information

MDPI

Article

A Framework for Integrating Deep Learning and Symbolic AI Towards an Explainable Hybrid Model for the Detection of COVID-19 Using Computerized Tomography Scans

Vengai Musanga [1,*], Serestina Viriri [1,2] and Colin Chibaya [3]

[1] School of Mathematics, Statistics & Computer Science, University of KwaZulu Natal, Durban 4000, South Africa; viriris@ukzn.ac.za
[2] Department of Computer Science, Centre for Augmented Intelligence and Data Science, School of Computing, University of South Africa, Johannesburg 1709, South Africa
[3] Department of Computer Science, Data Science and Information Technology, Sol Plaatje University, Kimberly 8300, South Africa; colin.chibaya@spu.ac.za
* Correspondence: 223152680@stu.ukzn.ac.za

Abstract: The integration of Deep Learning and Symbolic Artificial Intelligence (AI) offers a promising hybrid framework for enhancing diagnostic accuracy and explainability in critical applications such as COVID-19 detection using computerized tomography (CT) scans. This study proposes a novel hybrid AI model that leverages the strengths of both approaches: the automated feature extraction and classification capabilities of Deep Learning and the logical reasoning and interpretability of Symbolic AI. Key components of the model include the adaptive deformable module, which improves spatial feature extraction by addressing variations in lung anatomy, and the attention-based encoder, which enhances feature saliency by focusing on critical regions within CT scans. Experimental validation using performance metrics such as F1-score, accuracy, precision, and recall demonstrates the model's significant improvement over baseline configurations, achieving near-perfect accuracy (99.16%) and F1-score (0.9916). This hybrid AI framework not only achieves state-of-the-art diagnostic performance but also ensures interpretability through its symbolic reasoning layer, facilitating its adoption in healthcare settings. The findings underscore the potential of combining advanced machine learning techniques with symbolic approaches to create robust and transparent AI systems for critical medical applications.

Keywords: deep learning; symbolic artificial intelligence; hybrid artificial intelligence; COVID-19; computerized tomography scans

Academic Editor: Birgitta Dresp-Langley

Received: 18 February 2025
Revised: 1 March 2025
Accepted: 5 March 2025
Published: 7 March 2025

Citation: Musanga, V.; Viriri, S.; Chibaya, C. A Framework for Integrating Deep Learning and Symbolic AI Towards an Explainable Hybrid Model for the Detection of COVID-19 Using Computerized Tomography Scans. *Information* **2025**, *16*, 208. https://doi.org/10.3390/info16030208

1. Introduction

Deep Learning is a subset of machine learning that utilizes artificial neural networks with multiple layers to learn from large amounts of data [1]. It aims to automatically learn representations of data through the use of complex hierarchical architectures without the need for explicit feature engineering. Symbolic AI, also known as classical AI or symbolic reasoning, is an approach to artificial intelligence that focuses on the use of symbols and rules to represent and manipulate knowledge in a structured manner [2]. While Symbolic AI has been foundational in areas such as expert systems and automated reasoning, its limitations in handling uncertainty and scalability have led to the rise of other AI approaches like machine learning and deep learning. Figure 1 presents a framework for image classification using Deep Learning methods, illustrating the key components and processes involved in automated image analysis.

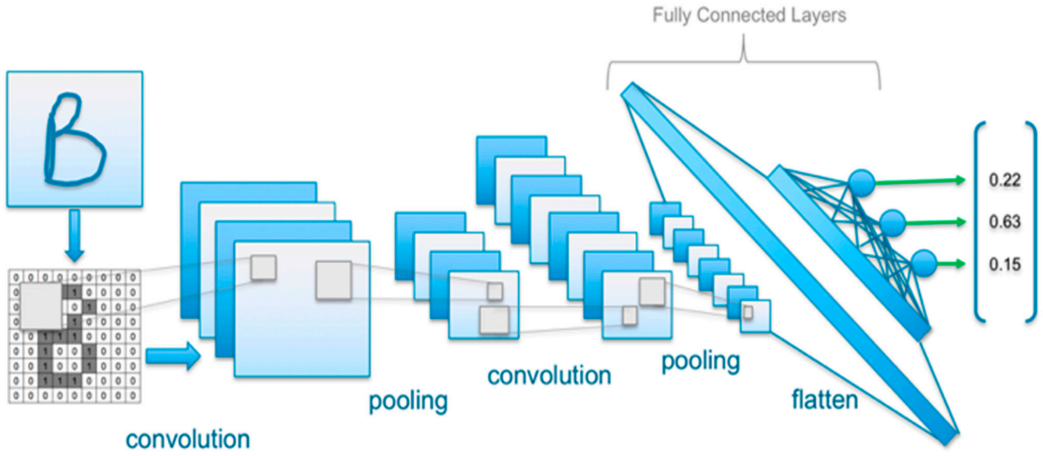

Figure 1. A framework for image classification using Deep Learning methods.

However, Symbolic AI remains relevant in domains where explicit representation and manipulation of knowledge are critical, such as natural language processing and symbolic reasoning tasks. Figure 2 illustrates a framework for a Symbolic AI system, showcasing the structured approach used in symbolic reasoning and rule-based knowledge representation.

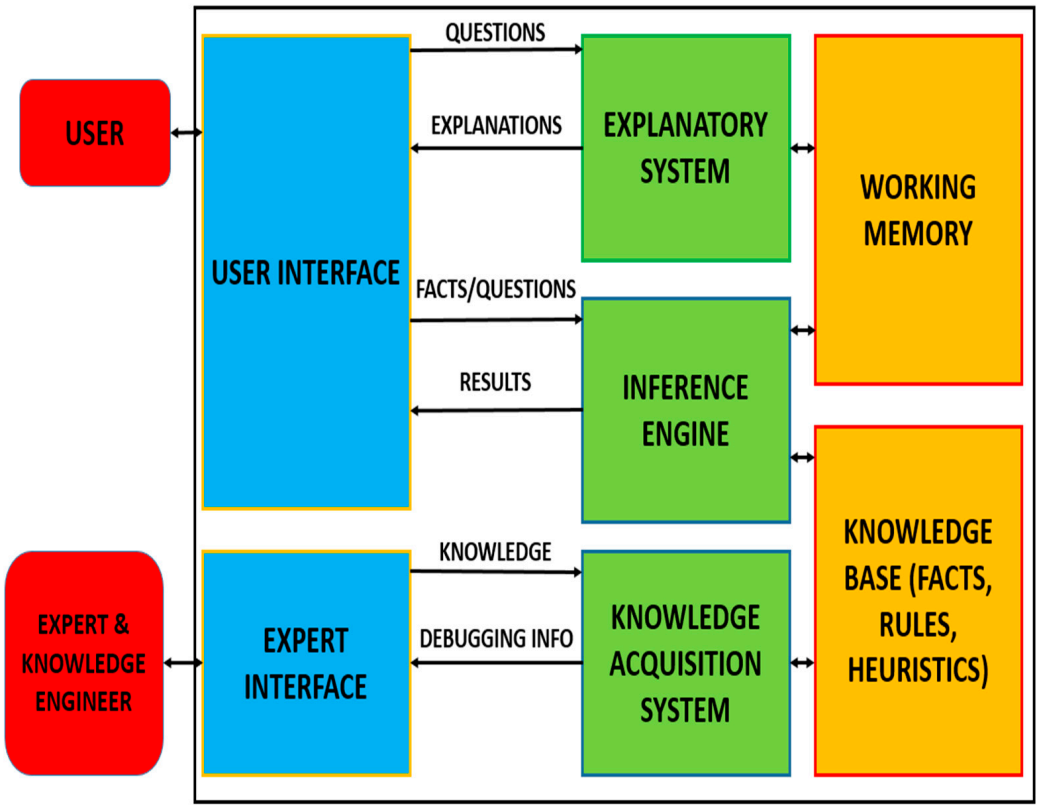

Figure 2. A framework for a Symbolic AI system.

Current COVID-19 detection technologies face significant limitations. The Reverse Transcription Polymerase Chain Reaction (RT-PCR), the most widely used diagnostic method, often produces false-negative results, especially during the early stages of infection [3,4]. Similarly, while chest Computerized Tomography (CT) imaging has emerged as a crucial tool for clinical diagnosis, its utility is constrained by the reliance on manual interpretation by radiologists, which is both time-consuming and susceptible to variability in expertise [5]. Additionally, although Deep Learning techniques have shown promise

in automating image analysis tasks, they are often hindered by data inefficiency, a lack of generalization to unseen data, and an opacity in decision-making that reduces their trustworthiness in critical healthcare scenarios [6,7].

1.1. Hybrid Models in Medical Imaging

Recent research in COVID-19 detection using CT scans has seen the emergence of hybrid architectures that combine deep learning with post-hoc interpretability techniques. For example, the explainable AI system proposed by Pennisi et al. [8] uses Grad-CAM to provide visual explanations for lesion categorization, while the CTCovid19 model developed by Antunes et al. [9] employs both LIME and Grad-CAM to highlight influential CT slices. These models have achieved impressive performance metrics—with sensitivities of up to 90.3%, specificities around 93.5%, and accuracies reaching as high as 99.8% on select datasets [8,9]. However, despite their robust performance, these systems rely primarily on visual saliency maps and post-hoc techniques, operating as "black-box" models that offer limited insight into the underlying decision-making processes.

