Vital Annex: International Journal of Novel Research in Advanced Sciences
ISSN: 2751-756X
Volume 04 Number 08 (2025)
https://innosci.org/IJNRAS

Article

Pneumonia Classification using SVM and GLCM, LBP and HOG Features

Abeer N. Faisal

1. Computer Information Systems Dep., College of Computer Science and Information Technology, University of Sumer, Thi-Qar, Rifai, Iraq

* Correspondence: abeernaser13@gmail.com

Abstract: A type of lung disease is Pneumonia that infect the respiratory system, and this disease can be recognized by X-ray images for the chest area. The presence of pneumonia or not need the medical professional to examine the X-ray image, and the rule of his decision not ignored, but may suffer from some problem that can led to mistake. Many studies introduced in this field by using new technology, artificial intelligent and deep-learning technics to reduce the error ratio. In this work we introduce a model that correctly classify the chest X-ray as a pneumonia or not, GLCM, LBP and HOG are used as a feature extraction methods after preprocess the used images. SVM used here as a classifier and the obtained results show the effectiveness and robustness SVM when used with a suitable preprocessing steps and prices feature extraction methods. The model Accuracy is 99.90% with Recall of 100%, Precision : 99.87 %, and F1-score: 99.94%. The used dataset is chest X-ray from kaggle with 5,216 chest X-ray images

Keywords: Pneumonia, SVM, GLCM, LBP, HOG.

Citation: Faisal, A. N. Pneumonia Classification using SVM and GLCM, LBP and HOG Features. Vital Annex: International Journal of Novel Research in Advanced Sciences 2025, 4(8), 324-331

Received: 15th Jun 2025
Revised: 29th Jul 2025
Accepted: 17th Aug 2025
Published: 10th Sep 2025

1. Introduction

The death may be causes by the age progress and backward for the vital body systems, and may cause by another reasons. The more popular reason is the infection by some diseases. People who suffer from weak immune system, are the most infected with disease. One of the human body systems is the respiratory system that can be the main system in the body. A serious disease that affect part of this system is pneumonia. A pneumonia is a dangerous disease that make the human ability to breathe through his lungs so difficult and this difficulty became worse if this disease not diagnose at its early stages. Pneumonia infects people and led to approximately many thousands of deaths every year, so patients that suffer from it may be under five years old or aged people with inefficient immune system [1]. Pulmonary infection led to pneumonia that may be caused by bacterial or viral invasion to lung tissues causing air sac inflammation and pleural effusion [2]. There are many methods to detect this disease that include (Chest X-rays (CXR) or lung CT scans, Lung function tests., Bronchoscopy, Sputum testing, Lung biopsy surgery) [3]. These ways need the medical professional to examine whether the patient suffer from pneumonia or not. A radiograph of the chest is created by X-ray penetration. The tissues that are hard such as bones appeared brighter than soft tissues that will be dark, and many infections (like cardiac infections, swollen blood vessels and cancer cells) can have detected if these tissues color appear darker. This X- ray is used to determine the location and the covered area of lung infection, and the researchers can be easily differing between free and these x-rays that have the disease by using modern processing and classification methods as well as the use of AI and deep learning [4]. The main rule of the radiologists to recognize the disease is not ignored, but it is may be led to a mistake that occur due to naked eye or work stress and effort. Over the past 25 years, researcher solve this problem by exchanging these traditional ways with more modern methods related with AI, deep learning and other image processing techniques [5]. The use and need to automated detection systems is essential issue to overcome subjective variability in pneumonia detection [6]. In this paper we introduce a model that can automatically recognize pneumonia infection. Faraway from using deep-learning and AI trained models, we use tradition image processing and features extraction methods to enhancing the images quality and extract the significant image features. The other parts of the paper are including: related works, the Methodology include: the used dataset, preprocessing, feature extraction, dimensionality reduction, classification and testing, after that evaluation introduced, then compare with other studies, discussion and conclusion.

