Pneumonia, and Hernia. However, the performance in diagnosing the rest of the pathology decreases as the red. Significant reduction falls into Pneumothorax the most. The five repeated experiments confirm that the result was in the same direction. The mean AUROC in the 3rd experiment is a 0.5901 percent increment from the baseline, which indicates the best overall performance in this combination.

	paper	myW	ogh1	ogh2	ogh3	ogh4	ogh5	Percent increment %ogh1	%ogh2	%ogh3	%ogh4	%ogh5
Atelectasis	0.8094	0.8107	0.8102	0.8128	0.8134	0.8132	0.8139	-0.0682	0.2500	0.3236	0.3072	0.3845
Cardiomegaly	0.9248	0.8914	0.8974	0.8941	0.9012	0.8947	0.8977	0.6739	0.3005	1.0985	0.3688	0.7065
Effusion	0.8638	0.8776	0.8791	0.8769	0.8772	0.8778	0.8761	0.1728	-0.0762	-0.0368	0.0267	-0.1642
Infiltration	0.7345	0.7152	0.7149	0.7165	0.7174	0.7169	0.7176	-0.0489	0.1798	0.3135	0.2437	0.3317
Mass	0.8676	0.8455	0.8427	0.8409	0.8490	0.8435	0.8466	-0.3330	-0.5398	0.4113	-0.2434	0.1304
Nodule	0.7802	0.7145	0.7324	0.7402	0.7360	0.7377	0.7398	2.5021	3.5905	3.0026	3.2377	3.5353
Pneumonia	0.7680	0.7721	0.7831	0.7754	0.7808	0.7839	0.7738	1.4214	0.4276	1.1282	1.5175	0.2112
Pneumothorax	0.8870	0.8761	0.8694	0.8727	0.8747	0.8720	0.8696	-0.7656	-0.3912	-0.1535	-0.4713	-0.7387
Consolidation	0.7901	0.7986	0.8004	0.7995	0.7982	0.8045	0.7973	0.2191	0.1091	-0.0599	0.7358	-0.1745
Edema	0.8878	0.8878	0.8895	0.8839	0.8892	0.8877	0.8873	0.1835	-0.4465	0.1511	-0.0178	-0.0590
Emphysema	0.9371	0.8967	0.9034	0.8963	0.9026	0.8932	0.8982	0.7495	-0.0341	0.6655	-0.3885	0.1691
Fibrosis	0.8047	0.7571	0.7570	0.7622	0.7579	0.7553	0.7532	-0.0118	0.6704	0.1092	-0.2380	-0.5125
Pleural_Thickening	0.8062	0.7835	0.7823	0.7849	0.7860	0.7919	0.7880	-0.1535	0.1798	0.3110	1.0688	0.5758
Hernia	0.9164	0.8727	0.8809	0.8772	0.8986	0.8774	0.8886	0.9378	0.5160	2.9670	0.5354	1.8237
mean auroc	0.8413	0.8214	0.8245	0.8238	0.8273	0.8250	0.8248	0.3734	0.2946	0.7185	0.4339	0.4175

Table 13: AUROC result from Transfer Learning and fine-tuning with gamma correction and Histogram Equalization combination compare with the AUROC from the original image model

4.2.3 Original image + Gamma correction + CLAHE

Table 14 shows the AUROC performance result from the combination. The AUROC improves clearly on Cardiomegaly, Nodules, Pneumonia, Emphysema, and Hernia. Fibrosis appeared to be one improvement in 5 experiments, and the result of it. The rest of these experiments on Fibrosis performance are decreased. 2nd experiment is the best overall in this combination. However, in the 2nd experiment, the model compensates for Hernia detection compared with the 4th experiment. The five repeated experiments confirm that most results were in the same direction. The mean AUROC in the 2nd experiment is 0.5812 percent increment from the baseline, which indicates the best overall performance in this combination. The interesting observation of this experiment is that overall performance remains to resemble. The difference between the experiment's results is close together.

	paper	myW	ogc1	ogc2	ogc3	ogc4	ogc5	%ogc1	%ogc2	%ogc3	%ogc4	%ogc5
Atelectasis	0.8094	0.8107	0.8102	0.8143	0.8079	0.8133	0.8114	-0.0682	0.4340	-0.3485	0.3169	0.0851
Cardiomegaly	0.9248	0.8914	0.8974	0.9001	0.8967	0.8982	0.8997	0.6739	0.9769	0.5916	0.7626	0.9303
Effusion	0.8638	0.8776	0.8791	0.8803	0.8788	0.8806	0.8787	0.1728	0.3142	0.1466	0.3484	0.1302
Infiltration	0.7345	0.7152	0.7149	0.7165	0.7158	0.7156	0.7173	-0.0489	0.1774	0.0845	0.0495	0.2926
Mass	0.8676	0.8455	0.8427	0.8503	0.8493	0.8451	0.8500	-0.3330	0.5680	0.4531	-0.0466	0.5336
Nodule	0.7802	0.7145	0.7324	0.7425	0.7381	0.7417	0.7419	2.5021	3.9117	3.2956	3.7965	3.8270
Pneumonia	0.7680	0.7721	0.7831	0.7805	0.7893	0.7893	0.7745	1.4214	1.0839	2.2171	2.2244	0.3092
Pneumothorax	0.8870	0.8761	0.8694	0.8720	0.8711	0.8754	0.8719	-0.7656	-0.4707	-0.5681	-0.0752	-0.4823
Consolidation	0.7901	0.7986	0.8004	0.8027	0.8009	0.8042	0.8000	0.2191	0.5064	0.2811	0.6941	0.1698
Edema	0.8878	0.8878	0.8895	0.8863	0.8880	0.8867	0.8844	0.1835	-0.1770	0.0204	-0.1314	-0.3929
Emphysema	0.9371	0.8967	0.9034	0.9011	0.9029	0.9032	0.8921	0.7495	0.4970	0.6954	0.7327	-0.5104
Fibrosis	0.8047	0.7571	0.7570	0.7652	0.7498	0.7479	0.7490	-0.0118	1.0652	-0.9700	-1.2132	-1.0657
Pleural_Thickening	0.8062	0.7835	0.7823	0.7963	0.7866	0.7885	0.7950	-0.1535	1.6281	0.3961	0.6394	1.4663
Hernia	0.9164	0.8727	0.8809	0.8731	0.8934	0.8911	0.8708	0.9378	0.0423	2.3712	2.1053	-0.2115
mean auroc	0.8413	0.8214	0.8245	0.8272	0.8263	0.8272	0.8241	0.3734	0.7076	0.5998	0.7059	0.3226

Table 14: AUROC result from Transfer Learning and fine-tuning with gamma correction and Contrast limited adaptive histogram equalization combination compared with the AUROC from the original image model

4.2.4 Original image + Gamma correction + MMCS

Table 15 shows the AUROC performance result from the combination. The AUROC improves clearly on Cardiomegaly, Nodules, Pneumonia, Pleural Thickening, and Hernia. In 4th experiment, the model compensates for Pneumothorax and Edema detection performance. The five repeated experiments confirm that most results were in the same direction. The mean AUROC in the 4th experiment is 0.4946 percent increment from the baseline, which indicates the best overall performance in this combination. The difference between experiment results is close together.

