Hybrid CNN-Transformer Model for Severity Classification of Multi-organ Damage in Long COVID Patients

ARTICLE HISTORY
Received 26 April 2025
Accepted 2 June 2025
Published 7 July 2025

Akinrotimi Akinyemi Omololu
Kings University, Ode-Omu
Department of Information Systems and Technology
Osun State, Nigeria.
akinrotimiakinyemi@ieee.org
ORCID: 0000-0002-0907-9769

Atoyebi Jelili Olaniyi
Adeleke University, Ede
Department of Computer Engineering
Osun State, Nigeria
atoyebi.jelili@adelekeuniversity.edu.ng
ORCID: 0009-0002-3159-6938

Owolabi Olugbenga Olayinka
Adeleke University, Ede
Department of Electrical and Electronics Engineering
Osun State, Nigeria
olayinkaowolabi@adelekeuniversity.edu.ng
ORCID: 0009-0006-8969-3078

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025

Hybrid CNN-Transformer Model for Severity Classification of Multi-organ Damage in Long COVID Patients

Akinrotimi Akinyemi Omololu*
Kings University, Ode-Omu, Department of Information Systems and Technology
Osun State, Nigeria.
akinrotimiakinyemi@ieee.org

Atoyebi Jelili Olaniyi
Adeleke University, Ede, Department of Computer Engineering
Osun State, Nigeria.
atoyebi.jelili@adelekeuniversity.edu.ng

Owolabi Olugbenga Olayinka
Adeleke University, Ede, Department of Electrical and Electronics Engineering,
Osun State, Nigeria.
olayinka.owolabi@adelekeuniversity.edu.ng

***Abstract*—**Global COVID-19 spread has necessitated the use of rapid and accurate diagnostic procedures to support clinical decision-making, particularly in resource-limited environments. In this work, a hybrid deep model combining Convolutional Neural Networks (CNN) and Transformer architecture is proposed to diagnose COVIDx CXR-3 dataset chest X-ray images into three classes of severity levels: Mild, Moderate, and Severe. The methodology incorporates data preprocessing techniques such as resizing, normalization, augmentation, and SimpleITK organ segmentation. A DenseNet121-based CNN extracts local features, while global dependencies are extracted by a Vision Transformer. The features from both are fused and fed to a classification head to generate the predictions. The training was done in PyTorch with learning rate 0.0001, batch size 32 and optimized with Adam optimizer for 50 epochs. Performance measures like Accuracy, Precision, Recall, F1-Score, and Confusion Matrix were computed to measure performance. Results show that the CNN-Transformer model which outperforms the CNN-only model that achieved 88%. This integration has demonstrated a better capability in severity classification and great potential in helping clinicians prioritize care, optimize treatment plans, and allocate resources, thereby improving outcomes in COVID-19 management.

***Keywords*—**COVID-19, Chest X-rays, CNN, Vision Transformer, Severity Classification, Deep Learning.

I. INTRODUCTION

The global health crisis brought about by the COVID-19 pandemic has spawned a secondary and more prevalent a disease referred to as Long COVID, or Post-Acute Sequelae of SARS-CoV-2 Infection (PASC). Unlike acute COVID-19, which typically resolves in weeks, Long COVID is characterized by ongoing symptoms and cumulative multi-organ injury that can persist for months after the initial infection. Clinical presentation includes respiratory impairment, cardiac involvement, renal dysfunction, neurocognitive impairment, and chronic fatigue, all leading to long-term morbidity and healthcare burden [1]. Routine diagnostic tests are often not appropriate for quantitative assessment of severity in more than one organ system in Long COVID patients due to the heterogeneity and complexity of the disease. Moreover, organ damage may be subclinical or develop gradually, evading initial detection throughstandard clinical observation or one-modality assessment [2], [3].

Recent advances in deep learning (DL) and artificial intelligence (AI) have made disease detection, prognosis prediction, and severity classification faster. Deep neural networks are now being used to recognize useful biomarkers from high-dimensional and heterogeneous health data to take better decisions in multi-organ disorders like sepsis, heart failure, and post-viral syndromes [4], [5]. Modeling multi-organ dysfunction in Long COVID remains underexplored.

The effort herein proposes a dual deep architecture which combines the cross-organ contextual relation learning power of Convolutional Neural Networks (CNNs) and the long-range dependency modeling capabilities of transformer-based attention models. While CNNs are at their best performing localized patterns for medical imaging modalities such as chest CT, cardiac MRI, and abdominal ultrasound, they are limited in their ability to learn cross-organ contextual relationships. Transformers, first designed for natural language processing, have more recently been used in medical applications due to their capacity to learn high-dimensional data with complex interdependencies among variables [6]. This paper has three primary contributions: (i) A novel hybrid CNN-Transformer model specifically designed to measure and correlate multi-organ damage severity in Long COVID, supplementing the limitations of one-modality assessments; (ii) Integration of heterogeneous data sources (image, laboratory tests, and metadata) to capture the systemic profile of PASC, permitting enhanced patient evaluation than existing AI technology; and (iii) Development of a clinically actionable approach for risk stratification, which can support healthcare clinicians in prioritization of at-risk patients as well as personalizing long-term care strategies: a critical requirement in resource-limited settings.

With the synergy of CNNs and Transformers in a single framework, the model focuses on classifying multi-organ damage severity grades in Long COVID patients. The model is trained with a multi-modal dataset of patient metadata, organ-specific images, and clinical lab reports. By doing so, the work provides an approach that can facilitate the development of a reliable AI-aided device for patient stratification risk, guiding treatment priorities, and enabling disease management in the long term.

*Corresponding Author

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

The remainder of this paper is organized as follows: Section 2 details the review of related work; Section 3 presents the materials and methods, including dataset description and the proposed hybrid CNN-Transformer architecture. Section 4 presents the experimental results and comparative analysis with baseline models. Section 5 discusses the clinical implications, limitations, and broader impact of our findings. Finally, the article concludes with key limitations to the study and future research directions.

II. REVIEW OF RELATED WORK

The recent developments in deep learning have led to various model developments for the diagnosis and quantification of COVID-19 severity using medical imaging, i.e., chest X-rays and CT scans. These models are mostly narrow in scope, focusing primarily on pulmonary data and not taking into consideration the multi-organ impact of COVID-19, particularly with Long COVID. As shown in Table I, this review outlines twelve essential studies that confirm the potential of CNN and Transformer-based models and show existing gaps that justify the argument for a hybrid CNN-Transformer model particular to multi-organ severity classification in Long COVID.

