Received: May 15, 2024. Revised: July 1, 2024.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Multiclass Classification of Chest X Ray Images Using Resnet-152 with Mix-Up Regularization Method

Roshan Shetty[1] Prasad Narasimha Sarappadi[2]* Neti Ramalingappa Pradeep[3]
Bittasandra Sachidananda Murthy Kavya[4] Udaya Kumar Sudhakar[1]

[1]*Department of ECE, Alva's Institute of Engineering & Technology, Moodbidri, India*
[2]*Department of Electrical and Electronics Engineering, Manipal Institute of Technology, Bengaluru, India*
[3]*Department of Electronics and Communication Engineering,*
Indian Institute of Information Technology, Dharwad, India
[4]*Department of Electronics and Communication, Manipal Institute of Technology, Bengaluru, India*
* Corresponding author's Email: sn.prasad@manipal.edu

Abstract: For the past few decades, thoracic disease classification using Chest X-Ray images has remained challenging in Artificial Intelligence (AI). The traditional Deep CNNs (DCNNs) always predict with overconfidence, so the accuracy of the architectures is likely to be lower than predictive scores. Also, DCNNs suffer from overfitting due to the large number of parameters, and they require immense data augmentation to reduce overfitting. To overcome these problems, this research implements ResNet-152 with a Mix-Up based regularization method to effectively classify multiclass Chest X-Ray images of the Chest X-ray14 dataset first time. It contains 112,120 Chest X-Ray images across 30,805 patients, with 14 common thoracic disease labels plus one Normal (Healthy) label. The Mix-Up technique creates new samples as a convex combination of training points and disease labels. This Mix-Up-based regularization trains a model by blending image pair and their associated labels. We found that Mix-Up is successful in improving the accuracy of the ResNet-152 model by reducing its overconfidence up to 0.002 and increasing average AUC to 99.2% across all 14 disease classes. The proposed model performed better than ResNet-152 with other regularization methods such as Cut-Out, Cut-Mix, and Aug-Mix Regularization. Compared to the other existing models such as CNN-ELM, PCSANet, and GWSA &LCD models, this model achieved higher average AUC values across all 14 pathologies.

Keywords: Chest X-ray14, Mix-up regularization, Multiclass classification, ResNet-152.

1. Introduction

The emergence of the coronavirus pandemic in December 2019 has led to a global health crisis in significant proportions [1-2]. The common thoracic disease pneumonia occurs due to some different pathogens like fungi, viruses, and bacteria [3]. A chest X-ray (CXR) is generated by exposing the chest to a small dose of ionizing radiation used for observing the state of the heart, mediastinum, lung tissue, thoracic cage, and chest cavity [4]. As the demand for thoroughly analyzed radiology units rises with an increasing number of patients, CXR images are highly sought after due to their considerable sensitivity across a wide range, thereby minimizing radiation exposure and cost-effectiveness [5]. Numerous diseases and medical disorders are also investigated with the CXR for screening, diagnosis, and treatment of lung illness [6-7]. A variety of diseases has already placed a huge pressure on the world's healthcare systems. Artificial intelligence (AI) systems provide an alternative method for the automated diagnosis of infections and the differentiation of such infections from other illnesses [7].

In medical image analysis, the deep learning (DL) method performs better than classical machine learning on tasks involving detection, segmentation, and classification of images. In DCNNs, accuracy is

likely to be lower than predicted score called as overconfidence of DCNN architectures. Most modern DNNs, models are trained to be overconfident because of their good supervised learning settings, and one hot encoded label with zero entropy signal. Approaches like label smoothing and its affects somehow reduced overconfidence [8]. The mix-up regularization is an efficient data diagnostic technique which solves the problem of corrupted label memorization [9]. A Mix-up is an implicit regularization effect that enables models to generalize better [10]. The proposed ResNet-152 model with Mix-up regularization is trained by convexly joining random pairs of images along with their labels. This process generates additional synthetic samples, serving as an effective method of data augmentation [11]. Here in this work, ResNet-152 with the Mix-up regularization method performed an effective classification of Chest X-ray14 multiclass images. This Chest X-Ray14 dataset has 14 thoracic disease labels and a Healthy label [12].

The main contributions of this research work are:

- The implemented ResNet-152 with the Mix-up regularization technique achieves a new state-of-the-art disease classification, as compared to other methods on the Chest X-Ray14 dataset.
- The overconfidence of the ResNet-152 increases due to data imbalance issue of Chest X-Ray14 dataset, which lowers classification accuracy of model. It is solved by using Mix-up regularization on ResNet-152.
- Superior results are obtained from the implemented ResNet-152 with Mix-up regularization method, as compared with other regularization methods such as ResNet-152 with AugMix regularization [13], Cutmix regularization [14], and cutout regularization [15] methods based on the 14 disease classes.

This research paper is organized as follows: Section 2 explains the literature review and Section 3 describes the proposed method. Section 4 demonstrates the outcomes and comparison, and Section 5 explains the conclusion.

2. Literature survey

Some literature surveys for multiclass classification of chest X-ray images using other methods are given below.

Shamrat [16] implemented a fine-tuned MobileLungNetV2 model from another DCNN transfer-learning model called MobileNetV2, which was utilized for the multiclass classification of chest X-ray14 images. This MobileLungNetV2 model extracted better features and successfully identified abnormalities in the lungs. The experimental results showed that this fine-tuned model achieved a higher classification accuracy of an average AUC of 91.6%. However, this model had time complexity issues with high computational demands.

