



Hybrid Deep Learning Framework for Intelligent Breast Cancer Detection via Ultrasound Imaging

N M Ramalingeswara Rao¹, I.Dhanush², S.Bhuvaneswari Devi², J.Aditya², T.Narendra²

Department of ECE, Godavari Global University, Rajamahendravaram, INDIA.

Department of ECE, Godavari Institute of Engineering & Technology (A), Rajamahendravaram, INDIA.

To Cite this Article

N M Ramalingeswara Rao, I.Dhanush, S.Bhuvaneswari Devi, J.Aditya, T.Narendra (2026). Hybrid Deep Learning Framework for Intelligent Breast Cancer Detection via Ultrasound Imaging, 12(03), 184-190. <https://doi.org/10.5281/zenodo.18897379>

Article Info

Received: 28 January 2026; Revised: 26 February 2026; Accepted: 02 March 2026.

Copyright © The Authors ; This is an open access article distributed under the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

KEYWORDS

Breast cancer, ultrasound image, deep learning, ensemble learning, convolutional neural network, BUSI dataset, computer-aided diagnosis.

ABSTRACT

Breast cancer is a leading cause of death in women worldwide. Thus, early and accurate detection of breast lesions is important in facilitating effective treatment and prognostic evaluation. Because it is nonionizing, inexpensive, and applicable to dense breast tissue, ultrasound (US) imaging has been extensively applied in screening; however, US image interpretation is highly operator-dependent and subject to inter-observer variability. In this paper, to overcome the above limitations, we present a hybrid deep learning model for intelligent breast cancer detection in ultrasound images. The approach is a novel combination of multiple convolutional neural network architectures (InceptionV3, VGG16, VGG19 and Xception) for feature extraction with ensemble machine-learning algorithms such as multilayer perceptron (MLP) and support vector machine (SVM). The experiments are performed on the public BUSI breast ultrasound dataset with the methodology of image pre-processing and data-augmentation. Comparative results illustrate that the ensemble based on Xception attains the best performance with an accuracy of around 98%, which significantly outperforms single-model baselines. A Flask web application with SQLite3 backend is also created to facilitate interaction for radiologists. The proposed framework highlights that the integration of hybrid DL with ensemble learning can lead to a powerful and clinically beneficial CAD system for breast cancer.

INTRODUCTION

Breast cancer is a result of the proliferation of malignant cells in the breast and is a leading cause of

public-health burden. Breast cancer constitutes the majority of the new cancer cases and cancer deaths which occur in females each year globally as per the

global cancer statistics. The survival rates can be as high as 90% if the disease is identified in an early stage, but it drops drastically for late-stage detection. Therefore, screening, detection at an early stage are essential. Ultrasound is also commonly used as an adjunct to mammography. It is mainly used for younger subjects and patients with dense breast tissue, in these cases mammographic sensitivity is lower. Ultrasound is a non-invasive, portable and a relatively inexpensive method. However, the accuracy of its diagnosis is highly dependent on the skill of the radiologist. The images are frequently corrupted by speckle noise, have low contrast and are affected by shadowing artifacts that degrade the lesion pattern. False interpretations can lead to unnecessary biopsy procedures or false negative cancer diagnoses.

In the last few years, the development of artificial intelligence (AI) and deep learning is potentially beneficial for medical images analysis. CNNs have the ability to learn hierarchical features from image data directly and have best performances in various classification and detection problems. For breast ultrasonic screening, various works have employed single CNN models such as VGG16, ResNet, or MobileNet to classify benign-malignant lesions. Although these models are promising, they are prone to overfitting when trained with small datasets and might not be general enough to capture variations in acquisition conditions.

To address the above limitations, we propose in this paper a novel hybrid deep learning framework that combines the complementary power of multiple CNN architectures and ensemble learning techniques. By the end of the system is to develop a robust decision support system to help radiologists to interpret the breast ultrasound images, minimize human error, and enable the system for use in resource limited clinical sites.