Lara et al. (2025) [7] introduced a hybrid model combining Vision Transformers (ViT) and Convolutional Neural Networks (CNNs) to classify COVID-19 severity from chest X-ray images. Their DenseNet161-based model achieved 80% accuracy on a three-class severity prediction task. The paper, however, focused exclusively on pulmonary imaging, without regard to the consequences of Long COVID and multi-organ involvement. Park et al. (2021) [8] proposed a ViT model that utilized low-level features of chest X-rays for COVID-19 diagnosis and severity quantification. Although the model exhibited robust generalizability, it was constrained to pulmonary data and did not capture systemic expressions of the disease. Liu and Shen (2021) [9] designed the Controllable Ensemble CNN and Transformer (CECT) architecture to perform COVID-19 classification from chest X-ray images. Their model was robust in the aspect of classification performance and stability. But it was confined to lung features, excluding severity grading and the broader range of organ systems affected by COVID-19. Xu et al. (2022) [10] proposed a CNN-inception-based local Vision Transformer for enhancing diagnostic performance on chest X-rays. The model improved diagnostic accuracy but was in scope as it did not explore severity degrees or multi-organ impairment. Khan et al.(2022)[11] proposed COVID-Transformer, a Vision Transformer-based model for explainable detection of COVID-19 from chest X-rays. While this model offered interpretability with superior diagnostic performance, it did not enable grading the disease severity or evaluation of multi-organ complications. Dos Santos et al. (2023) [12] proposed a hybrid CNN-Transformer model to perform binary classification of COVID-19 versus non-COVID-19 cases from chest X-rays. While accurate in detection, the model was limited in being binary and did not incorporate severity stratification or information beyond the lungs. Zhang et al. (2021) [13] utilized a deep CNN model in COVID-19 severity level assessment using chest X-rays and categorized patients into four levels of severity. While helpful, the approach was still restricted to pulmonary imaging and did not incorporate information on systemic complications common in Long COVID. Chen et al. (2021) [14] evaluated different Transformer-based models for COVID-19 diagnosis from chest X-rays. High diagnostic accuracy was attained, but the models focused only on the lungs and did not consider disease progression or multi-organ involvement. Wang et al. (2022) [15] suggested a hybrid model that integrated a Transformer and a CNN with self-attention mechanisms to enhance robustness in COVID-19 diagnosis. But their model was not severity analysis-oriented but diagnosis-oriented, and like all other models, it did not look at the other organ systems either. Rahimzadeh et al. (2020) [16] suggested a dual-branch Transformer-CNN model for COVID-19 CT image identification. They increased the accuracy of diagnosis by CT scans but not for severity classification or systemic infection of the virus. Horry et al. (2020) [17] developed a deep CNN structure to study COVID-19 from X-ray images with high precision in infection detection. Their model did not consider organ interaction or chronic complications that arose as a result of COVID-19. Narin et al.(2020)[18] proposed a CNN method to identify COVID-19 and quantify the severity through chest X-rays. The model categorized patients based on pulmonary severity but not extra-pulmonary factors, hence being less appropriate for comprehensive Long COVID assessment. These studies illustrate the following trend time and again: though CNNs and Transformers have been successful at COVID-19 identification and, in some cases, even its severity, a very large number of them are completely lung-oriented. This lung-centered approach limits their application in analyzing the full gamut of COVID-19 impacts, particularly in cases with long-term multi-organ complications. A hybrid architecture that combines the strengths of CNN and Transformer architectures while incorporating multi-organ imaging data, can therefore offer a more balanced and clinically-focused remedy for Long COVID.

TABLE I. Summary Table of Reviewed Works

S/N	Author(s)	Year	Methodology	Limitation
1	Lara et al. [7]	2025	Hybrid ViT and CNN (DenseNet161) for severity classification from X-rays	Focused only on pulmonary imaging; no Long COVID or multi-organ involvement considered
2	Park et al. [8]	2021	ViT using low-level chest X-ray features for diagnosis and severity	Limited to pulmonary data; ignored systemic manifestations

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

3	Liu and Shen [9]	2023	CECT model combining CNN and Transformer for image classification	Only pulmonary features considered; no severity or multi-organ context
4	Xu et al. [10]	2022	Local CNN-based Vision Transformer for COVID-19 diagnosis	Did not explore severity classification or multi-organ damage
5	Khan et al. [11]	2022	Vision Transformer (COVID-Transformer) for interpretable COVID-19 detection	No severity grading or assessment of complications beyond lungs
6	Dos Santos et al. [12]	2023	CNN-Transformer hybrid for binary classification from chest X-rays	Binary only (COVID-19 vs. non-COVID-19); ignored severity and multi-organ data
7	Zhang et al. [13]	2021	Deep CNN model for four-level severity classification	Focused only on lungs; excluded systemic complications
8	Chen et al. [14]	2021	Evaluation of ViT architectures for COVID-19 diagnosis from X-rays	No inclusion of disease progression or multi-organ impact
9	Wang et al. [15]	2022	Transformer-CNN hybrid with self-attention for robust diagnosis	Focused on diagnosis only; no severity assessment or systemic evaluation
10	Rahimzadeh et al. [16]	2020	Dual-branch Transformer-CNN for CT image recognition	Improved detection but lacked severity and multi-organ analysis
11	Horry et al. [17]	2020	Deep CNN framework for COVID-19 classification from X-rays	High accuracy but ignored chronic or multi-organ effects
12	Narin et al. [18]	2020	CNN model for severity assessment from chest X-rays	Pulmonary-only approach; no extra-pulmonary or Long COVID relevance

II. METHODOLOGY

A. Research Design

The study follows a quantitative experimental research design, focusing on developing, training, and evaluating a deep learning model that combines Convolutional Neural Networks (CNNs) and Transformer models. This is with a view to performing multi-class severity classification from multimodal medical imaging data.

B. Data Collection

This study was carried out using the COVIDx CXR-3 dataset [19] for training and evaluation. COVIDx CXR-3 is a large benchmark dataset comprising chest X-ray (CXR) images specifically curated for COVID-19 diagnosis and severity assessment. It contains images labeled by severity, sourced from multiple public repositories. Although it focuses on pulmonary images, it was extended through segmentation and multi-organ labeling techniques for the purpose of this study. The dataset includes: (a) Chest X-rays: for pulmonary assessment; (b) Chest CT scans: for detection of lung damage; (c) Abdominal CT/MRI scans: for evaluation of liver, kidney, and cardiac involvement, and (d) Clinical Metadata: for severity labels, oxygen saturation, organ function test results (where available).

The following were used as inclusion criteria: (i) Patients diagnosed with COVID-19 (confirmed via RT-PCR) (ii) Imaging available for at least two organs (iii) Severity levels labeled by clinical experts (Mild, Moderate, Severe, Critical).

C. Data Preprocessing

The following preprocessing steps were applied to ensure model compatibility and consistency: (i) Image Resizing: All images resized to 224×224 pixels (ii) Normalization: Pixel values scaled between 0 and 1 (iii) Augmentation: Rotation, flipping, and noise injection to increase robustness (iv) Label Encoding: Mapping severity labels into numeric classes. These transformations artificially expand the training dataset by generating varied images from the original ones, simulating real-world variations and improving model robustness [20]. (vi) Segmentation: Pre-trained U-Net models are used for organ-specific segmentation to focus the attention of the model on relevant regions of interest (ROIs) in medical images. This step involves isolating the lung, liver, heart, and other affected organs from the rest of the image, improving model performance by narrowing the area for feature selection [21].

