JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 07 Issue 01
2024

Research Article
https://doi.org/10.56979/701/2024

Integrating Data Augmentation with AdaBoost for Effective COVID-19 Pneumonia Classification

Muhammad Suliman[1], Fazal Malik[1*], Muhammad Qasim Khan[1], Irfan Ullah[1], and Abd Ur Rub[2]

[1]Department of Computer Science, Iqra National University Peshawar, Khyber Pakhtunkhwa (KPK), Pakistan.
[2]School of Electronics and Information, Northwestern Polytechnical University, Xi'an Shaanxi, China.
[*]Corresponding Author: Fazal Malik. Email: fazal.malik@inu.edu.pk

Received: December 19, 2023 **Accepted:** May 15, 2023 **Published:** June 01, 2024

Abstract: In the COVID-19 pandemic, the urgent need for precise pneumonia diagnosis has prompted this research to propose a customized AdaBoost algorithm tailored for classifying COVID-19 pneumonia cases. The study follows a structured framework encompassing four primary supervised learning stages: data acquisition, preprocessing, supervised learning with augmented data, and rigorous performance evaluation. Leveraging a comprehensive GitHub dataset and Python in Anaconda Jupyter Notebook, advanced preprocessing techniques are employed to optimize data for machine learning algorithms. The AdaBoost algorithm, enhanced with data augmentation methods, is deployed to bolster model robustness and performance. The model demonstrates superior effectiveness with an average accuracy rate of 84.49%, surpassing existing methodologies. This performance underscores its potential in addressing public health challenges associated with pneumonia diagnosis during the COVID-19 crisis. This research introduces an optimized application of AdaBoost for pneumonia classification, validated across diverse datasets, ensuring reliable disease classification and predictive modeling capabilities to anticipate future trends. These insights are pivotal for guiding public health interventions and optimizing resource allocation, marking significant advancements in diagnostic accuracy and patient care during the pandemic.

Keywords: COVID-19 Pneumonia; AdaBoost; Classification; Chest X-rays; Prediction.

1. Introduction

In global public health, pneumonia is regarded as a formidable challenge due to its inflammatory impact on the lungs, affecting individuals across various age groups but notably posing higher risks among newborns and those aged over 65 [1]. Understanding the diverse types of pneumonia becomes crucial for clinical decision-making and preventive measures. Pneumonia is caused by various infectious sources, including bacteria, viruses, and fungi, with each necessitating distinct diagnostic and therapeutic approaches for optimal patient care. A well-established classification system aids clinicians in categorizing pneumonia based on its source of infection, enabling tailored investigations, the selection of antimicrobial therapy, and the implementation of preventive strategies. Furthermore, complexity is added to the pneumonia landscape by walking pneumonia, characterized by milder symptoms akin to severe cold, and fungal pneumonia, which poses significant risks to immune-compromised individuals [2].

The condition is manifested through a spectrum of symptoms and potential complications, with severity often dictated by the underlying pathogens. While milder symptoms are typically exhibited by viral infections, grave consequences can arise from bacterial pneumonia, particularly with virulent strains, especially among immune-compromised populations such as newborns [3].

Effective management and improved patient outcomes are underscored by accurate and timely diagnosis. Treatment strategies are guided, broader public health interventions are informed by identifying specific pathogens, and the implementation of targeted antibiotic stewardship programs is

facilitated. Similarly, a diverse array of microbial origins, ranging from common viruses to fungal pathogens, is encompassed by community-acquired pneumonia (CAP) [4].

Pneumonia, a respiratory illness we often encounter, unfolds in four distinct stages that mark changes within the lungs. First comes congestion, which appears within the initial 24 hours of infection. This stage is defined by vascular congestion, fluid build-up in the alveoli, and noticeable lung prominence, often accompanied by coughing and difficulty breathing. Next is red Hepatization, where the lungs take on a liver-like color due to blood, neutrophils, and fibrin infiltrating the alveoli? Then comes, gray Hepatization, characterized by a grayish-brown or yellowish hue in the lungs because of vascular changes and fibrin deposition. Finally, resolution marks the healing phase. Here, enzymes break down exudates, allowing macrophages to absorb them or fibroblasts to clear them, ultimately restoring the lungs. Recognizing these stages is crucial for diagnosing and managing pneumonia-related issues effectively [5].

The landscape of the pneumonia types is comprehensively delineated, encompassing hospital-acquired pneumonia (HAP), community-acquired pneumonia (CAP), bacterial pneumonia, viral pneumonia, walking pneumonia, and fungal pneumonia. Each subtype carries unique clinical implications and challenges, demanding nuanced approaches to diagnosis, treatment, and prevention. Hospital-acquired pneumonia (HAP) represents a severe nosocomial infection, compounded by compromised immunity and the proliferation of multidrug-resistant pathogens within healthcare settings [6]. Bacterial pneumonia and viral pneumonia constitute the predominant forms of the disease, each presenting distinctive clinical features and management strategies.

The COVID-19 pandemic, triggered by the novel coronavirus 2019-nCoV, posed a major global health threat. COVID-19, short for "Coronavirus Disease 2019," belongs to the coronavirus family and spreads rapidly among humans. Medical imaging emphasizes the importance of using various phantom representations to improve disease detection accuracy [7].

Fever, cough, and shortness of breath are key symptoms [8]. A proposed approach employs deep transfer learning with models such as AlexNet, GoogleNet, and ResNet18 to classify pneumonia cases as normal or abnormal from X-ray images. COVID-19 presents significant risks, potentially leading to pneumonia, respiratory illnesses, and, in severe cases, organ failure and death [9].

Men and young individuals may be particularly vulnerable, with pneumonia cases often exhibiting faster breath rates compared to their healthy counterparts [10]. Among the diagnostic challenges, X-ray imaging serves as a crucial tool in identifying pneumonia, revealing distinctive patterns of haziness and opacity in lung structures [11]. As the world grapples with the multifaceted impact of COVID-19, understanding its symptoms and diagnostic methodologies remains paramount for effective disease management and prevention efforts.

Min Zhou et al. [12] developed early identification equipment for confirmed novel coronavirus pneumonia (NCP) cases and focal patients, demonstrating accurate results in chest examinations for COVID-19 and other pneumonia cases.

In recent years, medical imaging techniques have played a pivotal role in the diagnosis and classification of respiratory diseases such as COVID-19 and community-acquired pneumonia (CAP). Various studies have explored the application of advanced algorithms and machine learning models to improve disease detection and classification accuracy.

Feng Shi et al. [1] conducted a comprehensive study involving CT examinations on 1658 COVID-19 and 1027 CAP patients. They utilized a disease-size-Aware random forest technique (USAF) for classification, achieving notable performance.

Furthermore, Sung-mok Jung et al. [13] presented a method for predicting the increase in atypical pneumonia cases caused by new pathogens, using routine non-virological information.

The effectiveness of deep learning models in disease classification has been extensively explored. Mohamed Loey et al. [14] investigated deep learning models such as Googlenet and Alexnet, distinguishing COVID-19 and normal cases, utilizing Generative Adversarial Networks (GANs) and deep transfer models.

AdaBoost (Adaptive Boosting), developed in 1997, employs weak classifiers and an exponential loss function to enhance performance, with further research exploring its varied applications and capabilities [15]. AdaBoost excels in classifying output classes from input variables, providing rapid and superior mapping, a notable improvement over techniques like decision trees, regression, and nearest neighbors

[16]. Utilizing AdaBoost, this research enhances pneumonia detection amidst the COVID-19 epidemic, addressing symptoms such as fever and cough, primarily spread through saliva droplets. Other techniques are reviewed for comprehensive analysis [17].

