J. Electrical Systems 20-11s (2024): 396-408

Dr. M. Anusha[1]
P. Kiruthika[2*]

Multiclass COVID-19 Detection using Attention-Based 3D CNN with Grey Wolf Optimization

Abstract

The emergence and spread of COVID-19 have raised a need for accurate and quick diagnosis because of the high infection rate. Towards this goal, a brand-new solution for multiclass COVID-19 detection is presented in this study using deep learning. The methodology is a synergy of feature extraction of an advanced MobileNetV2 model trained on the COVID-19 data set and the classification strength of attention-based 3D Convolutional Neural Networks. MobileNetV2 also is acting as a feature extractor obtaining detailed information from various COVID-19 imaging datasets. Thereafter, an attention-based 3D CNN architecture is used for the classification process to distinguish COVID-19 classes effectively. To fine-tune the parameters of the presented model and increase modeling performance, the Grey Wolf Optimization (GWO) algorithm is added to improve the convergence to the best solution. The second concept; spatial attention is applied in order to enhance the feature selectivity in addition to enhancing the classification outcomes. The experimental data also prove the efficiency of the developed approach, which shows high accuracy as well as sensitivity and specificity in the multiclass COVID-19 detection problems. The proposed hybrid framework, has a clear potential to enhance diagnostic power in healthcare thus using the disease management and control systems.

Keywords:
Multiclass Classification, Covid-19 CT scan images, Enhanced MobileNetV2, Attention-based 3D CNN, Grey Wolf Optimization

1. INTRODUCTION

COVID-19 is a viral disease caused by the new coronavirus SARS-CoV-2, has turned into one of the biggest health crises of the current decades. After emerging in late December 2019, the virus continued to spread throughout the world and affected millions of people and caused significant death toll. The novelty and extensive reach of the pandemic have put issues related to prevention and consequences of the virus into focus [1]. Prominent among these is the creation and implementation of highly sensitive and rapid diagnostic techniques to identify COVID positive cases. Conventional procedures that have been useful include Polymerase Chain Reaction PCR tests which helps identify persons infected by the communicable disease. However, these methods often rely on specialized laboratory equipment, are rather time-consuming, and can be sensitive to supply chain disruption, which make them less easily implemented in large scale and resource constrained environments [2]. Besides, molecular tests, imaging-based approaches including the chest X-ray and/or CT scans have been used in COVID-19 diagnosis. All these modalities present with high turnover and could help in evaluating the pulmonary features of the disease. However, they are not without limitations such as differences in inter method variations, use of ionizing radiation and qualification of radiologist for image interpretation [3].

During these challenges, the advancement of deep learning techniques has proven remarkable for improving the diagnostics of COVID-19. Of the many subtypes of artificial neural networks, Convolutional Neural Network CNN has been proven to deliver high performance when used to process images including medical images. However, with the aid of Big datasets and automated computational power, CNNs are able to learn relevant features of raw image data for disease classification [4]. Nevertheless, the potential of those approaches is undoubted, however, the existing deep learning-based approaches for COVID-19 detection impose some limitations that are associated with generalization, scalableness, and interpretability. Additionally, most of these methods Centre their problem-solving on binary classification problems where the model is trained to either confirm or dispute the presence of COVID-19 without thought being given to its multiclass nature [5]. Therefore, it is possible to agree with the finding that there is a possibility to develop more sophisticated deep learning frameworks that will be viable to help diagnose COVID-19 across classes, including disease presentation severity and other forms [6]. To this end, the proposed work seeks to fill this gap through the development of a new hybrid model that employs the feature extraction from MobileNetV2 with CNNs trained with GWO for multiclass COVID-19 diagnosis. By incorporating the recent deep learning methodologies with optimization techniques, the proposed framework is intended to improve the effectiveness, time and size complexity of COVID-19 diagnostics which in turn support better disease management and control programs.

[1] Assistant Professor, PG & Research Department of Computer Science, National College (Autonomous), Affiliated to Bharathidasan University, Tamil Nadu, India, (E-mail-anusha260505@gmail.com)

[2*] Ph. D Research Scholar, PG & Research Department of Computer Science, National College (Autonomous), Affiliated to Bharathidasan University, Tamil Nadu, India, (E-mail-kiruthikashyne@gmail.com)

