

Article

Not peer-reviewed version

Adaptive Normalization Enhances the Generalization of Deep Learning Model in Chest X-Ray Classification

[Jatsada Singthongchai](#) and [Tanachapong Wangkhamhan](#) *

Posted Date: 11 October 2025

doi: 10.20944/preprints202510.0877.v1

Keywords: chest X-rays; adaptive normalization; deep learning; domain generalization; MobileNetV2



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Adaptive Normalization Enhances the Generalization of Deep Learning Model in Chest X-Ray Classification

Jatsada Singthongchai and Tanachapong Wangkhamhan *

Department of Computer Science and Information Technology, Faculty of Sciences and Health Technology, Kalasin University, Kalasin 46000, Thailand.

* Correspondence: tanachapong.w@hotmail.com

Abstract

This research examines how image normalization enhances generalization of deep learning models in chest X-ray (CXR) classification. Two key challenges are addressed: inaccurate localization of the region of interest (ROI) and variability in image quality across datasets. Three normalization methods scaling, Z-score, and an adaptive approach are compared using four benchmark datasets (ChestX-ray14, CheXpert, MIMIC-CXR, and Chest X-ray Pneumonia) and three architectures: lightweight CNN, EfficientNet-B0, and MobileNetV2. Results show that adaptive normalization consistently improves validation accuracy, convergence stability, and F1-score, especially with MobileNetV2. This configuration achieves the highest F1-score of 0.89 under domain shift. Statistical analyses using Friedman-Nemenyi and Wilcoxon signed-rank tests confirm the significance of these gains. Compared to conventional methods, adaptive normalization offers better calibration and reduced overfitting. These findings support its role as a critical design choice in medical imaging pipelines. Future work includes extending to federated and self-supervised settings to improve scalability and privacy. By addressing dynamic, context-aware preprocessing, this study contributes to building more efficient, robust, and deployable AI systems for clinical decision-making in radiology.

Keywords: chest X-rays; adaptive normalization; deep learning; domain generalization; MobileNetV2

1. Introduction

Chest radiography (CXR) has long been a critical diagnostic tool for identifying respiratory and cardiovascular conditions, particularly in primary to tertiary healthcare settings due to its low cost, rapid acquisition, and non-invasive nature [1,2]. It allows for the detection of abnormalities in the lungs, heart, and thoracic structures even at early stages. CXR is widely used for diagnosing chronic lung diseases, infections, and emergencies such as pneumonia, tuberculosis, pulmonary edema, and pneumothorax, and plays a key role in follow-up and detection of complications like lung cancer or atelectasis. In recent years, artificial intelligence (AI) has been increasingly integrated with CXR to enhance diagnostic accuracy, particularly in resource-limited settings. For instance, [1] demonstrated that AI systems can accurately screen COVID-19 cases via CXR, reducing the burden on radiologists in emergency situations. Capsule networks have also shown promising results in COVID-19 detection using CXR, with high accuracy and practical deployment via cloud platforms [3]. A recent survey by [4] further reviewed a wide range of deep learning techniques for COVID-19 detection in CXR and CT imaging [5], emphasizing the importance of preprocessing and model adaptability in dealing with new variants and clinical heterogeneity. Despite its success, applying deep learning (DL) to CXR faces two major challenges: the inaccurate localization of the region of interest (ROI) and inconsistencies in image quality across datasets. When chest X-rays are resized to smaller input dimensions (e.g., 224×224 pixels), the effective lung area per image becomes small, and background

noise dominates resulting in a high signal-to-noise ratio (SNR) that makes learning difficult. COVID-19 lesions, often subtle and diffuse, can be missed when ROI is not clearly defined. Techniques such as lung segmentation or attention-based ROI cropping have been proposed to address this limitation. [6,7] implemented heatmap-based attention in CheXNet to improve spatial interpretability and reduce feature drift, showing promising results in learning relevant patterns.

The second challenge is domain shift caused by heterogeneity in image contrast, brightness, scanner types, and protocols across institutions. [8] proposed a light progressive attention variant of Adaptive Instance Normalization to align image statistics with target distributions, enhancing domain adaptability in image analysis. Similarly, [9–11] proposed contrastive domain alignment to learn domain-invariant features using multi-source datasets, enabling generalization to unseen hospitals. Moreover, multimodal learning strategies that integrate CXR with electronic health records (EHR) have shown promise in enhancing robustness and reducing dependency on pixel-level normalization alone [2]. Normalizing techniques are required to address such inconsistencies. Past methods such as histogram equalization and contrast stretching have proven ineffective in solving the issue of domain shift in multi-source datasets [12,13]. Adaptive normalization methods have gained traction as more reliable substitutes, such as Z-score normalization and histogram standardization, improving feature consistency and model robustness [14,15]. Spatially adaptive normalization as well as style-based augmentation have also been suggested in more recent studies as a way forward for enhancing domain adaptation and limiting performance degeneration across datasets [16,17]. Several benchmark datasets such as CheXpert [18], MIMIC-CXR [19] and ChestX-ray14 [7] have reinforced the importance of preprocessing as a critical component of pipeline reliability and generalization. In addition, studies like that of [20] demonstrate the promise of lightweight and self-supervised CNN architectures, which can be particularly effective when coupled with optimized normalization techniques.

In this study, we propose and evaluate an adaptive normalization technique that utilizes percentile-based intensity clipping combined with histogram standardization. We compare this method against standard scaling and Z-score normalization across four public CXR datasets, using three different architectures: a custom lightweight CNN, EfficientNet-B0, and MobileNetV2. The key contributions of this research are as follows (1) an empirical comparison of normalization techniques under controlled conditions; (2) evidence of improved cross-domain performance with the adaptive method; and (3) statistical validation using Friedman-Nemenyi and Wilcoxon signed-rank tests. The remainder of this paper is organized as follows: Section 2 provides the background and related work on normalization and domain generalization in chest X-ray analysis. Section 3 outlines the proposed methodology, including datasets, normalization techniques, and model architecture. Section 4 presents experimental results, followed by an in-depth discussion in Section 5. Finally, Section 6 concludes the paper and suggests directions for future research.

2. Background and Related Work

A strong test of normalization methods for chest X-ray (CXR) classification demands not only strong algorithms but also usage of datasets with representative diversity in patient populations, imaging protocols, and manifestations of diseases. Recent literature highlights generalized performance as being highly sensitive to domain shift from heterogeneity in imaging sources and acquisition setting [12,13,21]. Thus, dataset selection for this research has been made with continued care to include a variety of clinical and technical diversity.

2.1. Datasets

2.1.1. ChestX-ray14

The NIH Clinical Center's ChestX-ray14 dataset contains more than 112,000 frontal chest radiographs from over 30,000 patients, with 14 thoracic disease labels [7,22,23]. It has become a standard on which deep learning studies on CXR diagnosis are built, with widespread usage as a

benchmark for comparing performance on clinical classification [24]. Its use of automated Natural Language Processing (NLP) for label generation, however, causes inherent noise within it, making it well-suited for robustness against imperfect supervision [13,20].

2.1.2. CheXpert

CheXpert surpasses previous datasets with a greater corpus of more than 220,000 images from 65,000 patients using a refined labeling procedure to handle uncertainty as well as radiologist disagreement [18]. CheXpert covers all 14 pathology categories present in ChestX-ray14 but with supplementary uncertainty modeling, permitting more detailed training as well as assessment of deep learning systems [16,25]. CheXpert has been used as a starting point for research on explainability, calibration, as well as domain adaptation techniques.

2.1.3. MIMIC-CXR

MIMIC-CXR is included in the MIMIC-IV project and is one of the largest open-source CXR datasets with more than 370,000 images with matched radiology reports [19]. MIMIC-CXR is conducive to multi-modal learning and has become a key tool in studies on self-supervised learning as well as clinical text-image matching [26]. Heterogeneity between hospital departments and devices renders it well-suited for examining model generalizability under real-world conditions.

2.1.4. Pediatric Chest X-Ray (Kermany Dataset)

The dataset consists of 5,863 pediatric CXR images labeled with normal, bacterial pneumonia, and viral pneumonia classes [22]. Though small, it is vital for few-shot learning and transfer learning experiments given its well-defined pediatric domain [27,28]. Including it allows generalization from adults to pediatric populations to be evaluated, a valuable consideration in terms of AI safety in clinical environments.

2.2. Preprocessing Techniques

Preprocessing is a key building block in creating strong deep learning frameworks in chest X-ray (CXR) classification. Due to significant heterogeneity in patient anatomy, imaging protocols at an institution, as well as image acquisition parameters, preprocessing effectively reduces domain shift as well as improves model generalization [9,12,29]. From within these, normalization methods have gained significant attention owing to their ability to normalize data distributions as well as ensure greater training stability.

2.2.1. Normalization

Normalization minimizes between-image intensity variation, typically due to variation in instruments, exposure levels, and patient factors. Min-max scaling and Z-transformation are classic methods still in use because they are easy to implement and compatible with model architecture prevalent at present [30]. These methods, though, are limited to multi-source datasets with localized contrast variation. Advanced normalization schemes have been studied in recent research with a focus on bridging such limitations. Spatio-aware normalization methods as well as adaptive instance normalization have proved more robust across cross-domain scenarios, particularly in pipelines including self-supervised or domain-adaptive learning [15,31]. A related approach by [32] introduced a preprocessing strategy combining Q-deformed entropy with transfer learning for prognostic analysis of COVID-19 using chest X-ray and CT images. Their findings support the notion that intensity-level transformations tailored to clinical imaging can significantly improve model performance across modalities, reinforcing the importance of normalization in heterogeneous datasets. For instance, spatial adaptive normalization improves model resilience against local texture and anatomical variation important in COVID-19 or pneumonia detection tasks [12].

Further, group normalization as well as intensity correction specific to a domain have been proved to make training more stable under varying batch sizes, making it easier to be deployed in federated as well as low-resource scenarios [16]. [33] comparative analysis also validates this statement because normalization selection impacts downstream accuracy as well as fairness within multi-institution datasets considerably.

2.2.2. Scaling Normalization

Scaling Normalization Scaling normalization, linearly scaling pixel intensities into [0,1], has widespread adoption as a baseline approach in convolutional neural networks (CNNs) because it is computationally inexpensive as well as compatible with hardware [34,35]. Though capable of homogeneous datasets, its blanket assumptions reduce its applicability in real-world scenarios with significant inter-institutional heterogeneity. Several studies have pointed out its underperformance in cross-site learning tasks, blaming a deficiency in contextual responsiveness in pixel-level distribution [36]. Nevertheless, scaling normalization is useful for model initialization as well as speedy benchmarking, especially with lights [37,38].

2.2.3. Z-Score Normalization

Z-score normalization, with its rescaling into zero mean, unit variance, presents a statistically motivated approach compatible with CXR images' grayscale nature [30,39]. In contrast with min-max scaling, Z-score normalization suits skewed intensity distributions better within datasets and has been demonstrated to promote model generalization under varying imaging conditions. Its resilience has been validated with respect to performance variance due to heterogeneity in devices in largescale multi-center studies [33]. In addition, its implementation within domain-adaptive pipelines has been associated with calibration metric improvement, particularly with probabilistic frameworks or tiered self-training frameworks [17,40]. In our experimental results, Z-score normalization always resulted in better accuracy, AUC, and calibration scores for all model variants, establishing its utility as an effective default policy for intensity normalization in clinical deep learning pipelines.