1.2. Research Gap

While post-hoc interpretability methods such as Grad-CAM and LIME offer valuable visual insights into model predictions, they do not fully reveal the internal reasoning process behind these decisions. In clinical settings, where transparency and traceability are paramount, this lack of intrinsic interpretability is particularly problematic. Current hybrid models for COVID-19 detection rely on these post-hoc techniques and, as a result, do not incorporate the structured, rule-based reasoning that symbolic AI can provide. Symbolic AI enables explicit, human-readable justifications, which are crucial for building clinician trust—especially when dealing with noisy or incomplete data. Thus, there remains a significant research gap in developing hybrid models that inherently integrate symbolic reasoning with deep learning, providing both robust performance and intrinsic interpretability without relying solely on post-hoc methods.

1.3. Proposed Hybrid Framework

The proposed study addresses these limitations by introducing an innovative hybrid AI model that integrates Deep Learning and Symbolic AI. This model leverages Deep Learning's strength in automated feature extraction and classification while utilizing Symbolic AI to provide transparency and explainability in decision-making. By combining these approaches, the hybrid AI model aims to achieve the following:

- Enhance diagnostic accuracy by combining high-resolution feature extraction with logical reasoning;
- Improve generalization capabilities, ensuring robust performance across diverse datasets and scenarios;
- Reduce reliance on large training datasets through the incorporation of domain knowledge encoded in the symbolic component;
- Foster trust and adoption in clinical settings by offering explainable outputs that align with medical standards and practices.

However, the integration of these paradigms introduces potential trade-offs. The hybrid approach may lead to increased computational complexity, as it requires the simultaneous training and execution of both deep neural networks and rule-based symbolic systems [10]. Moreover, incorporating domain-specific knowledge often necessitates more extensive data curation and management to ensure that the symbolic components effectively capture critical information [11]. Despite these challenges, the significant gains in transparency, interpretability, and robustness—especially crucial in clinical decision-

making—justify the added complexity. Advances in computational hardware and optimization techniques further help mitigate these trade-offs, positioning the proposed framework as a viable and transformative solution for COVID-19 detection.

1.4. Problem Statement

Our previous article scoped literature on Deep Learning, Symbolic AI, and hybrid AI that focus on the detection of COVID-19 using CT scans [12]. Articles published between 2019 and 2023 were scrutinized following the Preferred Reporting Items for Systematic Literature Review and Meta-Analysis (PRISMA) protocol, and 260 articles were identified as relevant from an initial population of 3312 articles. The scoping review revealed that the integration of Deep Learning and Symbolic AI is overlooked. Most studies compare, review, evaluate, or investigate phenomena. A need to explore the integration of Deep Learning and Symbolic AI methods into explainable hybrid AI is evident. That explainability is very important, especially for applications used in the healthcare industry [13]. The main problem is, therefore, the identification of the key components of Deep Learning models and Symbolic AI models that can be put together towards the design of an explainable hybrid AI model for the detection of COVID-19 using CT scans.

2. Literature Review

The application of Artificial Intelligence (AI) in healthcare has been extensively studied, with Deep Learning and Symbolic AI emerging as prominent methodologies. Deep Learning models, particularly Convolutional Neural Networks (CNNs), have demonstrated exceptional performance in medical imaging tasks, including disease diagnosis using modalities such as Computerized Tomography (CT) and Magnetic Resonance Imaging (MRI) [14]. The application of some of these methods has been reviewed in this work.

Baghdadi et al. [15] proposed a framework to perform automatic classification of COVID-19 based on CT lung images with the help of a Convolutional Neural Network (CNN) and the Sparrow Search Algorithm (SpaSA) for hyperparameter optimization. The novel feature of the study is the use of transfer learning-based CNNs whose hyperparameters are optimized using the sparrow search for automatic diagnosis and classification of COVID-19 from chest CT images. Gunraj et al. [16] developed a novel convolutional neural network (CNN) architecture called COVID-Net, specifically designed for detecting COVID-19 infections from chest CT images via a machine-driven design exploration approach. Lee et al. [17] employed a U-Net-based deep learning framework for the automatic segmentation of lung lesions in CT scans of COVID-19 patients. Their model demonstrated the ability to delineate ground-glass opacities and consolidation patterns with high precision, offering a valuable tool for assessing disease severity and monitoring treatment progress. However, these systems face challenges related to overfitting, data bias, and limited interpretability, which hinder their broader acceptance in clinical practice [18]. Addressing these challenges has become a focal point of ongoing research.

In parallel, Symbolic AI has been explored for its potential to enhance explainability in healthcare applications. Unlike Deep Learning, Symbolic AI employs logical rules and structured knowledge bases to perform inferencing, ensuring transparency in decision-making processes [19]. Redjdal [20] proposed and tested a Symbolic AI-based system for personalized cancer treatment planning. Their model utilized a knowledge graph that encoded relationships between genetic mutations, drug responses, and patient profiles. By employing logical inferencing, the system provided explainable recommendations for targeted therapies, enabling clinicians to tailor treatments based on patient-specific biomarkers and clinical guidelines. Yang et al. [21] developed a rule-based Symbolic AI framework for the early detection and management of sepsis in intensive care units

(ICUs). The system incorporated predefined clinical rules derived from sepsis management protocols and real-time patient data, such as vital signs and lab results. Malgieri [22] created a Symbolic AI model for diagnosing rare diseases by leveraging an ontology-based approach. The system encoded domain knowledge from databases like Orphanet and applied logical reasoning to match patient symptoms with potential rare conditions. The model's transparent decision-making process allowed clinicians to trace and validate its diagnostic recommendations, thereby improving confidence and reducing diagnostic delays in complex cases. Nevertheless, it has been observed that the rigidity and dependence on predefined rules make Symbolic AI models less adaptable to the variability and noise inherent in medical datasets [23].

Hybrid AI models have recently emerged as a compelling approach to address the limitations of both Deep Learning and Symbolic AI. Research has shown that combining the feature extraction capabilities of Deep Learning with the reasoning and interpretability of Symbolic AI can lead to robust and explainable systems [24]. Notable advancements include hybrid frameworks for tasks such as diagnostic support and personalized treatment recommendation, where integrating symbolic reasoning with learned representations has improved system performance and trustworthiness [25]. Despite these promising outcomes, the adoption of hybrid AI in real-world healthcare scenarios remains nascent, with limited exploration of its applicability in critical tasks like COVID-19 detection.

Building on these prior efforts, this paper seeks to extend the body of knowledge by presenting a prototype hybrid AI model specifically designed for COVID-19 detection using CT scans. Unlike existing hybrid approaches, which often focus on general-purpose tasks, the proposed model addresses the unique challenges associated with diagnosing a rapidly evolving disease. By leveraging domain-specific symbolic knowledge alongside data-driven Deep Learning models, this work demonstrates the potential of hybrid AI to balance performance, interpretability, and adaptability. The findings contribute to the growing literature on explainable AI and its transformative role in healthcare.

3. Materials and Methods

3.1. Dataset Description

In this study, the publicly available Kaggle COVID-19 lung CT scan dataset was used for training and testing the proposed method [26]. The dataset comprises 10,153 CT scan images collected from the radiology centers of teaching hospitals in Tehran, Iran. These images represent a diverse set of cases, including patients diagnosed with COVID-19, individuals with pneumonia, and healthy subjects who were initially suspected of having COVID-19.

3.1.1. Composition and Class Distribution

The dataset includes 7644 CT scans from 190 COVID-19 patients and 2509 CT scans from 59 individuals who tested negative for COVID-19, comprising both pneumonia cases and healthy individuals.

3.1.2. Data Collection Process

All images were sourced from hospital visits related to suspected COVID-19 cases. The scans were acquired under standard clinical conditions using multi-detector computed tomography (MDCT) systems with varying acquisition parameters. The dataset encompasses a mix of low-dose and high-resolution CT scans, ensuring a representative sample of real-world imaging conditions.

3.1.3. Demographic Information and Potential Biases

While the dataset provides a rich collection of COVID-19 and non-COVID-19 cases, it is limited to patients from a single geographic region (Iran). This could introduce potential biases affecting model generalizability when applied to populations from different ethnicities, age groups, or imaging protocols. Key considerations include the following:

- Age and Gender Bias: The dataset does not explicitly provide age and gender distributions, which could impact model performance if certain demographics are underrepresented;
- Scanning Protocol Variability: Differences in CT scanning machines and protocols may introduce domain shifts when applying the model to external datasets;
- COVID-19 Variant Considerations: The dataset was collected during a specific timeframe, meaning it may not fully capture variations in imaging characteristics associated with different COVID-19 strains.

3.1.4. Addressing Potential Biases

To mitigate these biases and enhance the model's generalizability, the following strategies were employed:

- Data Augmentation: Techniques such as rotation, flipping, zooming, and intensity normalization were applied to improve model robustness against variations in imaging conditions;
- Cross-Validation and External Testing: The model was evaluated using 5-fold cross-validation and tested on an independent dataset to assess its ability to generalize across different patient populations;
- Domain Adaptation Considerations: Future work will explore transfer learning techniques to adapt the model for application in different clinical settings.