***Corresponding Author:** P. Kiruthika,
*(E-mail-kiruthikashyne@gmail.com)

2. RELATED WORK

Hassan et al (2020) proposed a system to diagnoses COVID-19 using recurrent neural networks (RNNs). The authors developed this system for the recognition of COVID-19 based on medical data, that can be chest X-ray or a CT scan. Their approach utilizes temporal dependencies in the data with the help of which RNNs excel at to enhance the precision of detection. Machine learning techniques, primarily RNNs, have been identified to serve as a helping hand toward diagnosing COVID-19 in addition to increasing diagnostic achievement rates of 96.7 % during the pandemic. Amyar et al., approach a systematic approach in examining the CT images of COVID-19 pneumonia employing deep learning techniques. System to analyze medical data, such as chest X-rays or CT scans, to identify the presence of COVID-19. Their approach leverages the temporal dependencies in the data, which RNNs are well- suited to handle, to improve the accuracy of detection. The study highlights the potential of machine learning techniques, specifically RNNs, in aiding the diagnosis of COVID-19 and contributing to more efficient and achieved an accuracy rate of 96.7% during the pandemic. Amyar et al. (2020) undertook a comprehensive approach to analyze CT images of COVID-19 pneumonia using deep learning methods. The authors' proposed multi-task deep learning model for the segmentation and the classification tasks can be executed in tandem. Their model resulted in helping categorize COVID-19 cases and divide different parts of the lungs by analyzing CT scans and enabled estimation of the degree and severity of pneumonia in patients, which can help in their handling and treatments and reached an accuracy rate of 92.86%. That is, Fan et al. (2021) proposed a deep learning-based approach for the multi-class COVID-19 detection from X-ray images. Their solution presumed the usage of deep learning model, probably Convolutional Neural Network (CNN), that autonomously inevitably extracts features from X-ray images and classify images into multiple COVID classes that were possibly COVID-19, COVID-19 pneumonia and No Finding obtain accuracy margin from 93% to 98% based on the approach used and database information. As a result, different manifestations of the disease could be distinguished consistently with the help of discriminative power of deep learning in the X-ray based diagnosis, completed by the authors. In their study published in 2021, Khan et al. endeavored to improve COVID-19 case detection from the chest CT scans with the help of deep learning. Their strategy was to implement entropy-controlled firefly optimization in parallel with the feature fusion with deep learning model to enhance the COVID-19 diagnosis. As such, integrating these optimization techniques with the deep learning allowed the extraction of discriminative features from the CT images and in turn yielding a classification of COVID- 19 at a high accuracy rate, of 98%, to allow for early diagnosis and prompt management. Khan et al (2021) discuss a problem of ensuring proper classification of COVID-19 CT images as positive. They created a model based on Convolutional Neural Network (CNN) using CT images for the differentiation of COVID-19 positive and negative. With a help of discriminative features learned by the CNN the model reached the high accurate rate up to 98.6% as for COVID-19 cases detected by means of CT scans that fits better for COVID-19 time-oriented diagnosis. With deep learning models to improve the accuracy of COVID-19 diagnosis. By combining these optimization techniques with deep learning, their method effectively extracted discriminative features from CT images and achieved accurate classification of COVID-19 cases at rate of 98%, thereby facilitating timely diagnosis and patient management. Khan et al (2021) address the challenge of accurately classifying positive COVID-19 CT scans using deep learning techniques. They developed a Convolutional Neural Network (CNN)-based model trained on a dataset of CT images to distinguish between COVID-19 positive and negative cases. By leveraging the discriminative features learned by the CNN, the model achieved high accuracy rate of 98.6% in identifying COVID-19 cases from CT scans, thereby facilitating timely and accurate diagnosis of the disease. Majid et al. (2021) proposed a new ensemble system for COVID-19 classification from CT scans. Their approach included using several deep learning models trained to classify the data and to enhance the level of accuracy of classification. With help of the ensembles of deep learning models, that they used, their method improved the stability and the accuracy of COVID-19 diagnosis from CT scans and they got the accuracy rate of 95.6%. Canayaz et al. (2022) suggested using Bayes optimization-deep neural networks and machine learning algorithms in COVID-19 diagnosis based on CT images. Their approach thought used a framework known as Bayes optimization to tune deep neural network parameters to improve upon the classifiers' performance of distinguishing between CT images at the rate of 99.37%. Moreover, through incorporating machine learning algorithms, their method was helpful in distinguishing COVID-19 positive cases from the CT scans that were taken into consideration in disease control. Abdelhamid et al. (2022) proposed a deep learning model for multi-classification of posterior anterior chest X-rays for COVID-19 diagnosis and machine learning algorithms. Their approach leveraged Bayes optimization to optimize the parameters of deep neural networks, enhancing the model's performance in classifying CT images at the rate of 99.37%. By integrating machine learning algorithms, their method achieved accurate diagnosis of COVID-19 cases from CT scans, contributing to efficient disease management and control. AbdElhamid et al. (2022) developed a deep learning-based approach for multi-classification of chest X- rays for COVID-19 diagnosis. Their approach relied on deep learning to extract features directly from chest X-ray images and classify them in multiple COVID-19 classes. Thus, their method of correct X-ray images classification at the level of 99.3% help the clinicians and researchers to better understand the severity and symptoms of COVID-19 for further appropriate treatment and patients' management. According to Woan Ching et al. (2022) colleagues, the deep learning used CNN architecture was prospectively trained for multiclass COVID-19 CT image analysis. Although the CNN model, which the authors adopted, was trained with CT scans of different disease severity levels, it successfully categorised the images into multiple COVID-19 classes, including mild, moderate, or severe COVID- 19 at a 93.80% success rate. Their model helped them correctly estimate the severity of the disease from the CT images and get a tool for initial patient categorization and treatment planning. Samee et al (2022) developed a metaheuristic optimization algorithm incorporated with deep learning for the diagnosis of COVID-19 using chest X ray image. Their technique combined metaheuristic optimisation approach with deep learning algorithms to improve the classifier's performance of chest X-ray images. From these parameters, their approach ensured that the model yielded 99.8% accuracy on COVID-19 cases' classification to support an efficient diagnosis and management of the disease. Multiclass classification of COVID-19 CT images has thus been enhanced by the hybrid feature extraction method developed by Abubakar et al. (2024). To improve the discriminative ability of the classification model, they used few different feature extraction methods. These achievements have been made possible by capturing relevant features about COVID-19 from the CT images for the method achieved enhanced classification performance and passed through the rate of 99.4% for accurate diagnosis of COVID-19 cases regardless of the disease severity levels or multiple manifestations. At the end of the analysis of the literature review highlights significant advancements in deep learning- based approaches for COVID-19 diagnosis using medical imaging data, particularly chest CT scans and X-ray images. However, a notable research gap exists in the majority of the reviewed papers, as they primarily focus on binary classification tasks and often lack consideration for multiclass COVID-19 classification, which is crucial for assessing disease severity and guiding appropriate treatment strategies. The proposed methodology, which integrates MobileNetV2 as a feature extractor and employs attention-based 3D CNN with optimizing techniques such as grey wolf optimization for multiclass COVID-19 classification, addresses this research gap effectively. When tapping into MobileNetV2's affordable feature extractor and incorporating the spatial attention-based 3D CNN. In conclusion, the proposed methodology brings the solution to the mentioned research gap in the COVID-19 diagnostics by providing the comprehensive methodology for the multiclass COVID-19 classification. As a result, it is potential for state-of-art deep learning and optimization methods providing effective disease diagnosis and management towards better patient health condition in the fight against COVID-19 virus.

3. PROPOSED METHODOLOGY

3.1 Dataset description

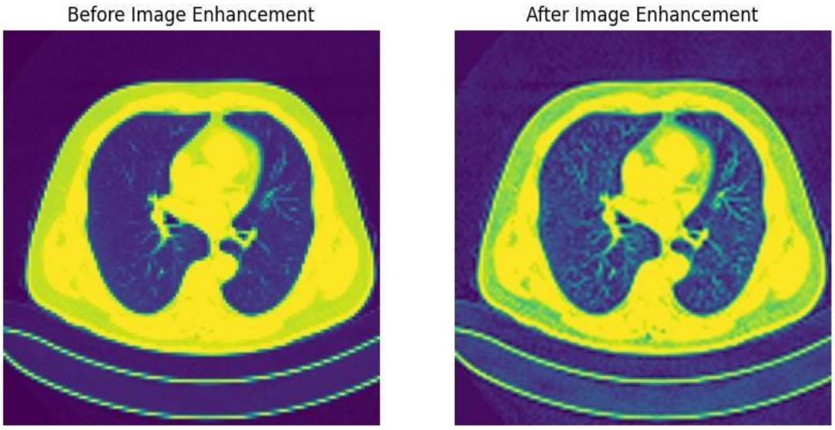
The dataset utilized in this research is sourced from various origins and is segregated into four distinct categories: These include normal transport, transport during the COVID-19 pandemic, transport in pneumonia cases, and transport in Omicron/Delta waves. Each category will denote a particular health condition or disease, and the number of images is relevant to the existing data for training and evaluating Machine Learning capabilities. In total, the dataset comprises 61,782 images of normal CT scans, 21,036 images of CT scans depicting Covid-19 cases, 21,191 images of CT scans illustrating pneumonia cases, and 12,200 images of CT scans representing the Omicron and Delta variants of the SARS-CoV-2 virus.

3.2 Image normalization

Image normalization is a form of image pre-processing, widely used in CT scans to enhance the contrast of images by equalizing the pixel intensity and Hounsfield Units (HU) across a scan. It has been observed that the intensity of CT scans varies because of causes such as settings of the scanner and the patient himself. Methods such as intensity normalization and Hounsfield unit transformation are used to bring the intensities of images into a uniform range in order to maintain consistency in presenting pixel values as image intensity and tissue density. Thus, normalizing influence removes significant differences in CT images and provides equally good images for machine learning algorithms. This preprocessing enhances model performance, facilitating accurate analysis and diagnosis of medical conditions from CT scans.

3.3 Image Augmentation

Image augmentation in the case of CT scan images using GANs would mean creating synthetic images that look like real images of CT scans with some variation in appearance [18],[19]. GANs, which compose of generator and discriminator network are trained in an adversarial manner where the generator gives back realistic images from random noise while the discriminator discerns between the realistic and fake images. In the domain of CT scans, the augmentation applying GAN entails realistic distortions comprising modulation of contrast, noise, texture, and anatomical structures producing more variety and richness of the training data set. Still addressing the need for improvement of the robustness and generalization capability of the models trained with CT images, GAN-based augmentation allows to create synthetic samples reflecting the variations often observed in the actual practice, as depicted in Fig. 1. However, GAN-based augmentation caters for the problems of limited or imbalanced dataset by creating more data that can balance the dataset and also reduce the likelihood of over fitting on the data set. Also, this technique may help resolve the problem of sharing medical imaging information by creating images that anonymize patient data, but retaining key features required to train the model. In summary, GAN based image augmentation provides a very effective method to extend training data set, enhancing diagnostic and analysis performance of medical imaging applications.


Before Image Enhancement

After Image Enhancement

Figure 1: Image Enhancement using GAN

3.4 Enhanced MobileNetV2 Deep features

Feature extraction using Enhanced MobileNetV2 for CT scan images involves leveraging the architecture and capabilities of MobileNetV2, a convolutional neural network (CNN) specifically engineered for efficient and accurate image analysis tasks. Enhanced MobileNetV2 incorporates various architectural innovations to optimize its performance, particularly in scenarios with resource constraints, such as processing CT scan images.

