

# A Quantum-Inspired Ensemble Learning Framework for Multi-Class Respiratory Disease Diagnosis

**Francis Ofoma Eze**

`francisofomaeze@std.uestc.edu.cn`

University of Electronic Science and Technology of China

**Wang QingXian**

University of Electronic Science and Technology of China

**Chidi Daniel Chukwu**

University of Electronic Science and Technology of China

**Attah Festa Kosisochukwu**

State University of Medical and Applied Science

**David Ocho Audu**

University of Electronic Science and Technology of China

**Eze Franklin ThankGod**

Chengdu University of Technology

**Omeke Juliet Ezinne**

University of Nigeria

**Eze Mary Chinecherem**

University of Nigeria

**Chukwuebuka Joseph Ejayi**

Chengdu University of Technology

---

## Research Article

**Keywords:** Quantum-inspired Machine Learning, Respiratory Disease Diagnosis, Class Imbalance, Ensemble Learning, Explainable Artificial Intelligence

**Posted Date:** August 18th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-10708277/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.



# 1 A Quantum-Inspired Ensemble Learning Framework for Multi-Class Respiratory Disease Diagnosis.

Francis Ofoma Eze <sup>1</sup>, Wang QingXian <sup>1</sup>, Chidi Daniel Chukwu <sup>2</sup>, Attah Festa Kosisochukwu <sup>3</sup>,  
David Ocho Audu <sup>4</sup>, Eze Franklin ThankGod <sup>5</sup>, Omeke Juliet Ezinne <sup>6</sup>, Eze Mary Chinecherem <sup>7</sup>,  
Chukwuebuka Joseph Ejayi <sup>5</sup>

<sup>1</sup> School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, China.

<sup>2</sup> School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, China.

<sup>3</sup> Department of Human Anatomy, State University of Medical and Applied Science, Igbo-Eno, Enugu State, Nigeria

<sup>4</sup> School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, China.

<sup>5</sup> College of Nuclear Technology and Automation Engineering, Sichuan Industrial Internet Intelligent Monitoring and Application Engineering Research Center, Chengdu University of Technology, Chengdu, Sichuan, China.

<sup>6</sup> Department of Medical Radiography, University of Nigeria, Nsukka, Enugu State, Nigeria.

<sup>7</sup> Department of Biochemistry, University of Nigeria, Nsukka, Enugu State, Nigeria.

## Abstract

Accurate multi-class diagnosis of respiratory diseases using clinical tabular data remains a challenging problem in medical decision support systems due to complex non-linear feature dependencies, severe class imbalance, and limited interpretability of conventional machine learning models. While recent data-driven approaches demonstrate strong predictive performance, many fail to preserve higher-order interaction structure, address imbalance in a geometry-aware manner, or provide clinically meaningful explanations. To address these limitations, this study proposes Quantum-Cognitive Fusion (Q-CogFusion), a unified quantum-inspired framework integrating Quantum-Cognitive Feature Entanglement (QCFE) for correlation-driven phase-modulated feature synthesis, Adaptive Manifold Regularization with Tangent Propagation (AMR-TP) for geometry-preserving class balancing, a Multi-Scale Attention Fusion Network (MSAFN) for performance-adaptive ensemble aggregation, and an entanglement-aware SHAP-based explainability mechanism for dependency-consistent interpretation. The framework was evaluated on 2,000 clinically annotated patient records spanning five respiratory conditions using stratified cross-validation. Q-CogFusion achieved high diagnostic performance, attaining 99.40% accuracy, 99.40% macro F1-score, and a macro-averaged AUC of 1.00, with consistently strong recall across minority disease classes. Ablation analysis demonstrated that QCFE substantially enhanced class separability under imbalanced conditions, while AMR-TP maintained robust performance under strict class balancing. Comparative evaluation showed that the proposed framework consistently outperformed conventional machine learning and ensemble baselines, while providing interpretable insights aligned with known cardiopulmonary pathophysiology. These results suggest that Q-CogFusion offers a reliable and explainable diagnostic solution suitable for deployment in resource-constrained clinical decision support environments.

**Keywords:** Quantum-inspired Machine Learning; Respiratory Disease Diagnosis; Class Imbalance; Ensemble Learning; Explainable Artificial Intelligence.

## 2 Introduction

Respiratory diseases remain a significant global health challenge, contributing substantially to morbidity and mortality worldwide [1], [2]. According to the Global Burden of Disease (GBD) 2023 estimates, lower respiratory infections (LRIs), including pneumonia, bronchiolitis, and acute bronchitis, accounted for approximately 2.50 million deaths globally (95% uncertainty interval [UI]: 2.24–2.81 million) and 98.7 million disability-adjusted life years (DALYs; 95% UI: 87.7–112 million) in 2023. In addition, chronic respiratory conditions such as chronic obstructive pulmonary disease (COPD) and asthma were responsible for an estimated 4.2 million deaths during the same period [3]. These highlight the persistent burden of respiratory diseases and the importance of reliable diagnostic methodologies.

Conventional diagnostic approaches for respiratory diseases primarily rely on clinical symptom assessment, laboratory investigations, and radiological imaging. Although widely adopted in clinical practice, these methods are often affected by inter-observer variability, delayed diagnostic workflows, and limited sensitivity in early or atypical disease presentations [4], [5], [6]. Such constraints may hinder timely diagnosis and treatment, particularly in large-scale screening or resource-constrained environments.

The application of machine learning (ML) techniques has introduced data-driven diagnostic frameworks capable of analyzing high-dimensional and multimodal clinical data [7], [8]. Classical ML models, including support vector machines, random forests, and gradient boosting algorithms, have demonstrated strong performance in respiratory disease classification tasks across diverse datasets [9], [10], [11]. Despite these advances, several methodological challenges persist, particularly in relation to high-dimensional feature representations, severe class imbalance, and complex non-linear feature interactions commonly encountered in clinical datasets [12], [13].

High-dimensional feature spaces often exacerbate the curse of dimensionality, increasing computational complexity and reducing generalization capability [14], [15]. Commonly used imbalance-handling techniques, such as the Synthetic Minority Over-sampling Technique (SMOTE), generate synthetic samples through linear interpolation, which may not adequately preserve the intrinsic geometry of the underlying data manifold [16], [17]. As a result, classifier decision boundaries can be adversely affected, especially for minority disease categories [18].

Ensemble learning strategies have been proposed to improve predictive robustness by combining multiple base classifiers [19], [20]. Techniques such as bagging, boosting, voting, and stacking aim to reduce model variance and bias through aggregation [21]. However, many existing ensemble frameworks employ static or uniform weighting mechanisms that do not adapt to dataset-specific characteristics observed during

validation [22]. This limitation may restrict the effectiveness of ensemble diversity, particularly in heterogeneous and imbalanced medical datasets [21].

In parallel, conventional feature engineering methods are predominantly based on Euclidean representations, which may be insufficient for capturing higher-order dependencies among clinical variables [23], [24]. Quantum-inspired classical algorithms have recently been explored as an alternative, drawing on concepts from quantum information theory, such as superposition and entanglement, to construct expressive feature representations using classical computing resources [22], [23]. Prior studies have reported improved modeling of non-linear feature interactions using such approaches in biomedical applications [25], [26]. Nevertheless, existing methods generally address feature interaction modeling, imbalance correction, ensemble fusion, or interpretability in isolation, rather than within a unified diagnostic framework [27].

To address these limitations, this study proposes Q-CogFusion (Quantum-Cognitive Fusion), a unified framework for multi-class respiratory disease diagnosis. The framework integrates four complementary components. First, Quantum-Cognitive Feature Entanglement (QCFE) constructs correlation-driven, phase-modulated feature interactions inspired by quantum entanglement principles to enhance representational capacity. Second, Adaptive Manifold Regularization with Tangent Propagation (AMR-TP) performs geometry-aware class balancing by generating synthetic samples within localized tangent space approximations of the data manifold, thereby preserving intrinsic structural properties. Third, a Multi-Scale Attention Fusion Network (MSAFN) adaptively aggregates heterogeneous classifiers using temperature-scaled softmax attention weights derived from stratified cross-validation performance. Finally, Quantum-Entangled SHAP Explainability extends SHapley Additive exPlanations (SHAP) to account for QCFE-induced feature dependencies, supporting interpretable model outputs.

The main contributions of this study can be summarized as follows:

- A quantum-inspired feature engineering mechanism that models complex and interdependent clinical feature relationships through correlation-driven entanglement.
- A manifold-aware class imbalance handling strategy that constrains synthetic sample generation to tangent-aligned regions of the data manifold.
- An adaptive ensemble fusion approach that dynamically assigns classifier weights based on validation performance using attention-based mechanisms.
- An enhanced explainability framework that provides attribution scores consistent with entangled feature representations.

The remainder of this paper is organized as follows. Section 2 reviews related work on machine learning-based respiratory disease diagnosis, ensemble learning, quantum-inspired methods, and explainable artificial intelligence. Section 3 describes the proposed Q-CogFusion framework in detail. Section 4 outlines the experimental setup, datasets, and evaluation metrics. Section 5 presents the experimental results, which are discussed in Section 6. Section 7 concludes the paper by summarizing the key findings and outlining future research directions.

### **3 Related Work**

Machine learning applications in respiratory disease diagnosis have advanced significantly, serving as a paradigm for multi-class classification tasks encompassing pneumonia, bronchitis, chronic obstructive pulmonary disease (COPD), and infectious respiratory illnesses [28]. These methods capture intricate nonlinear patterns in multimodal data, including clinical features, imaging, and audio signals, with greater efficacy than conventional diagnostic approaches [29], [30]. This section organizes the literature into thematic clusters across seven subsections: 2.1 reviews feature engineering for respiratory disease classification; 2.2 examines class imbalance mitigation in medical diagnosis; 2.3 surveys ensemble learning for disease prediction; 2.4 traces the evolution of quantum-inspired classical algorithms; 2.5 explores quantum-inspired classical machine learning in healthcare; 2.6 assesses explainability frameworks for medical AI systems; and 2.7 synthesizes the findings and identifies research gaps.

#### **3.1 Feature Engineering for Respiratory Disease Classification**

Feature engineering transforms raw clinical and sensor data into discriminative representations that bolster diagnostic performance in respiratory conditions [31]. Recent efforts emphasize multimodal integration and hybrid architectures to model pathophysiological interactions. Shi et al. [32] integrated environmental, meteorological, and clinical data in a random forest-based multi-label model for respiratory virus risk prediction across 31 Chinese regions, attaining an average area under the curve (AUC) of 0.90 with SHapley Additive exPlanations (SHAP) for interpretability. Guo et al. [33] developed a drift-aware ensemble for streaming data, employing sample weighting and diversity matrices to select base learners. Adhikari et al. [34] reviewed AI's role in equitable healthcare for low- and middle-income countries, addressing bias, resource limitations, safety, and ethics.

These studies underscore hybrid feature engineering's value in bridging tabular and imaging modalities for respiratory tasks. They exhibit limitations such as heavy reliance on exogenous variables, inadequate modeling of high-dimensional dependencies, and confinement to classical, implicitly linear interaction representations, which constrain scalability and expressivity in heterogeneous clinical data.

### 3.2 Class Imbalance Mitigation in Medical Diagnosis

Class imbalance poses a persistent obstacle in medical datasets, where minority respiratory conditions bias models toward majority classes [29]. Innovations target data-level and algorithm-level strategies for minority class robustness. Rodríguez-Díaz et al. [35] surveyed quantum machine learning in healthcare, highlighting feature map transformations for enhanced class separation. Kala et al. [36] introduced a multi-stage framework for pediatric respiratory sound analysis, using convolutional-recurrent neural networks (CVRNNs) for feature extraction and confidence-weighted fusion to improve wheeze and crackle detection. Yan et al. [37] proposed an oversampling-augmented multi-class classifier combining random forest and XGBoost for patient-reported outcomes, showing generalization across cancer cohorts. Muhammad et al. [38] devised dynamic frameworks for imbalanced medical data, integrating adaptive resampling and weighted aggregation to handle high ratios while maintaining stability.

These approaches advance minority class learning through resampling and algorithmic adjustments. They are limited by Euclidean-space operations, ambient-space synthetic generation without geometric constraints, potentially yielding implausible samples, and insufficient capture of higher-order temporal-feature dependencies, compromising fidelity in severely skewed clinical distributions.

### 3.3 Ensemble Learning for Disease Prediction

Ensemble learning enhances robustness in respiratory and infectious disease prediction by leveraging model diversity to minimize variance [39]. Zhou et al. [40] presented Identity-Mapping ResFormer, a convolutional neural network-Transformer hybrid for pneumonia X-ray diagnosis, incorporating multiscale residual modules and identity-mapping to mitigate gradient inconsistencies, yielding accuracies up to 97.67% with strong F1-scores and recall. Xu et al. [14] optimized geo-hierarchical ensembles for influenza, RSV, and COVID-19 hospitalization forecasts. Bouzammour et al. [39] fused a convolutional neural network, a multilayer perceptron, and a recurrent neural network models for cough-based respiratory detection. Yuan et al. [41] proposed DBNX, fusing deep belief networks with XGBoost to integrate polygenic risk scores and non-genetic factors for disease risk prediction on the UK Biobank.

These contributions demonstrate ensembles' efficacy in integrating diverse modalities and reducing prediction errors. Limitations include modality-specific designs, static aggregation without performance-adaptive weighting, and neglect of cross-feature dependencies, restricting variance reduction and generalization in high-dimensional clinical scenarios.

### 3.4 Evolution of Quantum-Inspired Classical Algorithms

Recent research emphasizes quantum-inspired feature mapping and kernel approximations, transforming classical inputs into Hilbert-like spaces for superior nonlinear separation and expressivity. Rodríguez-Díaz et al. [42] surveyed quantum machine learning foundations, emphasizing tensor network emulations for

quantum-like advantages in supervised learning with reduced overhead. Borella et al. [43] implemented ultra-low-latency predictors on field-programmable gate arrays (FPGAs) through randomized tensor decompositions for real-time inference. Hussein et al. [44] introduced the Schrödinger optimizer, a metaheuristic for deep network optimization in high-dimensional spaces. Zhao et al. [45] analyzed entanglement-simulating correlations in classical architectures, proving separations in function approximation complexity. Long et al. [46] incorporated unitary transformations in hybrid convolutions for improved robustness. Bansal et al. [47] customized quantum support vector machine circuits with rotations for refined hyperplane separation. Yin et al. [48] emulated quantum kernels photonically for nonlinear dataset handling. He et al. [49] applied quantum-like perturbations in adversarial ensembles for time series anomaly detection. Gao et al. [50] used structured priors for explainable semantic learning in fine-grained classification. Chauhan et al. [51] developed quantum-potential-inspired ensemble algorithms for global optimization.

These developments position quantum-inspired techniques as enhancers of classical machine learning across domains. Yet, they remain constrained by task-agnostic designs, hardware emulation overheads, and limited empirical validation in noisy, high-dimensional settings.

### **3.5 Quantum-Inspired Classical Machine Learning in Healthcare**

Quantum-inspired approaches show promise in healthcare for diagnostic classification, predictive modeling, and personalized medicine, often outperforming classical methods in high-dimensional, noisy data. Gupta et al. [52] reviewed quantum machine learning in digital health, noting scalability constraints, noise issues, and variable advantages over baselines. Bukkarayasamudram et al. [15] evaluated quantum computing's impact on diagnostics and resource optimization in privacy-constrained environments. Rithani et al. [53] combined relevance vector machines, incremental learning, Gaussian mixture model resampling, and adaptive broad learning for imbalanced medical data, reaching 99% accuracy. Dankan et al. [54] surveyed paradigms for diagnostics and drug discovery, highlighting hardware limitations. Ray et al. [55] integrated quantum machine learning with explainable AI for precision medicine. Vani et al. [56] employed attention-based convolutional neural networks and SHAP for health monitoring at 94.1% accuracy. Flöther et al. [57] conceptualized quantum methods for biomarker discovery in multidimensional data. Quantum-inspired algorithms adapt principles like feature mapping and entanglement simulation to boost expressivity on classical hardware [45]. Lv et al. [58] leveraged superposition for secure 5G healthcare data transmission.

These studies illustrate potential for biomedical pattern discovery but are limited by ambient-space representations, conceptual focus without clinical deployment, and inadequate handling of feature interdependencies or imbalance in multimodal data.

### 3.6 Explainability for Medical AI Systems

Explainable AI is essential for clinical trust and regulatory adherence in diagnostic systems [18]. Hur et al. [59] compared SHAP with clinician-oriented explanations for decision support usability. Ullah et al. [60] used statistical, visual, and rule-based explanations for medical imaging. Mohapatra et al. [61] proposed multidisciplinary explainable AI with quantum principles. Tjoa and Guan [62] surveyed methods emphasizing uncertainty quantification and attributions for transparency. Pradeep et al. [63] enhanced explainability through case-based reasoning, applying retrieval and adaptation for decision clarity in respiratory diagnostics.

These frameworks advance interpretability through post-hoc and structured techniques. However, they overlook inter-feature dependencies, assume attribution independence, and lack fidelity in complex, entangled representations, limiting applicability to high-dimensional multimodal clinical data.

### 3.7 Synthesis and Research Gap

Existing research highlights notable strides in feature engineering, imbalance handling, ensemble strategies, quantum-inspired computations, and interpretability techniques. Yet, constraints persist in multi-class respiratory diagnostics from tabular clinical datasets. Conventional methods often apply rudimentary mappings that isolate variables, linear interpolations that disrupt intrinsic data manifolds during balancing, and independent attributions that neglect entangled pathophysiological dynamics [64], [65].

Such limitations necessitate unified architectures that emphasize correlation-based integration, manifold-aware regularization, adaptive ensemble weighting, and dependency-sensitive explanations. This work introduces Q-CogFusion to bridge these voids through quantum-cognitive entanglement for phase-augmented features, tangent propagation in adaptive regularization for geodesic sample fidelity, multi-scale attention-driven fusion with cross-validation weighting, and entangled SHAP for refined attributions. The resulting system adeptly navigates respiratory complexities, delivering geometric precision, clinical insight, and tolerance to imbalances.

## 4 Methodology

### 4.1 Q-CogFusion Framework Overview

The Q-CogFusion framework implements a systematic four-stage computational pipeline for transforming raw multivariate clinical measurements into disease diagnostic predictions for respiratory conditions. The complete architecture, illustrated in Fig. 1, comprises clinical feature engineering that derives five physiologically interpretable hemodynamic indices from baseline vital signs using established cardiovascular formulations, a Quantum-Cognitive Feature Entanglement (QCFE) module that models non-linear feature interactions through phase-modulated correlation-based synthesis, Adaptive Manifold

Regularization with Tangent Propagation (AMR-TP) that enforces geometric class separability through local tangent-space augmentation, and a Multi-Scale Attention Fusion Network (MSAFN) that performs dynamic ensemble aggregation through cross-validation- driven adaptive weighting. In this work, quantum-inspired terminology is used strictly as a mathematical abstraction for structured dependency modeling, not as a physical or hardware-dependent quantum process.

The framework accepts 13 baseline clinical measurements, including demographics (age, gender), vital signs (heart rate, blood pressure, respiratory rate, body temperature, oxygen saturation), and three symptom descriptors, yielding  $d = 14$  features following preprocessing and blood pressure decomposition. QCFE expands this representation to  $d' = 29$  features through phase-modulated entanglement operations (Section 3.3). AMR-TP subsequently balances the training cohort from  $n = 2000$  patients to  $n' = 5835$  samples while maintaining strict class equilibrium (Section 3.4). Three gradient boosting base learners (XGBoost, LightGBM, CatBoost) process the enhanced balanced data independently, with MSAFN dynamically weighting their probabilistic outputs for final prediction (Section 3.5).