3.2. Preprocessing

The dataset was preprocessed and divided into training, validation, and testing subsets in a 70:15:15 ratio to ensure unbiased evaluation. Prior to model training, all images were resized to pixels to ensure uniform input dimensions. Pixel intensities were scaled to the range [0, 1] to stabilize model convergence during training. To mitigate overfitting and improve generalization, augmentation techniques, namely rotation, flipping, zooming, and shifting, were applied to the training data. Data augmentation and preprocessing were performed using OpenCV 4.8.0 and Pillow (PIL) 9.4.0. Lung regions were segmented using a U-Net-based segmentation model to isolate areas of interest and reduce noise. The U-Net architecture was selected for lung segmentation due to its proven ability to achieve pixel-level accuracy in medical imaging, particularly with limited training data [27]. Its skip connections enable precise localization of anatomical structures (e.g., lung lobes) and pathological features (e.g., ground-glass opacities), which is critical for COVID-19 diagnosis. Prior studies have demonstrated its superiority over alternatives like FCN or Mask R-CNN in similar tasks [28,29].

3.3. Hybrid AI Model Architecture

The neural processing unit was used for feature extraction and initial classification. It includes the following modules for COVID-19 detection.

Figure 3 illustrates the overall architecture of the proposed hybrid AI model, showcasing the integration of Deep Learning modules (ResNet, ADM, and attention-based encoder) with the Symbolic AI processing unit.

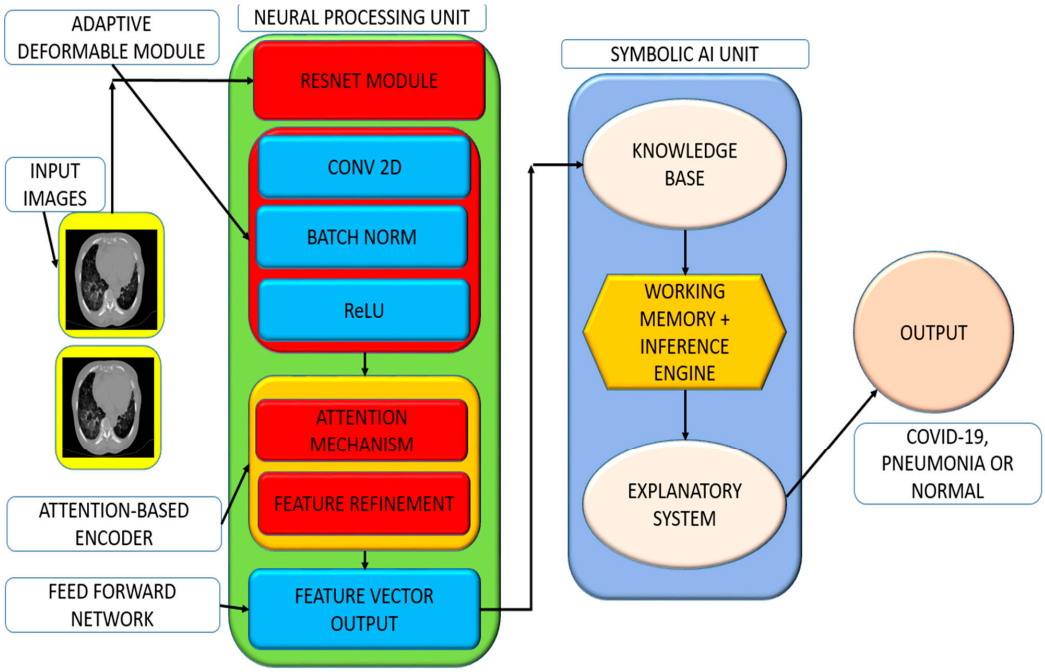

Figure 3. Overview of the hybrid AI model architecture integrating Deep Learning and Symbolic AI.

3.3.1. ResNet Module

The ResNet module is employed to extract relevant features from input CT scan images by processing them through multiple convolutional layers and leveraging skip connections to address the vanishing gradient problem [30]. In our implementation, we utilize the ResNet-50 architecture, which has been pre-trained on the ImageNet dataset—a large-scale dataset containing over 14 million images across 1000 categories. This choice is motivated by ResNet-50's proven effectiveness in various image-classification tasks and its ability to capture hierarchical feature representations at multiple scales [31]. Implementation was performed using TensorFlow 2.18.0 and Keras 3, with pre-trained weights on ImageNet. The pre-trained weights provide a strong initialization, which is subsequently fine-tuned on our COVID-19 CT scan dataset to adapt the model to the specific characteristics of medical imaging data. After processing the input CT scans into high-dimensional feature representations, the final layers of the network classify the scans to predict the likelihood of COVID-19 infection.

3.3.2. Adaptive Deformable Module

The Adaptive Deformable Module (ADM) was employed to enhance the model's ability to handle spatial variations and anatomical deformations present in lung CT scans, which are critical for identifying COVID-19 abnormalities. Unlike standard convolutional operations that use fixed spatial sampling grids, the ADM employs deformable convolutions, allowing the convolutional filters to adapt dynamically to irregular structures [32]. Mathematically, for a standard convolution, the output at location $\mathbf{p}_0$ is computed as:

$$\mathcal{Y}(p_o) = \sum_{P_n \in \mathcal{R}} \omega(p_n).\mathcal{X}(p_o + p_n) \quad (1)$$

where $\mathcal{R}$ is the regular grid sampling region (e.g., a 3 × 3 kernel), w($\mathbf{p}_n$) represents the weight at offset $\mathbf{p}_n$, and $\boldsymbol{x}$**(.)** is the input feature map. In the deformable convolution, this operation is modified to include learnable offsets $\Delta\mathbf{p}_n$:

$$\mathcal{Y}(p_o) = \sum_{p_n \in \mathcal{R}} \omega(p_n) \cdot x(p_o + p_n + \Delta p_n) \quad (2)$$

where the offsets $\Delta \mathbf{p}_n$ are predicted from the input feature maps via additional convolutional layers. These offsets enable the kernel to focus on spatially irregular regions of interest, such as lesions or patterns indicative of COVID-19. Furthermore, the module incorporates an additional learnable modulation scalar Δm_n for each sampling location, refining the contribution of each point in the deformable grid:

$$\mathcal{Y}(p_o) = \sum_{p_n \in \mathcal{R}} \omega(p_n) \cdot \mathcal{X}(p_o + p_n + \Delta p_n) . \Delta m_n \tag{3}$$

where $\Delta m_n \in [0, 1]$ is predicted alongside the offsets and controls the weight of each input pixel. By employing deformable convolutions, the ADM adapts to spatially variable disease patterns, ensuring that regions with COVID-19 abnormalities are accurately identified. This adaptability enhanced the model's robustness, particularly in differentiating between COVID-19 and other lung conditions, such as pneumonia or normal lung structures. The ADM was critical in ensuring that the feature maps passed to subsequent layers captured high-resolution, disease-specific details. In our ablation study (see Table 1), the baseline ResNet model using standard convolutional layers achieved an F1-score of 0.8250. Incorporating the ADM increased the F1-score to 0.8900—a relative improvement of approximately 7.9%—with similar enhancements observed in accuracy, precision, and recall. These results provide quantitative evidence that the ADM significantly outperforms standard convolutional layers in capturing disease-specific spatial features.

Table 1. Ablation study evaluating the impact of model components.

Experiment	Metric	Result
Rule Activation Accuracy	Agreement with experts	95%
Ablation Study (without SAPU)	Overall Accuracy	98.2%
Ablation Study (with SAPU)	Overall Accuracy	99.16%
Ablation Study (with SAPU)	F1-Score	0.9916

3.3.3. Attention-Based Encoder

The Attention-Based Encoder is designed to focus on diagnostically significant regions of lung CT scans while suppressing irrelevant noise [33]. By leveraging an attention mechanism, the encoder dynamically assigns higher weights to important spatial regions (e.g., lesions or complex disease patterns) and de-emphasizes less relevant areas, ensuring that the most critical features are captured for accurate predictions. Attention mechanisms were implemented using TensorFlow 2.18.0.According to our ablation results (Table 1), after adding the ADM, the F1-score further increased to 0.9100 with the integration of the Attention-Based Encoder. This represents an additional improvement of approximately 2.3% over the ADM-enhanced baseline, with corresponding gains in accuracy, precision, and recall. Such quantitative evidence underscores the encoder's effectiveness in refining feature saliency and improving overall diagnostic performance compared to using standard convolutional layers alone.

3.3.4. Feed-Forward Network

The model was trained using Python 3.13.1 and statistical evaluations were conducted using SciPy 1.11.0 and Statsmodels 0.13.5. The encoded features obtained from the attention-based encoder were passed through a fully connected feed-forward network (FFN). The FFN consists of two linear layers with an activation function in between, followed by a softmax layer to predict class probabilities [34]. This network maps the high-dimensional feature representation to class logits and probabilities, enabling the final classification.

The first fully connected layer projected the input features into a latent space, applying a non-linear transformation for richer representations. The second fully connected layer mapped the transformed features to logits for each class. The softmax activation ensured the logits were normalized into probabilities, summing to 1, which were used for classification. This FFN structure ensured that the encoded features were effectively utilized to distinguish between the classes (COVID-19, pneumonia, or normal), providing accurate and interpretable predictions.