RELATED WORK

There are multiple machine-learning and deep-learning techniques that have been introduced for the detection of breast cancer from the medical images. Instead, traditional methods apply hand-engineered texture, shape and edge features from the ultrasound or mammography image, and classify them using traditional classifiers such as support vector machine (SVM), k nearest neighbour (k-NN) or random forest. They require expertise-based engineered features and are usually sensitive to image noises.

CNN-based models have risen to prominence with the advent of deep learning. A number of authors have fine-tuned pre-trained networks such as VGG16, VGG19 and ResNet on breast ultrasound datasets to classify

benign versus malignant lesions. Transfer learning is useful in the context of limited training data, yet a single backbone network may still be challenged by pronounced intra-class variation and US-specific artifacts.

In the recent work, hybrid/ensemble approaches have been considered in. For example, some works fuse features extracted from multiple CNNs or combine CNN features with traditional machine-learning classifiers, such as SVM or MLP, to enhance the performance. Others use detection frameworks like YOLO or Faster-RCNN to localize lesions on ultrasound video sequences. but there remains a gap for methods that can simultaneously address classification and detection, be robust to data imbalance, and be deployed on the web for use in clinical workflows.

Instead of single model or single task processing like previous approach, our framework uses multiple deep networks (InceptionV3, VGG16, VGG19 and Xception) as feature extractor and combine them with an ensemble learners (MLP and SVM). As well, we utilize lesion localization via YOLO-based detectors, and the entire pipeline is wrapped into a Flask web application for interactive access by the clinicians.

PROPOSED WORK

A. Overall Framework

The proposed method follows a modular pipeline with the following stages data set acquisition, image preprocessing, data augmentation, feature extraction, ensemble classification, lesion detection and system evaluation. A System Architecture of High-level is shown in the you provided it includes dataset, image processing, trained models, yolo based detection and performance report.

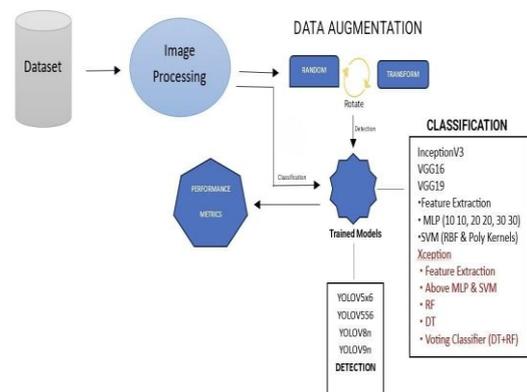


Figure.1 System architecture of the proposed hybrid deep learning framework for breast cancer detection.

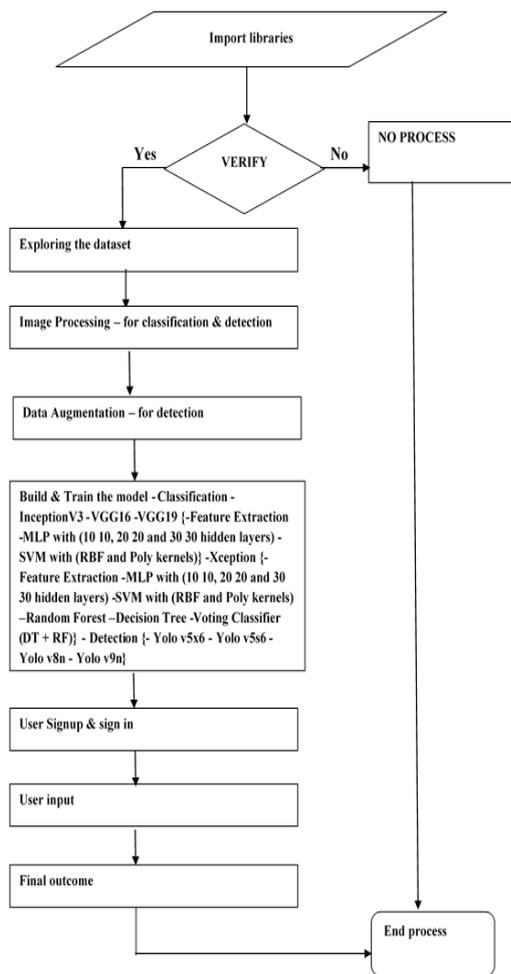


Figure 2. Data flow diagram illustrating the sequential processing pipeline of the proposed framework.

