HybridNet-SEVIT: Multilabel Classification for Epidemic Risk Management and Public Health Surveillance

Yanyan Jiao
Disease Prevention and Control Center of Hongdong County, Linfen City, Shanxi Province, 041600, China
E-mail: 13633430031@163.com

Keywords: Public health surveillance, epidemic risk management, multilabel classification, HybridNet-SEVIT, deep learning, advanced preprocessing, health informatics

Received: December 20, 2024

Effective public health monitoring and epidemic prevention are crucial in mitigating the impact of infectious diseases. The dynamic nature of disease transmission necessitates a comprehensive understanding of demographic, health, environmental, and behavioral data. Despite advances in data collection, conventional methods often fail to accurately model disease dynamics, leading to suboptimal predictive capabilities. To address these challenges, we integrate a dataset of 43,689 entries from multiple sources and employ innovative preprocessing techniques such as Adaptive Distribution Recalibration, Contextual Outlier Filtering, and Multiscale Variance Modulation to enhance prediction accuracy and dataset integrity. Our proposed HybridNet-SEVIT model, which incorporates depthwise separable convolutions, dense connectivity, and ghost feature generation, achieves superior classification performance, with an accuracy of 97.9% and AUC values ranging from 96% to 99%. Comparative analysis demonstrates that HybridNet-SEVIT outperforms SOTA models, including ResNet (90.0% accuracy), DenseNet (84.1%), and SVM (84.8%), showing a significant 7.9% improvement in accuracy over the best-performing baseline model. Additionally, novel evaluation metrics—Adaptive Variability Index for Classes (AVIC), Stability of Prediction Dynamics Measure (SPDM), and Confidence Level Weighted Score (CLWS)—offer deeper insights into model robustness and predictive confidence. This study contributes to enhancing risk detection and classification in public health, advocating for a robust, data-driven approach to epidemic outbreak management and emphasizing the need for targeted interventions and efficient resource allocation.

Povzetek: Predstavljen je hibridni model za večoznačno klasifikacijo epidemijskega tveganja, ki združuje napredno predobdelavo in globoke nevronske mreže za podporo javnozdravstvenemu spremljanju in ukrepanju.

1 Introduction

Artificial intelligence (AI) has recently transformed various fields, including public health, where it plays a crucial role in epidemic management, especially in light of global health crises like the COVID-19 pandemic [1]. AI-driven approaches, coupled with big data analytics and advanced computing, are reshaping how public health professionals monitor, respond to, and manage infectious diseases. This study highlights the significant role of AI in strengthening public health interventions and providing timely, data-driven responses to epidemic threats.

Effective epidemic management relies on comprehensive public health monitoring, analyzing diverse data sources to detect anomalies and identify disease outbreaks early. Traditional surveillance systems often struggle with both data source integration and robust analysis. Data sourcing challenges stem from identifying and operationalizing health data signals, while analytical difficulties arise in developing reliable computational frameworks for signal extraction and interpretation [2]. Recent advancements in AI have shown promise in integrating structured health records with unstructured sources like social media, providing a richer, real-time understanding of public health trends. For example, AI-based applications have been instrumental in anticipating disease transmission and detecting hotspots using unconventional sources such as smartphone movement patterns and online activity [3]. These novel data streams empower public health authorities to make well-informed decisions to combat disease spread.

AI also enables predictive modeling frameworks that can capture the complex, nonlinear dynamics of infectious disease transmission [4, 5, 6, 7]. Traditional statistical models, which rely on linear assumptions and aggregated data, often fall short in predicting disease propagation accurately. Machine learning models, however, can encapsulate intricate interactions within a population, providing insights into transmission patterns and intervention outcomes [8, 9]. Beyond predictive capabilities, AI applications extend to safeguarding health in public spaces; for instance, during the COVID-19 pandemic, AI solutions were deployed to monitor indoor air quality in high-footfall areas such as museums and historical sites [10, 11, 12, 13]. Using Internet of Things (IoT) sensors and machine learning, these systems enabled real-time surveillance for crowd management and disease risk mitigation [14]. Nevertheless, the growing use of AI in public health raises ethical considerations, particularly concerning data privacy and public trust. Health data collection and analysis must adhere to strict ethical guidelines to ensure privacy protection while enabling effective public health interventions [15]. As AI applications continue to expand, establishing rigorous ethical frameworks will be essential to maximize their benefits without compromising individual rights.

The main contributions of this article are:

1. **High-Accuracy Classification**: Achieving a classification accuracy of 97.9%, our HybridNet-SEVIT model establishes a new benchmark in public health predictive modeling. Incorporating depthwise separable convolutions, dense connectivity, and ghost feature generation, this model addresses unique challenges in epidemic-related classification tasks.

2. **Innovative Preprocessing Techniques**: This study pioneers state-of-the-art preprocessing methods, such as Adaptive Distribution Recalibration and Contextual Outlier Filtering with Dynamic Threshold Adjustment, which enhance data quality and prediction accuracy by addressing imbalances, noise, and other data inconsistencies.

3. **Advanced Evaluation Metrics**: The introduction of novel metrics—Adaptive Variability Index for Classes (AVIC), Stability of Prediction Dynamics Measure (SPDM), and Confidence Level Weighted Score (CLWS)—enables a more nuanced and comprehensive assessment of model performance, beyond traditional accuracy and precision metrics.

4. **Comparative Performance Analysis**: A thorough evaluation demonstrates that the HybridNet-SEVIT model significantly outperforms conventional machine learning models in terms of recall, precision, and other critical metrics, underscoring its efficacy in practical public health contexts.

5. **Robust Sensitivity Testing**: Comprehensive sensitivity analysis of essential hyperparameters demonstrates the model's robustness and adaptability in dynamic settings, ensuring reliable performance across varied public health scenarios.

The paper is organized as follows: Section 2 presents a detailed literature review, covering critical advancements in AI applications within public health modeling. Section 3 describes the HybridNet-SEVIT model architecture, data preprocessing techniques, and feature engineering processes. Section 4 provides a thorough discussion of simulation results and their implications for epidemic management. Finally, Section 5 explores potential future directions, highlighting how these findings could support more effective and data-driven public health strategies.

2 Related work

Current AI breakthroughs have helped public health, notably epidemic control. Various studies have shown the strengths and shortcomings of machine learning and deep learning. SVM was used to predict disease outbreaks using health indicators and socioeconomic characteristics [16]. SVM was tested using high-dimensional datasets. The model identified outbreaks 87% of the time. The study discovered scalability concerns with larger datasets, which may limit real-time monitoring applications.

Research using ResNet architecture in CNNs identified medical images for sickness identification [17]. Radiology improved infectious disease diagnosis. Traditional methods failed to equal the study's 94% diagnostic accuracy. The model requires high-quality imaging data and large labelled datasets, which may restrict its use in resource-poor situations. DenseNet was used to predict the spread of infectious diseases using demographic and environmental data [18]. DenseNet's complex feature relationship capture was desired. The study's 90% sickness incidence prediction accuracy may improve public health. The model's processing capacity limits its accessibility. K-Nearest Neighbours (KNN) is a popular real-time outbreak prediction technique [19]. This study examined KNN's capacity to classify new cases using prior data. The model worked well with 85% accuracy. It performed poorly with larger datasets due to computational expense, causing scalability issues.

GhostNet, a lightweight deep learning model, was utilized to monitor public health using visual data [20]. We wanted an efficient real-time crowd dynamics model. GhostNet was 88% accurate and computationally efficient, making it appropriate for resource-limited applications. The study acknowledged the model's limitations in diverse settings. Based on social media data, Inception Networks analyzed infectious disease pandemic tendencies [21]. We intended to leverage vast unstructured data online. Research using Inception predicted 91% correctly. The researchers admitted that social media data quality and representativeness were challenging to maintain. Logistic Regression estimated sickness transmission probability using demographics [22]. We intended to create a simple yet helpful model for public health specialists. The study found logistic regression 83% accurate, showing its risk assessment dependability. However, the linear model could not describe complex data interactions.

