

Texture-Based DNN for Pneumonia in Thorax X-Rays

Shubhra Prakash*, Ramamurthy B

CHRIST University, Bangalore, India. *Corresponding Author's Email: shubhra.prakash@res.christuniversity.in

Abstract

This paper introduces an innovative methodology for identifying pneumonia in thoracic X-ray images through the application of neural network classifiers. In our experiment, we employed a comprehensive training regimen involving multiple neural network classifiers, each trained on distinct sets of texture features meticulously extracted from thoracic X-ray images. Four different gray-level matrices and a neighboring gray-tone difference matrix (NGTDM) were used to generate these input features, guaranteeing a reliable depiction of the textural properties found in the X-ray pictures. We carried out an extensive evaluation utilizing a number of performance criteria to gauge the trained classifiers' efficacy. Classifying the thoracic X-ray pictures into two groups' pneumonia and healthy state was the assignment assigned to the classifiers. A thorough study of the classifiers' performance was provided by our assessment measures, which comprised accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC). The experimental findings showed that the suggested method accomplished a remarkable 91% overall test categorization accuracy, which was encouraging. This degree of precision highlights how well our approach works to accurately diagnose pneumonia from thoracic X-ray images. Furthermore, the consistent performance across different metrics highlights the robustness and generalizability of the proposed strategy.

Keywords: GLCM, GLDM, GLRLM, GLSZM, NGTDM, Texture Features.

Introduction

Pneumonia involves inflammation in the air sacs of the lungs and can lead to fluid buildup in the lungs, which can make it difficult to breathe. Chest pain, fever, chills, and shortness of breath, coughing, and trouble breathing are all signs of pneumonia. Acute and chronic pneumonia are both possible. Acute pneumonia is usually less severe and is typically caused by bacteria or viruses. Bacteria are typically the cause of chronic pneumonia, which can be more serious. Depending on the reason, antibiotics, antiviral drugs, and other drugs can be used to treat pneumonia. Treatment is usually effective, but in some cases, it can be fatal. People who have underlying medical conditions like asthma, COPD, diabetes, or heart disease are more vulnerable. It is possible to detect pneumonia from chest radiographic images using a computer aided detection (CAD) system. In a CAD system, a computer algorithm is trained on a dataset of chest radiographic images (mostly X-Ray and CT scans) to detect the presence of pneumonia-related abnormalities. After that, the system examines new X-ray images and finds regions that might suggest pneumonia. Image categorization,

the process of labeling medical images according to their content, is typically used in CAD systems. Image classification can be done using texture features by extracting features that represent the texture of an image and using them as inputs to a classification algorithm. Texture features are characteristics of the surface of an object that describe its visual texture. They can be determined by measuring the spatial arrangement of colors, intensities, and shapes in an image. Since X-ray images are grayscale, five Grey Level Matrices and features derived from them are used in the study. Each of these methods determines the texture of the image by examining the distribution of pixels with comparable gray levels. GLCM measures the occurrence of pairs of pixel values at a given offset or distance. GLSZM measures the distribution, size and shape of gray level zones. Count and length of gray level runs are measured by GLRLM. The number of neighboring pixels with differing levels is measured using NGTDM. GLDM determines the reliance between pixels with similar grey levels. The primary contributions of the manuscript are as follows: To classify pneumonia, a Deep neural

This is an Open Access article distributed under the terms of the Creative Commons Attribution CC BY license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

(Received 21st April 2025; Accepted 09th July 2025; Published 30th July 2025)

networks (DNN) classifier with fully connected layers was trained. The suggested method's input images were collected from several X-ray imaging databases, and appropriate pre-processing was carried out. Five different set of features (and a combined set) were extracted from gray level matrices. Six distinct DNN classifiers were trained using the retrieved features, and the integrated gradient approach was utilised to ascribe the features back. The classification performance of the individual sets were compared with the combined set based on various evaluation criterion

The structure of the remaining sections is as follows: The findings of the literature review are presented further in this section. Section "Materials and Methods" outlines the recommended models and techniques for utilizing a DNN to classify pneumonia. Section "Results and Discussion" presents and discusses the findings, accordingly. The manuscript is concluded in the last section.

Several other studies have been carried out by researcher where texture features were used. A novel tumor detection method for mammograms was proposed using a combination of GLCM features and an SVM classifier, incorporating a double threshold segmentation technique to distinguish between background and target regions. After optimization, the GLCM features were fed into the SVM classifier, and the method showed strong performance when tested against a mammogram database (1). The use of GLCM features for analyzing dental panoramic imagery to enable early detection of osteoporosis was explored, demonstrating that GLCM effectively captures texture relevant to underlying medical conditions and can differentiate between healthy and osteoporotic mandibular bones (2). Machine learning algorithms, including XGBoost and Random Forest, were applied to radiological feature analysis to distinguish between normal and COVID-19-affected chest X-ray images, achieving up to 82% accuracy and sensitivity, with XGBoost yielding the best results (3). A model combining LBP and HOG texture features with deep features (FC-VGG) was developed to enhance pneumonia diagnosis in pediatric chest X-ray images, achieving an accuracy of 92.19%, average precision of 93.44%, recall of 92.19%, and F1-score of 92.81% (4). An approach using an

open-source dataset and GLCM stride combinations demonstrated notable results across multiple metrics, highlighting its effectiveness in COVID-19 detection from X-ray images (5). A quadtree-based method was introduced to interpret nonlinear SVM predictions for medical image classification, validated through occlusion techniques, which helped in identifying discriminative image zones and improving interpretability (6). GLCM and GLRLM features were utilized to detect affected fissure regions in CT images, and a neural classifier trained on these features achieved classification accuracies of 86% and 87% (7). For accurate classification of teeth affected by caries, a binary classification method using optimized GLCM features and a DNN classifier was proposed, achieving an accuracy of 99% (8). Texture analysis of lower limb bone X-rays using GLCM and K-means clustering yielded an accuracy of 80%, supporting their combined application for fracture detection and structural analysis (9). The use of higher-order GLCM features for detecting malignant masses in mammogram images was demonstrated, marking a novel application of this method beyond its traditional use in remote sensing (10). A method was presented for detecting osteoarthritis in knee joints from MRI images using image processing and an SVM classifier, achieving an accuracy of 86.6% (11). A CAD system trained with GLCM features and an ANN classifier was designed to distinguish between cancerous and non-cancerous mammograms, achieving a sensitivity of 99.3% and a precision of 99.4% on the MIAS dataset (12). A tuberculosis detection system based on chest X-ray images was developed using optimized GLCM features and SVM classification, resulting in class-wise accuracies of 100% for normal, and 98.72% for both primary and secondary TB classes (13). Another study employed GLCM features and an SVM classifier to distinguish between benign and malignant mammography images, reporting 63.03% accuracy and 89% specificity (14). GLCM features and K-means clustering were used to classify contrast-enhanced dental radiographs, with performance evaluated based on accuracy metrics (15).