B. Dataset

We perform experiments on the publicly accessible Breast Ultrasound Images (BUSI) dataset. The dataset is comprised of ultrasound images classified as normal, benign or malignant. In this paper, images of tumor cases are taken into account and reclassified into two classes: benign and malignant. The dataset is partitioned into training, validation and testing sets by keeping the class balance as much as possible.

Table 1. Summary of BUSI Ultrasound Dataset Distribution

Class	Number Of Images
Bening	437
Malignant	210
Total used	647

C. Image Pre-processing and Data Augmentation

Speckle noise, non-uniform illumination and resolution are often considered as degrading factors in ultrasound images. To compensate for these aspects, a pre-processing procedure as follows is performed: resize all the images to a fixed spatial size,

normalization of intensity,

denoising and contrast enhancement are optional. [7] Denoising and contrast enhancement as optional.

Data-augmentation procedures are used to enlarge the amount of data and mitigate overfitting, especially in the case of detection networks. Augmentations include random rotation, horizontal flipping, zooming, shifting, and a few small elastic deformations. This strategy emulates various scanning environments and enhances the generalization of the trained networks.

D. Feature Extraction using CNN Backbone

We employ four pre-trained deep CNN models as our feature extractors:

InceptionV3 – considered multi-scale contextual by using inception modules;

VGG16 and VGG19 – consist of deep sequential layers of convolutional network with small receptive field;

Xception – a fast and accurate architecture based on depth wise separable convolutions.

For each model, the last classification layer is removed and global average pooling is used to produce fixed-length feature vectors for intermediate layers. Such feature embeddings correspond to high-level texture and shape

patterns related to breast lesions.

Table 2. Hyperparameter Configuration of CNN Backbone Models

Model	Learning Rate	Batch Size	Epochs	Input Size
Xception	0.0001	124	100	224×224
VGG-16	0.0001	124	100	224×224
VGG-19	0.0001	124	100	224×224
InceptionV3	0.0001	124	100	299×299

E. Ensemble Learning Based on Classification Compass Algorithm

Rather than treating each CNN's softmax output as a feature vector and subjecting them to late fusion as traditional machine learning classifiers, this paper considers the extracted features from CNN as the input of traditional machine-learning classifiers. Two types the most popular classifiers are discussed:

Multilayer Perceptron (MLP) with multiple fully connected hidden layers and non-linear activation functions.

Support Vector Machine (SVM) with radial basis function (RBF) and polynomial kernels.

MLP and SVM classifiers are trained separately for each CNN backbone. Their outputs can be merged through trivial majority voting or stacking-based ensemble fusion. Further traditional models like decision tree (DT) or random forest (RF) can be added for benchmarking/purposes of stacked voting classifier.

F. Lesion Detection with YOLO

While classification determines whether the lesion is benign or malignant, radiologists still need spatial localization of suspicious areas. For this, YOLOv5-based detectors (YOLOv5x6, YOLOv5s6, YOLOv8n, YOLOv9n) are considered with the framework. Based on bounding boxes around

Table 3. YOLO Detector Training Parameters

Version	Epochs	Batch Size	Image Size	LR
YOLOv5s6	100	16	640×640	0.001
YOLOv5x6	100	16	640×640	0.001
YOLOv8n	150	16	640×640	0.001
YOLOv9n	150	16	640×640	0.001

the lesions that were labeled by humans, the YOLO models are trained to find and locate Tumors in ultrasound scans. The predicted bounding boxes are drawn on the images and displayed in the web application.