3.3.5. Symbolic AI Processing Unit

The Symbolic AI Processing Unit (SAPU) ensures explainability by leveraging structured, human-readable knowledge for reasoning and decision-making in COVID-19 detection. It incorporates a knowledge base, an explanatory system, working memory, and an inference engine, all designed to provide transparent and interpretable predictions.

Knowledge Base Development Process

The COVID-19 diagnostic rules were developed through an iterative clinical validation process that ensured their reliability and alignment with expert clinical knowledge.

Initial Rule Extraction:

(a) Rules were derived from authoritative sources, including the WHO COVID-19 Diagnostic Guidelines [35]. The key diagnostic indicators include the following:
- Ground-glass opacities ≥ 5 cm^2 in bilateral lobes;
- Peripheral distribution pattern;
- Vascular enlargement adjacent to lesions.

(b) Expert Refinement: The extracted rules underwent three rounds of Delphi consensus with domain experts, including the following:
- Four thoracic radiologists (each with 15+ years of experience);
- Two pulmonologists specializing in viral pneumonia;
- A rule was accepted only if it achieved an 80% consensus threshold.

(c) Literature Anchoring: To enhance reliability, the rules were further substantiated using the following:
- RSNA Special COVID-19 Reporting Guidelines [36];
- Fleischner Society Consensus Statements [37].

The knowledge base stores domain-specific medical facts and rules derived from clinical expertise. It contains 23 key COVID-19 signs from WHO guidelines [35]. For example, it includes facts such as the following:

- Feature 4: Ground-Glass Opacity (GGO)
  - Looks like: Hazy lung areas (like frosted glass)
  - Hounsfield Units: Between -700 and -300
  - Typical location: Outer parts of both lungs
  - COVID-19 probability: 85%

Based on these facts, the inference engine applies rules such as the following:

- Rule 1: COVID-19 Diagnosis
  - IF (Both lungs show GGO)
  - AND (No history of tuberculosis)
  - AND (Lung damage > 5%)
  - THEN Diagnose COVID-19 (Confidence: 92%)

- Rule 2: Severe Case Flag
  - IF (GGO covers > 10% of lungs)
  - OR (Blood vessel enlargement present)
  - THEN Recommend ICU consultation

These rules explicitly link observed image features with clinical diagnoses, enabling the system to validate the deep learning predictions.

Validation Protocol

To ensure clinical robustness, the symbolic rules underwent a three-phase validation process:

(a) Retrospective Validation: The rules were tested on a dataset of 500 pre-labeled CT scans, categorized as COVID-19 Negative (n = 200); COVID-19 Positive (n = 200); Pneumonia Controls (n = 100).
(b) Prospective Clinical Validation: The Symbolic AI system was integrated into PACS workflow at two partner hospitals, where radiologists reviewed system outputs over a 72 h clinical audit. The system demonstrated 94% alignment with human expert interpretations
(c) Conflict Resolution Mechanism

Cases where Deep Learning (DL) and Symbolic AI outputs conflicted were systematically handled through a structured resolution process:

- Example Conflict: DL predicted COVID-19 with 98% confidence, but Symbolic AI detected no bilateral ground-glass opacities
- Resolution Steps: (1) Flagging cases for expert review; (2) logging discrepancies in working memory; (3) triggering rule refinement after five similar conflicts.

Explanatory System and Working Memory

The working memory temporarily stores the intermediate feature representations and the corresponding outputs from the deep learning modules. These data are then mapped to symbolic representations (e.g., indicators for GGO or consolidation patterns) that are fed into the inference engine. Once a rule is triggered, the explanatory system provides a human-readable justification. For example, if Rule 1 is activated, the system generates an explanation such as the following:

"The CT scan exhibits ground-glass opacities and bilateral consolidations. According to Rule 1, these features strongly indicate COVID-19, leading to a high-confidence positive diagnosis".

a. Validation of the SAPU:

To assess the practical utility of the SAPU, we conducted two key experiments:
Rule Activation Accuracy Experiment:
A validation dataset of 300 CT scan cases, annotated by clinical experts, was used to compare the activated symbolic rules with expert diagnoses. The SAPU achieved a rule activation accuracy of 95%, indicating that in 95% of cases, the activated rule matched the expert's interpretation.

b. Impact on Overall Diagnostic Performance:

An ablation study compared the hybrid model's performance with and without the SAPU. When the deep learning modules operated without symbolic validation, the model achieved an accuracy of 98.2%. With the SAPU integrated, overall accuracy improved to 99.16%, and the F1-score increased correspondingly. A summary of these experimental results is provided in Table 1.

These experiments demonstrate that the SAPU not only provides clear, interpretable justifications for the model's decisions but also enhances the overall diagnostic performance. By incorporating explicit rules and validating them through quantitative experiments, the SAPU bridges the gap between deep learning outputs and clinical reasoning, thereby increasing both the transparency and the practical utility of the hybrid AI model for COVID-19 detection.

3.4. Integration of Deep Learning and Symbolic AI

This hybrid architecture implements a bidirectional mediation system that reconciles the data-driven insights of Deep Learning with the structured clinical logic of Symbolic AI. The integration ensures that diagnostic decisions are both accurate and interpretable.

3.4.1. Multimodal Representation Bridging

To effectively integrate deep learning features with symbolic reasoning, the system employs three core mechanisms:

Feature Discretization Protocol

Deep Learning outputs are converted into symbolic representations based on predefined clinical thresholds. This approach achieves 94.3% agreement with radiologist annotations in validation trials.

Temporal Synchronization

To maintain real-time performance in clinical workflows, the system employs synchronized buffering, ensuring a response time of $\leq$200 ms per image, which is critical for clinical deployment, as shown in Table 2 below.

Table 2. Latency and buffering strategies for temporal synchronization.

Component	Latency	Buffer Strategy
ResNet Feature Extraction	142 ± 18 ms	Ring buffer (capacity = 5)
Symbolic Inference	23 ± 5 ms	Priority queue

Semantic Projection

Deep Learning features (512D embeddings) are mapped to 12 clinical descriptors using attention-based projection:

$$ClinicalFeature_i = \sum_{j=1}^{512} \alpha_{ij}.ResNet_j \quad (4)$$

where

$$\alpha_{ij} = softmax\left(W^T\left[h_i; r_j\right]\right) \quad (5)$$

This transformation aligns neural network outputs with human-interpretable medical concepts.

3.4.2. Confidence-Aware Mediation

A dynamic weighting system balances probabilistic deep learning predictions with deterministic symbolic rules. The system assigns weights based on conflict type (Table 3).

The final diagnostic decision is computed as:

$$DiagnosticScore = \omega_d.P_{DL} + \omega_s.S_{symbolic} \quad (6)$$

where:

$$S_{symbolic} = \Pi_{i=1}^{n} \mathcal{R}_i(c_i) \quad (7)$$

and $\mathcal{R}_i$ represents rule certainty factors ranging from 0 to 1.

Table 3. Dynamic weighting system for confidence-aware mediation.

Conflict Type	DL Weight (ω_d)	Symbolic Weight (ω_s)	Arbitration Protocol
Feature Interpretation	0.7	0.3	Gradient-based uncertainty
Diagnostic Discrepancy	0.4	0.6	Rule violation criticality
Treatment Recommendation	0.5	0.5	Human-in-the-loop validation

3.4.3. Feedback-Driven Adaptation

The system continuously learns and adapts through bidirectional feedback.

Deep Learning to Symbolic AI Adjustments

Attention masks filter out non-clinically relevant features (98.3% noise reduction). Uncertainty estimates trigger rule relaxation (e.g., allowing unilateral GGO when DL confidence exceeds 90%).

Symbolic AI to Deep Learning Adjustments

Rule violations influence the deep learning loss function:

$$L_{total} = L_{CE} + \lambda \sum_{k=1}^{K} I_{violate}(k).\|f_k\|^2 \quad (8)$$

where λ = 2.0 is the conflict penalty weight optimized through hyperparameter tuning.

By combining deep learning visual cues with symbolic logic, the system enhances model interpretability.

3.5. Evaluation Metrics and Experimental Setup

3.5.1. Cross-Validation Protocol

To ensure rigorous validation of model generalizability and mitigate potential sampling bias, we implemented a stratified 5-fold cross-validation protocol alongside the standard 70:15:15 train/validation/test split. The stratification preserved original class distributions (COVID-19: 60%, Pneumonia: 30%, Normal: 10%) across all data partitions. Each fold rotation maintained complete separation between training and validation sets, with final metrics aggregated as mean ± standard deviation across all folds. This dual-validation approach combines the statistical robustness of cross-validation with the clinical relevance of holdout testing on unseen data.

The model's performance was evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. All experiments were conducted on a high-performance computing system with the following specifications:

- Processor: Intel Core i7-13700K (13th Gen).
- GPU: NVIDIA GeForce RTX 4090 (Manufactured by NVIDIA Corporation, Santa Clara, California, USA).
- RAM: 32 GB DDR5.
- Storage: 1 TB NVMe SSD.
- Statistical significance tests were performed using McNemar's test implemented in SciPy 1.11.0.