2. Materials and Methods

Literature Review

The main role in this scope is the used datasets. In 2017, Wang et al. introduce a new dataset which contain (108,948 samples of X-ray) and 32,717 records for patient and obtained good results using CNN model [7].

In 2018 [8], the researchers proposed a network named (CheXNeXt), and employed a deep learning to avoid the small number of available data and produce an ensemble work of three models (DenseNet-121, ResNet-18 and GoogLeNet), and they used X-rays to recognize 14 chest illnesses (including pulmonary, pleural effusion, pneumonia and others).

While in 2019, A. K. Jaiswal et al. proposed a DL model which mixed global and local features for segmentation and enhancing images, the model called Mask-RCNN [9].

G. Liang and L. Zheng in 2020, proposed a framework based on the nature of child pneumonia image, they overcome problems such as over-fitting and the degradation by using the residual structure, and solve the loss of feature space information by implement dilated convolution [10].

In 2021, Adhiyaman Manickam et al. utilized a DL transfer for the detection process and accuracy improvement, ImageNet dataset and pre-trained models were used for detect pneumonia (normal or abnormal). They used two optimizers namely Adam and Stochastic Gradient Descent (SGD) to extract features and improve the accuracy for the pre-trained models [11], [12].

Rohit Kundu et al. in 2021, employed deep transfer learning to overcome broplem small available dataset and the used CNN models were: DenseNet-121, ResNet-18 and GoogLeNet [13].

Tuncer et al. used some traditional classifiers for local features extraction to classification and a fuzzy tree transformation applied to the X-rays. To improve the model performance, three features images generated and used them the work [14].

In 2021, a confidence-aware module developed by J. Zhang et al. that implemented on lung X-rays for detecting lung anomalies. Despite improving the performance, the obtained sensitivity on the dataset was too low for practical use [15].

Pramoth K M et al. used a CNN model to recognize if the X-ray image has Pneumonia to speed the treatment process, the model achieved high accuracy [15].

In 2023, the researcher produced a simple Rest API model used to classify normal or pneumonia (bacteria or virus) images (bacteria or virus), they used Multi-Layer Perceptron (MLP) with a training accuracy 73.89% [16].

In 2024, Bao Q. Bui et al. presented a detection and classification of lung infection, they proposed a new DL architecture composed of graph transformer networks and a traditional deep neural network module and recorded effective classification of silicosis and pneumonia infection [17].

The modified framework for recognize and classify the pneumonia is explained in this section, the traditional machine learning pipeline used here, which include (preprocessing steps, then the features extracted after that we use Principle Component Analysis (PCA) for reduce the dimension, the classification done using Support Vector Machine (SVM) [18] and finally the accuracy, confusion matrix and ROC curve are evaluate the work accuracy.

3.1 Dataset

The dataset we use to implement this study is Chest X-Ray Images (Pneumonia) with 5,216 JPEG X-ray images that are separated into two categories (1,341 normal and 3,875 pneumonia images with percentage of 80% for training and 20% for testing), figure1 below shows the count of used images.