Methodology

In this manuscript, a framework for pneumonia classification using DNN classifier is proposed. The images used in this framework are gathered from Guangzhou Women's Children's Hospital. The input images are preprocessed by the CLAHE method to enhance the contrast-value and clarity of the X-ray images. The texture features are then extracted using Gray Level Matrices. In order to classify chest images, the collected features are then given into a DNN classifier. Figure 1 illustrates the proposed framework and overall workflow. Figure 1 shows the image classification pipeline which includes input image pre-

processing, extracting texture features, training Deep Neural Network model, and applying interpretable AI techniques to better understand the classification process. Using individual matrices a $1 \times N$ vector was extracted for each image as well as a separate DNN classifier was trained on each. At last, for TexPneum Classifier all the features were concatenated and a separate DNN classifier had been trained for same. The architecture of the classifier can be seen in Figure 2, for the entire instance only the input layer was different (depending on the number of features) rest of the architecture remained the same.

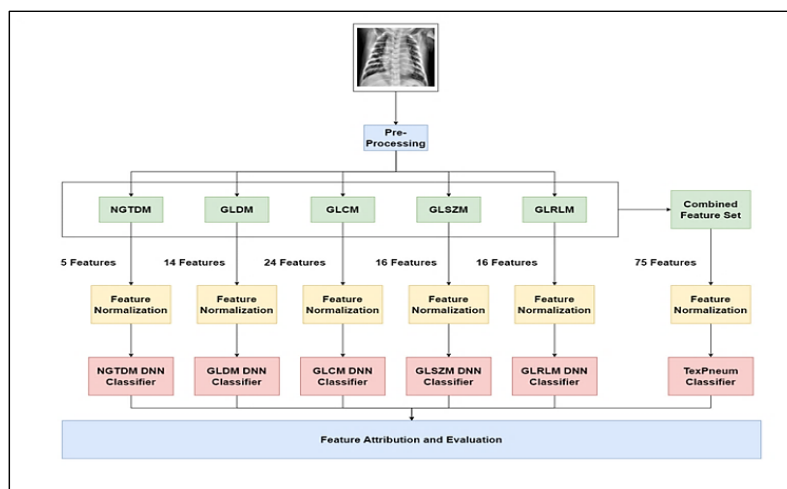


Figure 1: Flowchart of the Implemented Work

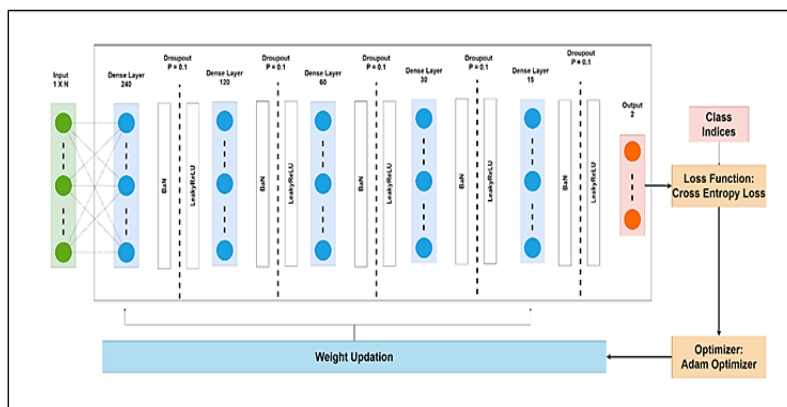


Figure 2: Architecture for the deep neural network

Dataset

A dataset of X-ray images from "Guangzhou Women's Children's Hospital" in Guangzhou, China, divided into two categories normal and pneumonia was used in this investigation (16). The labels were provided by expert radiologists as part of the original dataset. There were 5215 total images in the dataset, of which 25% were normal images and 75% were pneumonia images. To

create train, validation, and test, the dataset was stratified and split in a ratio of 70:15:15 for each fold. For kids between the ages of one and five, the images were collected as part of routine checkups. The images utilised in our study were in PNG format and varied in size, but most of them were 255x255 pixels with a bit depth of 24. Figure 3 shown below shows the label distribution for dataset.

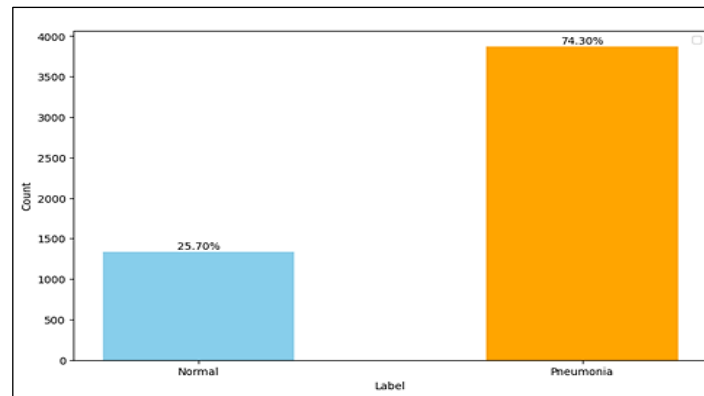


Figure 3: Label Distributions for Dataset

In our study, we encountered an imbalanced dataset, consisting of roughly 25% of samples representing the normal class (minority negative class) and 75% representing the pneumonia class (majority positive class). Despite this, we intentionally chose not to employ specific balancing techniques for several compelling reasons, all of which align with the overarching objectives and priorities of our research. First and foremost, our primary focus in this study was accurate pneumonia detection. Given the critical medical context and the potential consequences, our priority was to minimize false negatives. In this context, a false negative signifies the failure to identify a patient with pneumonia, which could lead to delayed treatment and adverse health outcomes. By abstaining from balancing the dataset, our strategy aimed to reduce the probability of classifying any pneumonia image as normal, thus minimizing false negatives. In many real-world healthcare settings, class distribution tends to be imbalanced, and it becomes paramount to ensure high sensitivity for detecting critical conditions, even if this necessitate accepting a higher false positive (FP) rate for the minority class (normal cases). Furthermore, our dataset was collected from the radiology department and is a reflection of real-world scenarios where the prevalence of specific medical conditions, such as pneumonia, can be substantially higher than others. By deliberately refraining from artificially balancing the dataset, our approach aimed to accurately mirror this real-world distribution, thereby rendering our results more applicable to practical clinical settings.

Preprocessing

Resizing

Resizing images to the same size is important for a neural network to properly process them. We set out to resize every image in our dataset to a standardised 224x224 pixel size in order to ensure uniformity and consistency. There were several reasons for this standardisation in the context of our study. Above all, it helped us create a standard that made it possible to compare our dataset directly. This comparability is a crucial prerequisite for the effective functioning of our feature extractors. Furthermore, our decision to adopt the 224x224 dimension aligns our study with research conducted by the authors and the broader research community. This alignment facilitates the comparison of our findings with existing literature, fostering a more comprehensive understanding of the field. Additionally, the use of smaller image sizes offers practical advantages by reducing the computational resources required.

CLAHE

CLAHE “(Contrast Limited Adaptive Histogram Equalization)” is an algorithm employed to improve the contrast of images. It works by redistributing the light and dark pixels in an image while also limiting the contrast increase of any given area to avoid over-saturation. This algorithm is important for x-ray images classification because it helps to improve the visibility of features that may otherwise be obscured or not easily seen. It can help to identify objects, or detect abnormalities that may otherwise be difficult to detect. Additionally, any image-based classification algorithm can benefit from CLAHE’s assistance in reducing image noise,

which can have a big impact on how well it performs.