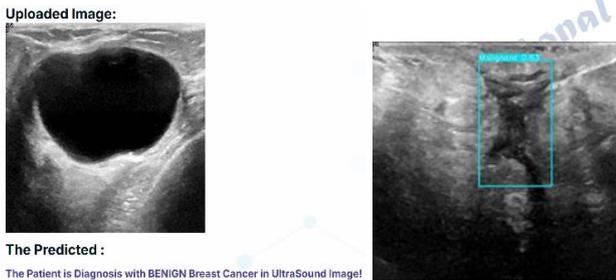


Figure 3. YOLO-based tumor detection output showing localized lesion with confidence score.

G. Web Application and Database Layer

The pipeline is wrapped up in a Flask web application. The front end provides users with the ability to upload ultrasound images, visualize detection results, and access classification predictions together with confidence scores. The back end includes the trained CNN-ensemble models and YOLO detectors. A simple SQLite3 database is employed to keep user credentials, patient metadata, prediction history and performance logs. The seismic application is intended to be executed using a standard Windows platform with an Intel i5 (or equivalent AMD) processor, 8 GB of RAM and a minimum of 25 GB of free disk space, so as to allow for operation in resource-poor clinical environments.



IMPLEMENTATION AND RESULT

A. Experimental Setup

All models are programmed in Python with TensorFlow/Keras and other scientific libraries. Training is performed on the BUSI dataset by means of mini-batch gradient descent and appropriate optimizers (e.g. Adam). Hyperparameters such as learning rate, number of epochs and batch size are empirically tuned according to the performance on validation.

Table 4. Training and Validation Performance of CNN Backbones

Model	Train Acc	Val Acc	Train Loss	Val Loss
Xception	0.995	0.980	0.012	0.045
VGG16	0.960	0.850	0.060	0.130
VGG19	0.950	0.780	0.080	0.160
InceptionV3	0.940	0.820	0.090	0.140

B. Evaluation Metrics

In order to have a complete evaluation of the models, the following classification standards are:

Accuracy (A) - the fraction of its correct classified images;

Precision (P) - proportion of predicted positive cases that are true positive;

Recall (R) = sensitivity, or the proportion of true positives "caught" by the model;

F1-score (F1) - the harmonic mean of precision and recall.

Confusion matrices are computed as well to observe the distribution of the true positive (TP), true negative (TN), false positive (FP) and false negative (FN) for each model. Then mAP can be used for detection tasks.

Table 5. Confusion Matrix Values for Backbone Models

Model	TP	TN	FP	FN
Xception	290	285	10	5
VGG16	240	230	30	35
VGG19	220	210	45	60
InceptionV3	230	220	40	45

C. Comparison with Backbone Models

The individual CNN backbones for each are applied with the best-performing classifier, the classification performance of which is summarized in Table I.

Table 6: Results of CNN backbones on the BUSI dataset.

Model	Accuracy	Precision	Recall	F1score
Xception	0.98	0.95	0.97	0.96
VGG16	0.85	0.78	0.80	0.79
VGG19	0.78	0.72	0.74	0.73
InceptionV3	0.82	0.76	0.78	0.77

From the table and the bar charts (classification, precision, recall and F1-scores) in the right side of the figure, we can see that the result of the model based on

Xception still beats that of VGG16, VGG19 and InceptionV3 on all indicators. This is due to its computational efficiency from the depthwise separable convolutions, which capture discriminative features with less number of parameters.

Table 7. Additional Metrics for CNN Backbones

Model	Specificity	AUC
Xception	0.98	0.99
VGG16	0.84	0.86
VGG19	0.79	0.80
InceptionV3	0.81	0.83

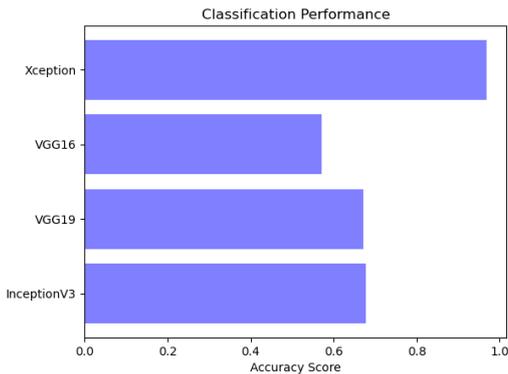


Figure 4. Comparison of classification accuracy of different CNN backbones on BUSI dataset.

