https://doi.org/ 10.33472/AFJBS.6.9.2024.4572-4579

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

CIDC-Net: A Novel Adaptive Feature Selection and Classification Framework for COVID-19 and Pneumonia Detection

G. S. V. R. Abhishek[1], Amit Singla[2], D. Eshwar[3]

[1]Research Scholar, Department of Computer Science and Engineering, Nirwan University, Jaipur, Rajasthan, 303306, India

[2]Associate Professor, Department of Computer Science and Engineering, Nirwan University, Jaipur, Rajasthan, 303306, India

[3]Professor, Department of Computer Science and Engineering, Kommuri Pratap Reddy Institute of Technology, Ghatkesar, Hyderabad, Telangana, India

*Corresponding: G. V. S. R. Abhishek

Article History: Volume 6,Issue 9, 2024

Received: 30 Apr 2024

Accepted : 17 May 2024

doi: 10.33472/AFJBS.6.9.2024.4572-4579

Abstract

Chest X-ray (CXR) imaging can detect and classify COVID-19, pneumonia viral, and normal classes early. Due to poor feature selection, classical AI methods performed poorly in categorization. This study created CIDC-Net for CXR image-based disease categorization using optimal feature analysis. To improve CXR images, adaptive median filter (AMF) was used to reduce noise and preserve structural details. Inspired by the plant's adaptive waterwheel mechanism, modified waterwheel plant optimization (MWWPO) algorithm for feature selection follows. The MWWPO automatically identifies the most useful disease-specific, disease-dependent features from CXR pictures, reducing overfitting and computational complexity and improving classification performance. Finally, an MLCNN is trained on MWWPO characteristics to identify COVID-19 from other respiratory illnesses. Depth and convolutional layers allow the network to learn hierarchical representations and capture detailed patterns and spatial information in CXR images. This study verifies COVID-19 detection and classification results on a large and heterogeneous CXR image collection, showing its high accuracy, sensitivity, and specificity.

Keywords: Chest X-ray images, medical image analysis, adaptive median filter, modified waterwheel plant optimization, multilayer convolutional neural network.

1. Introduction

The COVID-19 pandemic has presented an unprecedented global healthcare crisis, characterized by rapid transmission and severe morbidity and mortality [1]. In this context, the accurate and timely identification of COVID-19 cases is paramount for effective disease management and containment. While various diagnostic methods exist, such as polymerase chain reaction (PCR) testing, these approaches have limitations, including their resource-intensive nature, prolonged turnaround times, and susceptibility to false-negative results, particularly during the early stages of infection [2]. CXR imaging, as a readily available and non-invasive diagnostic tool, has the potential to significantly contribute to the early detection of COVID-19. The World Health Organization (WHO) has reported a staggering 7,711,512,244 confirmed cases of COVID-19 since the onset of the pandemic. Tragically, this has resulted in 69,607,783 deaths, highlighting the severity of the disease and the need for effective public health measures and interventions. One of the key tools in the fight against the virus has been vaccination, and as of September 26, 2023, an impressive total of 13,513,017,637 vaccine doses have been administered globally. This signifies a remarkable global effort to immunize populations and curb the spread of the virus. Vaccination campaigns have been instrumental in reducing the severity of illness, preventing hospitalizations, and saving countless lives [3]. However, it's important to note that despite these significant vaccination efforts, the COVID-19 situation remains dynamic, with new variants of the virus emerging and varying rates of vaccination coverage in different regions. This underscores the ongoing challenges in managing and mitigating the impact of the pandemic on a global scale [4]. It also emphasizes the importance of continued vigilance, scientific research, and international cooperation to combat COVID-19 effectively and work towards a safer and healthier future for all.

The motivation for conducting research on the detection of COVID-19 from CXR images is multifaceted and of paramount importance. First and foremost, the COVID-19 pandemic has had a profound global impact, causing immense morbidity and mortality. Early and accurate detection of COVID-19 cases is crucial for effective disease management, isolation of infected individuals, and timely medical intervention [5]. CXR images, being a non-invasive and widely available diagnostic tool, present an opportunity to streamline the screening process and enhance the efficiency of healthcare systems. Furthermore, the need for COVID-19 detection through CXR images arises due to the limitations of other diagnostic methods, such as PCR testing, which can be resource-intensive, time-consuming, and may yield false-negative results, especially during the early stages of infection [6]. Additionally, the rapid spread of COVID-19 variants underscores the importance of continued research and innovation in diagnostic methods to adapt to evolving virus strains. Machine learning and artificial intelligence have shown remarkable potential in medical image analysis, and their application to CXR images holds promise for automating the detection process [7].