One key architectural feature of MobileNetV2 [20],[21],[22] is its utilization of depth wise separable convolutions. Unlike traditional convolutions that operate on all input channels simultaneously, depth wise separable convolutions decompose the convolution operation into two separate steps: depth wise convolutions and pointwise convolutions. Depth wise convolutions apply a single filter per input channel, capturing spatial dependencies within each channel independently. Pointwise convolutions then combine the outputs of depth wise convolutions across channels to create the final feature map. This separation reduces the computational burden by factorizing the convolution operation, leading to a more lightweight and efficient network architecture.

Additionally, MobileNetV2's feature hierarchies enable the extraction of informative features at multiple levels of abstraction. The network comprises multiple layers, each progressively capturing more complex and abstract representations of the input images. This hierarchical feature extraction process allows MobileNetV2 to encode both fine-grained details and high-level semantic information present in CT scan images, facilitating accurate analysis and interpretation. Furthermore, Enhanced MobileNetV2 can benefit from transfer learning, a technique where pre-trained models are fine-tuned on target datasets. By initializing MobileNetV2 with weights learned from large-scale image datasets and fine- tuning it on CT scan images, the network can adapt its feature extraction capabilities to the specific characteristics of medical imaging data. This transfer of knowledge enables MobileNetV2 to effectively extract relevant features from CT scan images, enhancing its performance on tasks such as disease diagnosis, anomaly detection, or segmentation.

3.5 Classification

The classification process involves feeding CT scan images through the 3D CNN architecture and obtaining predictions for each image [23],[24]. During inference, the 3D CNN computes a probability distribution over the different COVID-19 classes or severity levels based on the extracted features. The class with the highest probability is assigned as the predicted class label for the input image. Apart from 3D CNN, RNN algorithm also was utilized for classification purpose for the purpose of comparative analysis.

3.5.1 Model Architecture of 3D CNN: For multiclass classification of COVID-19 CT scan images, a typical spatial attention-based 3D CNN architecture consists of several layers and it is described in the following stages:

1. **Input Layer**

The input to the network is a 3D volume representing the CT scan of a patient's chest. This volume is composed of multiple 2D slices stacked together to form a three-dimensional representation of the lung. Let X be the input 3D volume representing the CT scan with dimensions D×H×W, where D is the depth (number of slices), H is the height, and W is the width.

$$X \in R^{DXHXW} \tag{1}$$

2. **Initial 3D Convolutional Layers**

The initial layers of the network consist of several 3D convolutional layers. These layers use 3D kernels to perform convolutions across the volume, capturing spatial features in all three dimensions (height, width, and depth). Each convolutional layer is typically followed by a non-linear activation function such as ReLU (Rectified Linear Unit) and a 3D max-pooling layer to reduce the spatial dimensions while retaining important features.

Let W_{conv1} be the weights of the first 3D convolutional layer with kernel size $k_1 \times k_1 \times k_1$ and b_{conv1} be the biases.

$$Z_{conv} = ReLU(\text{Wconv1} * \text{X} + \text{bconv1}) \quad (2)$$

$$Z_{pool}1 = MaxPool3D(\text{Zconv1}, \text{pool\_size} = p_1 \quad (3)$$

3. Spatial Attention Mechanism

After the initial convolutional layers, the feature maps are passed through a spatial attention mechanism [25], [26], [27]. This mechanism consists of several steps:

- **Attention Map Generation**: A separate 3D convolutional layer is used to generate an attention map. This map highlights the regions within the 3D volume that are most relevant for detecting COVID-19.

Let W_{attn} be the weights for the 3D convolutional layer used for generating the attention map.

$$A = sigmoid(\text{W}_{\text{attn}} * z_{Conv}1 + \text{b}_{\text{attn}}) \quad (4)$$

- **Attention Weighing**: The attention map is applied to the original feature maps by element-wise multiplication. This process enhances the features from the most informative regions and suppresses the less relevant ones.

$$Z_{attn} = A \odot Z_{conv}1 \quad (5)$$

4. Refinement 3D Convolutional Layers

The weighted feature maps are then fed into additional 3D convolutional layers for further feature extraction and refinement. These layers help to capture more complex patterns and higher-level features that are indicative of COVID-19.

$$Z_{conv}2 = ReLU(\text{ W}_{\text{conv2}} * \text{ X } + \text{b}_{\text{conv2}}) \quad (6)$$

$$Z_{pool}2 = MaxPool3D(\text{Z}_{\text{conv1}}, \text{pool\_size} = p_2) \quad (7)$$

5. Flattening and Fully Connected Layers

The refined feature maps are flattened into a one-dimensional vector. This vector is then passed through several fully connected (dense) layers. These layers perform the final stages of feature integration and transformation to prepare for classification. Dropout layers may also be included to prevent overfitting by randomly deactivating neurons during training. Let z_{flat} be the flattened vector from Z_{pool2}.

$$Z_{flat} = \text{Flatten}(Z_{pool}2) \quad (8)$$

Fully connected layers:

$$\text{z}_{\text{fc1}} = \text{ReLU}(W_{\text{fc1}} z_{\text{flat}} + b_{fc1}) \quad (9)$$

$$z_{drop} = \text{Dropout}(z_{fc}1, \text{rate} = \text{r}) \quad (10)$$

6. Output Layer

The final layer of the network is a fully connected layer with a SoftMax activation function for multi- class classification. This layer outputs the probability of the input CT scan belonging to each class, typically indicating the presence or absence of COVID-19. Let W_{out} and b_{out} be the weights and biases of the output layer, respectively. The softmax function σ is applied to obtain the class probabilities.

$$y_{pred} = \sigma(W_{out} Z_{drop}) + b_{out} \quad (11)$$

The output y_{pred} represents the probability distribution over the classes for the given input CT scan.

3.5.2 Grey Wolf Optimization (GWO)

Grey Wolf Optimization (GWO) is a population-based optimization technique inspired by the social hierarchy and hunting behavior of grey wolves in nature. In GWO, the wolves are categorized into four groups: **alpha (α)**, **beta (β)**, **delta (δ)**, and **omega (ω)**, which represent the hierarchical structure of leadership and decision-making in the pack. The algorithm simulates the grey wolves' cooperative hunting behavior to search for an optimal solution in a multidimensional search space.

The **α wolves** represent the best solutions, followed by **β** and **δ** wolves, which guide the optimization process. **ω wolves** are the remaining population, exploring the search space.

GWO consists of three main stages: **encircling**, **hunting**, and **attacking** the prey (optimal solution). The positions of wolves are updated based on the positions of the top three wolves (α, β, and δ) using the following equations:

Encircling Prey:

$D = | C \cdot Xprey(t) - X(t) |$
$X(t+1) = Xprey(t) - A \cdot D$

where A and C are coefficient vectors that control the search dynamics, $Xprey$ is the prey's position (best solution), and X(t) is the current position of the wolf.

Hunting (Exploitation): The positions of the wolves are updated with respect to the α, β, and δ wolves, ensuring convergence towards the optimal solution.

Attacking (Exploration): As the wolves converge towards the prey, the vectors A and C are dynamically adjusted, allowing for a balance between exploration and exploitation to avoid local minima.

The GWO algorithm efficiently fine-tunes network parameters by balancing exploration and exploitation, which enhances convergence in deep learning models like the attention-based 3D CNN used for multiclass COVID-19 detection. This optimization technique contributes to better performance by refining hyperparameters, reducing training time, and improving accuracy.

3.5.3 Model architecture of RNN algorithm: The Recurrent Neural Network (RNN) [28],[29],[30] architecture for CT scan classification leverages the sequential nature of CT scan slices, capturing temporal dependencies and spatial features to improve classification accuracy.