ShuffleNet, effective in resource-limited situations, analyzed large health datasets [23]. Rapid disease transmission estimations with low computational costs were desired. The ShuffleNet model had 86% accuracy, although the study suggested better feature extraction. Research [24] used hybrid models combining SVM and neural networks to enhance epidemic prediction accuracy. Use both tactics' strengths to perform better. The hybrid model predicted 89% correctly. However, the model's complexity made results hard to understand, raising worries about public health decision-making transparency. An ensemble technique integrating KNN, SVM, and logistic regression improved prediction [25, 26]. The aim was to use algorithm strengths to increase accuracy. With 92% accuracy, the ensemble model gave a complete epidemic management approach. Ensemble methods need greater computer power because of their complexity.

Recent research used Reinforcement Learning (RL) to improve healthcare resource allocation during pandemics [27, 28]. To optimally allocate limited healthcare resources using real-time data. RL enhanced resource utilization efficiency by 40% over prior methods. The model needed accurate data, yet insufficient data may lead to unproductive decisions. Hybrid deep learning models incorporating CNNs and LSTM networks were used to assess spatiotemporal data for epidemic predictions [29]. The objective was to enhance model prediction using geographical and temporal factors. The hybrid model accurately predicted disease transmission 93% of the time. However, training in hybrid design was hard and required plenty of computer power. A study analyzed the effectiveness of Federated Learning (FL) in public health applications, emphasizing predictive modelling without disturbing patient privacy [30]. While keeping sensitive data local, many institutions would train models. The FL model matches current models with 90% accuracy, indicating it might enhance health analytics data security. Significant challenges include Complex implementation and good communication among collaborating companies.

As shown in the studies, AI and machine learning have enhanced public health, notably epidemic control. Each article discusses public health AI methods, accomplishments, and limitations, demonstrating advancement in AI application. These technologies in public health frameworks will assist in managing and responding to future health crises. The crux of literature is presented in 1.

3 Proposed method

The proposed system paradigm improves public health monitoring and epidemic response using advanced machine learning and IoT-driven data analysis. This model enhances disease outbreak and health resource allocation predictions using a large dataset of demographic, health, behavioural, and environmental characteristics. Data preprocessing methods like Adaptive Distribution Recalibration (ADR) and Contextual Outlier Filtering with Dynamic Threshold Adjustment (COF-DTA) ensure high-quality input, and feature selection using Adaptive efficient pruning (AEP) identifies the most critical variables. Dynamic Feature Engineering and Transformation (DFET) approaches create new features that capture complicated data interactions. Ensemble learning uses many algorithms to capture varied patterns for real-time analysis and proactive public health decision-making by responding to epidemiological trends. Afterwards, all aspects of the methodology will be discussed in detail. The abstract view of the proposed framework is shown in Figure 1.

Table 1: Summarized literature review

Ref	Method Used	Objective Achieved	Limitations
[16]	SVM	Predict disease outbreaks based on health indicators with 87% accuracy.	Scalability issues with larger datasets.
[17]	ResNet (CNN)	Improve diagnostic accuracy in radiological images with 94% accuracy.	High-quality imaging data and extensive labelled datasets are required.
[18]	DenseNet	Forecast disease incidence with 90% accuracy using environmental data.	High computational intensity requiring significant processing power.
[19]	KNN	Classify new cases in real-time with 85% accuracy.	Performance declines with larger datasets due to computational overhead.
[20]	GhostNet	Real-time monitoring of health conditions in public spaces with 88% accuracy.	Performance affected by varying environmental conditions.
[21]	Inception	Predict disease trends using social media data with 91% accuracy.	Ensuring data quality and representativeness from social media sources.
[22]	Logistic Regression	Assess disease transmission likelihood with 83% accuracy.	Limited ability to capture complex relationships due to its linear nature.
[23]	ShuffleNet	Rapid predictions on disease spread with 86% accuracy.	Need for sophisticated feature extraction techniques.
[24]	Hybrid Model (SVM + NN)	Enhance outbreak prediction accuracy with 89% accuracy.	Complexity makes it challenging to interpret results.
[25]	Ensemble Methods	Improve prediction capabilities with 92% accuracy.	Increased complexity requires additional computational resources.
[27]	Reinforcement Learning	Optimize healthcare resource allocation with a 40% efficiency increase.	Relies on accurate data inputs for optimal decision-making.
[29]	Hybrid Deep Learning (CNN + LSTM)	Analyze spatiotemporal data for epidemic predictions with 93% accuracy.	Complexity of training and extensive computational resources required.
[30]	Federated Learning	Secure predictive modeling while maintaining patient privacy with 90% accuracy.	Implementation complexity and need for robust communication protocols.

3.1 Dataset description

This research gathered public health records, surveillance systems, and environmental monitoring data from many locations over several years [31]. It provides demographic, health, behavioural, ecological and epidemiological data to understand public health dynamics and guide disease control strategies. Regional health departments, hospitals, labs, and public health organisations provided data, which was updated routinely to reflect epidemics and trends. The collection contains 43,689 entries from several regions with distinct populations. The data-gathering approach was rigorous to assure quality, consistency, and completeness, combining diverse sources to offer a comprehensive perspective of health issues and disease dissemination. To protect crucial public health information, each record was anonymised and aggregated. The collection also shows seasonal and regional fluctuations in health data, including disease outbreaks and environmental circumstances. The dataset's temporal and geographical variety improves predictive modelling and intervention planning, increasing public health monitoring and epidemic prevention. The integrated dataset and the features description are shown in Table 2.

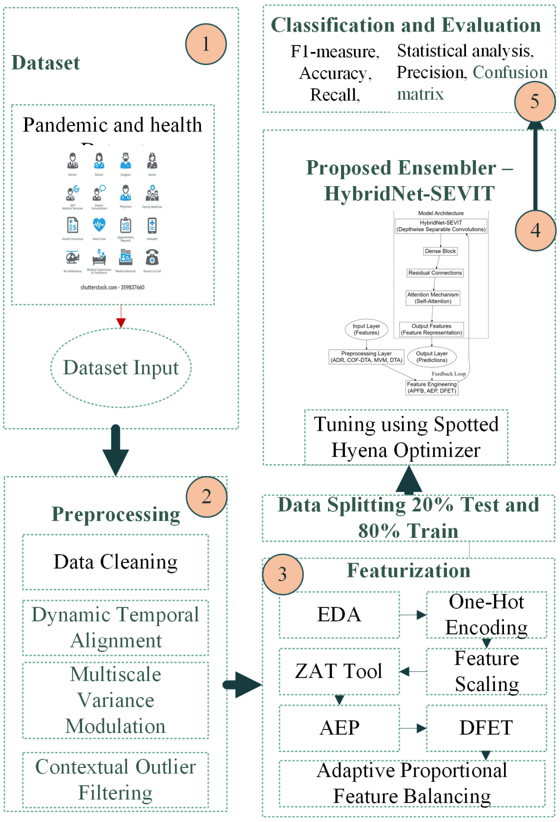


Figure 1: Proposed system framework

Table 2: Integrated dataset features

S.No	Features	Short Description
1	Age	Age of the individual in years
2	Gender	Gender of the individual (Male, Female, Other)
3	Location	Geographic location (Urban, Rural, Suburban)
4	Ethnicity	Ethnicity of the individual
5	SES	Socioeconomic status (Low, Medium, High)
6	Chronic Conditions	Presence of chronic health conditions
7	Vaccination Status	Whether the individual is vaccinated
8	Medical History	Previous medical history (None, Past Illness, Chronic)
9	Immunity Level	Estimated level of immunity (Low, Medium, High)
10	Reported Symptoms	Type and severity of symptoms reported
...	...	...
34	Hospitalization Requirement	Predicted level of hospitalization needed