	paper	myW	ogm1	ogm2	ogm3	ogm4	ogm5	Percent increment %ogm1	%ogm2	%ogm3	%ogm4	%ogm5
Atelectasis	0.8094	0.8107	0.8102	0.8109	0.8127	0.8112	0.8120	-0.0682	0.0163	0.2460	0.0519	0.1579
Cardiomegaly	0.9248	0.8914	0.8974	0.8925	0.8983	0.8974	0.8986	0.6739	0.1181	0.7783	0.6776	0.8060
Effusion	0.8638	0.8776	0.8791	0.8785	0.8775	0.8791	0.8775	0.1728	0.1105	-0.0062	0.1734	-0.0053
Infiltration	0.7345	0.7152	0.7149	0.7138	0.7157	0.7159	0.7172	-0.0489	-0.2022	0.0704	0.0961	0.2832
Mass	0.8676	0.8455	0.8427	0.8471	0.8448	0.8490	0.8477	-0.3330	0.1831	-0.0875	0.4160	0.2571
Nodule	0.7802	0.7145	0.7324	0.7375	0.7371	0.7376	0.7392	2.5021	3.2148	3.1594	3.2347	3.4457
Pneumonia	0.7680	0.7721	0.7831	0.7771	0.7849	0.7806	0.7803	1.4214	0.6462	1.6483	1.0991	1.0588
Pneumothorax	0.8870	0.8761	0.8694	0.8727	0.8712	0.8699	0.8736	-0.7656	-0.3865	-0.5607	-0.7102	-0.2864
Consolidation	0.7901	0.7986	0.8004	0.7977	0.7999	0.8001	0.7991	0.2191	-0.1192	0.1625	0.1785	0.0527
Edema	0.8878	0.8878	0.8895	0.8853	0.8809	0.8860	0.8861	0.1835	-0.2855	-0.7825	-0.2072	-0.1986
Emphysema	0.9371	0.8967	0.9034	0.8990	0.8974	0.9002	0.8956	0.7495	0.2571	0.0814	0.3969	-0.1216
Fibrosis	0.8047	0.7571	0.7570	0.7603	0.7575	0.7608	0.7582	-0.0118	0.4187	0.0553	0.4831	0.1469
Pleural_Thickening	0.8062	0.7835	0.7823	0.7944	0.7860	0.7900	0.7875	-0.1535	1.3866	0.3217	0.8221	0.5078
Hernia	0.9164	0.8727	0.8809	0.8848	0.8721	0.8911	0.8881	0.9378	1.3874	-0.0628	2.1125	1.7657
mean auroc	0.8413	0.8214	0.8245	0.8251	0.8240	0.8263	0.8258	0.3734	0.4504	0.3174	0.6021	0.5302

Table 15: AUROC result from Transfer Learning and fine-tuning with gamma correction and Min-Max Linear Contrast Stretching combination compare with the AUROC from the original image model

4.2.5 Original image + Gamma correction + BCET

Table 16 shows that The AUROC performance improves clearly on Cardiomegaly, Nodules, Pneumonia, and Hernia from the "ogb" combination. The First 3 experiment results stay in the same range in mean AUROC. The five repeated experiments confirm that most results were in the same direction, 3 out of 5. The mean AUROC in the 3rd experiment is 0.3313 percent increment from the baseline, which indicates the best overall performance in this combination. The overall performance in this combination is less than all the technique combinations above.

	paper	myW	ogb1	ogb2	ogb3	ogb4	ogb5	%ogb1	%ogb2	%ogb3	%ogb4	%ogb5
								Percent increment				
Atelectasis	0.8094	0.8107	0.8102	0.8069	0.8068	0.8092	0.8067	0.0682	-0.4803	-0.4926	-0.1893	-0.4933
Cardiomegaly	0.9248	0.8914	0.8974	0.8967	0.8985	0.8946	0.8918	0.6739	0.5917	0.7964	0.3549	0.0452
Effusion	0.8638	0.8776	0.8791	0.8767	0.8769	0.8768	0.8786	0.1728	-0.0953	-0.0760	-0.0896	0.1216
Infiltration	0.7345	0.7152	0.7149	0.7139	0.7165	0.7171	0.7151	-0.0489	-0.1753	0.1871	0.2584	-0.0156
Mass	0.8676	0.8455	0.8427	0.8429	0.8382	0.8434	0.8413	-0.3330	-0.3062	-0.8611	-0.2518	-0.4958
Nodule	0.7802	0.7145	0.7324	0.7373	0.7371	0.7421	0.7463	2.5021	3.1907	3.1589	3.8576	4.4423
Pneumonia	0.7680	0.7721	0.7831	0.7801	0.7851	0.7807	0.7674	1.4214	1.0286	1.6847	1.1159	-0.6194
Pneumothorax	0.8870	0.8761	0.8694	0.8682	0.8761	0.8719	0.8726	-0.7656	-0.8999	0.0045	-0.4792	-0.3987
Consolidation	0.7901	0.7986	0.8004	0.7978	0.8012	0.7999	0.7983	0.2191	-0.1022	0.3179	0.1542	-0.0424
Edema	0.8878	0.8878	0.8895	0.8863	0.8896	0.8851	0.8851	0.1835	-0.1735	0.1953	-0.3055	-0.3046
Emphysema	0.9371	0.8967	0.9034	0.8839	0.8941	0.8875	0.8934	0.7495	-1.4276	-0.2870	-1.0243	-0.3599
Fibrosis	0.8047	0.7571	0.7570	0.7550	0.7495	0.7454	0.7552	-0.0118	-0.2754	-0.9996	-1.5494	-0.2569
Pleural_Thickening	0.8062	0.7835	0.7823	0.7901	0.7869	0.7822	0.7827	-0.1535	0.8422	0.4336	-0.1710	-0.1100
Hernia	0.9164	0.8727	0.8809	0.8884	0.8894	0.8556	0.8710	0.9378	1.8007	1.9180	-1.9603	-0.1910
mean auroc	0.8413	0.8214	0.8245	0.8232	0.8247	0.8208	0.8218	0.3734	0.2142	0.4033	-0.0722	0.0512

Table 16: AUROC result from Transfer Learning and fine-tuning with gamma correction and Balance Contrast Enhancement Technique combination compared with the AUROC from the original image model

4.2.6 Multi-Channels image enhancement result comparison

We observe pathologies improve differently depending on each image enhancement technique. Based on the selection of the best mean AUROC from each Multi-Channel image enhancement model training, the comparison result is shown in table 17. The best overall can use mean AUROC to indicate from the table; the ogh3 combination is the best overall performance on 14 pathologies from these Multi-Channel input training experiments. For the pathology example of a Nodule alone, the ogc4 combination is the best technique to improve disease detection by 2.71 percent. The second best for Nodule detection is the ogm4 combination, as shown in Table 17.