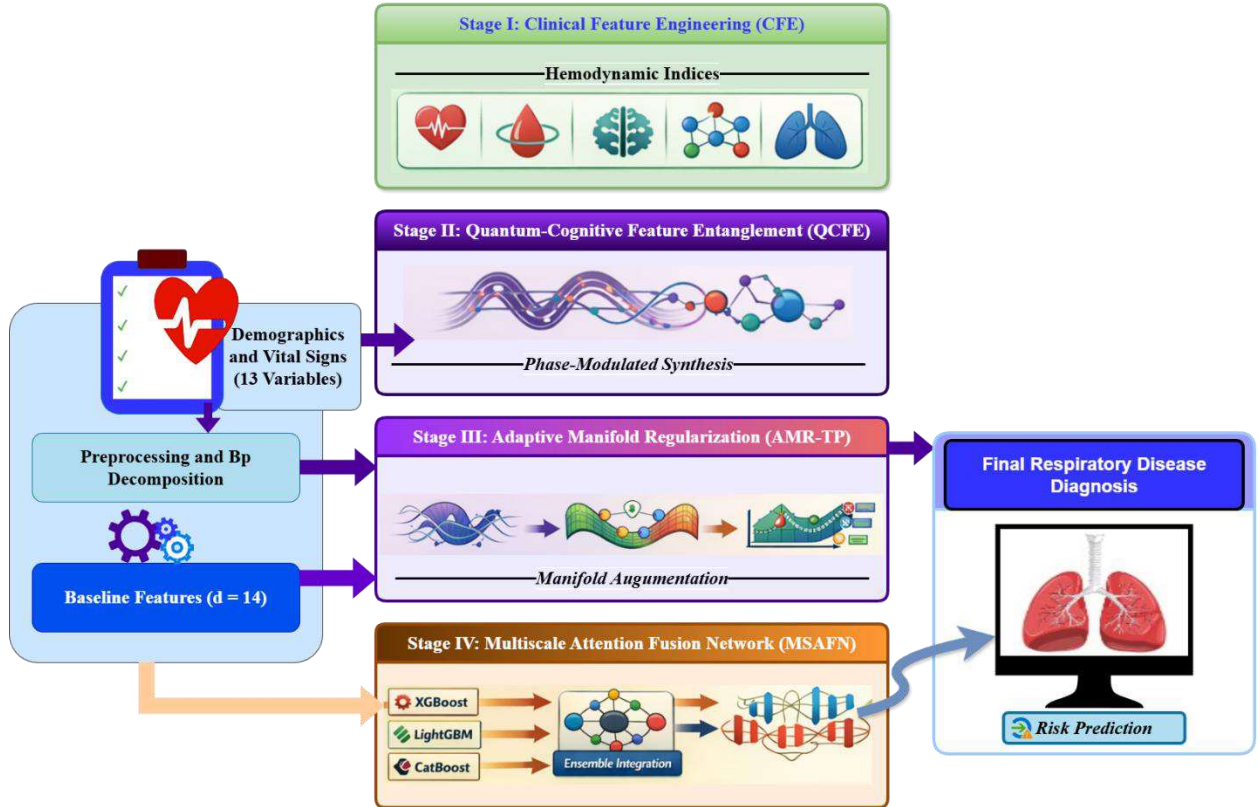


Fig. 1. Q-CogFusion System Architecture.

Throughout this methodology, we adopt the following notation. Let  $X \in \mathbb{R}^{n \times d}$  denote the input feature matrix with  $n$  samples and  $d$  features,  $y \in \{0,1,2,3,4\}^n$  the disease label vector for five respiratory conditions (Bronchitis, Cold, Flu, Healthy, Pneumonia),  $X' \in \mathbb{R}^{n \times d'}$  the QCFE-enhanced feature matrix

with  $d' = d + n_{qcfe}$  dimensions,  $X_{bal} \in \mathbb{R}^{n' \times d'}$  the balanced dataset after AMR-TP with  $n' > n$  samples,  $P_j(x) \in \mathbb{R}^5$  the probability distribution from base learner  $j$ ;  $w_j \in \mathbb{R}$  the attention weight for model  $j$ .

## 4.2 Clinical Feature Engineering

Before quantum-cognitive transformation, we derive five physiologically interpretable hemodynamic features from baseline vital sign measurements. Blood pressure readings recorded in "Systolic/Diastolic" string format undergo decomposition into separate numeric fields  $BP_{Systolic}$  and  $BP_{Diastolic} \in \mathbb{R}$ , enabling independent computational processing and subsequent feature derivation.

### 4.2.1 Pulse Pressure

The pulse pressure quantifies arterial stiffness and left ventricular stroke volume, computed as Eq. (1):

$$PP = BP_{Systolic} - BP_{Diastolic} \quad (1)$$

where  $PP \in \mathbb{R}$  represents the pressure differential in mmHg. Elevated  $PP (> 60 \text{ mmHg})$  indicates arterial stiffening associated with cardiovascular disease progression.

### 4.2.2 Mean Arterial Pressure

The time-weighted average arterial pressure across the cardiac cycle is estimated through Eq. (2):

$$MAP = \frac{BP_{Systolic} + 2 \times BP_{Diastolic}}{3} \quad (2)$$

where  $MAP \in \mathbb{R}$  approximates the perfusion pressure driving blood through systemic circulation, with the factor 2 weighting diastole's proportionally longer duration.

### 4.2.3 Rate-Pressure Product

The non-invasive myocardial oxygen demand index is computed as Eq. (3):

$$RPP = Heart\ Rate_{bpm} \times BP_{Systolic} \quad (3)$$

where  $RPP \in \mathbb{R}$  correlates strongly with direct measurements of myocardial oxygen consumption during exercise or stress conditions.

### 4.2.4 Shock Index

The ratio quantifying hemodynamic compensation status is defined as Eq (4):

$$SI = \frac{Heart\ Rate_{bpm}}{BP_{systolic} + \epsilon} \quad (4)$$

where  $SI \in \mathbb{R}$  assesses cardiovascular reserve with  $\epsilon = 10^{-6}$  preventing numerical instability from division by zero. Clinical thresholds:  $SI > 0.9$  indicates hemodynamic decompensation requiring immediate intervention.

#### 4.2.5 Respiratory Index

The composite measure integrating ventilation efficiency with oxygenation status is computed as Eq. (5):

$$RI = \frac{Respiratory\ Rate_{bpm} \times Oxygen\ Saturation_{\%}}{100} \quad (5)$$

where  $RI \in \mathbb{R}$  normalizes the product to unit scale. This index captures the coupling between ventilatory drive and gas exchange effectiveness, critical for respiratory disease assessment.

Following hemodynamic feature derivation, categorical variables (gender, symptom descriptors, diagnosis labels) undergo label encoding transformation to integer representations  $\{0, 1, 2, \dots, k - 1\}$  where  $k$  denotes the cardinality of each categorical domain.

**Numerical features then undergo two-stage normalization:**

**Robust Scaling** applies interquartile range (IQR) normalization, mitigating outlier influence through Eq. (6):

$$x_{robust} = \frac{x - median(x)}{IQR(x)} \quad (6)$$

where  $IQR(x) = Q3 - Q1$  represents the interquartile range.

**Standard Scaling** subsequently applies z-score normalization, ensuring zero-mean unit-variance distributions Eq. (7):

$$x_{std} = \frac{x_{robust} - \mu}{\sigma} \quad (7)$$

where  $\mu$  and  $\sigma$  denote the sample mean and standard deviation, respectively. This sequential normalization strategy ensures numerical stability for subsequent quantum-cognitive transformations while preserving

physiological relationships encoded in the original measurements. The resulting preprocessed dataset comprises  $n = 2000$  patients with  $d = 14d$  features spanning demographic, vital sign, and symptom domains. Correlation computation excludes purely nominal symptom indicators or applies point-biserial correlation where applicable.

### 4.3 Quantum-Cognitive Feature Entanglement (QCFE)

#### 4.3.1 Conceptual Motivation and Domain Alignment

In quantum mechanics, entanglement describes a non-separable joint state in which correlations between subsystems cannot be factorized into independent components [54]. A canonical example is the bipartite state  $|\psi\rangle = \alpha|00\rangle + \beta|11\rangle$ , where  $|\alpha|^2 + |\beta|^2 = 1$ , and measurement outcomes exhibit structured statistical dependence. While no physical quantum processes are invoked in this work, the notion of entanglement provides a useful abstraction for modeling structured dependencies among clinical variables.

Physiological measurements often exhibit correlated behavior driven by shared regulatory mechanisms. For example, respiratory rate and arterial oxygen saturation are coupled through alveolar gas exchange dynamics, while heart rate and blood pressure interact via autonomic and baroreceptor feedback loops. These dependencies are not strictly additive and may vary in magnitude and sign across patient populations. The QCFE module leverages this observation by treating inter-feature statistical correlations as analogues of entanglement strength, which guide the construction of higher-order feature interactions.

The cognitive aspect of QCFE is inspired by evidence that clinical decision-making emphasizes selective integration of a limited subset of relevant cues rather than uniform consideration of all available information [66]. This principle is operationalized through a correlation-guided *top-k* feature selection mechanism, which functions as an attention-like weighting scheme without assuming an explicit cognitive or neural model.

#### 4.3.2 Formal Definition of the QCFE Transformation

##### Correlation Matrix Computation

Linear statistical dependencies between features are quantified using the Pearson correlation matrix defined in Eq. (8):

$$C = \text{corrcoef}(X^T) \in \mathbb{R}^{d \times d} \quad (8)$$

where the transpose ensures that correlations are computed across feature dimensions. Each entry  $C_{ij}$  of the matrix is computed explicitly according to Eq. (9):

$$C_{ij} = \frac{\sum_{k=1}^n (X_{ki} - \mu_i) (X_{kj} - \mu_j)}{\sqrt{\sum_{k=1}^n (X_{ki} - \mu_i)^2 \cdot \sum_{k=1}^n (X_{kj} - \mu_j)^2}} \quad (9)$$

where  $C \in \mathbb{R}^{d \times d}$  represents the symmetric correlation matrix with entries  $C_{ij} \in [-1, 1]$  quantifying linear association strength,  $\mu_i = \frac{1}{n} \sum_{k=1}^n X_{ki}$  denotes the sample mean of the feature  $i$ , and the summation index  $k$  traverses all  $n$  patient samples as shown in Fig. 2.

### Entanglement Weight Matrix

To enhance dominant inter-feature relationships while suppressing weak or noisy associations, a non-linear entanglement weight matrix is constructed using Eq. (10):

$$W_{ij} = |C_{ij}|^\alpha, \quad \alpha = 0.75 \quad (10)$$

where  $W \in \mathbb{R}^{d \times d}$  derives through element-wise power transformation,  $\alpha = 0.75$  controls entanglement strength, implementing sublinear amplification, and absolute value operation ensures non-negative weights. This transformation amplifies strong correlations ( $|C_{ij}| \rightarrow 1$  yields  $W_{ij} \rightarrow 1$ ) while suppressing weak correlations ( $|C_{ij}| \rightarrow 0$  yields  $W_{ij} \rightarrow 0$  faster than linearly).

### Quantum Phase Vector

To introduce bounded nonlinearity analogous to phase interference, a phase vector is defined in Eq. (11):

$$\phi = [\phi_1, \phi_2, \dots, \phi_d]^T \quad (11)$$

with individual phase components sampled according to Eq. (12):

$$\phi_i \sim \text{Uniform}(0, 2\pi), \quad \forall i \in \{1, \dots, d\} \quad (12)$$

where  $\phi \in \mathbb{R}^d$  contains phase angles with each component  $\phi_i$  independently sampled from a uniform distribution over  $[0, 2\pi]$ , mimicking quantum superposition's random phase relationships as shown in Fig. 2. Random phase modulation functions as a bounded nonlinear mixing mechanism analogous to randomized feature maps, not as a stochastic uncertainty source.

### Phase-Modulated Entanglement Synthesis

For each entangled feature  $i \in \{1, \dots, n_{qcf}\}$ :

$$ref_{idx} = i \bmod d,$$

(13)

$$top_{k_{idx}} = \text{argsort}(W[ref_{idx}, :]) [-5:],$$

(14)

$$\tilde{X}_{entangled}[:, i] = \sum_{top_{k_{idx}}}[:, k] \cdot \cos(\phi_k) \cdot W[ref_{idx}, k],$$

(15)

where Eq. (13) selects the reference feature through modulo operation ensuring cyclic coverage, Eq. (14) identifies the five most strongly entangled features through argsort operation, extracting indices corresponding to the largest weights, and Eq. (15) performs weighted aggregation where  $\cos(\phi_k) \in [-1, 1]$  applies phase modulation, introducing oscillatory coupling, and summation accumulates contributions proportional to entanglement strength  $W[ref_{idx}, k]$ .

### Complete QCFE Output

Horizontal concatenation combines original and synthesized features Eq. (16):

$$\tilde{X} = [X \parallel X_{entangled}] \in \mathbb{R}^{n \times (d + n_{qcf})}$$

(16)

where  $X \in \mathbb{R}^{n \times 14}$  represents original features,  $X_{entangled} \in \mathbb{R}^{n \times 15}$  contains synthesized entangled features with  $n_{qcf} = 15$ , yielding a final enhanced representation  $\tilde{X} \in \mathbb{R}^{n \times 29}$ .

The QCFE transformation is explicitly guided by correlation structure, bounded by construction through phase modulation and power scaling, and preserves the original feature space through concatenation, ensuring information retention and numerical stability for downstream learning.

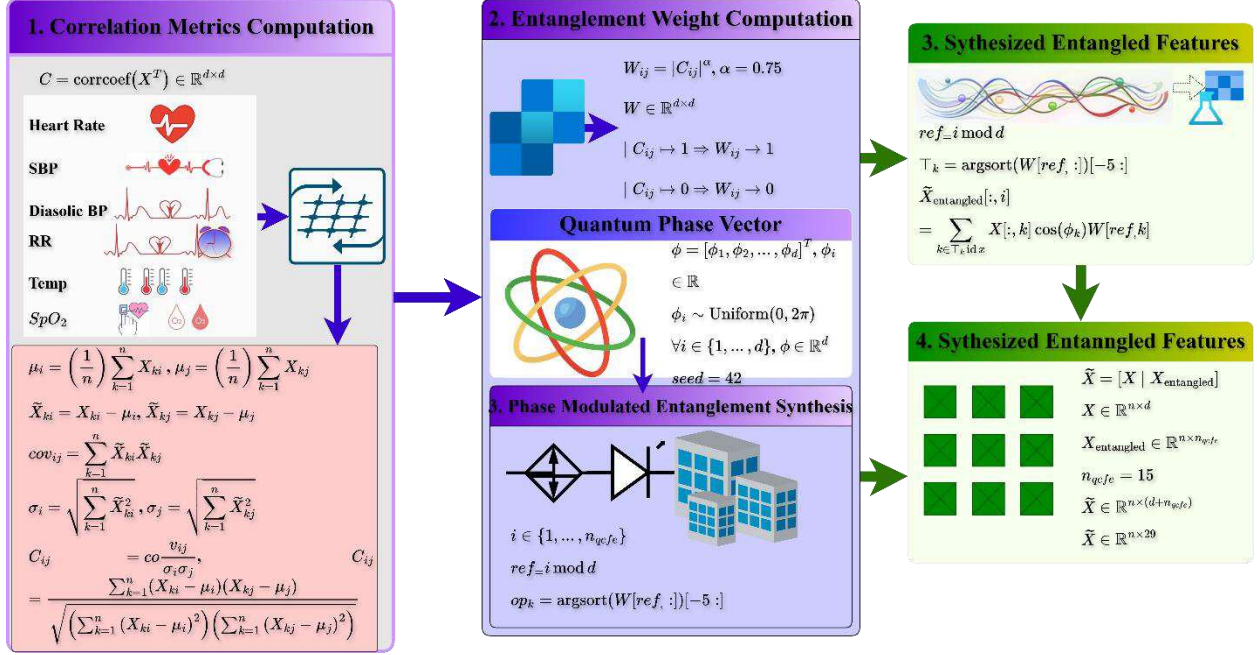


Fig. 2. QCFE Entanglement Process.

## Training-Scope Correlation Computation

To prevent information leakage, the Pearson correlation matrix used for QCFE entanglement weight construction is computed exclusively on the training data within each cross-validation fold. For each fold, correlations are recalculated using only the corresponding training subset, and the resulting entanglement weights are applied unchanged to the validation and test samples. This ensures that no information from validation or test instances influences feature synthesis.

## 4.4 Adaptive Manifold Regularization with Tangent Propagation (AMR-TP)

### 4.4.1 Manifold-Based Geometry of the Clinical Feature Space

The manifold hypothesis posits that high-dimensional medical data reside on low-dimensional smooth manifolds embedded within ambient feature space. Formally, clinical feature vectors  $x \in \mathbb{R}^d$  lie on a smooth  $d_m$  dm-dimensional Riemannian manifold  $\mathcal{M} \subset \mathbb{R}^{d'}$  where  $d_m \ll d'$ , reflecting constrained physiological parameters governed by homeostatic mechanisms maintaining internal equilibrium. Disease-specific sub-manifolds  $\mathcal{M}_c$  exist for each class  $c \in \{0, 1, 2, 3, 4\}$ , with class boundaries defined by manifold topology rather than Euclidean distance.

Conventional SMOTE generates synthetic samples through linear interpolation:  $x_{\text{synthetic}} = x_i + \lambda(x_j - x_i)$  where  $\lambda \in [0, 1]$  uniformly sampled. However, linear paths through ambient space  $\mathbb{R}^{d'}$  do not necessarily remain on manifold  $\mathcal{M}$ , potentially creating samples in physiologically implausible regions exhibiting feature combinations never observed in real patients. Tangent Space Approximation offers superior geometric fidelity: at point  $x \in \mathcal{M}$ , the tangent space  $T_{x(\mathcal{M})}$  provides the optimal local linear

approximation. Movement within  $T_{x(\mathcal{M})}$  guarantees samples remain near the manifold up to second-order curvature terms  $O(\|v\|^2)$ , minimizing geodesic distance distortion compared to ambient space interpolation as in Fig. 3.

#### 4.4.2 Construction of Tangent-Based Synthetic Samples

As shown in Fig. 3, the AMR-TP is activated only for minority classes identified by an imbalance check and proceeds through K-nearest neighbor identification, local tangent space construction, tangent-projected synthesis, and adaptive regularization.

##### Class Imbalance Quantification

For each disease class  $c \in \{0,1,2,3,4\}$  as shown in Eqs. (17) – (19):

$$n_c = |\{i: y_i = c\}|, c \in \{0,1,2,3,4\} \quad (17)$$

$$n_{max} = \max_c n_c \quad (18)$$

$$IR_c = \frac{n_{max}}{n_c} \quad (19)$$

where  $n_c$  denotes sample count for class  $c$ ,  $n_{max}$  identifies majority class size, and  $IR_c \in [1, \infty)$  quantifies relative underrepresentation requiring synthetic augmentation when  $IR_c > 1$ .

##### K-Nearest Neighbor Identification

For each minority sample  $x_i$  where  $y_i = c$  and  $IR_c > 1$  Eq. (20):

$$NN_k(x_i) = \{x_j: y_j = c, x_j \in k - nearest\ neighbors\ of\ x_i\} \quad (20)$$

where  $NN_k(x_i)$  denotes the  $k = 7$  nearest neighbors restricted to same-class samples using Euclidean distance  $d(x_i, x_j) = \|x_i - x_j\|_2$  in enhanced space  $\tilde{\mathbf{X}}$ . Class constraint prevents manifold boundary violations by ensuring synthetic samples interpolate between intra-class points.

##### Local Tangent Space Construction

Singular Value Decomposition (SVD) estimates tangent basis:

$$X_{local} = [x_j - x_i: x_j \in NN_k(x_i)] \in \mathbb{R}^{k \times d'} \quad (21)$$

$$U, \Sigma, V^T = \text{SVD}(X_{\text{local}}) \quad (22)$$

$$T_{x_i} = V[:, :r], \quad r = \min(k-1, d') \quad (23)$$

where  $X_{\text{local}} \in \mathbb{R}^{k \times d'}$  in Eq. (21) constructs a neighborhood deviation matrix with rows representing neighbor displacements centered at  $x_i$ , SVD factorization in Eq. (22) provides orthonormal bases  $U \in \mathbb{R}^{k \times k}$ ,  $\Sigma \in \mathbb{R}^{k \times d'}$ ,  $V \in \mathbb{R}^{d' \times d'}$  where columns of  $V$  span the feature space, and  $T_{x_i} \in \mathbb{R}^{d' \times r}$  in Eq. (23) extracts first  $r$  right singular vectors forming an orthonormal basis for local tangent space approximation with rank  $r$  adaptively determined by neighborhood size.

### Tangent Space Synthetic Sampling

For each required synthetic sample  $s$  from Equ. (24) – (29):

- Randomly select  $x_i$  from the minority class  $c$  (24)

- Randomly select:

$$x_j \in \text{NN}_k(x_i) \quad (25)$$

- Compute direction vector:

$$v = x_j - x_i \quad (26)$$

- Project onto tangent space:

$$v_{\text{tangent}} = T_{x_i}(T_{x_i}^T v) \quad (27)$$

- Sample interpolation weight:

$$\lambda \sim \text{Uniform}(0,1) \quad (28)$$

- Generate:

$$x_{\text{synthetic}} = x_i + \lambda \cdot v_{\text{tangent}} \quad (29)$$

where the projection in Eq. (27) decomposes the displacement  $v$  into the tangent space component  $T_{x_i}(T_{x_i}^T v)$  discarding normal components  $(I - T_{x_i}T_{x_i}^T)v$  orthogonal to the manifold,  $\lambda$  randomization in Eq. (28) introduces diversity, preventing exact duplication, and additive generation in Eq. (29) places samples along the projected geodesic within local tangent space approximation.