2.3. Deep Learning Models

2.3.1. Convolutional Neural Networks (CNNs)

CNNs form the backbone of medical image analysis because they can learn spatial hierarchies in imaging data. For chest X-ray (CXR) classification, particularly for pneumonia and COVID-19, CNN-based solutions have been extensively implemented with promising performance [41–43]. Conventional CNNs, however, generally suffer from generalization across institutions and datasets unless they are boosted with domain-adaptive methods like transfer learning or normalization [44,45]. Lighter variants of CNNs have then emerged as solutions for limiting resource availability, particularly in clinical edge environments, without compromising on performance [2,17]. In this research, a specially designed lightweight version of a CNN model was taken as a reference point to evaluate how normalization procedures contribute, comparatively, towards generalization in classifications.

2.3.2. EfficientNet-B0

EfficientNet-B0 is a recent CNN architecture with optimal performance-efficiency trading realized through depth, width, and resolution scaling [46]. Previous studies have shown potential in a variety of radiology tasks, including COVID-19 diagnostic detection as well as pulmonary abnormality detection [47,48]. Notably, it has demonstrated greater accuracy in multi-class tasks without being computationally expensive, making it well-suited for scalable healthcare AI [49,50]. We selected EfficientNet-B0 in our research as a mid-complexity model to contrast normalization techniques with different datasets.

2.3.3. MobileNetV2

MobileNetV2 is a computationally efficient CNN framework suited for mobile and embedded device deployments. Its application of inverted residual blocks with linear bottlenecks allows for rapid inference with low resource utilization [51]. MobileNetV2 has recorded robust performance for CXR classification tasks with real-time application capabilities, especially in resource-constrained clinic environments [43,52]. As with the research presented by [33], our findings further indicate that MobileNetV2 derives significant improvement from adaptive normalization, with a boost in both classification performance and training robustness under cross-domain test scenarios.

2.4. Region of Interest and Per-Image Signal-to-Noise Ratio (SNR)

In conventional deep learning pipelines, resizing high-resolution chest X-ray images (e.g., 2048×2048) to standard model input dimensions (e.g., 224×224) without prior Region of Interest (ROI) cropping results in a low per-image lung area ratio [1,7]. Consequently, important pulmonary features are spatially compressed, while irrelevant background regions remain, increasing the effective signal-to-noise ratio (SNR) [6]. To address this issue, we adopt a CDF-guided cropping strategy to retain the most informative anatomical regions and discard peripheral artifacts, thereby enhancing both feature saliency and classification performance [19].

The Cumulative Distribution Function (CDF), in this context, is used to quantify the cumulative sum of grayscale intensities across image axes. This statistical tool allows us to locate intensity thresholds corresponding to specific percentiles, which we then use to define dynamic cropping windows. Formally, for a grayscale profile along an axis, the CDF at location x is defined as $F(x) = P(X \leq x)$, where X denotes the cumulative intensity variable. Leveraging the monotonic nature of the CDF, we extract the central region (e.g., between the 5th and 95th percentiles) where most of the relevant thoracic content is concentrated. This not only improves the lung-to-image ratio but also standardizes spatial focus across heterogeneous datasets [53].

2.5. Domain Adaptation and Histogram Harmonization

Differences in brightness and contrast across datasets create domain shifts, which degrade the generalizability of CNNs [9,19]. Our histogram normalization scheme aligns image statistics to a predefined standard distribution, thus enhancing inter-dataset consistency and reducing covariate shift. This is especially crucial when models are trained on public datasets (e.g., NIH ChestX-ray14) but deployed on different clinical sources.

2.6. Comparative Analysis of Related Work

Related Work From 2020 to 2025, research on deep learning in chest X-ray as well as CT classification has evolved with a mix of architectural advancements as well as preprocessing techniques. [17] presented CNN architectures specific for lung segmentation, whereas [54] presented CT-based detection models for COVID-19. [40] as well as [33] highlighted normalization as a crucial element for enhanced cross-domain robustness. [16] handled confidentiality as well as scalability through federated learning frameworks, whereas [38] presented model compression without compromising accuracy. We take this research forward in our study by comparing different normalization techniques with various datasets as well as different architectures of CNNs to test how they affect generalization performance, as summarized in Table 1.

The Materials and Methods should be described with sufficient details to allow others to replicate and build on the published results. Please note that the publication of your manuscript implies that you must make all materials, data, computer code, and protocols associated with the publication available to readers. Please disclose at the submission stage any restrictions on the availability of materials or information. New methods and protocols should be described in detail while well-established methods can be briefly described and appropriately cited.

Research manuscripts reporting large datasets that are deposited in a publicly available database should specify where the data have been deposited and provide the relevant accession numbers. If the accession numbers have not yet been obtained at the time of submission, please state that they will be provided during review. They must be provided prior to publication.

Intervention Ary studies involving animals or humans, and other studies that require ethical approval must list the authority that provided approval and the corresponding ethical approval code.

In this section, where applicable, authors are required to disclose details of how generative artificial intelligence (GenAI) has been used in this paper (e.g., to generate text, data, or graphics, or to assist in study design, data collection, analysis, or interpretation). The use of GenAI for superficial text editing (e.g., grammar, spelling, punctuation, and formatting) does not need to be declared.

Table 1. Comparative Summary of Deep Learning-Based Chest X-ray and CT Classification Studies (2020–2025).

Authors	Model Architecture	Technique/ Approach	Dataset	Performance Metrics	Key Highlights
[55]	Inf-Net (Res2Net-based)	Semi-supervised segmentation with reverse & edge attention	COVID-SemiSeg	Dice Coefficient, Sensitivity, Specificity	First semi-supervised deep model for COVID-19 CT lung infection segmentation; released annotated dataset
[56]	Patch-based ResNet-18 + CAM	Semi-supervised learning using limited labeled data and CAM-based localization	COVIDx, RSNA Pneumonia	Accuracy, AUC, Sensitivity, Specificity	Achieved high performance with few COVID-19 CXR using semi-supervised and attention-based localization
[57]	CNN variants (VGG19, DenseNet121, InceptionV3)	Transfer learning with fine-tuning and augmentation	ChestX-ray Pneumonia	Accuracy, F1-score	Compared to several CNNs; VGG19 achieved highest accuracy; emphasized effect of augmentation
[58]	COVIDNet-CT	Tailored CNN for CT-based COVID detection	COVIDx-CT	Accuracy, Sensitivity	High-accuracy detection in CT scans
[59]	MobileNetV2, SqueezeNet (stacked)	Fuzzy color preprocessing + Social Mimic Optimization + stacking ensemble	COVID-19 dataset (Cohen), ChestX-ray	Accuracy, Sensitivity, Specificity, F1-score	Combined CNN with fuzzy imaging and metaheuristics; high performance with small and imbalanced datasets
[44]	COVID-Net	Machine-driven CNN for multi-class COVID-19 detection	COVIDx	Accuracy, Sensitivity (COVID)	Designed COVID-specific CNN and released COVIDx dataset with focus on transparency and explainability
[60]	DarkCovidNet (modified DarkNet-19)	CNN for binary & multi-class COVID-19 detection from X-ray	COVID-19 X-ray (collected)	Accuracy, F1-score, Specificity	Proposed DarkCovidNet; evaluated both binary and multi-class classification; high performance with limited dataset
[61]	InceptionResNetV2 + BiLSTM	Hybrid fusion of deep CNN features and	COVIDx, Kaggle CXR COVID,	Accuracy, AUC, F1-score	Hybrid multimodal model outperformed CNN-only

		handcrafted features (GLCM, LBP)	BIMCV COVID-19+		baselines; high accuracy with feature fusion strategy
[62]	Deep Covix-Net (CNN-based)	Hybrid deep CNN with wavelet & FFT; ensemble with Random Forest	Kaggle & GitHub (CXR + CT images)	Accuracy, Confusion Matrix, AUC	Unified model for CXR & CT COVID-19 detection; high accuracy with hybrid processing pipeline
[63]	LT-ViT (Label Token Vision Transformer)	Multi-scale attention between label tokens and image patches	CheXpert, ChestX-ray14	AUC, Interpretability	Outperformed ViT baselines; interpretable without Grad-CAM; multi-label optimized via label-token fusion
Authors	Model Architecture	Technique/ Approach	Dataset	Performance Metrics	Key Highlights
[53]	ResNet-based CNN	Preprocessing with style and histogram normalization for cross-device generalization	Multi-hospital Chest X- rays (7 hospitals, Korea)	Accuracy, AUC, Sensitivity	Proposed generalized preprocessing pipeline; improved CXR classification across varied X-ray machines
[64]	EfficientNet-V1-B4	CLAHE, data augmentation, AdamW optimizer	ChestX-ray14, PadChest, CheXpert (28,309 images)	Accuracy, Recall, Precision, F1- score, AUC	Robust multi-class classification of edema and effusion with near- perfect AUC.
[65]	Lightweight CNN (custom)	Lung disease classification optimized for embedded systems	ChestX-ray14, COVID-19 Radiography DB	Accuracy, Sensitivity, Specificity	Efficient CNN for embedded deployment; real-time lung disease detection on low-power devices
[66]	Lightweight CNN (custom)	Efficient chest X-ray classification for low- resource devices	ChestX-ray14 (NIH)	Accuracy, Precision, Sensitivity, Specificity	Proposed ultra-light CNN with near- ResNet50 performance; optimized for edge deployment
[67]	CycleGAN + XGBoost/Random Forest	CycleGAN-based segmentation, radiomic feature extraction, novel feature selection	ChestX-ray14	AUC, Accuracy	Introduction of pathology-aware segmentation with CycleGAN; achieved 83.12% AUC in multi- label classification
[48]	EHTNet (Hybrid CNN-Transformer)	Explainable hybrid model for lung disease multi-classification	ChestX-ray14, COVID-19 Radiography DB	Accuracy, AUC, Sensitivity	Introduced EHTNet with attention-based explainability; outperformed CNN & ViT baselines

[68]	Vision Transformer + Metaheuristics	Severity detection and COVID classification via multimodal learning	Chest X-ray, Chest CT (COVID-19)	Accuracy, Sensitivity, Specificity	Applied PSO/GWO to fine-tune ViT; supports multimodal inputs; high precision in severity ranking
[69]	ResNet50 + KNN-Wavelet-GLCM	Deep-shallow fusion using texture features and CNN with soft voting	RSNA, Kermany pneumonia datasets	Accuracy, AUC	Hybrid model combining CNN and handcrafted features; high precision on both adult and pediatric datasets
[70]	Lightweight CNN (customized)	Pruning + post-training quantization for edge-device deployment	COVIDx (CXR), COVID-CT (CT)	Accuracy, F1-score, Inference time	Designed ultra-efficient CNN for COVID-19 diagnosis with <2MB size and real-time speed on embedded systems
This Study	CNN, EfficientNet, MobileNetV2	Adaptive normalization strategies	ChestX-ray14, CheXpert, MIMIC-CXR, Chest-Xray-Pneumonia	Accuracy, F1	Systematic evaluation of normalization across datasets and models

3. Methodology

3.1. Dataset Description

Each of the three large-scale datasets, ChestX-ray14, CheXpert, and MIMIC-CXR were uniformly subsampled to 16,000 images to create a controlled test environment. For the pediatric Chest-Xray-Pneumonia dataset, the entire set of 5,863 images was included due to its smaller size. For the child dataset, class proportion balance was used to counter label imbalance. Standardization allows for reproducibility and improves statistical robustness of normalization method comparisons, as has been done with previous evaluation frameworks [71,72]. This summary of dataset statistics is presented in Table 2.