AdaBoost, an Artificial Intelligence (AI) ensemble technique, improves epidemic spread factor classification by selecting relevant attributes objectively. Its utilization in COVID-19 growth rate classification experiments showcases effective containment measure correlation with growth factors [18]. Introducing a COVID-19 Prediction Support System is utilizing AdaBoost, among other Machine Learning techniques, integrated with routine blood testing, demonstrating enhanced predictive accuracy over conventional methods [19]. Utilizing AdaBoost, a machine learning method, to differentiate COVID-19 from comparable respiratory ailments through lung Computed Tomography (CT) scans. Drawing on data from Imam Hussein Hospital, Tehran, it presents potential solutions for swift and accurate COVID-19 diagnosis [20].

Extensive public health challenges are posed by pneumonia, particularly during the COVID-19 pandemic, owing to its varied symptoms and the necessity for accurate diagnosis and classification. Drawbacks in current approaches to pneumonia diagnosis and classification, including the absence of advanced algorithms, limited validation across various datasets, less-than-optimal accuracy, and insufficient adaptation to changing disease patterns, are noted.

To address these challenges, this study proposes the development and implementation of a customized AdaBoost algorithm specifically designed for COVID-19 pneumonia classification. The primary contributions of this research can be outlined in four main phases: Firstly, the utilization of the latest GitHub dataset with Python and Anaconda Jupyter Notebook for data gathering (Data Acquisition). Secondly, employing techniques such as histograms and scatter plots to preprocess and analyze the acquired data, enhancing its suitability for machine learning algorithms (Data Processing and Analysis). Thirdly, the application of a supervised learning AdaBoost algorithm is supplemented by data augmentation methods to improve model robustness and performance (Supervised Learning AdaBoost Application with Data Augmentation). Lastly, the comparison of key metrics from the developed model with those of existing studies to evaluate its accuracy and effectiveness (Performance Evaluation).

These contributions ensure reliable and precise disease classification while also exploring predictive modeling capabilities to anticipate future disease trends. Such insights are crucial for guiding public health interventions and optimizing resource allocation.

The subsequent sections of the research paper will explore into a comprehensive discussion of current state-of-the-art methods (Section 2), introduce the proposed framework and technique (Section 3), present detailed experimental results (Section 4), and conclude with directions for future research (Section 5).

2. Literature Review

The literature review presents a comprehensive overview of studies in respiratory disease diagnosis, particularly focusing on COVID-19 and pneumonia.

An AdaBoost algorithm is introduced in a study, improving model performance by effectively handling rough knowledge and enhancing sensitivity and regression capacity [21].

Utilizing AdaBoost in 2019 Novel Coronavirus (nCoV) analysis, patient history, country, age, and gender are, integrated to enhance death prediction accuracy by over sevenfold, showcasing the potential of machine learning in early outcome estimation. This highlights the value of integrating disease history into predictive models, promising improved patient treatment and healthcare system relief [22].

Using AdaBoost, this study enhances COVID-19 case identification via symptom analysis, aiding in self-assessment and patient triage, while leveraging Non-dominated Sorting Genetic Algorithm II (NSGA-II) for feature selection to boost accuracy [23].

AdaBoost's performance in COVID-19 diagnosis surpasses Random Forest post-feature selection, indicating the potential for heightened disease classification accuracy [24].

Using AdaBoost alongside other techniques like Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), and k-nearest neighbors (KNN), this study on COVID-19 patient blood samples achieves top predictive performance, emphasizing age's significance and strong associations among Lactate Dehydrogenase (LD), C - reactive protein (CRP), and leukocytes [25].

Introducing Vulture Based Adaboost-Feedforward Neural (VbAFN), a novel AdaBoost-based scheme for early COVID-19 severity prediction from chest X-ray images, incorporating preprocessing, feature extraction, and segmentation for accurate data input [26]. Utilizing AdaBoost with decision tree estimators, this study enhances COVID-19 patient severity assessment with a novel parameter tuning process, showcasing competitive accuracy through extensive experimentation on the University of California Irvine (UCI) and COVID-19 datasets [27].

Proposing a Hybrid Disease Detection Principle (HDDP), that combines AdaBoost with Convolutional Neural Network (CNN) for COVID-19 detection from Lung Computed Tomography (CT) scans, offering enhanced accuracy and efficiency in disease identification [28].

Incorporating AdaBoost with CNN enhances COVID-19 detection from Chest X-Ray images, surpassing basic CNN and Residual Network-152 layers (ResNet-152) approaches in feature extraction. AdaBoost integrated with CNN effectively manages imbalanced class datasets, ensuring superior performance [29].

AdaBoost improves pneumonia detection from chest X-ray images, facilitating early diagnosis in regions with energy poverty and pollution-related health concerns, while evaluation metrics affirm its effectiveness alongside other machine learning algorithms [30]. Employing a customized AdaBoost algorithm, predicting COVID-19 severity with clinical markers and vital signs, aiding decision-making in resource-constrained settings, and potentially lowering mortality rates [31].

AdaBoost with pre-trained deep learning models such as ResNet-50 is employed in this study to enhance COVID-19 detection, surpassing traditional Reverse Transcription Polymerase Chain Reaction (RT-PCR) testing. Superior performance is exhibited by AdaBoost-ResNet-50 over other methods, offering a viable solution for effective virus detection during the pandemic [32].

Rehman, A., Naz, S., et al. [33] investigated the pathology of Hemagglutinin Type 1 and Neuraminidase Type 1 (H1N1), CAP, and sepsis, suggesting the potential of group mass disturbance for clinical outcomes in CAP and sepsis patients. Z. Neili, M. Fezari, et al. [34] applied volumetric Extraordinary Learning Machine (ELM) and k-Nearest Neighbor (K-nn) AI for breath sound analysis, demonstrating high accuracy through perceptual mode reduction and feature extraction.

Dimpy Varshni et al. [35] enhanced chest radiograph classification using AlexNet (MAN). Yu-Jie Zhang et al. [36] explored dataset preparation processes, proving effective in engineering and real datasets.

Nour Eldeen M. Khalifa et al. [37] demonstrated the effectiveness of Generative Adversarial Networks (GAN) for pneumonia chest X-ray recognition, outperforming related work in accuracy, recall, and F1 score. Lawrence O. Hall, Rahul Paul et al. [38] proposed an approach for diagnosing COVID-19 with chest X-rays, promising despite dataset limitations. Muhammad Ilyas et al. [39] emphasized procedural frameworks for COVID-19 identification. Ethan D. Evans et al. [40] explored AI's applicability in health status diagnosis. R. Nagamounika et al. [41] focused on pneumonia diagnosis through X-ray images. Muhaza Liebenlito et al. [42] and Shangjie Yao et al. [43] proposed efficient methods for tuberculosis and pneumonia detection, respectively.

Raman Chadha et al. [44] discussed machine learning's role in clinical areas, particularly pneumonia identification. Mariana Chumbita et al. [45] explored AI's potential in clinical decision-making. Lukas Ebner et al. [46] conducted a meta-study analyzing explicit tomography (CT) designs. Khan Maseeh Shuaib et al. [47] developed a web application for pneumonia differentiation. Lin Li et al. [48] established a fully automated system for COVID-19 identification. Yu-Hsuan Liao et al. [49] introduced evolutionary Neural Network (ENN) and SVM expectation models.

Cong Feng et al. [50] focused on pollution-related biomarkers and clinical symptoms for pneumonia prediction. AdaBoost, compared with Bagging, is examined for predicting COVID-19 Intensive Care Unit (ICU) admissions in Saudi Arabia, showing potential for targeted interventions and aiding in COVID-19 healthcare management [51].Utilizing AdaBoost for automating COVID-19, Tuberculosis (TB), and pneumonia classification from chest X-Ray (CXR) images during the pandemic, promising results were observed in enhancing classification accuracy alongside pre-processing techniques [52]. Using data from 1861 COVID-19 patients, AdaBoost demonstrates competitive performance alongside other methods, illustrating AI's promise in personalized medicine for infectious disease management [53].

3. Methodology

The study outlines a systematic approach consisting of three phases: data preprocessing, model training, and evaluation, for COVID-19 pneumonia classification utilizing AdaBoost, as shown in Figure 1 and outlined in Algorithm 01.This section provides an overview of each phase and its significance in the overall classification process.