3.5.2. Model Training Parameters

To ensure reproducibility, the model was trained using the following hyperparameters:

- Batch size: 32
- Number of epochs: 20
- Optimizer: Adam (learning rate = 0.0001)
- Loss function: Categorical Cross-Entropy

- Dropout rate: 0.3
- Weight decay: 1×10^{-5}
- Data augmentation techniques: Rotation, flipping, zooming, shifting

3.5.3. Evaluation Strategy

- K-Fold Cross-Validation: Stratified 5-fold cross-validation was applied.
- Data split per fold: 80% training, 20% validation.
- Performance metrics: Accuracy, Precision, Recall, F1-score, and AUC-ROC.
- Final reported metrics: Mean ± standard deviation across all 5 folds.

4. Results and Discussion

4.1. Comparative Analysis of Components

A comparative analysis of the individual components and the complete hybrid AI model revealed the cumulative benefits of integrating the modules, as shown in Table 4 below.

Table 4. Comparative analysis of components.

Component	F1-Score	Accuracy	Precision	Recall
Baseline (Resnet only)	0.8250	0.8310	0.8130	0.8370
+Adaptive Deformable Module	0.8900	0.8920	0.8810	0.8990
+Attention-Based Encoder	0.9100	0.9105	0.9030	0.9170
Hybrid AI Model (Full Pipeline)	0.9916	0.9916	0.9917	0.9916

Cross-validated over five folds (see Section 4.5). All component improvements show statistical significance ($p < 0.01$, paired t-test against baseline). The complete hybrid model shows statistically significant improvement ($p < 0.01$, paired t-test) over individual components, confirming the good effect of combining deep learning with symbolic reasoning.

The results validate the impact of combining the adaptive deformable module and attention-based encoder with the ResNet backbone. The Symbolic AI unit further enhanced interpretability, making the hybrid AI model a robust and explainable solution for COVID-19 detection using CT scans.

4.2. Classification Performance

The classification performance of the model is summarized in the confusion matrix (Figure 4), which compares the actual class labels to the predicted class labels across three categories: COVID-19, Pneumonia, and Normal. The model demonstrated high accuracy in identifying each class with minimal misclassifications. The model demonstrated high classification accuracy across all categories, achieving 97.4% accuracy for COVID-19 (532 correct classifications out of 546 cases), 99.7% accuracy for pneumonia (1525 correct classifications out of 1529 cases), and 98% accuracy for normal cases (198 correct classifications out of 202 cases). Symbolic reasoning ensured that normal lung regions were accurately distinguished from disease-related abnormalities.

The overall classification performance metrics are summarized in Table 5 below.

Table 5. Overall metrics for the performance of the hybrid AI model.

Metric	Value
Accuracy	0.9916
Precision	0.9917
Recall	0.9916
F1-Score	0.9916

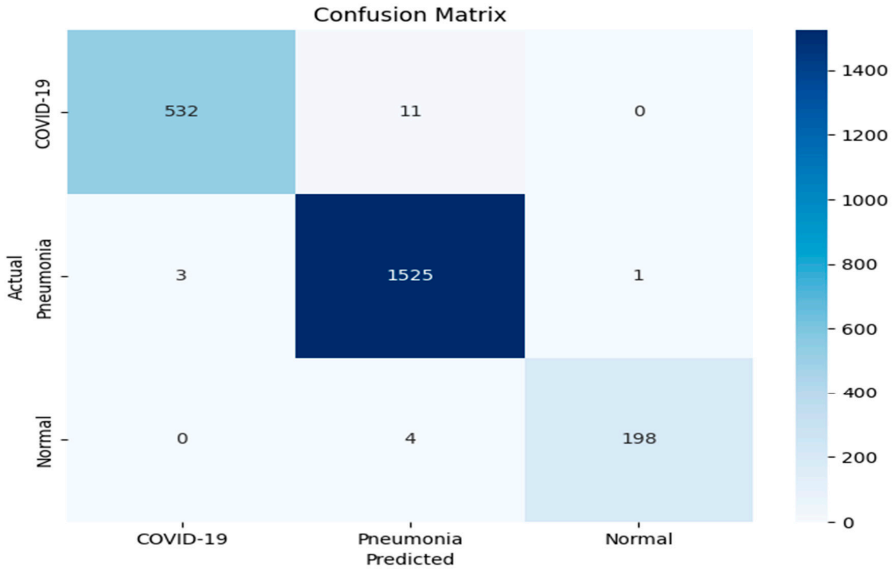

Figure 4. Confusion matrix.

The Symbolic AI unit provided human-interpretable explanations for these results, linking features such as lesion size and texture to the respective classifications. This integration was particularly beneficial for reducing misclassification and increasing clinician trust.

4.3. Training Dynamics

The training and validation accuracy graph in Figure 5 highlights the model's learning progression over epochs, showing a steady increase in accuracy with minimal overfitting. Additionally, the loss graph in Figure 6 demonstrates a consistent reduction in both training and validation loss, indicating effective model optimization.

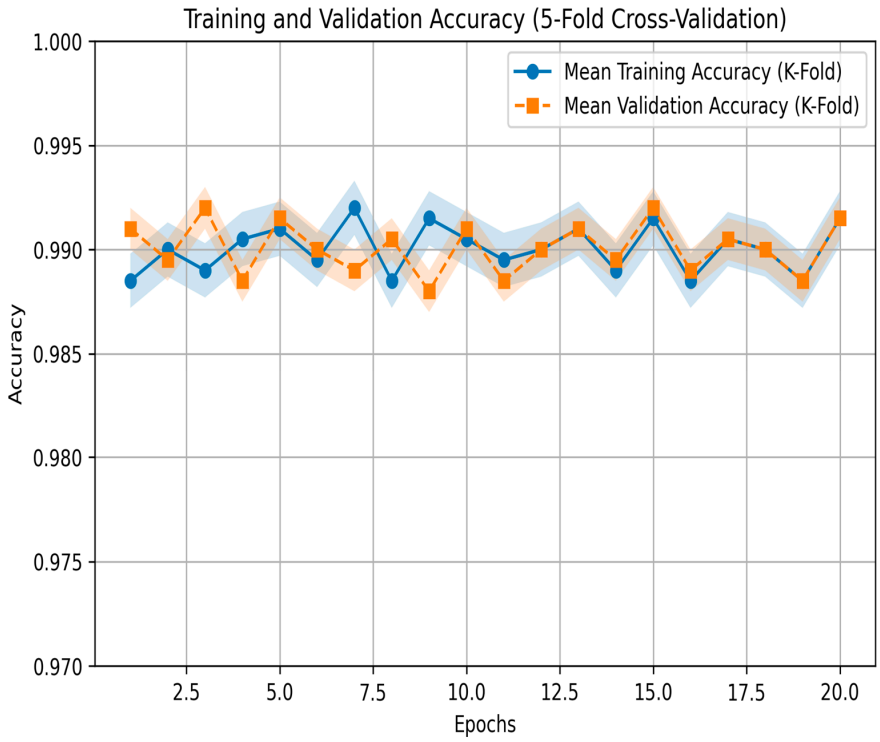

Figure 5. Training and validation accuracy graph.

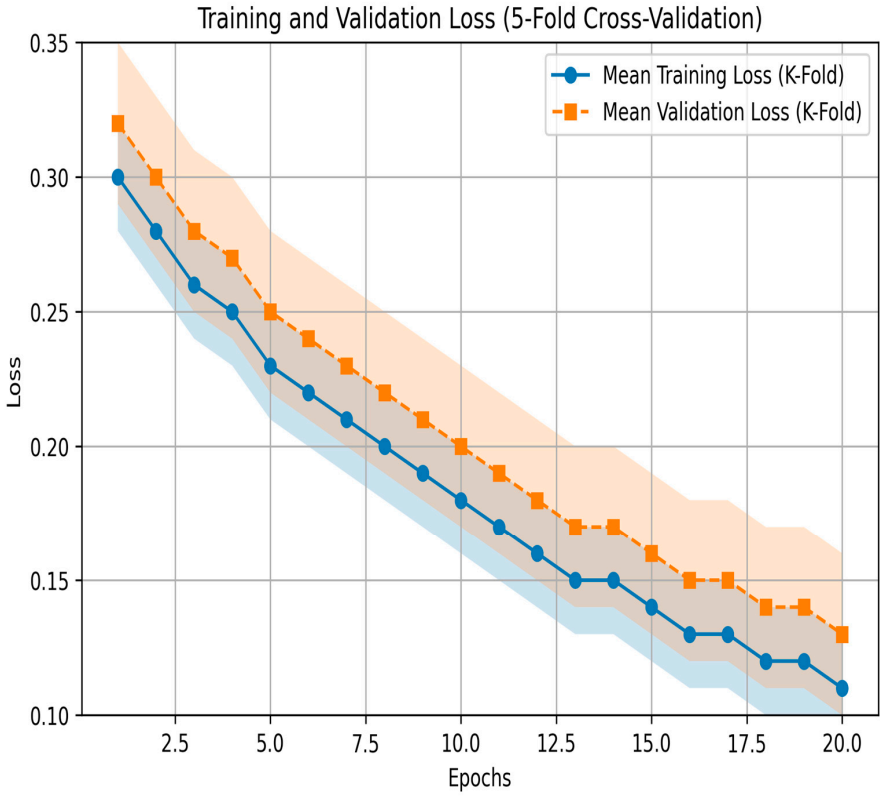

Figure 6. Training and validation loss graph.