1. **Input Layer**

The input layer receives a sequence of CT scan slices, treating each slice as a separate time step.

- **Input Shape:** (T,H,W,C)
- T: Number of slices in the sequence
- H: Height of each slice
- W: Width of each slice
- C: Number of channels (1 for grayscale)

2. **Preprocessing Layer**

Normalizes pixel values and applies data augmentations.

3. **Feature Extraction Layer**

A Convolutional Neural Network (CNN) extracts spatial features from each slice.

- **Convolutional Layer:**

$\mathrm{Conv}(X)=X*W+b \quad (12)$

X: Input image ,W: Convolutional filter, b: Bias term, *: Convolution operation

- **Activation (ReLU):**

$\mathrm{ReLU}(z)=\max(0,z) \quad (13)$

z: Input to the ReLU function

- **Pooling Layer:**

$\mathrm{MaxPool}(X)=\max(X) \quad (14)$

Applies max pooling to reduce spatial dimensions

4. **Recurrent Layer**

Processes the sequence of extracted features using LSTM or GRU layers.

5. **Fully Connected Layer**

Dense Layer: Transforms the RNN output into a format suitable for classification.

Dense Layer:$y = w_d . h + b_d \quad (15)$

$W_{d:}$ Weight matrix, b_d: Bias term, h: Hidden state or context vector

Dropout: Randomly sets a fraction of input units to zero during training to prevent overfitting.

6. Output Layer

The final layer produces classification probabilities using a softmax activation function.

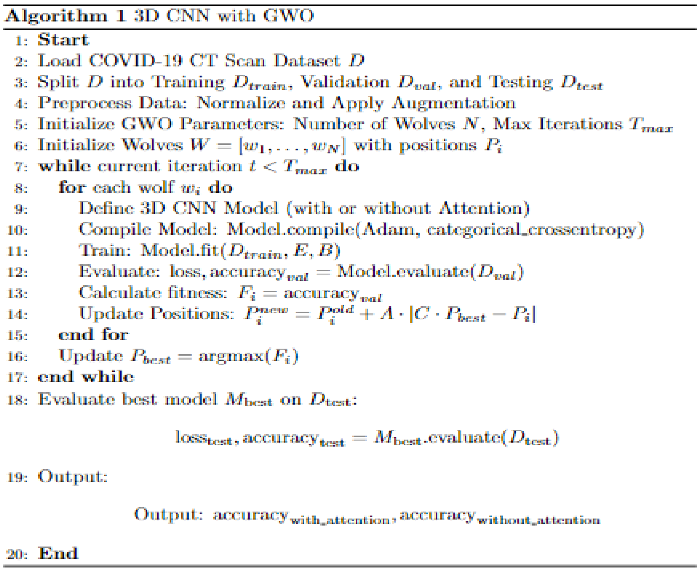
$$Softmax(z) = \frac{exp(z_i)}{\sum_j exp(z_j)}$$

z_i: Logits for class I, Softmax(z_i) : Probability of class i

3.5.4 Pseudocode

In this subsection pseudocode for each algorithm which are used for classification process has been elucidated below:

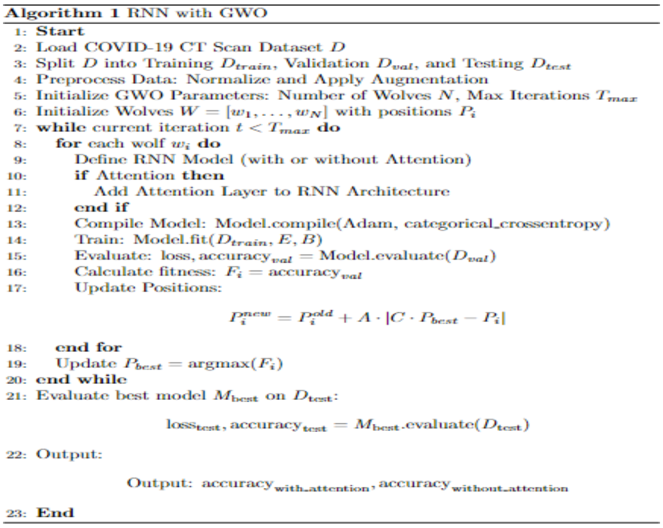
Algorithm 1: Pseudocode for 3D CNN with GWO



```
Algorithm 1 3D CNN with GWO
Start
Load COVID-19 CT Scan Dataset $D$
Split $D$ into Training $D_{train}$, Validation $D_{val}$, and Testing $D_{test}$
Preprocess Data: Normalize and Apply Augmentation
Initialize GWO Parameters: Number of Wolves $N$, Max Iterations $T_{max}$
Initialize Wolves $W = [w_1, \ldots, w_N]$ with positions $P_i$
while current iteration $t < T_{max}$ do
    for each wolf $w_i$ do
        Define 3D CNN Model (with or without Attention)
        Compile Model: Model.compile(Adam, categorical_crossentropy)
        Train: Model.fit($D_{train}, E, B$)
        Evaluate: $\text{loss}, \text{accuracy}_{val}$ = Model.evaluate($D_{val}$)
        Calculate fitness: $F_i = \text{accuracy}_{val}$
        Update Positions: $P_i^{new} = P_i^{old} + A \cdot |C \cdot P_{best} - P_i|$
    end for
    Update $P_{best} = \text{argmax}(F_i)$
end while
Evaluate best model $M_{\text{best}}$ on $D_{\text{test}}$:
        $\text{loss}_{\text{test}}, \text{accuracy}_{\text{test}} = M_{\text{best}}.\text{evaluate}(D_{\text{test}})$
Output:
        Output: $\text{accuracy}_{\text{with\_attention}}, \text{accuracy}_{\text{without\_attention}}$
End
```

The above pseudocode (algorithm 1) combines 3D CNN with Grey Wolf Optimization (GWO) for model optimization. It involves loading and preprocessing the dataset, defining GWO parameters, and iteratively training and evaluating 3D CNN models for each wolf in the GWO population. The wolves' positions are updated based on performance, and the best model is selected and evaluated on the test set for final accuracy and metrics calculation.

Algorithm 2: Pseudocode for RNN with GWO



```
Algorithm 1 RNN with GWO
Start
Load COVID-19 CT Scan Dataset $D$
Split $D$ into Training $D_{train}$, Validation $D_{val}$, and Testing $D_{test}$
Preprocess Data: Normalize and Apply Augmentation
Initialize GWO Parameters: Number of Wolves $N$, Max Iterations $T_{max}$
Initialize Wolves $W = [w_1, \ldots, w_N]$ with positions $P_i$
while current iteration $t < T_{max}$ do
    for each wolf $w_i$ do
        Define RNN Model (with or without Attention)
        if Attention then
            Add Attention Layer to RNN Architecture
        end if
        Compile Model: Model.compile(Adam, categorical_crossentropy)
        Train: Model.fit($D_{train}, E, B$)
        Evaluate: $\text{loss}, \text{accuracy}_{val}$ = Model.evaluate($D_{val}$)
        Calculate fitness: $F_i = \text{accuracy}_{val}$
        Update Positions:
                $P_i^{new} = P_i^{old} + A \cdot |C \cdot P_{best} - P_i|$
    end for
    Update $P_{best} = \text{argmax}(F_i)$
end while
Evaluate best model $M_{\text{best}}$ on $D_{\text{test}}$:
        $\text{loss}_{\text{test}}, \text{accuracy}_{\text{test}} = M_{\text{best}}.\text{evaluate}(D_{\text{test}})$
Output:
        Output: $\text{accuracy}_{\text{with\_attention}}, \text{accuracy}_{\text{without\_attention}}$
End
```

The pseudocode (algorithm 2) describes a hybrid approach using Recurrent Neural Networks (RNN) optimized with GWO. It begins by loading and preprocessing the dataset, followed by splitting it into training and testing sets. RNN models are created for each wolf in the GWO population, trained, and evaluated based on metrics like accuracy and loss.