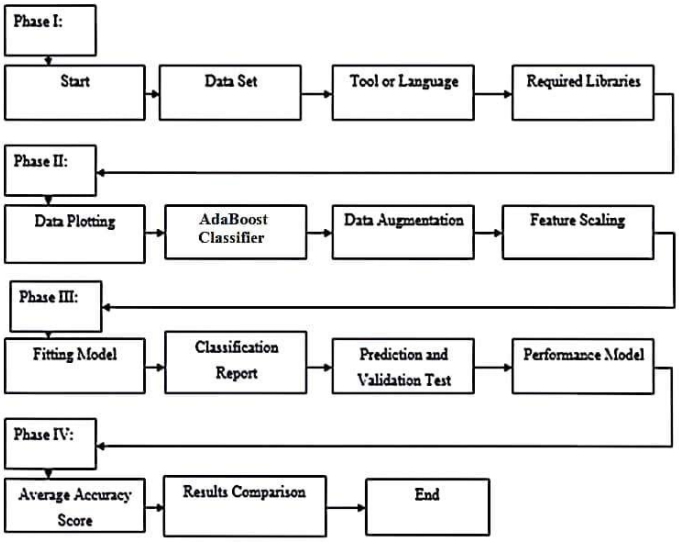

Figure 1. Block diagram of proposed work

Algorithm 01: Pneumonia Diagnosis Study
Input: X-ray (pneumonia) images from GitHub
Output: Model Evaluation Results

Step 1. Data Processing Phase
- 1.1. WHO_dataset ← LoadWHODataSet() // World Health Organization (WHO)
- 1.2. demog_info ← ExtractDemographicInfo (WHO_dataset)
- 1.3. preproc_data ← PreprocessData (WHO_dataset, demog_info)

Step 2. Data Visualization Phase
- 2.1. Demog_Distrib_Plot ← Visualize_DemographicDistribution(demog_info)
- 2.2. Symp_Preval_Plot ← PlotSymptomsPrevalence(preproc_data)

Step 3. Model Building for Feature Extraction
- 3.1. AdaBoost_model ← InitializeAdaBoostModel()
- 3.2. train_models ← TrainModels(AdaBoost_model, COVID19_XRay_images, demog_data)
- 3.3. optim_models ← OptimizeModels(train_models)

Step 4. Performance Evaluation Phase
- 4.1. train_set, test_set ← SplitDataset(preproc_data)
- 4.2. conf_matrix ← EvaluateModels(optim_models, test_set)
- 4.3. metrics ← CalculateMetrics(conf_matrix, accuracy, sensitivity, specificity)

Step 5. Comparative Analysis of Results
- 5.1. accu_compar ← CompareAccuracy(base_study, existing_research)
- 5.2. improv_areas ← IdentifyAreasOfSubstantialImprovement()
- 5.3. implic_and_applic ← DiscussImplicationsAndApplications()

Step 6. End

The methodology section outlines a comprehensive approach using AdaBoost, tailored for accurate classification of COVID-19 pneumonia cases. The contribution of the proposed study lies in several key aspects:

Data Preprocessing and Preparation: We emphasize the critical role of data cleaning, normalization, and preprocessing to ensure the suitability of chest X-ray images for the AdaBoost model. This step enhances the quality and compatibility of the input data, which is crucial for effective model training.

Model Training and Optimization: The methodology details our approach to training the AdaBoost algorithm on preprocessed data. We contribute by discussing our systematic parameter optimization and iterative improvement techniques. These efforts are aimed at enhancing the model's ability to learn intricate patterns indicative of COVID-19 pneumonia, thereby improving classification accuracy.

Model Evaluation and Performance Metrics: We highlight the significance of rigorous model evaluation using separate datasets. This includes assessing key performance metrics such as precision, recall, and F1-score. Our contribution lies in demonstrating how these metrics validate the effectiveness and reliability of our classification results.

Experimental Validation and Comparative Analysis: The methodology presents our validation strategy through experimental results and comparative analysis with existing methodologies. This step underscores our contribution by showcasing the robustness and accuracy of the AdaBoost model in accurately classifying COVID-19 pneumonia cases.

The proposed methodology makes a significant contribution by providing a structured framework that enhances the overall accuracy and reliability of COVID-19 pneumonia classification using AdaBoost. Through rigorous data preprocessing, optimized model training, thorough evaluation, and validation, our research aims to advance the field by offering a robust methodology for effective disease classification in clinical settings.

3.1. Phase I: Data Gathering and Tool Implementation

3.1.1. Data Gathering and Preprocessing

The study begins by acquiring the latest dataset from GitHub, comprising X-ray images of COVID-19 and other pneumonia cases. Over 1,000 chest X-ray images are included, providing a diverse sample for analysis [54]. Data preprocessing is crucial to ensuring the quality and compatibility of the dataset with the AdaBoost algorithm. Techniques such as resizing, normalization, and augmentation may be applied to enhance the dataset's suitability for training. Previous research studies have also utilized this dataset for classification purposes.

3.1.2. Tool Implementation: Python and Anaconda Jupyter Notebook

Python, in conjunction with the Anaconda Jupyter Notebook, serves as the primary tool for this investigation. Python's versatility in data mining, artificial intelligence, and machine learning makes it well-suited for developing and implementing algorithms like AdaBoost. Anaconda Jupyter Notebook provides an intuitive interface for coding and visualization, enhancing efficiency and ease of implementation. The compatibility between Anaconda and Python streamlines the development process, facilitating seamless integration of AdaBoost into the workflow.

3.1.3. Importing Required Libraries

Specialized libraries play a crucial role in implementing AdaBoost and conducting various functions within the study. Keras, a powerful deep learning library, is employed for constructing neural network models, complementing AdaBoost's ensemble learning approach. Additionally, libraries such as NumPy, Matplotlib, and Seaborn are utilized for data manipulation, numerical analysis, and data visualization, respectively. These libraries enhance the efficiency, accuracy, and interpretability of the results obtained from the AdaBoost algorithm.

The AdaBoost algorithm is implemented using Python and relevant libraries within the Anaconda Jupyter Notebook environment. It involves iteratively training weak learners on different subsets of the dataset and adjusting their weights based on their performance. The final classifier is a weighted combination of these weak learners, resulting in a robust and accurate model for pneumonia diagnosis.

3.2. Phase II: Data Processing and Model Analysis

3.2.1. Data Plotting

To comprehend the COVID-19 pneumonia dataset thoroughly, employing effective data plotting techniques is imperative. Various visualization methods, such as histograms, boxplots, scatter plots, and dimensionality reduction techniques like Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE), are utilized. These methods aim to visualize feature distribution, uncover relationships, and discern intricate patterns within the data. Additionally, grids of example

images are employed to emphasize characteristic features, facilitating a comprehensive understanding of the dataset [55].

The graphical representation of data is paramount in the analysis process. To achieve this, we utilized OpenCV, Matplotlib, and Seaborn libraries, which offer robust capabilities for data visualization. These libraries enable the creation of visually appealing and informative plots. As illustrated in Figure 2, an X-ray image is displayed for normal and abnormal classifications, providing a visual representation of the dataset. Our dataset comprises X-ray images sourced from GitHub, forming the cornerstone of our research endeavors.

A pivotal role is played by the graphical representation of data in the analysis process. To accomplish this task, OpenCV, Matplotlib, and Seaborn libraries, renowned for their robust capabilities in data visualization, were leveraged. These libraries enable visually appealing and informative plots to be created. For instance, as illustrated in Figure 2, an X-ray image is presented for normal and abnormal classification, offering a visual depiction of the dataset. It's noteworthy that the dataset comprises X-ray images sourced from GitHub, which serve as the bedrock of our research endeavors.