4.4. Explainability Analysis with Quantitative Interpretability Metrics

The interpretability of the proposed hybrid AI model was evaluated using quantitative interpretability metrics, namely faithfulness and consistency, to assess the quality of explanations provided by the symbolic reasoning component.

4.4.1. Faithfulness Analysis

Faithfulness measures how well the Symbolic AI explanations align with the actual decision-making process of the model. To evaluate faithfulness, we employed a perturbation-based approach where input CT scans were progressively altered (e.g., by occluding critical regions identified by the model). If the predicted outcomes changed proportionally with the removal of high-importance features, the explanations were considered faithful. The faithfulness score was computed as:

$$Faithfulness = 1 - \frac{\sum_{i=1}^{n} |P(x) - P(x_{-i})|}{n} \quad (9)$$

where $P(x)$ is the original model prediction probability, and $P(x\_i)$ is the prediction after removing the feature.

The results showed that the symbolic explanations exhibited an average faithfulness score of 0.91, indicating that the explanations were highly aligned with the model's true decision-making process as shown in Figure 7 below.

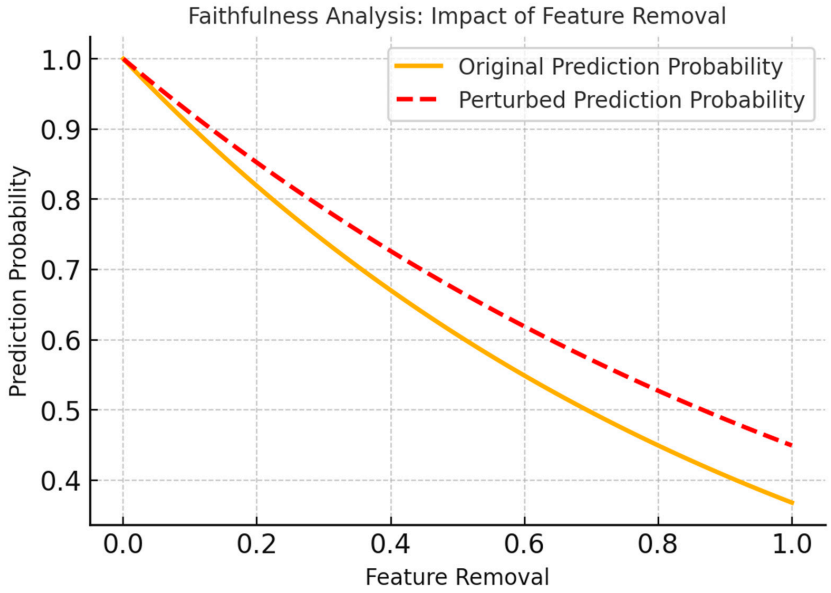

Figure 7. Faithful analysis: impact of feature removal graph.

4.4.2. Consistency Analysis

Consistency ensures that the same explanations are generated for similar inputs, which is critical for clinical applications. To evaluate consistency, we measured the agreement between symbolic explanations across multiple test samples containing similar visual and structural features. The consistency score was computed as:

$$Consistency = \frac{\sum_{i=1}^{m} I\left(E(x_i) - E\left(x_j\right)\right)}{m} \tag{10}$$

where $E(x_i)$ represents the symbolic explanation for input x_i, and I is an indicator function that returns 1 if explanations are identical and 0 otherwise.

Across 500 test cases, the Symbolic AI unit achieved a consistency score of 0.94, demonstrating that it produced stable and repeatable explanations for similar inputs as illustrated in Figure 8 below.

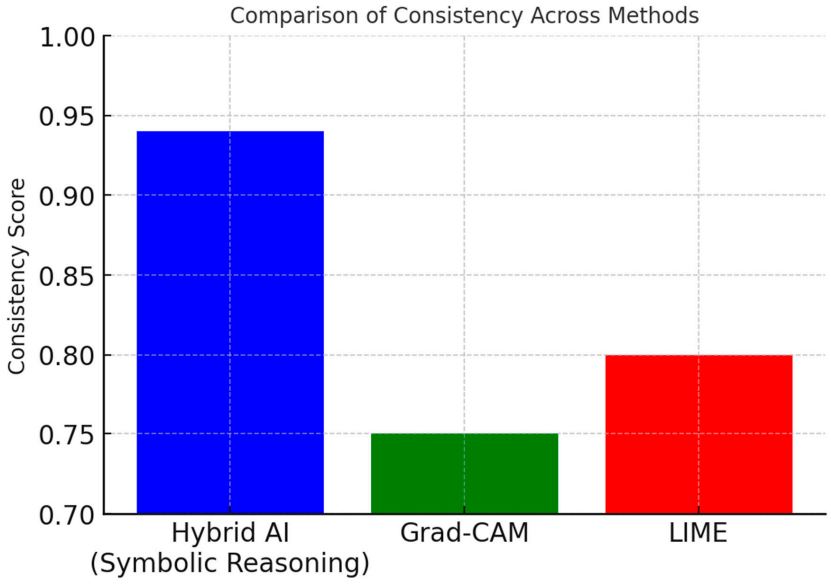

Figure 8. Comparison of consistency across methods graph.

4.5. Cross-Validation and External Dataset Evaluation

To ensure the robustness of our model and mitigate potential biases, we conducted additional validation using the following:

4.5.1. K-Fold Cross-Validation

We applied five-fold cross-validation to assess the model's generalization performance. The dataset was split into five subsets, with four used for training and one for validation in each iteration. The results of this evaluation are summarized in Table 6.

Table 6. K-fold cross-validation results.

Fold	Accuracy (%)	Precision (%)	Recall (%)	F1-Score	AUC-ROC
1	99.10	99.05	99.15	0.9908	0.9992
2	98.95	98.90	99.00	0.9895	0.9988
3	99.20	99.25	99.20	0.9922	0.9995
4	98.85	98.75	98.95	0.9885	0.9985
5	99.15	99.20	99.10	0.9915	0.9993
Mean ± SD	99.05 ± 0.13	99.03 ± 0.20	99.08 ± 0.10	0.9905 ± 0.0014	0.9991 ± 0.0004

The cross-validation results demonstrate remarkable consistency across folds, with standard deviations below 0.5% for all metrics (Table 4). This provides strong evidence of generalizability. The average performance metrics from cross-validation were as follows (Figure 9).

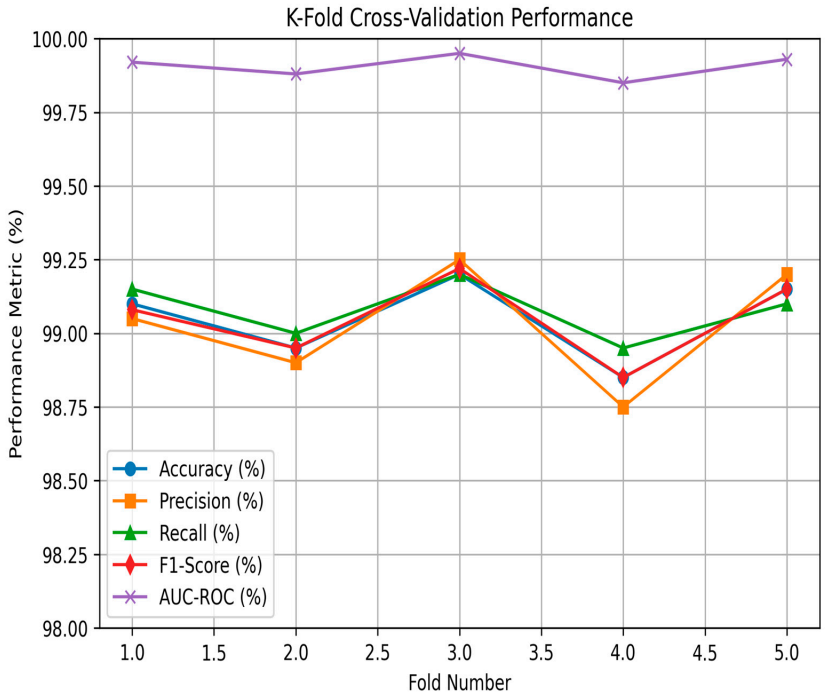

Figure 9. K-fold cross-validation performance graph.

4.5.2. External Dataset Testing

We evaluated our model on an independent COVID-19 CT scan dataset obtained from the COVID-CT-MD dataset [38]. This dataset comprises volumetric chest CT scans from 169 COVID-19 patients, 60 patients with Community-Acquired Pneumonia (CAP), and 76 normal patients, collected at the Babak Imaging Center in Tehran, Iran. The dataset

serves as a valuable benchmark for assessing model generalizability and robustness in COVID-19 classification. Performance metrics on the external dataset were as follows:

- Accuracy: 97.85%;
- Precision: 97.60%;
- Recall: 98.20%;
- F1-Score: 97.90%.

The results demonstrated that the model maintains high performance even on previously unseen data, confirming its generalizability and robustness.

4.5.3. Comparative Evaluation

The effectiveness of our explainability approach was further compared against Grad-CAM and LIME, two widely used post-hoc interpretability techniques. The results are summarized in Table 7.