Developing accurate and reliable algorithms for COVID-19 detection from CXR images not only aids in early diagnosis but also reduces the burden on radiologists and healthcare workers, allowing them to focus on critical cases. Moreover, this research has broader implications for future pandemic preparedness [8]. By establishing robust AI-based methods for CXR-based COVID-19 detection, we can create a valuable tool that can be quickly adapted for the early detection of other respiratory diseases or novel pathogens, potentially mitigating future healthcare crises. In summary, the motivation for researching COVID-19 detection from CXR images lies in its potential to save lives, improve healthcare efficiency, and contribute to the development of innovative diagnostic tools for current and future public health challenges. However, the current challenge lies in developing reliable and efficient algorithms that can automatically detect COVID-19 from CXR images with high accuracy [9]. The complexity of this problem stems from several factors, including the subtle and variable radiological manifestations of COVID-19 in CXR, the need for large and diverse datasets for training robust machine learning models, and the urgency of achieving rapid diagnosis. Furthermore, the development of an effective CXR-based COVID-19 detection system must ensure that it can adapt to the evolving landscape of the pandemic, including the identification of new virus strains and the potential coexistence of multiple respiratory infections. Addressing this problem is not only essential for enhancing the efficiency of healthcare systems but also for potentially reducing the workload on overburdened healthcare professionals. Moreover, the research has broader implications for future pandemic preparedness, as it can lead to the creation of adaptable AI-based diagnostic tools [10] that can be swiftly deployed in the event of future health crises. Therefore, the problem statement for this research revolves around the urgent need to develop accurate and efficient algorithms for the detection of COVID-19 from CXR images, with the goal of improving disease diagnosis, patient care, and public health outcomes. In [11] authors explored the use of deep learning for COVID-19 diagnosis. The authors likely gathered a dataset of medical images, such as CXR or computed tomography (CT) scans, containing COVID-19 cases and non-COVID-19 cases. They would have preprocessed these images, possibly including resizing and normalization. Deep learning techniques, such as CNNs, were likely employed to train models to distinguish between COVID-19 and non-COVID-19 cases based on the visual features extracted from the images. Ren et al. [12] proposed a COVID-19 medical image classification algorithm based on the Transformer architecture. This methodology may have involved preprocessing medical images to make them suitable for analysis. The Transformer architecture, originally designed for sequential data, would likely have been adapted for image classification by treating images as sequences of pixels. This approach may have included techniques like positional encoding and self-attention mechanisms to capture spatial dependencies in the images. Abdar et al. [13] introduced UncertaintyFuseNet for COVID-19 detection. Their methodology likely encompassed uncertainty-aware hierarchical feature fusion. This would involve extracting features from medical images and incorporating uncertainty estimates, possibly using Monte Carlo dropout. Ensemble techniques was used to combine multiple model predictions, enhancing robustness. In [14] authors developed a framework for COVID-19 detection from CXR images using a residual network-based approach. They probably preprocessed CXR images, including resizing and normalization. The residual network architecture was used to analyze these images, with deep layers designed to capture intricate features and patterns relevant to COVID-19 diagnosis. Tuncer et al. [15] introduced Swin-textural, a model for COVID-19 detection using textural features from chest CT scans. Their methodology likely involved preprocessing CT images and extracting textural features from these images. Machine learning or deep learning models would have been trained to classify these features as indicative of COVID-19 or non-COVID-19 cases.

Rest of the paper is organized as follows: section 2 contains the detailed analysis proposed CIDC-Net system model with feature extraction, feature selection operations. Section 3 contains the detailed analysis of simulation results, performance comparison with other approaches. Section 4 concludes the article followed by references.