Each pathology detection base on each image enhancement technique, from figure 26, shows the potential performance for the specific uses for one class of pathology detection. When observing different perspectives with all the experiments on image enhancement techniques specific to the disease, the disease which significantly improves is Cardiomegaly, Nodule, Pneumonia, Emphysema, Fibrosis, Pleural Thickening, and Hernia. To maximize improvement from these experiments for Cardiomegaly, specifically chooses the ogh3 combination and then select the ogc2 combination for fibrosis and pleural thickening. The most potent combination for detecting nodules and emphysema is ogb5 and ogh1. The most excellent combination to diagnose Pneumonia is the ogi3 combination, whereas the best combination to detect a hernia is the ogh3 combination. The best weight trained from the experiment is shown in figures 27-33, according to each pathology.

	paper	myW	ogh3	ogc4	ogm4	ogi3	ogb3	Percent increment				
								%oogh3	%ogc4	%ogm4	%oogi3	%oogb3
Atelectasis	0.8094	0.8107	0.8134	0.8133	0.8112	0.8112	0.8068	0.3236	0.3169	0.0519	0.0604	-0.4926
Cardiomegaly	0.9248	0.8914	0.9012	0.8982	0.8974	0.8929	0.8985	1.0985	0.7626	0.6776	0.1639	0.7964
Effusion	0.8638	0.8776	0.8772	0.8806	0.8791	0.8775	0.8769	-0.0368	0.3484	0.1734	-0.0099	-0.0760
Infiltration	0.7345	0.7152	0.7174	0.7156	0.7159	0.7162	0.7165	0.3135	0.0495	0.0961	0.1332	0.1871
Mass	0.8676	0.8455	0.8490	0.8451	0.8490	0.8524	0.8382	0.4113	-0.0466	0.4160	0.8134	-0.8611
Nodule	0.7802	0.7145	0.7360	0.7417	0.7376	0.7340	0.7371	3.0026	3.7965	3.2347	2.7224	3.1589
Pneumonia	0.7680	0.7721	0.7808	0.7893	0.7806	0.7914	0.7851	1.1282	2.2244	1.0991	2.4920	1.6847
Pneumothorax	0.8870	0.8761	0.8747	0.8754	0.8699	0.8721	0.8761	-0.1535	-0.0752	-0.7102	-0.4503	0.0045
Consolidation	0.7901	0.7986	0.7982	0.8042	0.8001	0.8005	0.8012	-0.0599	0.6941	0.1785	0.2268	0.3179
Edema	0.8878	0.8878	0.8892	0.8867	0.8860	0.8883	0.8896	0.1511	-0.1314	-0.2072	0.0532	0.1953
Emphysema	0.9371	0.8967	0.9026	0.9032	0.9002	0.8995	0.8941	0.6655	0.7327	0.3969	0.3228	-0.2870
Fibrosis	0.8047	0.7571	0.7579	0.7479	0.7608	0.7517	0.7495	0.1092	-1.2132	0.4831	-0.7154	-0.9996
Pleural_Thickening	0.8062	0.7835	0.7860	0.7885	0.7900	0.7879	0.7869	0.3110	0.6394	0.8221	0.5633	0.4336
Hernia	0.9164	0.8727	0.8986	0.8911	0.8911	0.8835	0.8894	2.9670	2.1053	2.1125	1.2351	1.9180
mean auroc	0.8413	0.8214	0.8273	0.8272	0.8263	0.8256	0.8247	0.7185	0.7059	0.6021	0.5165	0.4033