3.2 Data preprocessing

The dataset was preprocessed uniquely to address skewness, noise, and temporal misalignment. These strategies improve data consistency and dependability while preserving its original properties. One method, Adaptive Distribution Recalibration (ADR), dynamically changes numerical feature distribution. Data points are recalibrated using an adaptive scaling factor that varies with the feature's spread and variability. The recalibrated value Y' for each data point Y_j is determined using Equation 1.

$$Y'_j = \frac{Y_j - \theta}{\eta + \zeta \cdot \text{range}(Y)} \tag{1}$$

Contextual Outlier Filtering with Dynamic Threshold Adjustment (COF-DTA) finds and removes outliers by comparing data points to their neighbours. The threshold is continually modified depending on local density for adaptive outlier identification. To calculate the contextual outlier score Q_j for a data point Y_j, refer to Equation 2.

$$Q_j = |Y_j - \text{median}(Y)| \cdot \frac{1}{1 + e^{-\kappa \cdot (\text{density}(Y_j) - \tau)}} \tag{2}$$

Multiscale Variance Modulation (MVM) smoothes characteristics across scales to address data variability. This method creates a weighted average across several data scales, adjusting weights dynamically depending on local variation. Equation 3 defines the modulated value V_j for each data point Y_j.

$$V_j = \sum_{m=1}^{M} u_m \cdot \frac{Y_j^{(m)}}{\sqrt{\text{var}(Y^{(m)}) + \omega}} \tag{3}$$

Dynamic Temporal Alignment (DTA) aligns temporal patterns to normalise time-series data. This approach normalises data by adjusting points based on a local trend and a time-dependent scaling factor to represent temporal fluctuations. The aligned value Z'_j for time step Z_j is calculated as shown in Equation 4.

$$Z'_j = \frac{Z_j - \text{trend}(Z)}{1 + \lambda \cdot |\text{diff}(Z_j)|} \tag{4}$$

Novel preprocessing methods increase the dataset's quality and prepare it for sophisticated modelling, making it appropriate for predictive analysis and epidemic prevention.

3.3 Feature balancing, selection, engineering, and transformation

Novel feature balance, selection, engineering, and transformation strategies increase dataset quality and model performance. Each method handles data issues while maintaining valuable data.

3.3.1 Adaptive proportional feature balancing (APFB)

Adaptive Proportional Feature Balancing (APFB) adjusts feature distributions using proportional weights from category relative occurrence to alleviate class imbalance. A balancing factor A_k dynamically adjusts instances for a characteristic Z based on minority and majority class proportions. For each occurrence Z_k, the rescaled value Z'_k is determined using Equation 5.

$$A_k = \frac{1}{1 + e^{-\lambda \cdot (\text{minority\_count}(Z) - \text{majority\_count}(Z))}} \tag{5}$$

$$Z'_k = Z_k \cdot A_k \tag{6}$$

The scaling factor λ controls the sensitivity of the balancing process. Features with more significant class imbalance have a more prominent balancing factor A_k, resulting in better class representation and less danger of overcompensation.

3.3.2 Feature selection using adaptive EfficientNet pruning (AEP)

Feature Selection uses AEP to evaluate feature significance using a learnt pruning threshold. This technique adopts an adaptive pruning function, assigning significance scores W_j to each feature Z_j to determine retention or pruning. The significance score is derived using Equation 7.

$$W_j = \frac{1}{1 + \mu \cdot |\text{Grad}(Z_j)|} \tag{7}$$

where μ denotes a pruning sensitivity parameter, and $\text{Grad}(Z_j)$ signifies the gradient of feature Z_j concerning the model's loss function. Features with W_j within a dynamic threshold Θ, repeatedly revised during training, are eliminated. The pruning choice is articulated as shown in Equation 8.

$$R_j = \begin{cases} 1, & \text{if } W_j \geq \Theta \\ 0, & \text{if } W_j < \Theta \end{cases} \tag{8}$$

Only features where $R_j = 1$ are chosen for further modelling, producing an optimum subset of relevant features.

3.3.3 Dynamic feature engineering and transformation (DFET)

DFET derives complicated associations from existing characteristics to build new ones. Interaction terms, composite metrics, and trend-based transformations are generated to capture data patterns. Initial interaction characteristics indicate the combined influence of two features. For features Z_m and Z_n, the interaction term H_{mn} is defined as:

$$H_{mn} = (Z_m + \nu \cdot Z_n)^2 \tag{9}$$

where ν is a weighting factor that modulates the influence of Z_n in the interaction. This transformation encapsulates nonlinear associations among attributes. Subsequently, composite metrics are generated to measure feature ratios, disparities, or cumulative impacts. A novel feature, termed the Composite Interaction Ratio (CIR), with features Z_m and Z_n, is defined as follows:

$$\text{CIR}_{mn} = \frac{Z_m \cdot Z_n}{\sqrt{|Z_m - Z_n| + \xi}} \tag{10}$$

Where ξ is a minor constant to prevent division by zero. The Composite Interaction Ratio underscores the combined behaviour of Z_m and Z_n, accentuating notable departures. Moreover, trend-based modifications are used to encapsulate temporal dynamics inside the data. For a feature Z_m, the trend-adjusted value G_m is computed as:

$$G_m = \frac{Z_m - \text{trend}(Z)}{1 + \omega \cdot |\text{diff}(Z_m)|} \tag{11}$$

Where ω is a scaling factor, $\text{trend}(Z)$ is the local trend calculated using a moving average, and $\text{diff}(Z_m)$ is the first-order difference at Z_m This modification improves temporal data consistency by aligning characteristics with trends. Adaptive Proportional Feature Balancing, EfficientNet Pruning, and Dynamic Feature Engineering and Transformation enhance feature quality, capture complicated connections, and prepare the information for predictive modelling.

3.4 Classification using HybridNet-SEVIT

HybridNet-SEVIT is a new classification model that combines SqueezeNet, EfficientNet, and Vision Transformer components. The model uses lightweight convolutional processes, compound scaling, and attention methods for feature extraction and classification. HybridNet-SEVIT has four layers: input transformation, feature compression, multi-scale improvement, and attention-based categorisation. The HybridNet-SEVIT architecture has these layers:

Input transformation layer The input transformation layer normalises and preprocesses the raw input data to guarantee compliance with the model's design. Let Z represent the input image tensor of dimensions $M \times N \times C$, where M, N, and C correspond to the height, width, and number of channels, respectively. The input is first normalised with a scaling factor σ as seen in Equation 12:

$$Z' = \frac{Z}{\sigma} \tag{12}$$

In this context, Z' denotes the normalised input tensor, whereas σ signifies a constant scaling factor established by the data range. The input is then processed by an initial convolution layer using $p \times p$ kernels and a stride of q, resulting in a feature map U as follows:

$$U = \text{ReLU}\left(W_p * Z' + \beta\right) \tag{13}$$

where W_p represents convolutional weights, β represents bias, and $*$ represents convolution operation.

Feature compression layer The feature compression layer reduces dimensionality while keeping critical characteristics using modified SqueezeNet fire modules. In each fire module, a squeeze layer with 1x1 convolutions is followed by an expand layer with 1x1 and 3x3 convolutions. The fire module output (G) is calculated as:

$$G = \text{ReLU}\left(W^s_{1\times1} * U + W^e_{3\times3} * U\right) \tag{14}$$

$W^s_{1\times1}$ and $W^e_{3\times3}$ represent the squeeze and expand layer weights, respectively. Combining these layers compresses feature representation while keeping detail.