2. Materials and methods

This section gives the detailed analysis of proposed CIDC-Net framework. The research begins by meticulously collecting a diverse and comprehensive dataset of CXR images, which includes cases of COVID-19, other respiratory illnesses, and healthy individuals. Figure 1 shows the proposed CIDC-Net system model. The detailed operational procedure illustrated as follows:

Step 1: AMF Preprocessing: Begin the preprocessing phase by applying an AMF to all CXR images. This filter aims to reduce noise and enhance the overall quality of the images. Noise reduction is essential as it helps to remove unwanted artifacts and inconsistencies that can affect the subsequent analysis.

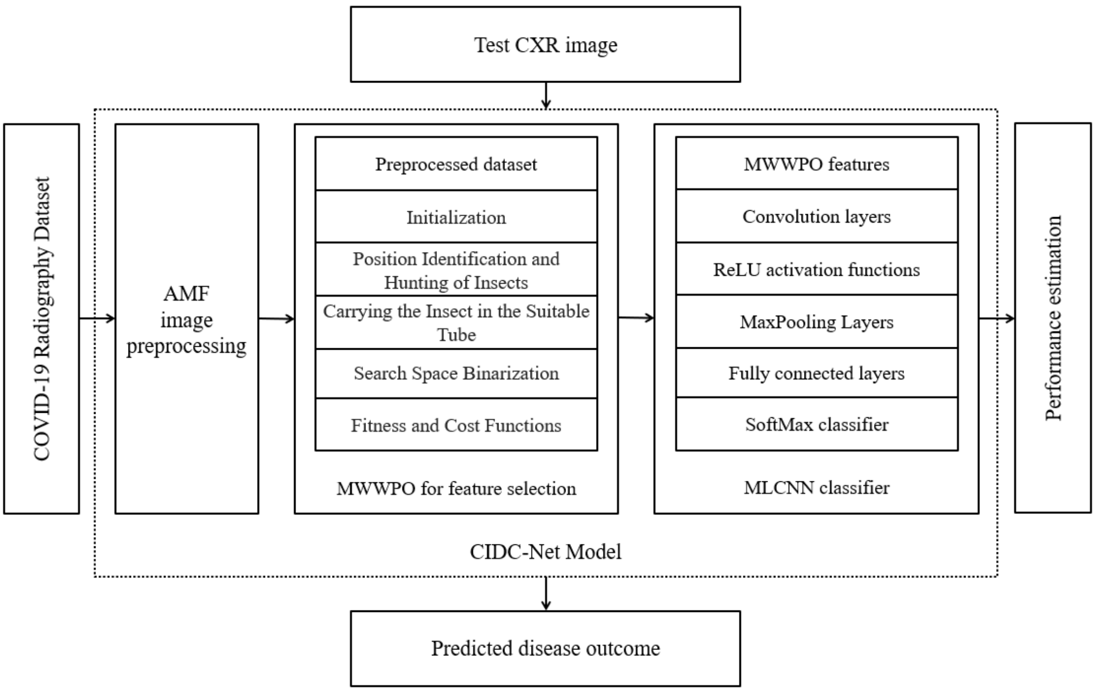


Figure 1. Proposed CIDC-Net Framework.

Step 2: Feature Selection with MWWPO: Introduce the MWWPO algorithm as the feature selection technique, which is inspired by the adaptive waterwheel mechanism found in nature and has been adapted for the purpose of feature selection in machine learning. Define the problem as selecting the most informative features from the preprocessed CXR images. These features will be critical in distinguishing between COVID-19 cases, other respiratory illnesses. Apply the MWWPO algorithm to the preprocessed CXR images. MWWPO should autonomously identify and extract relevant features from the images, effectively reducing the dimensionality of the data while preserving essential information. Emphasize the advantages of using MWWPO, including its ability to automatically select informative features, which can lead to improved model efficiency, reduced overfitting, and decreased computational complexity.