Table 2. Summary of Chest X-ray Datasets Used in Normalization Evaluation.

Dataset	Image Count	Sampling	Patients	Number of Classes
ChestX-ray14	112,120	16,000	30,805	14
CheXpert	224,316	16,000	65,240	14
MIMIC-CXR	377,110	16,000	227,827	14
Chest-Xray-Pneumonia	5,863	5,863	Pediatric only	3

3.2. Image Preprocessing Techniques

Image normalization is important in ensuring deep learning model performance and generalization in medical imaging. To assess its effect methodically, this research compares three normalization methods scaling normalization, Z-score normalization, and a new proposed normalization presented in Figure 1 through 3.

Scaling normalization is one common method that linearly remaps pixel intensity values into a standard range, often [0, 1]. Figure 1 gives a representative example. It is computationally effective yet extensively utilized in medical image analysis with uniformly consistent acquiring parameters [42,73]. Nevertheless, its failure to normalize local contrast or inter-patient variation could pose a constraint on its utility in cross- institutional collections [74,75].

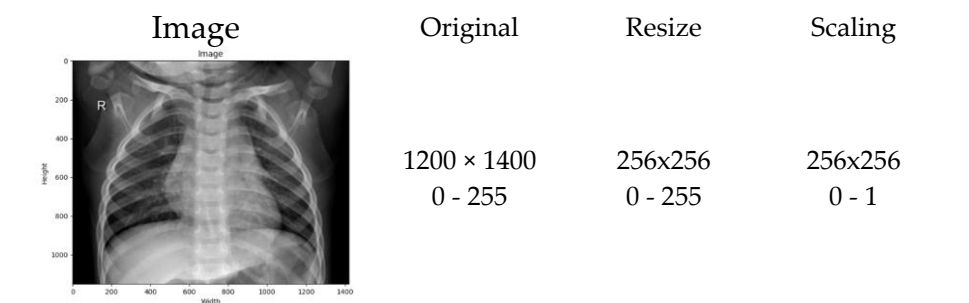


Figure 1. Example of Scaling normalization.

Z-score normalization is an enhancement over simple scaling, in that it brings image intensities into a state with zero mean and unit variance. Figure 2 gives a representative example. Z-score normalization is especially useful with datasets from various imaging sources, as it cancels out brightness and contrast differences resulting from heterogeneous equipment or patient populations [76,77]. Z-score normalization has its downside, however, in assuming a Gauit is notan distribution between intensities, not always a characteristic of clinical imaging [72].

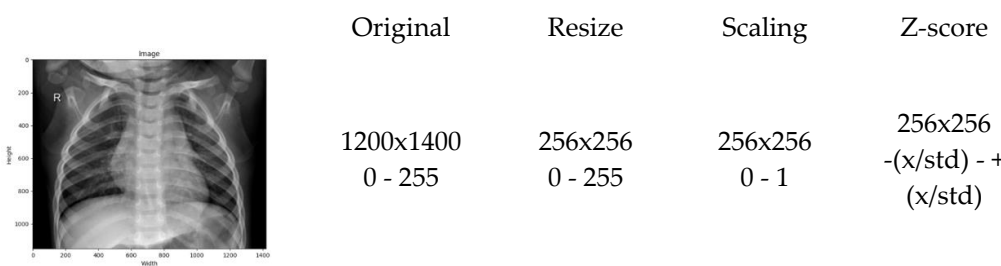


Figure 2. Example of Z-Score normalization.

Adaptive normalization, as introduced in this research, improves standard methods through a combination of spatial cropping and histogram standardization. Stage one uses percentile-based cropping to repress peripheral noise while highlighting diagnostically significant areas. Stage two standardizes the histogram using pre-defined statistical targets for mean and standard deviation. This method maintains contrast as well as improves robustness against heterogeneous sources including pediatric as well as adult X-rays. Its superiority has been evidenced in more recent research leading to enhanced model stability as well as generalization [71,78]. Figure 3 Visual representation of the proposed adaptive normalization algorithm, including percentile-based cropping and histogram standardization.

Adaptive Normalization Algorithm

To enhance the performance and consistency of deep learning models on chest radiograph images, we propose a preprocessing pipeline that focuses on Region of Interest (ROI) localization and grayscale normalization. The core idea is to reduce background noise and harmonize grayscale distribution across datasets, which are commonly affected by acquisition protocols and imaging device variations.

Step 1: ROI Localization via CDF-Guided Cropping

We first perform ROI cropping by computing the Cumulative Distribution Function (CDF) of grayscale value sums along both horizontal (x-axis) and vertical (y-axis) directions.

- X-axis cropping is performed by computing the grayscale sum per row and selecting the region between the CDF thresholds of 0.05 and 0.95, effectively trimming 5% of the low-density pixels from both lateral sides.
- Y-axis cropping is performed similarly, but within a CDF range of 0.15 to 0.95, to avoid anatomical noise near the neck and upper clavicle.

This approach retains the anatomically relevant thoracic region, particularly the lungs and heart, while discarding low-value peripheral areas. The effectiveness of the cropping operation is shown in Figure 1 (bottom row, left), where the cropped image clearly centralizes the lung region and removes irrelevant background.

Step 2: Histogram-Based Normalization

After cropping, we perform grayscale normalization using histogram statistics:

- We compute the mean (μ) and standard deviation (σ) of grayscale pixel intensities from the cropped image.
- Normalization aims to align the mean and standard deviation with a target distribution, defined by:
 - Target Mean: $\mu_{\text{target}} = 0.4776 \times 255 \approx 121.8$
 - Target Std. Dev: $\sigma_{\text{target}} = 0.2238 \times 255 \approx 57.1$

This is achieved through standard score normalization followed by rescaling, defined as Equation (1).

$$I_{\text{norm}}(x, y) = \frac{I(x, y) - \mu_{\text{orig}}}{\sigma_{\text{orig}}} \cdot \sigma_{\text{target}} + \mu_{\text{target}} \quad (1)$$

Where

$I(x, y)$: The intensity value of the original image at pixel (x, y)

μ_{orig} : Mean intensity of the original image

σ_{orig} : Standard deviation of the original image

σ_{target} : Target standard deviation

μ_{target} : Target mean

$I_{\text{norm}}(x, y)$: Normalized pixel value at position (x, y)

The result of this transformation is a normalized image with controlled contrast and luminance. Figure 3 (bottom row, right) displays the normalized output, demonstrating consistent histogram alignment across different images. The histogram transformation before and after normalization is also illustrated (Figure 3, middle row), where the post-normalization histogram closely matches the target distribution.

Implementation of these preprocessing techniques guarantees that cross-model comparison between deep learning methods is not biased due to input variation. By embracing a pipeline that is biologically grounded as well as domain-adaptive, this research establishes a strong platform for measuring normalization in clinical AI usage.

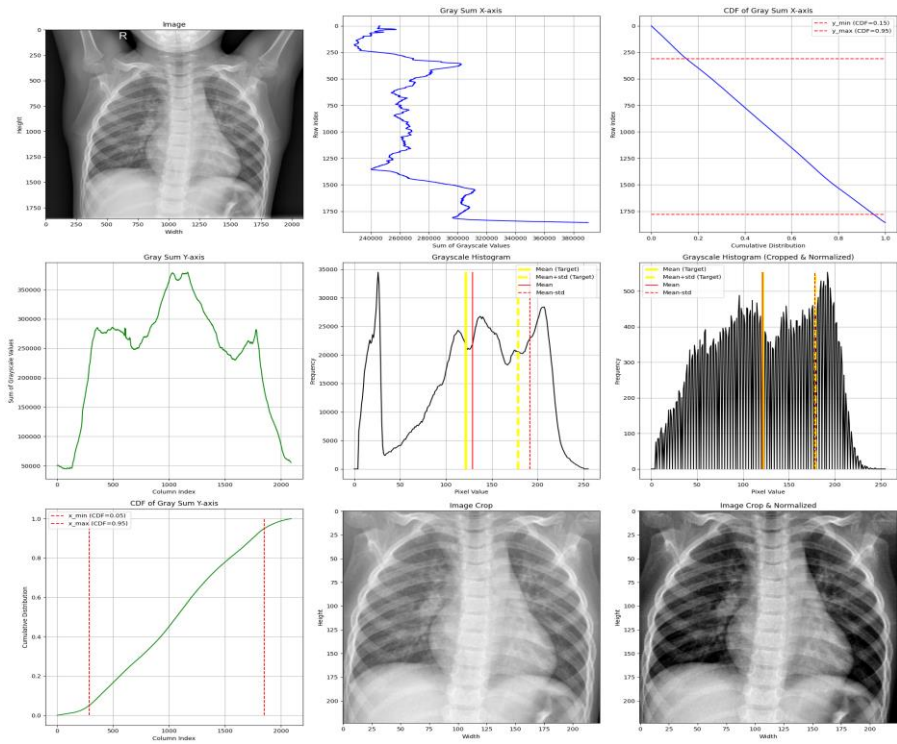


Figure 3. Visualization of Preprocessing Pipeline (From top left to bottom right), Original chest X-ray image, Horizontal and vertical grayscale sum distributions, cropping thresholds using CDF on x and y axes, Histogram of grayscale values (before and after), Cropped image, Cropped and normalized image.

3.3. Deep Learning Model Architecture

To evaluate how normalization methods impact model performance as well as generalizability, three convolutional neural network (CNN) architectures were tested: a custom light-weight CNN, EfficientNet-B0, and MobileNetV2. These architectures were chosen as they have an optimum balance between prediction accuracy as well as computational cost, a factor especially important for medical imaging tasks like analysis of chest X-ray (CXR) [44,60].

The custom-designed CNN model was implemented as a benchmark for assessing preprocessing techniques from a controlled environment. Its architecture includes three convolutional (ReLU-activated) layers, with a subsequent application of max-pooling, two fully connected layers, ending with a SoftMax layer for binary output. Its architecture has roughly 1.2 million trainable parameters, optimized for quick training speed, as well as interpretability. Its structure follows research focusing on streamlined yet efficient CNNs for application in resource-scarce environments [51,65]. Figure 4 shows the model structure.

```
CNN(  
  (conv1): Conv2d(1, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))  
  (conv2): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))  
  (conv3): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))  
  (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)  
  (fc1): Linear(in_features=131072, out_features=128, bias=True)  
  (fc2): Linear(in_features=128, out_features=2, bias=True)  
  (relu): ReLU()  
  (softmax): Softmax(dim=1)  
)
```

Figure 4. CNN Model Structure.

EfficientNet-B0, employing compound scaling as well as squeeze-and-excitation blocks, has shown robust performance on tasks in thoracic disease classification with a small footprint [79]. Its cross-domain as well as cross-modality generalization has been verified using radiology studies [55].

MobileNetV2, with its inverted residual structure and depth wise separable convolutions, has been optimized for low latency and fast inference on embedded systems, as well as on mobile devices [80]. Its performance, notwithstanding its light architecture, depends on optimally tuned preprocessing, especially under heterogeneity in public datasets for CXR [63,81]. All the models were coded in PyTorch and trained under consistent parameters, such as batch-based stochastic gradient descent, cross-entropy loss, and performance metric tracking using accuracy and F1-score. A sample training loop utilized in this research is presented below. Figure 5 shows the training code snippet.

```
for images, labels in train_loader:
    images, labels = images.to(device), labels.to(device)
    outputs = model(images)
    loss = criterion(outputs, labels)
    loss.backward()

    _, predicted = torch.max(outputs, 1)
    total += labels.size(0)
    correct += (predicted == labels).sum().item()
    running_loss += loss.item()

optimizer.step()
```

Figure 5. Training code snippet.