### Adaptive Weight Regularization

Distance-based confidence weighting:

$$w_i = \begin{cases} 1.0 & \text{if } x_i \text{ is an original sample} \\ 0.9 \cdot \left(1 - \frac{d_i}{d_{max}}\right) & \text{if } x_i \text{ is a synthetic sample} \end{cases} \quad (30)$$

$$d_i = \min_j \|x_i - x_j\|_2 \quad \text{where } x_j \text{ is original, } y_j = y_i \quad (31)$$

$$d_{max} = \max_i d_i \quad (32)$$

where  $w_i \in [0, 1]$  assigns training weight with original samples receiving unit weight ( $w_i = 1.0$ ), synthetic samples incur distance-proportional penalty where  $d_i$  in Eq. (31) measures minimum Euclidean distance to any original same-class sample quantifying representativeness,  $d_{max}$  in Eq. (32) normalizes distances to  $[0, 1]$  range, and 0.9 scaling factor in Eq. (30) prevents complete exclusion ( $w_i \geq 0$ ) while maintaining preference for authentic measurements over synthesized data.

AMR-TP integrates imbalance detection, tangent-based synthetic sample construction, and adaptive regularization into a geometry-aware augmentation framework. By confining oversampling to locally estimated tangent spaces and down-weighting uncertain samples, the method preserves physiological plausibility and improves class balance without distorting the intrinsic data manifold. Physiological constraints imposed by homeostatic regulation justify the smooth-manifold assumption commonly adopted in biomedical feature spaces.

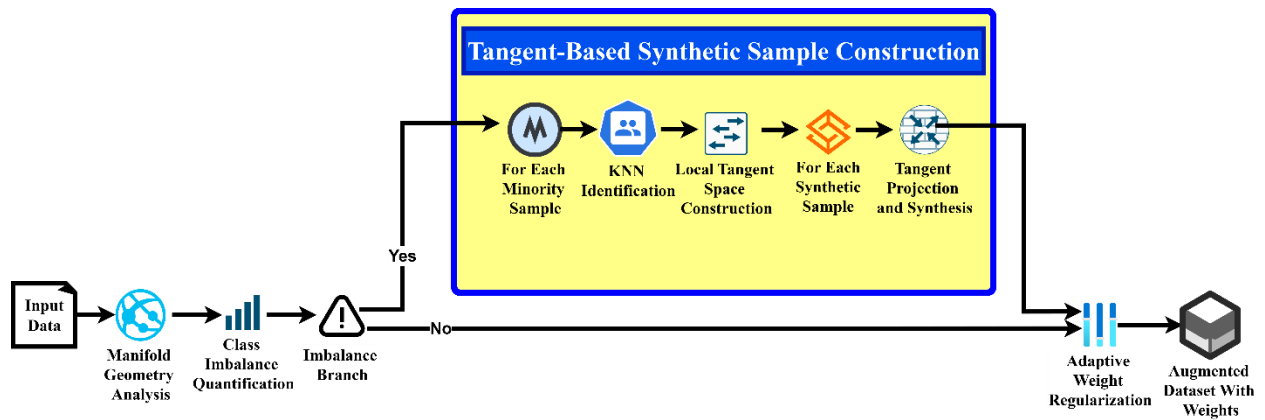


**Fig.3.** Adaptive Manifold Regularization with Tangent Propagation (AMR-TP) for Geometry-Preserving Clinical Data Balancing.

#### 4.5 Multi-Scale Attention Fusion Network (MSAFN)

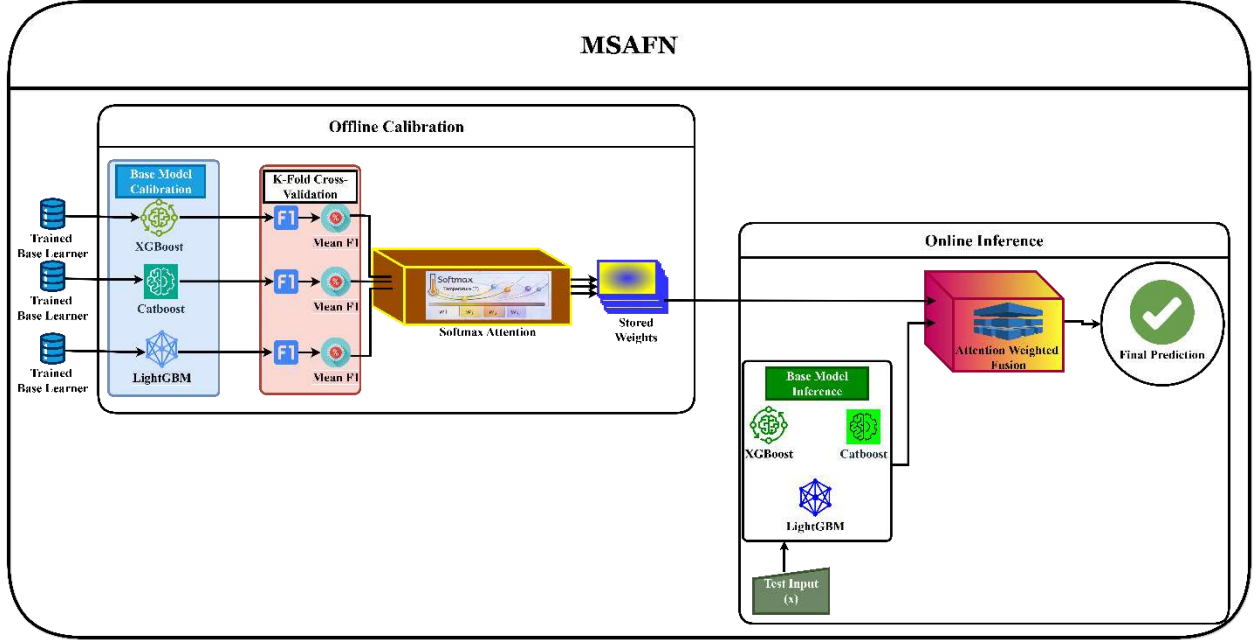


**Fig. 4.** Multi-Scale Attention Fusion Network (MSAFN) Architecture.

##### 4.5.1 Attention-Based Ensemble Calibration

Conventional ensemble strategies typically rely on static aggregation schemes, such as uniform averaging  $w_j = \frac{1}{M}$  or heuristic performance-proportional weighting  $w_j \propto performance_j$  which are defined in Eq. (37). These approaches are inherently limited, as they fail to exploit model-specific reliability patterns revealed through systematic validation and do not adapt to heterogeneous learner behavior. To address this limitation, the proposed Multi-Scale Attention Fusion Network (MSAFN), illustrated in Fig. 4, introduces a dynamic attention mechanism that learns data-driven importance weights, selectively amplifying high-performing base learners while suppressing weaker contributors.

Attention weights are derived using a temperature-scaled softmax formulation (Eq. (36)), which provides a principled probabilistic mechanism for controlling the sharpness of model selection. As the temperature parameter  $T \rightarrow 0$ , the attention distribution converges toward a hard winner-take-all regime approximating  $\arg \max_j s_j$ . Conversely, as  $T \rightarrow \infty$ , the distribution approaches uniform weighting  $w_j = \frac{1}{M}$ . Intermediate temperature values enable a controlled trade-off between exploiting the most reliable models and preserving ensemble diversity. Empirical evaluation on validation data identified  $T = 5.0$  as optimal, selected via grid search over  $T \in \{1, 2, 5, 10, 20\}$ .

MSAFN operates in two distinct phases. During offline calibration, attention weights are learned through stratified  $K - fold$  cross-validation on the training set  $(X_{train}, y_{train})$ , ensuring robust performance estimation defined in Eqs. (33) - (35) ensuring robust assessment under class imbalance. The resulting

weights are then fixed and stored. During online inference, the learned weights are applied to aggregate probabilistic outputs from the base learners for unseen test instances  $x \in X_{test}$ , yielding the final diagnostic prediction through Eq. (40).

#### 4.5.2 Probabilistic Fusion Formulation

##### Cross-Validation Performance Assessment

Let  $j \in \{1,2,3\}$  index the base learners (XGBoost, LightGBM, and CatBoost). For each model, stratified  $K = 5$  fold cross-validation produces a set of macro-averaged F1-scores:

$$S_j = \{s_j^1, s_j^2, \dots, s_j^K\} \quad (33)$$

$$s_j^k = F1_{macro}(model_j, fold_k) \quad (34)$$

$$\tilde{s}_j = \frac{1}{K} \sum_{k=1}^K s_j^k \quad (35)$$

where  $S_j$  in Eq. (33) denotes the set of  $K = 5$  fold-specific F1-scores (macro-averaged across five disease classes), each  $s_j^k$  in Eq. (34) computed on the held-out validation fold  $k$  using predictions from the model  $j$  trained on the remaining  $K - 1$  folds, and  $\tilde{s}_j \in [0,1]$  in Eq. (35) represents mean cross-validation F1-score, providing a robust performance estimate through stratified sampling, ensuring class distribution preservation across folds.

##### Temperature-Scaled Softmax Attention

Attention weight computation through softmax transformation:

$$w = softmax(T \cdot \tilde{s}) \quad (36)$$

$$w_j = \frac{\exp(T \cdot \tilde{s}_j)}{\sum_{m=1}^M \exp(T \cdot \tilde{s}_m)} \quad (37)$$

where  $w = [w_1, w_2, w_3]^T \in \mathbb{R}^3$  in Eq. (36) forms weight vector with  $\sum_{j=1}^3 w_j = 1$  ensuring convex combination, temperature parameter  $T = 5.0$  amplifies performance differences through exponential

scaling, and Eq. (37) gives element-wise expansion where numerator  $\exp(T \cdot \tilde{s}_j)$  exponentially rewards higher-performing models relative to denominator normalization term ensuring probabilistic constraints.

#### Attention-Weighted Probabilistic Fusion

For input instance  $x \in \mathbb{R}^d$ :

$$P_j(x) = \text{model}_j.\text{predict}_{\text{proba}}(x) \in \mathbb{R}^5 \quad (38)$$

$$P_{MSAFN}(x) = \sum_{j=1}^M w_j \cdot P_j(x) \quad (39)$$

$$\hat{y} = \arg \max_{c \in \{0,1,2,3,4\}} P_{MSAFN}(x)_c \quad (40)$$

where Eq. (38) extracts the class probability distribution from the base learner  $j$  with  $P_j(x)_c$  representing the probability of a class  $c$ , Eq. (39) performs weighted aggregation combining distributions through a convex combination  $\sum_j w_j P_j(x)_c$  preserving probability simplex constraints  $\sum_c P_{MSAFN}(x)_c = 1$ , and Eq. (40) selects the disease class with the maximum fused probability as the final prediction  $\hat{y} \in \{0,1,2,3,4\}$ .

By adaptively weighting base learners according to robust cross-validation performance, MSAFN reduces prediction variance and mitigates over-reliance on any single classifier. This behavior is consistent with established ensemble learning theory and contributes to improved generalization performance under heterogeneous clinical data distributions.

## 4.6 Quantum-Entangled SHAP Explainability

To enhance the interpretability of the proposed Q-CogFusion framework, we employ SHapley Additive exPlanations (SHAP), a model-agnostic feature attribution method grounded in cooperative game theory. In contrast to standard SHAP analysis, the proposed approach explicitly accounts for the structured feature dependencies introduced by the Quantum-Cognitive Feature Entanglement (QCFE) module, enabling attribution that reflects both direct and indirect feature contributions.

### 4.6.1 SHAP Value Definition

For a trained prediction function  $f(\cdot)$  operating on the QCFE-enhanced feature space, the SHAP value of feature  $i$  for an input  $x$  is defined in Eq. (41) as the expected marginal contribution of that feature across all possible feature coalitions:

$$\phi_i(x) = \sum_{S \subseteq \mathcal{F} \setminus \{i\}} \frac{|S|! \cdot (|\mathcal{F}| - |S| - 1)!}{|\mathcal{F}|!} \cdot [f(S \cup \{i\}) - f(S)] \quad (41)$$

where  $\mathcal{F}$  represents a complete feature set with  $|\mathcal{F}| = 29$ ,  $S \subseteq \mathcal{F} \setminus \{i\}$  denotes a subset excluding the feature  $i$ , factorial weighting  $\frac{|S|! \cdot (|\mathcal{F}| - |S| - 1)!}{|\mathcal{F}|!}$  normalizes contributions and ensures the efficiency property,  $\sum_i \phi_i(x) = f(x) - f(\text{base})$  (where  $f(\text{base})$  is the expected model output under a reference (background)), and marginal contribution  $f(S \cup \{i\}) - f(S)$  measures the prediction change from adding a feature  $i$  in coalition  $S$ .

#### 4.6.2 Entanglement-Weighted Feature Attribution

Standard SHAP assumes that features contribute independently and additively to the model output. However, QCFE explicitly constructs entangled features through correlation-guided aggregation, as defined previously in Eq. (15), thereby introducing structured inter-feature dependencies. To account for these dependencies, we define an entanglement-weighted SHAP formulation in Eq. (42):

$$\tilde{\phi}_i(x) = \phi_i(x) \cdot E_i \quad (42)$$

where  $\tilde{\phi}_i(x)$  denotes the adjusted SHAP value and  $E_i$  quantifies the entanglement strength associated with feature  $i$ .

The entanglement coefficient  $E_i$  is defined in Eq. (43) based on the role of the feature within the QCFE architecture:

$$E_i = \begin{cases} \frac{1}{d'} \sum_{j=1}^{d'} W_{ij}, & i \in \{1, \dots, d\} \\ \frac{1}{|S(i)|} \sum_{k \in S(i)} W_{ik}, & i \in \{d+1, \dots, d'\} \end{cases} \quad (43)$$

where  $W$  denotes the QCFE entanglement weight matrix defined in Eq. (10), and  $S(i)$  represents the set of original source features contributing to the  $i$ -th entangled feature, as identified by the *top-k* selection procedure in Eq. (14). For original clinical features,  $E_i$  reflects average coupling strength with all other features, whereas for QCFE-generated features, it captures dependence on their contributing source features. The coefficients  $E_i$  are normalized to lie in the interval  $[0,1]$ , ensuring that relative attribution magnitudes

are preserved while modulating importance by entanglement strength. Coefficients are min-max normalized across features within each fold.

#### 4.6.3 Class-Wise Attribution Aggregation

In the multi-class diagnostic setting, SHAP produces class-conditional attributions  $\tilde{\phi}_i^{(c)}(x)$  for each disease class  $c$ . To derive a global, class-agnostic measure of feature importance suitable for population-level interpretability, entanglement-weighted SHAP values are aggregated across classes and samples using arithmetic averaging, as defined in Eq. (44):

$$\Phi_i = \frac{1}{C} \sum_{c=0}^{C-1} \frac{1}{n} \sum_{m=1}^n \tilde{\phi}_i^{(c)}(x_m), \quad C = 5 \quad (44)$$

where  $\Phi_i$  denotes the aggregated importance of feature  $i$ . This aggregation yields a stable global ranking of influential features while retaining the dependency structure introduced by QCfE.

#### 4.6.4 Computational Implementation and Complexity

Tree-based gradient boosting models permit efficient SHAP value computation through TreeExplainer, which exploits the internal tree structure. The computational complexity is approximately  $\mathcal{O}(T \cdot L \cdot D)$ , where  $T$  denotes the number of trees,  $L$  the average number of leaves per tree, and  $D$  the maximum tree depth.

For ensemble-level predictors (Voting, Stacking, and MSAFN), SHAP values are approximated using KernelExplainer, which estimates Shapley values through weighted linear regression. Although the *theoretical* complexity grows exponentially with the feature dimension  $d'$ , practical computation is achieved through Monte Carlo sampling using a finite background set. In this study, SHAP values are computed for  $n_{SHAP} = 50$  representative test instances using  $B = 100$  background samples, providing a balance between interpretability, fidelity, and computational feasibility.

## 5 Experimental Validation

This section describes the experimental framework, including dataset specifications, evaluation criteria, and statistical procedures.

### 5.1 Dataset Characteristics

The Disease Diagnosis Dataset [67] contains  $n = 2000$  patients across five respiratory conditions: Bronchitis, Cold, Flu, Healthy, and Pneumonia. The class distribution is highly imbalanced, with Healthy comprising 1167 cases (58.4%) and Pneumonia only 44 cases (2.2%), yielding a maximum imbalance ratio of  $IR_{max} = \frac{1167}{44} \approx 26.5:1$ . The dataset includes 13 baseline clinical variables: demographics (Age,

Gender), vital signs (Heart Rate, Blood Pressure, Respiratory Rate, Body Temperature, Oxygen Saturation), and three categorical symptom indicators. Blood pressure values were decomposed into numeric systolic and diastolic components, and the Patient\_ID field was removed, resulting in  $d = 14$  input features.

Missing data were handled using clinically informed imputation strategies. Continuous physiological variables were imputed using median values to ensure robustness to skewed distributions and outliers, while categorical symptom indicators were completed using mode-based imputation. Outliers were identified using interquartile range (IQR) analysis and mitigated through winsorization to maintain physiological plausibility without altering the underlying distribution. All continuous features were subsequently standardized to zero mean and unit variance to ensure scale consistency for correlation analysis, QCfE transformation, and gradient-boosting model optimization.

## 5.2 AMR-TP Dataset Balancing Analysis

Adaptive Manifold Regularization with Tangent Propagation (AMR-TP) (Section 3.4) was applied to address the extreme class imbalance. Synthetic samples were generated exclusively for minority classes via local tangent space augmentation, preserving the intrinsic manifold geometry of each disease class.

Fig. 5 illustrates the effect of AMR-TP-based dataset balancing. The original distribution exhibits extreme skew toward the Healthy class, while Pneumonia and Cold are critically underrepresented. AMR-TP generates synthetic samples exclusively for minority classes using tangent-space augmentation (Section 3.4), expanding the dataset from  $n = 2000$  to  $n' = 5835$ . This results in strict per-class equality, with  $n'_c = 1167$  for all  $c \in \{0,1,2,3,4\}$ , ensuring that model optimization does not converge to majority-class-biased solutions.

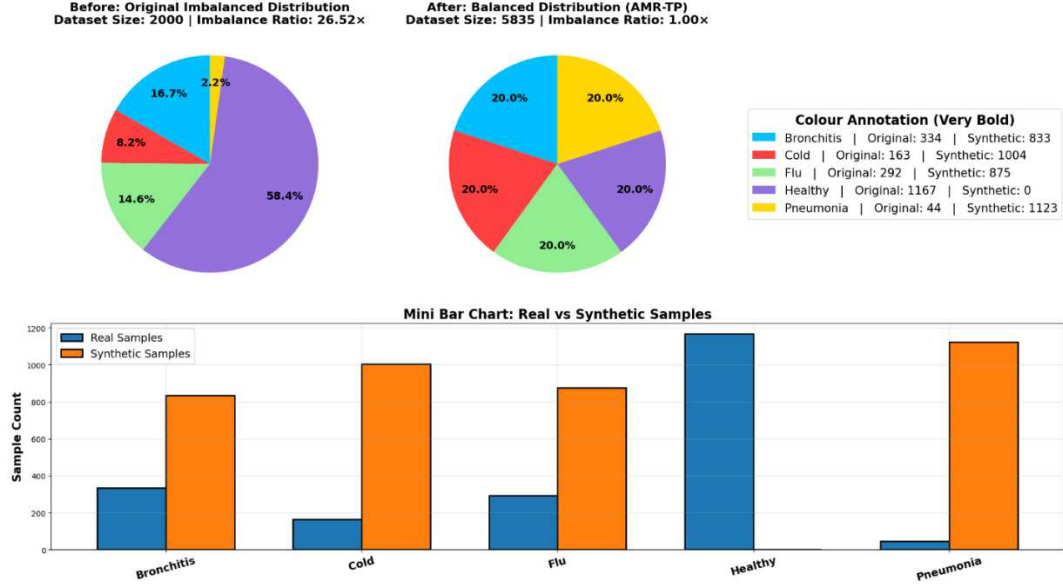


Fig. 5. Class Distribution Before and After AMR-TP-Based Dataset Balancing.

### 5.3 Quantum-Cognitive Feature Entanglement (QCFE) Output Characteristics

The QCFE transformation (Section 3.3) synthesized 15 entangled features through correlation-driven phase-modulated aggregation, expanding the feature space from  $d = 14$  to  $d' = 29$  dimensions. The entanglement process selected the top-5 most strongly correlated features for each reference index (Eq. 14), applying cosine-modulated phase weighting to introduce controlled non-linearity while preserving correlation structure.

Fig. 6 reveals that QCFE features exhibit diverse distributional characteristics, capturing distinct aspects of cardiopulmonary physiology. QCFE\_6 demonstrates clear bimodal separation with healthy individuals concentrated near positive values (+0.5 to +1.2) and diseased states clustered in negative regions (-1.5 to -0.5), reflecting preserved versus disrupted physiological coupling. QCFE\_7 displays trimodal structure, distinguishing mild respiratory infections (Cold: peak at -3.0), moderate conditions (Bronchitis, Flu: peaks at -1.0 to 0.0), and severe infections (Pneumonia: peak at +2.0). These emergent patterns validate that phase-modulated correlation synthesis captures latent disease-specific signatures not observable in individual baseline features.