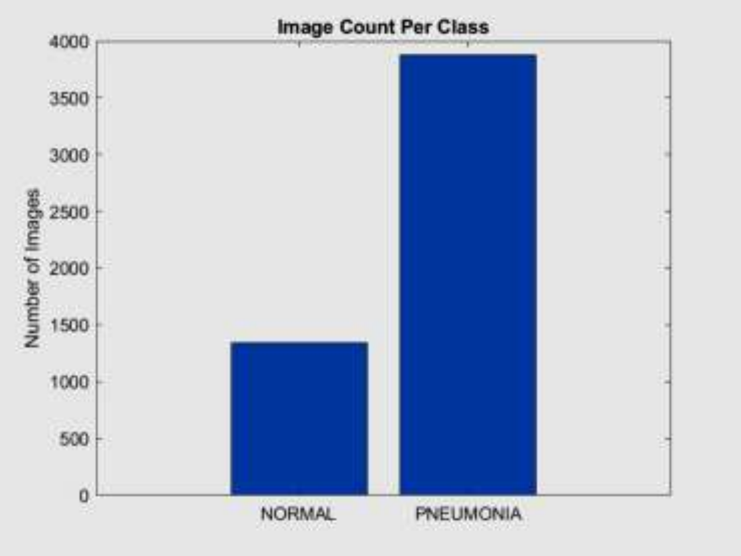

Figure 1 : Images count.

3.2 Preprocessing

To enhance the used images quality and extract more useful features, the preprocessing is an essential phase. The first step is resize images to be with a fixed size to have a consistent feature dimensions. The next step is convert RGB images to gray-scale (the X-ray image contain intensity structural information). After that median filter used to reduce salt-pepper noise and preserve edges, this filter commonly used in medical images. Contrast Limited Adaptive Histogram Equalization (CLAHE) is implemented to adjust visibility of lung structures. To emphasize important structures sharpening is used here, and finally wiener filter is used to reduce noise without blurring image edge [19, 20, 21]. Figure 2 shows X-ray image before preprocessing and after it.

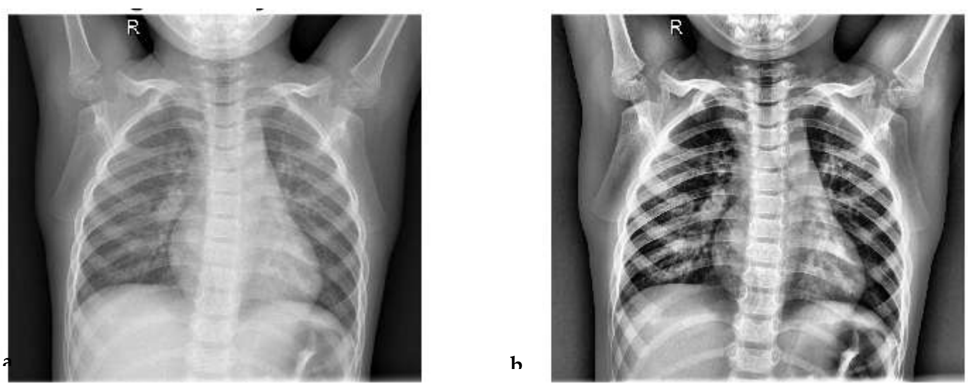

Figure 2: (a) X-ray image before preprocessing, (b): after preprocessing.

3.3.Feature Extraction

The useful features are extracted from preprocessed images by apply three techniques that have the ability to capture different image features (like texture, edge and shape information). These techniques are Gray-Level Co-occurrence Matrix (GLCM), Local

Binary Patterns (LBP) and Histogram Oriented Gradient (HOG), the combine of them create a rich and descriptive feature vector for image classification.

- **a)** Gray-Level Co-occurrence Matrix (GLCM): GLCM used for second-order statistical texture features extraction, it can find how often pixel pairs in an image with specific values in a specified spatial relationship occur, useful for distinguishing tissues in X-ray images [22,23].
- **b)** Local Binary Patterns (LBP): LBP robust against illumination variation, extract fine texture patterns (such as edges, flat areas and spots) by divide the image into 32x32 patches, compare each pixel with its neighbors and produce the histogram from encoded local texture (0 or 1) [24, 25].
- **c)** Histogram Oriented Gradient (HOG): it use local 16x16 pixels cell to find edge orientations and produce a histogram of direction and magnitudes of edges, this technique robust for find shape and structure (like lung outlines or lesions) [26, 27].