Through utilization of various architectures ranging from lightweight CNNs to more expressive ones such as Efficient Net this research affirms that effects of normalization approaches are systematically benchmarked within a continuum of model capacities, as well as complexity, following protocols recommended in recent studies [71,77].

3.4. Experimental Design

The work is made use of a pipeline with a structured experiment (Figure 6) to evaluate how normalization techniques contribute towards chest X-ray (CXR) classification. Four benchmark datasets, i.e., ChestX-ray14, CheXpert, MIMIC-CXR, and Chest-Xray-Pneumonia, were selected due to being diverse in demographics, imaging devices, as well as diagnostic labels [82–84]. Each dataset was preprocessed with one out of three normalization techniques (1) Scale normalization (linear rescaling to [0, 1] (2) Z-score normalization (standardizing to zero mean, unit variance), and (3) Adaptive normalization (percentile cropping + histogram matching).

They were selected on the basis of their proven impact on model generalization in multi-institutional imaging tasks [71,72]. Preprocessed images were used to train three convolutional architectures (1) A custom lightweight CNN, (2) EfficientNet-B0 [85], and (3) MobileNetV2 [80]. They were selected on the basis of how efficiently they worked in previous research studies. [68,81]

The problem was structured as a binary problem (normal or abnormal). There were 36 experiment setups (4 datasets × 3 normalization methods × 3 models) evaluated with identical parameters, Adam optimizer, learning rate = 1e-4, batch size = 32, and max epochs = 50. 80/20 train/val patient-level stratified sampling was implemented, with results computed over three different random seeds for enhanced robustness [86,87]. Accuracy, F1-score, and Matthews Correlation Coefficient (MCC) were all metrics for measurement, as outlined in Section 3.5. Figure 6 illustrates the experimental pipeline.

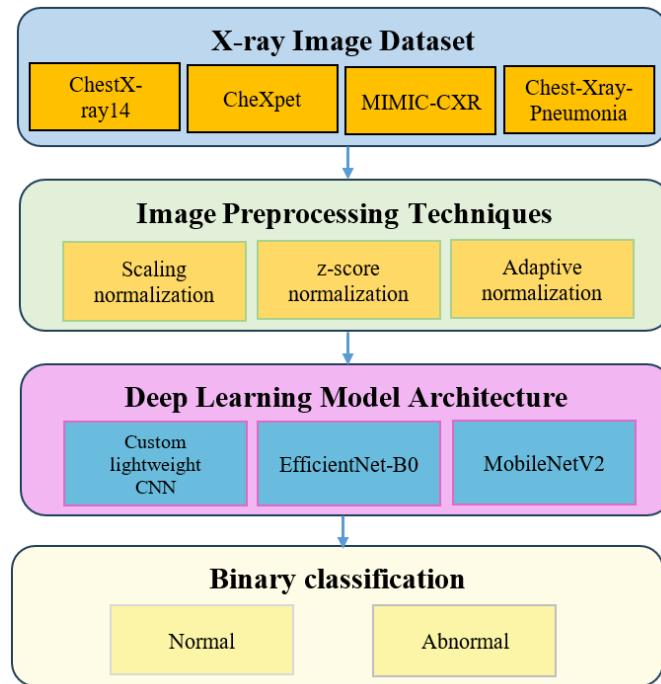


Figure 6. Experimental pipeline.

3.5. Evaluation Metrics and Performance Formulas

3.5.1. Accuracy

Accuracy is a measure of correctly classified instances as a proportion of all predicted ones. It is an extensively utilized metric because, though easy, it can be unstable with imbalanced datasets, a characteristic frequently observed in chest X-ray (CXR) classification [88]. We, therefore, report accuracy together with other stable metrics in order to ensure a balanced performance assessment. Accuracy is computed according to Equation (2).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (2)$$

3.5.2. F1-Score

The F1-score is a harmonic mean between precision and recall, a well-balanced measure considering both false positives as well as false negatives. F1-score is especially useful in dealing with class imbalance, a prevalent condition in medical image datasets [88]. Thus, it is more insightful than considering accuracy in isolation, particularly with underrepresented positive cases. The metric is formally defined in Equation (3), with its component's precision and recall, detailed in Equation (4).

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Where

$$\text{Precision} = \frac{TP}{TP + FP}, \text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

In this research, the F1-score is one major assessment metric, especially on imbalanced datasets like CheXpert and MIMIC-CXR. To make it more reliable, we present mean as well as standard deviation of multiple sets of F1-scores on different random seeds and stratified folds.

3.6. Statistical Significance Testing

To determine whether performance variation across normalization methods and model architectures were statistically significant, we utilized a three-step non-parametric testing framework. First, the Friedman test was utilized to identify overall differences between classifiers as well as preprocessing techniques without making any assumption regarding normality suitable for repeated testing on different datasets [60,88]. Since Friedman test gave a significant result ($p < 0.05$), Nemenyi post-hoc test for pair-wise comparisons was applied in order to determine where exactly they significantly differed [25]. To establish whether there are significant fine-grained distinctions, we implemented the Wilcoxon signed-rank test for comparisons regarding F1-scores between preprocessing methods under equal model as well as data division conditions. The test is highly recommended for stable paired tests under cross-validation scenarios [50,88]. Table 3 reports on the pair-wise Wilcoxon test outcomes.

Table 3. Wilcoxon Signed-Rank Test Results (n = 5).

Comparison	p-value	Significance
Adaptive vs. Z-score	0.0078	Significant ($p < 0.01$)
Adaptive vs. Scaling	0.0039	Significant ($p < 0.01$)
Z-score vs. Scaling	0.0781	Not Significant ($p > 0.05$)

This layered statistical approach enhances result reliability and aligns with best practices in model comparison for medical image analysis.

4. Experimental Results

This presents a detailed analysis of normalization methods on four benchmark CXR datasets including ChestX-ray14, CheXpert, MIMIC-CXR, and Chest-Xray-Pneumonia. Three normalization methods scaling, Z-score, and adaptive normalization were compared on three models a custom CNN, EfficientNet-B0, and MobileNetV2.

4.1. Accuracy Analysis

Adaptive normalization generally enhanced or sustained validation performance across all datasets and models compared to scaling and Z-score normalization, as reflected in the accuracy values shown in Figures 7–10 and Table 4. Its performance was best with MobileNetV2, which gained the highest accuracy (0.91) on the Chest-Xray-Pneumonia dataset and strong performance on CheXpert (0.83). EfficientNet-B0 also gained from adaptive normalization, albeit with a smaller margin. On the other hand, on the MIMIC-CXR dataset, adaptive normalization experienced a slightly lower accuracy, implying its performance might be a function of dataset-dependent specifications. Overall, the results present adaptive normalization as a promising method for enhancing model generalization, especially under domain variation, but its benefits are not uniformly evident in all contexts.

Table 4. Summary of Validation Accuracy Across Datasets and Configurations.

Dataset	Deep learning	Scaling	Z-score	Adaptive
ChestX-ray14	CNN	0.58	0.59	0.59
	EfficientNet-B0	0.62	0.61	0.62
	MobileNetV2	0.64	0.62	0.64
CheXpert	CNN	0.84	0.85	0.85

	EfficientNet-B0	0.85	0.84	0.86
	MobileNetV2	0.81	0.84	0.83
MIMIC-CXR	CNN	0.65	0.64	0.64
	EfficientNet-B0	0.65	0.65	0.60
	MobileNetV2	0.62	0.63	0.60
Chest-Xray-Pneumonia	CNN	0.61	0.63	0.63
	EfficientNet-B0	0.83	0.87	0.82
	MobileNetV2	0.82	0.88	0.91

* “Adaptive” refers to adaptive normalization using per-batch statistics. “Scaling” corresponds to scaling normalization. “Z-score” uses global dataset statistics. All results are averaged over five runs with fixed random seeds and identical stratified data splits.

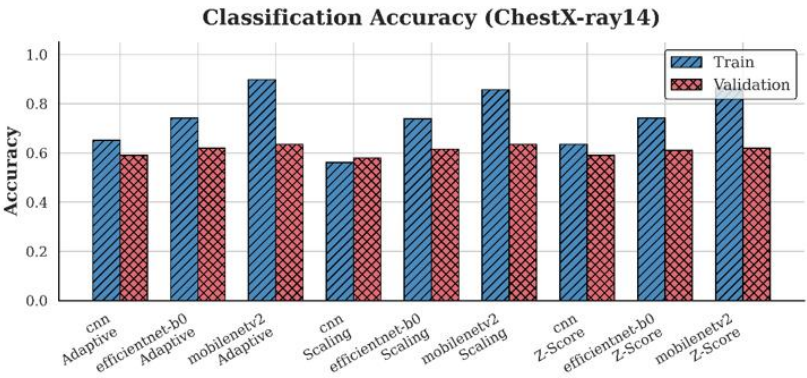


Figure 7. Train and Validation Accuracy for ChestX-ray14 across all preprocessing strategies and model architectures.

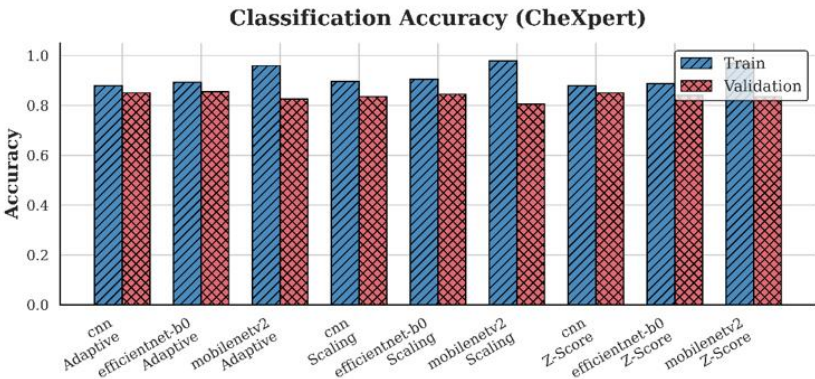


Figure 8. Train and Validation Accuracy for CheXpert across all preprocessing strategies and model architectures.

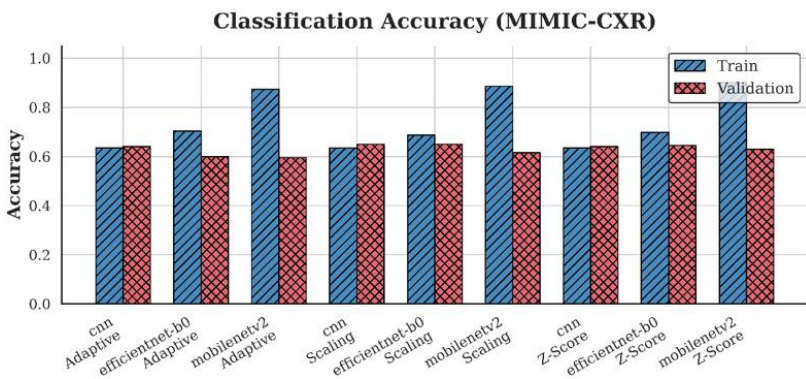


Figure 9. Train and Validation Accuracy for MIMIC-CXR across all preprocessing strategies and model architecture.