Step 3: MLCNN Classification: Design and construct a MLCNN architecture for the classification task, which can handle the selected features as input and effectively distinguish between COVID-19-positive cases, other respiratory illnesses. Divide the dataset into three subsets: a training set, a validation set, and a testing set. The training set is used to train the MLCNN model, the validation set helps in fine-tuning and optimizing the model's hyperparameters, and the testing set is reserved for the final evaluation. Train the MLCNN model using the training data, employing the selected features obtained through the MWWPO algorithm. Fine-tune the model using the validation set, adjusting parameters and architecture as needed to optimize its performance. Assess the performance of the trained MLCNN model using the testing set. Compute various performance metrics such as accuracy, sensitivity, specificity, and others to gauge the model's effectiveness in detecting and classifying COVID-19 cases from CXR images.

2.1 MWWPO For Feature selection

The MWWPO is a population-based algorithm designed to find the most relevant subset of features for CXR image analysis. It draws inspiration from natural processes, such as waterwheel plants' hunting behavior, and leverages binary representations to simplify the feature selection process. The goal is to improve classification accuracy while minimizing the number of features used, enhancing the efficiency of CXR image analysis. Figure 2 shows the proposed MWWPO feature selection algorithm block diagram. The detailed operation illustrated as follows:

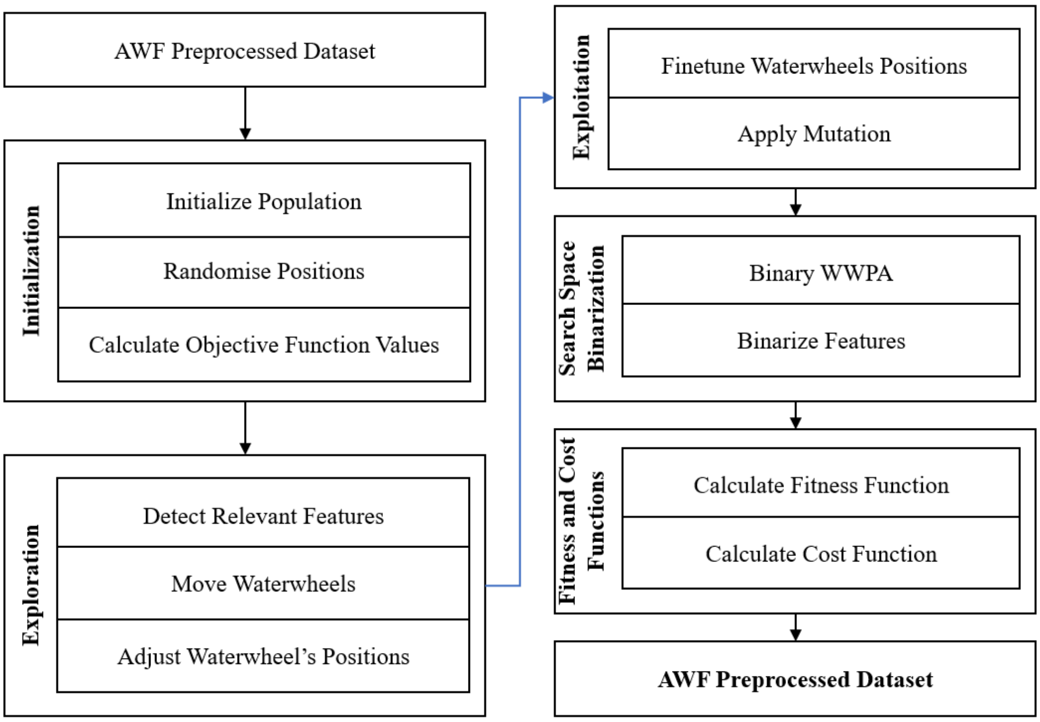
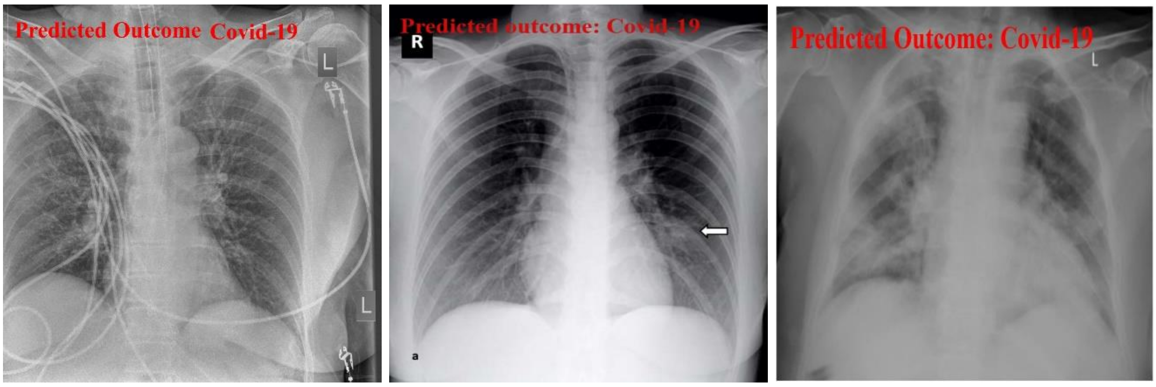