Nahiduzzaman [17] implemented a convolutional neural network with an Extreme Learning Machine (CNN-ELM) method for multiclass categorization of CXR images. The implemented method is integrated with the lightweight parallel capability of CNN and ELM classification. As a result, this integrated CNN-ELM method was successfully achieved and the 17 kind of lung diseases are identified. Nonetheless, after the integration with multiple Chest X-ray datasets, some categories of images suffered in the CNN-ELM method due to an insufficient quantity of training data.

Mann [18] implemented the Multi-Modal Fusion of Deep Transfer Learning (MMF-DTL) method employed to classify CXR images into six classes. This model used 3 DL specifically VGG16, Inception v3, and ResNet 50 for extensive feature extraction and softmax as the classifier. The outcomes established that the MMF-DTL method suffered from less training data especially while using ResNet-152, which is prone to overfitting.

Chen [19] implemented a pyramidal convolution module and shuffle attention module (PCSANet) based on a residual network model for 14 thorax disease classification. Pyramid convolution was utilized to extract features from images and shuffle attention enabled a focus on pathological features. This PCSANet model efficiently increased the thoracic disease classification performance. Acquiring additional multi-scale discriminative features of pathological abnormality improves the overall classification efficiency. Nonetheless, the AUC score of infiltration was identified as quite low.

Xu [20] implemented a Group-Wise Spatial Attention (GWSA) and Label Co-Occurrence Dependency (LCD) module for the classification of multiple diseases. The GWSA increased spatial features inside dissimilar groups and the LCD module did correlations between different thoracic abnormalities. This implemented method was employed with a DCNN called DenseNet. This model showed high categorization for distinct diseases once databases were adequate and composed. Due to low specificity, this implemented method had less performance in correctly identifying a negative class.

Wang [21] developed a Thorax-Net model for 14 thorax diseases by exploiting the Chest X-Ray14 dataset. This approach contained the ResNet-152

classification block and performed feature extraction, supported by the attention branch. The ResNet was exploited as a classification branch assisted by Gradient-weight Class Activation Maps (GradCAM). The performance of this ResNet-152-based Thorax-Net model was affected by overfitting due to imbalanced disease class distribution in the Chest X-Ray14 dataset.

Overall, the above literature surveys showed that most DNN architectures suffered from misclassification, imbalanced data distribution, overfitting, time complexity issues and high computational demands. While some categories of image results were still restricted due to an insufficient quantity of training data. On the other hand, the AUC score and specificity was identified as relatively low while classifying a negative class. So, they failed to obtain adequate biased energy to process efficient classification of 14 diseases. In order to overcome these drawbacks, this research implemented an overfitting prone ResNet-152 DCNN model with Mix-up regularization technique to address multiclass image classification problems of the Chest X-ray14 dataset.

3. Methodology

The implemented method's block diagram is illustrated in Fig. 1. In this paper, a chest X-ray14 dataset is used as the input image and a Gaussian filter is employed for image denoising, while CLAHE is used for image enhancement. The data augmentation is performed to enhance the size of the dataset. At last, ResNet-152 with mix-up regularization is utilized for the CXR image categorization process.

3.1 Dataset

This work utilized an early large-scale CXR database issued by the National Institutes of Health (NHI) known as the ChestX-ray14 dataset [22]. This Chest X-ray14 contains a total of 112,120 CXR images across from 30,805 patients, with 14 common thoracic disease labels plus one Normal (Healthy) label.

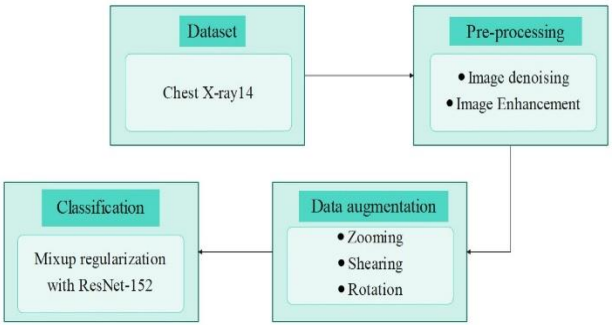


Figure. 1 Block diagram of the implemented method

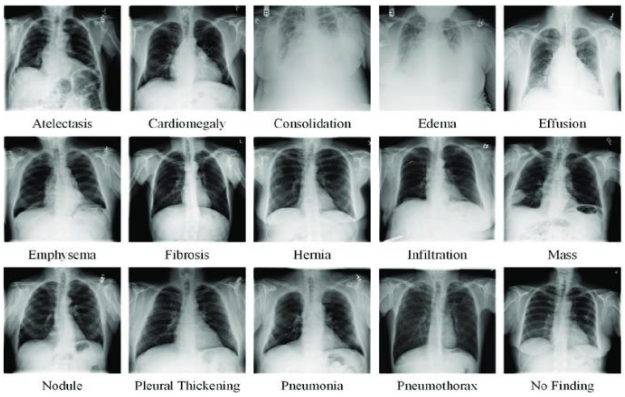


Figure. 2 Sample images of Chest X-ray14 dataset

Table 1. NHI dataset properties

Properties	Data
Total classes number	15
Normal data	60412
Overall data	112120
Total multi-disease	49991
Total Single-Disease	31191
Gender (Female: Male)	43.51%, 56.49%
Splits	Patient-wise official, Image-wise Random Split (utilized here)

Each image in the dataset has a resolution of 1024×1024 pixels. This dataset has a 10% validation set, 70% training set, and 20% test set which are divided randomly at the patient level. The small dataset size in comparison to the computer vision field is a significant disadvantage of any deep learning approach to medical image processing. Therefore, it is frequently impractical to train a CNN from scratch. One solution is transfer-learning, where model was initialized with pre-trained ImageNet weight and then again trained on the ChestX-ray14 target dataset. Table 1 represents the detailed dataset properties, and Fig. 2 illustrates the chest X-ray14 dataset's sample images.