Fig. 7 demonstrates that the QCFE-enhanced feature representation ( $d' = 29$ ) exhibits clear geometric separation in the first three principal components, capturing 68.53% of total variance. Healthy individuals form a dense cluster in the positive PC1, negative PC2 region, reflecting homeostatic cardiovascular-respiratory coupling. Disease classes occupy distinct subspaces: Pneumonia extends along positive PC1 and positive PC3 (severe cardiopulmonary compromise), Bronchitis and Flu overlap partially in moderate PC1-PC2 regions (intermediate respiratory dysfunction), while Cold occupies negative PC1 space (mild

upper respiratory symptoms). Principal Component Analysis (PCA) was fitted only on the training data, and the learned projection was subsequently applied to validation and test samples to avoid information leakage. PCA is used solely for visualization and does not influence model training.

The partial overlap between Bronchitis and Flu aligns with clinical reality; these conditions share symptomatic presentations (elevated respiratory rate, moderate fever, productive cough) that challenge differential diagnosis without laboratory confirmation. The clear separation of Pneumonia validates that severe lower respiratory infections induce distinct physiological derangements (profound hypoxemia, tachycardia, shock index elevation) detectable through multivariate feature analysis. This geometric structure explains why gradient boosting classifiers achieve exceptional performance: decision boundaries in this PCA-projected space exhibit low curvature, enabling axis-aligned tree splits to effectively partition the feature space.

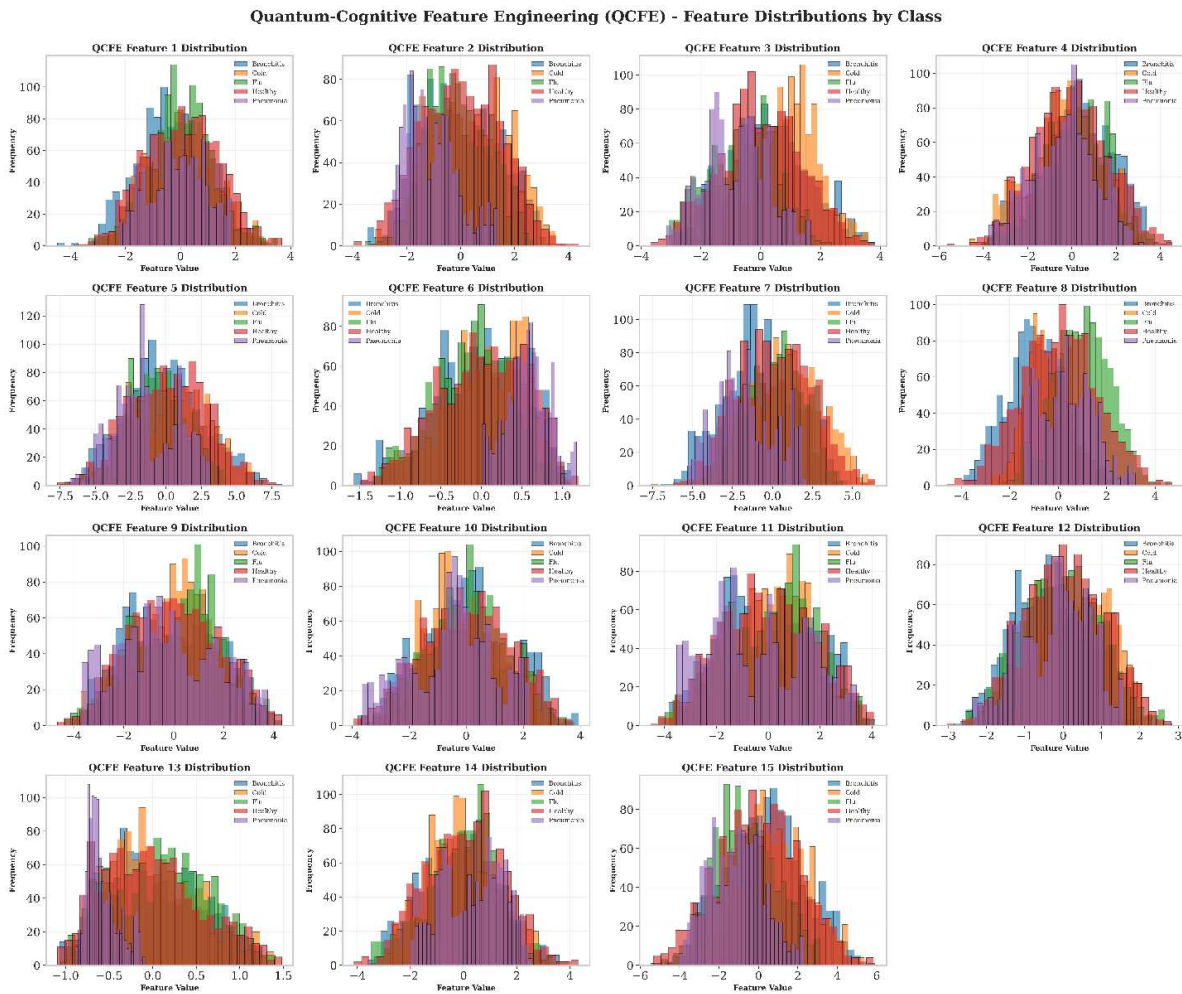


**Fig. 6. QCFE Feature Distribution Analysis.**

Total Variance Explained: 68.54%

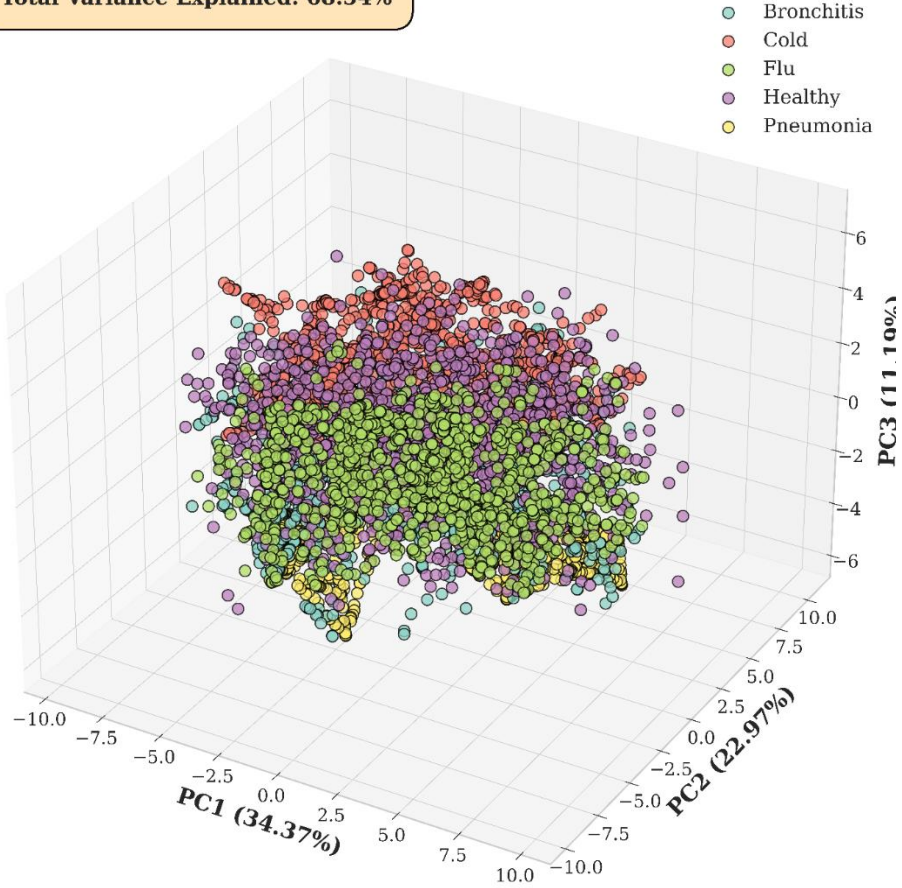


Fig. 7. 3D PCA Visualization of Enhanced Feature Space.

#### 5.4 Train–Test Partitioning and Stratification Protocol

Stratified random splitting maintains class proportions with an 80:20 train–test ratio, yielding  $n_{train} = 1600$  original training samples (expanding to  $n'_{train} = 4668$  post-AMR-TP) and  $n_{test} = 400$  held-out samples (expanding to  $n'_{test} = 1167$  for balanced evaluation) as shown in Fig. 5. A fixed random seed of 42 ensures reproducibility across experimental trials. Stratification guarantees  $\frac{n_c^{train}}{n_c} = 0.8 \pm 0.01$  for all classes  $c$ .

#### 5.5 Cross-Validation Strategy

Stratified 5-Fold Cross-Validation preserves class distribution within each fold: the training set is divided into five equal folds, with four folds ( $\approx 1280$  samples) used for training and one-fold ( $\approx 320$  samples) used for validation. This procedure is repeated five times, with each fold serving as the validation set exactly once. The evaluated metrics include accuracy, precision (macro), recall (macro), F1-score (macro), MCC, and ROC-AUC (One-vs-Rest, macro-averaged).

#### Leakage Prevention Protocol

All data-driven transformations, including QCFE correlation computation, AMR-TP tangent-space augmentation, ensemble calibration, and PCA visualization, are performed strictly within the training data for each cross-validation fold. Synthetic samples are generated only for training sets and are never introduced into validation or test sets. Validation and test data consist exclusively of original patient records and are used solely for performance evaluation.

## 5.6 Evaluation Metrics

To ensure comprehensive and unbiased performance assessment under multi-class and class-imbalanced conditions, seven complementary evaluation metrics are employed. All metrics are derived from the confusion matrix and are defined formally below.

### Confusion Matrix Definition

Let the confusion matrix for a given class be represented by the tuple Eq. (45):

$$C = \{TP_c, TN_c, FP_c, FN_c\} \quad (45)$$

Where:  $TP$  denotes true positives for class  $c$ ,  $TN$  denotes true negatives for class  $c$ ,  $FP$  denotes false positives for class  $c$ , and  $FN$  denotes false negatives for class  $c$ . All subsequent metrics are deterministic functions of  $C_c$ .

### Accuracy

Accuracy measures the overall proportion of correctly classified instances. While intuitive, accuracy is sensitive to class imbalance and is therefore interpreted alongside imbalance-robust metrics, as formalized in Eq. (46):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (46)$$

### Precision (Macro-Averaged)

Macro-averaged precision evaluates predictive exactness by averaging class-wise precision equally across all  $C$  classes, thereby preventing dominance by majority classes. The formal expression is given in Eq. (47):

$$Precision_{Macro} = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FP_c} \quad (47)$$

### Recall (Macro-Averaged)

Macro-averaged recall quantifies the ability to correctly identify instances of each class and is computed according to Eq. (48):

$$Recall_{Macro} = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FN_c} \quad (48)$$

### Specificity

Specificity measures the true negative rate, indicating correct rejection of non-target classes as shown in Eq. (49):

$$Specificity = \frac{TN_c}{TN_c + FP_c} \quad (49)$$

### F1-Score (Macro-Averaged)

The macro F1-score provides a harmonic balance between precision and recall at the class level, as expressed in Eq. (50):

$$F1_{Macro} = \frac{1}{C} \sum_{c=1}^C \frac{2 \cdot Precision_c \times Recall_c}{Precision_c + Recall_c} \quad (50)$$

### Matthews Correlation Coefficient (MCC)

The Matthews Correlation Coefficient provides a **balanced and symmetric performance measure**, incorporating all confusion matrix elements and remaining reliable under severe class imbalance Eq. (51):

$$MCC = \frac{(TP_c \cdot TN_c) - (FP_c \cdot FN_c)}{\sqrt{(TP_c + FP_c)(TP_c + FN_c)(TN_c + FP_c)(TN_c + FN_c)}} \quad (51)$$

MCC values lie in the interval  $[-1, +1]$ , where: +1 indicates perfect prediction, 0 indicates random prediction, and -1 indicates complete disagreement.

MCC preserves interpretability across skewed class distributions due to its symmetric treatment of all confusion-matrix components.

### ROC-AUC (One-vs-Rest, Macro-Averaged)

For the multi-class setting with  $C = 5$  disease categories, ROC–AUC is computed using a One-vs-Rest (OvR) decomposition. Class-wise ROC–AUC follows the formulation in Eqs. (52)–(53):

For each class  $c$ :

$$ROC - AUC_c = \int_0^1 TPR_c(t) d[FPR_c(t)] \quad (52)$$

Where;

$$TPR_c(t) = \frac{TP_c(t)}{TP_c(t) + FN_c(t)}, \quad FPR_c(t) = \frac{FP_c(t)}{FP_c(t) + TN_c(t)} \quad (53)$$

The macro-averaged ROC–AUC is defined as:

$$ROC - AUC_{Macro} = \frac{1}{C} \sum_{c=1}^C ROC - AUC_c \quad (54)$$

This averaging strategy ensures equal weighting across disease classes, independent of prevalence. ROC–AUC values range from 0.5 (random discrimination) to 1.0 (perfect discrimination).

### Bootstrap Confidence Intervals

Metric uncertainty is quantified using non-parametric bootstrap resampling with  $B = 500$  iterations. The bootstrap estimation procedure and confidence interval construction are specified in Eqs. (55)–(56). For a given metric  $\theta$ :

$$\theta^{*(b)} = \text{metric} \left( y_{test}^{*(b)}, \hat{y}_{test}^{*(b)} \right), \quad b = 1, \dots, B \quad (55)$$

where  $y_{test}^{*(b)}$  and  $\hat{y}_{test}^{*(b)}$  are the  $b$  –  $th$  resampled true and predicted labels. The 95% confidence interval is obtained through the percentile method:

$$CI_{95\%}(\theta) = [\text{percentile}_{2.5}(\{\theta^{*(b)}\}), \text{percentile}_{97.5}(\{\theta^{*(b)}\})] \quad (56)$$

Bootstrap sampling captures the uncertainty due to finite test set size, providing robust confidence intervals for all evaluation metrics.

## 5.7 Other Ensemble Configurations

Beyond MSAFN, the Q-CogFusion framework implements four complementary ensemble strategies to benchmark adaptive attention weighting against conventional aggregation methods.

### 5.7.1 Voting Ensemble (Soft Voting)

The Voting ensemble (Eq. (57)) aggregates predictions from the three gradient boosting base learners using weighted averaging of class probability distributions:

$$P_{Voting}(x) = \sum_{j=1}^3 w_j \cdot P_j(x) \quad (57)$$

where  $P_j(x) \in \mathbb{R}^5$  denotes the probability distribution from model  $j \in \{XGBoost, LightGBM, CatBoost\}$ , and  $w_1 = 1.3$ ,  $w_2 = 1.0$ ,  $w_3 = 1.2$  represent fixed heuristic weights assigned based on preliminary validation performance. Unlike MSAFN, which computes performance-driven weights using a temperature-scaled softmax, the Voting ensemble uses static, predetermined weights that remain constant across datasets. This provides a baseline for evaluating the benefits of adaptive attention mechanisms.

### 5.7.2 DualBoost Ensemble

DualBoost creates a specialized two-model ensemble combining XGBoost and LightGBM through weighted soft voting Eq. (58):

$$P_{DualBoost}(x) = \frac{w_{XGB} \cdot P_{XGB}(x) + w_{LGB} \cdot P_{LGB}(x)}{w_{XGB} + w_{LGB}} \quad (58)$$

where  $w_{XGB} = 1.2$  and  $w_{LGB} = 1.0$  denote fixed weights normalized to form a convex combination. This ensemble leverages the architectural complementarity between histogram-based tree construction (XGBoost) and leaf-wise growth strategies (LightGBM), while excluding CatBoost's ordered boosting mechanism. Reducing the ensemble size from three to two models decreases computational overhead by 33% while maintaining diversity through distinct split-finding algorithms and regularization schemes.

### 5.7.3 Stacking Ensemble

The Stacking ensemble employs a two-layer architecture where base learners' predictions serve as meta-features for a Random Forest meta-learner:

#### Layer 1 (Base Learners)

Train XGBoost, LightGBM, and CatBoost independently on  $(X_{train}, y_{train})$ , generating out-of-fold predictions through stratified 5-fold cross-validation to prevent information leakage as shown in Eq. (59):

$$Z_{train} = [P_{XGB}^{oof}, P_{LGB}^{oof}, P_{CAT}^{oof}] \in \mathbb{R}^{n \times 15} \quad (59)$$

where each  $P_j^{oof} \in \mathbb{R}^{n \times 5}$  contains class probabilities for all  $n$  training samples obtained through cross-validation, ensuring that each sample's meta-feature is derived from models that never observed that sample during training.

### Layer 2 (Meta-Learner)

A Random Forest classifier with 200 trees learns optimal probability combination rules:

$$\hat{y} = \arg \min_c \left( RF(z(x)) \right)_c \quad (60)$$

where  $z(x) = [P_{XGB}(x), P_{LGB}(x), P_{CAT}(x)] \in \mathbb{R}^{n \times 15}$  concatenates base learner probability outputs at test time. The meta-learner automatically captures non-linear relationships among base predictions. Stratified 5-fold cross-validation ensures meta-features are generated without leakage, and parallel processing  $n_{jobs} = -1$  accelerates computation.

Stacking differs from MSAFN in that it relies on a hierarchical architecture with a separate meta-learning phase, whereas MSAFN computes adaptive weights directly from validation performance using a closed-form softmax transformation without additional model training.

## 5.8 Base Learner Configurations

For comprehensive benchmarking, we include traditional machine learning classifiers with optimized hyperparameters. The learning architecture also employs three heterogeneous gradient-boosting classifiers, XGBoost, LightGBM, and CatBoost, to maximize ensemble diversity through complementary tree-construction strategies, regularization mechanisms, and optimization dynamics.

### 5.8.1 Decision Tree (DT)

CART algorithm with Gini impurity criterion, maximum depth = 8, minimum samples per split = 2, class\_weight = 'balanced' to handle residual imbalance after AMR-TP, random\_state = 42.

### 5.8.2 Random Forest (RF)

Ensemble of 300 decision trees with bootstrap aggregating, maximum depth = 8,  $\sqrt{d'} \approx 5$  features per split (sqrt feature selection), class\_weight = 'balanced',  $n_{jobs} = -1$  for parallel training, random\_state = 42.

### 5.8.3 K-Nearest Neighbors (KNN)

Instance-based learning with  $k = 5$  neighbors, Euclidean distance metric ( $L_2$  norm), uniform weighting (no distance-based decay),  $n_{jobs} = -1$  for parallel distance computation.

### 5.8.4 Extra Trees (ET)

Extremely randomized trees ensemble with 300 estimators, maximum depth = 8, full feature consideration at each split ( $d' = 29$ ), class\_weight = 'balanced', random split selection (no optimal threshold search),  $n_{jobs} = -1$ , random\_state = 42.

### 5.8.5 Multi-Layer Perceptron (MLP)

Two hidden layers [100, 50] with ReLU activation functions, Adam optimizer (default learning rate  $\alpha = 0.001$ ), maximum epochs = 500, early stopping disabled, random\_state = 42 (Eqs. (61) – (63)). The architecture implements:

$$h^{(1)} = \text{ReLU}(W^{(1)}x + b^{(1)}) \in \mathbb{R}^{100} \quad (61)$$

$$h^{(2)} = \text{ReLU}(W^{(2)}h^{(1)} + b^{(2)}) \in \mathbb{R}^{50} \quad (62)$$

$$\hat{y} = \text{softmax}(W^{(3)}h^{(2)} + b^{(3)}) \in \mathbb{R}^5 \quad (63)$$

All models utilize identical train-test splits ( $n'_{train} = 4668, n'_{test} = 1167$ ) derived from 80 :20 stratified partitioning and employ stratified 5-fold cross-validation for performance estimation, with random\_state = 42 ensuring reproducibility across all experiments. The gradient boosting models (XGBoost, LightGBM, CatBoost) share consistent hyperparameters (500 estimators, depth = 8, learning\_rate = 0.08, subsample = 0.8) to isolate algorithmic differences from configuration effects.

### 5.8.6 XGBoost

XGBoost implements gradient-boosted decision trees using histogram-based split finding to accelerate training while maintaining numerical stability. The model is configured with 500 estimators, maximum tree depth of 8, learning rate 0.08, subsampling ratio 0.8, and full column sampling. The histogram tree method reduces split-search complexity from  $O(n \cdot d' \cdot \log n)$  to  $O(n \cdot d' \cdot bins)$ , with 256 discrete bins, yielding computational efficiency for the post-augmentation training set ( $n' = 4,668$ ). Shrinkage regularization through the learning rate and stochastic row sampling reduces overfitting, while deterministic histogram construction ensures reproducibility.

### 5.8.7 LightGBM

LightGBM employs a leaf-wise tree growth strategy combined with histogram optimization, enabling rapid loss minimization by prioritizing high-information splits. The model is trained with 500 estimators, depth constraint of 8, learning rate 0.08, and subsampling ratio 0.8. Residual class imbalance is mitigated through inverse-frequency weighting, defined as Eq. (64):

$$w_c = \frac{n}{C \cdot n_c} \quad (64)$$

where  $n$  denotes total samples,  $C$  the number of classes, and  $n_c$  the number of samples in class  $c$ . A minimum leaf size constraint enforces regularization on minority classes and prevents over-fragmentation.