4 RESULT AND DISCUSSION

This section presents the outcomes and analysis of the proposed framework for multiclass COVID-19 detection using deep learning techniques. The study aimed to address the need for accurate and efficient diagnostic tools amidst the global COVID-19 pandemic. Leveraging the combined capabilities of MobileNetV2 for feature extraction and Attention based 3D Convolutional Neural Networks (CNNs) for classification, the methodology was designed to enhance diagnostic accuracy across diverse COVID- 19 imaging datasets. The accuracy of a classification model is calculated using the following equation:

4.1 Explanation of Terms:

4.1.1 Number of Correct Predictions: This is the sum of true positive (TP) and true negative (TN) predictions made by the model. In the context of COVID-19 detection, this includes correctly identifying the presence (positive) or absence (negative) of a specific class (e.g., COVID-19, Pneumonia, Normal, Omicron/Delta).

4.1.2 Total Number of Predictions: This is the sum of all predictions made by the model, including true positives, true negatives, false positives (FP), and false negatives (FN). Essentially, it's the total number of samples in the dataset.

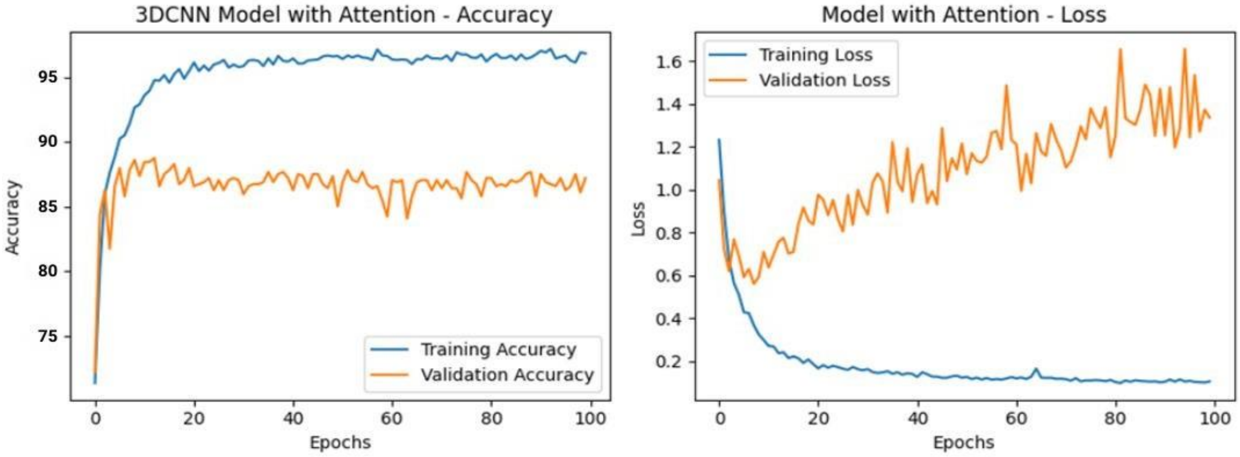


Figure 2: Performance analysis of 3DCNN model with attention

The figure 2 represents the training and the validation loss of the proposed 3D CNN with attention used in the experiments. The training accuracy in the accuracy plot increases sharply and reaches approximately 98% which shows good learning capability. Validation accuracy rises and it oscillates before finalizing minutely lower than the best percentage, which points towards the good generalization not excellent. The loss plot displays that training loss is low and stable which means better fit to the training set while validation loss drops initially and relatively higher than training loss and it increases after few epochs, which assume over fitting. Overall, the model is highly accurate, but regularization could improve its generalization.

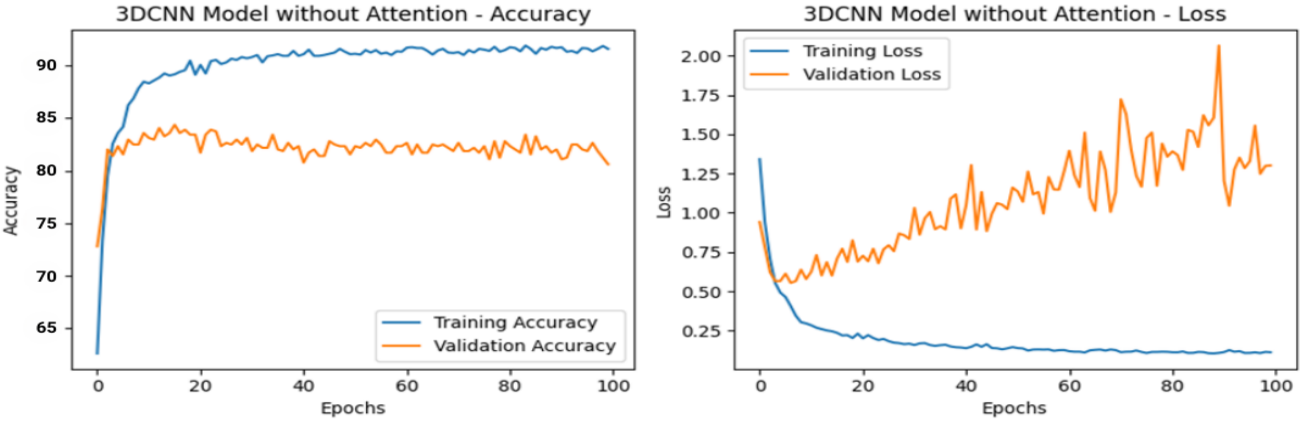


Figure 3: Performance analysis of 3DCNN model without attention

The figure 3 shows the training and validation performance of a 3D CNN model without attention over 100 epochs. In the accuracy plot (left), the training accuracy (blue line) improves and stabilizes around 90%, while the validation accuracy (orange line) fluctuates around 80%, indicating moderate generalization. In the loss plot (right), the training loss decreases and remains low, but the validation loss fluctuates significantly after an initial drop, suggesting overfitting. The model learns the training data well but struggles to generalize effectively to unseen data, as indicated by the divergence between validation loss and accuracy.

Model	**Accuracy**
3DCNN with Attention	98%
3DCNNwithout Attention	92%

Table 1: Performance analysis of 3DCNN

The table 1 shows a comparison between 3D CNN models with attention mechanisms and those without them when used for detecting multiclass COVID-19 with an emphasis on accuracy. The system that integrates spatial-based attention presents a much higher accuracy of 98%, than the 3D CNN sole model, which yields 92% accuracy. The attention mechanism enhances the network by restricting attention to the critical features of the CT scan images so as to achieve better differentiation between multiple COVID-19 classes. In contrast, the model without attention lacks this targeted focus, which limits its ability to capture the subtle, crucial details required for high-precision classification.

