Preprints.org

Article

Not peer-reviewed version

Optimizing IoMT Security: Performance Trade-Offs Between Neural Network Architectural Design, Dimensionality Reduction, and Class Imbalance Handling

Heyfa Ammar * and Asma Cherif *

Posted Date: 22 October 2025

doi: 10.20944/preprints202510.1779.v1

Keywords: intrusion detection; IoT; healthcare security; neural networks; ANN; autoencoders; class imbalance

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Optimizing IoMT Security: Performance Trade-Offs Between Neural Network Architectural Design, Dimensionality Reduction, and Class Imbalance Handling

Heyfa Ammar [1,2,*] **and Asma Cherif** [3,4,*]

[1] RIOTU Lab, Computer Science Department, College of Computer and Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia
[2] RISC-ENIT Lab, National College of Engineers, University of Tunis-ElManar, Tunisia
[3] Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia
[4] Center of Excellence in Smart Environment Research, King Abdulaziz University, Jeddah, Saudi Arabia
* Correspondence: hammar@psu.edu.sa (H.A.); acherif@kau.edu.sa (A.C.)

Abstract

The proliferation of Internet of Medical Things (IoMT) devices in healthcare requires robust intrusion detection systems to safeguard sensitive data and ensure patient safety. This study presents a comprehensive evaluation of advanced Artificial Neural Network (ANN) architectures and autoencoder (AE) preprocessing for intrusion detection in IoMT environments. Using the WUSTL-EHMS-2020 dataset, we investigate whether feature dimensionality reduction through autoencoders enhances detection performance when combined with sophisticated neural network designs. Our research implements and evaluates several ANN variants, including standard feedforward networks, dual-branch models with addition and concatenation operations, and architectures incorporating shortcut connections. To address the prevalent class imbalance, we compare three distinct approaches: Synthetic Minority Over-sampling Technique (SMOTE), weighted loss functions, and a hybrid over-under sampling strategy. Performance evaluation reveals that while autoencoders achieve high precision (93.83% with SMOTE), they sacrifice recall (74.85%) due to information loss during compression. The dual-branch ANN with addition operations, combined with weighted loss function and without dimensionality reduction, achieves superior overall performance with 94.03% accuracy and 0.8716 F1-score.

Keywords: intrusion detection; IoT; healthcare security; neural networks; ANN; autoencoders; class imbalance

1. Introduction

The pervasive integration of Internet-connected devices into healthcare systems has given rise to the Internet of Medical Things (IoMT), a specialized network that has transformed patient care and medical data management [1]. This technological evolution, while beneficial, introduces significant security vulnerabilities that could compromise patient safety and data confidentiality [2]. The sensitive nature of healthcare data makes these systems particularly attractive targets for cyber attackers, necessitating robust security mechanisms [3].

Intrusion Detection Systems (IDSs) serve as critical defensive components in healthcare cybersecurity frameworks, continuously monitoring network communications to identify suspicious activities and security breaches [4]. These systems employ various detection methodologies, including signature-based, behavior-based, and anomaly-based approaches [5–8]. Among these, anomaly-based detection has emerged as particularly valuable for identifying novel threats by detecting deviations from established baseline behaviors. However, this approach faces challenges related to high false-positive rates and vulnerability to sophisticated attack vectors [9].

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

The advent of advanced machine learning techniques has created new opportunities for enhancing intrusion detection capabilities. Previous research has demonstrated the efficacy of Extreme Learning Machines (ELMs) for IoT security applications [10,11], highlighting their fast training speeds and reasonable generalization capabilities. However, more sophisticated Artificial Neural Network (ANN) architectures may offer superior performance for the complex patterns characteristic of network intrusions, particularly in specialized IoMT environments.

Contemporary neural network research has produced numerous architectural innovations, including multi-branch networks [12], concatenation operations, and shortcut connections [13], which have demonstrated remarkable success in computer vision and natural language processing. These advanced designs have yet to be thoroughly explored in the context of healthcare intrusion detection, where their capacity to model complex relationships between features could potentially yield significant performance improvements [14].

A parallel challenge in developing effective intrusion detection systems is managing the high dimensionality of network traffic data. Autencoders (AEs), a class of neural networks designed for unsupervised learning, offer a promising approach to dimensionality reduction by learning compressed representations of input data. Their ability to capture essential patterns while filtering noise makes them potentially valuable for preprocessing network traffic features before classification [15].