3.2 Pre-processing

Following the data collection, pre-processing is performed in this section. This pre-processing process has two techniques such as image denoising using Gaussian filter and image enhancement using CLAHE, as described below.

3.2.1. Image denoising using Gaussian filter

The Gaussian filter is utilized to filter an image before categorization. The images are rescaled into 256 × 256 to reduce computation time and complexity. Based on the Gaussian function's shape, this approach selects a linear filter with a weighted value for each image component. This approach was selected because it suppresses the visual noise, and thus improves visual structures at various scales. The values of each component in the Gaussian smoothing filter is computed using the Eq. (1),

$$h(x,y) = \frac{1}{c} e^{\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

Where, the normalization constant is denoted as c and the Gaussian kernel's standard deviation is represented as σ.

3.2.2. Image enhancement using CLAHE

The contrast of the image is enhanced or improved by adjusting the range of the distribution of pixel intensity. Contrast enhancement is used to enhance the overall quality of the image in various computer vision tasks. The contrast improvement of the image is carried out using CLAHE (Contrast Limited Adaptive Histogram Equalization) Algorithm, which also avoids noise over enhanced images. The pre-processed images are passed as input to the next stage called the image augmentation process.

3.3 Image augmentation

Following the pre-processing, the image data augmentation technique is used to produce modified versions of the images to artificially improve the dataset size. Data augmentation helps cope with the "not enough data" problem, avoids overfitting and advances the model's ability to generalize better. To produce the point coordinates using the data augmentation technique on the image, the transformation matrix is employed. Zooming, shearing, and rotation methods are adopted in this image augmentation technique.

- **Zooming:** This is applied to develop images with different zooming levels. This method zooms image and includes new pixels for each image randomly which is described in Eq. (2),

$$\text{Zooming Matrix} = \begin{bmatrix} C_x & 0 & 0 \\ 0 & C_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Where, the zoom factor along the y-axis and x-axis are determined by C_y, and C_x.

- **Shearing:** Shifting one image portion by transforming image orientation is utilized in shearing method. The matrix for shearing is given in Eq. (3).

$$\text{Shearing} = \begin{bmatrix} 1 & sh_y & 0 \\ sh_x & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

Where, shear factor along the y-axis and x-axis are demonstrated by sh_y and sh_x.

- **Rotation:** The image rotating along axis 1° and 359° is accomplished by rotation method, either clockwise or counter-clockwise. Data labels are no longer preserved when the rotation degree increases. The matrix for rotation is provided in Eq. (4).

$$\text{Rotation} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

Where, θ is represented as the rotation angle and the rotation angle is selected randomly. Data augmentation technique not only increases the model's performance but also helps to reduce overfitting which is a general issue in DL methods. The augmented images are passed as input to the classification process with Mix-up regularization. It also does augmentation internally and boost the classification performance, which is explained in next section.

3.4 Classification using Mix-up regularization with ResNet-152

Following the initial data augmentation, the classification process is performed by utilizing the Mix-up regularization with the ResNet-152 model. The ResNet-152 is used as a backbone network to extract the features.

3.4.1. ResNet-152

The deep CNN known as ResNET, and an ensemble of these residual nets won the ILSVRC 2015 contest in image classification with a top-5 error rate of 3.57%. The ResNets known as Residual networks where many small individual Residual networks are linked to form the larger network.

The ResNet-152 [22] as a feature extractor and classifier as shown in Fig. 3. ResNet-152 is already

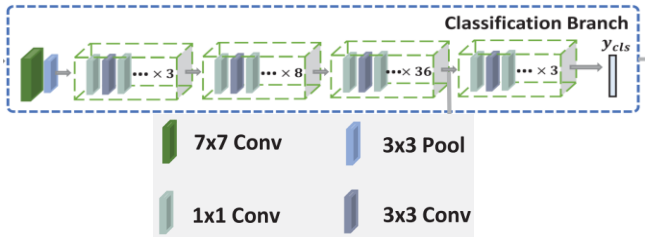


Figure. 3 ResNet-152 networks with residual blocks

trained on the ImageNet dataset which contains all types of images. The ResNet-152 contains 152 layers, having a convolutional layer with the filer size 7 × 7, a max pooling layer with size 3 × 3 and four residual blocks. This residual block is a 3, 8, 36, and 3 triple-layer block. As per our problem definition, we not used softmax layer and replaced the last fully connected layer with 14 neurons, using the sigmoid activation. The predictions of 14 neurons are termed label prediction vector y_{cls}, indicating the possibility of each disease. If all values are zero, the image signifies no result or a healthy image.