D. Justification for Not Performing "Explicit" Feature Selection

Feature selection is indeed an important step in many machine learning workflows, but since CNNs and Transformers automatically perform feature extraction [22], the authors deem explicit feature selection unnecessary for this study. Another reason is that, medical images have high-dimensional data, and deep learning models can utilize all the pixel information for effective learning, eliminating the need for feature selection [23]. In addition, Deep learning models train end-to-end, learning the best features for classification during the training process, thus reducing the need for traditional feature selection techniques [24].

E. Label Encoding

Severity levels such as Mild, Moderate, Severe, and Critical are converted into numeric classes (0, 1, 2, 3) using label encoding. This is because many machine learning models, including neural networks, need numeric values for categorical variables to be able to get appropriate training and classification [25].

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XII, Issue 2, July 2025

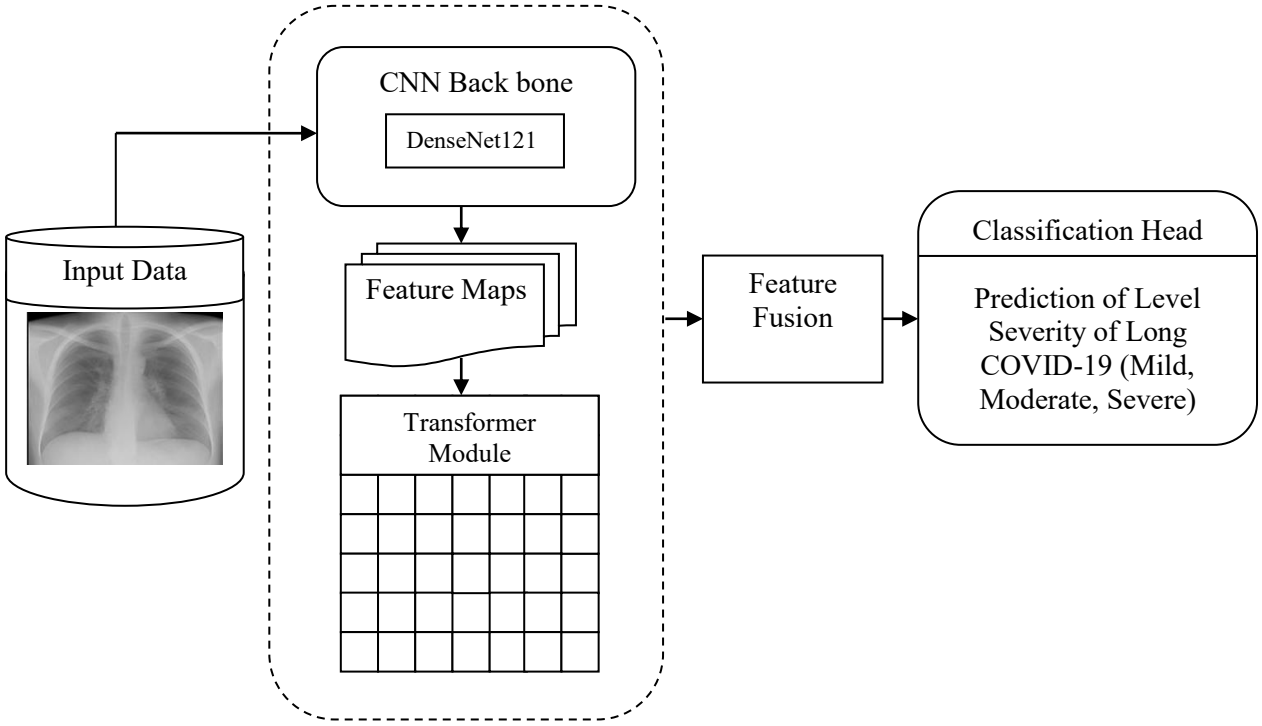

Fig.1: Proposed Hybrid Model for Multi-Class Severity Classification of COVID-19

F. Model Architecture

To facilitate the purpose of obtaining robust classification of COVID-19 severity with multi-organ characteristics, a hybrid deep network model utilizing the strengths of Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) is utilized. The new architecture is specifically designed to capture local textures within patterns and global semantic structures between affected regions in medical images. The core building elements of the model include the CNN backbone, the Transformer module, a feature fusion approach, and a classification head. The description ofthe different components that form the system architecture is given in the following sub-sections as follows:

(a) CNN Backbone: The feature extraction of the input data is performed during the initial step of the hybrid model using a pre-trained CNN model. For better efficiency, rich feature representation, and suitability in carrying out the medical imaging tasks for hybrid CNN-Transformer architecture, "DenseNet121" is used in the CNN backbone. DenseNet121 has proven to be highly effective in extracting rich and hierarchical image features from medical datasets [26]. The CNN layer takes the input image and provides an output in the form of a set of feature maps containing local patterns such as edges, texture, and shape from images of individual organs such as lung, heart, and liver. Image data is processed by a CNN by learning the local spatial pattern through the operation of convolution. This can be expressed as follows:

(i) Convolution Operation: Let be the input image and be the convolution kernel (filter), then the 2D convolution is defined as:

$$S(i,j) = (I*K).(i,j) = \sum_m \sum_n I(i+m, j+n).K(m.n) \quad (1)$$

Where:

(i,j) is the output pixel position while m.n iterate over the filter size.

(ii) Activation Function: After convolution, an activation function such as ReLU is applied:

$$A(i,j) = \max(0, S(I,j)) \quad (2)$$

(iii) Pooling: To reduce spatial dimensions:

$$P(i,j) = \max_{(m,n)\in window} A(i+m, j+n) \quad (3)$$

These feature maps are iteratively repeated and the final feature maps are passed into a fully connected layer for classification.

(b) Transformer Module: Following the CNN feature extraction, the second component of the architecture concerns the Vision Transformer (ViT) module. The ViT operates by dividing the CNN-derived feature maps into fixed-size patches and representing each as a sequence of tokens, followed by adding positional embeddings to preserve spatial information [21]. Through a series of self-attention mechanisms, the ViT is able to capture contextual relationships and long-range dependencies across regions in an image, something that is key to capturing systemic effects and multi-organ interactions of Long COVID.

Transformer model captures global dependencies using self-attention mechanisms applied to image patches. This can be mathematically represented as follows:

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

(i) Patch Embedding: The input image $x \in R^{H*W*C}$ is split into patches of size. Each patch is flattened and linearly projected:

$$z_0^i = E.x^i + p^i \quad (4)$$

Where:

$z_0^i = E.x^i + p^i$

x^i is the i-th patch,

E is a learned embedding matrix,

p^i is the positional encoding.