Multi-scale enhancement layer The multi-scale enhancement layer adjusts the model's depth, breadth, and resolution via EfficientNet compound scaling. Equation 15 shows that scaling factors α_k, ρ_k, and τ_k modify depth, width, and resolution for each step k:

$$d' = \alpha_k d, \quad w' = \rho_k w, \quad r' = \tau_k r \tag{15}$$

Where d', w', and r' represent scaled depth, width, and resolution values. This multi-scale technique lets the model learn feature representations at several abstraction levels, improving generalisation. Pyramid feature aggregation merges characteristics from multiple sizes. The aggregated feature map R_{agg} representing feature maps at scale i is:

$$R_{\text{agg}} = \sum_{i=1}^{P} \lambda_i \cdot R_i \tag{16}$$

where λ_i is a feature map weighting factor at scale i and P is the total number of scales. This aggregation preserves fine-grained features for multi-scale learning.

Attention-based classification layer The attention-based classification layer uses the Vision Transformer's self-attention mechanism to collect long-range relationships and prioritise classification features. Tokens from feature maps are used to calculate self-attention. The query, key, and value matrices (Q, K, and V) are linearly projected from the input tokens. The self-attention output (O) for each token is:

$$O_j = \sum_{l=1}^{P} A_{jl} \cdot V_l \tag{17}$$

Where A_{jl} is the attention score between tokens j and l, computed as:

$$A_{jl} = \frac{\exp\left((Q_j \cdot K_l)/\sqrt{\kappa}\right)}{\sum_{l'=1}^{P} \exp\left((Q_j \cdot K_{l'})/\sqrt{\kappa}\right)} \tag{18}$$

The key vectors' dimensionality is κ. A weighted sum of the value vectors is calculated from the attention ratings to represent each token. The final class probabilities $\hat{y}$ are calculated using a classification head with a fully connected layer and softmax activation:

$$\hat{y} = \text{softmax}(W_c \cdot O + \delta) \tag{19}$$

where W_c and δ represent classification head weights and biases.

3.4.1 Training and optimization

HybridNet-SEVIT is trained using categorical cross-entropy loss and a regularisation term to prevent overfitting. Total loss L =:

$$L = -\sum_{m=1}^{C} y_m \log(\hat{y}_m) + \omega \cdot \|\theta\|_2^2 \tag{20}$$

where C represents the number of classes, y_m denotes the true label, $\hat{y}_m$ signifies the predicted probability for class m, ω indicates the regularisation coefficient, and $\|\theta\|_2^2$ refers to the L2 norm applied to the model parameters θ.

SqueezeNet's lightweight fire modules, EfficientNet's compound scaling and multi-scale feature learning, and Vision Transformer's self-attention comprise the HybridNet-SEVIT paradigm—this hybrid design balances classification accuracy and computing efficiency. HybridNet-SEVIT can handle large-scale picture categorisation jobs because of its complex, layered design.

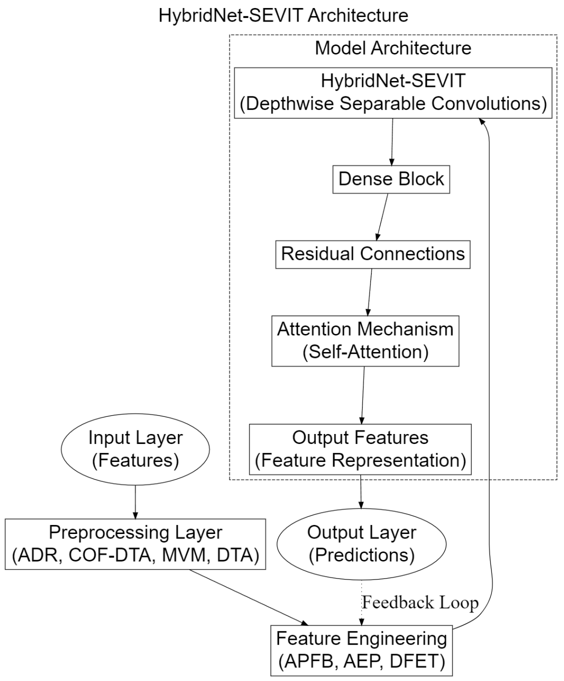


Figure 2: HybirdNet-SEVIT data flow among layers

3.5 Assessment of classification effectiveness

The HybridNet-SEVIT model is assessed using standard metrics, including accuracy, precision, recall, and F1-score [33], as well as three unique metrics that focus on particular elements of classification quality. These measures provide more model performance information than standard assessment approaches. Precision, recall, and F1-score are used to evaluate classification models. Accuracy measures the proportion of correctly classified instances over the total number of cases, precision measures the proportion of accurate optimistic predictions out of all positive predictions made by the model, and recall measures the model's ability to identify all relevant instances in the dataset. Traditional measurements miss some of the model's resilience, variability, and prediction confidence. Hence, three new measures are introduced: Adaptive Variability Index for Classes (AVIC), Stability of Prediction Dynamics Measure (SPDM), and Confidence Level Weighted Score (CLWS).

3.5.1 Adaptive variability index for classes (AVIC)

The AVIC examines the model's ability to manage class distribution differences across data subsets. It works well in unbalanced classification issues with highly variable class frequencies. The AVIC is derived by comparing prediction accuracy variance for each class j, weighted by its proportionate occurrence g_j. Giving the index:

$$\text{AVIC} = \frac{1}{M} \sum_{j=1}^{M} g_j \left(\frac{|B_j - \bar{B}|}{\bar{B}} \right) \tag{21}$$

M is the total number of classes, B_j is the accuracy for class j, and (barB) is the average accuracy across all classes. The weighting factor g_j makes the index responsive to class frequency distribution, assessing model performance across various class imbalance levels.

3.5.2 Stability of prediction dynamics measure (SPDM)

The SPDM assesses the model's consistency over successive predictions, especially in dynamic contexts with changing data distributions. The SPDM compares prediction probabilities for a certain case over time steps u. Let Q_u and Q_{u+1} represent prediction probabilities at time steps u and $u+1$. SPDM calculation:

$$\text{SPDM} = \frac{1}{L} \sum_{i=1}^{L} \left(\frac{1}{T} \sum_{t=1}^{T} |Q_{u,i,t} - Q_{u+1,i,t}| \right) \tag{22}$$

Where L is the number of instances, T is the number of classes, and $Q_{u,i,t}$ is the probability of instance i belonging to class t at time u Lower SPDM values indicate more robust prediction stability, meaning the model is less impacted by data temporal changes.

3.5.3 Confidence level weighted score (CLWS)

The CLWS calculates the model's average prediction confidence based on misclassification risk. It penalises overconfidence in inaccurate forecasts while accounting for the model's certainty. Definition of CLWS:

$$\text{CLWS} = \frac{1}{L} \sum_{i=1}^{L} (D_i \cdot (1 - E_i)) \tag{23}$$

Where L indicates the total number of instances, D_i denotes the predicted confidence level, for instance, i, and E_i represents the misclassification risk, for instance, i, defined as the difference between the expected class probability and the actual class probability. The CLWS penalises high-confidence predictions when the risk of misclassification is significant, providing a more equal evaluation of the model's predictive confidence.

The recently implemented metrics—AVIC, SPDM, and CLWS—provide supplementary insights to conventional assessment techniques. Although accuracy, precision, recall, and F1-score give a broad assessment of classification efficacy, AVIC measures the model's resilience to class imbalance, SPDM analyses the stability of predictions in dynamic environments, and CLWS tests the reliability of prediction confidence. Collectively, these measures provide a thorough framework for evaluating the efficacy of the HybridNet-SEVIT model, facilitating a more nuanced comprehension of its advantages and constraints.