3.4.2. Mix-up-based Regularization

The amount of annotated data is limited in every domain, which results in marginal distributions and aligning causing mismatching. To reduce these harmful effects, this research implements a category called Mix-up [9] based on interpolation regularizations on ResNet. The Mix-up is a data augmentation technique that creates a new sample by convex combinations of training points from random pairs of images and associated disease labels.

Vicinal Risk Minimization is the foundational concept of mix-up training [23]. In this case, the classifier is trained both on training data and in the environment surrounding every initial sample. Eqs. (5) and (6) provide an extra straightforward rule that determines how the vicinal points are created. The $\lambda \in [0, 1]$ is called a linear interpolator used to calculate the mixing ratio from symmetric Beta distribution $\text{Beta}(\alpha, \alpha)$ on all training iterations.

$$\tilde{x} = \lambda x_i + (1 - \lambda)x_j \quad (5)$$

$$\tilde{y} = \lambda y_i + (1 - \lambda)y_j \quad (6)$$

Where, (x_i, x_j) and (y_i, y_j) are input points and their associated encoded labels, respectively. Both are affected by empirical Dirac delta distribution as per Eq. (7). Where (x_i, y_i) are at center and they are being replaced by empirical vicinal distribution as per Eq. (8),

$$P_\delta(x, y) = \frac{1}{n}\sum_i^n \delta(x = x_i, y = y_i) \quad (7)$$

$$P_v(\tilde{x}, \tilde{y}) = \frac{1}{n}\sum_i^n v(\tilde{x}, \tilde{y} \mid x_i, y_i) \quad (8)$$

The vicinity distribution v gives the probability value P_v to find virtual feature-target pair $(\tilde{x}, \tilde{y})$ in the vicinity of the original pair(x_i, y_i). Eqs. (5) and (6) express the numerical expression of the method to generate new vicinal samples $(\tilde{x}, \tilde{y})$.

The empirical vicinal risk R_v as per Eq. (9) is used to perform minimization during training with the help of new vicinal data $(\tilde{x_i}, \tilde{y_i})$.

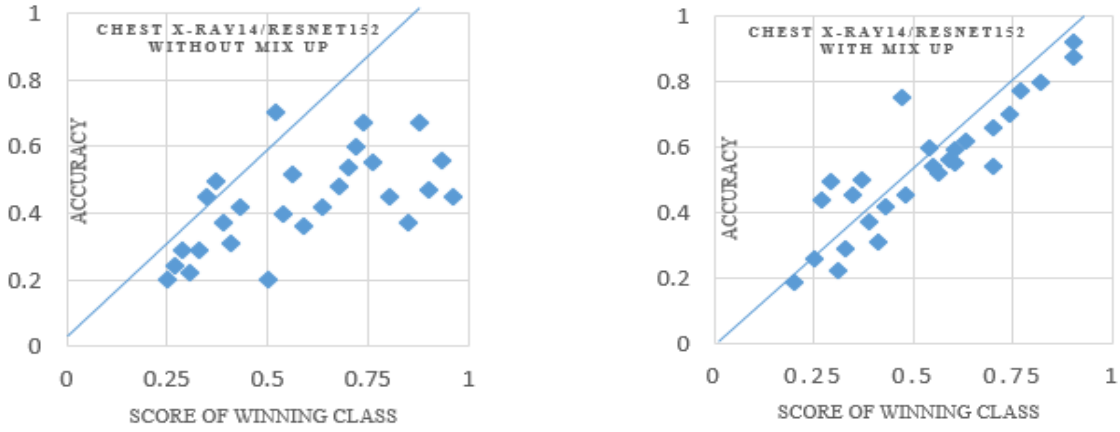
$$R_v(f) = \frac{1}{m}\sum_{i=1}^m L(f(\tilde{x}_i), \tilde{y}_i) \quad (9)$$

Mix-up generates new input/output and minimizes the corresponding empirical vicinal risk, following Mix-up risk R_m as shown in Eq. (10). The R_m is minimized by a typical stochastic gradient descent algorithm, for this gradient is obtained with the help of sampled value of λ at each iteration. A sample mini-batch of training pairs is created and later, Mix-up random pairs within this mini-batch are generated, in turn producing a stochastic gradient for Eq. (10).

$$R_m = \frac{1}{n^2}\sum_{i=1}^n \sum_{j=1}^n L(\lambda y_i + (1 - \lambda)y_j, f(\lambda x_i + (1 - \lambda)x_j)) \quad (10)$$

The L indicates cross-entropy loss on soft-labels $\tilde{y}_i$, rather than hard labels. This training augments the new feature set and induces a new set of soft-labels. The Mix-up makes this model learn from modified new inputs of $\tilde{x}$ and outputs $\tilde{y}$. It encourages classification regions to vary linearly between input samples. Hence, Mix-up implicitly shrinks inputs towards their mean (x=y line), while also improving the classification accuracy as shown in Fig. 4 in this experiment while training with Chest X-ray14 images. In our model, the confidence matches the accuracy most of the points will be near x = y line in right side of Fig. 4. The Accuracy is calculated as the ratio of the correctly predicted instances to the total instances in the dataset.