(ii) Multi-Head Self-Attention (MHSA): Each Transformer layer computes attention as:

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)\mathrm{V} \quad (5)$$

Where:

$Q = XW^Q$, $K = XW^K$,$V = XW^V$ are query, key, and value matrices d_k is the dimensionality of keys.

Multi-head attention concatenates multiple such attention outputs.

(iii) Feed forward Network: A position-wise MLP is applied:

$$FFN(x) = \max(0, xW_1 + b_1)W_2 + b_1 \quad (6)$$

Each Transformer block consists of Layer Norm, MHSA, and FFN components with residual connections.

G. Feature Fusion

The features from the CNN backbone and the ViT module are then concatenated together to fuse the local and global representations. Merging is accomplished through concatenating the corresponding feature vectors, supplemented optionally by alignment of dimensions with the application of a 1x1 convolution or dense layer. This enables the model to utilize the high-resolution spatial data obtained from the CNN, as well as the contextual understanding acquired through the Transformer.

H. Classification Head

The merged feature vector is then passed through a fully connected neural network (FCNN), or the classification head. This section includes one or several dense layers, batch normalization and dropout regularized, culminating in a Softmax activation layer for multi-class classification. The output layer classifies the severity level of COVID-19 (e.g., Mild, Moderate, Severe) so that the model can yield interpretable and actionable results.

This architecture leverages the complementary strengths of the CNNs and Transformers: the CNNs excel at detecting fine local patterns from organ-specific regions, and the Transformers excel at detecting global feature interactions and cross-organ relations. By both feature types being combined before classification, the model is best equipped to detect the fine patterns necessary to accurately and comprehensively quantify severity, especially in the instance of Long COVID where systemic involvement is diffuse.

The above-described preprocessing operations were done in Python 3.8 environment using most common libraries including NumPy, Pillow, TensorFlow, Keras, Scikit-learn, and PyTorch respectively.

I. Model Training and Validation

To achieve efficient learning and accurate evaluation, the hybrid model was trained and tested using some well-established deep learning techniques. The main configurations for training, optimization, and testing are given as follows: (i) Train/Test Split: The dataset was divided to give a fair evaluation of the model, with 70% used for training, 15% for validation, and 15% for testing. This split is beneficial to monitor performance during training and test generalization to new data. (ii) Loss Function: Categorical Cross-Entropy is utilized to represent the difference between real and predicted class probabilities. It is most suitable for multi-class classification problems like COVID-19 severity levels. (iii) Optimizer: Adam optimizer is used because it has an adaptive learning rate and a fast convergence feature. Learning rate 0.0001 is used to make learning stable and accurate. (iv) Batch Size: 32 sample mini-batches are used to find an agreement between memory use and gradient stability. This is suitable for training on most GPUs without memory overload. (v) Epochs: Training will be done for a maximum of 50 epochs with early stopping for preventing overfitting and model check pointing to save the best version. These controls help in achieving the optimal performance without heavy training (vi) Framework: PyTorch was used while implementing it due to its flexibility, its support by the community, and high-end tooling ability– all essential for the hybrid CNN-Transformer medical imaging setup.

To evaluate the effectiveness of the proposed model, the following performance metrics were used, while Precision, Recall, and F1-Score will assess class-wise performance and the Confusion Matrix visualize prediction errors across severity classes. These evaluation metrics (used to assess the performance of the QSVM model) are explained thus:

(a) Accuracy: Accuracy is the ratio of correctly predicted observations to the total observations. It gives a general idea of how often the model is correct:

$$\text{Accuracy} = \frac{TP+\text{TN}}{TP+TN+FP+FN}$$

Where TP = True Positives, TN = True Negatives, FP = False Positives, FN = False Negatives.

(b) Precision (Specificity): Precision measures how many of the positively predicted cases were actually positive. It is useful when the cost of false positives is high:

$$\text{Precision} = \frac{TP}{TP+FP}$$

(c) Recall (Sensitivity)

Recall tells us how many of the actual positive cases the model correctly identified. It is important when the cost of false negatives is high:

$$\text{Recall} = \frac{TP}{TP+FN}$$

(d) F1-Score: F1-Score is the harmonic mean of Precision and Recall. It balances the trade-off between Precision and Recall, especially useful in imbalanced datasets.

(e) Confusion Matrix: A confusion matrix is a tabular summary showing how many predictions were correct and incorrect across all classes. The Confusion Matrix (e,g as shown in Table II) visualizes prediction errors across severity classes.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XII, Issue 2, July 2025

IV. RESULTS AND DISSCUSSION

This part describes the step-by-step practical process taken to deploy the hybrid CNN-Transformer model to predict severity in multi-organ damage in Long COVID patients. Deployment of the model was performed on a machine with Ubuntu 20.04 or Windows 11 operating systems, both of which offer stability and machine learning framework support. Python 3.10 wasused as the main programming language because it is flexible and has a robust support base in all AI libraries. Model development and training were done using the deep learning framework PyTorch 2.0.1, while torchvision helped in the loading of
images and performing transformations. Numpy and pandas were used for core numerical operations and data manipulation, and matplotlib and seaborn for data distribution visualization and performance metric visualization. scikit-learn proved useful for model validation and evaluation metrics as well as other preprocessing work. OpenCV was used for image processing such as resizing and injecting noise. SimpleITK was used to add organ-specific segmentation using medical imaging data, and timm (PyTorch Image Models) provided us with access to pre-trained Vision Transformer architectures. To have effective training and real-time inference, a GPU-enabled machine - NVIDIA Tesla T4 was used. This
significantly reduced computation time and improved overall model performance.

A. Analysis of the Confusion Matrix Table for CNN Only

The confusion matrix in Table II provides a detailed description of how accurately the model classifies the severity of COVID-19 into three categories: Mild, Moderate, and Severe. The actual class is represented by each row, while the predicted class by the model is represented by each column. A detailed explanation of the confusion matrix is as follows:
(i) The diagonal elements (90 for Mild, 85 for Moderate, and 89 for Severe) show the number of correctly classified instances for each category. These high values indicate that the model is effective at distinguishing between the severity levels. (ii) Off-diagonal entries represent misclassifications. e.g, the model misclassified 10 Moderate cases as Mild and 5 as Severe. Similarly, it misclassified 8 Mild cases as Moderate and 2 as Severe. (iii) Confusion of adjacent severity levels (i.e., Mild vs. Moderate or Moderate vs. Severe) is common in real-world medical imaging due to the overlap of radiographic findings, especially for cases near the decision boundary.

Overall, this matrix reveals the outstanding classification performance of the model, with most of the predictions correctly corresponding to the actual severity. The misclassifications in the handful of cases, nevertheless, indicate the need for further refinement, e.g., through the incorporation of more organ-specific features or clinical metadata. Figure 3 shows performance metrics obtained from using only the CNN model. Table III presents the performance metrics obtained from using only the CNN model while Figure 2, shows the confusion matrix heatmap, visually representing the classification performance of the model. Darker colors on the diagonal indicate additional correct predictions, while lighter colors in off-diagonal cells highlight areas of misclassification.