Despite these advances, several critical challenges persist in IoMT intrusion detection. First, class imbalance—where normal traffic significantly outnumbers attack instances—presents a major obstacle to developing effective detection models. This imbalance typically biases models toward the majority class, resulting in poor detection of the minority attack instances that are of greatest interest. Second, the high dimensionality and potential redundancy in network traffic features may impede model performance and increase computational requirements. Third, the relationship between feature dimensionality reduction and detection accuracy remains insufficiently explored, particularly in healthcare contexts.

This study addresses these challenges by presenting a comprehensive evaluation of advanced ANN architectures for intrusion detection in IoMT environments, with particular attention to the impact of autoencoder preprocessing for dimensionality reduction. We leverage the WUSTL-EHMS-2020 dataset [16], specifically designed for IoMT systems, to ensure the relevance and applicability of our findings. Our research focuses on implementing and evaluating multiple ANN architectures, from standard feedforward networks to sophisticated designs featuring dual-branch inputs, concatenation operations, and shortcut connections. We compare these architectures both with and without autoencoder dimensionality reduction to assess the trade-offs between feature compression and detection performance.

To mitigate the class imbalance inherent in network traffic data, we employ and compare three distinct approaches: Synthetic Minority Over-sampling Technique (SMOTE), weighted loss functions, and a hybrid over- and under-sampling strategy [17]. This comprehensive methodology enables us to identify optimal combinations of preprocessing techniques, neural architectures, and class balancing strategies for IoMT intrusion detection.

Contributions. This work advances the development of secure IoMT frameworks through improved intrusion detection techniques based on sophisticated neural network approaches. The key contributions of this research include:

1. A systematic evaluation of multiple advanced ANN architectures for intrusion detection in IoMT environments, including standard feedforward networks, dual-branch models with addition and concatenation operations, and networks incorporating shortcut connections.
2. Comprehensive assessment of autoencoder preprocessing for dimensionality reduction in intrusion detection, revealing critical trade-offs between feature compression and detection performance.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

3. Comparative analysis of three class imbalance mitigation strategies (SMOTE, weighted loss functions, and hybrid sampling) across different neural architectures, identifying optimal combinations for effective attack detection.

Outline. The remainder of this paper is structured as follows. Section 2 discusses related work in intrusion detection, with particular focus on neural network approaches and dimensionality reduction techniques. Section 3 details our methodology, including dataset description, class imbalance strategies, neural network architectures, and experimental setup. Section 4 presents our experimental results and analysis. Finally, Section 5 concludes with a summary of findings and directions for future research.

2. Related Work

The growing adoption of Internet of Medical Things (IoMT) devices has heightened the need for robust security mechanisms, particularly intrusion detection systems. This section reviews significant research contributions in this domain, highlighting various approaches to enhance intrusion detection in IoT and healthcare-specific environments.

2.1. Traditional Machine Learning Approaches

Early research on network intrusion detection extensively utilized the KDD'99 dataset [18]. Zhang et al. [19] proposed a Random Forest (RF) model for anomaly detection, achieving 95% accuracy with a 1% false-positive rate. However, as noted by Hady et al. [20], this dataset lacks healthcare specificity and fails to represent contemporary network environments.

Li et al. [21] employed an integrated approach combining clustering methods, ant colony algorithms, and support vector machines (SVM) on the KDD'99 dataset. Their classifier achieved 98.6% accuracy in cross-validation with a Matthews correlation coefficient of 0.861. To address the high dimensionality challenge, researchers explored feature reduction techniques. Tesfahun et al. [22] applied Information Gain (IG) with random forests, while Shah et al. [23] demonstrated that reduced feature sets improved Back Propagation Neural Network performance in terms of size, complexity, and generalization ability.

The NSL-KDD dataset [24] was introduced to overcome KDD'99 limitations, though it still doesn't fully represent modern network environments. Kale et al. [25] utilized this dataset in their three-phase deep learning framework combining K-means clustering, GANomaly, and Convolutional Neural Networks. Albulayhi et al. [26] proposed an innovative feature selection approach using mathematical set theory, achieving remarkable 99.98% classification accuracy.

Iwendi et al. [27] integrated RF with genetic algorithms for feature optimization, achieving a 98.81% detection rate with only a 0.8% false alarm rate across multiple datasets. Their approach demonstrated the critical importance of feature selection in enhancing intrusion detection system performance.