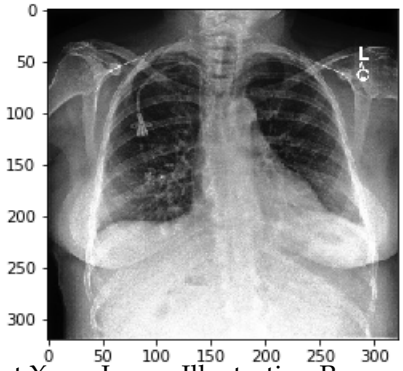

Figure 2. Chest X-ray Image Illustrating Pneumonia Condition

After data visualization, the subsequent step involves the construction of an AdaBoost model for data analysis. AdaBoost is recognized as a formidable machine learning algorithm known for its efficiency and performance in handling structured data. Unlike convolutional neural networks (CNNs), which are predominantly excelled in image processing tasks, AdaBoost is distinguished in tabular data analysis. Leveraging its ensemble learning approach and gradient boosting framework, AdaBoost facilitates the development of predictive models adept at capturing intricate relationships within the data [56].

3.2.2. Supervised Learning with AdaBoost

Supervised learning with AdaBoost involves training a model on a well-defined dataset containing input features and corresponding output labels. These features and labels play a crucial role in various applications, such as COVID-19 pneumonia classification using chest X-ray images. In the context of COVID-19 pneumonia classification, the training dataset consists of input features extracted from chest X-ray images. These features typically represent pixel values, serving as descriptors for the image's characteristics. The output labels indicate whether the image depicts COVID-19 pneumonia or not.

Using the supervised learning approach, the AdaBoost algorithm learns the underlying patterns and relationships between input features and output labels. This learning process enables AdaBoost to accurately classify images based on the learned patterns, making it effective for medical imaging analysis and similar tasks. AdaBoost's ability to learn from labeled data and identify relevant patterns makes it particularly valuable in medical imaging analysis. By leveraging AdaBoost, healthcare professionals can improve the accuracy and efficiency of pneumonia diagnosis, especially in scenarios like the COVID-19 pneumonia classification [57].

- *AdaBoost Classifier*

AdaBoost, an ensemble learning algorithm known as Adaptive Boosting, is widely used in machine learning for classification tasks, particularly binary classification problems. Data is categorized into two classes, and AdaBoost combines multiple weak classifiers to create a strong classifier, improving model accuracy overall. The COVID-19 pneumonia classification, crucial for timely patient diagnosis and treatment, involves categorizing chest X-ray images into COVID-19 pneumonia and non-COVID-19 pneumonia classes. AdaBoost operates by iteratively training weak classifiers on various subsets of the

training data, learning from misclassified data points from previous classifiers. It assigns higher weights to misclassified points, leading to a final model that combines predictions from all weak classifiers. In the context of COVID-19 pneumonia classification with chest X-ray images, AdaBoost effectively identifies subtle patterns and features indicative of COVID-19 pneumonia, enhancing sensitivity and specificity for more reliable classification results. The training process divides chest X-ray image datasets into training and testing sets, with AdaBoost iteratively training weak classifiers on the training data and adjusting their weights based on performance. While AdaBoost excels at handling complex datasets with high-dimensional features, such as those in medical image analysis, it may be sensitive to noisy data and outliers, potentially affecting performance. Additionally, the computational complexity of training multiple weak classifiers could be a limiting factor in large-scale applications.

3.2.3. Data Amplification

Data augmentation serves as a crucial technique within machine learning, notably in applications such as COVID-19 pneumonia classification via chest X-ray images. Specifically, in the context of AdaBoost models, data augmentation plays a pivotal role in enriching the training dataset through the generation of synthetic data from existing samples. Various augmentation techniques are employed, including rotation, flipping, scaling, cropping, adjustment of brightness or contrast, and introduction of noise to chest X-ray images. These diverse transformations introduce variability into the dataset, thereby enhancing the model's capability to generalize and make accurate predictions on unseen images. Notably, data augmentation serves to bolster model performance and robustness while also serving as a potent tool to mitigate overfitting, as it introduces variability into the training data without necessitating additional real-world samples.

3.2.4. Feature Scaling for Optimal AdaBoost Performance

Feature scaling is a crucial preprocessing step essential for optimizing the performance of AdaBoost, especially in the context of the COVID-19 pneumonia classification. This section introduces the concept of feature scaling and its importance in machine learning algorithms like AdaBoost.

Unlike deep learning models that handle feature scaling automatically, AdaBoost requires manual preprocessing to ensure optimal performance. This subsection discusses why feature scaling is particularly important for AdaBoost, emphasizing its role in preventing certain features from dominating others and ensuring uniformity in scale across all features. In the case of COVID-19 pneumonia classification using chest X-ray images, specific feature scaling techniques are required. This section explores the various techniques used to scale pixel intensity values to a smaller range, such as 0 to 1 or -1 to 1. Additionally, it discusses the scaling of demographic or clinical features like age, body temperature, and blood oxygen levels to make them comparable.

Feature scaling significantly influences the performance and robustness of AdaBoost in accurately classifying COVID-19 pneumonia cases. This subsection delves into how proper feature scaling aids the AdaBoost algorithm in effectively learning from diverse features, leading to improved classification accuracy and generalization ability.

3.3. Phase III: Analyzing Classification and Performance Models

3.3.1. Model Training Using the AdaBoost Technique

In this phase, we delve into the process of fitting a model using the AdaBoost technique for COVID-19 pneumonia classification.

Data preparation starts with assembling a dataset containing labeled chest X-ray images, distinguishing between COVID-19 pneumonia and non-COVID-19 pneumonia cases. This dataset is then partitioned into training and testing subsets to facilitate model evaluation. Subsequently, feature extraction is performed to extract relevant characteristics from the chest X-ray images. These features serve as the basis for training the AdaBoost model and play a crucial role in accurately classifying pneumonia cases. The AdaBoost model is trained using an ensemble of decision trees. Each decision tree independently predicts the output class based on the extracted features from the chest X-ray images. The ensemble approach enhances the model's predictive accuracy and robustness.

Once the model is trained, it undergoes evaluation using various performance metrics such as accuracy, precision, recall, and F1-score. Parameter tuning may be conducted to optimize the model's performance further.

- *Data Preparation*

In this initial phase, a dataset comprising labeled chest X-ray images is assembled, distinguishing between COVID-19 pneumonia and non-COVID-19 pneumonia cases. This dataset is then divided into training and testing subsets to facilitate model evaluation.

● *Feature Extraction*

Following data preparation, feature extraction is executed to derive relevant characteristics from the chest X-ray images. These extracted features form the foundation for training the AdaBoost model and are pivotal in accurately classifying pneumonia cases.

● *Model Training with AdaBoost*

The AdaBoost technique is employed to train the model, utilizing an ensemble of decision trees. Each decision tree operates independently, predicting the output class based on the extracted features from the chest X-ray images. This ensemble approach enhances the model's predictive accuracy and robustness.

● *Model Evaluation*

Once the model is trained, it undergoes evaluation using various performance metrics, including accuracy, precision, recall, and F1-score. This evaluation stage provides insights into the model's effectiveness and aids in identifying areas for improvement.

● *Parameter Tuning*

Parameter tuning may be performed to optimize the model's performance further. This involves adjusting the parameters of the AdaBoost algorithm to achieve the best possible results, ensuring the model is well-suited for the COVID-19 pneumonia classification.

3.3.2. AdaBoost-Based Classification and Prediction Outcomes

In this subsection, we delve into the assessment of the AdaBoost model's performance in classifying and predicting COVID-19 pneumonia cases. Metrics such as accuracy, precision, recall, and F1-score are meticulously examined to provide a holistic understanding of the model's effectiveness. These metrics serve as benchmarks to gauge the model's ability to differentiate between normal and abnormal pneumonia cases accurately.

● *Real-Time Validation Testing*

The AdaBoost model's performance is further scrutinized through real-time validation tests conducted on chest X-ray images. These tests are essential to validate the model's efficacy and reliability in clinical settings, ensuring its practical applicability. By subjecting the model to real-world scenarios, we assess its ability to accurately classify pneumonia cases and its potential for clinical integration.