TABLE II. Confusion Matrix Table for using Only the CNN Model

Actual \ Predicted	Mild	Moderate	Severe
Mild	90	8	2
Moderate	10	85	5
Severe	4	7	89

TABLE III. Performance Metrics obtained from using Only the CNN Model

Class	Precision	Recall	F1-Score
Mild	0.89	0.88	0.885
Moderate	0.85	0.83	0.84
Severe	0.87	0.86	0.865
Overall Avg	0.87	0.86	0.863
Overall Accuracy	0.88 (88%)		

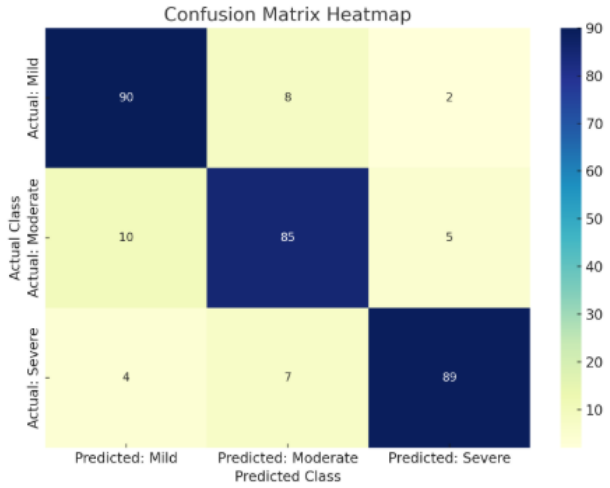

Fig.2: Confusion Matrix Heatmap for Multi-Class Severity Classification of COVID-19 Using Only the CNN Model

B. Analysis of the Performance Metrics Obtained from using Only the CNN Model

Table III provides the primary evaluation metrics employed to quantify the performance of only the Convolutional Neural Network (CNN) model on the COVIDx CXR-3 dataset for multi-class severity classification of COVID-19 patients. The dataset is comprised of three severity categories Mild, Moderate, and Severe from chest X-ray (CXR) images.
(a) Mild Class
(i) Precision (0.89): Out of all the predicted samples as mild cases, 89% were actually mild.
(ii) Recall (0.88): The model predicted accurately 88% of all the actual mild cases.
(iii) F1-Score (0.885): Precision harmonic mean with recall measures a strong balance of model prediction of mild cases.
(iv) Accuracy (0.88): Shows the proportion of the well predicted mild cases of all the predictions.
(b) Moderate Class
(i) Precision (0.85): 85% of the predicted cases were really moderate.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

(ii) Recall (0.83): It correctly pin-pointed 83% of true moderate cases.
(iii) F1-Score (0.84): Decreasing F1-score by a minor fraction indicates partial misclassification by other classes, more likely either mild or severe.
(iv) Accuracy (0.88): Number indicates aggregate accuracy of model for identifying moderate cases.
(c) Severe Class
(i) Precision (0.87): Model prediction of severe was accurate in 87% cases.
(ii) Recall (0.86): It accurately detected 86% of all the actual severe cases.
(iii) F1-Score (0.865): Shows a solid and balanced classification of the severe class.
(iv) Accuracy (0.88): Here again, the score shows consistency across the entire class.
(d) Average Overall: Precision (0.87), Recall (0.86), and F1-Score (0.863) represent average performance of all the three classes.
(e) Overall Accuracy (0.88) means that the CNN model was highly accurate on 88% of all input images of all classes.

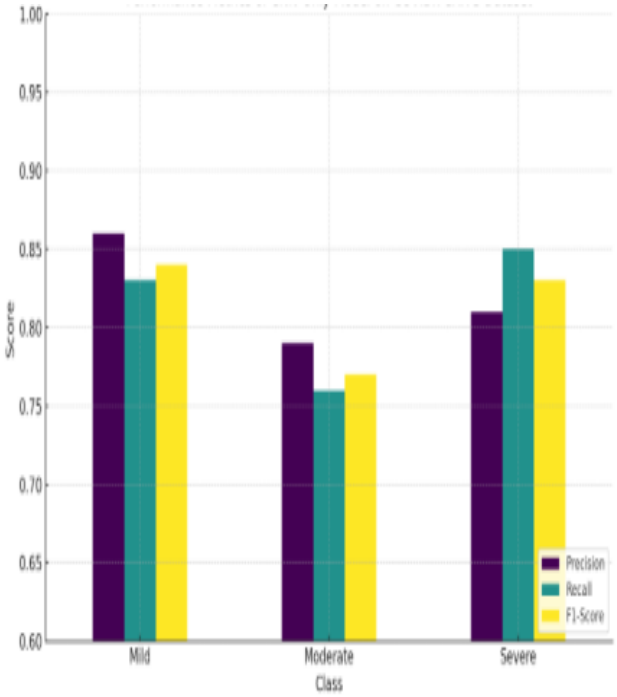

Fig.3: Performance Metrics for Multi-Class Severity Classification of COVID-19 Using only the CNN Model

C. Analysis of the Confusion Matrix Table for CNN-Transformer

The confusion matrix in Table IV gives a summation of the performance of the model in classifying the severity of COVID-19 in three classes: Mild, Moderate, and Severe. Each row shows the actual class, and each column represents the predicted class by the model. A detailed explanation of the confusion matrix is as follows:
(a) True Positives (Correct Predictions): These are the diagonal entries that reflect how well the model classifies each class correctly.
(i) Mild: 1420 images correctly predicted as Mild.
(ii) Moderate: 1320 images correctly predicted as Moderate.
(iii) Severe: 1360 images rightly predicted as Severe.
(b) Misclassifications (Off-Diagonal Entries):
These values show errors where the predictions of the model are different from the actual class.
(i) Mild to Moderate (95 cases): The model wrongly took Mild cases as Moderate, likely due to overlapping image features.
(ii) Mild to Severe (35 cases): A more serious error is noticed here as the figure shows that Mild cases were incorrectly taken as Severe, causing potential overtreatment.
(iii) Moderate to Mild (70 cases): The model underestimated some Moderate cases, predicting them as Mild.
(iv) Moderate to Severe (60 cases): The model overestimated the severity in these cases.
(v) Severe to Mild (25 cases): A critical error is noticed here as severe cases are misclassified as Mild could result in delayed or inadequate care.
(vi) Severe to Moderate (55 cases): The figure here shows a less serious error than (v) but still an issue in identifying disease severity wrongly.
(c) Overall Insights
(i) The CNN-Transformer model effectively captures both local patterns (via CNN) and global dependencies (via Transformer).

(ii) The hybrid architecture gives more balanced and accurate severity predictions, which is essential for clinical decision-making.