3.4 Dimensionality Reduction

To reduce the dimensionality of the feature space, decrease noise , speed up training and testing and high efficiency Principal Component Analysis (PCA) is applied here. PCA is a statistical approach that transform a high-dimensional feature space into a low-dimensional with the retain of variance (important features), see figure 3. PCA standardize the feature data by zero-means, find the features covariance matrix, eigenvectors and eigenvalues are computed and sorted, compute PCA transformation matrix by select the top k-eigenvectors and transform the original features into the new PCA space [28].

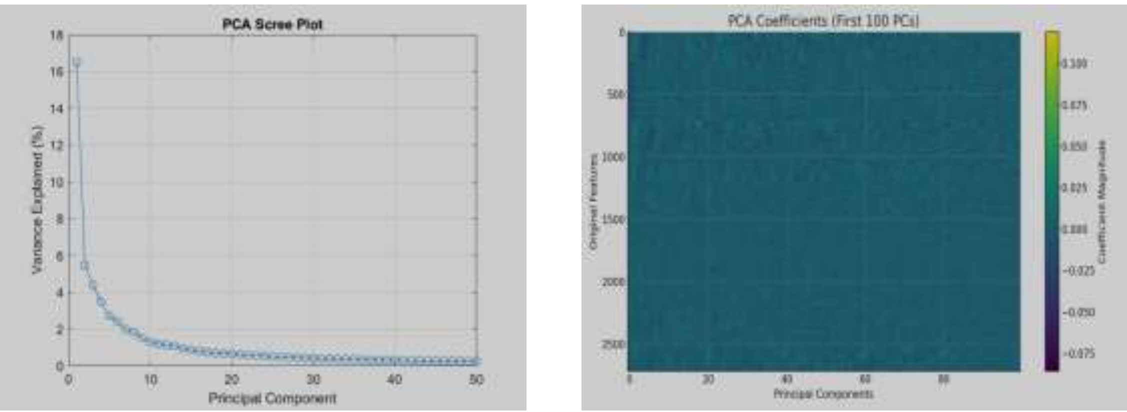

Figure 3: (a) PCA scree plot, (b): PCA coefficients (first 100 PCs).

b

3.5 Classification

For classification process, a linear Support Vector Machine (SVM) is used here to create a classifier that can correctly predict the testing data [28]. The classifier is trained on the training set (80% of the total images). For testing process, we use the testing dataset to evaluate the model accuracy and the obtained results explained in the confusion matrix. The produced SVM classification model will use a hyperplane to split the two image types as shown in the figure 4.

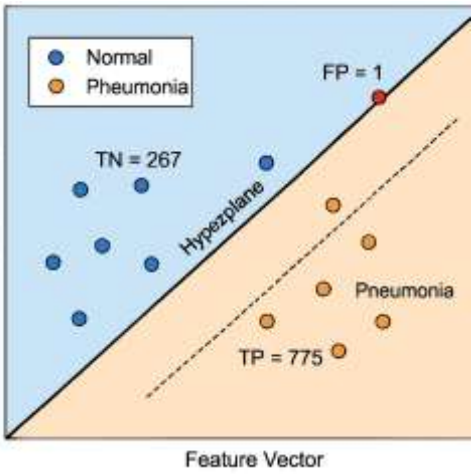

Figure 4: SVM separation

3.5 Testing

In this stage the confusion matrix shows the results of model testing, which achieve excellent accuracy. The result confusion matrix have four parts as follow, figure 5:

- True Positive (TP): show the correct classification of pneumonia images (775).
- True Negative (TN): show the correct classification of normal images (267).
- False Positive (FP): show the normal that classified as a pneumonia images (1).
- False Negative (FN): show the pneumonia that classified as a normal images (0).