● *Interpretation of Classification Results*

This subsection delves into a detailed interpretation of the AdaBoost model's classification results. We analyze the model's decision-making process and identify the key features contributing to accurate predictions. By understanding the underlying mechanisms of the model, we gain valuable insights into its predictive capabilities and the factors influencing its classification decisions.

● *Visualization of Classification and Prediction Results*

Figure 3 presents a visual representation of the classification and prediction outcomes derived from the AdaBoost model.

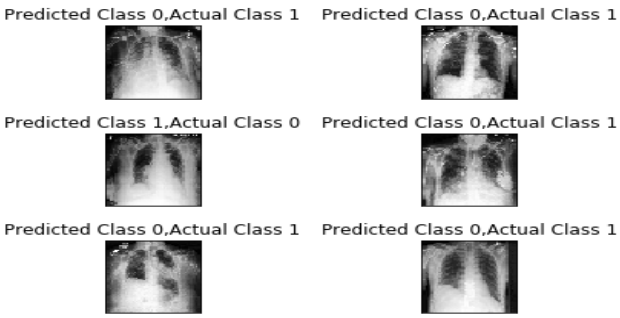

Figure 3. Classification and Prediction results derived from the proposed model.

This visualization provides a clear depiction of the model's effectiveness in accurately classifying COVID-19 pneumonia cases based on chest X-ray images. Through visual analysis, we gain insights into the model's performance and its ability to differentiate between normal and abnormal pneumonia cases.

3.3.3. Assessing Model Performance through Confusion Matrix Analysis

The performance of classification algorithms is assessed using a confusion matrix, a fundamental tool. It provides a detailed examination of accuracy and error patterns by comparing the model's predictions to actual labels. Constructed by organizing model predictions and actual labels into a matrix format, it comprises four main components: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), with each cell representing the count of instances classified accordingly. The interpretation of the confusion matrix is crucial for evaluating model performance, with true positives and true negatives indicating correct classifications and false positives and false negatives indicating misclassifications. Metrics such as accuracy, precision, recall, and F1-score can be derived from the confusion matrix to quantify the model's performance. The effectiveness of the classification algorithm is assessed by analyzing the distribution of predictions across the confusion matrix, with patterns such as high false positive or false negative rates indicating areas for model refinement or feature engineering. The insights gained from confusion matrix analysis can guide model optimization strategies, enhancing the overall performance and reliability of the classification algorithm.

- *Performance Evaluation Using the Confusion Matrix*

In the realm of machine learning, the confusion matrix stands as a fundamental tool for evaluating model performance. It offers a detailed comparison between actual and predicted labels, providing insights into classification accuracy and error types. Through the identification of specific errors, like false positives and false negatives, the confusion matrix guides strategies for refining and optimizing models, ultimately improving diagnostic accuracy and clinical interpretation [59].

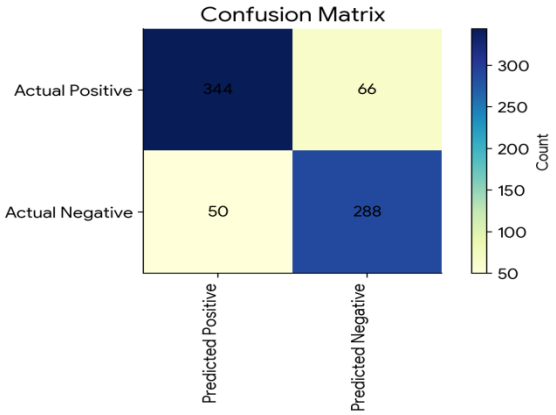

Figure 4. Confusion Matrix Generated by AdaBoost

The confusion matrix, as depicted in Figure 4, is generated by our AdaBoost model. It encapsulates the total count of actual and predicted labels, including true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These components are deemed pivotal in computing various performance metrics and delineating the model's predictive accuracy [60].

- *Confusion Matrix Components*

True Positive (TP): Accurate positive predictions, totaling 344 instances.

True Negative (TN): Accurate predictions of negative cases, amounting to 288 instances.

False Positive (FP): Erroneous positive predictions, occurring 50 times.

False Negative (FN): Inaccurate negative predictions, comprising 66 cases.

By utilizing these matrix elements, precise calculations of our model's precision, recall, and overall accuracy are facilitated. This comprehensive analysis of the confusion matrix provides a thorough evaluation of our proposed AdaBoost model's effectiveness in classification tasks, guiding further optimizations for enhanced performance [61].

4. Results and Discussion

4.1. Phase IV: Assessing Performance and Conducting Comparative Analysis

In this phase, the research aims to thoroughly evaluate the performance of the proposed AdaBoost model and conduct a comparative analysis with existing studies. During this stage, key performance metrics are calculated to assess the accuracy of the AdaBoost model. Metrics such as average accuracy scores, recovery rates, and F1 scores are computed. These metrics play a pivotal role in determining the

precision and recall capabilities of the model, offering crucial insights into its overall effectiveness in classification tasks.

4.2. Average Accuracy Score to assess the overall performance of the model

The average accuracy score, essential in evaluating the AdaBoost model's performance, is calculated from a thorough examination of precision accuracy, recall, and F1 scores, crucial in assessing the model's effectiveness. Precision accuracy denotes the proportion of relevant outcomes accurately classified by the model, while recall signifies the percentage of relevant outcomes correctly identified. The F1 score, amalgamating average precision and recall, provides a comprehensive measure of the model's efficacy. These computations stem from the confusion matrix, offering a comprehensive overview of the model's classification performance across diverse categories.

4.3. Evaluation Metrics for COVID-19 Pneumonia Classification Using AdaBoost

In the evaluation of our proposed AdaBoost model for COVID-19 pneumonia classification, assessing its effectiveness in medical diagnosis is crucial. Key performance metrics, including accuracy, precision, recall, and F1-score, are instrumental in this evaluation process.

Accuracy, representing the overall proportion of correctly classified cases, is calculated using a formula that considers true positive (TP) and true negative (TN) cases.

$$Accuracy(AC) = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision, on the other hand, measures the proportion of true positive predictions for COVID-19 pneumonia.

$$Precision\ (PR) = \frac{TP}{TP+FP} \quad (2)$$

Recall quantifies the proportion of true positive predictions within all actual COVID-19 cases.

$$Recall\ (RE) = \frac{TP}{TP+FN} \quad (3)$$

The F1-score, a harmonic mean of precision and recall, provides a balanced measure of both metrics.

$$F1\ Score = 2 \times \frac{PR \times RE}{PR+RE} \quad (4)$$

4.3.1. Model Evaluation Results

Upon calculating these metrics using our AdaBoost model's performance, derived from the confusion matrix, we found precision, recall, and F1-score to be 87.31%, 83.90%, and 85.57%, respectively. The accuracy of the model reached 84.49%.

4.3.2. Interpretation of Results

These results underscore the effectiveness of the AdaBoost model, with an accuracy of 84.49% indicating its ability to correctly classify cases in the test set. The balanced F1-score of 85.57% showcases robust performance in identifying positive cases while minimizing false positives. However, the recall rate of 83.90% suggests a potential oversight of true positive cases, indicating the need for further refinements to enhance sensitivity and improve positive case detection, as shown in Table 1.

Table1. Performance Measures of the Model: Accuracy, Precision, Recall, and F1 Score (%)

Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
87.31	83.90	85.57	84.49

4.4. Performance Analysis of the Proposed Model Utilizing the AdaBoost Technique

The performance analysis of our proposed model, which utilizes the AdaBoost technique, is considered a critical aspect of evaluating its effectiveness in classifying COVID-19 pneumonia cases. Promising performance has been demonstrated by the model, with an impressive overall accuracy of 84.49% achieved. This accuracy metric, as depicted in Table 1, indicates the proportion of correctly classified cases out of the total cases evaluated.

4.4.1. Key Performance Metrics: Precision, Recall, and F1-Score

The provided analysis focuses on key performance metrics, namely precision, recall, and F1-score, which are crucial in evaluating the effectiveness of a classification model.