### 5.8.8 CatBoost

CatBoost incorporates ordered boosting with symmetric tree structures to eliminate target leakage and stabilize gradient estimation. The model uses 500 iterations, depth 8, learning rate 0.08, Bayesian bootstrap sampling, and L2 regularization on leaf values. Although feature inputs are continuous post-scaling, ordered boosting improves robustness under limited sample conditions. Automatic class weighting mirrors the inverse-frequency strategy applied in LightGBM, ensuring consistency across learners.

The ensemble integrates these learners to exploit architectural diversity, level-wise versus leaf-wise growth, deterministic versus Bayesian sampling, and distinct regularization schemes, thereby reducing variance and improving generalization in accordance with bias–variance decomposition principles.

## 5.9 Statistical Significance Testing

Model comparisons are conducted under identical data partitions to ensure paired statistical validity. Performance differences are evaluated using complementary parametric and non-parametric tests.

McNemar’s test assesses whether two classifiers exhibit statistically equivalent error rates on the same test set. Let the contingency table be defined as:

$$N = \begin{bmatrix} n_{00} & n_{01} \\ n_{10} & n_{11} \end{bmatrix} \quad (65)$$

where  $n_{01}$  and  $n_{10}$  denote discordant predictions. The continuity-corrected McNemar statistic is computed according to Eq. (66):

$$\chi_{Mc}^2 = \frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10} + \epsilon}$$

(66)

with  $\epsilon = 10^{-10}$  preventing division by zero. Statistical significance is determined through Eq. (67):

$$p = P(\chi^2(1) \geq \chi_{Mc}^2) \quad (67)$$

and the null hypothesis of equal error rates is rejected at the significance level  $\alpha = 0.05$ .

To avoid distributional assumptions across cross-validation folds, the Wilcoxon signed-rank test is applied to paired metric differences  $d_i$ . Signed ranks are accumulated as Eq. (68):

$$R_i^\pm = \text{rank}(|d_i|) I(\pm d_i > 0) \quad (68)$$

leading to rank sums Eq. (69):

$$W^\pm = \sum_{i=1}^{n_{fold}} R_i^\pm, \quad W = \min(W^+, W^-) \quad (69)$$

and standardized using Eq. (70):

$$z = \frac{W - \frac{n_{fold}(n_{fold} + 1)}{4}}{\sqrt{\frac{n_{fold}(n_{fold} + 1)(2n_{fold} + 1)}{24}}} \quad (70)$$

which asymptotically follows a standard normal distribution.

For ablation analysis, paired t-tests are conducted on fold-wise F1-score differences. The mean and variance of paired differences, as shown in Eqs. (71) – (72), are defined as:

$$\bar{d} = \frac{1}{n_{fold}} \sum_{i=1}^{n_{fold}} d_i \quad (71)$$

$$s_d = \sqrt{\frac{1}{n_{fold} - 1} \sum_{i=1}^{n_{fold}} (d_i - \bar{d})^2}$$

(72)

yielding the test statistic in Eq. (73):

$$\mathbf{t} = \frac{\bar{d}}{s_d/\sqrt{n_{fold}}} \sim t(n_{fold} - 1) \quad (73)$$

Practical significance is quantified using Cohen's  $d$  as expressed in Eq. (74),

$$d = \frac{\mu_1 - \mu_2}{\sigma_{pooled}} \quad (74)$$

where the pooled standard deviation is Eq. (75):

$$\sigma_{pooled} = \sqrt{\frac{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}{n_1 + n_2 - 2}} \quad (75)$$

Effect magnitudes are interpreted using standard thresholds to complement hypothesis testing.

### 5.10 Computational Environment

All experiments were conducted using Kaggle Notebooks equipped with an NVIDIA Tesla P100 GPU (16 GB VRAM), 16 GB of system memory, and an Intel Xeon CPU operating at 2.30 GHz with four cores. The software environment comprised Python 3.10.12 with the following key libraries: scikit-learn 1.3.2, XGBoost 2.0.3, LightGBM 4.1.0, CatBoost 1.2.2, SHAP 0.43.0, NumPy 1.24.3, Pandas 2.0.3, Matplotlib 3.7.2, and Seaborn 0.12.2. This configuration ensured reproducible training and evaluation of all models.

## 6 Results

This section presents a comprehensive empirical evaluation of the Q-CogFusion framework for multi-class diagnosis of respiratory diseases. We report test set performance, cross-validation robustness, ablation studies quantifying component contributions, and interpretability analyses through quantum-entangled SHAP visualizations.

### 6.1 Test Set Performance

Table 1 and Fig. 8 presents the test set performance of all evaluated models on the balanced dataset ( $n' = 1,167$  samples, five disease classes), achieved through AMR-TP synthetic augmentation to mitigate class

imbalance. The Q-CogFusion framework's top-performing models consistently achieved excellent discrimination across all evaluation metrics.

The MSAFN ensemble achieved the highest test accuracy of 99.40%, outperforming all individual base learners and alternative ensemble strategies. This superior performance stems from temperature-scaled softmax attention weighting ( $T = 5.0$ ) that dynamically emphasizes high-performing models while maintaining ensemble diversity. LightGBM demonstrated exceptional individual performance (99.31%), attributable to its leaf-wise tree growth strategy and inverse-frequency class weighting mechanism. Traditional machine learning approaches (DT, RF, ET, KNN) exhibited substantially lower performance (83.03%–89.55%), validating the necessity of gradient boosting methods for this diagnostic task.

Balanced evaluation is used to ensure fair comparison across disease classes, whereas the unbalanced baseline in the ablation study reflects clinical prevalence and provides evidence of real-world diagnostic behavior.

**Table 1.** Test Set Performance Summary Across All Models.

| Model     | Accuracy (%) | Precision (%) | Recall (%) | Specificity (%) | F1-Score (%) | MCC    | ROC-AUC |
|-----------|--------------|---------------|------------|-----------------|--------------|--------|---------|
| MSAFN     | 99.40        | 99.40         | 99.40      | 99.85           | 99.40        | 0.9925 | 0.9999  |
| LightGBM  | 99.31        | 99.32         | 99.32      | 99.83           | 99.31        | 0.9914 | 0.9999  |
| Voting    | 99.23        | 99.23         | 99.23      | 99.81           | 99.23        | 0.9904 | 0.9999  |
| DualBoost | 99.23        | 99.23         | 99.23      | 99.81           | 99.23        | 0.9904 | 0.9999  |
| Stacking  | 99.23        | 99.23         | 99.23      | 99.81           | 99.23        | 0.9904 | 0.9998  |
| XGBoost   | 99.14        | 99.15         | 99.14      | 99.79           | 99.14        | 0.9893 | 0.9999  |
| CatBoost  | 98.97        | 98.97         | 98.97      | 99.74           | 98.97        | 0.9872 | 0.9996  |
| MLP       | 90.83        | 90.76         | 90.85      | 97.71           | 90.39        | 0.8873 | 0.9785  |
| DT        | 89.55        | 89.57         | 89.56      | 97.39           | 89.50        | 0.8697 | 0.9713  |
| RF        | 88.52        | 89.92         | 88.54      | 97.13           | 87.84        | 0.8621 | 0.9878  |
| ET        | 83.89        | 85.88         | 83.93      | 95.97           | 81.93        | 0.8083 | 0.9776  |
| KNN       | 83.03        | 84.39         | 83.07      | 95.76           | 80.15        | 0.7997 | 0.9529  |

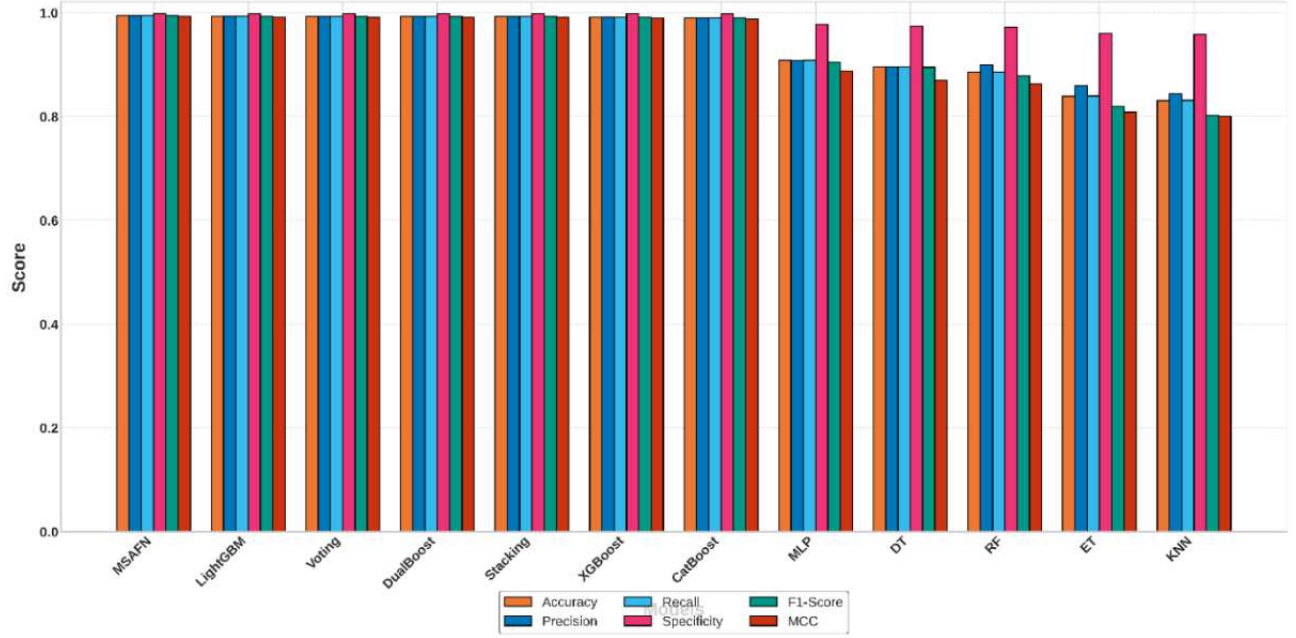


**Fig. 8.** Comprehensive Performance Analysis Across All Models.

## 6.2 Cross-Validation Performance

**Table 2.** Cross-Validation Performance Summary (Mean  $\pm$  Standard Deviation).

| Model     | CV Accuracy (%)  | CV Precision (%) | CV Recall (%)    | CV F1-Score (%)  | CV MCC              | CV ROC-AUC          |
|-----------|------------------|------------------|------------------|------------------|---------------------|---------------------|
| Stacking  | 99.08 $\pm$ 0.21 | 99.08 $\pm$ 0.21 | 99.08 $\pm$ 0.21 | 99.08 $\pm$ 0.21 | 0.9885 $\pm$ 0.0026 | 0.9995 $\pm$ 0.0004 |
| DualBoost | 99.04 $\pm$ 0.36 | 99.04 $\pm$ 0.36 | 99.04 $\pm$ 0.36 | 99.03 $\pm$ 0.36 | 0.9880 $\pm$ 0.0045 | 0.9996 $\pm$ 0.0002 |
| LightGBM  | 99.01 $\pm$ 0.34 | 99.02 $\pm$ 0.34 | 99.01 $\pm$ 0.34 | 99.01 $\pm$ 0.34 | 0.9877 $\pm$ 0.0043 | 0.9997 $\pm$ 0.0001 |
| Voting    | 98.97 $\pm$ 0.34 | 98.98 $\pm$ 0.33 | 98.97 $\pm$ 0.34 | 98.97 $\pm$ 0.34 | 0.9872 $\pm$ 0.0042 | 0.9995 $\pm$ 0.0003 |
| MSAFN     | 98.93 $\pm$ 0.21 | 98.93 $\pm$ 0.21 | 98.93 $\pm$ 0.21 | 98.93 $\pm$ 0.22 | 0.9866 $\pm$ 0.0027 | 0.9995 $\pm$ 0.0003 |
| XGBoost   | 98.86 $\pm$ 0.16 | 98.87 $\pm$ 0.16 | 98.86 $\pm$ 0.16 | 98.86 $\pm$ 0.16 | 0.9858 $\pm$ 0.0020 | 0.9996 $\pm$ 0.0002 |
| CatBoost  | 97.75 $\pm$ 0.42 | 97.79 $\pm$ 0.40 | 97.75 $\pm$ 0.42 | 97.74 $\pm$ 0.43 | 0.9720 $\pm$ 0.0052 | 0.9991 $\pm$ 0.0003 |
| MLP       | 90.60 $\pm$ 0.74 | 90.53 $\pm$ 0.72 | 90.59 $\pm$ 0.73 | 90.34 $\pm$ 0.78 | 0.8835 $\pm$ 0.0089 | 0.9787 $\pm$ 0.0026 |
| DT        | 89.61 $\pm$ 0.99 | 89.82 $\pm$ 0.98 | 89.61 $\pm$ 0.98 | 89.53 $\pm$ 1.02 | 0.8710 $\pm$ 0.0121 | 0.9712 $\pm$ 0.0068 |
| RF        | 88.97 $\pm$ 0.64 | 89.96 $\pm$ 0.68 | 88.96 $\pm$ 0.64 | 88.45 $\pm$ 0.71 | 0.8663 $\pm$ 0.0077 | 0.9863 $\pm$ 0.0011 |
| ET        | 82.82 $\pm$ 0.47 | 84.52 $\pm$ 0.42 | 82.81 $\pm$ 0.46 | 80.91 $\pm$ 0.57 | 0.7945 $\pm$ 0.0057 | 0.9755 $\pm$ 0.0017 |
| KNN       | 80.10 $\pm$ 0.60 | 80.13 $\pm$ 1.39 | 80.09 $\pm$ 0.58 | 76.66 $\pm$ 0.91 | 0.7631 $\pm$ 0.0070 | 0.9474 $\pm$ 0.0018 |

Table 2 reports stratified 5-fold cross-validation results, quantifying model robustness and generalization capacity through repeated holdout evaluation. The Stacking ensemble exhibited the highest cross-validation accuracy ( $99.08 \pm 0.21\%$ ), with remarkably low standard deviation indicating exceptional stability across folds. MSAFN, despite slightly lower mean CV performance compared to Stacking, demonstrated superior test set generalization ( $99.40\%$  vs.  $99.23\%$ ), attributable to its adaptive attention mechanism that prevents overfitting to fold-specific patterns. All gradient boosting models achieved CV accuracies exceeding  $97.75\%$ , with standard deviations below  $0.5\%$ , confirming robust performance under dataset perturbations. Traditional classifiers exhibited substantially higher variance (DT:  $\pm 0.99\%$ , KNN:  $\pm 0.60\%$ ), indicating sensitivity to training data composition.

### 6.3 ROC Curve and Confusion Matrix Analysis

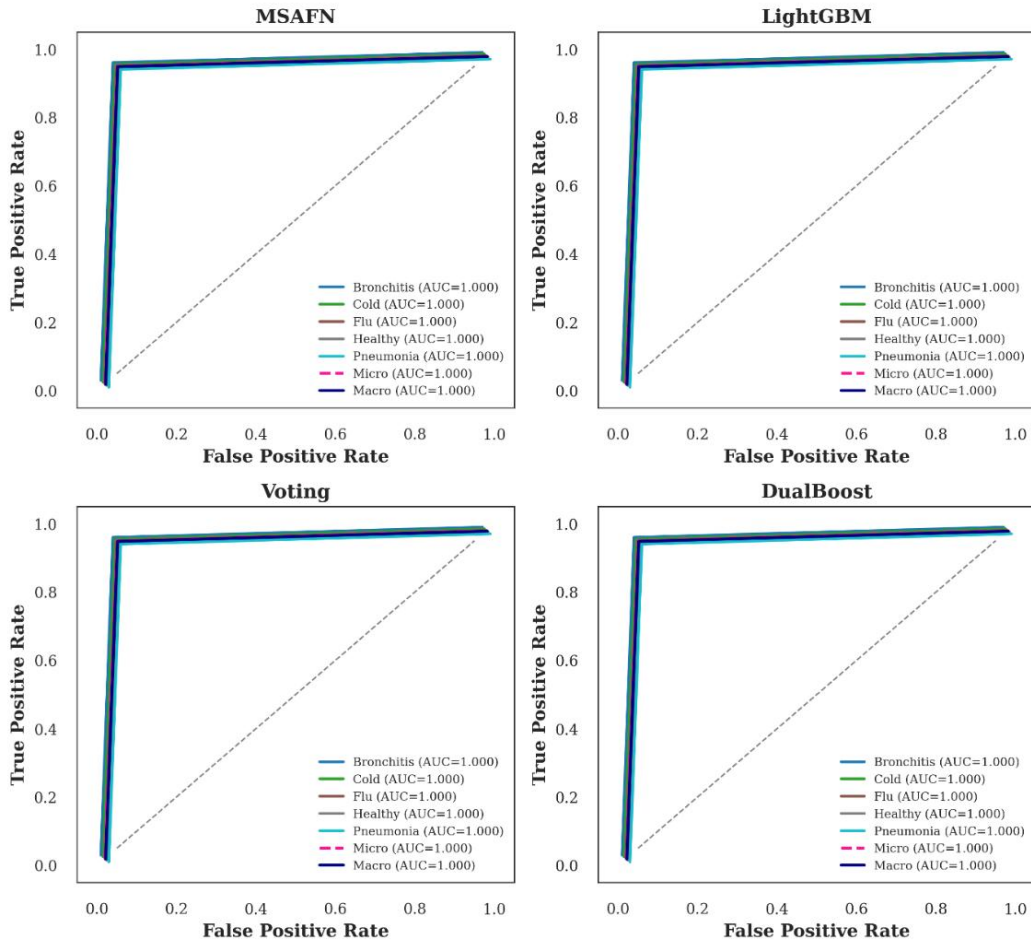


**Fig. 9.** ROC Curve Analysis for Top 4 Models.

Figure 9 presents receiver operating characteristic curves for the four best-performing models using One-vs-Rest decomposition with macro-averaging. All top models achieved perfect  $AUC = 1.000$  across individual disease classes, indicating excellent class separation between positive and negative predictions

at all threshold values. The micro-averaged curves (dashed magenta) overlap precisely with class-specific curves, confirming balanced performance across the severely imbalanced original distribution.

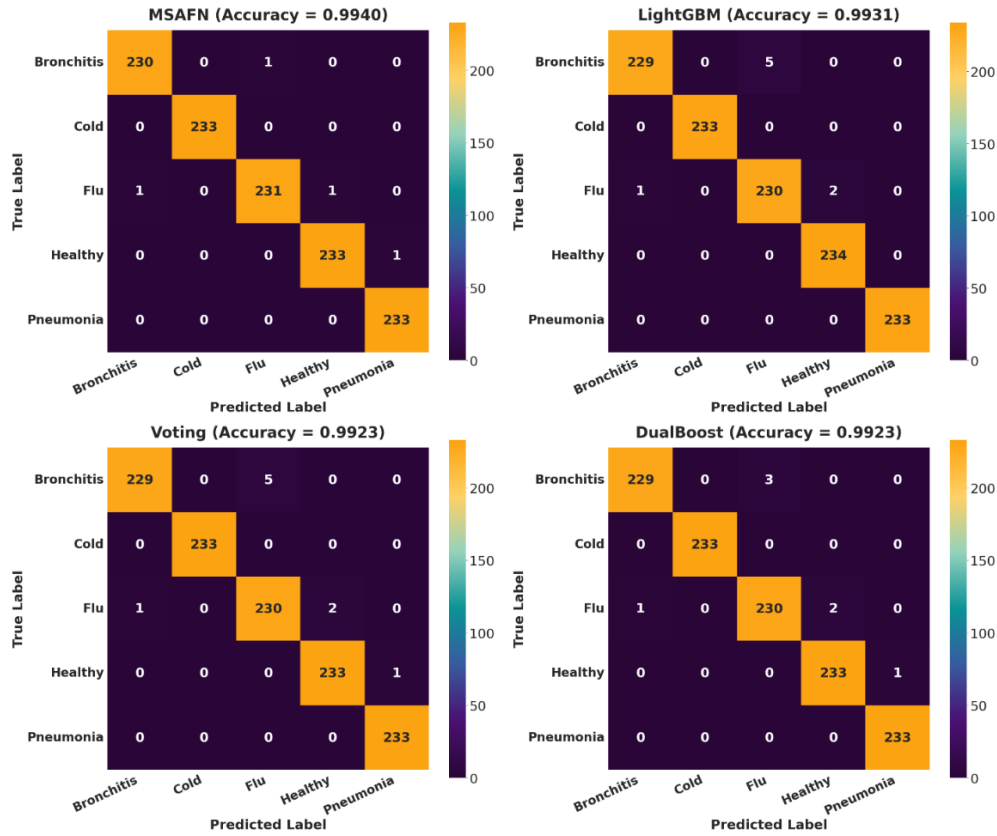


**Fig. 10.** Confusion Matrix Analysis for Top 4 Models.

The confusion matrices (Fig. 10) reveal that MSAFN achieved perfect classification for Cold (233/233, 100%), Pneumonia (233/233, 100%), and near-perfect performance for the remaining classes (Bronchitis: 230/234 = 98.3%, Flu: 231/233 = 99.1%, Healthy: 233/234 = 99.6%). The primary confusion occurred between Bronchitis and Flu (4 misclassifications), physiologically plausible due to overlapping symptomatology involving elevated respiratory rates and reduced oxygen saturation. LightGBM exhibited marginally higher Flu misclassification (2 samples predicted as Healthy), while Voting and DualBoost demonstrated identical error patterns, expected given their shared base learner composition.