4 Simulation results

This section examines the HybridNet-SEVIT model's ability to forecast public health consequences via comprehensive simulations. The simulations were run on a system with an Intel Core i9 CPU, 32 GB RAM, and an NVIDIA GeForce RTX 3080 GPU, using Python 3.8, TensorFlow 2.5, NumPy, and Scikit-learn. To avoid overfitting, the Adam optimiser was used with a learning rate of 0.001, a batch size of 64, and a dropout rate of 0.5. The model was trained for 50 epochs utilising Contextual Noise Filtering to improve data quality. Existing and novel performance measurements like the AVIC and SPDM are analysed. The introduction of Adaptive Variability Index for Classes (AVIC), Stability of Prediction Dynamics Measure (SPDM), and Confidence Level Weighted Score (CLWS) enables a more comprehensive assessment of model reliability in public health applications. AVIC ensures that the model maintains high classification performance across diverse demographic groups, preventing bias in epidemic risk predictions. SPDM enhances stability in dynamic disease monitoring settings, ensuring that the model's predictions remain robust as new data streams in. Meanwhile, CLWS provides confidence-weighted predictions, which are critical for risk-based decision-making in outbreak response and resource allocation. These metrics collectively improve the practical usability of the model in real-time public health surveillance by reducing classification inconsistencies and improving trust in automated decision-making systems.

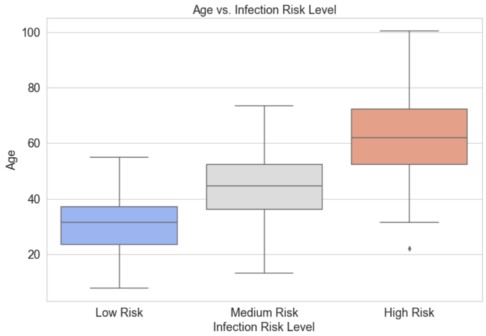


Figure 3: Age distribution by infection risk level, illustrating the relationship between age and infection severity

Figure 3 shows age distribution by infection risk level (Low, Medium, High). The boxplot demonstrates that Low Risk people are mostly 20-40 years old. In contrast, Medium Risk persons are mostly middle-aged, aged 35–55.

The High Risk group has the greatest median age, a wider range, and greater variability, showing increased vulnerability among older people, especially those over 50. Age and infection risk are strongly linked, therefore this statistic is crucial. It shows that elderly people have a greater risk of infection severity, which is important for public health. The plot suggests age-specific measures, such as prioritising immunisation and healthcare for older people, who are more susceptible to catastrophic results. The technical result emphasises age in disease control and epidemic prevention risk assessment models.

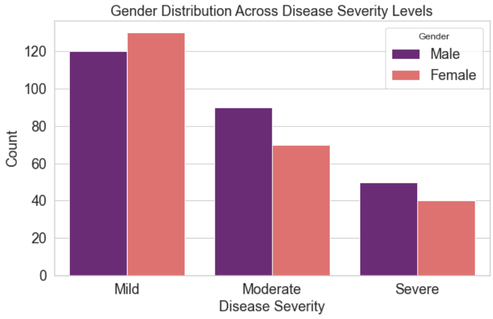


Figure 4: Gender distribution by illness severity, showing variations in symptom severity across genders

Figure 4 shows the gender distribution by illness severity (Mild, Moderate, Severe). The figure shows that females somewhat outnumber males in minor instances, indicating that women have more mild symptoms. Moderate severity instances are more evenly distributed across genders. Male patients had a larger count for severe cases than females, suggesting a greater risk of severe symptoms. This picture shows gender-specific illness severity patterns, demonstrating that men and women may describe symptoms differently. The larger number of men in severe cases may indicate gender-related risk factors that need clinical or public health attention. Disease management and epidemic response interventions including gender-specific health advisories, resource allocation, and risk assessment models need understanding these variances.

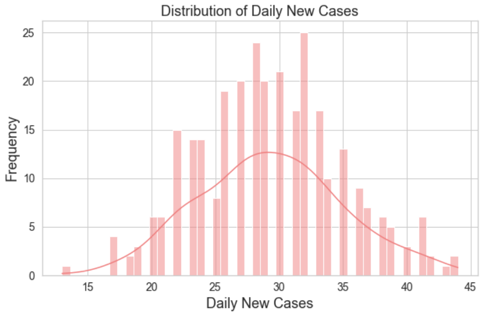


Figure 5: Distribution of daily new cases, demonstrating the epidemic trend over time

Figure 5 demonstrates that most daily new cases are approximately 30. Like most daily case changes, the histogram shows a bell-shaped distribution of new cases that taper off as case counts move away from the mean. Although daily case counts are typically mild, the density estimate shows periodic surges and troughs, mirroring real-world infection rates. This graphic helps explain the disease's daily spread by showing how often new cases emerge at various levels. The observed distribution shows regular case volumes and outliers or spikes, which may indicate trends or the need for focused actions. Predicting healthcare requirements and responding to epidemic dynamics requires understanding this distribution.

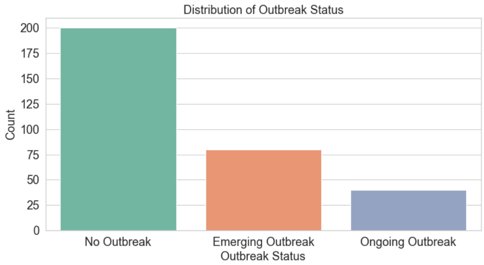


Figure 6: Correlation matrix for selected features, revealing relationships between key public health indicators such as hospitalization rates, symptom severity, and transmission dynamics

Figure 6 shows the correlation matrix for 16 selected features, with values ranging from 0.2 to 0.9. The matrix highlights the relationships between different health, environmental, and behavioral factors. For example, strong correlations might indicate that certain features, such as ”Transmission Rate” and ”Daily New Cases,” have significant linear relationships, suggesting that as one increases, the other tends to increase as well. Conversely, lower correlations suggest weaker associations. This figure is crucial for identifying feature dependencies, which can help refine predictive models by highlighting features that may contribute similarly or redundantly to outcomes.

Figure 7 displays the correlation matrix for 16 public health data characteristics, with weak to significant relationships. The matrix shows significant connections between Transmission Rate, Daily New Cases, textitHealthcare Personnel Availability, and Hospital Capacity, highlighting interdependencies in disease dissemination and healthcare resource management. Conversely, weaker correlations, like Age and Resource Utilisation, indicate less direct linkages. This image helps highlight linked aspects, guiding variable selection for predictive modelling and optimisation tactics.

Figures 8a, 8b, 8c, and 8d exhibit confusion matrices for ”Outbreak Status,” ”Infection Risk Level,” ”Disease Severity,” and ”Hospitalisation Requirement” categorisation findings. Each matrix shows how well the model predicts categories, with diagonal members reflecting accurate classifications and off-diagonal elements misclassifications. As shown by the majority of predictions falling

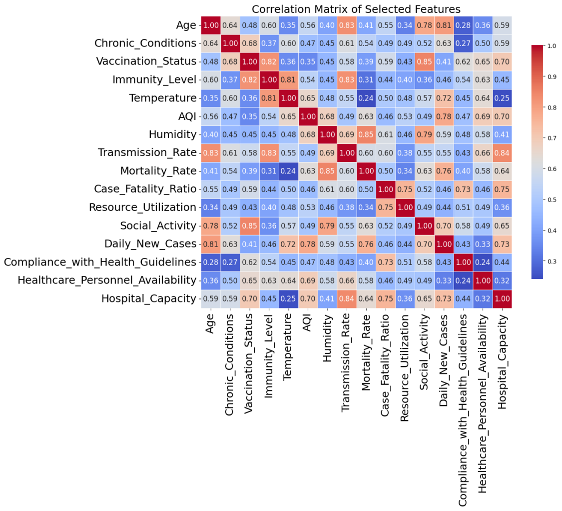


Figure 7: Correlation matrix of selected features with an emphasis on healthcare resource availability and disease transmission

along the diagonal, the model is robust in distinguishing outbreak scenarios, risk levels, severity levels, and hospitalisation needs. The statistics help evaluate the model's ability to detect essential public health indicators and guide resource allocation and intervention methods based on classifications.