Precision measures the proportion of correctly identified positive cases among all predicted positive cases. In this case, a precision of 87.31% indicates that a vast majority of predicted positive cases are indeed true positives. This high precision suggests that the model is adept at minimizing false positives, reducing the risk of misdiagnoses and unnecessary interventions.

On the other hand, recall measures the proportion of correctly identified positive cases among all actual positive cases. A recall of 83.90% suggests that while the majority of positive cases are detected,

there is a possibility that a small portion may go undetected. This highlights the importance of investigating potential reasons for this discrepancy, such as data characteristics or model architecture, to improve sensitivity and ensure that all positive cases are identified.

The F1-score provides a balanced measure of precision and recall, combining both metrics into a single value. With an F1-score of 85.57%, the model demonstrates effectiveness in accurately classifying both normal and abnormal cases while maintaining a high level of precision. This balanced metric underscores the model's capability to effectively capture positive cases without compromising precision. Overall, the analysis suggests that the model performs well in accurately classifying cases, but there is room for improvement in terms of recall to ensure comprehensive detection of positive cases.

4.4.2. Distinguishing Normal vs. Abnormal Classification

The text emphasizes that the model categorizes images as either normal or abnormal without explicitly differentiating between "normal" and "pneumonia" cases. This strategy ensures that the model's primary focus is on identifying any abnormalities present in the images, which may include pneumonia-related indicators, within the broader "abnormal" category.

The AdaBoost model is highlighted as a significant advancement in COVID-19 pneumonia classification, demonstrating high accuracy and balanced performance metrics. While acknowledging room for improvement, particularly in enhancing recall rates, the overall performance indicates considerable potential for improving diagnostic accuracy and patient care.

The text underscores the ongoing research and development efforts in AI-powered medical diagnosis, indicating promising prospects for the future. It suggests that continued advancements in this domain could lead to further enhancements in diagnostic accuracy and patient care, reflecting the evolving landscape of AI in healthcare.

4.5. Comparative Analysis with Existing Research

Our investigation into COVID-19 pneumonia recognition involves a thorough comparative analysis with methodologies presented by Khalifa et al. [9] and Ieracitano et al. [58].

Khalifa et al. utilized deep transfer learning methods employing models like Alexnet, Googlenet, and Restnet18 to recognize pneumonia in X-ray images. Their approach is primarily aimed at binary classification, distinguishing normal from abnormal pneumonia cases. Despite achieving a commendable 78.70% accuracy, their focus was primarily on binary classification.

Ieracitano et al. introduced the CovNNet model, integrating fuzzy logic with deep learning techniques to differentiate COVID-19 pneumonia from other types of interstitial pneumonia. CovNNet achieved 81% accuracy using image features and fuzzy edge data from various datasets. Their primary objective was to accurately classify COVID-19 pneumonia cases among other interstitial pneumonia types.

Our proposed AdaBoost model outperforms existing methodologies, achieving an impressive accuracy of 84.49%. This surpassed the models presented by Khalifa et al. and Ieracitano et al., as evidenced in Table 2.

Precision measures the proportion of true positive predictions among all positive predictions made by the model. In this context, precision metrics revealed a reduction in false positives, meaning that the model made fewer incorrect predictions of pneumonia cases when they were not actually present. This reduction in false positives contributed to an impressive F1-score of 85.57%. The F1-score considers both false positives and false negatives, making it particularly useful in situations where there is an imbalance between the classes being predicted. Therefore, the notable reduction in false positives and the resulting high F1-score of 85.57% underscore the model's exceptional ability to accurately categorize pneumonia cases. This is especially significant within the context of COVID-19, where accurate and reliable diagnosis is crucial for effective patient management and public health interventions.

The results emphasize the significance of the proposed approach in advancing pneumonia classification methodologies. It highlights the observed superiority in performance, particularly in terms of accuracy and precision, as evidenced by the results presented in Table 2. This superiority underscores the valuable contribution of the proposed approach to the field, reaffirming its potential to enhance existing methodologies for pneumonia classification.

Table 2. Comparative Analysis of Proposed and Existing Research Models for X-ray Image Classification

Authors	Data Set	Algorithm	Accuracy
NourEldeen M. Khalifa et al. [9]	X-ray Images	Alexnet, Googlenet, and Restnet18	78.70%

Ieracitano, Cosimo, et al. [58]	X-ray Images	Fuzzy-CovNNet	81.00%
Our propose research work	X-ray Images	AdaBoost Classifier	84.49%

Furthermore, the discussion emphasizes the specific focus of the proposed approach on distinguishing between normal and abnormal pneumonia cases, with a particular emphasis on COVID-19. This specialization contrasts with the broader classification approaches employed by previous studies. Despite a slightly lower recall compared to one of the existing models, the nuanced strategy of the proposed approach has the potential to improve precision and overall performance by facilitating more precise learning.

Overall, the proposed study signifies significant advancements in COVID-19 pneumonia classification, addressing critical gaps in existing research and ultimately leading to improved healthcare outcomes, especially for vulnerable demographics. The systematic framework employed ensures reproducibility and sets a solid foundation for future investigations in this important area of study. Additionally, the technical excellence demonstrated through the meticulous design of the AdaBoost model architecture and the strategic utilization of programming tools further enhances the reliability and applicability of the proposed approach for future research endeavors in pneumonia classification, particularly within the context of COVID-19.

5. Conclusion and Future Directions

This research tackles the significant challenge of COVID-19 pneumonia classification through the development and application of an optimized AdaBoost algorithm. The study employs a comprehensive three-phase methodology, covering data acquisition, processing, analysis, and model evaluation, effectively demonstrating the efficacy of AdaBoost in distinguishing between normal and abnormal pneumonia cases.

Through a detailed analysis of key performance metrics, including average accuracy scores, recovery rates, and F1 scores, the model's precision and recall capabilities are highlighted. With an accuracy rate of 84.49%, the model demonstrates proficiency in accurately classifying relevant cases, surpassing the performance of existing methodologies. The success of the model can be attributed to careful consideration of technical intricacies, such as architecture design and training specifics, which enhance the model's reliability and reproducibility.

The systematic framework presented in this study not only advances the COVID-19 pneumonia classification but also lays the groundwork for future investigations, particularly concerning vulnerable populations. Acting as a cornerstone, this research contributes to enhancing diagnostic accuracy and patient care in the COVID-19 pneumonia classification.

Future research endeavors should prioritize enhancing sensitivity, improving recall rates, and validating datasets on a larger scale to further enhance model performance. Efforts should also focus on broadening the model's applicability across diverse demographics and healthcare settings by integrating real-time data for continuous optimization and adaptation.