Texture Features Extraction

In our study, we use the whole X-ray image to create Texture based features and we do not select a region of interest (ROI). The decision to use the whole image was made for practical purposes and based on the results of our experiments. Initially, we attempted to perform lung segmentation using Unet-based techniques to isolate regions of interest, particularly lung regions. However, during our experiments, we encountered difficulties in correctly segmenting the lungs, especially in the case of pneumonia. The presence of opacities and varying degrees of lung involvement in pneumonia made the segmentation technique unreliable. We deliberately chose to use the full X-ray image for the extraction of textural features in light of these segmentation difficulties and the possibility of introducing segmentation errors that could degrade the quality of our analysis. Below is a discussion of the features that were taken from the images.

Gray Level Co-occurrence Matrix (GLCM) Features

One kind of statistical technique is called GLCM, and it is used to extract texture information from digital images. It is based on estimating the probability that two pixels of a specific gray level would appear together. This 2-D matrix shows the probability that a pair of pixels with a particular gray level will occur for each element. The GLCM can be used to measure an image's texture or pixel layout. It is used on X-ray pictures to measure the texture of a certain area. By counting the occurrences of each pair of Gray levels in a specific region of the image, the GLCM is determined. It then divides the total number of pixel pairs in the region by each matrix element. Using the resulting matrix, other texture features are then calculated, such as entropy, contrast, correlation, as well as angular second moment. It is therefore possible to use these qualities to distinguish between different textures in an image.

Gray Level Size Zone Matrix (GLSZM) Features

The features used in this method are based on a matrix that shows the number of regions in an

image that are a specific size and are gray tones. The number of voxels (volumetric pixels) in a picture with a specific gray level and size is counted to create the matrix. The resulting matrix is then used to calculate other statistics, such as the size zone non-uniformity (SZN), the Gray level non-uniformity (GLN), and coarseness. These statistics can be used to evaluate the texture of the tissue and to distinguish between different tissue types.

Gray Level Run Length Matrix (GLRLM) Features

It is a statistical technique used in image analysis to look at the grayscale intensity and spatial distribution of pixel values. It is a matrix of counts of the quantity of runs (consecutive pixels) of a specific gray level in a particular direction of an image. GLRLM is particularly helpful for spotting subtle changes in an image and can be used to quantify texture in medical and other images.

Neighboring Gray Tone Difference Matrix (NGTDM) Features

The differences in the Gray tones of neighbouring pixels in an image are stored using NGTDM. The texture of a picture is quantified by analysing the geographical changes in Grayscale values. An image's homogeneity, complexity, and regularity can all be determined using the NGTDM.

Gray Level Dependence Matrix (GLDM) Features

The relationship between two or more Gray levels in an image is measured using GLDM. It measures the degree of similarity between adjacent pixels. In contrast to GLCM, the GLDM takes into account the pixel's location within the image as well as its Gray level. The linear correlation of the Gray levels of neighbouring pixels is measured by GLDM, whereas the second-order statistical features of a picture are measured by GLCM. Figure 4 shows the comparison of randomly sampled normal and pneumonia cases based on radiomics features. In Table 1 we have shown the features extracted from all the matrices with the formula used to derive these features. These extracted features will act as the input for our classifier models.

Table 1: List of Features and Respective Formulas

SNo.	Matrix	Features	Formula	Eq
	GLCM			
1		AutocorrelationValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)ij$	[1]
2		JointAverageValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)i$	[2]
3		ClusterProminenceValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^4 p(i, j)$	[3]
4		ClusterShadeValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^3 p(i, j)$	[4]
5		ClusterTendencyValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^2 p(i, j)$	[5]
6		ContrastValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 p(i, j)$	[6]
7		CorrelationValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)ij - \mu_x \mu_y}{\sigma_x(i) \sigma_y(j)}$	[7]
8		DifferenceAverageValue	$\sum_{k=0}^{N_g-1} k p_{x-y}(k)$	[8]
9		DifferenceEntropyValue	$\sum_{k=0}^{N_g-1} p_{x-y}(k) \log_2 (p_{x-y}(k) + \epsilon)$	[9]
10		DifferenceVarianceValue	$\sum_{k=0}^{N_g-1} (k - DA)^2 p_{x-y}(k)$	[10]
11		JointEnergyValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p(i, j))^2$	[11]
12		JointEntropyValue	$-\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log_2 (p(i, j) + \epsilon)$	[12]
13		IMCValue	$\frac{HXY - HXY1}{\max\{HX, HY\}}$	[13]
14		IMC2Value	$\sqrt{1 - e^{-2(HXY2 - HXY)}}$	[14]
15		IDMValue	$\sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + k^2}$	[15]
			$\sqrt{\text{second largest eigenvalue of } QQ(i, j)}$	
16		Maximal Correlation Coefficient (MCC)Value	$= \sum_{k=0}^{N_g} \frac{p(i, k)p(j, k)}{p_x(i)p_y(k)}$	[16]
17		IDMNValue	$\sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + (\frac{k^2}{N_g^2})}$	[17]
18		InverseDifference(ID)Value	$\sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + k}$	[18]
19		InverseDifferenceNormalized (IDN)Value	$\sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + (\frac{k}{N_g})}$	[19]

20		InverseVarianceValue	$\sum_{k=1}^{N_g-1} \frac{p_{x-y}(k)}{k^2}$	[20]
21		MaximumProbabilityValue	$\max(p(i, j))$	[21]
22		SumAverageValue	$\sum_{k=2}^{2N_g} p_{x+y}(k)k$	[22]
23		SumEntropyValue	$\sum_{k=2}^{2N_g} p_{x+y}(k) \log_2 (p_{x+y}(k) + \epsilon)$	[23]
24	GLSZM	SumSquaresValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 p(i, j)$	[24]
25		SmallAreaEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i, j)}{j^2}}{N_z}$	[25]
26		LargeAreaEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i, j)j^2}{N_z}$	[26]
27		GrayLevelNon-UniformityValue	$\frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_s} P(i, j))^2}{N_z^2}$	[27]
28		GrayLevelNon-UniformityNormalizedValue	$\frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_s} P(i, j))^2}{N_z^2}$	[28]
29		Size-ZoneNon-UniformityValue	$\frac{\sum_{j=1}^{N_s} (\sum_{i=1}^{N_g} P(i, j))^2}{N_z^2}$	[29]
30		Size-ZoneNon-UniformityNormalizedValue	$\frac{\sum_{j=1}^{N_s} (\sum_{i=1}^{N_g} P(i, j))^2}{N_z^2}$	[30]
31		ZonePercentageValue	$ZP = \frac{N_z}{N_p}$	[31]
32		GrayLevelVarianceValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i, j)(i - \mu)^2$	[32]
33		ZoneVarianceValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i, j)(j - \mu)^2$	[33]
34		ZoneEntropyValue	$-\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i, j) \log_2 (p(i, j) + \epsilon)$	[34]
35		LowGrayLevelZoneEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i, j)}{i^2}}{N_z}$	[35]
36		HighGrayLevelZoneEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i, j)i^2}{N_z}$	[36]
37		SmallAreaLowGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i, j)}{i^2 j^2}}{N_z}$	[37]
38		SmallAreaHighGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i, j)i^2}{j^2}}{N_z}$	[38]
39		LargeAreaHighGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i, j)j^2}{i^2}}{N_z}$	[39]
40	GLRLM	LargeAreaLowGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i, j)i^2 j^2}{N_z}$	[40]
41		ShortRunEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i, j \theta)}{j^2}}{N_r(\theta)}$	[41]
42		LongRunEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i, j \theta)j^2}{N_r(\theta)}$	[42]