Table 17: AUROC comparison from Transfer Learning and fine-tuning with five multi-channel combination image enhancement techniques and the AUROC from the original image(myW column) based on the best overall AUROC (mean AUROC).
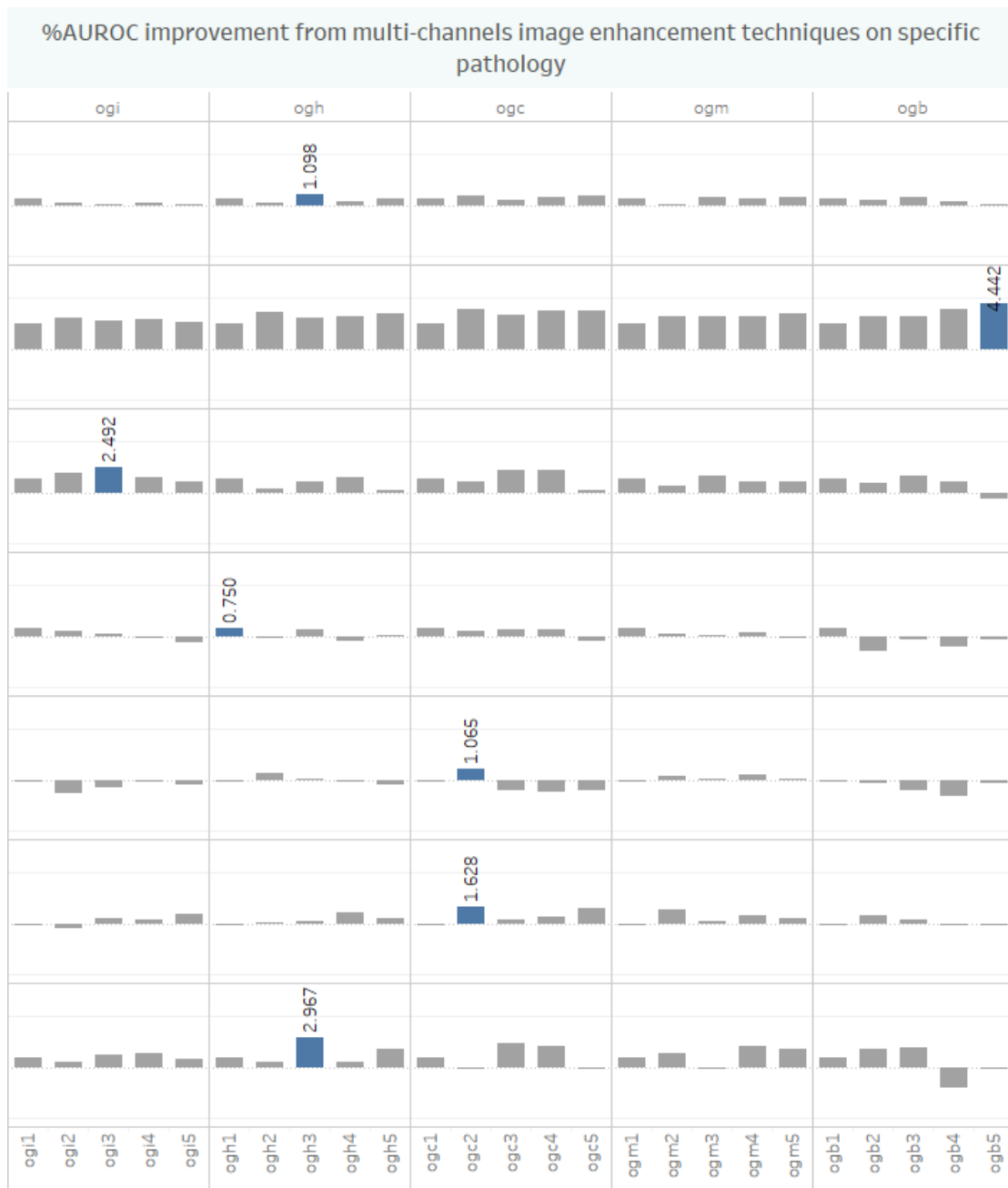
Figure 26: The percentage of AUROC improvement from multi-channel combination image enhancement techniques on the specific pathology