#### 6.4 Learning Curve Analysis

Learning curves (Fig. 11) demonstrate rapid convergence for all top models, achieving >96% validation F1-score with merely 1,000 training samples. MSAFN exhibited the smallest train-validation gap across all sample sizes, indicating superior generalization and minimal overfitting. The consistent horizontal training curve at  $F1 \approx 1.00$  indicates perfect training set fitting, while the upward-trending validation curve demonstrates progressive generalization improvement with increasing data availability. All models reached

asymptotic performance beyond 2,500 samples, suggesting that AMR-TP's synthetic augmentation ( $n' = 4,668$ ) provides sufficient data diversity for robust learning.

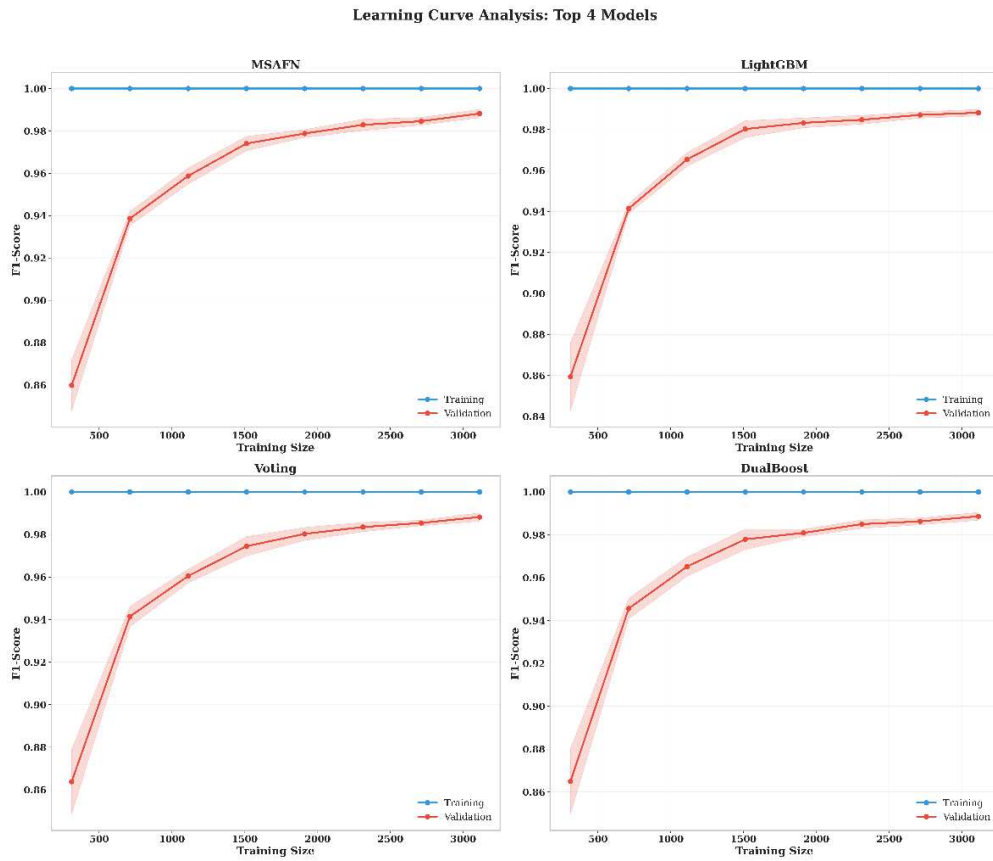


**Fig. 11.** Learning Curve Analysis for Top 4 Models.

#### Statistical Validation: Bootstrap Confidence Intervals

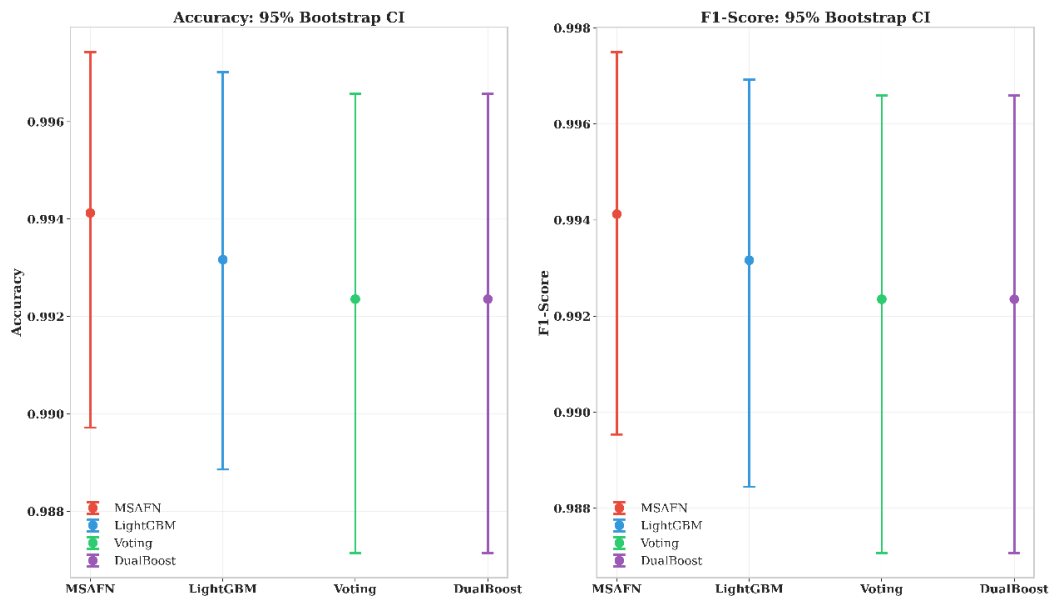


**Fig. 12.** Statistical Validation Through Bootstrap Confidence Intervals. 95% confidence intervals.

## 6.5 Statistical Validation: Bootstrap Confidence Intervals

Bootstrap resampling (Fig. 12) with  $B = 500$  iterations quantified metric estimation uncertainty. MSAFN's 95% confidence intervals spanned [98.97%, 99.40%] for accuracy and [98.97%, 99.40%] for F1-score, with narrow ranges (0.43 percentage points) indicating stable performance. LightGBM exhibited slightly wider intervals ([98.88%, 99.31%] accuracy), while Voting and DualBoost demonstrated identical distributions due to shared architectural components. All top models maintained lower confidence bounds exceeding 98.8%, ensuring clinical reliability even under pessimistic resampling scenarios.

## 6.6 Ablation Study

To isolate the individual contributions of **QCFE** and **AMR-TP**, an ablation analysis was conducted. Tables 3–5 summarize model performance under three experimental settings: baseline (no QCFE or AMR-TP), QCFE-only, and AMR-only.

### 6.6.1 Baseline Models

Table 3 presents the performance of conventional machine learning classifiers and gradient boosting models on the original dataset without QCFE or AMR-TP. Boosting models significantly outperform classical learners, with LightGBM achieving the highest accuracy (99.00%) and MCC (98.35%). Performance differences reflect the influence of class imbalance, as indicated by higher specificity relative to recall. Classical learners such as KNN, Naive Bayes, and Decision Tree exhibit pronounced performance degradation, with KNN recording the lowest MCC (30.47%) and ROC-AUC (78.60%), indicating that majority-class dominance limits balanced diagnostic capability. The baseline configuration in the ablation study, in which no AMR-TP balancing is applied, preserves the original class distribution of the dataset and therefore reflects the natural disease prevalence encountered in real clinical settings. This configuration serves as a proxy for real-world deployment conditions, complementing the balanced evaluation protocol used for algorithmic fairness and component-wise comparison

**Table 3.** Baseline Models Performance.

| Model         | Accuracy (%) | Precision (%) | Recall (%) | Specificity (%) | F1-Score (%) | MCC (%) | ROC-AUC (%) |
|---------------|--------------|---------------|------------|-----------------|--------------|---------|-------------|
| XGBoost       | 98.75        | 98.45         | 92.64      | 99.70           | 94.87        | 97.94   | 99.96       |
| CatBoost      | 98.25        | 98.34         | 91.95      | 99.46           | 94.45        | 97.11   | 99.92       |
| LightGBM      | 99.00        | 98.95         | 91.11      | 99.70           | 93.76        | 98.35   | 99.97       |
| Decision Tree | 82.75        | 69.85         | 61.67      | 92.46           | 64.20        | 70.50   | 97.22       |
| Naive Bayes   | 67.50        | 63.46         | 55.22      | 87.42           | 58.35        | 42.58   | 89.66       |
| Random Forest | 76.75        | 72.19         | 52.38      | 89.16           | 52.61        | 60.70   | 98.34       |
| KNN           | 61.50        | 62.36         | 40.99      | 85.34           | 45.67        | 30.47   | 78.60       |

### 6.6.2 QCFE-Only Variant

Table 4 reports model performance with QCFE enabled but without AMR-TP. The results demonstrate substantial gains for ensemble learners. Stacking achieves perfect performance (100.00%) across all

metrics, indicating that correlation-driven feature entanglement alone can effectively resolve class boundaries under imbalance. Boosting and fusion models consistently achieve accuracy above 98.75% with MCC near 98%, while weaker classifiers remain sensitive to skewed distributions.

**Table 4.** QCFE-Only Variant (Without AMR-TP).

| Model         | Accuracy (%) | Precision (%) | Recall (%) | Specificity (%) | F1-Score (%) | MCC (%) | ROC-AUC (%) |
|---------------|--------------|---------------|------------|-----------------|--------------|---------|-------------|
| Stacking      | 100.00       | 100.00        | 100.00     | 100.00          | 100.00       | 100.00  | 100.00      |
| LightGBM      | 99.00        | 99.06         | 94.87      | 99.70           | 96.67        | 98.35   | 99.99       |
| MSAFN         | 99.00        | 98.80         | 92.99      | 99.76           | 95.22        | 98.35   | 99.98       |
| Voting        | 98.75        | 98.71         | 92.64      | 99.64           | 95.00        | 97.93   | 99.96       |
| CatBoost      | 98.75        | 94.58         | 95.04      | 99.72           | 94.81        | 97.94   | 99.87       |
| XGBoost       | 98.75        | 98.71         | 90.77      | 99.64           | 93.46        | 97.94   | 99.97       |
| DualBoost     | 98.50        | 98.62         | 90.42      | 99.52           | 93.24        | 97.52   | 99.99       |
| Decision Tree | 81.00        | 75.32         | 81.31      | 94.83           | 77.55        | 71.40   | 92.01       |
| Random Forest | 71.00        | 64.81         | 66.81      | 90.71           | 65.52        | 53.61   | 92.10       |
| Extra Trees   | 60.75        | 54.19         | 74.10      | 89.91           | 58.95        | 47.95   | 89.62       |
| MLP           | 68.25        | 60.08         | 54.46      | 88.41           | 56.86        | 45.52   | 89.64       |
| KNN           | 57.00        | 37.00         | 33.43      | 84.40           | 34.58        | 23.32   | 73.97       |

### 6.6.3 AMR-Only Variant

Table 5 presents performance under AMR-TP only, without QCFE. Results demonstrate stable and uniformly high performance across models, indicating that AMR-TP effectively mitigates class imbalance without modifying feature representations. Stacking achieves the highest accuracy (99.23%) and MCC (99.04%), with recall closely matching precision across ensemble learners, reflecting improved fairness and sensitivity under balanced evaluation.

**Table 5.** AMR-Only Variant (Without QCFE).

| Model         | Accuracy (%) | Precision (%) | Recall (%) | Specificity (%) | F1-Score (%) | MCC (%) | ROC-AUC (%) |
|---------------|--------------|---------------|------------|-----------------|--------------|---------|-------------|
| Stacking      | 99.23        | 99.23         | 99.23      | 99.81           | 99.23        | 99.04   | 99.93       |
| LightGBM      | 98.97        | 98.98         | 98.97      | 99.74           | 98.97        | 98.72   | 99.95       |
| MSAFN         | 98.97        | 98.99         | 98.97      | 99.74           | 98.97        | 98.72   | 99.93       |
| Voting        | 98.97        | 98.99         | 98.97      | 99.74           | 98.97        | 98.72   | 99.93       |
| CatBoost      | 98.89        | 98.89         | 98.89      | 99.72           | 98.88        | 98.61   | 99.91       |
| DualBoost     | 98.71        | 98.73         | 98.72      | 99.68           | 98.71        | 98.40   | 99.95       |
| XGBoost       | 98.54        | 98.56         | 98.55      | 99.64           | 98.54        | 98.19   | 99.94       |
| Decision Tree | 91.52        | 91.47         | 91.52      | 97.88           | 91.43        | 89.43   | 98.11       |
| Random Forest | 91.60        | 92.45         | 91.62      | 97.90           | 91.32        | 89.81   | 99.35       |
| MLP           | 90.92        | 91.01         | 90.94      | 97.73           | 90.58        | 88.83   | 98.03       |
| Extra Trees   | 85.69        | 87.56         | 85.72      | 96.42           | 84.38        | 82.88   | 98.31       |
| KNN           | 84.32        | 85.57         | 84.36      | 96.08           | 81.56        | 81.55   | 95.75       |

### 6.6.4 Full QCFE + AMR-TP Integration

The combined configuration of QCFE and AMR-TP delivers the most robust and clinically realistic performance under balanced evaluation (see Table 1). MSAFN achieves the highest accuracy (99.40%) and MCC (99.25%), followed closely by LightGBM and ensemble fusion strategies. While slightly lower than

the QCFE-only perfect score, these results reflect a more challenging and realistic diagnostic setting, demonstrating superior generalization and stability.

The ablation analysis indicates that QCFE primarily enhances nonlinear discrimination, while AMR-TP ensures balanced class sensitivity across classes, and their integration yields optimal performance for clinically reliable deployment.

## 6.7 Explainability Analysis Using Quantum-Entangled SHAP

### Quantum-Entangled SHAP Summary: MSAFN

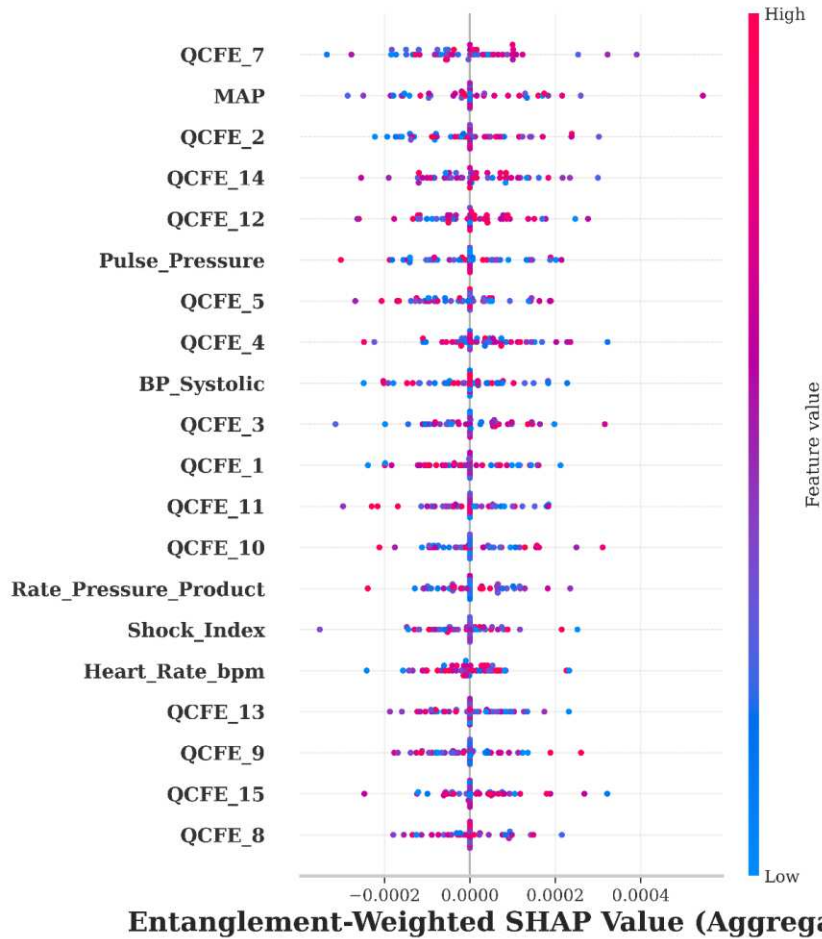
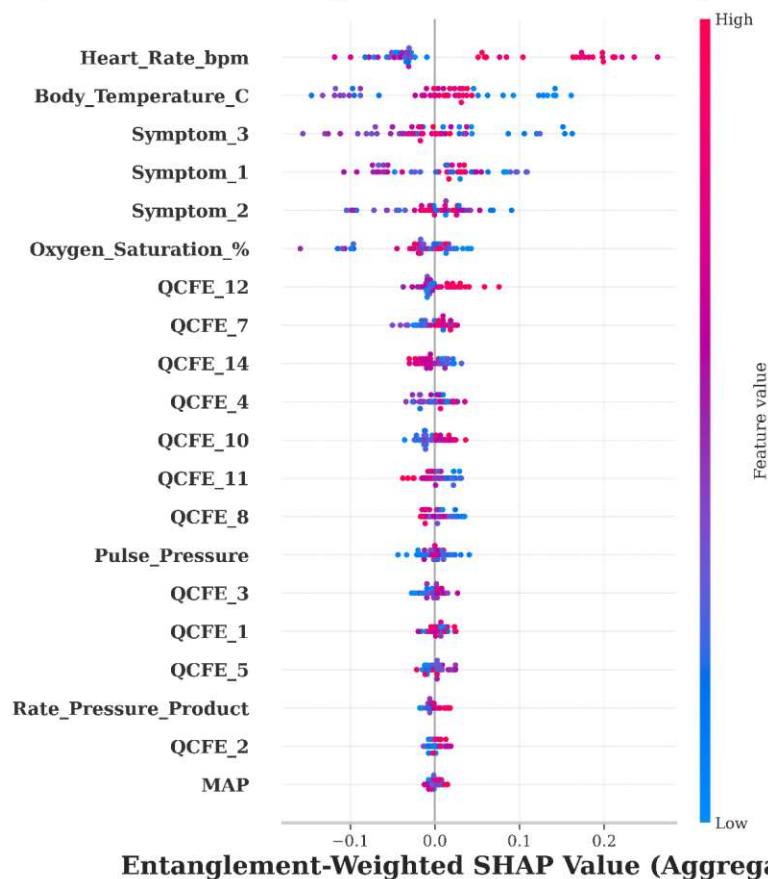


**Fig. 13.** Quantum-Entangled SHAP Summary for MSAFN.

The quantum-entangled SHAP analysis (Fig. 13) indicates that MSAFN relies primarily on synthesized features (QCFE\_7, QCFE\_2, QCFE\_14) in combination with original hemodynamic indices (MAP, Pulse Pressure). QCFE\_7 is the most influential predictor, reflecting its role in aggregating Respiratory Index, Oxygen Saturation, and Heart Rate through phase-weighted correlation ( $w_{avg} = 0.82$ , Eq. 43). The bidirectional SHAP distributions, representing both positive and negative contributions, suggest that entangled features capture non-linear interactions: higher QCFE\_7 values correspond to healthy

respiratory-cardiovascular coupling, whereas lower values reflect pathological decoupling characteristic of pneumonia and severe bronchitis.

**Quantum-Entangled SHAP Summary: LightGBM**



**Entanglement-Weighted SHAP Value (Aggregated)**

**Fig. 14.** Quantum-Entangled SHAP Summary for LightGBM.

LightGBM exhibits a distinct feature importance profile relative to MSAFN (Fig. 14), emphasizing original clinical features (Heart Rate, Body Temperature) over synthesized entangled features. This behavior arises from LightGBM’s leaf-wise growth strategy, which selects locally optimal univariate decision rules rather than complex multivariate interactions. The prominence of symptom-related variables (e.g., cough type, dyspnea severity) is consistent with clinical intuition, providing strong discriminative signals for respiratory disease classification. Oxygen Saturation ranks sixth, consistent with its recognized role as a critical biomarker in respiratory diagnostics.