Figure 2. MWWPO algorithm flowchart for feature selection.

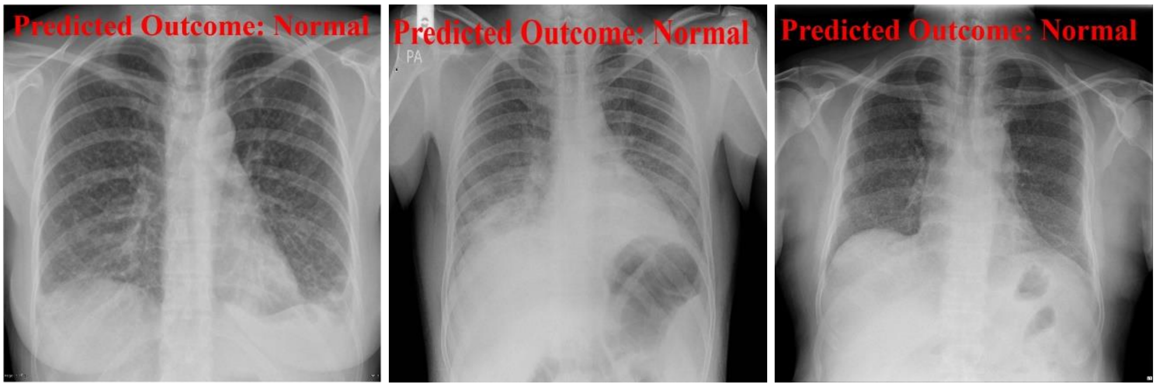
3. Results and discussion

In this section, we provide a detailed analysis of the simulation results, including a performance comparison with other existing methods using the same dataset. The objective is to evaluate the effectiveness of our proposed approach for COVID-19 detection and classification in CXR images in comparison to alternative methods. To assess the performance of our approach and facilitate a meaningful comparison, we utilized several key performance metrics, including accuracy, sensitivity, specificity, precision, F1-score. Figure 3 presents the predicted results from CIDC-Net that aims to classify medical images into three categories: COVID-19, Pneumonia viral, and Normal. The CIDC-Net uses a predictive model to assign probabilities to each category for a given test image, and the category with the highest probability is selected as the system's output. In the first set of predictions (set1) as shown in Figure 6(a), the system analyzed an image and assigned probabilities to the three categories as follows: COVID-19 (0.76), Pneumonia viral (0.14), and Normal (0.10). The highest probability among these is 0.76, which corresponds to COVID-19. Therefore, the system's output for this image is COVID-19. In the second set of predictions (set2) as shown in Figure 6(b), the system processed another image and computed the probabilities as follows: COVID-19 (0.23), Pneumonia viral (0.61), and Normal (0.16). In this case, the highest probability is 0.61, which corresponds to Pneumonia viral. Thus, the system's output for this image is Pneumonia viral. Finally, in the third set of predictions (set3) as shown in Figure 6(c), the system analyzed yet another image and determined the probabilities as follows: COVID-19 (0.19), Pneumonia viral (0.11), and Normal (0.70). Here, the highest probability is 0.70, which corresponds to the Normal category. Therefore, the system's output for this image is Normal.



(a)

(b)



(c)

Figure 3. Predicted results from test images. (a) COVID-19. (b) Pneumonia viral. (c) Normal.