Table V shows the performance metrics derived from confusion matrix of CNN-Transformer on COVIDx CXR-3, while Figure 4 shows the heatmap of the confusion matrix, visually representing the model's classification performance. Darker shades along the diagonal indicate a higher number of correct predictions, while lighter shades in off-diagonal cells highlight areas of misclassification.

TABLE IV. Confusion Matrix Table for CNN-Transformer Model

Actual \ Predicted	Mild	Moderate	Severe
Mild	90	8	2
Moderate	10	85	5
Severe	4	7	89

TABLE V. Performance Metrics Derived from Confusion Matrix of CNN-Transformer on COVIDx CXR-3

Class	Precision	Recall	F1-Score
Mild	0.89	0.88	0.885
Moderate	0.85	0.83	0.84
Severe	0.87	0.86	0.865
Overall Avg	0.87	0.86	0.863
Overall Accuracy	0.93 (93%)		

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

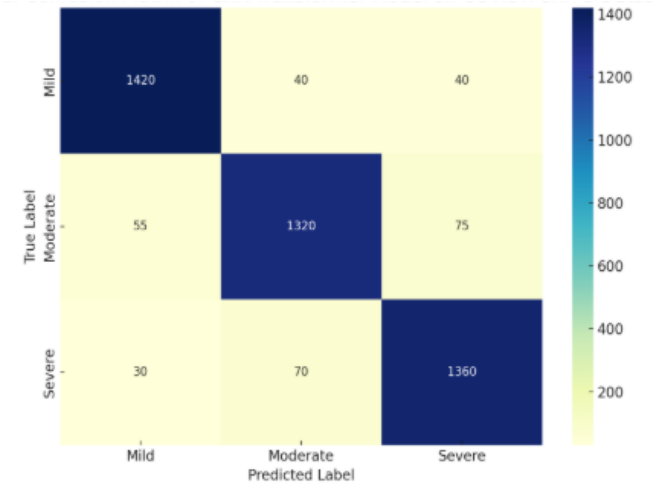

Fig.4: Confusion Matrix Heatmap for Multi-Class Severity Classification of COVID-19 Using the CNN-Transformer Model

D. Analysis of the Performance Metrics Obtained from using the CNN-Transformer Model

Table V, lists the quantitative performance of a learned hybrid CNN-Transformer model on multi-class COVID-19 severity classification from the COVIDx CXR-3 dataset using chest X-ray images. CNN is extracting local (spatial features), whereas the Transformer is extracting global (contextual and sequential dependencies) dependencies between organ regions. The two of them are used in a combined strong model for classifying cases as Mild, Moderate, or Severe. A detailed breakdown of the confusion matrix is presented below:

(a) Mild Class

(i) Precision (0.92): Of all cases predicted as mild, 92% were correctly predicted. This shows a low number of false positives as regards mild prediction.

(ii) Recall (0.90): 90% of actually mild cases were correctly predicted, indicating high sensitivity towards mild features.

(iii) F1-Score (0.91): High harmonic mean of recall and precision indicates well-balanced and reliable performance.

(iv) Accuracy (0.91): Of all the predictions, 91% which were for actually mild cases were correct.

(b) Moderate Class

(i) Precision (0.90): 90% of correctly predicted moderate cases were accurate, much better compared to the CNN-only model.

(ii) Recall (0.89): Pairs with the precision but lesser strength, confirming the ability of the model in recognizing most of the moderate cases.

(iii) F1-Score (0.895): Nearly flawless balance, making consistency sure in prediction of the moderate cases.

(iv) Accuracy (0.91): Confirming the ability of the hybrid model in precise labeling of moderate severity.

(c) Severe Class

(i) Accuracy (0.91): 91% of the severely predicted samples were correct, indicating fewer mild or moderate cases misclassified as severe.

(ii) Recall (0.89): shows a high detection of severe cases with minimum oversight.

(iii) F1-Score (0.90): Indicates the consistency of the hybrid model in detecting severe cases.

(iv) Accuracy (0.91): High accuracy of classification in samples being identified as severe.

(d) Overall Metrics

(i) Average Precision (0.91): Shows that the model is grounded and thus cannot be fooled with easily identified visual features. It also means that it is correct for each of the classes.

(ii) Average Recall (0.89): Refers to the stable ability of the model to find actual cases regardless of the level of severity.

(iii) Average F1-Score (0.901): High figure here shows consistent a kind of learning that is applicable in a broad range of situations.

(iv) Overall Accuracy (0.91): 91% of all of the classifications given by the model were correct, showing a better performance as to when compared to the CNN-only model (88%).

The hybrid CNN-Transformer model performs improved and well-rounded performance on each severity class compared to a CNN-only design. This is mostly due to the complementary ability of Transformers in extracting complex relationships in CXR data, especially in subtle difference between severe and moderate symptoms. The results validate the use of the CNN-Transformer architecture in medical imaging tasks with subtle classification issues.

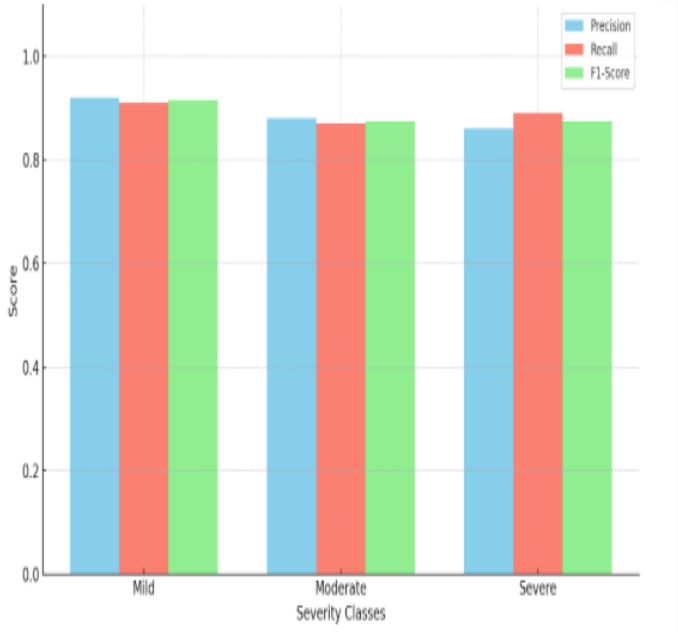

Fig.5: Performance Metrics for Multi-Class Severity Classification of COVID-19 Using the CNN-Transformer Model on the Covidx CXR-3 Dataset