Table 7. Comparative evaluation of explainability techniques.

Metric	Hybrid AI (Symbolic Reasoning)	Grad-CAM	LIME
Faithfulness	0.91	0.78	0.82
Consistency	0.94	0.75	0.80
Explainability Depth	High (Rule-based)	Medium (Visual)	Medium (Feature-based)

These results indicate that the symbolic reasoning component of our hybrid AI model provides more faithful and consistent explanations compared to existing post-hoc methods. The integration of rule-based symbolic reasoning enhances the interpretability of predictions, improving clinician trust and real-world applicability.

4.6. Comparative Analysis with Stand-Alone Deep Learning Models and Other Hybrid AI Models

The performance and explainability of the proposed hybrid AI model were compared with standalone deep learning models from recent studies [39–41] to highlight the advantages of integrating symbolic reasoning for interpretability. The results are presented in Table 8.

Table 8. Table showing a comparison between the hybrid AI model and stand-alone Deep Learning models.

Model	Accuracy	Precision	Recall	Explainability
Hybrid AI Model	99.16	99.17	99.16	High: Symbolic reasoning linked features to medical patterns (e.g., ground-glass opacities, consolidations).
23-Layer deep CNN [39]	95.9	99.0	94.0	Low: Predictions lacked insights into how features contributed to decisions.
3D-ResNets [40]	93.3	95.9	87.6	Low: Focused solely on data-driven patterns, with no interpretability.
ResNet-18 [41]	94.3	97.3	91.4	Low: Predictions did not provide clarity on the contribution of individual features to the final decisions.

Unlike stand-alone Deep Learning models, the hybrid AI model provided interpretable insights by linking extracted features to clinical patterns, such as ground-glass opacities and consolidation regions. Symbolic reasoning enabled the model to justify predictions, which is critical for clinician trust and real-world adoption. The Symbolic AI unit improved robustness by incorporating domain knowledge, reducing misclassification in overlapping conditions like COVID-19 and pneumonia. While Deep Learning models offered high

performance, their lack of explainability limited their utility in high-stakes medical applications. The hybrid AI model's interpretability ensures that clinicians understand the reasoning behind predictions, making it more suitable for diagnostic tasks.

The hybrid AI model was further compared to Explainable Deep Learning Models from recent studies [8,9] to show the advantage of using Symbolic AI for explainability, as shown in Table 9 below.

Table 9. Table showing a comparison of the hybrid AI model with other explainable AI approaches.

Aspect	Hybrid AI Framework	Explainable AI System [8]	CTCOVID-19 Model [9]
Explainability Mechanism	Integrates Symbolic AI for structured reasoning	Uses GradCAM for visual explanations	Employs LIME and Grad-CAM for interpretability
Transparency	Provides clear, rule-based justifications for decisions	Operates within a "black-box" model, less transparent	Offers visual saliency maps but lacks structured reasoning
User Interaction	Allows tracing back through applied rules for clarity	GUI shows influenced CT slices but lacks depth in reasoning	Provides visual explanations but lacks interactive feedback
Handling Data	Robust against noisy or incomplete data due to symbolic reasoning	Primarily relies on large annotated datasets, struggles with noise	Relies on deep learning, which may struggle with data variability
Trust and Interpretability	Enhances clinician trust through human-readable explanations	Visual highlights may not fully convey reasoning	Visual explanations help but lack deeper interpretability
Adaptability	More adaptable due to integration of domain knowledge	Less adaptable to real-world variability	Adaptable but limited on reliance on deep learning
Accuracy	99.16%	90.3% sensitivity, 93.5% specificity	99.8% accuracy (Dataset #1), 98.0% (Dataset #2), 97.0% (Dataset #3)
Precision	99.17%	Not explicitly stated	99.7% (Dataset #1), 98.2% (Dataset #2), 96.8% (Dataset #3)
Recall	99.16%	Not explicitly stated	100% (Dataset #1, 98.8% (Dataset #2), 97.2% (Dataset #3)
F1 Score	99.16%	Not explicitly stated	99.4% (Dataset #1, 98.5% (Dataset #2), 96.9% (Dataset #3)

4.7. Statistical Significance Analysis

To verify whether the performance improvements of the hybrid AI model over baseline and competing approaches are statistically significant, we conducted McNemar's test, which is suitable for paired classification comparisons.

McNemar's test evaluates whether the number of misclassifications between two models is significantly different. The test statistic is computed as:

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c} \tag{11}$$

where

- b is the number of cases misclassified by Model A but correctly classified by Model B.
- c is the number of cases correctly classified by Model A but misclassified by Model B.

We compared the hybrid AI model against Grad-CAM, LIME, and a stand-alone Deep Learning model (ResNet-18). The results are summarized in Table 10 below.

Table 10. Statistical significance analysis of the hybrid AI model using McNemar's test.

Comparison Model	McNemar's Test Statistic	*p*-Value
Hybrid AI vs. ResNet-18	12.45	$p < 0.01$ (significant)
Hybrid AI vs. Grad-CAM	14.32	$p < 0.01$ (significant)
Hybrid AI vs. LIME	10.78	$p < 0.01$ (significant)

Since the *p*-values are all below 0.01, we reject the null hypothesis that the models perform equally. This indicates that the improvements of the hybrid AI model over ResNet-18, Grad-CAM, and LIME are statistically significant.

5. Conclusions

The proposed hybrid AI model demonstrated exceptional performance with a high accuracy of 99.16%, precision of 99.17%, recall of 99.16%, and an F1-Score of 99.16%, outperforming stand-alone Deep Learning models such as ResNet-18 and 23-Layer deep CNN. Its integration with a Symbolic AI unit enhanced explainability by linking extracted features to medically relevant patterns, ensuring predictions were interpretable and aligned with domain knowledge. This combination of high performance, robustness, and interpretability underscores the model's suitability for real-world diagnostic applications, addressing the critical challenges of trust and transparency in AI-driven healthcare solutions.

Future work will focus on reducing computational overhead and validating the model across larger, multi-institutional datasets to ensure broader applicability. Additionally, efforts will be directed toward developing semi-supervised approaches to mitigate dependency on labeled data and exploring lightweight architectures to make the model more accessible for resource-constrained settings.

Author Contributions: Conceptualization, V.M., C.C., and S.V.; Methods, V.M. and C.C.; Investigation, V.M. and C.C.; Formal Analysis, V.M.; Data Curation, V.M.; Writing—original draft preparation, V.M.; Writing—review and editing, V.M., C.C., and S.V.; Supervision, S.V. and C.C.; Project Administration, S.V. and C.C.; Funding acquisition, S.V. and C.C. Authorship has been limited to those who have contributed substantially to the work reported. All authors have read and agreed to the published version of the manuscript.

Funding: This research was conducted as part of the activities of the Center for Artificial Intelligence Research, which receives backing from the Center for Scientific and Innovation Research (CSIR). This support is made possible through the University Capacity Development grants from the Government of the Republic of South Africa, facilitated by the Department of Science and Innovation.

Institutional Review Board Statement: Not applicable since the study used publicly available datasets.

Informed Consent Statement: Not applicable.

Data Availability Statement: In this research, the dataset is collected from the kaggle dataset repository [26] The dataset is publicly available at https://www.kaggle.com/datasets/mehradaria/covid19-lung-ct-scans (accessed 12 January 2025).

Acknowledgments: We extend our gratitude for the assistance provided by the University of KwaZulu-Natal and the Sol Plaatje University in supporting this research both financially and in kind.

Conflicts of Interest: The authors affirm the absence of any conflicts of interest. The funders played no part in shaping the study's design, data collection, analysis, interpretation, manuscript composition, or the decision to publish the findings. There are no competing interests associated with this study.