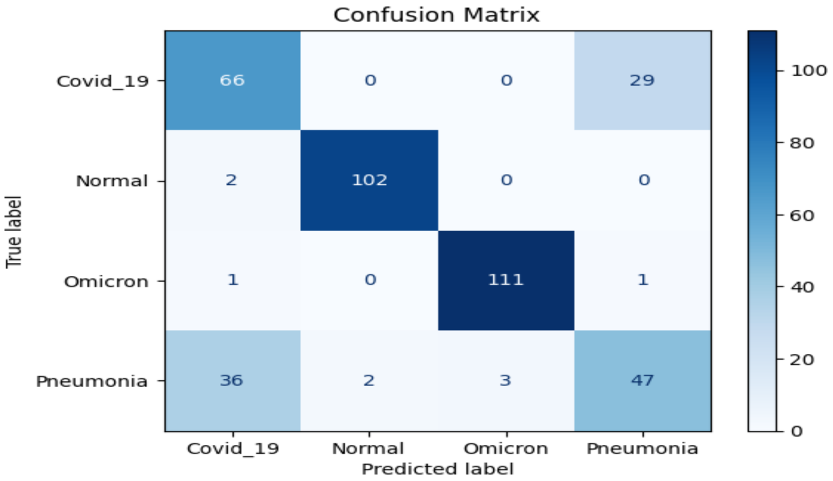


Figure 4a: Confusion matrix for 3DCNN with attention

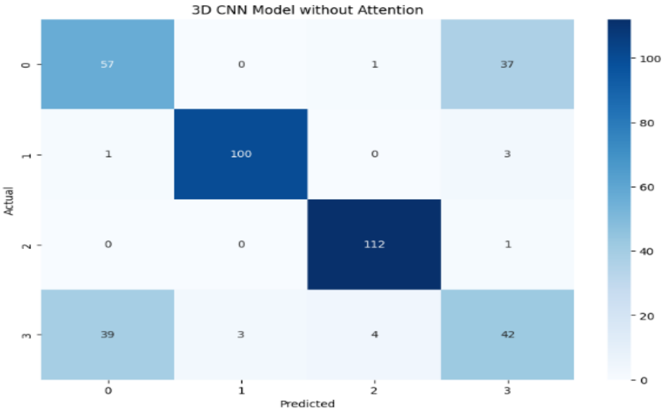


Figure 4b: Confusion matrix for 3DCNN without attention

Two confusion matrices depicted in Figures 4a and 4b show the efficiency of 3D CNNs with and without attending mechanisms when it comes to the multiclass COVID-19 detection. Figure 4a depicts correct classifications where the model with attention records more correct classification values such as 103, 113 but at the same time the model misclassifies instance, for instance class 4 instance 38 misclassified as class 1. In figure 4b, the model which does not incorporate attention also fares well with figures such as 100 and 112, however, wrongly classifies 39 samples belonging to class 3 as belonging to class 0. Both models perform well overall, but misclassifications occur, particularly between certain classes.

4.2 RNN Model

In the table 2 the accuracy of the RNN models with and without attention mechanisms is shown. The RNN with Attention yields an accuracy of 91%, namely due to paying attention to the significant features. However, the accuracy of the model, named RNN without Attention, is 83%, because it does not pay attention to the most important parts of the input. As we can see from the above comparison, it would be more appropriate to enhance the classification ability of the RNN by including attention. In general, it has been identified that through the use of attention mechanisms, the accuracy of a model is enhanced.

Model	**Accuracy**
RNN with Attention	91%
RNN without Attention	83%

Table 2: Performance analysis of RNN

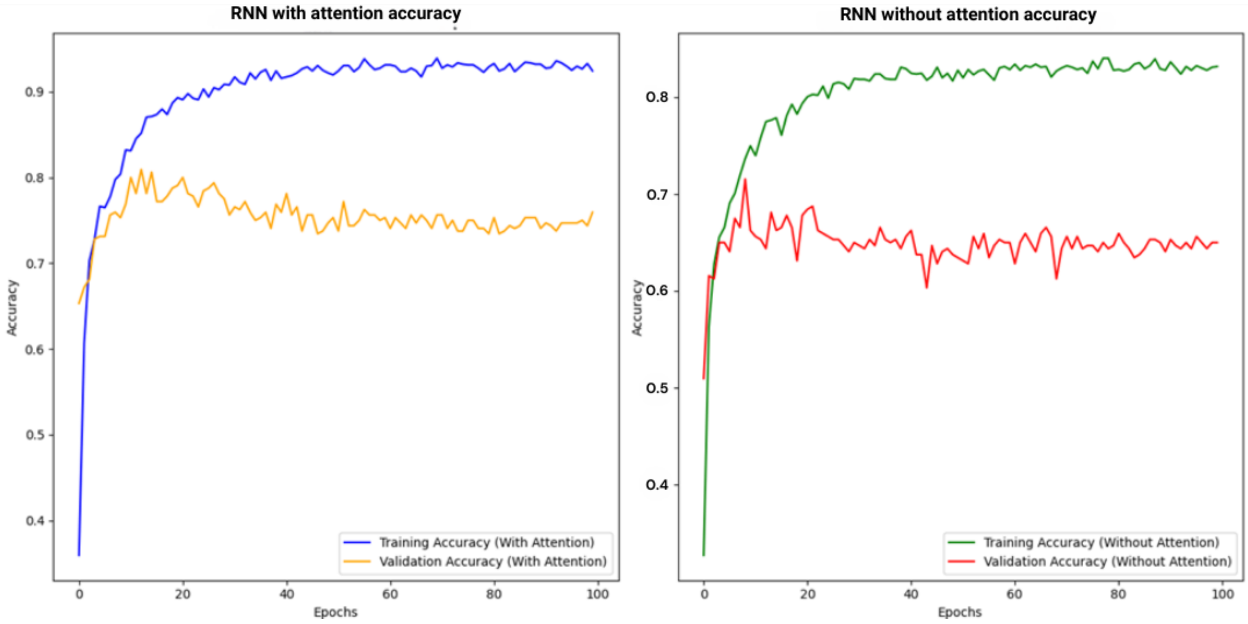


Figure 5: Performance analysis of RNN model with attention

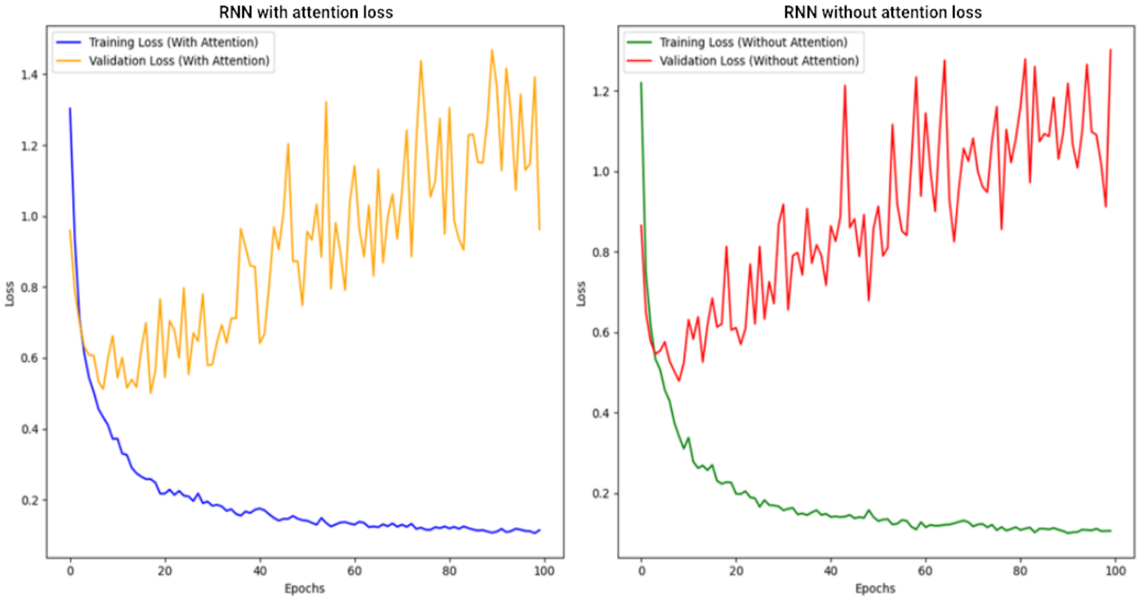


Figure 6: Performance analysis of RNN model without attention

The graph (figure 5) shows the training/ validation accuracy curve for RNN models with/ without layer attention in 100 epochs. As shown in the left plot of the RNN with attention, for training accuracy (blue line), it gradually rises and reaches up to about 90% where it seems to plateau, for validation accuracy (orange line), the model improves which shows moderate generalization and fluctuates between 75% and 76%. In the right plot, based on the training curve, the model without attention gets the training accuracy of around 85%, and the validation accuracy fluctuates around 70%, indicating less generalization compared with the attention model. In the case of the attention mechanism, it can be seen that the position significantly improves model performance especially the validation accuracy.