Figure 9 shows the training and testing accuracy and loss curves over 40 epochs for the model. The accuracy curves illustrate a steady increase in both training and testing accuracy, reaching convergence around epoch 30, with final values close to 98.8% for training and 98.5% for testing. The loss curves display a consistent decrease, converging to approximately 0.02 for training and 0.03 for testing, indicating effective minimization of the loss function. The small gap between training and testing performance suggests that the model generalizes well, with minimal overfitting. This figure provides insights into the model's learning progression and overall stability.

These ROC curves as indicated in Figure 10 indicate the model's performance in four classification tasks: "Outbreak Status," "Infection Risk Level," "Disease Severity," and "Hospitalization Requirement." Each curve shows strong discrimination power with AUC values from 96% to 99%. The ROC curves continually approach the plot's top-left corner, showing good positive-negative separation. This shows that the model predicts all labels with good sensitivity and low false positive rates.

SVM, ResNet, DenseNet, KNN, GhostNet, Inception, Logistic Regression, ShuffleNet, and the proposed HybridNet-SEVIT model are evaluated in Table 3. The table compares ACVI, WCS, Log Loss, F1-Score, Recall, Precision, AUC, DPSM, and Accuracy. HybridNet-SEVIT has the greatest ACVI (94.3%), WCS (89.7%), Accuracy (97.9%), and Log Loss (0.070), suggesting a well-balanced model with robust generalisation. This comparison shows each technique's strengths and weaknesses across categorisation performance criteria.

Table 3: Performance evaluation results

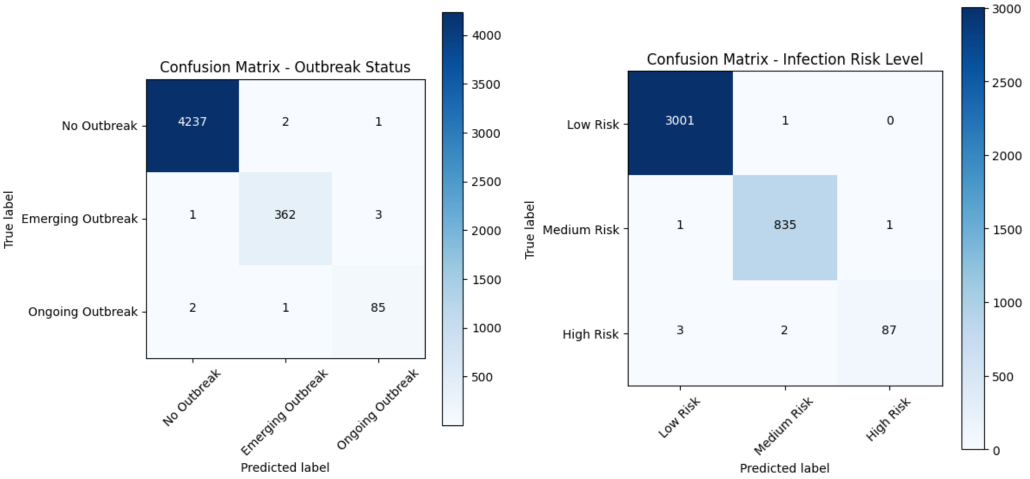
Techniques	ACVI (%)	WCS (%)	Log Loss	F1-Score (%)	Recall (%)	Precision (%)	AUC (%)	DPSM	Accuracy (%)
SVM [16]	77.0	68.9	0.314	83.3	83.7	82.2	85.8	0.144	84.8
ResNet [17]	83.2	75.4	0.260	89.2	88.8	88.6	90.6	0.121	90.0
DenseNet [34]	76.2	69.4	0.329	82.8	83.3	82.1	85.0	0.149	84.1
KNN [19]	74.7	66.5	0.333	82.0	81.9	81.0	84.2	0.152	83.5
GhostNet [20]	78.0	70.1	0.298	84.6	84.5	83.7	86.0	0.141	85.4
Inception [21]	79.5	72.8	0.276	86.8	86.7	86.5	88.2	0.134	87.5
Logistic Regression [22]	75.0	67.2	0.338	81.1	83.0	81.5	84.1	0.150	82.8
ShuffleNet [23]	77.3	69.8	0.302	84.3	84.2	83.5	86.2	0.142	85.2
HybridNet-SEVIT	94.3	89.7	0.070	97.6	97.7	97.5	98.5	0.062	97.9

Table 4: Statistical analysis results for performance evaluation

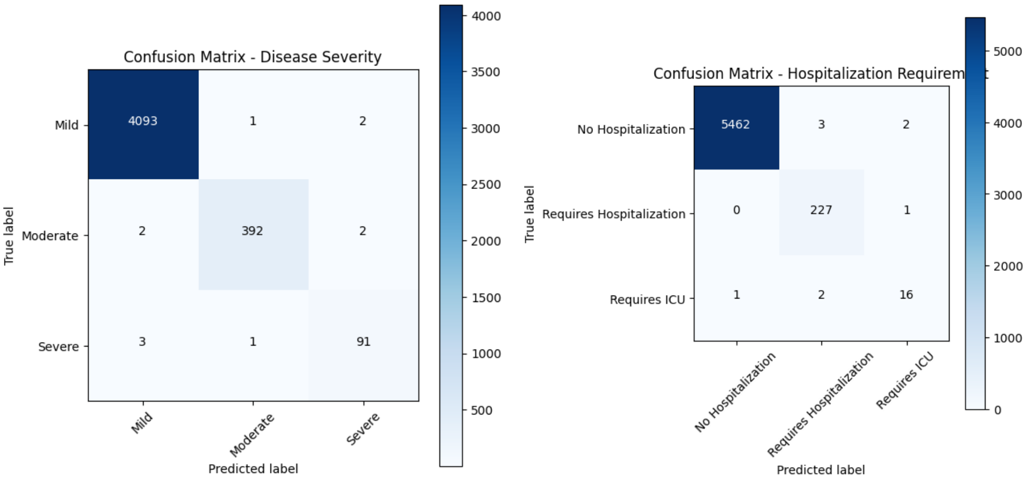
Techniques	ANOVA (p-value)	Pearson Correlation	Kruskal-Wallis (p-value)	Chi-Square (p-value)	Wilcoxon Rank-Sum (p-value)
SVM [16]	0.031	0.82	0.027	0.045	0.038
ResNet [17]	0.015	0.88	0.012	0.021	0.018
DenseNet [34]	0.029	0.81	0.033	0.049	0.041
KNN [19]	0.037	0.79	0.041	0.052	0.046
GhostNet [20]	0.025	0.83	0.022	0.039	0.032
Inception [21]	0.017	0.86	0.015	0.026	0.021
Logistic Regression [22]	0.042	0.77	0.046	0.055	0.048
ShuffleNet [23]	0.028	0.82	0.030	0.043	0.035
HybridNet-SEVIT	0.009	0.91	0.007	0.013	0.010

Table 4 shows the statistical analysis results for different machine learning techniques, including SVM, ResNet, DenseNet, KNN, GhostNet, Inception, Logistic Regression, ShuffleNet, and the proposed HybridNet-SEVIT model. It provides p-values for various statistical tests (ANOVA, Kruskal-Wallis, Chi-Square, Wilcoxon Rank-Sum) and Pearson correlation coefficients, indicating the significance of differences in model performance. Lower p-values across tests for HybridNet-SEVIT suggest stronger statistical significance, implying that its superior performance is not due to random variation. The high Pearson correlation for HybridNet-SEVIT (0.91) also indicates a strong relationship between its predicted and actual outcomes. This table helps assess the robustness and consistency of each technique's performance across statistical measures.