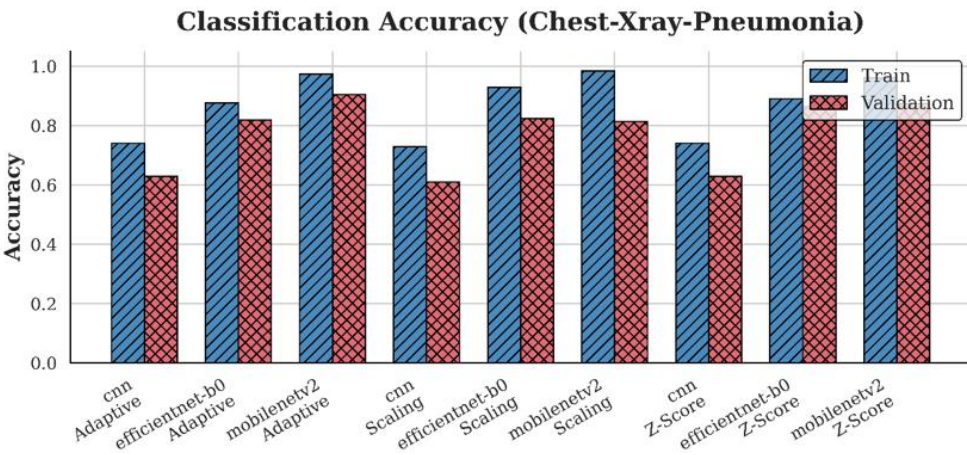


Figure 10. Train and Validation Accuracy for Chest-Xray-Pneumonia across all preprocessing strategies and model architectures.

4.2. Loss Analysis

Figures 11–14 and Table 5 compare training and validation loss for all experimental configurations. Adaptive normalization uniformly had lower or similar validation loss in most scenarios, with very strong performance on MobileNetV2 on ChestX-ray14 (0.71) and Chest-Xray-Pneumonia (0.29), the best overall value. This reflects better model calibration as well as convergence. While Z-score normalization sometimes matched or outperformed adaptive normalization (for example, MobileNetV2 on MIMIC-CXR with 0.69 compared with 0.74), it also had greater variance between loss on training and validation, potentially suggesting overfitting. In general, adaptive normalization supported more stable training behavior as well as better generalization in most setups, particularly with MobileNetV2.

Table 5. Summary of Validation Loss Across Datasets and Configurations.

Dataset	Preprocessing	Scaling	Z-score	Adaptive
ChestX-ray14	CNN	0.68	0.67	0.69
	EfficientNet-B0	0.66	0.67	0.68
	MobileNetV2	0.66	0.73	0.71
CheXpert	CNN	0.47	0.46	0.46
	EfficientNet-B0	0.42	0.40	0.44
	MobileNetV2	0.43	0.39	0.37
MIMIC-CXR	CNN	0.66	0.65	0.65
	EfficientNet-B0	0.70	0.66	0.68
	MobileNetV2	0.70	0.69	0.74
Chest-Xray-Pneumonia	CNN	0.72	0.70	0.70
	EfficientNet-B0	0.36	0.40	0.40
	MobileNetV2	0.45	0.40	0.29

*Lower loss values indicate better model calibration and convergence. All loss values represent averages over five experimental runs.

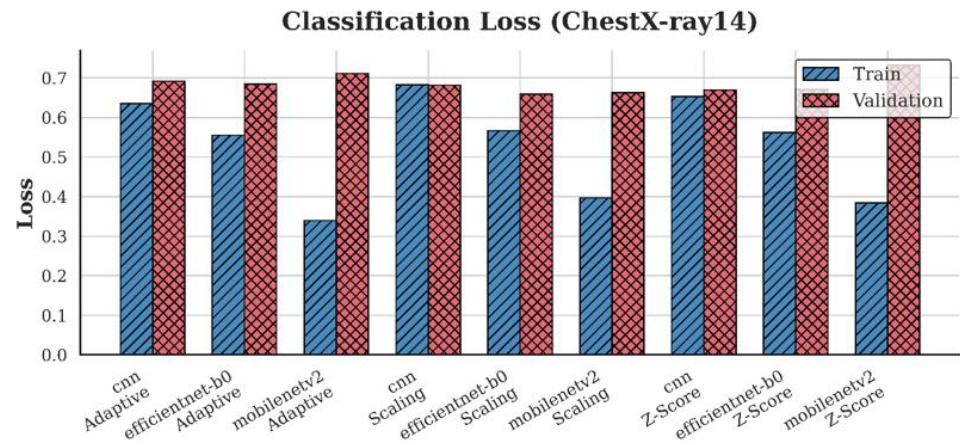


Figure 11. Train and Validation Loss for ChestX-ray14 across all preprocessing strategies and model architectures.

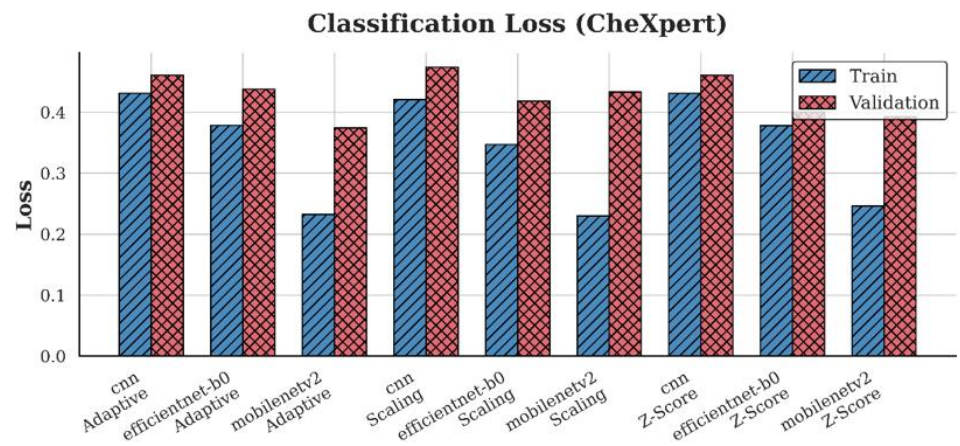


Figure 12. Train and Validation Loss for CheXpert across all preprocessing strategies and model architectures.

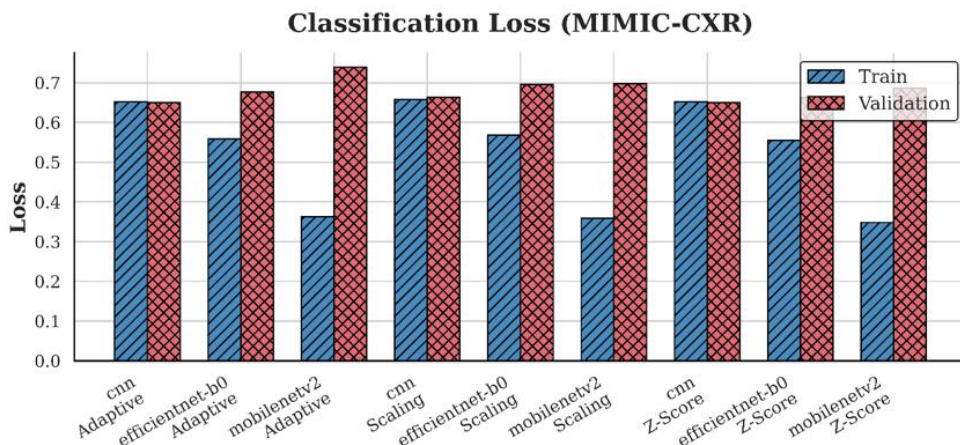


Figure 13. Train and Validation Loss for MIMIC-CXR across all preprocessing strategies and model architectures.

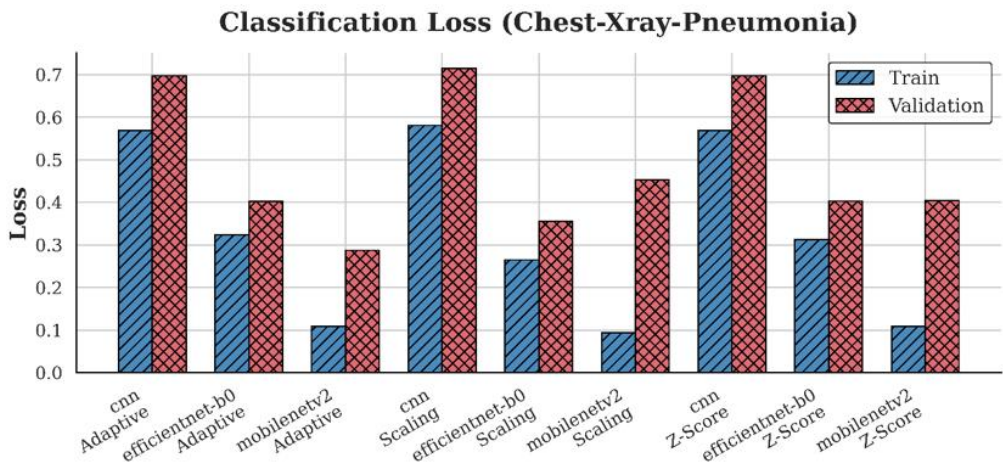


Figure 14. Train and Validation Loss for Chest-Xray-Pneumonia across all preprocessing strategies and model architectures.

4.3. F1-Score Analysis

Figures 15–18, as well as Table 6, show F1-performance on a variety of normalization methods and architectures. Adaptive normalization saw its best F1-performance (0.89) with MobileNetV2 on Chest-Xray-Pneumonia but was competitive on CheXpert (0.82) even if slightly worse than Z-score normalization (0.83). Though adaptive normalization demonstrated equivalent or better scores in some configurations specifically with CNN on ChestX-ray14 (0.59) its benefits were not across-the-board. In MIMIC-CXR, adaptive normalization’s F1-scores worsened somewhat on EfficientNet-B0 and MobileNetV2 (0.60), as compared with Z-score (0.64, 0.63, respectively). Across the board, then, what these results show is that there is potential for adaptive normalization to improve performance on F1, particularly with MobileNetV2, but its benefits are partly model and data-set dependent.

Table 6. Summary of Validation F1-scores Across Datasets and Configurations.

Dataset	Preprocessing	Scaling	Z-score	Adaptive
ChestX-ray14	CNN	0.57	0.58	0.59
	EfficientNet-B0	0.62	0.60	0.60
	MobileNetV2	0.64	0.62	0.62
CheXpert	CNN	0.84	0.85	0.85
	EfficientNet-B0	0.84	0.84	0.84
	MobileNetV2	0.79	0.83	0.82
MIMIC-CXR	CNN	0.65	0.64	0.64
	EfficientNet-B0	0.64	0.64	0.60
	MobileNetV2	0.62	0.63	0.60
Chest-Xray-Pneumonia	CNN	0.61	0.63	0.63
	EfficientNet-B0	0.81	0.84	0.81
	MobileNetV2	0.80	0.85	0.89

*Higher F1-scores indicate better model calibration and classification performance. All values represent averages over five experimental runs.