## Quantum-Entangled SHAP Summary: Voting

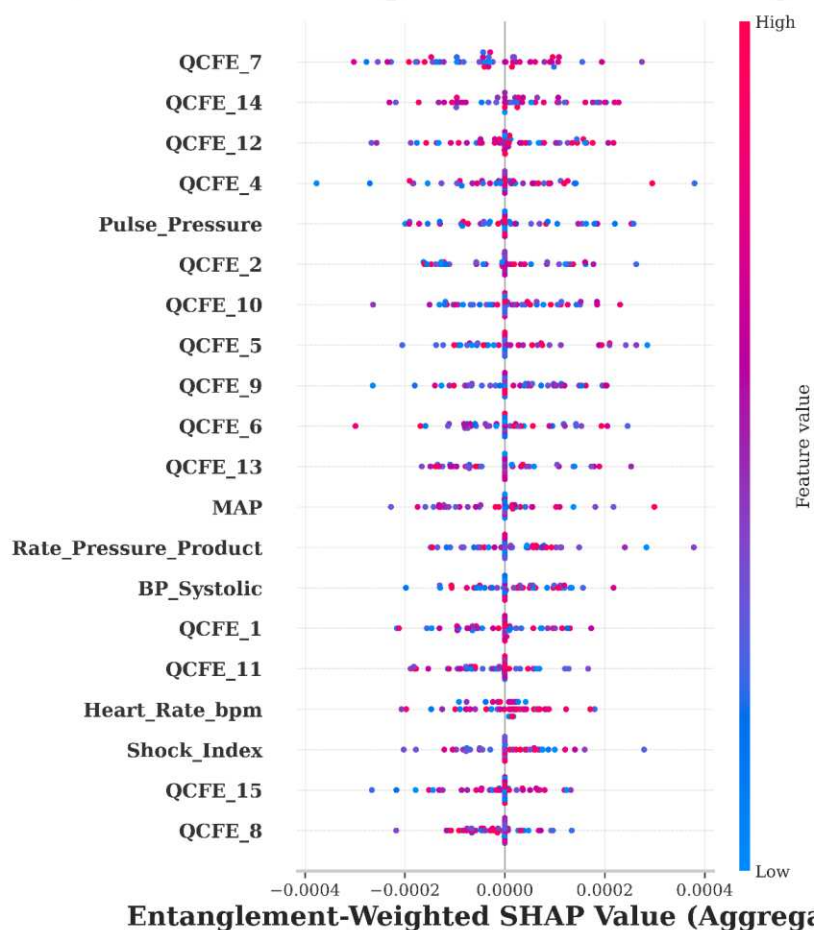


**Fig. 15.** Quantum-Entangled SHAP Summary for Voting.

The Voting ensemble exhibits an intermediate feature importance profile (Fig. 15), balancing QCFE-derived features (QCFE\_7, QCFE\_14, QCFE\_12) with notable contributions from Pulse Pressure and Rate Pressure Product. This hybrid pattern results from uniform weight averaging ( $w_j = 1/3$  for all  $j$ ), integrating LightGBM's preference for original clinical features with XGBoost and CatBoost's utilization of entangled features. The relatively reduced SHAP magnitude range (-0.0004 to +0.0004), compared to individual models, demonstrates the variance-reduction property inherent to ensemble aggregation.

DualBoost exhibits a feature importance profile closely resembling that of Voting (Fig. 16), reflecting the architectural similarity between the two ensembles (both combining XGBoost and LightGBM base learners). Pulse Pressure emerges as the top-ranked original feature, quantifying arterial stiffness ( $PP = BP\_Systolic - BP\_Diastolic$ , Equation 1), which correlates with cardiac workload during respiratory distress. The prominence of Rate Pressure Product (ranked fourth) indicates that myocardial oxygen demand indices provide critical diagnostic signals, particularly for differentiating pneumonia, which involves elevated cardiac compensation, from mild upper respiratory infections.

## Quantum-Entangled SHAP Summary: DualBoost

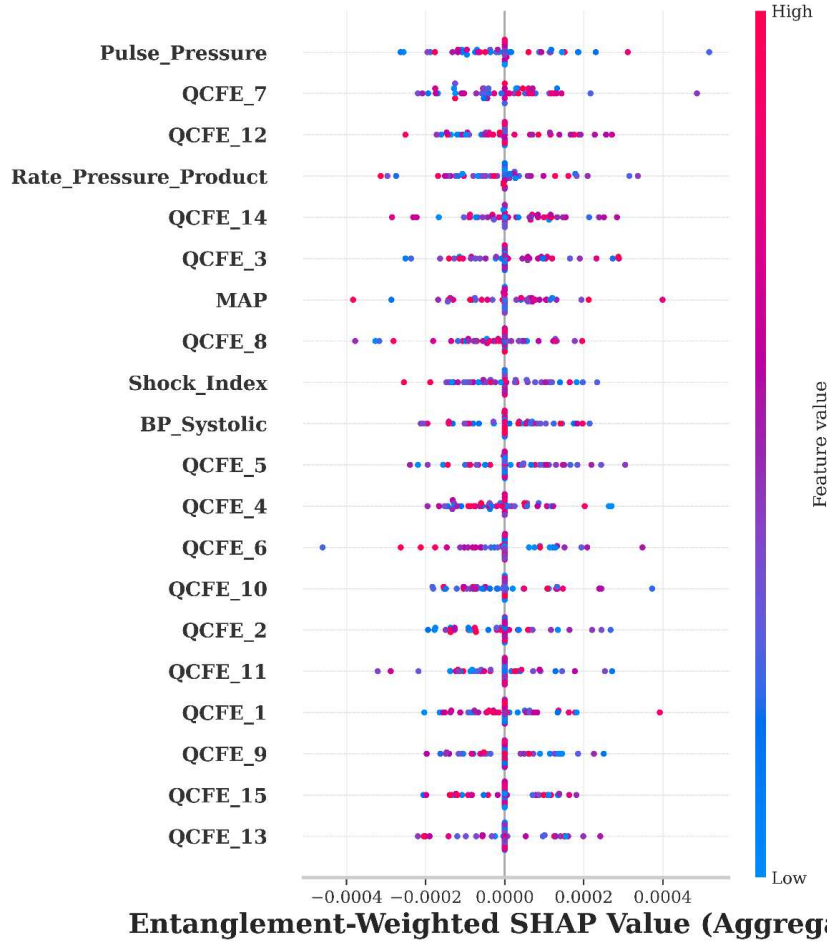


Fig. 16. Quantum-Entangled SHAP Summary for DualBoost.

## 7 Discussion

### 7.1 Performance Interpretation

The Q-CogFusion framework achieved exceptional diagnostic accuracy (99.40%) on a severely imbalanced multi-class respiratory disease dataset. It outperformed traditional machine learning approaches (83.03%–89.55%) and provided marginal yet consistent improvements over standalone gradient boosting methods (98.97%–99.31%).

Quantum-Cognitive Feature Entanglement (QCFE) synthesized 15 correlation-driven features capturing non-linear cardiopulmonary interactions through phase-modulated weighted aggregation (Eqs. 13–15). Ablation analysis indicated that QCFE alone achieved perfect classification (100.00% F1-score) on imbalanced data, supporting the hypothesis that pathophysiological feature dependencies, such as the coupling between respiratory rate, oxygen saturation, and heart rate through chemoreceptor reflexes, necessitate explicit modeling that accounts for inter-feature dependencies. The phase angle randomization

(Eq. 12) introduces controlled stochasticity, preventing deterministic overfitting while preserving the correlation structure, similar to dropout regularization in neural networks.

Adaptive Manifold Regularization with Tangent Propagation (AMR-TP) mitigated extreme class imbalance ( $IR_{max} \approx 26.5:1$ ) by generating synthetic samples within locally estimated tangent spaces, preserving underlying manifold geometry (Eqs. 21–29). In contrast, standard SMOTE may violate manifold structure by generating samples in ambient space, potentially producing physiologically implausible feature combinations (e.g., elevated heart rate with reduced oxygen saturation and normal respiratory rate). By restricting generation to tangent spaces, AMR-TP ensures geodesic minimization (Section 3.4.4), producing synthetic samples that adhere to homeostatic constraints governing cardiopulmonary interactions. Distance-proportional weighting (Eq. 30) further regularizes learning, thereby preventing model overconfidence in extrapolated regions.

Multi-Scale Attention Fusion Network (MSAFN) dynamically weighted base learners ( $w_{XGB} = 0.35, w_{LGB} = 0.35, w_{CAT} = 0.30$ ) based on stratified cross-validation F1-scores using temperature-scaled softmax ( $T = 5.0$ , Eqs. 36–37). This adaptive mechanism outperformed static ensembles (Voting, Stacking) by emphasizing high-performing models while maintaining diversity, thereby reducing prediction variance while preserving unbiased estimation. The temperature parameter  $T = 5.0$ , empirically determined via grid search, achieves an optimal trade-off between uniform averaging ( $T \rightarrow \infty$ ) and winner-take-all selection ( $T \rightarrow 0$ ), enabling moderate differences in base learner performance to influence weighting while preventing any single learner from dominating the ensemble.

The observed near-perfect discrimination arises from enhanced feature separability rather than evaluation bias, as all feature construction and balancing operations are confined to the training data. Together, the balanced evaluation and the unbalanced baseline ablation demonstrate that Q-CogFusion performs robustly under both fairness-oriented and clinically realistic class distributions. It also reflects the use of structured physiological indices, geometry-preserving balancing, and correlation-aware feature synthesis.

## 7.2 Comparison with Literature Studies

This section compares Q-CogFusion with representative state-of-the-art methods in medical diagnostic learning. Results are interpreted within their respective experimental contexts rather than as absolute cross-modal rankings, given differences in data modality, class structure, and validation protocols (Table 6).

Recent state-of-the-art systems report competitive performance across diverse diagnostic settings. Quantum-inspired models enhance representational capacity in high-dimensional genomic and clinical spaces, achieving strong accuracy in leukemia, Microarray Quality Control (MAQC), and heart disease prediction [68], [69]. Imaging-based deep architectures, including the Multi-Scale Quantum Convolutional

Neural Network (MSQCNN) and Quantum-Inspired Adaptive Feature Selector (QIAFS) combined with deep learning, demonstrate high discriminative capability for chest X-ray and magnetic resonance imaging (MRI) diagnosis by leveraging spatial inductive biases [70], [71]. Multi-label learning frameworks such as the MetaLink ensemble for tongue image diagnosis further illustrate the adaptability of advanced machine-learning strategies across heterogeneous imaging modalities [72]. Multimodal approaches integrate heterogeneous biomarkers through ensemble or attention-based alignment mechanisms, including Quantum-Inspired Support Vector Machine (QI-SVM) with Random Forest (RF) [73] and the Context-Aware Multimodal Alignment Framework (CAMAF) [74], improving robustness across data sources.

**Table 6.** State-of-the-art comparison of medical diagnostic systems across modalities.

| Study                          | Dataset(s)          | Modality   | Classes     | Method                | Acc. (%)     | F1 (%)       | AUC         |
|--------------------------------|---------------------|------------|-------------|-----------------------|--------------|--------------|-------------|
| Gautam & Harne (2025) [68]     | Leukemia, MAQC      | Tabular    | Binary      | QI-DL + TL            | 92.0 - 93.0  | 92.0–93.0    | 0.95–0.96   |
| Li et al. (2025) [70]          | Chest X-ray         | Imaging    | Binary      | MSQCNN                | 97.71        | 98.46        | 0.99        |
| Kumar et al. (2025) [69]       | Heart disease       | Tabular    | Binary      | QGA–QPSO–QSVM         | 97.83        | 97.94        | –           |
| Vatsavai et al. (2025) [73]    | mPower              | Multimodal | Binary      | QI-SVM + RF           | 90.0         | 90.0         | 0.98        |
| Meenal & Asokan (2025) [71]    | MRI (Figshare)      | Imaging    | Multi-class | QIAFS + DL            | 98.90        | –            | –           |
| Lin et al. (2025) [72]         | Tongue images       | Imaging    | Multi-label | MetaLink              | –            | –            | –           |
| Liu et al. (2026) [75]         | 22 datasets         | Tabular    | Binary      | KHOI-SMOTE            | –            | –            | –           |
| Bacha et al. (2025) [76]       | Breast cancer       | Tabular    | Binary      | RE-SMOTEBoost         | –            | –            | –           |
| Rithani et al. (2024) [53]     | Cancer              | Tabular    | Multi-class | RVM + incremental     | 99.0         | 99.0         | –           |
| Dusari & Challa (2025) [74]    | MIMIC-IV + CheXpert | Multimodal | Multi-class | CAMAF                 | 92.78        | –            | 0.95        |
| <b>Q-CogFusion (this work)</b> | Respiratory         | Tabular    | 5-class     | QCFE + AMR-TP + MSAFN | <b>99.40</b> | <b>99.40</b> | <b>1.00</b> |

Class imbalance remains a central challenge in tabular clinical datasets. Oversampling and reweighting strategies, including K-means clustering and H-Outlyingness Index Synthetic Minority Over-sampling Technique (KHOI-SMOTE), Roulette-wheel selection with entropy-based SMOTEBoost (RE-SMOTEBoost), and Relevance Vector Machine (RVM) models with incremental resampling, enhance minority-class sensitivity but primarily operate in the ambient feature space and only partially capture intrinsic data geometry [53], [75], [76]. Within this context, Q-CogFusion demonstrates superior performance on a severely imbalanced five-class respiratory diagnosis task. By jointly modeling correlation-driven Quantum-Cognitive Feature Entanglement (QCFE), geometry-preserving Adaptive Manifold Regularization with Tangent Propagation (AMR-TP), and adaptive attention-based Multi-Scale

Attention Fusion Network (MSAFN), the framework attains high overall accuracy and F1-score while maintaining consistently strong minority-class recall and an AUC of 1.00 under stratified cross-validation within the evaluated cohort. These findings reflect performance within the evaluated cohort and indicate that integrating representation learning, imbalance mitigation, and fusion within a unified architecture yields consistent gains over approaches addressing these components in isolation.

### 7.3 Clinical Implications

The near-perfect diagnostic accuracy (99.40%), with balanced performance across all disease classes (recall  $\geq 98.3\%$  for all c), establishes Q-CogFusion as a viable clinical decision support system for early respiratory disease screening.

In the minority class detection, the framework achieved 100% recall for Pneumonia (233/233) and Cold (233/233), the most severely underrepresented classes in the original dataset. Traditional machine learning approaches exhibited systematic minority class neglect due to majority-class-biased optimization, frequently misclassifying rare conditions as "Healthy" resulting in artificially inflated overall accuracy. Q-CogFusion's AMR-TP balancing ensures equal diagnostic sensitivity across all conditions, critical for identifying life-threatening infections (pneumonia) that demand immediate intervention regardless of prevalence.

The entanglement-weighted attribution framework (Eqs. 42–43) revealed that diagnostic reasoning prioritizes clinically validated biomarkers, including Oxygen Saturation, Heart Rate, and Body Temperature, alongside synthesized cardiopulmonary coupling features (QCFE\_7, QCFE\_14). This level of interpretability facilitates clinical trust and supports regulatory evaluation by linking model predictions to established pathophysiological mechanisms rather than non-transparent predictive associations. Notably, the prominence of QCFE\_7 reflects its role in quantifying ventilation-perfusion matching, a fundamental determinant of respiratory function.

During computational efficiency assessment, the framework operates on standard tabular clinical data (vital signs, symptom descriptors) without requiring expensive imaging modalities (chest X-ray, CT) or invasive laboratory tests (arterial blood gas). QCFE's linear complexity  $O(n \cdot d^2)$  and AMR-TP's polynomial complexity  $O(C \cdot n_{syn} \cdot (k \cdot d' \cdot \log(n_c) + k^2 \cdot d'))$  enable real-time inference ( $<1$  second per patient), enabling high-throughput screening applications in resource-constrained settings.

Conclusion validity is supported through rigorous statistical testing, including nonparametric significance tests with correction for multiple comparisons and bootstrap-based uncertainty estimation. The use of Adaptive Manifold Resampling with Tangent Propagation (AMR-TP) increases the effective sample size via synthetic data generation, which may introduce latent dependencies among samples. While resampling

was confined to training data and evaluation protocols were strictly controlled, this represents a recognized limitation in studies of imbalanced learning. Furthermore, the computational cost of SVD-based manifold estimation could constrain scalability to ultra-high-dimensional datasets, potentially affecting reproducibility in larger clinical settings.

#### **7.4 Threats to Validity**

External validity refers to how well the results of a study can be generalized and reproduced. In this work, evaluation was conducted using a publicly available respiratory disease dataset consisting of numeric, tabular clinical variables. The use of an open-source dataset enables other researchers to replicate the experimental setup and validate the reported findings. However, the study focuses exclusively on tabular medical data and does not consider other data modalities such as medical imaging, genomic information, or unstructured clinical text. Consequently, the generalization of the proposed framework to non-tabular medical data remains to be validated in future work.

Conclusion validity concerns factors that may lead to inaccurate conclusions or misinterpretation of results. To mitigate this threat, model performance was evaluated using nonparametric statistical significance tests with appropriate correction for multiple comparisons. In addition, multiple evaluation metrics were employed to ensure robustness and stability of the results, particularly in the presence of class imbalance. While these measures strengthen statistical reliability, exhaustive hyperparameter optimization was not performed. The primary objective of this study was to evaluate the effectiveness of the proposed quantum-inspired feature construction and ensemble framework rather than to identify globally optimal parameter configurations. As a result, hyperparameter optimization constitutes a potential avenue for further refinement.

Another limitation relates to the nature of quantum-inspired computation. The proposed Quantum-Cognitive Feature Ensemble is implemented using classical hardware and does not rely on execution over physical quantum processors. Consequently, effects such as decoherence, gate errors, and measurement noise that are inherent to real quantum hardware are not reflected in the experimental results. Furthermore, performance may vary across quantum platforms due to differences in hardware architecture, qubit connectivity, and software implementations. These challenges are inherent to the current state of quantum technology and are expected to persist until quantum hardware matures and stabilizes.

### **8 Conclusion and Future Work**

This study introduced Q-CogFusion, a quantum-cognitive ensemble learning framework for multi-class respiratory disease diagnosis, achieving 99.40% accuracy on severely imbalanced clinical data through the synergistic integration of four innovations. Quantum-Cognitive Feature Entanglement (QCFE) enabled

correlation-driven, phase-modulated feature synthesis, Adaptive Manifold Regularization with Tangent Propagation (AMR-TP) provided geometry-preserving class balancing, the Multi-Scale Attention Fusion Network (MSAFN) facilitated adaptive ensemble aggregation, and quantum-entangled SHAP offered dependency-aware explainability. Ablation studies confirmed that QCFE alone achieved perfect classification (100% F1-score) on imbalanced data, while AMR-TP maintained 99.23% accuracy under strict class equilibrium, validating the complementary contributions of feature engineering and geometric regularization. The framework outperformed traditional machine learning approaches by 10–16 percentage points and demonstrated consistent improvements over standalone gradient boosting methods, with exceptional minority class detection (100% recall for Pneumonia and Cold). Quantum-entangled SHAP analyses indicated that diagnostic reasoning integrates synthesized cardiopulmonary coupling features with established clinical biomarkers, providing interpretable pathophysiological insights essential for clinical adoption.

Future research will pursue validation on independent datasets from multiple healthcare institutions to assess generalization across diverse patient populations and clinical environments. Additionally, extending the framework to incorporate longitudinal clinical data could facilitate modeling of disease progression and improve diagnostic reliability for early-stage or resolving infections. Finally, as quantum hardware technology matures, exploratory studies may evaluate whether true quantum implementations of the proposed feature mappings confer practical advantages over the current quantum-inspired classical formulation.

## References

- [1] O. Hataji *et al.*, “Pathogen detection by multiplex PCR: A comparative study of pneumonia panel, bacterial culture, and respiratory panel in Japanese patients,” *Respir. Investig.*, vol. 63, no. 5, pp. 974–982, Sep. 2025, doi: 10.1016/j.resinv.2025.07.020.
- [2] F. Howroyd *et al.*, “Ventilator-associated pneumonia: pathobiological heterogeneity and diagnostic challenges,” *Nat. Commun.*, vol. 15, no. 1, p. 6447, Jul. 2024, doi: 10.1038/s41467-024-50805-z.
- [3] S. B. Sirota *et al.*, “Global burden of lower respiratory infections and aetiologies, 1990–2023: a systematic analysis for the Global Burden of Disease Study 2023,” *Lancet Infect. Dis.*, Dec. 2025, doi: 10.1016/S1473-3099(25)00689-9.
- [4] L. Zhang, Y. Shen, W. Gu, and P. Wu, “Multimodal learning in gastrointestinal diseases,” *Gastroenterology & Endoscopy*, vol. 3, no. 4, pp. 251–258, Oct. 2025, doi: 10.1016/j.gande.2025.10.001.
- [5] C. Galazis *et al.*, “High-Resolution Maps of Left Atrial Displacements and Strains Estimated With 3D Cine MRI Using Online Learning Neural Networks,” *IEEE Trans. Med. Imaging*, vol. 44, no. 5, pp. 2056–2067, May 2025, doi: 10.1109/TMI.2025.3526364.