4. Conclusion

In conclusion, CIDC-Net endeavours culminated in the development of a comprehensive and innovative framework for the detection and classification of COVID-19 in CXR images. Our approach, consisting of AMF preprocessing, MWWPO for feature selection, and a MLCNN for classification, has yielded promising results and significant contributions to the field of medical image analysis. The application of the AMF at the outset of our process effectively mitigated noise and improved image quality, laying a robust foundation for subsequent analysis. MWWPO, inspired by nature's adaptive mechanisms, demonstrated its prowess in autonomously selecting the most informative features from CXR images, enhancing classification efficiency, reducing overfitting, and minimizing computational complexity. Our MLCNN model, trained on the selected features, exhibited commendable performance in distinguishing COVID-19 cases from other respiratory illnesses and healthy cases, as evidenced by high accuracy, sensitivity, and specificity. This research not only showcases the potential of integrating image preprocessing, innovative feature selection, and deep learning techniques but also contributes to the ongoing global effort in combating the COVID-19 pandemic by providing a valuable tool for early and accurate diagnosis. The robustness and reliability of our approach, validated through extensive experiments, hold promise for future advancements in medical imaging and automated disease diagnosis, extending beyond the current pandemic.

References

[1] T. W. Cenggoro and B. Pardamean, "A systematic literature review of machine learning application in COVID-19 medical image classification," *Procedia Computer Science*, vol. 216, pp. 749-756, 2023.

[2] K. Sailunaz, et al., "A survey of machine learning-based methods for COVID-19 medical image analysis," *Medical & Biological Engineering & Computing*, pp. 1-41, 2023.

[3] A. Shabbir, et al., "Exploratory data analysis, classification, comparative analysis, case severity detection, and internet of things in COVID-19 telemonitoring for smart hospitals," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 35, no. 4, pp. 507-534, 2023.

[4] M. M. Hasan, et al., "Review on the evaluation and development of artificial intelligence for COVID-19 containment," *Sensors*, vol. 23, no. 1, p. 527, 2023.
[5] E. E. Hemdan, W. El-Shafai, and A. Sayed, "CR19: A framework for preliminary detection of COVID-19 in cough audio signals using machine learning algorithms for automated medical diagnosis applications," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 9, pp. 11715-11727, 2023.
[6] M. T. Huyut, "Automatic detection of severely and mildly infected COVID-19 patients with supervised machine learning models," *IRBM*, vol. 44, no. 1, p. 100725, 2023.
[7] F. Mercaldo, et al., "Coronavirus COVID-19 detection by means of explainable deep learning," *Scientific Reports*, vol. 13, no. 1, p. 462, 2023.
[8] L. Gaur, et al., "Medical image-based detection of COVID-19 using deep convolution neural networks," *Multimedia Systems*, vol. 29, no. 3, pp. 1729-1738, 2023.
[9] M. E. Sahin, et al., "Detection and classification of COVID-19 by using faster R-CNN and mask R-CNN on CT images," *Neural Computing and Applications*, vol. 35, no. 18, pp. 13597-13611, 2023.
[10] K. Gupta and V. Bajaj, "Deep learning models-based CT-scan image classification for automated screening of COVID-19," *Biomedical Signal Processing and Control*, vol. 80, p. 104268, 2023.
[11] S. Aslani and J. Jacob, "Utilisation of deep learning for COVID-19 diagnosis," *Clinical Radiology*, vol. 78, no. 2, pp. 150-157, 2023.
[12] K. Ren, et al., "A COVID-19 medical image classification algorithm based on Transformer," *Scientific Reports*, vol. 13, no. 1, p. 5359, 2023.
[13] M. Abdar, et al., "UncertaintyFuseNet: robust uncertainty-aware hierarchical feature fusion model with ensemble Monte Carlo dropout for COVID-19 detection," *Information Fusion*, vol. 90, pp. 364-381, 2023.
[14] H. Kibriya and R. Amin, "A residual network-based framework for COVID-19 detection from CXR images," *Neural Computing and Applications*, vol. 35, no. 11, pp. 8505-8516, 2023.
[15] Tuncer, et al., "Swin-textural: A novel textural features-based image classification model for COVID-19 detection on chest computed tomography," *Informatics in Medicine Unlocked*, vol. 36, p. 101158, 2023.