The figure 6 shows the training and validation loss curves for two RNN models: a focused and an unfocused one. The left plot is the training and validation loss for the RNN with attention where the training and validation loss both reduce with epochs and there is minimal overfitting. For instance, the training loss may go down from 1.4 to 0.2 and the validation loss going down from 1.2 to 0.3. On the right side, we can observe a similar graph, but without paying attention; after a specific number of epochs, the accuracy on the validation set begins to grow due to overfitting. For example, the validation loss may rise from 0.3 to 0.6, yet the training loss still goes on sinking. This comparison is aimed to show that the use of attention mechanisms in RNNs has advantages and can enhance the quality of the resulting model and reduce the risk of overfitting.

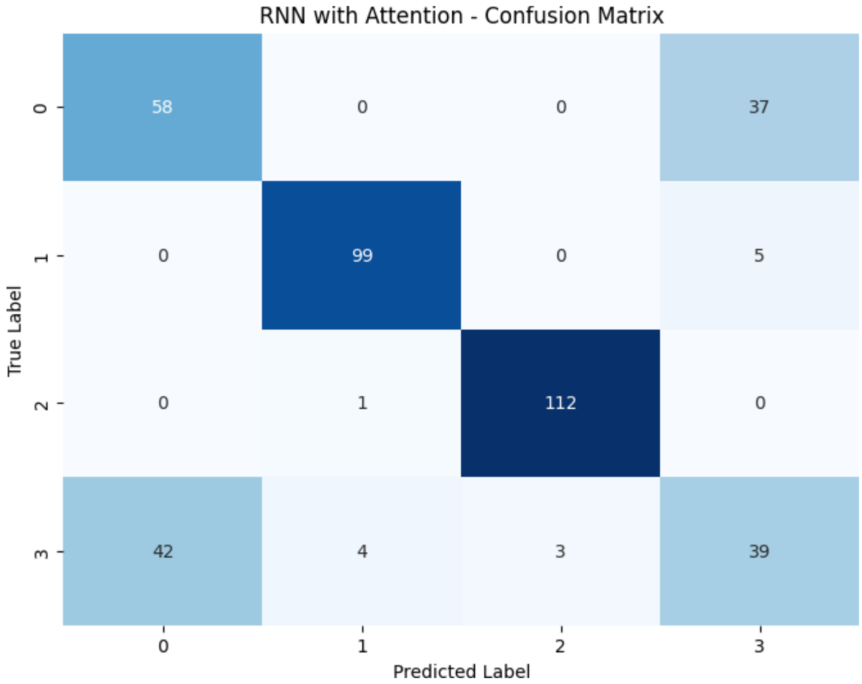


Figure 7a: Confusion matrix for RNN with attention

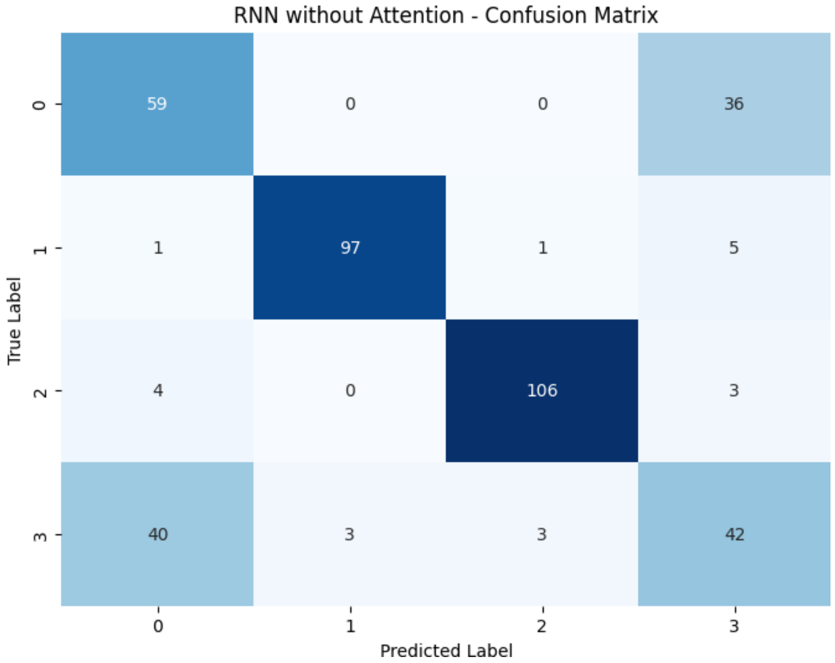


Figure 7b: Confusion matrix for RNN without attention

Figures 7a and 7b are the confusion matrices on multiclass COVID-19 detection using the RNN model with and without the attention mechanism, respectively. From figure 7a, we can see that the Model with attention has high correct classifications such as 99 and 112 of class 0 but also instances such 42 from class 3 misclassified as class 0. Fig 7b shows that even the model without attention has good accuracy, with values as high as 97 and 106, but incorrectly classifies 40 samples of class 3 to class 0. The generalization capability of the proposed p-RNN model is also better but the values are extremely low compared to that of CNN.

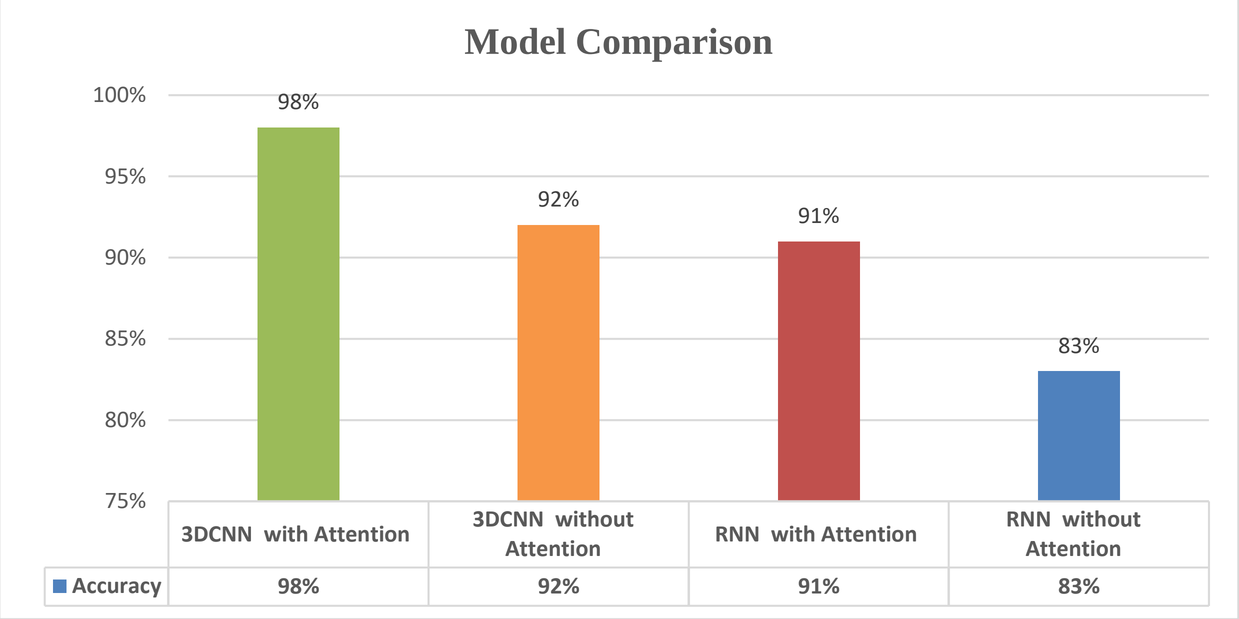


Figure 8: Overall performance analysis in terms of accuracy

The above graph analyzed the accuracy of performance of four deep learning models to distinguish different classes of COVID-19 from CT scans. They comprised 3D CNN with and without attention and 3D CNN with and without RNN attention. This deployment exhibited 3D CNNs with attention having the best accuracy of 98%, 3D CNNs without attention at 92%, RNNs with attention at 91% and RNNs without attention at 83%. These results show that the use of 3D CNNs and attention mechanisms is specifically helpful with this task, and the attention mechanism also appears to enhance the model's performance.