References

1. Shi, Feng, Liming Xia, Fei Shan, Bin Song, Dijia Wu, Ying Wei, Huan Yuan et al. "Large-scale screening to distinguish between COVID-19 and community-acquired pneumonia using infection size-aware classification." Physics in medicine & Biology 66, no. 6 (2021): 065031.
2. Yu, Xinxin, Shuai Zhang, Jingxu Xu, Yong Huang, Hao Luo, Chencui Huang, Pei Nie et al. "Nomogram using CT radiomics features for differentiation of pneumonia-type invasive mucinous adenocarcinoma and pneumonia: multicenter development and external validation study." American Journal of Roentgenology 220, no. 2 (2023): 224-234.
3. Wang, Kefan, and Yanping Shi. "Discussion on the syndrome differentiation treatment of pneumonia and cough in children." MEDS Chinese Medicine 5, no. 8 (2023): 121-126.
4. File Jr, Thomas M., and Julio A. Ramirez. "Community-acquired pneumonia." New England Journal of Medicine 389, no. 7 (2023): 632-641.
5. Ma, Hai-Ran, Bi-Ying Deng, Jing Liu, Peng Jiang, Yan-Lei Xu, Xiu-Yun Song, Jie Li et al. "Lung ultrasound to diagnose infectious pneumonia of newborns: a prospective multicenter study." Pediatric Pulmonology 58, no. 1 (2023): 122-129.
6. Alshammari, M. K., Alotaibi, M. A., AlOtaibi, A. S., Alosaime, H. T., Aljuaid, M. A., Alshehri, B. M., ... & Alotaibi, A. A. . Prevalence and Etiology of Community- and Hospital-Acquired Pneumonia in Saudi Arabia and Their Antimicrobial Susceptibility Patterns: A Systematic Review. Medicina, 2023, 59(4), 760.
7. Aditya Kakde, Nitin Arora, Durgansh Sharma, & Subhashchander Sharma. Multi-spectral classification and recognition of breast cancer and pneumonia. Polish Journal of Medical Physics and Engineering, 2020, 26(1), doi:10.2478/pjmpe-2020-0001.
8. Chen, Kuan-Fu, Tsai-Wei Feng, Chin-Chieh Wu, Ismaeel Yunusa, Su-Hsun Liu, Chun-Fu Yeh, Shih-Tsung Han et al. "Diagnostic accuracy of clinical signs and symptoms of COVID-19: A systematic review and meta-analysis to investigate the different estimates in a different stage of the pandemic outbreak." Journal of global health 13 (2023).
9. NE, M. Khalifa, F. Smarandache, and M. Loey. "COVID-19 Chest X-Ray Images Diagnosis: A Neutrosophic and Deep Transfer Learning Approach." (2020).
10. Xu, Xiaowei, Xiangao Jiang, Chunlian Ma, Peng Du, Xukun Li, Shuangzhi Lv, Liang Yu et al. "Deep Learning System to Screen Coronavirus Disease 2019 Pneumonia (preprint)." (2020).
11. Moorthy, Chinnadurai Ganesa, and Ganesamoorthy Udhaya Sankar. "Analysis on Electromagnetic Waves of CT Scanners and MRI Scanners for Applications." World Scientific News 188 (2024): 1-14.
12. Zhou, Min, Dexiang Yang, Yong Chen, Yanping Xu, Jin-Fu Xu, Zhijun Jie, Weiwu Yao et al. "Deep learning for differentiating novel coronavirus pneumonia and influenza pneumonia." Annals of translational medicine 9, no. 2 (2021).
13. Jung, Sung-mok, Ryo Kinoshita, Robin N. Thompson, Natalie M. Linton, Yichi Yang, Andrei R. Akhmetzhanov, and Hiroshi Nishiura. "Epidemiological identification of a novel pathogen in Real Time: analysis of the atypical pneumonia outbreak in Wuhan, China, 2019–2020." Journal of clinical medicine 9, no. 3 (2020): 637.
14. Loey, Mohamed, Florentin Smarandache, and Nour Eldeen M. Khalifa. "Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning." Symmetry 12, no. 4 (2020): 651.
15. Souza, Erico N., and Stan Matwin. "Improvements to AdaBoost dynamic." Advances in Artificial Intelligence (2012):293–298..
16. Shhadat, Ihab, Amena Hayajneh, and Ziad A. Al-Sharif. "The use of machine learning techniques to advance the detection and classification of unknown malware." Procedia Computer Science 170 (2020): 917-922.
17. Lai, Chih-Cheng, Tzu-Ping Shih, Wen-Chien Ko, Hung-Jen Tang, and Po-Ren Hsueh. "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges." International journal of antimicrobial agents 55, no. 3 (2020): 105924.
18. Singer, Gonen, and Matan Marudi. "Ordinal decision-tree-based ensemble approaches: The case of controlling the daily local growth rate of the COVID-19 epidemic." Entropy 22, no. 8 (2020): 871.
19. El Gannour, Oussama, Soufiane Hamida, Bouchaib Cherradi, and Abdelhadi Raihani. "Enhancing early detection of COVID-19 with machine learning and blood test results." Multimedia Tools and Applications (2024): 1-31.
20. Rad, Houman Bahrami, Mehrdad Kargari, Meysam Alavi, and Amir Behnam Kharazmy. "Hybrid Approach for Automated Discrimination between COVID-19 and Similar Respiratory Diseases." In 2024 13th Iranian/3rd International Machine Vision and Image Processing Conference (MVIP), pp. 1-7. IEEE, 2024.

21. Bai, Yun, Jingjing Xie, Dongqiang Wang, Wanjuan Zhang, and Chuan Li. "A manufacturing quality prediction model based on AdaBoost-LSTM with rough knowledge." Computers & Industrial Engineering 155 (2021): 107227.
22. Khan, Koffka, and Emilie Ramsahai. "Maintaining proper health records improves machine learning predictions for novel 2019-nCoV." BMC Medical Informatics and Decision Making 21, no. 1 (2021): 172.
23. Soui, Makram, Nesrine Mansouri, Raed Alhamad, Marouane Kessentini, and Khaled Ghedira. "NSGA-II as feature selection technique and AdaBoost classifier for COVID-19 prediction using patient's symptoms." Nonlinear dynamics 106, no. 2 (2021): 1453-1475.
24. Aremu, T. B., Akinyemi Moruff OYELAKIN, M. B. Akanbi, H. F. Sulaimon, and K. M. Odeyale. "Machine Learning Techniques for Early Classification of COVID-19 Disease in Patients." Journal of Information Communication Technologies and Robotic Applications 13, no. 1 (2022): 8-14.
25. Mazloumi, Rahil, Seyed Reza Abazari, Farnaz Nafarieh, Amir Aghsami, and Fariborz Jolai. "Statistical analysis of blood characteristics of COVID-19 patients and their survival or death prediction using machine learning algorithms." Neural Computing and Applications 34, no. 17 (2022): 14729-14743.
26. Mary, S. Roselin, Vinit Kumar, KJ Prasanna Venkatesan, R. Satish Kumar, Naga Padmaja Jagini, and Amedapu Srinivas. "Vulture-based AdaBoost-feedforward neural frame work for COVID-19 prediction and severity analysis system." Interdisciplinary Sciences: Computational Life Sciences 14, no. 2 (2022): 582-595.
27. Sevinç, Ender. "An empowered AdaBoost algorithm implementation: A COVID-19 dataset study." Computers & Industrial Engineering 165 (2022): 107912.
28. Prabha, B., Sandeep Kaur, Jaspreet Singh, Praful Nandankar, Sanjiv Kumar Jain, and Harikumar Pallathadka. "Intelligent predictions of Covid disease based on lung CT images using machine learning strategy." Materials Today: Proceedings 80 (2023): 3744-3750.
29. Darıcı, Muazzez Buket. "Performance analysis of combination of cnn-based models with adaboost algorithm to diagnose covid-19 disease." Politeknik Dergisi 26, no. 1 (2023): 179-190.
30. Godbin, A. Beena, and S. Graceline Jasmine. "A Machine Learning Based Approach for Diagnosing Pneumonia with Boosting Techniques." In Machine Intelligence for Smart Applications: Opportunities and Risks, pp. 145-160. Cham: Springer Nature Switzerland, 2023.
31. Khanna, Varada Vivek, Krishnaraj Chadaga, Niranjana Sampathila, Srikanth Prabhu, and Rajagopala Chadaga. "A machine learning and explainable artificial intelligence triage-prediction system for COVID-19." Decision Analytics Journal (2023): 100246.
32. Mubarak, Auwalu Saleh, Sertan Serte, Zubaida Sa'id Ameen, Chadi Altrjman, and Fadi Al-Turjman. "COVID-19 Detection Based on Deep Learning Feature Extraction and AdaBoost Ensemble Classifier." In Artificial Intelligence of Health-Enabled Spaces, pp. 47-61. CRC Press, 2023.
33. Rehman, Arshia, Saeeda Naz, Ahmed Khan, Ahmad Zaib, and Imran Razzak. "Improving coronavirus (COVID-19) diagnosis using deep transfer learning." In Proceedings of International Conference on Information Technology and Applications: ICITA 2021, pp. 23-37. Singapore: Springer Nature Singapore, 2022.
34. Neili, Z., M. Fezari, and A. Redjati. "ELM and K-nn machine learning in classification of Breath sounds signals." Int J Electr Comput Eng 10, no. 4 (2020): 3528-3536.
35. Varshni, Dimpy, Kartik Thakral, Lucky Agarwal, Rahul Nijhawan, and Ankush Mittal. "Pneumonia detection using CNN based feature extraction." In 2019 IEEE international conference on electrical, computer and communication technologies (ICECCT), pp. 1-7. IEEE, 2019.
36. Zhao, Peng, Jia-Wei Shan, Yu-Jie Zhang, and Zhi-Hua Zhou. "Exploratory machine learning with unknown unknowns." Artificial Intelligence 327 (2024): 104059.
37. Khalifa, Nour Eldeen M., Mohamed Hamed N. Taha, Aboul Ella Hassanien, and Sally Elghamrawy. "Detection of coronavirus (COVID-19) associated pneumonia based on generative adversarial networks and a fine-tuned deep transfer learning model using chest X-ray dataset." In International Conference on Advanced Intelligent Systems and Informatics, pp. 234-247. Cham: Springer International Publishing, 2022.
38. Hall, Lawrence O., Rahul Paul, Dmitry B. Goldgof, and Gregory M. Goldgof. "Finding covid-19 from chest x-rays using deep learning on a small dataset." arXiv preprint arXiv:2004.02060 (2020).
39. Ilyas, Muhammad, Hina Rehman, and Amine Naït-Ali. "Detection of covid-19 from chest x-ray images using artificial intelligence: An early review." arXiv preprint arXiv:2004.05436 (2020).
40. Evans, Ethan D., Claire Duvallet, Nathaniel D. Chu, Michael K. Oberst, Michael A. Murphy, Isaac Rockafellow, David Sontag, and Eric J. Alm. "Predicting human health from biofluid-based metabolomics using machine learning." Scientific reports 10, no. 1 (2020): 17635.