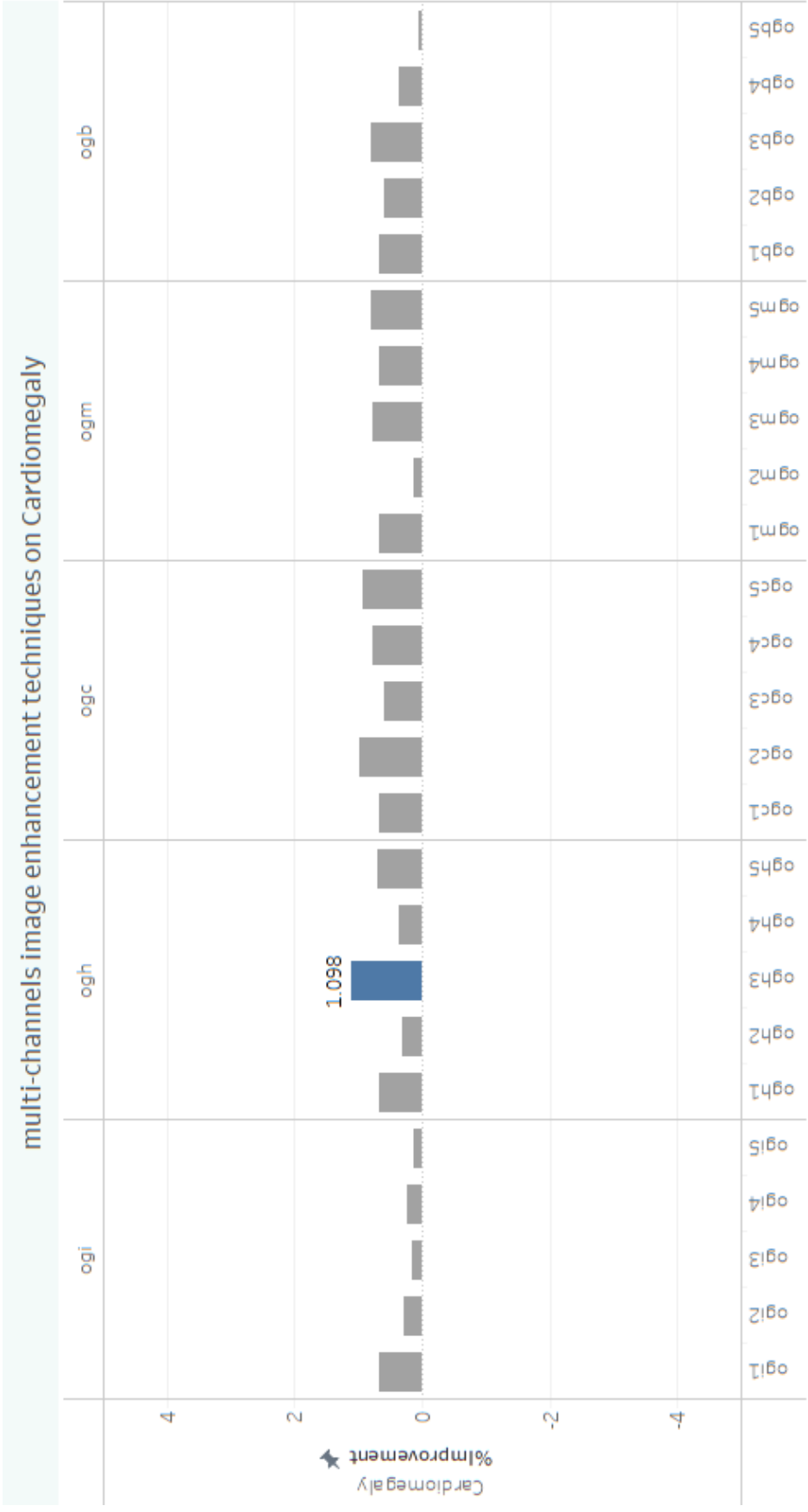
Figure 27: The percentage of AUROC improvement from multi-channel combination image enhancement techniques on Cardiomegaly.
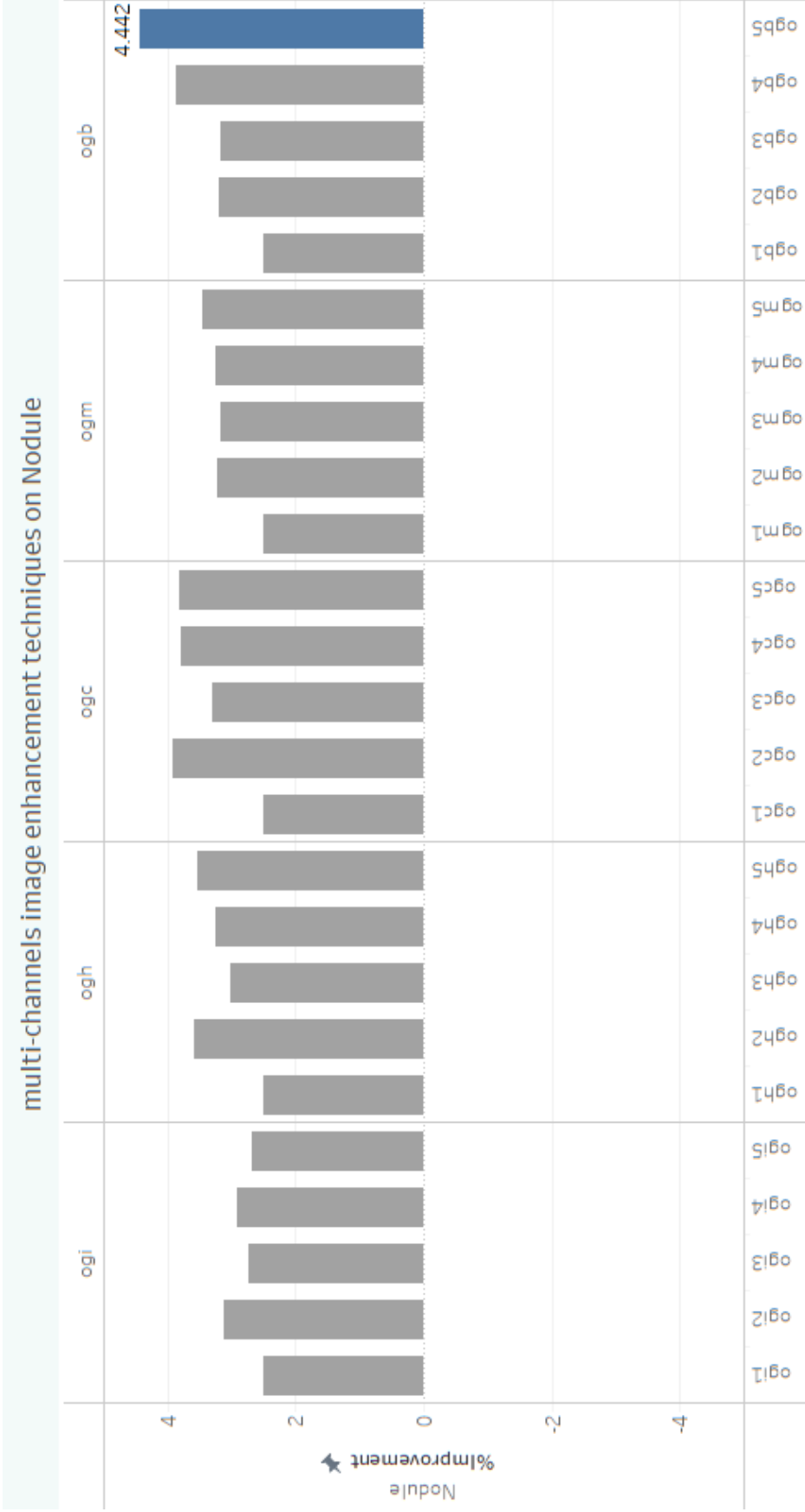
Figure 28: The percentage of AUROC improvement from multi-channel combination image enhancement techniques on Nodule.
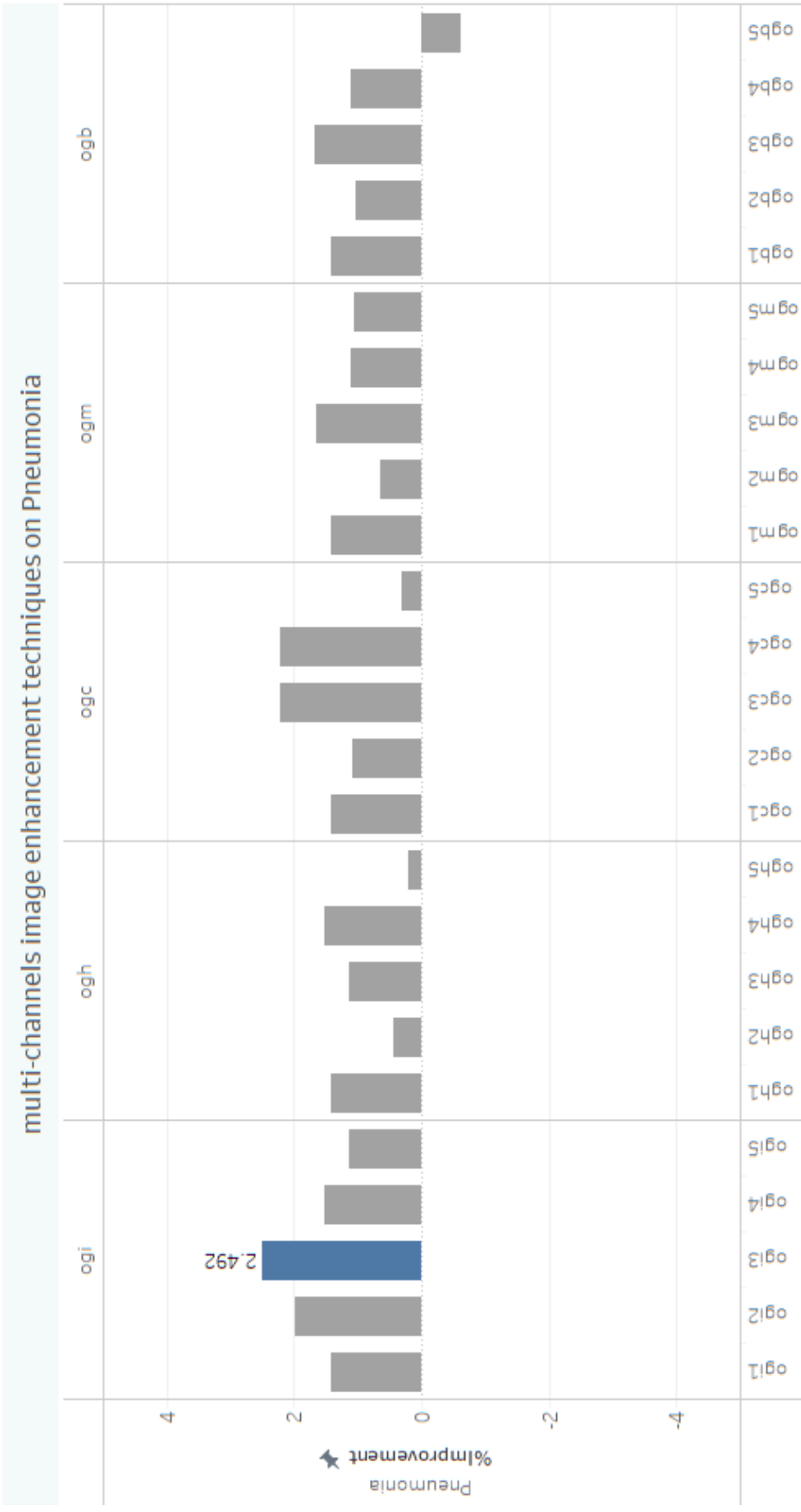