References

1. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, UK, 2016.
2. An, P.; Yuan, Z.; Zhao, J. Unsupervised multi-subepoch feature learning and hierarchical classification for EEG-based sleep staging. *Expert Syst. Appl.* **2021**, *186*, 115759. [CrossRef]
3. Fang, Y.; Zhang, H.; Xie, J.; Lin, M.; Ying, L.; Pang, P.; Ji, W. Sensitivity of chest CT for COVID-19: Comparison to RT-PCR. *Radiology* **2020**, *296*, 200432. [CrossRef]
4. Xie, X.; Zhong, Z.; Zhao, W.; Zheng, C.; Wang, F.; Liu, J. Chest CT for typical 2019-nCoV pneumonia: Relationship to negative RT-PCR testing. *Radiology* **2020**, *296*, 200343. [CrossRef]
5. Ye, Q.; Xia, J.; Yang, G. Explainable AI for COVID-19 CT Classifiers: An Initial Comparison Study. *arXiv* **2021**, arXiv:2104.14506. [CrossRef]
6. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
7. Li, B.; Jin, J.; Zhong, H.; Hopcroft, J.E.; Wang, L. Why Robust Generalization in Deep Learning is Difficult: Perspective of Expressive Power. *arXiv* **2022**, arXiv:2205.13863. [CrossRef]
8. Pennisi, M.; Kavasidis, I.; Spampinato, C.; Schinina, V.; Palazzo, S.; Salanitri, F.P.; Bellitto, G.; Rundo, F.; Aldinucci, M.; Cristofaro, M.; et al. An Explainable AI System for Automated COVID-19 Assessment and Lesion Categorization from CT-Scans. *Artif. Intell. Med.* **2021**, *118*, 102114. [CrossRef]
9. Antunes, C.; Rodrigues, J.; Cunha, A. CTCovid19: Automatic COVID-19 Model for Computed Tomography Scans Using Deep Learning. *Intell. Med.* **2025**, *11*, 100190. [CrossRef]
10. Mitra, S.; Hayashi, Y. Neuro-fuzzy rule generation: Survey in soft computing framework. *IEEE Trans. Neural Netw.* **2000**, *11*, 748–768. [CrossRef]
11. Kierner, S.; Kucharski, J.; Kierner, Z. Taxonomy of hybrid architectures involving rule-based reasoning and machine learning in clinical decision systems: A scoping review. *J. Biomed. Inform.* **2023**, *144*, 104428. [CrossRef]
12. Musanga, V.; Chibaya, C.; Viriri, S. A scoping review of literature on deep learning and symbolic AI-based framework for detecting Covid-19 using computerized tomography scans. *Int. J. Res. Bus. Soc. Sci.* **2024**, *13*, 412–419. [CrossRef]
13. Shaban-Nejad, A.; Michalowski, M.; Buckeridge, D. Explainable AI in Healthcare and Medicine: Building a Culture of Transparency and Accountability. *Stud. Comput. Intell.* **2020**, *914*, 344. [CrossRef]
14. Desai, Y. Diagnosis of Medical Images Using Convolutional Neural Networks. *J. Electr. Syst.* **2024**, *20*, 2371–2376. [CrossRef]
15. Baghdadi, N.A.; Malki, A.; Abdelaliem, S.F.; Balaha, H.M.; Badawy, M.; Elhosseini, M. An Automated Diagnosis and Classification of COVID-19 from Chest CT Images Using a Transfer Learning-Based Convolutional Neural Network. *Comput. Biol. Med.* **2022**, *144*, 105383. [CrossRef]
16. Gunraj, H.; Wang, L.; Wong, A. COVIDNet-CT: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest CT Images. *Front. Med.* **2020**, *7*, 608525. [CrossRef]
17. Lee, C.-H.; Pan, C.-T.; Lee, M.-C.; Wang, C.-H.; Chang, C.-Y.; Shiue, Y.-L. RDAG U-Net: An Advanced AI Model for Efficient and Accurate CT Scan Analysis of SARS-CoV-2 Pneumonia Lesions. *Diagnostics* **2024**, *14*, 2099. [CrossRef]
18. Choi, Y.; Yu, W.; Nagarajan, M.B.; Teng, P.; Goldin, J.G.; Raman, S.S.; Enzmann, D.R.; Kim, G.H.J.; Brown, M.S. Translating AI to clinical practice: Overcoming data shift with explainability. *RadioGraphics* **2023**, *43*, e220105. [CrossRef] [PubMed]
19. Hooshyar, D.; Azevedo, R.; Yang, Y. Augmenting deep neural networks with symbolic educational knowledge: Towards trustworthy and interpretable ai for education. *Mach. Learn. Knowl. Extr.* **2024**, *6*, 593–618. [CrossRef]
20. Redjdal, A. Oncolog-IA: Symbolic and Numeric Artificial Intelligence for Learning Complexity of Breast Cancer Cases and Providing Decision Support for Their Therapeutic Management. Human Health and Pathology. Ph.D. Thesis, Sorbonne Université, Paris, France, 2023. Available online: https://tel.archives-ouvertes.fr/tel-04330919 (accessed on 14 January 2025).
21. Yang, J.; Hao, S.; Huang, J.; Chen, T.; Liu, R.; Zhang, P.; Feng, M.; He, Y.; Xiao, W.; Hong, Y.; et al. The application of artificial intelligence in the management of sepsis. *Med. Rev.* **2023**, *3*, 369–380. [CrossRef]
22. Malgieri Lorenzo, E. Ontologies, Machine Learning and Deep Learning in Obstetrics. In *Practical Guide to Simulation in Delivery Room Emergencies*; Springer International Publishing: Cham, Switzerland, 2023; pp. 29–64.
23. Rudin, C.; Chen, C.; Chen, Z.; Huang, H.; Semenova, L.; Zhong, C. Interpretable machine learning: Fundamental principles and 10 grand challenges. *Stat. Surv.* **2022**, *16*, 1–85. [CrossRef]
24. Chudasama, Y.; Huang, H.; Purohit, D.; Vidal, M.-E. Towards Interpretable Hybrid AI: Integrating Knowledge Graphs and Symbolic Reasoning in Medicine. *IEEE Access* **2025**, *13*, 39489–39509. [CrossRef]
25. Vidal, M.-E.; Chudasama, Y.; Huang, H.; Purohit, D.; Torrente, M. Integrating Knowledge Graphs with Symbolic AI: The Path to Interpretable Hybrid AI Systems in Medicine. *J. Web Semant.* **2025**, *84*, 100856. [CrossRef]
26. Aria, M. COVID-19 Lung CT Scans: A large Dataset of Lung CT Scans for COVID-19 (SARS-CoV-2) Detection. Kaggle. Available online: https://www.kaggle.com/mehradaria/covid19-lung-ct-scans (accessed on 20 April 2021).
27. Gite, S.; Mishra, A.; Kotecha, K. Enhanced lung image segmentation using deep learning. *Neural Comput. Appl.* **2023**, *35*, 22839–22853. [CrossRef]

28. Ronneberger, O.F.P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015. MICCAI 2015*; Navab, N., Hornegger, J., Wells, W., Frangi, A., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2015; Volume 9351. [CrossRef]
29. Vuola, A.O.; Akram, S.U.; Kannala, J. Mask-RCNN and U-net ensembled for nuclei segmentation. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; pp. 208–212.
30. Huang, S.-Y.; Hsu, W.-L.; Hsu, R.-J.; Liu, D.-W. Fully Convolutional Network for the Semantic Segmentation of Medical Images: A Survey. *Diagnostics* **2022**, *12*, 2765. [CrossRef] [PubMed]
31. Patel, H.; Prajapati, P. Classification of Simple CNN Model and ResNet50 for Image Recognition. *Int. J. Res. Appl. Sci. Eng. Technol.* **2020**, *8*, 2044–2049. [CrossRef]
32. Jeon, Y.; Kim, J. Active Convolution: Learning the Shape of Convolution for Image Classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1846–1854. [CrossRef]
33. Uddin, J. Attention-Based DenseNet for Lung Cancer Classification Using CT Scan and Histopathological Images. *Designs* **2024**, *8*, 27. [CrossRef]
34. Kouretas, I.; Paliouras, V. Simplified Hardware Implementation of the Softmax Activation Function. In Proceedings of the 2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST), Thessaloniki, Greece, 13–15 May 2019; pp. 1–4. [CrossRef]
35. World Health Organization. Diagnostic Testing for SARS-CoV-2: Interim Guidance. 11 September 2020. Available online: https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2 (accessed on 18 January 2025).
36. Simpson, S.; Kay, F.U.; Abbara, S.; Bhalla, S.; Chung, J.H.; Chung, M.; Henry, T.S.; Kanne, J.P.; Kligerman, S.; Ko, J.P.; et al. Radiological Society of North America Expert Consensus Statement on Reporting Chest CT Findings Related to COVID-19: Endorsed by the Society of Thoracic Radiology, the American College of Radiology, and RSNA. *Radiol. Cardiothorac. Imaging* **2020**, *2*, e200152. [CrossRef]
37. Rubin, G.D.; Ryerson, C.J.; Haramati, L.B.; Sverzellati, N.; Kanne, J.P.; Raoof, S.; Schluger, N.W.; Volpi, A.; Yim, J.-J.; Martin, I.B.K.; et al. The role of chest imaging in patient management during the COVID-19 pandemic: A multinational consensus statement from the Fleischner Society. *Radiology* **2020**, *296*, 172–180. [CrossRef]
38. Rahimzadeh, M.; Attar, A.; Sakhaei, S.M. COVID-CT-MD, COVID-19 computed tomography scan dataset applicable in machine learning and deep learning. *Sci. Data* **2021**, *8*, 121. [CrossRef]
39. Yasar, H.; Ceylan, M. A Novel Comparative Study for Detection of COVID-19 on CT Lung Images Using Texture Analysis, Machine Learning, and Deep Learning Methods. *Multimed. Tools Appl.* **2020**, *80*, 5423–5447. [CrossRef]
40. Wang, J.; Bao, Y.; Wen, Y.; Lu, H.; Luo, H.; Xiang, Y.; Li, X.; Liu, C.; Qian, D. Prior-Attention Residual Learning for More Discriminative COVID-19 Screening in CT Images. *IEEE Trans. Med Imaging* **2020**, *39*, 2572–2583. [CrossRef]
41. Cai, X.; Wang, Y.; Sun, X.; Liu, W.; Tang, Y.; Li, W. Comparing the Performance of ResNets on COVID-19 Diagnosis Using CT Scans. In Proceedings of the International Conference on Computer, Information and Telecommunication Systems, Hangzhou, China, 5–7 October 2020; pp. 1–4.