43	GrayLevelNon-UniformityValue	$\frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_r} P(i,j \theta))^2}{N_r(\theta)^2}$	[43]
44	GrayLevelNon-UniformityNormalizedValue	$\frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_r} P(i,j \theta))^2}{N_r(\theta)^2}$	[44]
45	RunLengthNon-UniformityValue	$\frac{\sum_{j=1}^{N_r} (\sum_{i=1}^{N_g} P(i,j \theta))^2}{N_r(\theta)}$	[45]
46	RunLengthNon-UniformityNormalizedValue	$\frac{\sum_{j=1}^{N_r} (\sum_{i=1}^{N_g} P(i,j \theta))^2}{N_r(\theta)^2}$	[46]
47	RunPercentageValue	$\frac{N_r(\theta)}{N_p}$	[47]
48	GrayLevelVarianceValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j \theta)(i-\mu)^2$	[48]
49	RunVarianceValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j \theta)(j-\mu)^2$	[49]
50	RunEntropyValue	$-\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j \theta) \log_2 (p(i,j \theta) + \epsilon)$	[50]
51	LowGrayLevelRunEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j \theta)}{i^2}}{N_r(\theta)}$	[51]
52	HighGrayLevelRunEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j \theta) i^2}{N_r(\theta)}$	[52]
53	ShortRunLowGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j \theta)}{i^2 j^2}}{N_r(\theta)}$	[53]
54	ShortRunHighGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j \theta) i^2}{j^2}}{N_r(\theta)}$	[54]
55	LongRunLowGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j \theta) j^2}{i^2}}{N_r(\theta)}$	[55]
56	LongRunHighGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j \theta) i^2 j^2}{N_r(\theta)}$	[56]
57	CoarsenessValue	$\frac{1}{\sum_{i=1}^{N_g} p_i s_i}$	[57]
58	ContrastValue	$\left(\frac{1}{N_{g,p}(N_{g,p}-1)} \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_i p_j (i-j)^2 \right) \left(\frac{1}{N_{v,p}} \sum_{i=1}^{N_g} s_i \right), \text{ where } p_i \neq 0, p_j \neq 0$	[58]
59	BusynessValue	$\frac{\sum_{i=1}^{N_g} p_i s_i}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} ip_i - jp_j }, \text{ where } p_i \neq 0, p_j \neq 0$	[59]
60	ComplexityValue	$\frac{1}{N_{v,p}} \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i-j \frac{p_i s_i + p_j s_j}{p_i + p_j}, \text{ where } p_i \neq 0, p_j \neq 0$	[60]
61	StrengthValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p_i + p_j)(i-j)^2}{\sum_{i=1}^{N_g} s_i}, \text{ where } p_i \neq 0, p_j \neq 0$	[61]
62	SmallDependenceEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{P(i,j)}{i^2}}{N_z}$	[62]

63	LargeDependenceEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} P(i,j)j^2}{N_z}$	[63]
64	GrayLevelNon-UniformityValue	$\frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_d} P(i,j))^2}{N_z}$	[64]
65	DependenceNon-UniformityValue	$\frac{\sum_{j=1}^{N_d} (\sum_{i=1}^{N_g} P(i,j))^2}{N_z}$	[65]
66	DependenceNon-UniformityNormalizedValue	$\frac{\sum_{j=1}^{N_d} (\sum_{i=1}^{N_g} P(i,j))^2}{N_z^2}$	[66]
67	GrayLevelVarianceValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i,j)(i - \mu)^2, \text{ where } \mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} ip(i,j)$	[67]
68	DependenceVarianceValue	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i,j)(j - \mu)^2, \text{ where } \mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} jp(i,j)$	[68]
69	DependenceEntropyValue	$-\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i,j) \log_2(p(i,j) + \epsilon)$	[69]
70	LowGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{P(i,j)}{i^2}}{N_z}$	[70]
71	HighGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} P(i,j)i^2}{N_z}$	[71]
72	SmallDependenceLowGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{P(i,j)}{i^2 j^2}}{N_z}$	[72]
73	LargeDependenceLowGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{P(i,j)j^2}{i^2}}{N_z}$	[73]
74	LargeDependenceHighGrayLevelEmphasisValue	$\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} P(i,j)i^2 j^2}{N_z}$	[74]

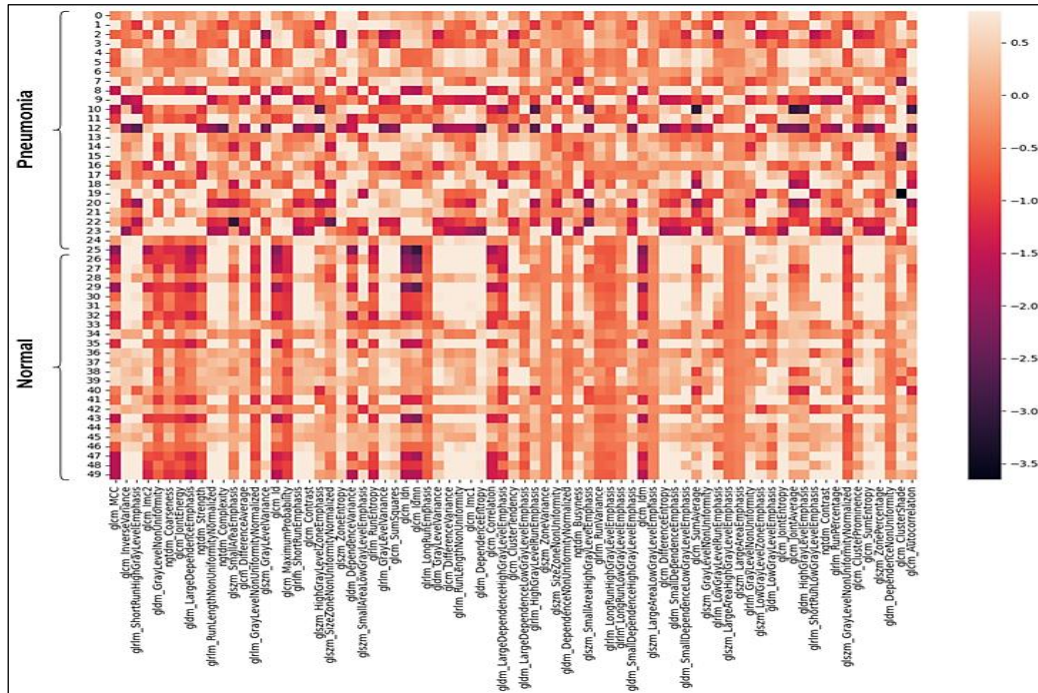


Figure 4: Comparison of Randomly Sampled Normal and Pneumonia Cases based on Radiomics Features

Deep Neural Network Classifier

In our study, we employed a multilayer perceptron, which is a type of “artificial neural network”, as our classifying model. The term “deep” in the context of neural networks generally implies architecture with multiple layers designed to capture intricate patterns and hierarchical representations from the input data. Despite varying interpretations, we deliberately opted for a five-layer network structure. While this choice may appear modest when compared to the complex multi-level deep learning models, it was made based on several considerations. These considerations included the level of

representation training required, the kind of the input data we were dealing with, and the complexity of the task we were doing. It's worth noting that while our model doesn't reach the depth of some other models, it is significantly more intricate than a shallow neural network with just one or two layers. It's important to stress that the definition of “deep” in machine learning is subjective and might vary according to the application and area. In our investigation, the five-layer structure we selected worked well for identifying chest X-ray pictures using attributes that were extracted.