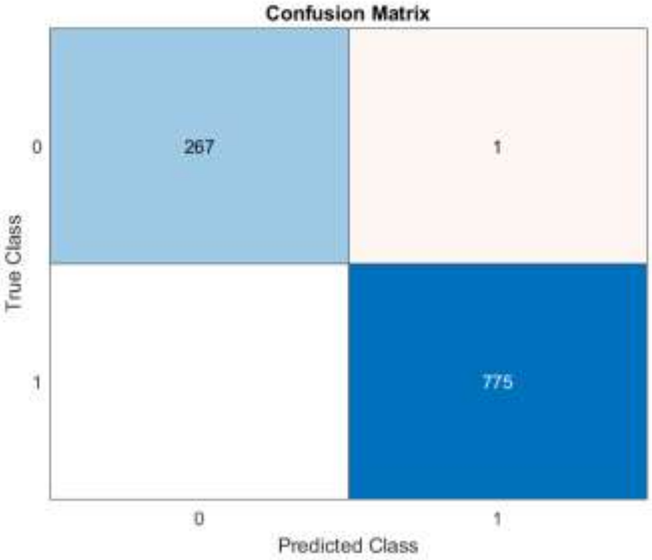

Figure 5: Confusion Matrix

3. Results and Discussion

Evaluation

From the above result matrix, the model performance can be evaluated using some important metrics (accuracy, precision, recall and F1-score) that show how the model separate the normal and pneumonia images. In this work, the values of these metrics are as follow: A**ccuracy : 99.90 %, Precision : 99.87 %, Recall: 100 %** and **F1-score: 99.94%.** These results indicate superior predictive performance. Figure 6 shows the ROU Curve and evaluation metrics [29], [30], [31]. The calculation of these metrics as below:

$$\boldsymbol{Accuracy} = (\boldsymbol{TP}+\boldsymbol{TN})/(\boldsymbol{TP}+\boldsymbol{TN}+\boldsymbol{FP}+\boldsymbol{FN}) \quad (1)$$

$$\textbf{Precision} = \boldsymbol{TP}/(\boldsymbol{TP}+\boldsymbol{FP}) \quad (2)$$

$$\textbf{Recall} = \boldsymbol{TP/(TP + FN)} \quad (3)$$

$$\textbf{F1} = \boldsymbol{2 * (Precision * Recall)/(Precision + RecallN)} \quad (4)$$

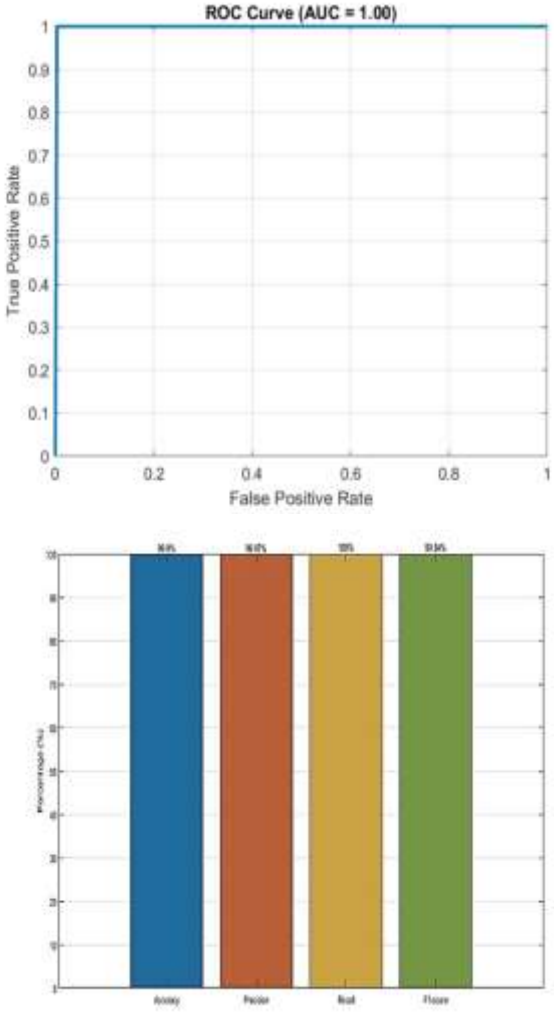

Figure 6: ROU Curve and key metrics

Comparison with other studies

The developed model show superior results with accuracy of **99.90** that overcome other previous studies. The use of multi descriptor such as (GLCM, LBP, and HOG) make the extracted features contain more prices and important image details. A comparison with the other studies explained in the table 1 below .