The interpolation among a pair of images and related smoothed labels are controlled by hyperparameter α from symmetric Beta distribution $\text{Beta}(\alpha, \alpha)$. The $\alpha = 0$ is the base case with zero-entropy training labels. Here, the value of $\alpha > 0$ produces the average of inputs and labels. The selected smaller range values of $\alpha \in [0.1, 0.5]$ in the



(a) (b)

Figure. 4 Scatterplots for accuracy and confidence on Chest X-ray14 image datasets for: (a) Mix-up and and (b) baseline (no Mix-up)

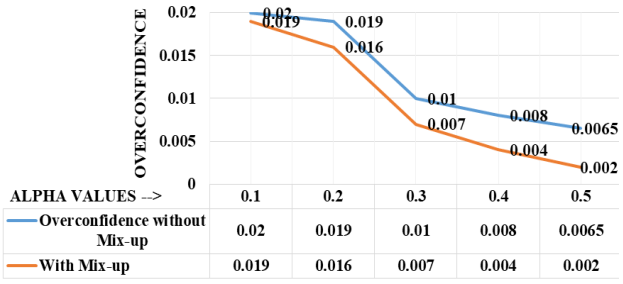


Figure. 5 Variation overconfidence for Mix-up and no Mix-up on Chest X-ray14 image datasets.

classification task as overconfidence decrease monotonically with α. The high values of $\alpha = 0.5$ result in significant under-fitting with the lowest overconfidence of 0.002 in this experiment, as shown in Fig. 5, which is related to manifold intrusion [24]. A high value of $\alpha = 0.5$ without using Mix-up resulted in a significant overconfidence of 0.0065 in Fig. 5. The observed Mix-up regularization technique provides a prediction consistency in between training instances.

4. Result and Discussion

The system requirements for this research method are as follows: i9 processor, 32 GB RAM, and Windows 11, exploited to classify the multiclass categorization of CXR images with a learning rate of 0.1. A network is trained with a weight decay of 5 ×104 using SGD. The learning rate is halved at 2 for 60,120, 180, and 240 epochs.

4.1 Evaluation metrics

The implemented ResNet-152 with Mix-up regularization method's performance is estimated in a Chest X-Ray14 utilizing measures of Area under Curve (AUC), Receiver Operating Curve (ROC), and Accuracy.

Table 2. AUC performance analysis of implemented method based on 14 classes

Diseases	ResNet 50 without Mix-Up	ResNet 50 with Mix-Up	ResNet-152 without Mix-Up	ResNet-152 with Mix-Up
Cardiomegaly	0.783	**0.845**	0.842	**0.992**
Atelectasis	0.741	**0.850**	0.856	**0.995**
Mass	0.789	**0.821**	0.899	**0.992**
Pneumonia	0.835	**0.795**	0.787	**0.984**
Consolidation	0.721	**0.826**	0.822	**0.989**
Effusion	0.824	**0.763**	0.836	**0.991**
Emphysema	0.768	**0.854**	0.892	**0.994**
Edema	0.739	**0.833**	0.865	**0.988**
Fibrosis	0.816	**0.798**	0.913	**0.987**
Infiltration	0.796	**0.742**	0.985	**0.997**
Nodule	0.820	**0.865**	0.919	**0.996**
Pleural Thickening	0.799	**0.777**	0.920	**0.997**
Pneumothorax	0.750	**0.848**	0.867	**0.998**
Hernia	0.855	**0.860**	0.888	**0.989**
Avg AUC	**0.784**	**0.819**	**0.877**	**0.992**

4.2 Quantitative and Qualitative Analysis

Table 2 displays the AUC performance evaluations of the implemented model based on 14 classes. Specifically, these classes include Infiltration, Emphysema, Pneumonia, Mass, Pneumothorax, Pleural Thickening, Hernia, Atelectasis, Edema, Effusion, Pulmonary Fibrosis, Cardiomegaly, Consolidation, and Nodule. The performance of both

Table 3. AUC performance analysis of ResNet-152 with different regularization methods

Diseases	ResNet-152 with Cutout Regularization	ResNet-152 with CutMix Regularization	ResNet-152 with AugMix Regularization	ResNet-152 with Mix-up regularization method
Cardiomegaly	0.721	0.821	0.817	**0.992**
Atelectasis	0.758	0.898	0.828	**0.995**
Mass	0.735	0.865	0.839	**0.992**
Pneumonia	0.798	0.878	0.848	**0.984**
Consolidation	0.740	0.749	0.859	**0.989**
Effusion	0.719	0.815	0.826	**0.991**
Emphysema	0.799	0.726	0.830	**0.994**
Edema	0.738	0.735	0.851	**0.988**
Fibrosis	0.777	0.824	0.862	**0.987**
Infiltration	0.763	0.868	0.884	**0.997**
Nodule	0.789	0.890	0.895	**0.996**
Pleural Thickening	0.723	0.891	0.897	**0.997**
Pneumothorax	0.756	0.873	0.864	**0.998**
Hernia	0.798	0.782	0.831	**0.989**
Avg AUC	**0.750**	**0.829**	**0.852**	**0.992**

ResNet50 and ResNet-152 are evaluated with and without Mix-up regularization method based on these 14 disease classes as shown in Table 2. The obtained outcomes express that the implemented ResNet-152 with Mix-up regularization method achieves better classification results when compared to others. This implemented ResNet-152 with the Mix-up regularization method achieves a high average AUC value of 0.992 (99.2%), as opposed to ResNet-50 with and without Mix-up. One of the reasons is that ResNet-152 is deeper than ResNet-50. As opposed to ResNet-50 with and without Mix-Up, the results show significantly commendable performance with Mix-Up-based ResNet 50, as indicated in Table 2.