Table 2: Model Summary

	Input Shape	Output Shape	Param
Linear-1	[1,N]	[1,240]	(N×240)+240
Linear-2	[1, 240]	[1,120]	28,920
Linear-3	[1, 120]	[1,60]	7,260
Linear-4	[1, 60]	[1,30]	1,830
Linear-5	[1, 30]	[1,15]	465
Linear-6	[1,15]	[1,2]	32

The structure of our neural network can be better understood by considering that it contains input layer, five hidden layers, and output layer. N in Table 2 indicates that the first layer corresponds to size of the features in the training data. In the GLCM-DNN Classifier, for instance, we extracted 24 features from the GLCM matrix (see Table 1 for a list of features). As a result, the DNN model will have 24 neurons in the input layer; similarly, the number will change based on how many features are utilised in the experiment. The rest of the architecture remains the same. The last layer consists of two output neurons, each representing the pneumonic class or normal class. For hidden layers, as you move from the first to the last layer, the number of neurons starts from 240 (which is three times the max features used) and the density of neurons progressively decreases by half in each hidden layer, leading to 120, 60, 30 in subsequent layers culminating in 15 neurons in the final layer (excluding output layer). The total trainable parameters for each layer are dependent on number of input and output neurons. In terms of activation functions, we employed the rectified linear unit (ReLU) and Leaky ReLU throughout this neural network.

The formula for Leaky ReLU is:

$$f(x) = \{ x, \text{if } x > 0 \text{ and } ax, \text{if } x < 0 \} [75]$$

Where a is a small positive value.

Because of its simplicity and efficiency, this function is often used in multilayer perceptrons. To avoid overfitting, the network also uses dropout technique that removes certain neurons from the network during training. This increases the generalization performance and makes it more likely that the model can be applied to new data. The Loss criterion and the optimizer are the network's two additional crucial components. The Cross Entropy Loss and Adam optimizer are both employed while training this classifier. The optimizer is in charge of changing the neural network's parameters to reduce loss, while the criterion is in charge of calculating the neural network's loss or error during training. Cross Entropy Loss is a gauge of how well the model can categorize the data. Because it allows you to measure the probability of each class and calculate the loss based on the difference between the true and projected class probabilities, cross-entropy loss is a common option when training classification models. The following is the formula for cross entropy:

$$\text{Loss} = - \sum (y * \log(p) + (1-y) * \log(1-p)) [76]$$

Where y is ground truth and p is probability of positive class.

To reduce the loss, Adam adjusts the network parameters by applying gradient descent. The neural network here trains with a learning rate (lr) of 0.01. A neural network learns faster with higher learning rate and slower with lower learning rate. If the learning is too high, the best solution will be skipped. If it is too small, it will take too long to reach the optimal value.

Performance Metrics

Accuracy: It is a metric for determining a classification model's effectiveness. The percentage of all cases in the dataset that are correctly classified is described. It is calculated by taking the total number of instances and dividing it by the number of correctly classified cases.

Precision: It indicates the percentage of cases that were successfully classified out of all the positive predictions made by the model. It can be computed by dividing the total number of cases the model has projected to be true by the number of accurately classified true cases.

Recall: Out of all the occurrences in the dataset that are truly positive, it indicates the percentage of correctly identified instances. It can be computed as the sum of true positives (TP) and false negatives divided by the number of TPs.

F1 Score: It is a weighted average of recall and precision. It is computed by taking the accuracy and efficiency harmonic means. When your data set includes a different class, the F1 score is helpful because it accounts for both FP and false negatives.

Explainable Features

In the paper Integrated Gradients have been used

to find the feature importance for each prediction. Integrated Gradients is an explainable AI (XAI) algorithm developed by Microsoft Research, which provides an interpretable view of a model by attributing each input feature's contribution to the model's output. It serves as an explanation for a deep learning model's forecast. A popular attribution technique that calculates the gradients of the output w.r.t the inputs of a neural network is extended by Integrated Gradients.

Results and Discussion

Five set of features based on GLCM, GLSZM, GLRLM, NGTDM and GLDM were used from all the radiographs. Based on the matrix used the number of features in each set varied accordingly. Multiple DNN classifiers were trained individually on these feature sets and the performance was recorded. At the end all the features were merged and a separate DNN was re-trained to check if there was any improvement. The architecture of the DNN remained fixed throughout for each training, only the number of input features changed.

GLCM-DNN Classifier

The first classifier GLCM-DNN, was trained using 24 features (1 to 24 in feature list Table 1) extracted from the GLCM. The input layer for the classifier consisted of 24 neurons, while keeping the number of neurons intact in rest of the layers as shown in Figure 2. To create train, validation, and test sets, the dataset was stratified and split in a ratio of 70:15:15 for each fold. Shown in Table 3 is the average of 5-fold training and testing.

Table 3: Results for DNN Trained on GLCM Features

GLCM: 24 Features	Accuracy	Precision	Recall	F1Score
Training	0.897	0.931	0.929	0.930
Testing	0.880	0.925	0.913	0.919

The feature importance, as well as internal attribution to comprehend the network function was done using the integrated gradient method. Shown in Figure 5 are the top five features that had the highest magnitude for mean attribution score.

GLSZM-DNN Classifier

The second classifier GLSZM DNN, was trained using 16 features (25 to 40 in feature list Table 1) extracted from the GLSZM. The input layer for the classifier consisted of 16 neurons, while keeping the number of neuron intact in rest of the layers

as shown in Figure 2. Table 4 is the average of 5 fold training and testing.

Shown in Figure 6 are the top five features that had the highest magnitude for mean attribution score.

GLRLM-DNN Classifier

The third classifier GLRLM -DNN, was trained using 16 features (41 to 56 in feature list Table 1) extracted from the GLRLM. The input layer for the classifier consisted of 16 neurons, while keeping the number of neuron intact in rest of the layers

as shown in Figure 2. Shown in Table 5 is the average of 5 fold training and testing.

Table 4: Results for DNN trained on GLCM Features

GLSZM: 16 Features	Accuracy	Precision	Recall	F1Score
Training	0.906	0.953	0.918	0.935
Testing	0.865	0.933	0.881	0.906

Table 5: Results for DNN trained on GLRLM Features

GLRLM: 16 Features	Accuracy	Precision	Recall	F1Score
Training	0.884	0.921	0.923	0.922
Testing	0.857	0.900	0.908	0.904

Shown in Figure 7 are the top five features that had the highest magnitude for mean attribution score.

NGTDM-DNN Classifier

The fourth classifier NGTDM –DNN, was trained using 5 features (57 to 61 in feature list Table 1)

extracted from the NGTDM. The input layer for the classifier consisted of 5 neurons, while keeping the number of neurons intact in rest of the layers as shown in Figure 2. Shown in Table 6 is the average of 5-fold training and testing.

Table 6: Results for DNN trained on NGTDM Features

NGTDM: 5 Features	Accuracy	Precision	Recall	F1Score
Training	0.743	0.743	1.000	0.852
Testing	0.743	0.743	1.000	0.853

Table 7: Results for DNN trained on GLDM Features

GLDM: 14 Features	Accuracy	Precision	Recall	F1Score
Training	0.908	0.952	0.922	0.937
Testing	0.892	0.934	0.919	0.926

Shown in Figure 8 are the top five features that had the highest magnitude for mean attribution score.