Table 1. Comparison with other studies

Study	Method	datasets	Accuracy
Rahman et al. , 2020,	SVM with wavelet features	1,200 X-rays	95.3%
Minaee et al. ,2020,	Deep learning	5,000 X-rays	97.8%
Risha Ambar Wati et al. ,2020. [28]	SVM and GLCM	5,853 X-rays	62.66
Kermany et al. ,2018,	Transfer Learning (CNN)	5,856 X-rays	92.8%
Proposed model ,2025	SVM	5,216 X-rays	99.90%

Discussion

The obtained results show excellent performance of the model using (GLCM, LBP and HOG) as a features extraction methods as well as using a suitable preprocessing steps the support the obtained accuracy. The GLCM find texture information, the local texture patterns captured using LBP, while the edge and shape can be represented by HOG. All these image information produce a rich feature vector that used in the classification process between normal and pneumonia X-ray chest-images. As the confusion matrix shows, all the pneumonia images are correctly labeled (TP=775), only one normal image labeled as pneumonia, and other normal images are labeled normal (TN=267). The recorded evaluation metrics are: **99.90 %, Precision : 99.87 %, Recall: 100 %** and **F1-score: 99.94%,** such results explain that using SVM with a rich feature vectors and appropriate preprocessing, is proper methods for classification tasks in medical work where consistency and classification are crucial.

4. Conclusion

The accuracy of the proposed model using SVM as a classifier achieve near-full results, that show the efficiency of using traditional ML methods to extract precise and effect information. The study prove that SVM is an efficient and robust classifier in the medical field when it used with the correct helpful features. For future improvement we suggest the following:

- Increasing the dataset.
- Testing the model with larger testing set.
- Comparing the model with DL models.

REFERENCES

[1] Shahzad, M. A. Arshed, F. Liaquat, M. Tanveer, M. Hussain, and R. Alamdar, "Pneumonia classification from chest X-ray images using pre-trained network architectures," *AWKUM Transactions on Computer Sciences*, vol. 10, no. 2, pp. 34–44, Jul.–Dec. 2022.

[2] T. Rajasenbagam, S. Jeyanthi, and J. A. Pandian, "Detection of pneumonia infection in lungs from chest X-ray images using deep convolutional neural network and content-based image retrieval techniques," *Journal of Ambient Intelligence and Humanized Computing*, 2021, doi: 10.1007/s12652-021-03075-2.

[3] B. Q. Bui, T. T. T. Nguyen, D. M. Le, C. Tran, and C. Pham, "Innovative silicosis and pneumonia classification: Leveraging graph transformer post-hoc modeling and ensemble techniques," *IEEE Journal of Biomedical and Health Informatics*, arXiv:2501.00520 [cs.CV], Dec. 31, 2024.

[4] A. Manickam, J. Jiang, Y. Zhou, A. Sagar, R. Soundrapandiyan, and R. D. J. Samuel, "Automated pneumonia detection on chest X-ray images: A deep learning approach with different optimizers and transfer learning architectures," *Measurement*, vol. 184, Nov. 2021, Art. no. 109953, doi: 10.1016/j.measurement.2021.109953.

[5] R. Kundu, R. Das, Z. W. Geem, G. T. Han, and R. Sarkar, Pneumonia detection in chest X-ray images using an ensemble of deep learning models, PLoS One, vol. 16, no. 9 September, Sep. 202 1 , doi:10.1371/joumal.pone.0256630.

[6] G. Williams *et al.*, "Variability and accuracy in interpretation of consolidation on chest radiography for diagnosing pneumonia in children under 5 years of age," *Pediatric Pulmonology*, vol. 48, pp. 1195–1200, 2013, doi: 10.1002/ppul.22806.