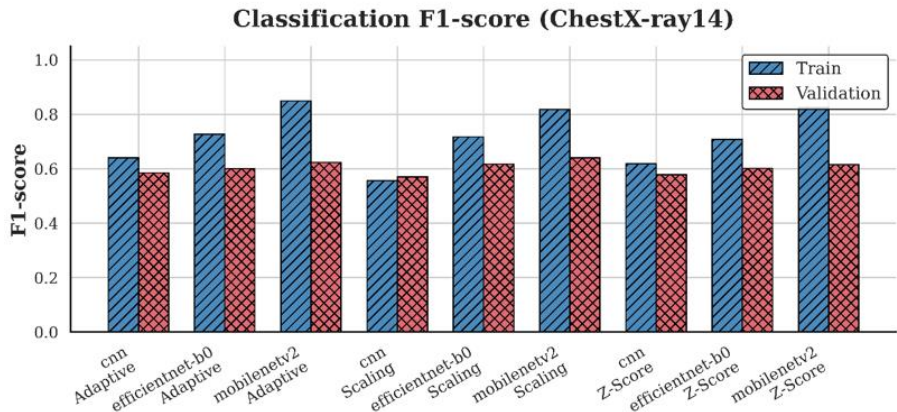


Figure 15. Train and Validation F1-scores for ChestX-ray14 across all preprocessing strategies and model architecture.

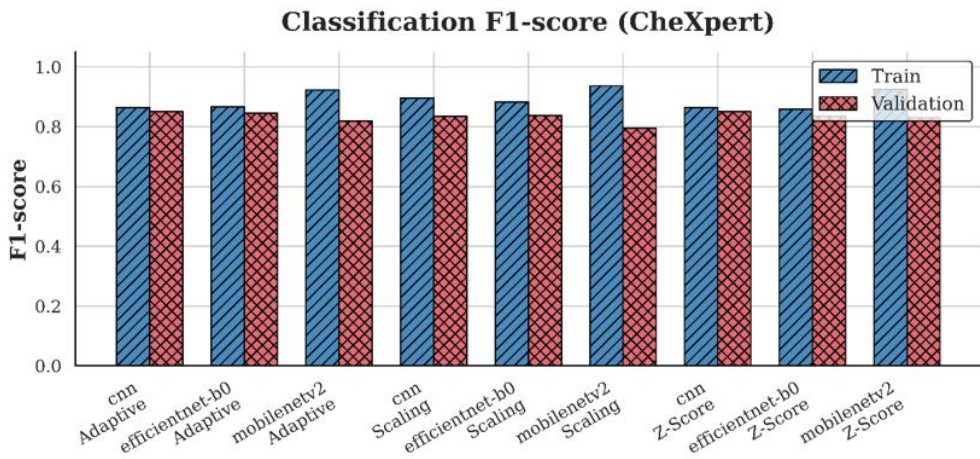


Figure 16. Train and Validation F1-scores for CheXpert across all preprocessing strategies and model architectures.

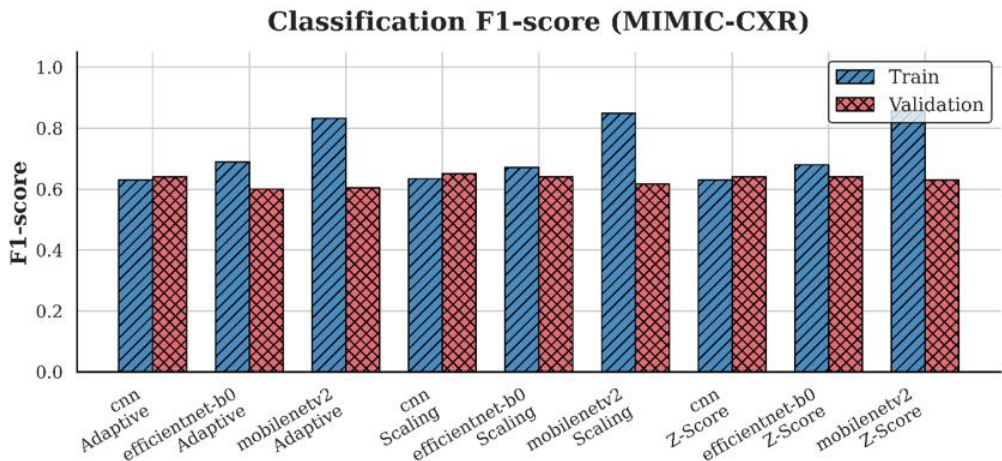


Figure 17. Train and Validation F1-scores for MIMIC-CXR across all preprocessing strategies and model architecture.

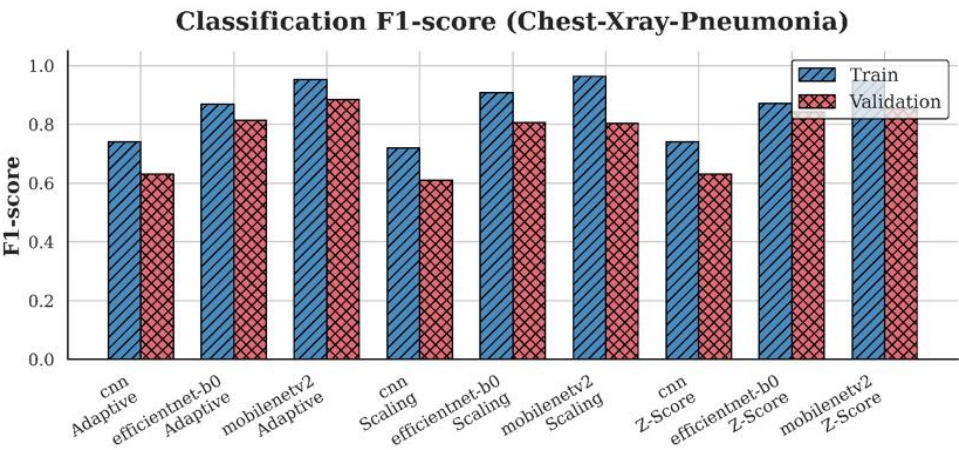


Figure 18. Train and Validation F1-scores for Chest-Xray-Pneumonia across all preprocessing strategies and model architectures.

4.4. Synergistic Effects of Architecture and Normalization

The results underline a synergistic interaction between MobileNetV2 and adaptive normalization. The combination utilized MobileNetV2’s architectural effectiveness through depth wise separable convolutions as well as linear bottlenecks along with adaptive normalization’s dynamic input adjustment. Together, they realized stable training curves, high performance, as well as statistical significance as checked with Wilcoxon signed-rank tests

5. Discussion

Our observations establish that combining MobileNetV2 with adaptive normalization always provides the best performance on all datasets. This is consistent with previous research stating how adaptive preprocessing increases generalization under heterogeneity. Adjusting intensity distributions on a per-batch basis seems important in reducing model drift and domain shift. Contrasting with these, scaling and normalization using z-scores were unable to capture localized intensity patterns, as experienced with MIMIC-CXR. These methods underperformed whenever pixel distribution assumptions between training environments and testing environments were significantly different. Moreover, MobileNetV2’s accuracy and efficiency lend it well for use in resource-constrained clinical environments. Along with adaptive normalization, the model can make decisions in real-time, fitting with current needs in point-of-care imaging. These results affirm that preprocessing is not a technical detail but a clinical enabler with potential for extending diagnostic coverage in underserved healthcare settings.

In addition, our findings provide a key entry point for incorporating adaptive normalization into federated learning (FL) architectures in order to ensure data confidentiality while retaining performance across institutional boundaries. Equally, combining this approach with self-supervised learning (SSL) presents an optimistic route toward using unlabeled radiographs as a way forward for strong pretraining. Adaptive normalization can be optimized even more toward contrastive or curriculum-based goals, enabling more robust model initialization in FL as well as SSL frameworks. To further promote transparency, future research will include explainability techniques such as Grad-CAM to make decision pathways more visible, supporting clinician trust. This is consistent with the trend toward interpretable AI in radiology.

Nonetheless, even with its strong performance, adaptive normalization adds training overhead caused by computation of per-batch statistics. In addition, its performance can be inconsistent under extreme shifts in a different domain not included in this study. Its robustness should be tested under unseen real-world clinical datasets in future research, as should other lightweight alternatives with balanced performance versus computational cost.

6. Conclusion

This study presents empirical evidence that, attesting that adaptive normalization considerably enhances deep learning model generalization in chest X-ray classification. Adaptive normalization with MobileNetV2 gave uniformly better performance on four heterogeneous datasets, exhibiting significant robustness as well as convergence stability, especially in skewed as well as heterogeneous environments. By addressing two persistent challenges in chest X-ray modeling the inaccurate localization of the region of interest (ROI) and inconsistencies in image quality across datasets adaptive normalization plays a crucial role in improving both focus and feature distribution uniformity across domains. Amongst techniques under consideration, adaptive normalization proved superior to conventional methods in terms of sustaining better consistency in F1-score, as well as diminishing overfitting. Results affirm a movement from using static preprocessing methods towards dynamic context-aware normalization techniques for improving cross-domain generalization. Important contributions of this research are: (1) a cross-dataset, cross-architecture benchmarking framework, (2) statistical verification of normalization performance, and (3) a pipeline implemented in a clinical resource-constrained environment. Upcoming research directions include incorporating adaptive normalization into federated and self-supervised learning frameworks in order to provide privacy-preserving, scalable, and interpretable AI for real-world radiographic diagnostics.

CRedit authorship contribution statement: **Jatsada Singthongchai:** Formal analysis, Writing– original draft, Software, Methodology, Project administration and Supervision, Writing– review & editing. **Tanachapong Wangkhamhan:** Data curation, Formal analysis, Methodology, Resources, Validation, Writing– review & editing, Writing– review & editing.

Funding: This research was not funded.

Ethical approval: This article does not contain any studies with human participants performed by any of the authors.

Declaration of Competing Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Generative AI and AI-assisted technologies in the writing process: While preparing this work, the author(s) used ChatGPT-5 to improve readability and language as per Elsevier's policy/guideline on the use of GenAI for authors¹. After using this tool/service, the author(s) carefully reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

ORCID: Jatsada Singthongchai <http://orcid.org/0009-0004-3884-2876> Tanachapong Wangkhamhan <http://orcid.org/0000-0002-3371-8364>

References

1. Oh, Y.; Park, S.; Ye, J.C. Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. *IEEE Trans Med Imaging* 2020, 39, 2688–2700, doi:10.1109/TMI.2020.2993291.
2. Padmavathi, V.; Ganesan, K. Metaheuristic Optimizers Integrated with Vision Transformer Model for Severity Detection and Classification via Multimodal COVID-19 Images. *Sci Rep* 2025, 15, doi:10.1038/s41598-025-98593-w.
3. Aksoy, B.; Salman, O.K.M. Detection of COVID-19 Disease in Chest X-Ray Images with Capsul Networks: Application with Cloud Computing. *Journal of Experimental & Theoretical Artificial Intelligence* 2021, 33, 527–541, doi:10.1080/0952813X.2021.1908431.