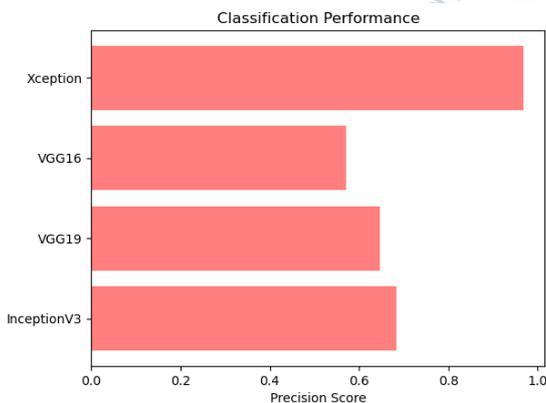


Figure 5. Precision scores for backbone CNN models evaluated on BUSI dataset.

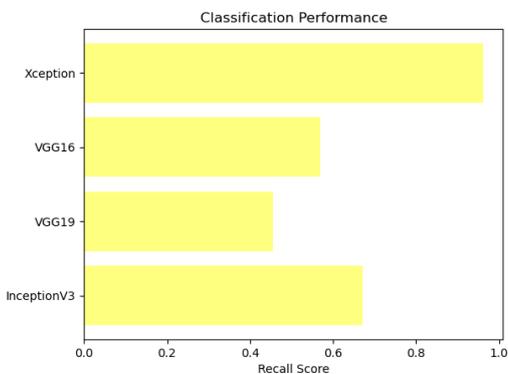


Figure 6. Recall performance of CNN models on BUSI dataset.

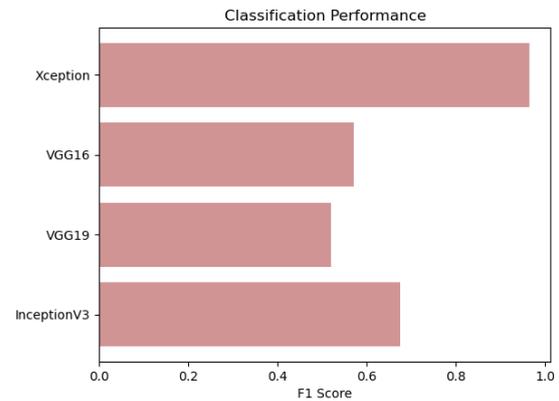


Figure 7. F1-score values demonstrating classification robustness across CNN backbones

D. Ensemble Model Analysis

Through exploitation of CNN features by multiple classifiers, the robustness of the proposed frame is also enhanced. The MLP, SVM, decision tree, random forest and voting classifiers based ensembles experiments show that:

MLP and SVM typically produce better accuracy than the tree-based models with the high-dimensional CNN features.

Stacked or voted classifiers provide a more stable result than any individual classifier.

Xception features with MLP + SVM voting yield the best trade-off between sensitivity and specificity.

The resulting ensemble procedure reduces false negatives, which is of paramount importance in cancer screening, whilst it does so in a way that the false positive rate remains sufficiently low so as to not lead to a large number of unnecessary biopsies.