[7] X. Wang *et al.*, "Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 3462–3471.

[8] P. Rajpurkar *et al.*, "Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists," *PLoS Medicine*, vol. 15, no. 11, p. e1002686, 2018.

[9] A. K. Jaiswal *et al.*, "Identifying pneumonia in chest X-rays: A deep learning approach," *Measurement*, vol. 145, pp. 511–518, 2019.

[10] G. Liang and L. Zheng, "A transfer learning method with deep residual network for pediatric pneumonia diagnosis," *Computer Methods and Programs in Biomedicine*, vol. 187, p. 104964, 2020.

[11] T. Tuncer, F. Ozyurt, S. Dogan, and A. Subasi, "A novel Covid-19 and pneumonia classification method based on F-transform," *Chemometrics and Intelligent Laboratory Systems*, vol. 210, Mar. 2021, Art. no. 104256.

[12] J. Zhang *et al.*, "Viral pneumonia screening on chest X-rays using confidence-aware anomaly detection," *IEEE Transactions on Medical Imaging*, vol. 40, no. 3, pp. 879–890, Mar. 2021, doi: 10.1109/TMI.2020.3040950.

[13] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York, NY, USA: Springer, 2002.

[14] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[15] Y. Zhang, Z. Dong, L. Wu, and S. Wang, "A hybrid method for MRI brain image classification," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10049–10053, 2011.

[16] C. Bergmeir and J. M. Benítez, "On the use of cross-validation for time series predictor evaluation," *Information Sciences*, vol. 191, pp. 192–213, 2012.

[17] N. Sharma, A. K. Ray, S. Sharma, and K. K. Shukla, "Automated medical image segmentation techniques," *Journal of Medical Physics*, vol. 35, no. 1, pp. 3–14, 2010.

[18] Kaggle, "Chest X-ray pneumonia," [Online]. Available: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia.

[19] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. Upper Saddle River, NJ, USA: Pearson, 2018.

[20] W. K. Pratt, *Digital Image Processing: PIKS Scientific Inside*, 4th ed. Hoboken, NJ, USA: Wiley-Interscience, 2007.

[21] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*, P. S. Heckbert, Ed. San Diego, CA, USA: Academic Press, 1994, pp. 474–485.

[22] S. M. Suryanarayana and A. N. J. Raj, "Texture-based medical image classification using GLCM and deep learning techniques," *Biomedical Signal Processing and Control*, vol. 83, p. 104642, 2023.

[23] L. Alzubaidi *et al.*, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, p. 53, 2021.

[24] S. Yasmin, S. Sharmin, S. Akter, and M. Paul, "Local binary pattern-based feature extraction for lung disease classification using chest X-ray images," *Computers in Biology and Medicine*, vol. 153, p. 106552, 2023.

[25] N. S. Alghamdi and G. Muhammad, "Automated detection of pneumonia in chest X-ray images using local binary patterns and deep learning," *IEEE Access*, vol. 10, pp. 25305–25315, 2022.

[26] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2005, pp. 886–893.

[27] M. T. Islam, M. M. Rahman, and S. Mahmud, "Lung disease detection from chest X-ray images using HOG features with machine learning," *Informatics in Medicine Unlocked*, vol. 38, p. 101168, 2023.

[28] R. A. Wati, H. Irsyad, and M. E. A. Rivan, "Klasifikasi pneumonia menggunakan metode support vector machine," *Jurnal Algoritme*, vol. 1, no. 1, pp. 21–32, 2020.

[29] T. Rahman *et al.*, "Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray," *Applied Sciences*, vol. 10, no. 9, p. 3233, 2020.

[30] S. Minaee *et al.*, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," *Medical Image Analysis*, vol. 65, p. 101794, 2020.

[31] D. Kermany *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.