4. Khan, A.; Khan, S.H.; Saif, M.; Batool, A.; Sohail, A.; Waleed Khan, M. A Survey of Deep Learning Techniques for the Analysis of COVID-19 and Their Usability for Detecting Omicron. *Journal of Experimental & Theoretical Artificial Intelligence* 2024, 36, 1779–1821, doi:10.1080/0952813X.2023.2165724.
5. Attaullah, M.; Ali, M.; Almufareh, M.F.; Ahmad, M.; Hussain, L.; Jhanjhi, N.; Humayun, M. Initial Stage COVID-19 Detection System Based on Patients' Symptoms and Chest X-Ray Images. *Applied Artificial Intelligence* 2022, 36, 2055398, doi:10.1080/08839514.2022.2055398.
6. Marikkar, U.; Atito, S.; Awais, M.; Mahdi, A. LT-ViT: A Vision Transformer for Multi-Label Chest X-Ray Classification. In Proceedings of the 2023 IEEE International Conference on Image Processing (ICIP); 2023; pp. 2565–2569.
7. Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K.; et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. *Computer Vision and Pattern Recognition* 2017, doi:10.48550/arXiv.1711.05225.
8. Zhu, Q.; Bai, H.; Sun, J.; Cheng, C.; Li, X. LPAdaIN: Light Progressive Attention Adaptive Instance Normalization Model for Style Transfer. *Electronics (Switzerland)* 2022, 11, doi:10.3390/electronics11182929.
9. Shin, H.; Kim, T.; Park, J.; Raj, H.; Jabbar, M.S.; Abebaw, Z.D.; Lee, J.; Van, C.C.; Kim, H.; Shin, D. Pulmonary Abnormality Screening on Chest X-Rays from Different Machine Specifications: A Generalized AI-Based Image Manipulation Pipeline. *Eur Radiol Exp* 2023, 7, doi:10.1186/s41747-023-00386-1.
10. Karaki, A.A.; Alrawashdeh, T.; Abusaleh, S.; Alksasbeh, M.Z.; Alqudah, B.; Alemerien, K.; Alshamaseen, H. Pulmonary Edema and Pleural Effusion Detection Using EfficientNet-V1-B4 Architecture and AdamW Optimizer from Chest X-Rays Images. *Computers, Materials and Continua* 2024, 80, 1055–1073, doi:10.32604/cmc.2024.051420.
11. Wangkhamhan, T. Adaptive Chaotic Satin Bowerbird Optimization Algorithm for Numerical Function Optimization. *Journal of Experimental & Theoretical Artificial Intelligence* 2021, 33, 719–746, doi:https://doi.org/10.1080/0952813X.2020.1785018.
12. Demircioğlu, A. The Effect of Feature Normalization Methods in Radiomics. *Insights Imaging* 2024, 15, doi:10.1186/s13244-023-01575-7.
13. Rayed, M.E.; Islam, S.M.S.; Niha, S.I.; Jim, J.R.; Kabir, M.M.; Mridha, M.F. Deep Learning for Medical Image Segmentation: State-of-the-Art Advancements and Challenges. *Inform Med Unlocked* 2024, 47.
14. Gangwar, S.; Devi, R.; Mat Isa, N.A. Optimized Exposer Region-Based Modified Adaptive Histogram Equalization Method for Contrast Enhancement in CXR Imaging. *Sci Rep* 2025, 15, doi:10.1038/s41598-025-90876-6.
15. Tomar, D.; Lortkipanidze, M.; Vray, G.; Bozorgtabar, B.; Thiran, J.-P. Self-Attentive Spatial Adaptive Normalization for Cross-Modality Domain Adaptation. *IEEE Trans Med Imaging* 2021, 40, 2926–2938, doi:10.1109/TMI.2021.3059265.
16. Luo, Z.; Luo, X.; Gao, Z.; Wang, G. An Uncertainty-Guided Tiered Self-Training Framework for Active Source-Free Domain Adaptation in Prostate Segmentation. *Computer Vision and Pattern Recognition* 2024, 1–11, doi:https://doi.org/10.48550/arXiv.2407.02893.
17. Fan, D.P.; Zhou, T.; Ji, G.P.; Zhou, Y.; Chen, G.; Fu, H.; Shen, J.; Shao, L. Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images. *IEEE Trans Med Imaging* 2020, 39, 2626–2637, doi:10.1109/TMI.2020.2996645.
18. Irvin, J.; Rajpurkar, P.; Ko, M.; Yu, Y.; Ciurea-Ilcus, S.; Chute, C.; Marklund, H.; Haghighoo, B.; Ball, R.; Shpanskaya, K.; et al. CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison. *AAAI'19/IAAI'19/EAAI'19: Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI*

- Symposium on Educational Advances in Artificial Intelligence* 2019, 33, 590–597, doi:https://doi.org/10.1609/aaai.v33i01.3301590.
19. Johnson, A.E.W.; Pollard, T.J.; Berkowitz, S.J.; Greenbaum, N.R.; Lungren, M.P.; Deng, C. ying; Mark, R.G.; Horng, S. MIMIC-CXR, a de-Identified Publicly Available Database of Chest Radiographs with Free-Text Reports. *Sci Data* 2019, 6, doi:10.1038/s41597-019-0322-0.
 20. Gazda, M.; Plavka, J.; Gazda, J.; Drotar, P. Self-Supervised Deep Convolutional Neural Network for Chest X-Ray Classification. *IEEE Access* 2021, 9, 151972–151982, doi:10.1109/ACCESS.2021.3125324.
 21. Rahman, T.; Chowdhury, M.E.H.; Khandakar, A.; Islam, K.R.; Islam, K.F.; Mahbub, Z.B.; Kadir, M.A.; Kashem, S. Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-Ray. *Applied Sciences (Switzerland)* 2020, 10, doi:10.3390/app10093233.
 22. Kermayn, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.S.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* 2018, 172, 1122–1131.e9, doi:10.1016/j.cell.2018.02.010.
 23. Wang, L.; Lin, Z.Q.; Wong, A. COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images. *Sci Rep* 2020, 10, doi:10.1038/s41598-020-76550-z.
 24. Ait Nasser, A.; Akhloufi, M.A. A Review of Recent Advances in Deep Learning Models for Chest Disease Detection Using Radiography. *Diagnostics* 2023, 13.
 25. Öztürk, Ş.; Turah, M.Y.; Çukur, T. HydraViT: Adaptive Multi-Branch Transformer for Multi-Label Disease Classification from Chest X-Ray Images. *Biomed Signal Process Control* 2023, 100, 1–10, doi:https://doi.org/10.1016/j.bspc.2024.106959.
 26. Dede, A.; Nunoo-Mensah, H.; Tchao, E.T.; Agbemenu, A.S.; Adjei, P.E.; Acheampong, F.A.; Kponyo, J.J. Deep Learning for Efficient High-Resolution Image Processing: A Systematic Review. *Intelligent Systems with Applications* 2025, 26.
 27. Ahmad, I.S.; Li, N.; Wang, T.; Liu, X.; Dai, J.; Chan, Y.; Liu, H.; Zhu, J.; Kong, W.; Lu, Z.; et al. COVID-19 Detection via Ultra-Low-Dose X-Ray Images Enabled by Deep Learning. *Bioengineering* 2023, 10, doi:10.3390/bioengineering10111314.
 28. Oltu, B.; Güney, S.; Yuksel, S.E.; Dengiz, B. Automated Classification of Chest X-Rays: A Deep Learning Approach with Attention Mechanisms. *BMC Med Imaging* 2025, 25, doi:10.1186/s12880-025-01604-5.
 29. Saad, M.M.; Rehmani, M.H.; O'Reilly, R. Addressing the Intra-Class Mode Collapse Problem Using Adaptive Input Image Normalization in GAN-Based X-Ray Images. In Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); 2022; pp. 2049–2052.
 30. Reinhold, J.C.; Dewey, B.E.; Carass, A.; Prince, J.L. Evaluating the Impact of Intensity Normalization on MR Image Synthesis. *Proc SPIE Int Soc Opt Eng* 2018, 10949, doi:10.1117/12.2513089.
 31. Huang, X.; Belongie, S. Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV); 2017; pp. 1510–1519.
 32. R, V.; Kumar, A.; Kumar, A.; Ashok Kumar, V.D.; K, R.; Kumar, V.D.A.; Jilani Saudagar, A.K.; A, A. COVIDPRO-NET: A Prognostic Tool to Detect COVID 19 Patients from Lung X-Ray and CT Images Using Transfer Learning and Q-Deformed Entropy. *Journal of Experimental & Theoretical Artificial Intelligence* 2023, 35, 473–488, doi:10.1080/0952813X.2021.1949755.
 33. Albert, S.; Wichtmann, B.D.; Zhao, W.; Maurer, A.; Hesser, J.; Attenberger, U.I.; Schad, L.R.; Zöllner, F.G. Comparison of Image Normalization Methods for Multi-Site Deep Learning. *Applied Sciences (Switzerland)* 2023, 13, doi:10.3390/app13158923.

34. Al-Waisy, A.S.; Mohammed, M.A.; Al-Fahdawi, S.; Maashi, M.S.; Garcia-Zapirain, B.; Abdulkareem, K.H.; Mostafa, S.A.; Kumar, N.M.; Le, D.N. COVID-DeepNet: Hybrid Multimodal Deep Learning System for Improving COVID-19 Pneumonia Detection in Chest X-Ray Images. *Computers, Materials and Continua* 2021, 67, 2409–2429, doi:10.32604/cmc.2021.012955.
35. Nuruddin Bin Azhar, A.; Sani, N.S.; Luan, L.; Wei, X. Enhancing COVID-19 Detection in X-Ray Images Through Deep Learning Models with Different Image Preprocessing Techniques. *IJACSA) International Journal of Advanced Computer Science and Applications* 2025, 16, 633–644, doi:http://dx.doi.org/10.14569/IJACSA.2025.0160161.
36. Toğaçar, M.; Ergen, B.; Cömert, Z. COVID-19 Detection Using Deep Learning Models to Exploit Social Mimic Optimization and Structured Chest X-Ray Images Using Fuzzy Color and Stacking Approaches. *Comput Biol Med* 2020, 121, doi:10.1016/j.compbiomed.2020.103805.
37. Sanida, T.; Dasygenis, M. A Novel Lightweight CNN for Chest X-Ray-Based Lung Disease Identification on Heterogeneous Embedded System. *Applied Intelligence* 2024, 54, 4756–4780, doi:10.1007/s10489-024-05420-2.
38. Sriwiboon, N. Efficient and Lightweight CNN Model for COVID-19 Diagnosis from CT and X-Ray Images Using Customized Pruning and Quantization Techniques. *Neural Comput Appl* 2025, 37, 13059–13078, doi:10.1007/s00521-025-11219-0.
39. Chicco, D.; Jurman, G. The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. *BMC Genomics* 2020, 21, doi:10.1186/s12864-019-6413-7.
40. Bani Baker, Q.; Hammad, M.; Al-Smadi, M.; Al-Jarrah, H.; Al-Hamouri, R.; Al-Zboon, S.A. Enhanced COVID-19 Detection from X-Ray Images with Convolutional Neural Network and Transfer Learning. *J Imaging* 2024, 10, doi:10.3390/jimaging10100250.
41. Apostolopoulos, I.D.; Mpesiana, T.A. Covid-19: Automatic Detection from X-Ray Images Utilizing Transfer Learning with Convolutional Neural Networks. *Phys Eng Sci Med* 2020, 43, 635–640, doi:10.1007/s13246-020-00865-4.
42. Chowdhury, N.K.; Rahman, Md.M.; Kabir, M.A. PDCOVIDNet: A Parallel-Dilated Convolutional Neural Network Architecture for Detecting COVID-19 from Chest X-Ray Images. 2020, doi:10.1007/s13755-020-00119-3.
43. Öztürk, Ş.; Turalı, M.Y.; Çukur, T. HydraViT: Adaptive Multi-Branch Transformer for Multi-Label Disease Classification from Chest X-Ray Images. 2023.
44. Wang, L.; Lin, Z.Q.; Wong, A. COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images. *Sci Rep* 2020, 10, doi:10.1038/s41598-020-76550-z.
45. Rajpurkar, P.; Irvin, J.; Ball, R.L.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.P.; et al. Deep Learning for Chest Radiograph Diagnosis: A Retrospective Comparison of the CheXNeXt Algorithm to Practicing Radiologists. *PLoS Med* 2018, 15, e1002686-.
46. Tan, M.; Le, Q. V *EfficientNetV2: Smaller Models and Faster Training*; 2021;
47. Çallı, E.; Sogancioglu, E.; van Ginneken, B.; van Leeuwen, K.G.; Murphy, K. Deep Learning for Chest X-Ray Analysis: A Survey. *Med Image Anal* 2021, 72, 102125, doi:https://doi.org/10.1016/j.media.2021.102125.
48. Fu, X.; Lin, R.; Du, W.; Tavares, A.; Liang, Y. Explainable Hybrid Transformer for Multi-Classification of Lung Disease Using Chest X-Rays. *Sci Rep* 2025, 15, doi:10.1038/s41598-025-90607-x.
49. Bani Baker, Q.; Hammad, M.; Al-Smadi, M.; Al-Jarrah, H.; Al-Hamouri, R.; Al-Zboon, S.A. Enhanced COVID-19 Detection from X-Ray Images with Convolutional Neural Network and Transfer Learning. *J Imaging* 2024, 10, doi:10.3390/jimaging10100250.