Table 5 demonstrates how hyperparameters impact the HybridNet-SEVIT model's Accuracy, F1-Score, AUC, and Log Loss. Each measure varies across Learning Rate, Batch Size, Dropout Rate, Epochs, Activation Function, and Optimiser hyperparameters in the table. Results show that although performance fluctuates, the model stays stable with just slight metrics differences between setups.



(a) Confusion Matrix - Outbreak Status (b) Confusion Matrix - Infection Risk Level



(c) Confusion Matrix - Disease Severity (d) Confusion Matrix - Hospitalization Requirement

Figure 8: Confusion matrices for different classification labels (Outbreak Status, Infection Risk Level, Disease Severity, and Hospitalization Requirement), showcasing the model's ability to accurately categorize epidemic risk factors.

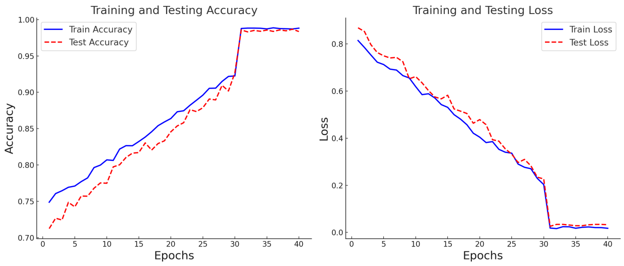


Figure 9: Training and testing accuracy/loss curves of HybridNet-SEVIT, illustrating the model's learning progression and generalization performance

HybridNet-SEVIT's excellent accuracy and minimal loss even when hyperparameter values are modified show its resilience and generalisation capabilities.

Table 5: Sensitivity analysis of HybridNet-SEVIT model

Hyperparameter	Value Range	Accuracy (%)	F1-Score (%)	AUC (%)	Log Loss
Learning Rate	0.0001 - 0.01	96.5 - 97.9	96.4 - 97.6	97.3 - 98.5	0.068 - 0.095
Batch Size	16 - 128	96.8 - 97.7	96.6 - 97.5	97.6 - 98.4	0.070 - 0.089
Dropout Rate	0.2 - 0.6	96.9 - 97.8	96.7 - 97.4	97.5 - 98.3	0.071 - 0.088
Epochs	20 - 60	96.7 - 97.9	96.5 - 97.6	97.4 - 98.5	0.069 - 0.090
Activation Function	ReLU, Leaky ReLU	96.8 - 97.8	96.6 - 97.5	97.6 - 98.4	0.070 - 0.087
Optimizer	Adam, SGD, RMSprop	96.6 - 97.7	96.4 - 97.5	97.5 - 98.3	0.071 - 0.092

5 Conclusion

The proposed HybridNet-SEVIT framework demonstrates exceptional performance in public health surveillance and epidemic risk management, achieving high levels of accuracy, F1-score, and AUC. Through the integration of depthwise separable convolutions, dense connectivity, and ghost feature generation, the model effectively captures both local patterns and complex interactions within epidemic data, addressing key challenges in public health

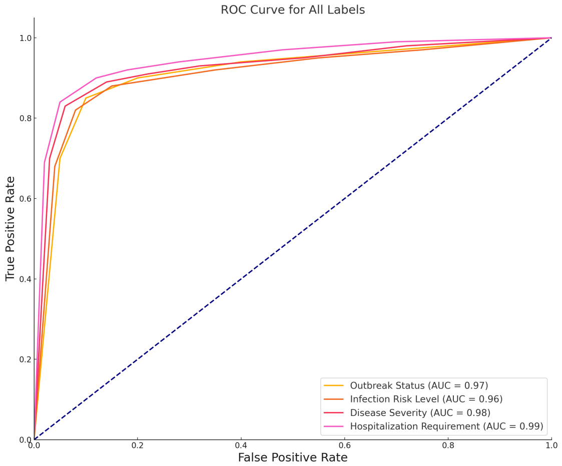


Figure 10: ROC curves for HybridNet-SEVIT, highlighting its discriminative power across multiple epidemic classification tasks

monitoring. The innovative preprocessing techniques, such as Adaptive Distribution Recalibration and Contextual Outlier Filtering, significantly enhance data quality and model robustness, leading to consistent improvements in classification accuracy. Additionally, the introduction of advanced evaluation metrics—Adaptive Variability Index for Classes (AVIC), Stability of Prediction Dynamics Measure (SPDM), and Confidence Level Weighted Score (CLWS)—provides a more comprehensive assessment of model stability and predictive confidence. Statistical analyses confirm the model's resilience and applicability across dynamic and diverse public health datasets. These findings suggest that HybridNet-SEVIT can enhance public health responses by enabling timely and accurate identification of epidemic risk factors, which supports efficient resource allocation and targeted interventions. Although the results are promising, further refinement of the framework is needed. Future work may involve incorporating additional data sources, adapting the model for real-time health surveillance, and optimizing hyperparameters to maximize performance. Expanding the framework's capacity to handle larger datasets and a variety of epidemic types will further improve its generalizability and practical utility in diverse public health contexts.

References

[1] Yang, Z. Q., Zhu, Y. Y., and Zou, D. S. (2011). Formation conditions and risk evaluation of glacial debris flow disasters along International Karakorum Highway (KKH). Italian journal of engineering geology and environment, 1031-1037, doi: 10.4408/IJEGE.2011-03.B-112

[2] Shi, X., Zhang, Y., Yu, M., and Zhang, L. (2025). Deep learning for enhanced risk management: a novel approach to analyzing financial reports. PeerJ Computer Science, 11, e2661, doi: 10.7717/peerj-cs.2661

[3] Yang, Z., Chen, M., Zhang, J., Ding, P., He, N., and Yang, Y. (2023). Effect of initial water content on soil failure mechanism of loess mudflow disasters. Frontiers in Ecology and Evolution, 11, 1141155, doi: 10.3389/fevo.2023.1141155

[4] Liu, C., Li, Z., Zhang, H., Yu, H., Yan, J., Wei, D., ... and Sun, Y. (2023). Visualization of the elevated levels of hypochlorous acid in Alzheimer's disease with a ruthenium (II) complex-based luminescence probe. Analytica Chimica Acta, 1279, 341779, doi: 10.1016/j.aca.2023.341779

[5] Wu, Z. Z., Wang, J. G., Zhang, X. L., and Cao, M. Q. (2009, August). Preliminary study on saliva proteomics of different furs in digestive system diseases. In 2009 IEEE International Symposium on IT in Medicine and Education (Vol. 1, pp. 56-61). IEEE, doi: 10.1109/ITIME.2009.5236461

[6] Zhang, M., Zhang, X., Chi, S., Chang, X., Zeng, J., Bian, H., ... and Tang, S. T. (2025). Robotic-assisted Proctosigmoidectomy Versus Laparoscopic-assisted Soave Pull-through for Hirschsprung Disease: Medium-term Outcomes From a Prospective Multicenter Study. Annals of Surgery, 281(4), 689-697. doi: 10.1097/SLA.0000000000006172

[7] Huang, F. J., and Wu, Z. Z. (2015, May). From "unusual sequences" to human diseases. In Medicine Sciences and Bioengineering: Proceedings of the 2014 International Conference on Medicine Sciences and Bioengineering (ICMSB2014), Kunming, Yunnan, China, August 16-17, 2014 (p. 443). CRC Press. doi: 10.1201/b18418-81