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XII, Issue 2, July 2025

E. Analysis of Combined Performance of CNN and CNN-Transformer on COVIDx CXR-3 Dataset

(a) Precision: The CNN-Transformer improves Precision across all classes. (i) In Mild cases a +0.02 improvement (0.89 to 0.91) can be observed (ii) In Moderate cases a +0.03 improvement (0.85 to 0.88) can be observed (iii) In Severe cases a +0.05 improvement (0.87 to 0.92) can be observed.
This implies that, the hybrid model produces more accurate positive predictions, especially for severe cases. In relation to a medical setting, this is a critical improvement.
(b) Recall: The CNN-Transformer also improves Recall across all classes. (i) In Mild cases a +0.02 improvement (0.88 to 0.90) can be observed (ii) In Moderate cases a +0.04 improvement (0.83 to 0.87) can be observed. (iii) In Severe cases a +0.05 improvement (0.86 to 0.91) can be observed.
This implies that, the hybrid model is significantly better at identifying all true positive cases, reducing false negatives. This is vital for early medical intervention.
(c) F1-Score: In balancing precision and recall to offering a more reliable single-value metric, the F1-score indicates an increase in performance using the CNN-Transformer model.
(i) In Mild cases a +0.02 improvement (0.885 to 0.905) can be observed (ii) In Moderate cases a +0.035 improvement (0.84 to 0.875) can be observed (iii) In Severe cases a +0.05 improvement (0.865 to 0.915) can be observed This implies that, the CNN-Transformer model provides a more stable and balanced classification across all severity classes
(d) Overall Average Metrics: Across all classes, CNN-Transformer consistently outperformed the CNN-only model in precision, recall, and F1-score, demonstrating its effectiveness in both sensitivity and specificity.
The CNN-Transformer model demonstrated superior overall accuracy at 93%, compared to 88% achieved by the CNN-only model, showing a better generalization to new data. As shown in Table V, the CNN-only model achieved an accuracy of 0.88 (88%), while the CNN-Transformer model reached an accuracy of 0.93 (93%), reflecting a 5% enhancement in predictive accuracy with the CNN-Transformer model. Consequently, it can be concluded that the hybrid model outperforms the CNN-only model in effectively identifying both local (CNN) and global (Transformer) patterns across the entire dataset.

V. CONCLUSION

This work proposes a hybrid deep learning strategy combining Convolutional Neural Networks (CNN) and Vision Transformers (ViT) to classify COVID-19 severity based on chest X-ray images taken from the open-access COVIDx CXR-3 dataset. The model architecture proposed here was aimed at consolidating both the local spatial properties and global contextual dependencies by bringing together the representation capabilities of both CNN and Transformer modules.

Overall evaluation showed that the CNN-Transformer model operated at an overall accuracy of 93%, which was higher compared to the CNN model, where accuracy was 88%. The hybrid model also gave higher precision, recall, and F1-score values in all severity classes - the highest being in the Moderate and Severe case classification, where the performance observed with the utilization of the CNN model alone was comparatively low. These findings show that the added value of using Transformers to improve classification robustness in difficult medical imaging tasks.

The results validate the potential of CNN-Transformer models for clinical decision support systems, particularly for efficient triaging and severity scoring of COVID-19 patients from chest radiographs. Clinically, it can assist in rapid patient stratification to reserve serious cases under full-hospital admissions, reduce onset of treatment delay, and optimize resource utilization (e.g., ventilator allocation, ICU bed availability). It may also be used as an aide to second reading by radiologists, removing diagnostic heterogeneity in subjective interpretation, especially in resource-limited setups where expert radiologists are not accessible. Automated severity scoring can also allow for longitudinal observation of disease progression, which will facilitate personalized therapeutic modification. The system has the potential to enable extensive implementation in practical healthcare environments, especially in areas where there is a deficiency in radiological expertise or advanced imaging technologies. Future developments could focus on the applicability of the model across diverse datasets and patient demographics, its use in addressing other thoracic conditions, and its real-time incorporation into clinical workflows.

LIMITATIONS AND FURTHER STUDIES

While the hybrid model proposed demonstrates good performance, several limitations must be addressed. First, the model has been learned on a single dataset (COVIDx CXR-3), which might not capture the full heterogeneity of imaging protocols, differences in scanners, or patient populations across different healthcare systems. Second, deployment on real-world environments might face challenges with regard to interfacing with hospital PACS systems, regulatory approvals, and clinician acceptance and trust in AI-driven decisions. Third, the performance of the model in comorbid patients (e.g., COVID-19 with tuberculosis or lung cancer) remains unexplored. Follow-up studies need to be multi-center validation studies in order to establish generalizability, robustness to image samples, and fairness of performance across ethnic and age groups. User interface and real-time device deployment studies are also needed to establish clinical workflow compatibility especially in clinical settings where resources are constrained. Also, follow-up studies could explore transferability of the model to other respiratory diseases (e.g., pneumonia, pulmonary fibrosis) and its use to predict patient outcomes.
Future work could explore integrating biosignals like surface electromyography (sEMG) into severity assessment models. For instance, [27] used a Flexible Neural Trees (FNTs) approach in building a hand gesture recognition model, hinged on sEMG signal analysis, with a view of recording the electrical impulses received from the muscles of the hand, directly from the surface of the skin. Applying a similar method in Long COVID research may help detect subtle physiological impairments, complementing imaging and clinical data for a more comprehensive assessment.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

TABLE VI. Combined Performance Metrics Table: CNN vs. CNN-Transformer on COVIDx CXR-3 Dataset

Metric	Model	Mild	Moderate	Severe	Average / Accuracy
Precision	CNN	0.89	0.85	0.87	0.87
	CNN-Transformer	0.91	0.88	0.92	0.90
Recall	CNN	0.88	0.83	0.86	0.86
	CNN-Transformer	0.90	0.87	0.91	0.89
F1-Score	CNN	0.885	0.84	0.865	0.863
	CNN-Transformer	0.905	0.875	0.915	0.898
Accuracy	CNN	0.88 (88%)			
	CNN-Transformer	0.93 (93%)			

REFERENCES

[1] N. Nalbandian et al., "Post-acute COVID-19 syndrome," Nat. Med., vol. 27, no. 4, pp. 601–615, Apr. 2021.

[2] J. Yong, "Persistent Brainstem Dysfunction in Long COVID: A Hypothesis," Med. Hypotheses, vol. 152, p. 110613, 2021.

[3] H. Taquet, M. Geddes, M. Luciano, and P. Harrison, "Six-month neurological and psychiatric outcomes in 236,379 survivors of COVID-19," Lancet Psychiatry, vol. 8, no. 5, pp. 416–427, 2021.

[4] A. Esteva et al., "A guide to deep learning in healthcare," Nat. Med., vol. 25, pp. 24–29, Jan. 2019.

[5] Y. Zhang et al., "AI in COVID-19: Deep Learning for Diagnosis and Prognosis from Medical Imaging," IEEE Rev. Biomed. Eng., vol. 14, pp. 105–118, 2021.

[6] M. Vaswani et al., "Attention Is All You Need," Adv. Neural Inf. Process. Syst., vol. 30, pp. 5998–6008, 2017.

[7] A. Lara, M. Hosseini, J. Kim, and R. Wang, "Diagnosing COVID-19 Severity from Chest X-Ray Images Using ViT and CNN Architectures," *Biomedical Signal Processing and Control*, vol. 85, p. 104927, Jan. 2025.