GLDM-DNN Classifier

The fifth classifier GLDM –DNN, was trained using 14 features (62 to 74 in feature list Table 1) extracted from the GLDM. The input layer for the classifier consisted of 14 neurons, while keeping the number of neuron intact in rest of the layers as shown in Figure 2. Shown in Table 7 is the average of 5 fold training and testing. Shown in Figure 9 are the top five features that had the highest magnitude for mean attribution score. Figures 5 to 10 present the feature importance visualizations generated using Integrated Gradients for each of the trained DNN classifiers. Figures 5 through 9 correspond to models trained individually on GLCM, GLSZM, GLRLM, NGTDM, and GLDM features, respectively, while Figure 10 illustrates the feature attributions for the combined TexPneum classifier. Each figure

highlights the top five features contributing most significantly to the model's predictions. For instance, the GLSZM-based classifier (Figure 6) emphasizes features like Small Area Emphasis and Zone Entropy, which reflect the presence of small, dense, high-intensity regions—typical markers of pneumonia. Similarly, the GLDM classifier (Figure 9) identifies Dependence Variance and Gray Level Variance as dominant features, indicating structural irregularities in lung textures. The TexPneum classifier (Figure 10), which integrates all feature sets, shows that the most influential predictors are drawn from GLSZM, GLDM, and GLRLM matrices, confirming that combining diverse texture representations enhances performance and interpretability. These visualizations provide a transparent understanding of the decision-making process and support the clinical trustworthiness of the proposed model.

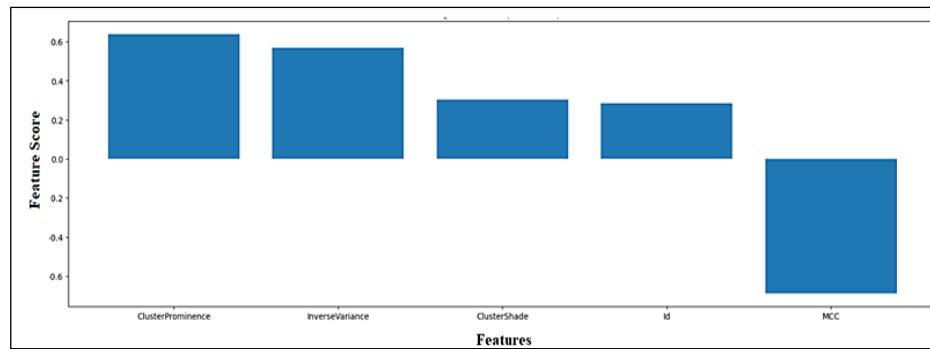


Figure 5: Top Five Features Based on Integrated Gradients for DNN trained on GLCM Features

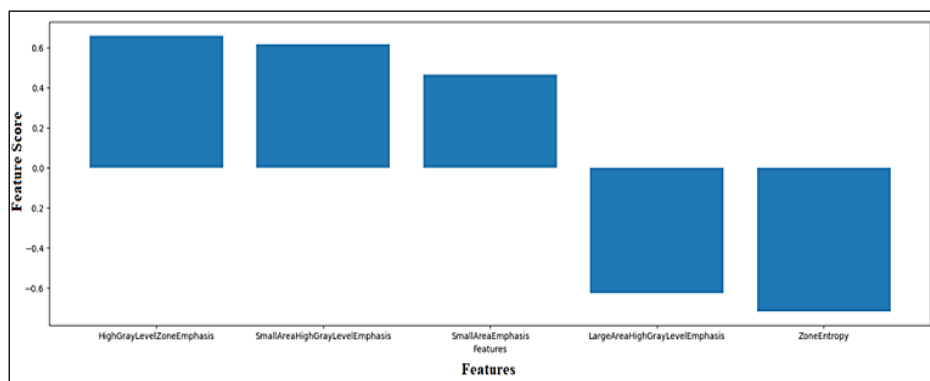


Figure 6: Top Five Features Based on Integrated Gradients for DNN trained on GLSZM Features

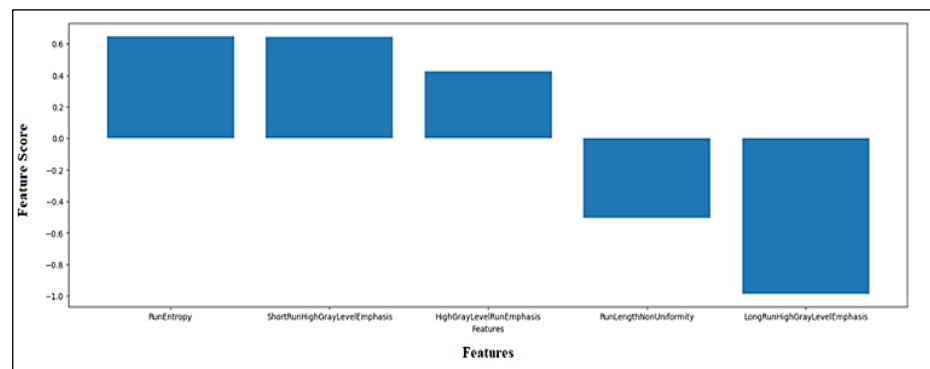


Figure 7: Top Five Features Based on Integrated Gradients for DNN trained on GLRLM Features

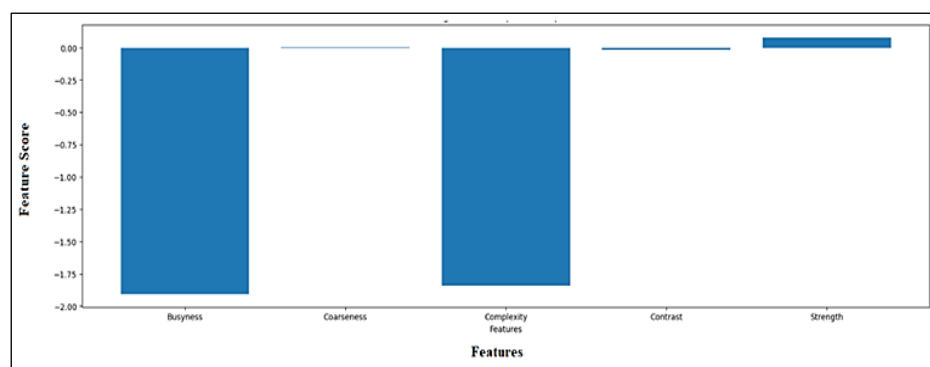


Figure 8: Top Five Features Based on Integrated Gradients for DNN trained on NGTDM Features

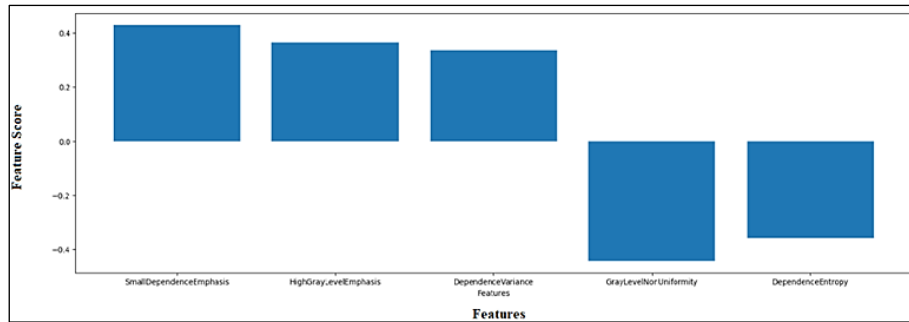


Figure 9: Top Five Features Based on Integrated Gradients for DNN trained on GLDM Features