50. Yan, Z.; Li, X.; Li, M.; Zuo, W.; Shan, S. Shift-Net: Image Inpainting via Deep Feature Rearrangement. In Proceedings of the Computer Vision – ECCV 2018; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Springer International Publishing: Cham, 2018; pp. 3–19.
51. Gazda, M.; Plavka, J.; Gazda, J.; Drotar, P. Self-Supervised Deep Convolutional Neural Network for Chest X-Ray Classification. *IEEE Access* 2021, 9, 151972–151982, doi:10.1109/ACCESS.2021.3125324.
52. Pavlova, M.; Terhijan, N.; Chung, A.G.; Zhao, A.; Surana, S.; Aboutaleb, H.; Gunraj, H.; Sabri, A.; Alaref, A.; Wong, A. COVID-Net CXR-2: An Enhanced Deep Convolutional Neural Network Design for Detection of COVID-19 Cases From Chest X-Ray Images. *Front Med (Lausanne)* 2022, 9, doi:10.3389/fmed.2022.861680.
53. Shin, H.; Kim, T.; Park, J.; Raj, H.; Jabbar, M.S.; Abebaw, Z.D.; Lee, J.; Van, C.C.; Kim, H.; Shin, D. Pulmonary Abnormality Screening on Chest X-Rays from Different Machine Specifications: A Generalized AI-Based Image Manipulation Pipeline. *Eur Radiol Exp* 2023, 7, doi:10.1186/s41747-023-00386-1.
54. Gunraj, H.; Wang, L.; Wong, A. COVIDNet-CT: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases From Chest CT Images. *Front Med (Lausanne)* 2020, 7, doi:10.3389/fmed.2020.608525.
55. Fan, D.P.; Zhou, T.; Ji, G.P.; Zhou, Y.; Chen, G.; Fu, H.; Shen, J.; Shao, L. Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images. *IEEE Trans Med Imaging* 2020, 39, 2626–2637, doi:10.1109/TMI.2020.2996645.
56. Oh, Y.; Park, S.; Ye, J.C. Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. *IEEE Trans Med Imaging* 2020, 39, 2688–2700, doi:10.1109/TMI.2020.2993291.
57. Rahman, T.; Chowdhury, M.E.H.; Khandakar, A.; Islam, K.R.; Islam, K.F.; Mahbub, Z.B.; Kadir, M.A.; Kashem, S. Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-Ray. *Applied Sciences (Switzerland)* 2020, 10, doi:10.3390/app10093233.
58. Gunraj, H.; Wang, L.; Wong, A. COVIDNet-CT: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases From Chest CT Images. *Front Med (Lausanne)* 2020, 7, doi:10.3389/fmed.2020.608525.
59. Toğaçar, M.; Ergen, B.; Cömert, Z. COVID-19 Detection Using Deep Learning Models to Exploit Social Mimic Optimization and Structured Chest X-Ray Images Using Fuzzy Color and Stacking Approaches. *Comput Biol Med* 2020, 121, doi:10.1016/j.compbiomed.2020.103805.
60. Ozturk, T.; Talo, M.; Yildirim, E.A.; Baloglu, U.B.; Yildirim, O.; Rajendra Acharya, U. Automated Detection of COVID-19 Cases Using Deep Neural Networks with X-Ray Images. *Comput Biol Med* 2020, 121, 103792, doi:https://doi.org/10.1016/j.compbiomed.2020.103792.
61. Al-Waisy, A.S.; Mohammed, M.A.; Al-Fahdawi, S.; Maashi, M.S.; Garcia-Zapirain, B.; Abdulkareem, K.H.; Mostafa, S.A.; Kumar, N.M.; Le, D.N. COVID-DeepNet: Hybrid Multimodal Deep Learning System for Improving COVID-19 Pneumonia Detection in Chest X-Ray Images. *Computers, Materials and Continua* 2021, 67, 2409–2429, doi:10.32604/cmc.2021.012955.
62. Vinod, D.N.; Jeyavadhanam, B.R.; Zungeru, A.M.; Prabakaran, S.R.S. Fully Automated Unified Prognosis of Covid-19 Chest X-Ray/CT Scan Images Using Deep Covix-Net Model. *Comput Biol Med* 2021, 136, doi:10.1016/j.compbiomed.2021.104729.
63. Marikkar, U.; Atito, S.; Awais, M.; Mahdi, A. LT-ViT: A Vision Transformer for Multi-Label Chest X-Ray Classification. 2023, doi:10.1109/ICIP49359.2023.10222175.
64. Karaki, A.A.; Alrawashdeh, T.; Abusaleh, S.; Alksasbeh, M.Z.; Alqudah, B.; Alemerien, K.; Alshamaseen, H. Pulmonary Edema and Pleural Effusion Detection Using EfficientNet-V1-B4 Architecture and AdamW Optimizer from Chest X-Rays Images. *Computers, Materials and Continua* 2024, 80, 1055–1073, doi:10.32604/cmc.2024.051420.

65. Sanida, T.; Dasygenis, M. A Novel Lightweight CNN for Chest X-Ray-Based Lung Disease Identification on Heterogeneous Embedded System. *Applied Intelligence* 2024, 54, 4756–4780, doi:10.1007/s10489-024-05420-2.
66. Yen, C.T.; Tsao, C.Y. Lightweight Convolutional Neural Network for Chest X-Ray Images Classification. *Sci Rep* 2024, 14, doi:10.1038/s41598-024-80826-z.
67. Hage Chehade, A.; Abdallah, N.; Marion, J.M.; Hatt, M.; Oueidat, M.; Chauvet, P. Reconstruction-Based Approach for Chest X-Ray Image Segmentation and Enhanced Multi-Label Chest Disease Classification. *Artif Intell Med* 2025, 165, doi:10.1016/j.artmed.2025.103135.
68. Padmavathi, V.; Ganesan, K. Metaheuristic Optimizers Integrated with Vision Transformer Model for Severity Detection and Classification via Multimodal COVID-19 Images. *Sci Rep* 2025, 15, doi:10.1038/s41598-025-98593-w.
69. Shati, A.; Hassan, G.M.; Datta, A. A Comprehensive Fusion Model for Improved Pneumonia Prediction Based on KNN-Wavelet-GLCM and a Residual Network. *Intelligent Systems with Applications* 2025, 26, doi:10.1016/j.iswa.2025.200492.
70. Sriwiboon, N. Efficient and Lightweight CNN Model for COVID-19 Diagnosis from CT and X-Ray Images Using Customized Pruning and Quantization Techniques. *Neural Comput Appl* 2025, doi:10.1007/s00521-025-11219-0.
71. Albert, S.; Wichtmann, B.D.; Zhao, W.; Maurer, A.; Hesser, J.; Attenberger, U.I.; Schad, L.R.; Zöllner, F.G. Comparison of Image Normalization Methods for Multi-Site Deep Learning. *Applied Sciences (Switzerland)* 2023, 13, doi:10.3390/app13158923.
72. Demircioğlu, A. The Effect of Feature Normalization Methods in Radiomics. *Insights Imaging* 2024, 15, doi:10.1186/s13244-023-01575-7.
73. Oh, Y.; Park, S.; Ye, J.C. Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. *IEEE Trans Med Imaging* 2020, 39, 2688–2700, doi:10.1109/TMI.2020.2993291.
74. Chicco, D.; Jurman, G. The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. *BMC Genomics* 2020, 21, doi:10.1186/s12864-019-6413-7.
75. Demšar, J. *Statistical Comparisons of Classifiers over Multiple Data Sets*; 2006; Vol. 7;.
76. Reinhold, J.C.; Dewey, B.E.; Carass, A.; Prince, J.L. Evaluating the Impact of Intensity Normalization on MR Image Synthesis. 2018.
77. Banik, P.; Majumder, R.; Mandal, A.; Dey, S.; Mandal, M. A Computational Study to Assess the Polymorphic Landscape of Matrix Metalloproteinase 3 Promoter and Its Effects on Transcriptional Activity. *Comput Biol Med* 2022, 145, 105404, doi:https://doi.org/10.1016/j.compbiomed.2022.105404.
78. Saad, M.M.; Rehmani, M.H.; O'reilly, R. Addressing the Intra-Class Mode Collapse Problem Using Adaptive Input Image Normalization in GAN-Based X-Ray Images. In Proceedings of the Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS; Institute of Electrical and Electronics Engineers Inc., 2022; Vol. 2022-July, pp. 2049–2052.
79. Tan, M.; Le, Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Mach Learn* 2020, doi:https://doi.org/10.48550/arXiv.1905.11946.
80. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2018; pp. 4510–4520.
81. Bougourzi, F.; Dornaika, F.; Distant, C.; Taleb-Ahmed, A. D-TrAttUnet: Toward Hybrid CNN-Transformer Architecture for Generic and Subtle Segmentation in Medical Images. *Comput Biol Med* 2024, 176, 108590, doi:https://doi.org/10.1016/j.compbiomed.2024.108590.

82. Irvin, J.; Rajpurkar, P.; Ko, M.; Yu, Y.; Ciurea-Illcus, S.; Chute, C.; Marklund, H.; Haghighi, B.; Ball, R.; Shpanskaya, K.; et al. *CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison*;
83. Johnson, A.E.W.; Pollard, T.J.; Berkowitz, S.J.; Greenbaum, N.R.; Lungren, M.P.; Deng, C. ying; Mark, R.G.; Horng, S. MIMIC-CXR, a de-Identified Publicly Available Database of Chest Radiographs with Free-Text Reports. *Sci Data* 2019, 6, doi:10.1038/s41597-019-0322-0.
84. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.S.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* 2018, 172, 1122-1131.e9, doi:10.1016/j.cell.2018.02.010.
85. Tan, M.; Le, Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. 2019.
86. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A Survey on Deep Learning in Medical Image Analysis. *Med Image Anal* 2017, 42, 60–88, doi:https://doi.org/10.1016/j.media.2017.07.005.
87. Zhou, Z.; Sodha, V.; Pang, J.; Gotway, M.B.; Liang, J. Models Genesis. *Med Image Anal* 2021, 67, 101840, doi:https://doi.org/10.1016/j.media.2020.101840.
88. Rayed, Md.E.; Islam, S.M.S.; Niha, S.I.; Jim, J.R.; Kabir, M.M.; Mridha, M.F. Deep Learning for Medical Image Segmentation: State-of-the-Art Advancements and Challenges. *Inform Med Unlocked* 2024, 47, 101504, doi:https://doi.org/10.1016/j.imu.2024.101504.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.