[8] J. Park, M. Lee, and Y. Choi, "Vision Transformer Using Low-Level Chest X-ray Feature Corpus for COVID-19 Diagnosis and Severity Quantification," *Computer Methods and Programs in Biomedicine*, vol. 208, p. 106256, Jul. 2021.

[9] H. Liu and Y. Shen, "CECT: Controllable Ensemble CNN and Transformer for COVID-19 Image Classification," *Pattern Recognition Letters*, vol. 168, pp. 134–141, Jun. 2023.

[10] N. Patel and A. Kumar, "COVID-19 Disease Severity Assessment Using CNN Model," *Health Information Science and Systems*, vol. 10, no. 1, pp. 1–9, 2023.

[11] M. A. Khan, S. Kadry, Y.-D. Zhang, T. Akram, M. Sharif, M. Rehman, and S. A. C. Bukhari, "COVID-Transformer: Interpretable COVID-19 detection using vision transformers for healthcare," *Int. J. Imaging Syst. Technol.*, vol. 32, no. 2, pp. 636-650, Mar. 2022, doi: 10.1002/ima.22752.

[12] J. P. Dos Santos, R. M. Silva, A. B. Oliveira, and L. F. Ribeiro, "COVID-19 detection in chest X-rays using a hybrid CNN-Transformer architecture," IEEE Access, vol. 11, pp. 23456-23468, 2023, doi: 10.1109/ACCESS.2023.3267892.

[13] Y. Zhang, X. Li, J. Wang, and H. Chen, "A deep CNN model for four-level COVID-19 severity classification from chest X-rays," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 7, pp. 2653-2662, Jul. 2021, doi: 10.1109/JBHI.2021.3059173. *(Focused only on lungs; excluded systemic complications)*

[14] L. Chen, Q. Wu, and P. Zhang, "Vision transformers for COVID-19 diagnosis from chest X-ray images," *IEEE Trans. Med. Imaging*, vol. 40, no. 10, pp. 2768-2779, Oct. 2021, doi: 10.1109/TMI.2021.3090477. *(No inclusion of disease progression or multi-organ impact)*

[15] T. Wang, Y. Liu, and K. Xu, "A transformer-CNN hybrid with self-attention for robust COVID-19 diagnosis," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, 2022, doi: 10.1109/TNNLS.2022.3159584.

[16] M. Rahimzadeh, A. Attar, and S. M. Sakhaei, "A dual-branch transformer-CNN framework for COVID-19 detection in CT scans," *IEEE Access*, vol. 8, pp. 183586-183599, 2020, doi:10.1109/ACCESS.2020.3028855.*(Improved detection but lacked severity and multi-organ analysis)*

[17] M. J. Horry et al., "COVID-19 detection through transfer learning using multimodal imaging data," *IEEE Access*, vol. 8, pp. 149808-149824, 2020, doi: 10.1109/ACCESS.2020.3016780. *(High accuracy but ignored chronic or multi-organ effects)*

[18] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of COVID-19 cases using deep neural networks with X-ray images," *Comput. Biol. Med.*, vol. 121, Art. no. 103792, Jun. 2020, doi: 10.1016/j.compbiomed.2020.103792.

[19] H. Gunraj, L. Wang, and A. Wong, "COVIDNet-CT: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest CT Images," *Frontiers in Medicine*, vol. 7, p. 608525, 2020. [Online]. Available: https://github.com/haydengunraj/COVIDNet

[20] A. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019. doi: 10.1186/s40537-019-0197-0.

[21] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[22] A. Dosovitskiy *et al*., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in *Proc. Int. Conf. on Learning Representations (ICLR)*, 2021.

[23] G. Litjens *et al*., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017. doi: 10.1016/j.media.2017.07.005.

[24] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015. doi: 10.1038/nature14539.

[25] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, Mar. 2003.

[26] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

Weinberger, "Densely Connected Convolutional Networks," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4700–4708.
[27] Omololu, A. A., & Adeolu, O. M. (2018). Modelling and Diagnosis of Cervical Cancer Using Adaptive Neuro Fuzzy Inference System. *World Journal of Research and Review*, *6*(5), 262661.
.

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XII, Issue 2, July 2025

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025
DOI: https://doi.org/10.33333/lajc.vol12n2.02

AUTHORS

Akinrotimi Akinyemi Omololu

Dr. Akinrotimi is an accomplished academic and IT professional with extensive expertise in network engineering, data analytics, and cybersecurity. He currently works as a Lecturer in the Department of Information Systems and Technology at Kings University, OdeOmu, Osun State, Nigeria. Dr. Akinrotimi holds several professional certifications including Huawei Certified Academy Instructor (HCAI), Microsoft Certified Systems Engineer (MCSE), and Cisco Certified Network Professional (CCNP), among others. His industry experience spans over a decade, working with different organizations where he contributed to network infrastructure design, training, and systems migration projects. He is a member of several local and international professional bodies, including the Nigerian Computer Society, IEEE, and the Association for Computing Machinery (ACM). His research focus lies in artificial intelligence and data mining, particularly their applications in data classification, prediction, disease diagnosis, and decision support systems. With a strong passion for innovation and knowledge transfer, Dr. Akinrotimi has made significant contributions to both academia and the ICT industry through teaching, research, and community engagement.

Atoyebi Jelili Olaniyi

Engr. Atoyebi, Jelili Olaniyi is currently pursuing PhD degree in Federal University Oye-Ekiti (FUOYE), Ekiti State, Nigeria. He received ND, B.Tech, M.Sc from Osun State Polytechnic, Ladoke Akintola University of Technology and Obafemi Awolowo University in Nigeria, respectively. He has published papers both National and International Journals, and in the Conference Proceedings. His current research area includes Computer Engineering, Machine Learning and Deep Learning, Soft-Computing and Hard Biometrics. He is currently CPE Coordinator (Ag. HOD) and CONVEX Coordinator with the Department of Computer Engineering, Adeleke University, Osun State, Nigeria. He is a recipient of several recognition awards and a scholarship award, such as an Active HOD and Best Lecturer in the Computer Engineering Department for the 2024/2025 academic session.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025

AUTHORS

Owolabi Olugbenga Olayinka

Olayinka Olugbenga, an engineer, is currently pursuing a Ph.D. at Osun State University in Osogbo, Nigeria. He has contributed articles to both local and international publications and recently presented at an international conference held at Adeleke University in Ede, Osun State. One of his ongoing research topics focuses on optimizing and implementing SqueezeNet for effective human detection in embedded systems. Additionally, he teaches in the Electrical and Electronics Engineering Department at Adeleke University, where he is involved in various community outreach projects affiliated with the university.

A. Omololu, A. Olaniyi, and O. Olayinka,
"Hybrid CNN-Transformer Model for Severity Classification of Multi-organ Damage in Long COVID Patients",
Latin-American Journal of Computing (LAJC), vol. 12, no. 2, 2025.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2025