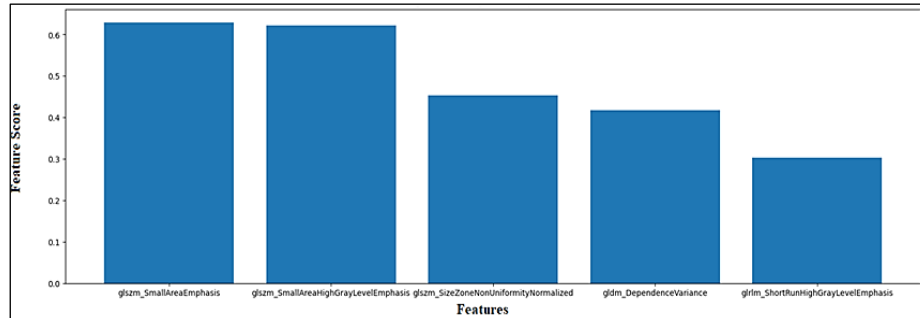


Figure 10: Top Five Features based on Integrated Gradients for TexPneum Classifier

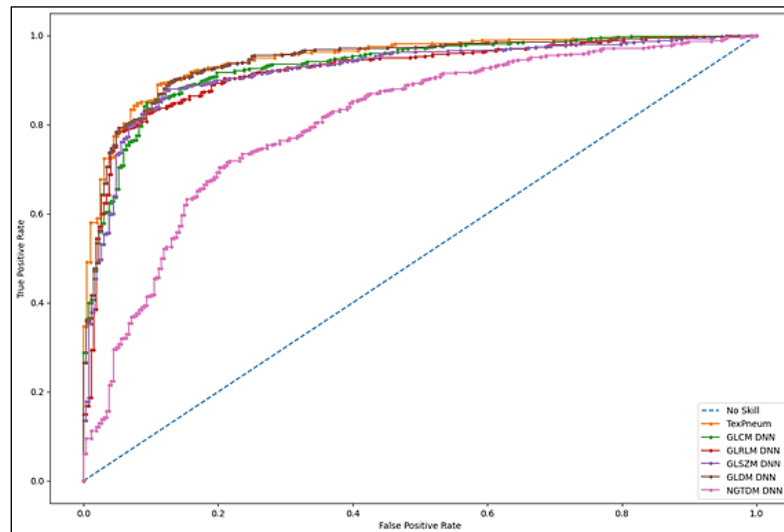


Figure 11: ROC Curve for all the Classifiers

Table 8: Comparison of Results for DNN Trained on Different Features and TexPneum Classifier

Feature Matrix	No. of Features	Training			F1 Score	Testing			F1 Score
		Accuracy	Precision	Recall		Accuracy	Precision	Recall	
GLCM-DNN Classifier	24	0.897	0.931	0.929	0.930	0.880	0.925	0.913	0.919
GLSZM-DNN Classifier	16	0.906	0.953	0.918	0.935	0.865	0.933	0.881	0.906
GLRLM-DNN Classifier	16	0.884	0.921	0.923	0.922	0.857	0.900	0.908	0.904
NGTDM-DNN Classifier	5	0.743	0.743	1.000	0.852	0.743	0.743	1.000	0.853
GLDM-DNN	14	0.908	0.952	0.922	0.937	0.892	0.934	0.919	0.926

Classifier

TexPneum

Classifier	75	0.946	0.970	0.958	0.964	0.907	0.937	0.938	0.937
------------	----	-------	-------	-------	-------	-------	-------	-------	-------

TexPneum Classifier

This classifier, TexPneum, was trained using all the features combined from all the matrices. The input layer for the classifier consisted of 75 neurons, while keeping the number of neurons intact in rest of the layers as shown in Figure 2. Shown in Table 8 is the comparison with other classifiers.

TexPneum Classifier was the best model in comparison. The GLDM and GLSZM classifiers also showed good results with test accuracy of 0.892 and 0.865. The combined feature set only provides a slight improvement in the test accuracy, and it can be seen in the combined feature set importance chart below that out of 5 top features, 3 are from GLSZM matrix and 2 from GLDM and GLRLM matrices. The ROC curves for all the classifiers are shown in Figure 11.

As seen in the above sections, the texture based features are able to distinguish between Normal and Pneumonia images. In Figure 10, we can see that TexPneum classifier does better than rest of the classifiers and slightly better performance than GLDM and GLSZM-based DNN classifiers, which are the close challengers. Figure 9 displays the results of the analysis on the GLSZM-based area features, as well as one GLDM and one GLRLM feature. The distribution of small size






zones, their corresponding gray-level values, their variability, the variance in dependence size, and the joint distribution of shorter run lengths with greater gray-level values are clearly the key features that ranked higher in predicting pneumonia and can be used to distinguish between the x-ray images, according to the Integrated Gradients. Table 9 illustrates this concept by displaying six randomly selected photos, three of which are from each class and have feature values that the TexPneumClassifier evaluated as the highest.

As discussed earlier, we can see there is a clear separation boundary between first 4 features in the sampled images GLSZM small area emphasis, GLSZM small area high gray level emphasis, GLSZM size zone non-uniformity normalized and GLDM dependence variance. Higher values for these features lead to a higher prediction score for pneumonia class.

Model Comparison

In this part, we compare our suggested model's performance to other Texture feature-based methods being used in the field to detect pneumonia. In comparison to the other research, ours shows significant progress, but with just slight improvements.

Table 9:Sample Pneumonia and Normal Images with top Five Features from TexPneum Classifier

Class	Sample Images	GLSZM Small Area Emphasis	GLSZM Small Area High Gray Level Emphasis	GLSZM Size Zone Non-Uniformity Normalized	GLDM Dependence Variance	GLRLM Short Run High Gray Level Emphasis	Prediction Score(Pneumonia)
Pneumonia		0.5384	24.5657	0.2692	4.9234	25.1884	0.97066146
Pneumonia		0.5752	24.1918	0.3060	4.5891	21.1753	0.9291196
Pneumonia		0.5236	29.2045	0.2553	4.7401	29.9130	0.9992085
Normal		0.4870	20.5730	0.2219	4.3713	23.7380	0.27677116
Normal		0.4958	21.2272	0.2295	4.3050	25.4438	0.18925022


Normal		0.4972	20.4960	0.2328	4.0692	28.1906	0.11549588
--------	---	--------	---------	--------	--------	---------	------------