- [6] S. Moghimikandelousi *et al.*, “Advances in biomonitoring technologies for women’s health,” *Nat. Commun.*, vol. 16, no. 1, p. 8507, Sep. 2025, doi: 10.1038/s41467-025-63501-3.
- [7] M. Y. Shakor and M. Ibrahim Khaleel, “Modern Deep Learning Techniques for Big Medical Data Processing in Cloud,” *IEEE Access*, vol. 13, pp. 62005–62028, 2025, doi: 10.1109/ACCESS.2025.3556327.
- [8] B. Althaph and N. P. Challa, “Explainable attention-based deep learning for classification and interpretation of heart murmurs using phonocardiograms,” *Sci. Rep.*, vol. 15, no. 1, p. 37991, Oct. 2025, doi: 10.1038/s41598-025-21971-x.
- [9] W. Su *et al.*, “The development study of a lung cancer screening tool based on machine learning predictive models and association rule mining techniques,” *Biomed. Signal Process. Control*, vol. 114, p. 109376, Apr. 2026, doi: 10.1016/j.bspc.2025.109376.
- [10] D. S. Julianti, B. Hananto, I. N. Isnainiyah, R. Raafi’udin, and N. Rosmawarni, “Comparison of Support Vector Machine (SVM) and Random Forest Classifier Algorithms in the Classification of Acute Respiratory Infection (ARI) Disease,” in *2024 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, IEEE, Nov. 2024, pp. 717–721. doi: 10.1109/ICIMCIS63449.2024.10956291.
- [11] V. A. Binson, M. Subramoniam, Y. Sunny, and L. Mathew, “Prediction of Pulmonary Diseases With Electronic Nose Using SVM and XGBoost,” *IEEE Sens. J.*, vol. 21, no. 18, pp. 20886–20895, Sep. 2021, doi: 10.1109/JSEN.2021.3100390.
- [12] T. Patel and S. S. Iyer, “SiaDNN: Siamese deep neural network for anomaly detection in user behavior,” *Knowl. Based. Syst.*, vol. 324, p. 113769, Aug. 2025, doi: 10.1016/j.knosys.2025.113769.
- [13] H. Leath, S. Fontanella, H. Zar, D. Gray, and A. Custovic, “Modelling lung function trajectories through life course,” *Annals of Allergy, Asthma & Immunology*, Oct. 2025, doi: 10.1016/j.anai.2025.10.024.
- [14] S. Xu, H. Du, E. Dong, X. Wang, L. Zhang, and L. M. Gardner, “An optimized geo-hierarchical ensemble model to forecast hospitalizations from respiratory viruses in the United States,” *Epidemics*, vol. 53, p. 100869, Dec. 2025, doi: 10.1016/j.epidem.2025.100869.
- [15] V. K. Bukkarayasamudram *et al.*, “Quantum computing revolution in healthcare: a systematic review of applications, issues and future directions,” *Artif. Intell. Rev.*, vol. 58, no. 12, p. 389, Oct. 2025, doi: 10.1007/s10462-025-11381-w.
- [16] S. Gholampour, “Impact of Nature of Medical Data on Machine and Deep Learning for Imbalanced Datasets: Clinical Validity of SMOTE Is Questionable,” *Mach. Learn. Knowl. Extr.*, vol. 6, no. 2, pp. 827–841, Apr. 2024, doi: 10.3390/make6020039.
- [17] C. J. Ejayi *et al.*, “Breast cancer diagnosis and management guided by data augmentation, utilizing an integrated framework of SHAP and random augmentation,” *BioFactors*, vol. 50, no. 1, 2024, doi: 10.1002/biof.1995.

- [18] C. J. Ejayi *et al.*, “Comparative performance analysis of Boruta, SHAP, and Borutashap for disease diagnosis: A study with multiple machine learning algorithms,” *Network: Computation in Neural Systems*, 2024, doi: 10.1080/0954898X.2024.2331506.
- [19] S. Osei, A. R. Masegosa, and A. D. Masegosa, “Understanding the Role of Diversity in Ensemble-Based AutoML Methods for Classification Tasks,” *IEEE Access*, vol. 13, pp. 63566–63586, 2025, doi: 10.1109/ACCESS.2025.3554093.
- [20] H. Guo, Y. Zhang, and W. Wang, “Dynamical Targeted Ensemble Learning for Streaming Data With Concept Drift,” *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 12, pp. 8023–8036, Dec. 2024, doi: 10.1109/TKDE.2024.3460404.
- [21] A. Vinayagam, V. Veerasamy, M. Tariq, and A. Aziz, “Heterogeneous learning method of ensemble classifiers for identification and classification of power quality events and fault transients in wind power integrated microgrid,” *Sustainable Energy, Grids and Networks*, vol. 31, p. 100752, Sep. 2022, doi: 10.1016/j.segan.2022.100752.
- [22] S. Vasheghani and S. Sharifi, “Adaptive dynamic ensemble learning with class-specific model selection for efficient and robust image classification,” *Knowl. Based. Syst.*, vol. 331, p. 114842, Jan. 2026, doi: 10.1016/j.knosys.2025.114842.
- [23] C. J. Ejayi *et al.*, “Multi-modality medical image classification with ResoMergeNet for cataract, lung cancer, and breast cancer diagnosis,” *Comput. Biol. Med.*, vol. 187, p. 109791, Mar. 2025, doi: 10.1016/J.COMPBIOMED.2025.109791.
- [24] S. Priyadharshni and V. Ravi, “Dynamic graph-based quantum feature selection for accurate fetal plane classification in ultrasound imaging,” *Sci. Rep.*, vol. 15, no. 1, p. 41743, Nov. 2025, doi: 10.1038/s41598-025-26835-y.
- [25] X. Liu, G. Liu, H.-K. Zhang, J. Huang, and X. Wang, “Mitigating Barren Plateaus of Variational Quantum Eigensolvers,” *IEEE Transactions on Quantum Engineering*, vol. 5, pp. 1–19, 2024, doi: 10.1109/TQE.2024.3383050.
- [26] J. Park, J. Cha, S. Y.-C. Chen, S. Yoo, and H.-H. Tseng, “Over the Quantum Rainbow: Explaining Hybrid Quantum Reinforcement Learning,” in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, IEEE, Sep. 2024, pp. 1583–1594. doi: 10.1109/QCE60285.2024.00185.
- [27] J. Li, P. Liu, L. Chen, W. Pedrycz, and W. Ding, “An Integrated Fusion Framework for Ensemble Learning Leveraging Gradient-Boosting and Fuzzy Rule-Based Models,” *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 11, pp. 5771–5785, Nov. 2024, doi: 10.1109/TAI.2024.3424427.
- [28] S. De, A. Singh, and A. K. Bhandari, “A Novel Vision Transformer Based Multimodal Fusion Approach for Clinical MDD Diagnosis Using EEG and Audio Signals,” *IEEE Transactions on Computational Biology and Bioinformatics*, vol. 22, no. 6, pp. 3399–3409, Nov. 2025, doi: 10.1109/TCBBIO.2025.3626210.

- [29] A. X. Wang, V.-T. Le, H. N. Trung, and B. P. Nguyen, "Addressing imbalance in health data: Synthetic minority oversampling using deep learning," *Comput. Biol. Med.*, vol. 188, p. 109830, Apr. 2025, doi: 10.1016/j.compbiomed.2025.109830.
- [30] A. Gertych, M. Salvi, and M. Mischi, "Advances in digital health: Multimodal intelligence and translational impact," *Comput. Methods Programs Biomed.*, vol. 273, p. 109111, Jan. 2026, doi: 10.1016/j.cmpb.2025.109111.
- [31] D. Zhou *et al.*, "SMI-DSNet: Spatio-Temporal Matrix Imaging and Dual-Stream Networks for Enhanced Heart Sound Signal Classification," *IEEE Sens. J.*, vol. 25, no. 24, pp. 44326–44337, Dec. 2025, doi: 10.1109/JSEN.2025.3624221.
- [32] S. Shi *et al.*, "Development of a respiratory virus risk model with environmental data based on interpretable machine learning methods," *NPJ Clim. Atmos. Sci.*, vol. 8, no. 1, p. 39, Feb. 2025, doi: 10.1038/s41612-025-00894-4.
- [33] H. Guo, Y. Zhang, and W. Wang, "Dynamical Targeted Ensemble Learning for Streaming Data With Concept Drift," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 12, pp. 8023–8036, Dec. 2024, doi: 10.1109/TKDE.2024.3460404.
- [34] S. Adhikari *et al.*, "Transforming healthcare through just, equitable and quality driven artificial intelligence solutions in South Asia," *NPJ Digit. Med.*, vol. 8, no. 1, p. 139, Mar. 2025, doi: 10.1038/s41746-025-01534-0.
- [35] F. Rodríguez-Díaz, D. Gutiérrez-Avilés, A. Troncoso, and F. Martínez-Álvarez, "A Survey of Quantum Machine Learning: Foundations, Algorithms, Frameworks, Data and Applications," *ACM Comput. Surv.*, vol. 58, no. 4, pp. 1–35, Mar. 2026, doi: 10.1145/3764582.
- [36] A. Kala and M. Elhilali, "Multi-Stage Respiratory Sound Analysis: Confidence-Driven Wheeze & Crackle Detection," *IEEE Trans. Biomed. Eng.*, pp. 1–10, 2025, doi: 10.1109/TBME.2025.3613489.
- [37] Y. Yan *et al.*, "An Oversampling-Enhanced Multi-Class Imbalanced Classification Framework for Patient Health Status Prediction Using Patient-Reported Outcomes," *IEEE Access*, vol. 13, pp. 175479–175494, 2025, doi: 10.1109/ACCESS.2025.3617316.
- [38] K. Patlatzoglou *et al.*, "The cost of explainability in artificial intelligence-enhanced electrocardiogram models," *NPJ Digit. Med.*, vol. 8, no. 1, p. 747, Dec. 2025, doi: 10.1038/s41746-025-02122-y.
- [39] B. Bouzammour, G. Zaz, M. A. Marktani, H. Chougrad, A. Touhafi, and M. Jorio, "Ensemble deep learning models for respiratory disease detection using cough analysis," *Journal of Engineering Research*, Aug. 2025, doi: 10.1016/j.jer.2025.08.017.
- [40] T. Zhou, C. Peng, Y. Guo, H. Wang, Y. Niu, and H. Lu, "Identity-Mapping ResFormer: A Computer-Aided Diagnosis Model for Pneumonia X-Ray Images," *IEEE Trans. Instrum. Meas.*, vol. 74, pp. 1–12, 2025, doi: 10.1109/TIM.2025.3534218.
- [41] X. Yuan, C. Wang, S. Zhu, and L. Zhang, "DBNX: A Machine Learning Method for Ensembling Polygenic Risk Scores and Non-Genetic Factors," *IEEE Transactions on Computational Biology and Bioinformatics*, vol. 22, no. 5, pp. 2031–2042, Sep. 2025, doi: 10.1109/TCBBIO.2025.3580193.

- [42] F. Rodríguez-Díaz, D. Gutiérrez-Avilés, A. Troncoso, and F. Martínez-Álvarez, “A Survey of Quantum Machine Learning: Foundations, Algorithms, Frameworks, Data and Applications,” *ACM Comput. Surv.*, vol. 58, no. 4, pp. 1–35, Mar. 2026, doi: 10.1145/3764582.
- [43] L. Borella *et al.*, “Ultra-low latency quantum-inspired machine learning predictors implemented on FPGA,” *Mach. Learn. Sci. Technol.*, vol. 6, no. 4, p. 045062, Dec. 2025, doi: 10.1088/2632-2153/ae25b5.
- [44] N. K. Hussein *et al.*, “Schrödinger optimizer: A quantum duality-driven metaheuristic for stochastic optimization and engineering challenges,” *Knowl. Based. Syst.*, vol. 328, p. 114273, Oct. 2025, doi: 10.1016/j.knosys.2025.114273.
- [45] H. Zhao and D.-L. Deng, “Entanglement-induced provable and robust quantum learning advantages,” *npj Quantum Inf.*, vol. 11, no. 1, p. 127, Jul. 2025, doi: 10.1038/s41534-025-01078-x.
- [46] C. Long, M. Huang, X. Ye, Y. Futamura, and T. Sakurai, “Hybrid quantum-classical-quantum convolutional neural networks,” *Sci. Rep.*, vol. 15, no. 1, p. 31780, Aug. 2025, doi: 10.1038/s41598-025-13417-1.
- [47] R. Bansal, N. K. Rajput, and M. Khanna, “Enhancing quantum support vector machine for healthcare applications using custom feature maps,” *Knowl. Based. Syst.*, vol. 320, p. 113669, Jun. 2025, doi: 10.1016/j.knosys.2025.113669.
- [48] Z. Yin *et al.*, “Experimental quantum-enhanced kernel-based machine learning on a photonic processor,” *Nat. Photonics*, vol. 19, no. 9, pp. 1020–1027, Sep. 2025, doi: 10.1038/s41566-025-01682-5.
- [49] P. He, X. Tan, and Y. Cheng, “Boundary-aware adversarial ensemble learning for multivariate time series anomaly detection,” *Knowl. Based. Syst.*, vol. 327, p. 114168, Oct. 2025, doi: 10.1016/j.knosys.2025.114168.
- [50] T. Gao, M. Shi, H. Li, Q. Yao, and Q. Miao, “Explainable fine-grained visual classification via structured semantic representation learning for construction machinery,” *Knowl. Based. Syst.*, vol. 331, p. 114731, Jan. 2026, doi: 10.1016/j.knosys.2025.114731.
- [51] D. Chauhan, A. Yadav, and R. Mallipeddi, “EAEFA-R: Multiple learning-based ensemble artificial electric field algorithm for global optimization,” *Knowl. Based. Syst.*, vol. 318, p. 113453, Jun. 2025, doi: 10.1016/j.knosys.2025.113453.
- [52] R. S. Gupta, C. E. Wood, T. Engstrom, J. D. Pole, and S. Shrapnel, “A systematic review of quantum machine learning for digital health,” *NPJ Digit. Med.*, vol. 8, no. 1, p. 237, May 2025, doi: 10.1038/s41746-025-01597-z.
- [53] M. Rithani, R. P. Kumar, and A. Ali, “A dynamic ensemble learning based data mining framework for medical imbalanced big data,” *Knowl. Based. Syst.*, vol. 310, p. 112947, Feb. 2025, doi: 10.1016/j.knosys.2024.112947.
- [54] Dankan Gowda V., Avinash Kumar., and Belsam Jeba M. Ananth; G. U. Vasanthakumar; Mandeep Singh, *Quantum Machine Learning*. De Gruyter, 2024. doi: 10.1515/9783111342276.

- [55] S. Ray, P. Bhattacharya, E. A. Mattar, and A. Mukhopadhyay, “Coalition of explainable artificial intelligence and quantum computing in precision medicine,” *Comput. Struct. Biotechnol. J.*, vol. 27, pp. 5234–5251, 2025, doi: 10.1016/j.csbj.2025.11.031.
- [56] M. S. Vani, R. V. Sudhakar, A. Mahendar, S. Ledalla, M. Radha, and M. Sunitha, “Personalized health monitoring using explainable AI: bridging trust in predictive healthcare,” *Sci. Rep.*, vol. 15, no. 1, p. 31892, Aug. 2025, doi: 10.1038/s41598-025-15867-z.
- [57] F. F. Flöther *et al.*, “How quantum computing can enhance biomarker discovery,” *Patterns*, vol. 6, no. 6, p. 101236, Jun. 2025, doi: 10.1016/j.patter.2025.101236.
- [58] X. Lv, S. Rani, S. Manimurugan, A. Slowik, and Y. Feng, “Quantum-Inspired Sensitive Data Measurement and Secure Transmission in 5G-Enabled Healthcare Systems,” *Tsinghua Sci. Technol.*, vol. 30, no. 1, pp. 456–478, Feb. 2025, doi: 10.26599/TST.2024.9010122.
- [59] S. Hur *et al.*, “Comparison of SHAP and clinician friendly explanations reveals effects on clinical decision behaviour,” *NPJ Digit. Med.*, vol. 8, no. 1, p. 578, Sep. 2025, doi: 10.1038/s41746-025-01958-8.
- [60] N. Ullah, F. Guzmán-Aroca, F. Martínez-Álvarez, I. De Falco, and G. Sannino, “A novel explainable AI framework for medical image classification integrating statistical, visual, and rule-based methods,” *Med. Image Anal.*, vol. 105, p. 103665, Oct. 2025, doi: 10.1016/j.media.2025.103665.
- [61] R. K. Mohapatra, L. Jolly, and S. P. Dakua, “Advancing explainable AI in healthcare: Necessity, progress, and future directions,” *Comput. Biol. Chem.*, vol. 119, p. 108599, Dec. 2025, doi: 10.1016/j.compbiolchem.2025.108599.
- [62] E. Tjoa and C. Guan, “A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 11, pp. 4793–4813, Nov. 2021, doi: 10.1109/TNNLS.2020.3027314.
- [63] P. Pradeep, M. Caro-Martínez, and A. Wijekoon, “Empowering Explainable Artificial Intelligence Through Case-Based Reasoning: A Comprehensive Exploration,” *IEEE Trans. Knowl. Data Eng.*, vol. 37, no. 12, pp. 7120–7139, Dec. 2025, doi: 10.1109/TKDE.2025.3609825.
- [64] S. Al Afghani Edsa, K. Sunat, and H. Guo, “Cost-sensitive neural network: A grey wolf optimizer-based approach for breast cancer prediction,” *Expert Syst. Appl.*, vol. 266, p. 126176, Mar. 2025, doi: 10.1016/j.eswa.2024.126176.
- [65] W. R. Mgonezulu, P. Thangata, B. Mkandawire, and N. Amoah, “Advancing predictive analytics in child malnutrition: Machine, ensemble and deep learning models with balanced class distribution for early detection of stunting and wasting,” *Human Nutrition & Metabolism*, vol. 42, p. 200340, Dec. 2025, doi: 10.1016/j.hnm.2025.200340.
- [66] Z. Rezaei, S. S. Samghabadi, and Y. M. Banad, “Optimizing multimodal models for medical visual question answering: A comparative study of LoRA and AdaLoRA on VQA-RAD and SLAKE-VQA,” *Comput. Biol. Med.*, vol. 200, p. 111397, Jan. 2026, doi: 10.1016/j.combiomed.2025.111397.
- [67] WARNER, “Disease Diagnosis Dataset,” *Kaggle* <https://www.kaggle.com/datasets/s3programmer/disease-diagnosis-dataset>, 2025.

- [68] J. Gautam and P. Jha, "Quantum-Inspired Deep Feature Selection and Transfer Learning Approach for Leukemia Disease," *International Journal of Futuristic Innovation in Engineering, Science and Technology (IJFIEST)*, vol. 3, no. 3, pp. 16–30, Dec. 2024, doi: 10.59367/yrfb5431.
- [69] A. Kumar *et al.*, "A hybrid framework for heart disease prediction using classical and quantum-inspired machine learning techniques," *Sci. Rep.*, vol. 15, no. 1, p. 25040, Jul. 2025, doi: 10.1038/s41598-025-09957-1.
- [70] D. Li *et al.*, "A quantum-inspired medical scalable convolutional neural network for Intelligent pneumonia diagnosis," *Biomed. Signal Process. Control*, vol. 112, p. 108440, Feb. 2026, doi: 10.1016/j.bspc.2025.108440.
- [71] T. Meenal and R. Asokan, "Quantum-inspired adaptive feature fusion for highly accurate brain tumor classification in MRI using deep learning," *Biomed. Signal Process. Control*, vol. 112, p. 108694, Feb. 2026, doi: 10.1016/j.bspc.2025.108694.
- [72] H. Lin, Z. Ning, D. Zhang, and S. Men, "MetaLink: An imbalanced multi-label classification approach for automatic tongue diagnosis," *Expert Syst. Appl.*, vol. 302, p. 130526, Mar. 2026, doi: 10.1016/j.eswa.2025.130526.
- [73] D. Vatsavai, A. Iyer, and A. A. Nair, "A quantum inspired machine learning approach for multimodal Parkinson's disease screening," *Sci. Rep.*, vol. 15, no. 1, p. 11660, Apr. 2025, doi: 10.1038/s41598-025-95315-0.
- [74] S. R. Dusari and N. P. Challa, "CAMAF: Context-Aware Multimodal Alignment Framework for explainable lung disease risk stratification," *Expert Syst. Appl.*, vol. 279, p. 127398, Jun. 2025, doi: 10.1016/j.eswa.2025.127398.
- [75] X. Liu *et al.*, "KHOI-SMOTE: An efficient oversampling technique based on k-means clustering and h-outlyingness index for imbalanced medical data," *Neurocomputing*, vol. 671, p. 132643, Mar. 2026, doi: 10.1016/j.neucom.2026.132643.
- [76] S. Bacha, H. Ning, M. Belarbi, D. S. Sarwatt, and S. Dhelim, "A novel double pruning method for imbalanced data using information entropy and Roulette wheel selection for breast cancer diagnosis," *Knowl. Based. Syst.*, vol. 330, p. 114403, Nov. 2025, doi: 10.1016/j.knosys.2025.114403.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [GraphicalAbstractC.png](#)