Table 3 displays the AUC performance analysis of ResNet-152 with different regularization methods. The performance of ResNet-152 with Cutout Regularization [15], ResNet-152 with CutMix Regularization [14], and ResNet-152 with AugMix Regularization [13] are with implemented ResNet-152 with Mix-up regularization method based on these 14 disease classes. The Cutout is an image augmentation and regularization technique that randomly masks out square regions of input during training. Which can be used to improve the robustness and overall performance of convolutional neural networks**.** Cut-mix uses a patch from a different image to replace the areas that are taken out. Following the total number of pixels in the combined images, the ground truth labels are likewise merged. The AugMix data processing method combines augmentations created at random and apply a Jensen-Shannon loss [13] to ensure continuity. The attained results show that the implemented ResNet-152 with the Mix-up regularization method achieves preferable classification outcomes in relation to other regularization models. Fig. 6 represents the resulting ROC curves attained by implementing ResNet-152 with Mix-up regularization. The major difference between regularization schemes is shown in Table 4.

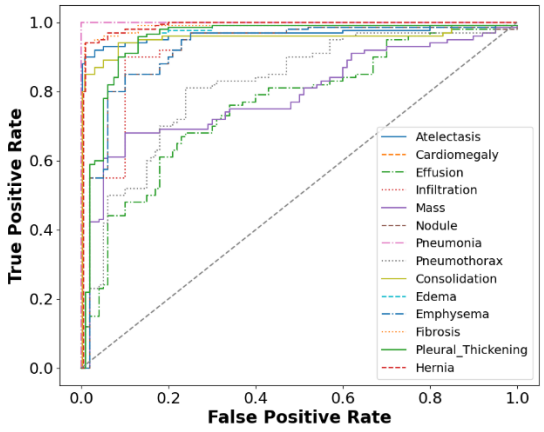


Figure. 6 Implemented ResNet-152 with Mix-up model ROC curves for the Chest X-Ray14 dataset

Table 4. Mix-Up, CutMix, & CutOut

	Mix-Up	Cutout	CutMix
Full image region	✓	✗	✓
Regional dropout	✗	✓	✓
Mixed image & label	✓	✗	✓

4.3 Comparison with Other Existing Models

This section demonstrates the comparative analysis of the ResNet-152 with a Mix-up regularization classifier using Chest X-ray14 dataset as exhibited in Table 5. The implemented ResNet-152 with the Mix-up regularization has the potential to learn more complex features and representations which leads to superior performance results, especially when being trained on large datasets. Similarly, the existing methods such as CNN-ELM [17], PCSANet [19] and GWSA&LCD [20] also used the same Chest X-ray14 dataset with 112,120 CXR images from 30,805 patients with 1 Healthy label and 14 thoracic disease labels. All the images present in this dataset comprises a resolution quality of 1024×1024 pixels. Fig. 7 represents the comparative analysis of Average AUC values' graphical representation. When compared to existing CNN-ELM [17], PCSANet [19] and GWSA&LCD [20], this implemented ResNet-152 with the Mix-up regularization method achieved a high 0.992 average AUC value which is tabulated in table 5.

Table 5. Comparative analysis of AUC values of existing and implemented methods

Diseases	CNN-ELM [17]	PCSANet [19]	GWSA&LCD [20]	Our method
Cardiomegaly	0.9789	0.910	0.877	**0.992**
Atelectasis	0.9711	0.807	0.770	**0.995**
Fibrosis	0.8047	0.812	0.839	**0.987**
Mass	0.9425	0.824	0.821	**0.992**
Pneumonia	0.9123	0.750	0.732	**0.984**
Effusion	0.9742	0.879	0.827	**0.991**
Consolidation	0.7901	0.802	0.746	**0.989**
Emphysema	0.9371	0.890	0.924	**0.994**
Edema	0.8878	0.888	0.847	**0.988**
Infiltration	0.9722	0.698	0.701	**0.997**
Nodule	0.702	0.750	0.790	**0.996**
Pleural Thickening	0.8062	0.768	0.782	**0.997**
Pneumothorax	0.8887	0.850	0.870	**0.998**
Hernia	0.9164	0.915	0.921	**0.989**
Avg AUC	0.891	0.824	0.817	**0.992**