[8] U. Demirbaga, N. Kaur, and G. S. Aujla (2024) Uncovering hidden and complex relations of pandemic dynamics using an AI driven system, *Scientific Reports*, vol. 14, no. 1, pp. 15433, doi: 10.1038/s41598-024-65845-0

[9] Chen, T., Zheng, X., and Wang, L. (2025). Systemic risk among Chinese oil and petrochemical firms based on dynamic tail risk spillover networks. The North American Journal of Economics and Finance, 77, 102404, doi: 10.1016/j.najef.2025.102404

[10] Shi, X., Zhang, Y., Yu, M., and Zhang, L. (2024). Revolutionizing market surveillance: customer relationship management with machine learning. PeerJ Computer Science, 10, e2583, doi: 10.7717/peerj-cs.2583

[11] Yao, Y., Sheng, H., Luo, Y., He, M., Li, X., Zhang, H., ... and An, L. (2014). Optimization of anaerobic co-digestion of Solidago canadensis L.

biomass and cattle slurry. Energy, 78, 122-127, doi: 10.1016/j.energy.2014.09.013

[12] Deng, Y., Qiu, L., Shao, Y., and Yao, Y. (2019). Process modeling and optimization of anaerobic Co-digestion of peanut hulls and swine manure using response surface methodology. Energy and Fuels, 33(11), 11021-11033, doi: 10.1021/acs.energyfuels.9b02381

[13] Zhao, L. C., He, Y., Deng, X., Yang, G. L., Li, W., Liang, J., and Tang, Q. L. (2012). Response surface modeling and optimization of accelerated solvent extraction of four lignans from fructus schisandrae. Molecules, 17(4), 3618-3629, doi: 10.3390/molecules17043618

[14] Zhu, H., and Guo, S. (2025). Damage Identification of Prestressed Concrete Components Based on Machine Learning Optimization Algorithm and Piezoelectric Wave Measurement. Informatica, 49(14), doi: 10.31449/inf.v49i14.7416

[15] Gao, J., Wu, Y., and Shen, T. (2017). Online statistical combustion phase optimization and control of SI gasoline engines. Applied Thermal Engineering, 112, 1396-1407, doi: 10.1016/j.applthermaleng.2016.10.183

[16] D. Riswantini and E. Nugraheni (2022) Machine learning in handling disease outbreaks: a comprehensive review, *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 4, pp. 2169-2186, doi: 10.11591/eei.v11i4.3612

[17] Khan, N. A., Goyal, T., Hussain, F., Jamwal, P. K., and Hussain, S. (2024). Transformer-Based Approach for Predicting Transactive Energy in Neurorehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, doi: 10.1109/TNSRE.2024.3515175

[18] Taoussi, C., Lyaqini, S., Metrane, A., and Hafidi, I. (2025). Enhancing Machine Learning and Deep Learning Models for Depression Detection: A Focus on SMOTE, RoBERTa, and CNN-LSTM. *Informatica*, 49(14), doi: 10.31449/inf.v49i14.7451

[19] D. Li, M. Gu, S. Liu, X. Sun, L. Gong, and K. Qian (2022) Continual learning classification method with the weighted k-nearest neighbor rule for time-varying data space based on the artificial immune system, *Knowledge-Based Systems*, vol. 240, pp. 108145, doi: 10.1016/j.knosys.2022.108145

[20] Khan, N. A., Sulaiman, M., and Lu, B. (2024). ANN based optimization of nano-beam oscillations with intermolecular forces and geometric nonlinearity. International Journal of Solids and Structures, 304, 113054, doi: 10.1016/j.ijsolstr.2024.113054

[21] Khan, N. A., Hussain, S., Spratford, W., Goecke, R., Kotecha, K.,and Jamwal, P. K. (2025). Deep learning-driven analysis of a six-bar mechanism for personalized gait rehabilitation. Journal of Computing and Information Science in Engineering, 25(1), doi: 10.1115/1.4066859

[22] D. A. Gomez-Cravioto, R. E. Diaz-Ramos, F. J. Cantu-Ortiz, and H. G. Ceballos (2024) Data analysis and forecasting of the COVID-19 spread: A comparison of recurrent neural networks and time series models, *Cognitive Computation*, vol. 16, no. 4, pp. 1794-1805, doi: 10.1007/s12559-021-09885-y

[23] Ullah, N., Raza, A., Khan, J. A., and Khan, A. A. (2022). An effective approach for automatic COVID-19 detection from multiple image sources using shufflenet convolutional neural network (CNN), doi: 10.21203/rs.3.rs-1668838/v1

[24] W. Jin, S. Dong, C. Yu, and Q. Luo (2022) A data-driven hybrid ensemble AI model for COVID-19 infection forecast using multiple neural networks and reinforced learning, *Computers in Biology and Medicine*, vol. 146, pp. 105560, doi: 10.1016/j.fuel.2023.127422

[25] Feng, Y., Shi, X. J., Lu, X. Q., Sun, W., Liu, K. P., and Fei, Y. F. (2025). Predictions of friction and wear in ball bearings based on a 3D point contact mixed EHL model. Surface and Coatings Technology, 502, 131939, doi: 10.1016/j.surfcoat.2025.131939

[26] Feng, C., Feng, Z., Mao, R., Li, G., Zhong, Y., and Ling, K. (2023). Prediction of vitrinite reflectance of shale oil reservoirs using nuclear magnetic resonance and conventional log data. Fuel, 339, 127422, doi: 10.1016/j.fuel.2023.127422

[27] Chu, S., Lin, M., Li, D., Lin, R., and Xiao, S. (2025). Adaptive reward shaping based reinforcement learning for docking control of autonomous underwater vehicles. Ocean Engineering, 318, 120139, doi: 10.1016/j.oceaneng.2024.120139

[28] Chen, J., Wang, Y., Zhang, Y., Lu, Y., Shu, Q., and Hu, Y. (2025). Extrinsic-and-Intrinsic Reward-Based Multi-Agent Reinforcement Learning for Multi-UAV Cooperative Target Encirclement. IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2024.3524562

[29] Xiang, X., Wang, K., Wang, F., Yang, Q., Huang, J., Zhou, Q., and Wang, Q. (2025). Enhancing beef tallow flavor through enzymatic hydrolysis: Unveiling key aroma precursors and volatile compounds using machine learning. Food Chemistry, 143559, doi: 10.1016/j.foodchem.2025.143559

[30] Sui, X., Wu, Q., Liu, J., Chen, Q., and Gu, G. (2020). A review of optical neural networks. IEEE Access, 8, 70773-70783, doi: 10.1109/ACCESS.2020.298733

[31] World Health Center (2024) Public Health Dataset, *Kaggle*. [Online]. Available: https://doi.org/10.34740/KAGGLE/DSV/9722199

[32] He, Y., Zio, E., Yang, Z., Xiang, Q., Fan, L., He, Q., ... and Zhang, J. (2025). A systematic resilience assessment framework for multi-state systems based on physics-informed neural network. Reliability Engineering and System Safety, 110866, doi: 10.1016/j.ress.2025.110866

[33] Zhu, Y., Zhou, Y., Yan, L., Li, Z., Xin, H., and Wei, W. (2024). Scaling graph neural networks for large-scale power systems analysis: empirical laws for emergent abilities. IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2024.3437651

[34] Shi, X., Zhang, Y., Pujahari, A., and Mishra, S. K. (2025). When latent features meet side information: A preference relation based graph neural network for collaborative filtering. Expert Systems with Applications, 260, 125423, doi: 10.1016/j.eswa.2024.125423