Figure 29: The percentage of AUROC improvement from multi-channel combination image enhancement techniques on Pneumonia.
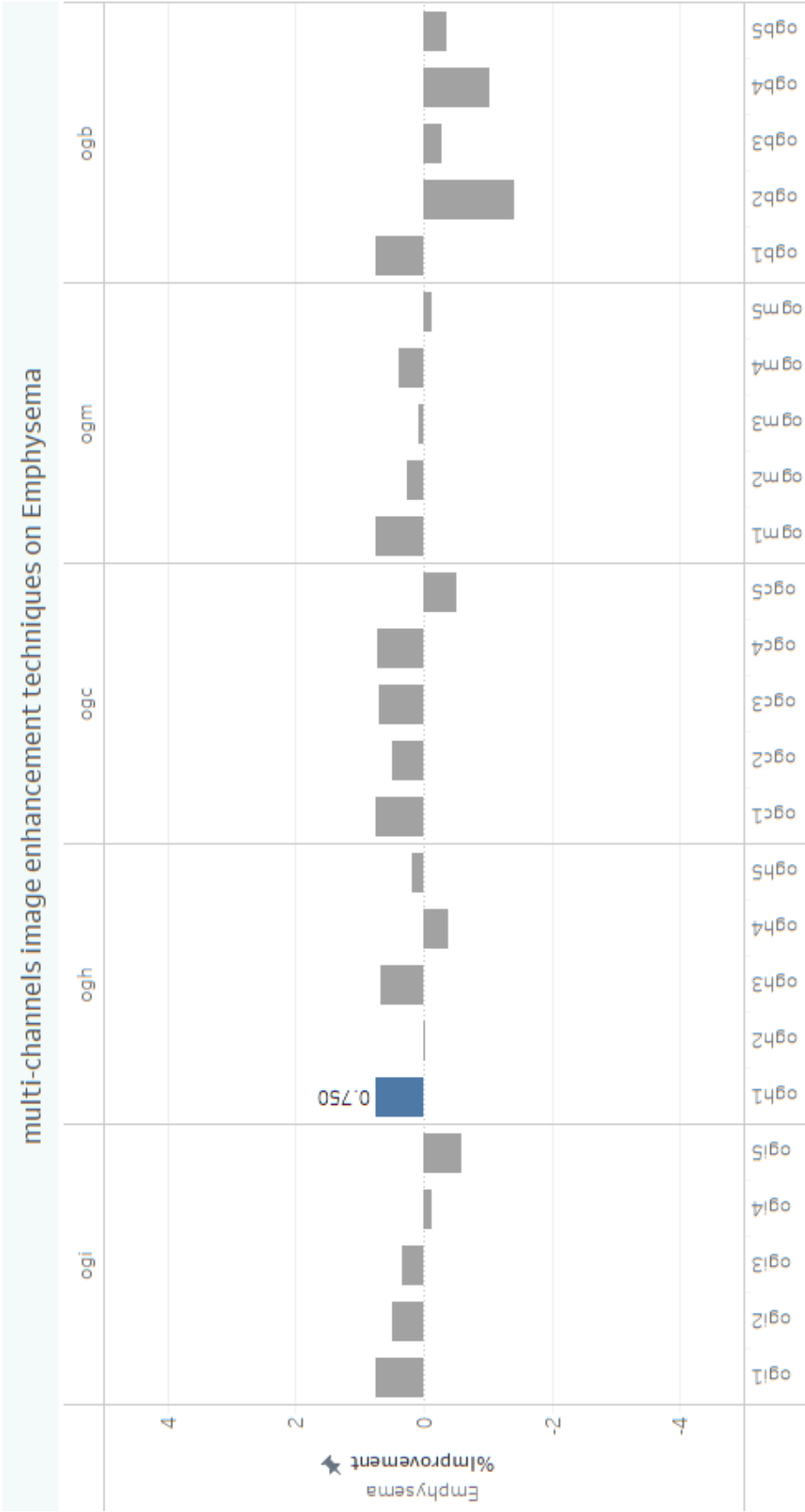
Figure 30: The percentage of AUROC improvement from multi-channel combination image enhancement techniques on Emphysema.
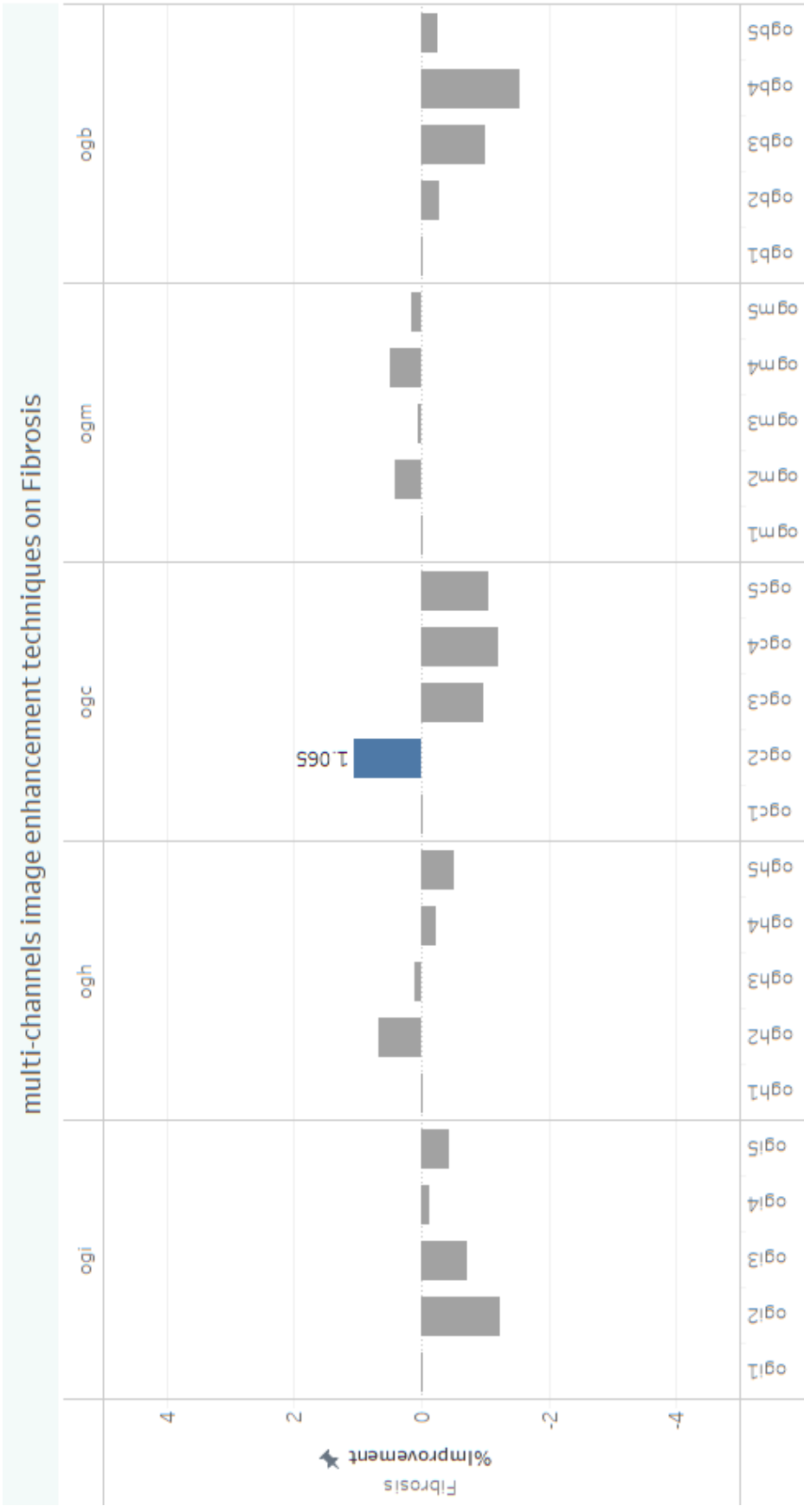
Figure 31: The percentage of AUROC improvement from multi-channel combination image enhancement techniques on Fibrosis.