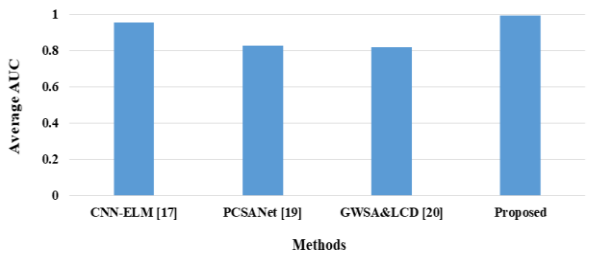


Figure. 7 Comparative analysis of Average AUC values' graphical representation

4.4 Discussion

This study shows the efficacy of Mix-Up training on the Chest X-ray14 dataset using a proposed methodology. Through numerous trials, this research investigates the influence of Mix-Up training on ResNet when applied to Chest X-ray14 data. Throughout the experiments, this method applies Mix-Up to pairs of both images and labels. The research findings elucidate how and when Mix-Up aids DNN models like ResNet-152 in the improved classification of Chest X-ray14 images. Importantly, this research demonstrates that training with Mix-Up is tantamount to learning on modified data, thereby mitigating overfitting in the ResNet model. Hence, Mix-up is proved to be a simple and cheap data augmentation technique. The Mix-Up brought a second set of augmentation after the initial augmentation which boosted the obtained results. The outcomes indicate that the mix-up-based label smoothing is a type of entropic regularization, which prevents DCNNs from committing mistakes to a large extent. Nonetheless, the limitation of Mix-Up is that Mix-up samples induce confusion in the model due to their regional contradiction and inappropriateness, especially when localization is being performed. Mix-up tends to cause a classifier to concentrate on limited regions and reduces the initial model's localization accuracy. Mix-up regularization help to improve the model's capability to accurately classify rare abnormalities or disease in chest X-ray images, also to address class imbalance issue of dataset by generating more synthetic samples for minority pathology classes. We have found $\alpha = 0.5$ in which Mix-Up shows showed better performances with low overconfidence. We know that training with hard labels is one of the contributors for overconfidence in all DCNNs

5. Conclusion

A novel method of ResNet-152 with a Mix-Up regularization method is implemented for multiclass classification in Chest X-ray14 medical images. In the pre-processing stage, Gaussian filter is employed to reduce the high-frequency noise and CLAHE is used for image enhancement to improve the image contrast. The initial data augmentation including zooming, rotation, and shearing is applied to reduce overfitting. The augmented images are passed to ResNet-152 for feature extraction and image classification. The Mix-up based regularization is utilized to train a model by blending image pairs and their labels. The proposed ResNet-152 with Mix-Up regulation shows improved performance as opposed

to ResNet50, with and without Mix-Up. The proposed ResNet-152 model with Mix-Up regularization is also compared with the ResNet-152 model with Cutout Regularization, CutMix Regularization, and AugMix Regularization across these 14 disease classes. The attained results show that implemented ResNet-152 with the Mix-up regularization method accomplishes superior classification outcomes when compared to other regularization models. When compared to the existing DCNN methods on the Chest X-ray14 dataset including CNN-ELM, PCSANet, and GWSA&LCD, this implemented ResNet-152 with Mix-up regularization method accomplishes a high 0.992 (99.2%) average AUC value. So, Mix-Up is proven as one of the effective types of data augmentation by generating additional synthetic samples along with label smoothing. The research experiments demonstrate that the Mix-Up is successful in improving the accuracy of the ResNet-152 model on Chest X-ray14 by reducing its overconfidence up to 0.002 and average AUC of 99.2%. Only the Image-wise Random Split of the Chest X-ray14 dataset is utilized here, in future it will be interesting to see results with a Patient-wise-official split where Chest X-ray of the same patient can appear either in the training set or the testing set but not on both.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper formal analysis, resources, methodology, software, writing—original draft preparation done by 1st author, writing—review and visualization, editing, have been done by 4th and 5th author. The investigation, validation, supervision, is done by Corresponding and 3rd Author and project administration have been done by 2nd author.

Acknowledgments

The authors extend their gratitude to Alva's Institute of Engineering and Technology - Moodbidri and Manipal Institute of Technology- Bangalore, for their provision of laboratory facilities, as well as their moral and ethical support, which enabled the successful completion of this research work.

Notation List

symbol	Description
σ	Standard deviation of gaussian kernel
$h(x, y)$	Gaussian smoothing filter
C_x	Zoom factor with x-axis
C_y	Zoom factor with y-axis
sh_x	Shear factor with x-axis
sh_y	Shear factor with y-axis
θ	Rotation Angle
λ	Linear Interpolator
α	Symmetric beta distribution
(x_i, x_j)	Input points
(y_i, y_j)	Output points
$P_\delta(x, y)$	Empirical Dirac delta distribution
$P_v(\tilde{x}, \tilde{y})$	Empirical vicinal distribution
$(\tilde{x}, \tilde{y})$	Virtual feature-target pair
(x_i, y_i)	Vicinity of the original pair
R_v	Empirical vicinal risk
R_m	Mix-up risk
L	Cross entropy

References

[1] A. AbdElhamid, E. AbdElhalim, M. A. Mohamed, and F. Khalifa, "Multi-Classification of Chest X-Rays for COVID-19 Diagnosis Using Deep Learning Algorithms", *Applied Sciences*, Vol. 12, No. 4, p. 2080, 2022.

[2] Nillmani, P. K. Jain, N. Sharma, M. K. Kalra, K. Viskovic, L. Saba, and J. S. Suri, "Four types of multiclass frameworks for pneumonia classification and its validation in X-ray scans using seven types of deep learning artificial intelligence models", *Diagnostics*, Vol. 12, No. 3, p. 652, 2022.

[3] G. H. Huang, Q. J. Fu, M. Z. Gu, N. H. Lu, K. Y. Liu, and T. B. Chen, "Deep Transfer Learning for The Multilabel Classification of Chest X-Ray Images", *Diagnostics*, Vol. 12, No. 6, p. 1457, 2022.