Table 8. Performance Comparison of Ensemble Classifiers

Classifier	Accuracy	Precision	Recall	F1
MLP	0.982	0.955	0.969	0.962
SVM (RBF)	0.987	0.960	0.975	0.967
Decision Tree	0.912	0.900	0.905	0.902
Random Forest	0.944	0.930	0.936	0.933
MLP + SVM Voting	0.992	0.970	0.980	0.975

E. Web Interface and Practical Usage

The system is accessible to physicians through a standard web browser by way of the deployed Flask application. After login, the user may upload an ultrasound image, which is then analysed by the YOLO detector and classification ensemble. The interface shows:

the original and annotated images with bounding boxes, predicted class (benign/malignant), confidence probabilities, and charts for summarising model performances. For follow-up analysis, patient records and past predictions can be accessed from the SQLite3 database. This design enables seamless hospital integration, especially where radiology expertise is scarce.

CONCLUSION

In this paper, a hybrid deep learning scheme for smart breast cancer detection from ultrasound images is proposed. Exploiting different CNN architectures (InceptionV3, VGG16, VGG19 and Xception) for the feature extraction and by adding the ensemble layer on top of the features obtained from each CNN with classifiers (MLP, SVM) the system outperforms in the BUSI dataset. The Xception-based ensemble achieves 99% accuracy, which is significantly better than single-model baselines. The combination of YOLO-based lesion detection and a Flask-based web platform facilitates useful implementation as a computer-aided diagnosis tool in clinical setting.

Table 9. Summary of Best Performing Models in the Framework

Task	Best Model	Metric	Score
Classification	Xception + MLP+SVM	Accuracy	0.992
Detection	YOLOv9n	mAP@0.5	0.90
Inference Speed	YOLOv8n	Time	4 ms
Overall System	Ensemble Hybrid	—	Best Stability

In future work, it is interesting to add segmentation of tumor regions, test with larger multi-center dataset, and

develop the system for multi-class classification (normal vs. benign vs. malignant). And, it is said that the explainable AI approach also can be incorporated to show the regions which most largely affect the decision in the model to trust more in clinicians.

ACKNOWLEDGMENT

The authors would like to thank Mr. N. M. Ramalingeswara Rao, Assistant Professor, Department of ECE, GIET (A), for his continuous guidance, encouragement and valuable suggestions throughout the course of this project.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

- [1] Hashmi, M. F., Husain, A. A., Kotambkar, D. M., & Kumar, K. D. (2025). Hybrid deep learning and handcrafted feature fusion for pneumonia detection in chest X-rays. *Journal of Electrical Systems and Information Technology*, 12(1), 94.
- [2] Shankar, K., & Perumal, E. (2021). A novel hand-crafted with deep learning features based fusion model for COVID-19 diagnosis and classification using chest X-ray images. *Complex & Intelligent Systems*, 7(3), 1277-1293.
- [3] Bassi, P. R., & Attux, R. (2022). COVID-19 detection using chest X-rays: Is lung segmentation important for generalization?. *Research on Biomedical Engineering*, 38(4), 1121-1139.
- [4] Siddiqi, R., & Javaid, S. (2024). Deep learning for pneumonia detection in chest x-ray images: A comprehensive survey. *Journal of imaging*, 10(8), 176.
- [5] Bal, U., Bal, A., Moral, Ö. T., Düzgün, F., & Gürbüz, N. (2024). A deep learning feature extraction-based hybrid approach for detecting pediatric pneumonia in chest X-ray images. *Physical and Engineering Sciences in Medicine*, 47(1), 109-117.
- [6] Ukwuoma, C. C., Qin, Z., Heyat, M. B. B., Akhtar, F., Smahi, A., Jackson, J. K., ... & Nneji, G. U. (2022). Automated lung-related pneumonia and COVID-19 detection based on novel feature extraction framework and vision transformer approaches using chest X-ray images. *Bioengineering*, 9(11), 709.
- [7] Abubakar, H., Al-Turjman, F., Ameen, Z. S., Mubarak, A. S., & Altrjman, C. (2024). A hybridized feature extraction for COVID-19 multi-class classification on computed tomography images. *Heliyon*, 10(5).
- [8] Ou, C. Y., Chen, I. Y., Chang, H. T., Wei, C. Y., Li, D. Y., Chen, Y. K., & Chang, C. Y. (2024). Deep learning-based classification and semantic segmentation of lung tuberculosis lesions in chest X-ray images. *Diagnostics*, 14(9), 952.
- [9] Li, Y., Zhang, L., Yu, H., Wang, J., Wang, S., Liu, J., & Zheng, Q. (2024). A comprehensive segmentation of chest X-ray improves deep learning-based WHO radiologically confirmed pneumonia diagnosis in children. *European Radiology*, 34(5), 3471-3482.
- [10] Win, K. Y., Maneerat, N., Hamamoto, K., & Sreng, S. (2020). Hybrid learning of hand-crafted and deep-activated features using particle