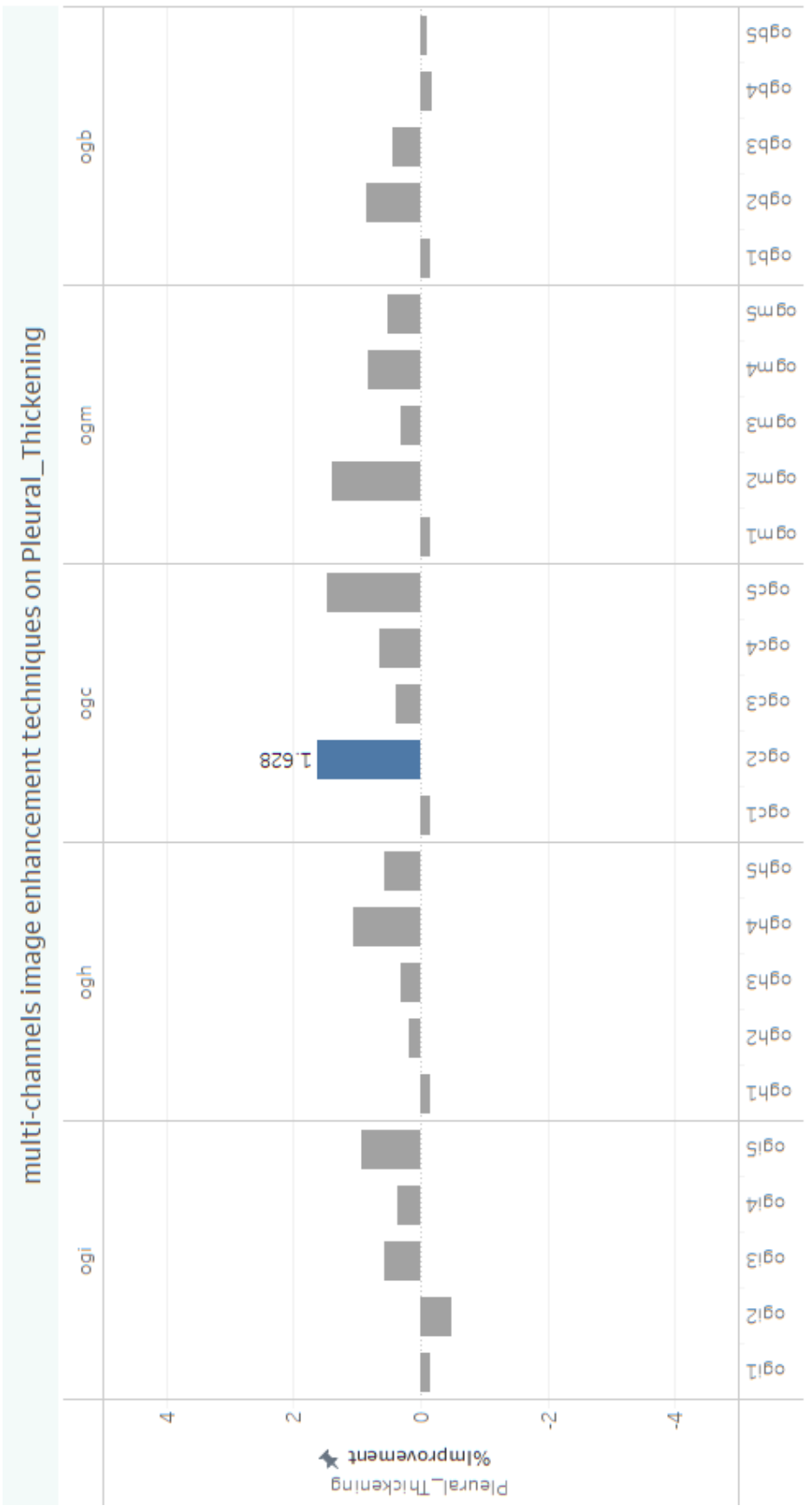
Figure 32: The percentage of AUROC improvement from multi-channel combination image enhancement techniques on Pleural Thickening.
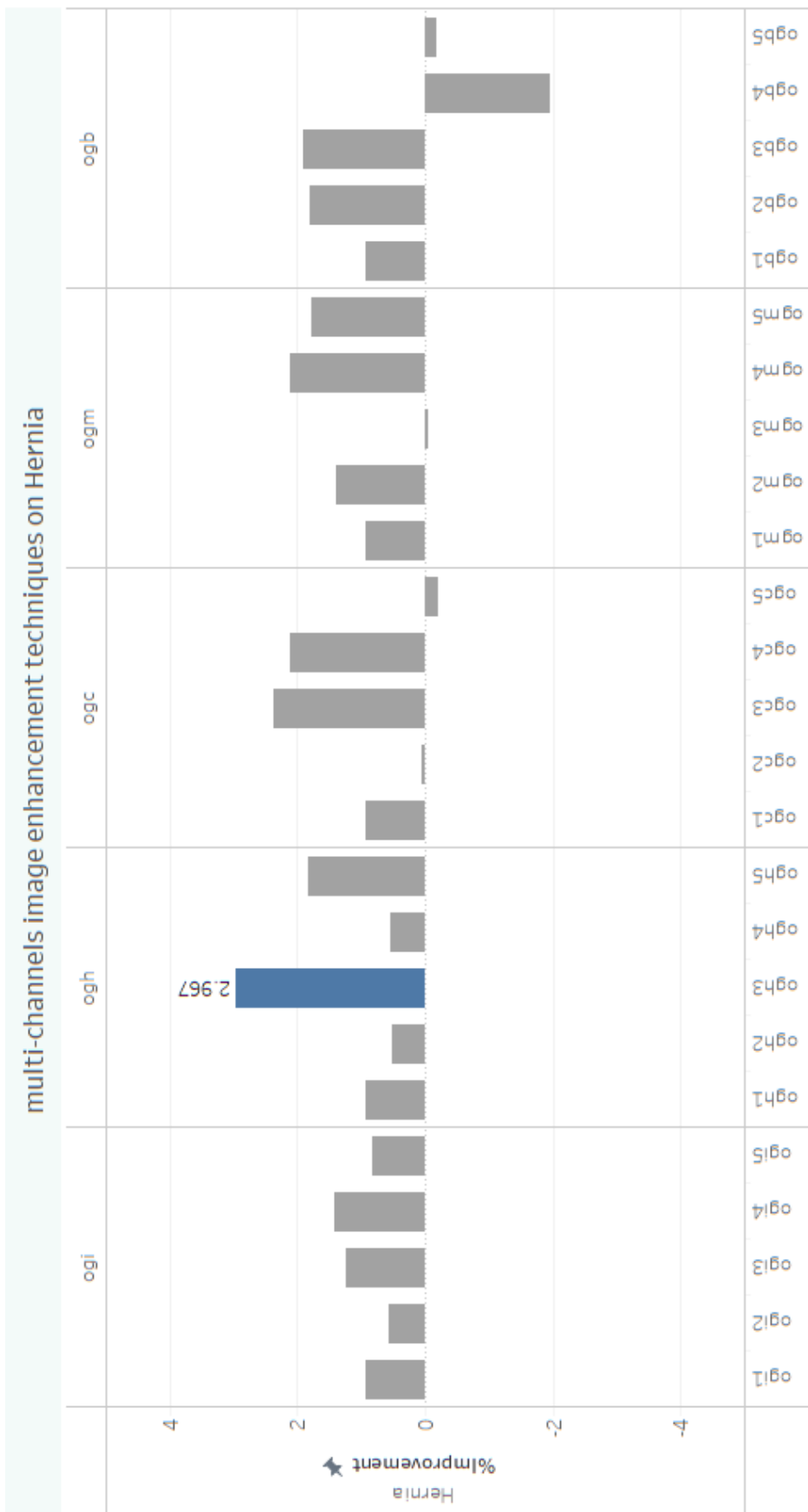
Figure 33: The percentage of AUROC improvement from multi-channel combination image enhancement techniques on Hernia.