[4] E. Khan, M. Z. U. Rehman, F. Ahmed, F. A. Alfouzan, N. M. Alzahrani, and J. Ahmad, "Chest X-Ray Classification for The Detection of COVID-19 Using Deep Learning Techniques", *Sensors*, Vol. 22, No. 3, p. 1211, 2022.

[5] A. Hussain, M. Imad, A. Khan, and B. Ullah, "Multi-Class Classification for The Identification of COVID-19 In X-Ray Images Using Customized Efficient Neural Network", *AI and IoT for Sustainable Development in Emerging Countries: Challenges and Opportunities*, Cham: Springer International Publishing, pp. 473-486, 2022.

[6] H. Wu, J. Sun, and Q. You, "Semi-Supervised Learning for Medical Image Classification

Based on Anti-Curriculum Learning", *Mathematics*, Vol. 11, No. 6, p. 1306, 2023.

[7] F. M. J. M. Shamrat, S. Azam, A. Karim, R. Islam, Z. Tasnim, P. Ghosh, and F.D. Boer, "Lungnet22: A Fine-Tuned Model for Multiclass Classification and Prediction of Lung Disease Using X-Ray Images", *Journal of Personalized Medicine*, Vol. 12, No. 5, p. 680, 2022.

[8] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision", In: *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818-2826, 2016.

[9] H. Zhang, M. Cissé, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: Beyond Empirical Risk Minimization", In: *Proc. of 6th International Conference on Learning Representations, ICLR 2018*, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings, 2018.

[10] L. Zhang, Z. Deng, K. Kawaguchi, A. Ghorbani, and J. Zou, "How Does Mixup Help with Robustness and Generalization?", *arXiv preprint arXiv:2010.04819*, 2020.

[11] S. Thulasidasan, G. Chennupati, J. A. Bilmes, T. Bhattacharya, and S. Michalak, "On Mixup Training: Improved Calibration and Predictive Uncertainty for Deep Neural Networks", *Advances in neural information processing systems*, Vol. 32, 2019.

[12] R. Shetty, and P. N. Sarappadi, "Self-Sequential Attention Layer based DenseNet for Thoracic Diseases Detection", *International Journal of Intelligent Engineering & Systems*, Vol. 14, No. 4, pp. 158-165, 2021, doi: 10.22266/ijies2021.0831.15.

[13] N. Hendrycks, E. D. Mu, B. Cubuk, J. Zoph, Gilmer, and B. Lakshminarayanan, "Augmix: A Simple Data Processing Method to Improve Robustness and Uncertainty", *arXiv preprint arXiv:1912.02781*, 2019.

[14] S. Yun, D. Han, S. J. Oh, S. Chun, J. Choe, and Y. Yoo, "Cutmix: Regularization Strategy to Train Strong Classifiers with Localizable Features", In: *Proc. of the IEEE/CVF International Conference on Computer Vision*, pp. 6023-6032, 2019.

[15] T. DeVries, and G. W. Taylor, "Improved Regularization of Convolutional Neural Networks with Cutout", *arXiv preprint arXiv:1708.04552*, 2017.

[16] F. M. J. M. Shamrat, S. Azam, A. Karim, K. Ahmed, F. M. Bui, and F. De Boer, "High-Precision Multiclass Classification of Lung Disease Through Customized Mobilenetv2 from Chest X-Ray Images", *Computers in Biology and Medicine*, Vol. 155, p. 106646, 2023.

[17] M. Nahiduzzaman, M. O. F. Goni, R. Hassan, M. R. Islam, M. K. Syfullah, S. M. Shahriar, M. S. Anower, M. Ahsan, J. Haider, and M. Kowalski, "Parallel CNN-ELM: A Multiclass Classification of Chest X-Ray Images to Identify Seventeen Lung Diseases Including COVID-19", *Expert Systems with Applications*, Vol. 229A, p. 120528, 2023.

[18] M. Mann, R. P. Badoni, H. Soni, M. Al-Shehri, A. C. Kaushik, and D. Q. Wei, "Utilization of deep convolutional neural networks for accurate chest X-ray diagnosis and disease detection", *Interdisciplinary Sciences: Computational Life Sciences*, Vol. 15, No. 3, pp. 374-392, 2023.

[19] K. Chen, X. Wang, and S. Zhang, "Thorax Disease Classification Based On Pyramidal Convolution Shuffle Attention Neural Network", *IEEE Access*, Vol. 10, pp. 85571-85581, 2022.

[20] Y. Xu, H. K. Lam, X. Bao, and Y. Wang, "Learning Group-Wise Spatial Attention and Label Dependencies for Multi-Task Thoracic Disease Classification", *Neurocomputing*, Vol. 573, p. 127228, 2024.

[21] H. Wang, H. Jia, L. Lu, and Y. Xia, "Thorax-Net: An Attention Regularized Deep Neural Network for Classification of Thoracic Diseases On Chest Radiography", *IEEE Journal of Biomedical and Health Informatics*, Vol. 24, No. 2, pp. 475-485, 2020.

[22] 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: *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2097-2106, 2017.

[23] O. Chapelle, J. Weston, L. Bottou, and V. Vapnik, "Vicinal risk minimization", *Advances in neural information processing systems*, pp. 416–422, 2001.

[24] H. Guo, Y. Mao, and R. Zhang, "Mixup as Locally Linear Out-Of-Manifold Regularization", *arXiv preprint arXiv:1809.02499*, 2018.