5 CONCLUSION

The rapid spread of COVID-19 has highlighted the critical need for accurate and efficient diagnostic tools. This study presents a novel approach for multiclass COVID-19 detection, combining the feature extraction power of an enhanced MobileNetV2, pretrained with a COVID-19 dataset, and the classification capabilities of attention-based 3D CNNs. By employing spatial-based attention to fine-tune the network parameters, the model effectively captures intricate patterns from diverse imaging datasets, achieving high accuracy, sensitivity, and specificity in COVID-19 detection tasks. Additionally, Grey Wolf Optimization (GWO) is applied to optimize the hyperparameters of the 3D CNN, ensuring more efficient convergence and improving model performance by mimicking the leadership hierarchy and hunting strategies of grey wolves. The proposed hybrid framework, enhanced with GWO, shows great potential for improving diagnostic accuracy in healthcare systems and aiding in better disease management strategies. Future work can focus on expanding the dataset for more comprehensive imaging coverage, ensuring the model's robustness across various populations and regions. Furthermore, integrating the framework with other diagnostic inputs, such as clinical symptoms and lab results, can enhance its performance. Developing real-time applications and user-friendly interfaces would make this approach more accessible for healthcare professionals. Extending this hybrid methodology to other infectious diseases could pave the way for advancements in medical diagnostics and treatment strategies.

REFERENCES

1. Jebril, N. (2020). World Health Organization declared a pandemic public health menace: a systematic review of the coronavirus disease 2019 "COVID-19". *Available at SSRN 3566298*.
2. Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K., ... & Drosten, C. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, *25*(3), 2000045.
3. Rubin, G. D., Ryerson, C. J., Haramati, L. B., Sverzellati, N., Kanne, J. P., Raoof, S., ... & Leung, A. N. (2020). The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society. *Radiology*, *296*(1), 172-180.
4. Suganyadevi, S., Seethalakshmi, V., & Balasamy, K. (2022). A review on deep learning in medical image analysis. *International Journal of Multimedia Information Retrieval*, *11*(1), 19- 38.

5. Jin, C., Chen, W., Cao, Y., Xu, Z., Tan, Z., Zhang, X., ... & Feng, J. (2020). Development and evaluation of an AI system for COVID-19.
6. Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., & Acharya, U. R. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Computers in biology and medicine*, *121*, 103792.
7. Khan, M. A., Hussain, N., Majid, A., Alhaisoni, M., Chan Bukhari, S. A., Kadry, S., ... & Zhang, Y. D. (2021). Classification of Positive COVID-19 CT Scans Using Deep Learning. *Computers, Materials & Continua*, *66*(3).
8. Amyar, A., Modzelewski, R., Li, H., & Ruan, S. (2020). Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation. *Computers in biology and medicine*, *126*, 104037.
9. 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).
10. Woan Ching, S. L., Lai, K. W., Chuah, J. H., Hasikin, K., Khalil, A., Qian, P., ... & Dhanalakshmi, S. (2022). Multiclass Convolution Neural Network for Classification of COVID- 19 CT Images. *Computational Intelligence & Neuroscience*.
11. Fan, Z., He, S., Ruan, S., Wang, X., & Li, H. (2021, February). Deep learning-based multi-class COVID-19 classification with x-ray images. In *Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling* (Vol. 11598, pp. 286-296). SPIE.
12. Khan, M. A., Alhaisoni, M., Tariq, U., Hussain, N., Majid, A., Damaševičius, R., & Maskeliūnas, R. (2021). COVID-19 case recognition from chest CT images by deep learning, entropy- controlled firefly optimization, and parallel feature fusion. *Sensors*, *21*(21), 7286.
13. Canayaz, M., Şehribanoğlu, S., Özdağ, R., & Demir, M. (2022). COVID-19 diagnosis on CT images with Bayes optimization-based deep neural networks and machine learning algorithms. *Neural Computing and Applications*, *34*(7), 5349-5365.
14. AbdElhamid, A. A., AbdElhalim, E., Mohamed, M. A., & Khalifa, F. (2022). Multi- classification of chest X-rays for COVID-19 diagnosis using deep learning algorithms. *Applied Sciences*, *12*(4), 2080.
15. Majid, A., Khan, M. A., Nam, Y., Tariq, U., Roy, S., Mostafa, R. R., & Sakr, R. H. (2021). COVID19 Classification Using CT Images via Ensembles of Deep Learning Models. *Computers, Materials & Continua*, *69*(1).
16. Hassan, A., Shahin, I., & Alsabek, M. B. (2020, November). COVID-19 detection system using recurrent neural networks. In *2020 International conference on communications, computing, cybersecurity, and informatics (CCCI)* (pp. 1-5). IEEE.
17. Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 4489-4497).
18. Motamed, S., Rogalla, P., & Khalvati, F. (2021). Data augmentation using Generative Adversarial Networks (GANs) for GAN-based detection of Pneumonia and COVID-19 in chest X-ray images. *Informatics in medicine unlocked*, *27*, 100779.
19. Gulakala, R., Markert, B., & Stoffel, M. (2022). Generative adversarial network based data augmentation for CNN based detection of Covid-19. *Scientific Reports*, *12*(1), 19186.
20. Xiang, Q., Wang, X., Li, R., Zhang, G., Lai, J., & Hu, Q. (2019, October). Fruit image classification based on Mobilenetv2 with transfer learning technique. In *Proceedings of the 3rd international conference on computer science and application engineering* (pp. 1-7).
21. Nguyen, H. (2020). Fast object detection framework based on mobilenetv2 architecture and enhanced feature pyramid. *J. Theor. Appl. Inf. Technol*, *98*(05).
22. Srinivasu, P. N., SivaSai, J. G., Ijaz, M. F., Bhoi, A. K., Kim, W., & Kang, J. J. (2021). Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. *Sensors*, *21*(8), 2852.
23. Ji, S., Xu, W., Yang, M., & Yu, K. (2012). 3D convolutional neural networks for human action recognition. *IEEE transactions on pattern analysis and machine intelligence*, *35*(1), 221-231.
24. Hara, K., Kataoka, H., & Satoh, Y. (2018). Can spatiotemporal 3d cnns retrace the history of 2d cnns and imagenet?. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 6546-6555).
25. Bello, I., Zoph, B., Vaswani, A., Shlens, J., & Le, Q. V. (2019). Attention augmented convolutional networks. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 3286-3295).
26. Qin, Z., Liu, Z., Guo, Q., & Zhu, P. (2022). 3D convolutional neural networks with hybrid attention mechanism for early diagnosis of Alzheimer’s disease. *Biomedical Signal Processing and Control*, *77*, 103828.
27. Elbaz, K., Shaban, W. M., Zhou, A., & Shen, S. L. (2023). Real time image-based air quality forecasts using a 3D-CNN approach with an attention mechanism. *Chemosphere*, *333*, 138867.
28. Dhruv, P., & Naskar, S. (2020). Image classification using convolutional neural network (CNN) and recurrent neural network (RNN): A review. *Machine learning and information processing: proceedings of ICMLIP 2019*, 367-381.
29. Mou, L., Ghamisi, P., & Zhu, X. X. (2017). Deep recurrent neural networks for hyperspectral image classification. *IEEE transactions on geoscience and remote sensing*, *55*(7), 3639-3655.
30. Agarwal, P., Nachappa, M. N., & Gautam, C. K. (2024, January). Multi-Scale Recurrent Neural Networks for Medical Image Classification. In *2024 International Conference on Optimization Computing and Wireless Communication (ICOCWC)* (pp. 1-6). IEEE.