4.3 All result comparison

Multi-channel image enhancement compensates performance on detection-specific pathology instead of overall pathologies. Table 18 all results are shown in this table. The single channel image enhancement considered gamma correction experiment has stable and the highest overall performance (mean auroc).

Furthermore, multi-channel image enhancement considered the "ogc" combination experiment has stable and the highest overall performance. The standard deviation of the "gam5" is less than the "ogc2" combination, which indicates that the "gam5" has better overall pathologies detection than the "ogc2" combination.

Table 18:: AUROC comparison from Transfer Learning and fine-tuning with single channel image enhancement techniques and the

AUROC from the original image(myW column)

Table 19: AUROC comparison from Transfer Learning and fine-tuning with five multi-channel combination image enhancement

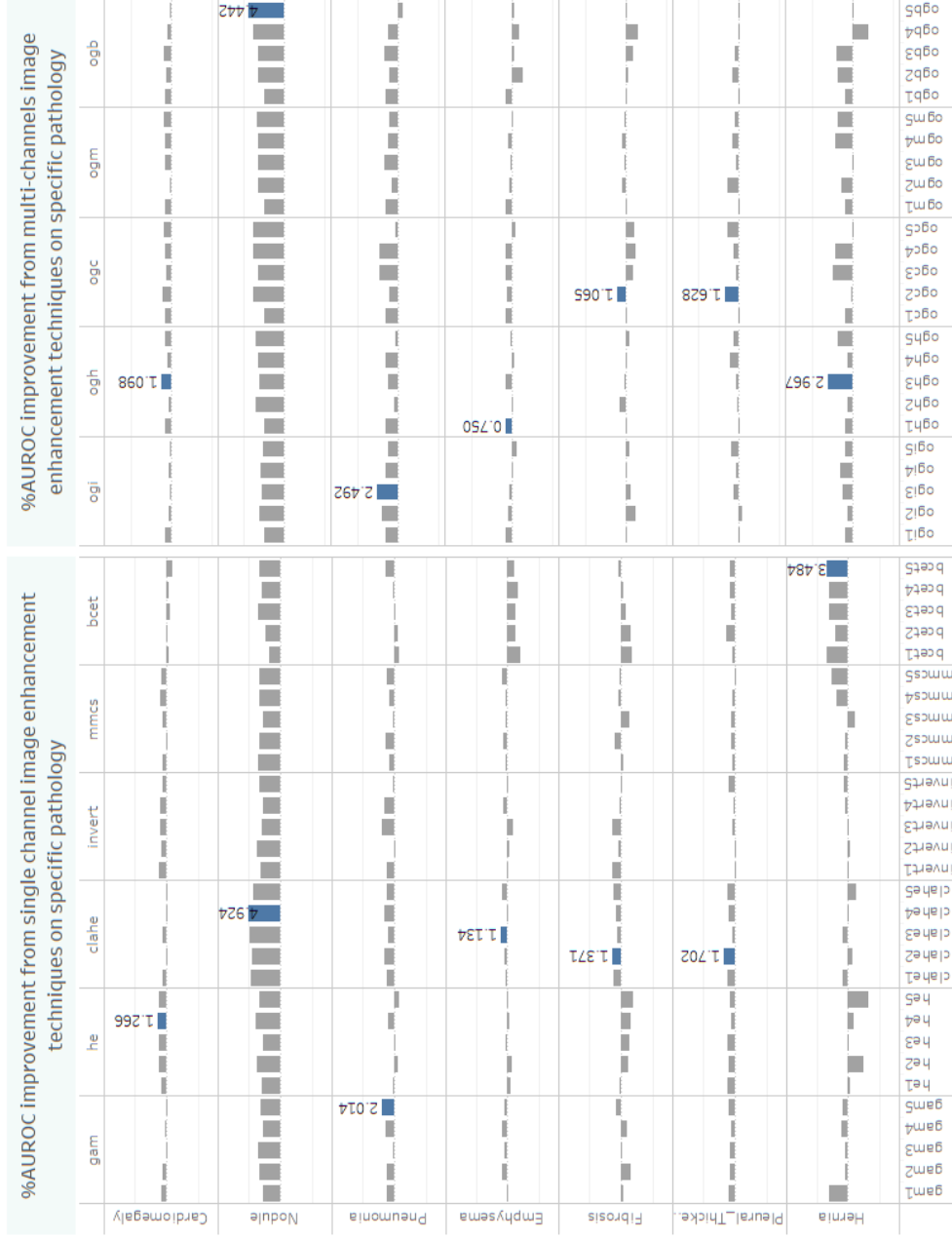techniques and the AUROC from the original image(myW column)

Figure 34: The percentage AUROC improvement on seven specific pathologies of all image enhancement techniques

REFERENCES

[1]   O. Er, N. Yumusak, and F. Temurtas, "Chest diseases diagnosis using artificial neural networks," *Expert Systems with Applications,* vol. 37, no. 12, pp. 7648-7655, 2010, doi: 10.1016/j.eswa.2010.04.078.

[2]   R. Noumeir, "Radiology interpretation process modeling," *Journal of Biomedical Informatics,* vol. 39, no. 2, pp. 103-114, 2006/04/01/ 2006, doi: https://doi.org/10.1016/j.jbi.2005.07.001.

[3]   A. A. El-Solh, C. B. Hsiao, S. Goodnough, J. Serghani, and B. J. Grant, "Predicting active pulmonary tuberculosis using an artificial neural network," *chest,* vol. 116, no. 4, pp. 968-73, Oct 1999, doi: 10.1378/chest.116.4.968.

[4]   A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems,* vol. 25, pp. 1097-1105, 2012.

[5]   M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International Conference on Machine Learning*, 2019: PMLR, pp. 6105-6114.

[6]   A. Kesner, R. Laforest, R. Otazo, K. Jennifer, and T. Pan, "Medical imaging data in the digital innovation age," *Medical physics,* vol. 45, no. 4, pp. e40-e52, 2018.

[7]   P. Rajpurkar *et al.*, "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," pp. 3-9, 2017. [Online]. Available: http://arxiv.org/abs/1711.05225.

[8]   L. A. Lehmann *et al.*, "Generalized image combinations in dual KVP digital radiography," *Med Phys,* vol. 8, no. 5, pp. 659-67, Sep-Oct 1981, doi: 10.1118/1.595025.

[9]   C. A. A. Vinhais, "Medical X-ray Images of the Human Thorax Segmentation, Decomposition and Reconstruction," PhD Thesis 2007. [Online]. Available: https://repositorio-aberto.up.pt/bitstream/10216/12167/2/Texto%20integral.pdf.

[10]   S. Sanada, K. Doi, X. W. Xu, F. F. Yin, M. L. Giger, and H. MacMahon, "Comparison

of imaging properties of a computed radiography system and screen–film systems," *Medical physics,* vol. 18, no. 3, pp. 414-420, 1991.

[11]   M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European conference on computer vision*, 2014: Springer, pp. 818-833.

[12]   O. Russakovsky *et al.*, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision,* vol. 115, no. 3, pp. 211-252, 2015/12/01 2015, doi: 10.1007/s11263-015-0816-y.

[13]   I.-M. Baltruschat, "Deep learning for automatic lung disease analysis in chest x-rays " 5-May-2021 2021, doi: 10.15480/882.3511.

[14]   G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700-4708. [Online]. Available: https://scholar.google.com/scholar_lookup?title=Densely%20Connected%20Convolutional%20Networks&author=Gao%20Huang&author=Zhuang%20Liu&publication_date=2018/01/28&arxiv_id=1608.06993. [Online]. Available: https://scholar.google.com/scholar_lookup?title=Densely%20Connected%20Convolutional%20Networks&author=Gao%20Huang&author=Zhuang%20Liu&publication_date=2018/01/28&arxiv_id=1608.06993

[15]   N. D. Uvarov, "Multi-label classification of a real-world image dataset," 2017.

[16]   L. J. Guo, "Balance contrast enhancement technique and its application in image colour composition," *International Journal of Remote Sensing,* vol. 12, no. 10, pp. 2133-2151, 1991/10/01 1991, doi: 10.1080/01431169108955241.

[17]   T. Rahman *et al.*, "Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images," *Comput Biol Med,* vol. 132, p. 104319, May 2021, doi: 10.1016/j.compbiomed.2021.104319.

[18]   S. S. Al-amri, N. V. Kalyankar, and S. D. Khamitkar, "Contrast Stretching Enhancement in Remote Sensing Image," (in English), *International Journal of Computer Science Issues (IJCSI),* vol. 7, no. 2, pp. 26-29, Mar 2010 2010.

[Online]. Available: https://www.proquest.com/scholarly-journals/contrast-stretching-enhancement-remote-sensing/docview/89069664/se-2?accountid=44800.

[19]   X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "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)*, 21-26 July 2017 2017, pp. 3462-3471, doi: 10.1109/CVPR.2017.369. [Online]. Available: https://ieeexplore.ieee.org/document/8099852/

[20]   J. Irvin *et al.*, "CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison," *33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019,* pp. 590-597, 2019, doi: 10.1609/aaai.v33i01.3301590.

[21]   S. Kornblith, J. Shlens, and Q. V. Le, "Do better imagenet models transfer better?," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2661-2671.

[22]   H.-C. Shin, K. Roberts, L. Lu, D. Demner-Fushman, J. Yao, and R. Summers, *Learning to Read Chest X-Rays: Recurrent Neural Cascade Model for Automated Image Annotation*. 2016, pp. 2497-2506.

VITA