41. Nagamounika, R., C. N. S. V. Sri, A. Harshitha, K. L. Tejaswi, and P. R. S. M. Lakshmi. "Prediction of pneumonia disease by using deep convolutional neural networks." J Eng Sci Criterion 17 (2020): 18.
42. Liebenlito, Muhaza, Yanne Irene, and Abdul Hamid. "Classification of tuberculosis and pneumonia in human lung based on chest x-ray image using convolutional neural network." InPrime: Indonesian Journal of Pure and Applied Mathematics 2, no. 1 (2020): 24-32.
43. Yao, Shangjie, Yaowu Chen, Xiang Tian, Rongxin Jiang, and Shuhao Ma. "An improved algorithm for detecting pneumonia based on YOLOv3." Applied Sciences 10, no. 5 (2020): 1818.
44. Chadha, Raman. "A Novel Approach for Detecting Pneumonia in Machine Learning.", (2019).
45. Chumbita, Mariana, Catia Cillóniz, Pedro Puerta-Alcalde, Estela Moreno-García, Gemma Sanjuan, Nicole Garcia-Pouton, Alex Soriano, Antoni Torres, and Carolina Garcia-Vidal. "Can artificial intelligence improve the management of pneumonia." Journal of clinical medicine 9, no. 1 (2020): 248.
46. Ebner, Lukas, Stergios Christodoulidis, Thomai Stathopoulou, Thomas Geiser, Odile Stalder, Andreas Limacher, Johannes T. Heverhagen, Stavroula G. Mougiakakou, and Andreas Christe. "Meta-analysis of the radiological and clinical features of Usual Interstitial Pneumonia (UIP) and Nonspecific Interstitial Pneumonia (NSIP)." PLoS One 15, no. 1 (2020): e0226084.
47. Shuaib, Khan Maseeh, Solkar Ahmed Shahid, Ansari Almas Javed, and Mohammed Zaid. "Pneumonia detection through x-ray using deep learning." IOSR Journal of Computer Engineering (IOSR-JCE) 22, no. 1 (2020): 08-11.
48. Li, Lin, Lixin Qin, Zeguo Xu, Youbing Yin, Xin Wang, Bin Kong, Junjie Bai et al. "Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT." Radiology (2020).
49. Liao, Yu-Hsuan, Zhong-Chuang Wang, Fu-Gui Zhang, Maysam F. Abbod, Chung-Hung Shih, and Jiann-Shing Shieh. "Machine learning methods applied to predict ventilator-associated pneumonia with Pseudomonas aeruginosa infection via sensor array of electronic nose in intensive care unit." Sensors 19, no. 8 (2019): 1866.
50. Feng, Cong, Lili Wang, Xin Chen, Yongzhi Zhai, Feng Zhu, Hua Chen, Yingchan Wang et al. "A Novel triage tool of artificial intelligence-assisted diagnosis aid system for suspected COVID-19 pneumonia in fever clinics." MedRxiv (2020): 2020-03.
51. Ghandorh, Hamza, Mohammad Zubair Khan, Mehshan Ahmed Khan, Yousef M. Alsofayan, Ahmed A. Alahmari, and Anas A. Khan. "Predicting ICU Admission for COVID-19 Patients in Saudi Arabia: A Comparative Study of AdaBoost and Bagging Methods." International Journal of Advanced Computer Science & Applications 15, no. 3 (2024).
52. Amin, Sareer Ul, Sher Taj, Adnan Hussain, and Sanghyun Seo. "An automated chest X-ray analysis for COVID-19, tuberculosis, and pneumonia employing ensemble learning approach." Biomedical Signal Processing and Control 87 (2024): 105408.
53. Shaima, Mujiba, Norun Nabi, Md Nasir Uddin Rana, Ahmed Ali Linkon, Md Shohail Uddin Sarker, Nishat Anjum, and Hammed Esa. "Machine Learning Models for Predicting Corticosteroid Therapy Necessity in COVID-19 Patients: A Comparative Study." Journal of Computer Science and Technology Studies 6, no. 1 (2024): 217-224.
54. Bhosale, Yogesh H., and K. Sridhar Patnaik. "PulDi-COVID: Chronic obstructive pulmonary (lung) diseases with COVID-19 classification using ensemble deep convolutional neural network from chest X-ray images to minimize severity and mortality rates." Biomedical Signal Processing and Control 81 (2023): 104445.
55. Garcea, Fabio, Alessio Serra, Fabrizio Lamberti, and Lia Morra. "Data augmentation for medical imaging: A systematic literature review." Computers in Biology and Medicine 152 (2023): 106391.
56. Ni, Ansong, Srini Iyer, Dragomir Radev, Veselin Stoyanov, Wen-tau Yih, Sida Wang, and Xi Victoria Lin. "Lever: Learning to verify language-to-code generation with execution." In International Conference on Machine Learning, pp. 26106-26128. PMLR, 2023.
57. Tyagi, Amit Kumar, ed. Automated Secure Computing for Next-Generation Systems. John Wiley & Sons, 2023.
58. Ieracitano, Cosimo, Nadia Mammone, Mario Versaci, Giuseppe Varone, Abder-Rahman Ali, Antonio Armentano, Grazia Calabrese et al. "A fuzzy-enhanced deep learning approach for early detection of Covid-19 pneumonia from portable chest X-ray images." Neurocomputing 481 (2022): 202-215.
59. Ahmed, F., Sumra, I. A., & Jamil, U. (2024). A Comprehensive Review on DDoS Attack in Software-Defined Network (SDN): Problems and Possible Solutions. Journal of Computing & Biomedical Informatics, 7(01).
60. Abbas, F., Iftikhar, A., Riaz, A., Humayon, M., & Khan, M. F. (2024). Use of Big Data in IoT-Enabled Robotics Manufacturing for Process Optimization. Journal of Computing & Biomedical Informatics, 7(01), 239-248.
61. Munir, A., Sumra, I. A., Naveed, R., & Javed, M. A. (2024). Techniques for Authentication and Defense Strategies to Mitigate IoT Security Risks. Journal of Computing & Biomedical Informatics, 7(01).