- swarm optimization and optimized support vector machine for tuberculosis screening. *Applied Sciences*, 10(17), 5749.
- [11] Kulkarni, A., Parasnis, G., Balasubramanian, H., Jain, V., Chokshi, A., & Sonkusare, R. (2023). Advancing diagnostic precision: Leveraging machine learning techniques for accurate detection of covid-19, pneumonia, and tuberculosis in chest x-ray images.
- [12] Mabrouk, A., Diaz Redondo, R. P., Dahou, A., Abd Elaziz, M., & Kayed, M. (2022). Pneumonia detection on chest X-ray images using ensemble of deep convolutional neural networks. *Applied Sciences*, 12(13), 6448..
- [13] Wang, C. H., Chang, W., Lee, M. R., Tay, J., Wu, C. Y., Wu, M. C., ... & Huang, C. H. (2024). Deep learning-based diagnosis of pulmonary tuberculosis on chest X-ray in the emergency department: A retrospective study. *Journal of Imaging Informatics in Medicine*, 37(2), 589-600.
- [14] Kabir, M. M., Mridha, M. F., Rahman, A., Hamid, M. A., & Monowar, M. M. (2024). Detection of COVID-19, pneumonia, and tuberculosis from radiographs using AI-driven knowledge distillation. *Heliyon*, 10(5).
- [15] Usman, C., Rehman, S. U., Ali, A., Khan, A. M., & Ahmad, B. (2025). Pneumonia Disease Detection Using Chest X-Rays and Machine Learning. *Algorithms*, 18(2), 82.
- [16] Guo, K., Cheng, J., Li, K., Wang, L., Lv, Y., & Cao, D. (2023). Diagnosis and detection of pneumonia using weak-label based on X-ray images: a multi-center study. *BMC Medical Imaging*, 23(1), 209.
- [17] Billah, M. M., Al Rakib, A., Ahamed, A. S., Chowdhury, S., & Mitro, S. (2024). A Comparative Study on the Detection of Pneumonia in Chest X-Ray Images Utilizing Deep Learning Models. *European Journal of Computer Science and Information Technology*, 12(7), 1-11.
- [18] Ahmed, A. M., Babikir, G. A., & Osman, S. M. (2022). Classification of pneumonia using deep convolutional neural network. *american journal of computer science and technology*, 5(2), 26-33.
- [19] Garstka, J., & Strzelecki, M. (2020, September). Pneumonia detection in X-ray chest images based on convolutional neural networks and data augmentation methods. In *2020 signal processing: algorithms, architectures, arrangements, and applications (SPA)* (pp. 18-23). IEEE.
- [20] Xia, Y. (2024). Enhanced pneumonia detection in chest X-rays based on integrated Denoising Autoencoders and Convolutional neural networks. In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence* (pp. 799-803).
- [21] Vijayakumar, K., Maziz, M. N. H., Ramadasan, S., Kadry, S., & Arunmozhi, S. (2025). Classification of chest radiographs into healthy/pneumonia using Harris-Hawks Algorithm optimized deep-features. *Discover Computing*, 28(1), 115.
- [22] Mwendu, I., Gikunda, K., & Maina, A. (2023). Deep transfer learning for detecting Covid-19, Pneumonia and Tuberculosis using CXR images--A Review. *arXiv preprint arXiv:2303.16754*.