Table 10: Comparison of Results from Different Studies

	Model	Accuracy	Precision	Recall	F1Score
Moura <i>et al.</i> , (3)	XGB	0.82	0.82	0.82	0.82
	Random Forest	0.77	0.75	0.81	0.78
Wang <i>et al.</i> , (4)	LBP+SVM	77.59	76.18	77.59	76.87
	HOG+SVM	83.49	85.91	83.49	84.68
	VGG16	72.01	53.53	72.01	61.4
	MobileNet	85.32	82.68	85.32	83.97
	Inception V3	67.54	68.88	67.54	68.2
	ChexNet	79.8	88.05	79.8	83.72
	Proposed Model	92.19	93.44	92.19	92.81
Thepade <i>et al.</i> , (5)	Random Forest + MLP	84.814	88.835	56.7	74.784
	MLP	82.22	86.8	51.7	70.685
Shukla <i>et al.</i> , (6)	SVM	87.79	92.43	90.16	91.28
Panwar <i>et al.</i> , (17)	nConvNet	88.10	97.62	82	89.13
Konar <i>et al.</i> , (18)	Shallow Learning				
	Network	93.1	89	83.5	82.6
Sousa <i>et al.</i> , (19)	SVM	77			
	KNN	70			
Depeursinge <i>et al.</i> , (20)	Naïve Bayes	78.73			
	SVM	77.31			
Fernandez <i>et al.</i> , (21)	SVM (HOG + GLCM)	93.73			
	KNN (HOG + GLCM)	93.43			
PneumTexClassifier		94.6	97.0	95.8	96.4

The aforementioned outcome in Table 10 demonstrates how our method may improve the efficiency of medical image classification.

Conclusion

The study conducted in this research paper has demonstrated the potential of GLCM, GLSZM, GLRLM, NGTDM, and GLDM for classifying pneumonia. It was observed that these five image texture features can accurately differentiate lung images of pneumonia patients from those of healthy individuals, achieving good test classification accuracy. All the matrices except NGTDM showed promising results, with test accuracies exceeding 85%. Furthermore, the combination of all five feature sets proved to be more effective than any individual set, yielding an overall accuracy of 91%. This suggests that when used as part of a classifier, the combined texture features offer an effective and reliable method for identifying pneumonia in lung images. Additionally, due to its lightweight architecture and explainable output, the proposed system can serve as a triage tool or second-reader to assist radiologists in low-resource settings. However, a key limitation of this study is that the model has

not yet been evaluated across diverse radiography settings, equipment types, and patient populations. Further validation in varied clinical environments is essential to assess its generalizability and practical utility.

Abbreviation

None.

Acknowledgement

We extend our gratitude to the university for offering the necessary resources to carry out this research.

Author Contributions

Shubhra Prakash: Conceptualization, methodology, software, validation, resources, data curation, writing original draft preparation writing, review and editing, Ramamurthy B: formal analysis, investigation, visualization, supervision.

Conflict of Interest

The authors declare that there are no conflicts of interest related to this work.

Ethics Approval

Not applicable.

Funding

This study did not receive any financial support from government agencies or private organizations.

References

- Unni A, Eg N, Vinod S, Nair LS. Tumour Detection in Double Threshold Segmented Mammograms Using Optimized GLCM Features fed SVM. In: 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI). 2018:554–9.
- Suprijanto, Juliastuti E, Diputra Y, Mayantasari M, Azhari. Dental panoramic image analysis on mandibular bone for osteoporosis early detection. In: 2013 3rd International Conference on Instrumentation Control and Automation (ICA). 2013:138–43.
- Moura LV de, Mattjie C, Dartora CM, Barros RC, Marques da Silva AM. Explainable Machine Learning for COVID-19 Pneumonia Classification With Texture-Based Features Extraction in Chest Radiography. *Frontiers in Digital Health*. 2022; 3:662343. doi:10.3389/fdgth.2021.662343
- Wang S, Wen J, Huizhu Y, Xue Y, Tuo H. Experimental investigation on heat transfer enhancement of a heat exchanger with helical baffles through blockage of triangle leakage zones. *Applied Thermal Engineering*. 2014 Jun 1;67:122–30.
- Thepade SD, Jha H. COVID-19 Identification using Machine Learning Classifiers with GLCM Features of Chest X-ray Images. *Trends in Sciences*. 2021 Nov 20;18(23):46–46.
- Shukla P, Verma A, Abhishek, Verma S, Kumar M. Interpreting SVM for medical images using Quadtree. *Multimed Tools Appl*. 2020 Oct 1;79(39):29353–73.
- Wei Q, Hu Y. A Study on Using Texture Analysis Methods for Identifying Lobar Fissure Regions in Isotropic CT Images. *Conference proceedings: . Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference*. 2009 Sep 1;2009:3537–40.
- Sornam M, Prabhakaran M. A new linear adaptive swarm intelligence approach using back propagation neural network for dental caries classification. In: 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI). 2017:2698–703.
- Prihatini RS, Setyaningrum AH, Shofi IM. Texture analysis and fracture identification of lower extremity bones X-ray images. In: 2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI). 2017:1–5.
- Gaike V, Akhter N, Kale KV, Deshmukh P. Application of higher order GLCM features on mammograms. In: 2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT). 2015:1–3.
- Kumar VA, Jayanthi AK. Classification of MRI images in 2D coronal view and measurement of articular cartilage thickness for early detection of knee osteoarthritis. In: 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). 2016:1907–11.
- Loganathan GB, M. P, D. JR. Intelligent classification technique for breast cancer classification using digital image processing approach. In: 2019 International Conference on Smart Structures and Systems (ICSSS). 2019:1–6.
- Junaedi I, Yudaningtyas E, Rahmadwati R. Tuberculosis Detection In Chest X-Ray Images Using Optimized Gray Level Co-Occurrence Matrix Features. In: 2019 International Conference on Information and Communications Technology (ICOIACT). 2019:95–9.
- Sarosa SJA, Utaminingrum F, Bachtiar FA. Mammogram Breast Cancer Classification Using Gray-Level Co-Occurrence Matrix and Support Vector Machine. In: 2018 International Conference on Sustainable Information Engineering and Technology (SIET). 2018:54–9.
- FarzanaShaharBanu A, Kayalvizhi M, Arumugam B, Gurunathan U. Texture based classification of dental cysts. In: 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT). 2014:1248–53.
- Kermany D, Zhang K, Goldbaum M. Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification. 2018 Jan 6;2. <https://data.mendeley.com/datasets/rscbjbr9sj/2>
- Panwar H, Gupta PK, Siddiqui MK, Morales-Mendez R, Singh V. Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. *Chaos, Solitons& Fractals*. 2020 Sep 1;138:109944.
- Konar D, Panigrahi BK, Bhattacharyya S, Dey N, Jiang R. Auto-Diagnosis of COVID-19 Using Lung CT Images With Semi-Supervised Shallow Learning Network. *IEEE Access*. 2021;9:28716–28.
- Sousa RT, Marques O, Soares FAAMN, Sene IIG, de Oliveira LLG, Spoto ES. Comparative Performance Analysis of Machine Learning Classifiers in Detection of Childhood Pneumonia Using Chest Radiographs. *Procedia Computer Science*. 2013 Jan 1;18:2579–82.
- Depeursinge A, Iavindrasana J, Hidki A, Cohen G, Geissbuhler A, Platon A, et al. Comparative performance analysis of state-of-the-art classification algorithms applied to lung tissue categorization. *J Digit Imaging*. 2010 Feb;23(1):18–30.
- Fernandez-Grandon C, Soto I, Zabala-Blanco D, Alavia W, Garcia V. SVM and ANN classification using GLCM and HOG features for COVID-19 and Pneumonia detection from Chest X-rays. In: 2021 Third South American Colloquium on Visible Light Communications (SACVLC). 2021:01–6.