2.2. Neural Network and Advanced Learning Approaches

Neural networks have gained significant traction in intrusion detection research due to their adaptive learning capabilities. Cherif et al. [11] conducted a comprehensive evaluation of Extreme Learning Machine (ELM) models. Nayak et al. [28] proposed a hybrid model combining Bayesian optimization and ELM for IoMT security. Using the ToN_IoT dataset [29], their approach achieved impressive results: 0.990 for precision, recall, and F1 score. However, their study did not address the critical class imbalance problem inherent in intrusion detection datasets.

2.3. IoMT-Specific Approaches

Research specifically targeting IoMT environments has gained momentum recently. Hady et al. [20] explored the integration of both medical and network data to enhance intrusion detection in healthcare. They developed a real-time Enhanced Healthcare Monitoring System (EHMS) testbed and created a specialized IoMT dataset comprising over 16,000 records. Their findings demonstrated that

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

utilizing both network and biometric metrics as features in IDSs yields superior performance compared to using either type alone, with improvements ranging from 7% to 25%. To address class imbalance, they employed SMOTE as a resampling technique, with their SVM implementation achieving 92.46% accuracy while ANN exhibited a 92.98% AUC score.

Mohammed et al. [30] introduced an ensemble-based IDS for IoMT environments using the WUSTL-EHMS-2020 dataset. To mitigate class imbalance, they implemented random over-sampling, creating a balanced corpus of 28,540 entries. Their system achieved a 99.80% accuracy rate in testing and 99.96% average accuracy through 10-fold cross-validation, with a 0.9980 F1 score. However, their use of random oversampling potentially introduces overfitting risks by duplicating existing minority samples without introducing new information.

2.4. Research Gaps and Opportunities

Despite significant advances in intrusion detection systems for IoMT, several research gaps remain:

1. Most studies focus on either traditional machine learning or basic neural network structures, with limited exploration of advanced neural architectures specifically designed for healthcare intrusion detection.
2. While class imbalance is acknowledged as a challenge, comprehensive comparisons of different balancing techniques and their impact on various neural network architectures are scarce.
3. The interaction between dimensionality reduction techniques and different neural network designs remains underexplored, particularly in healthcare-specific contexts.
4. The impact of feature normalization and channel number optimization on neural network performance for intrusion detection has received insufficient attention.

Our research addresses these gaps by providing a systematic evaluation of multiple neural network architectures across different class balancing techniques, offering insights into optimization strategies for IoMT intrusion detection systems.

Table 1 summarizes key related works in this domain.

Table 1. Summary of Related Work.

Authors	Dataset	Methodology	Results
Zhang et al. [19]	KDD 1999	Random Forest (RF) for anomaly detection	95% accuracy, 1% false-positive rate
Li et al. [21]	KDD 1999	Clustering, Ant Colony Algorithm, SVM	98.62% accuracy, MCC of 0.861
Shah et al. [23]	KDD 1999	Information Gain (IG) for feature reduction	Improved model performance with reduced dataset
Tesfahun et al. [22]	KDD 1999	Random Forest with IG	Enhanced generalization capacity
Kale et al. [25]	NSL-KDD, CIC-IDS2018, TON IoT	Three-stage deep learning framework (K-means, GANomaly, CNN)	91.6% accuracy on NSL-KDD
Albulayhi et al. [26]	NSL-KDD	Feature selection using set theory	99.98% classification accuracy
Iwendi et al. [27]	NSL-KDD	RF with Genetic Algorithm for feature optimization	98.81% detection rate, 0.8% false alarm rate
Nayak et al. [28]	ToN_IoT	Bayesian Optimization and ELM	High precision and recall, but no class imbalance solution

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

Table 1. *Cont.*

Authors	Dataset	Methodology	Results
Hady et al. [20]	Custom dataset WUSTL-EHMS-2020 (16,000 records)	Integration of medical and network data using EHMS testbed	Improved performance by 7% to 25%; SVM accuracy 92.46%, ANN AUC 92.98%
Mohammed M. et al. [30]	WUSTL-EHMS-2020	Ensemble learning and explainable AI with random over sampling	99.96% accuracy and 0.998 F1 score
Cherif A. [11]	WUSTL-EHMS-2020	Multiple neural network architectures with three class balancing approaches	Dual-branch model: 94.03% accuracy, 0.8716 F1-score with weighted loss

3. Methodology

Our research methodology follows a systematic approach to evaluate the effectiveness of various ANN architectures and autoencoder preprocessing for intrusion detection in IoMT environments. Figure 1 illustrates the comprehensive workflow, highlighting the parallel processing paths with and without dimensionality reduction.

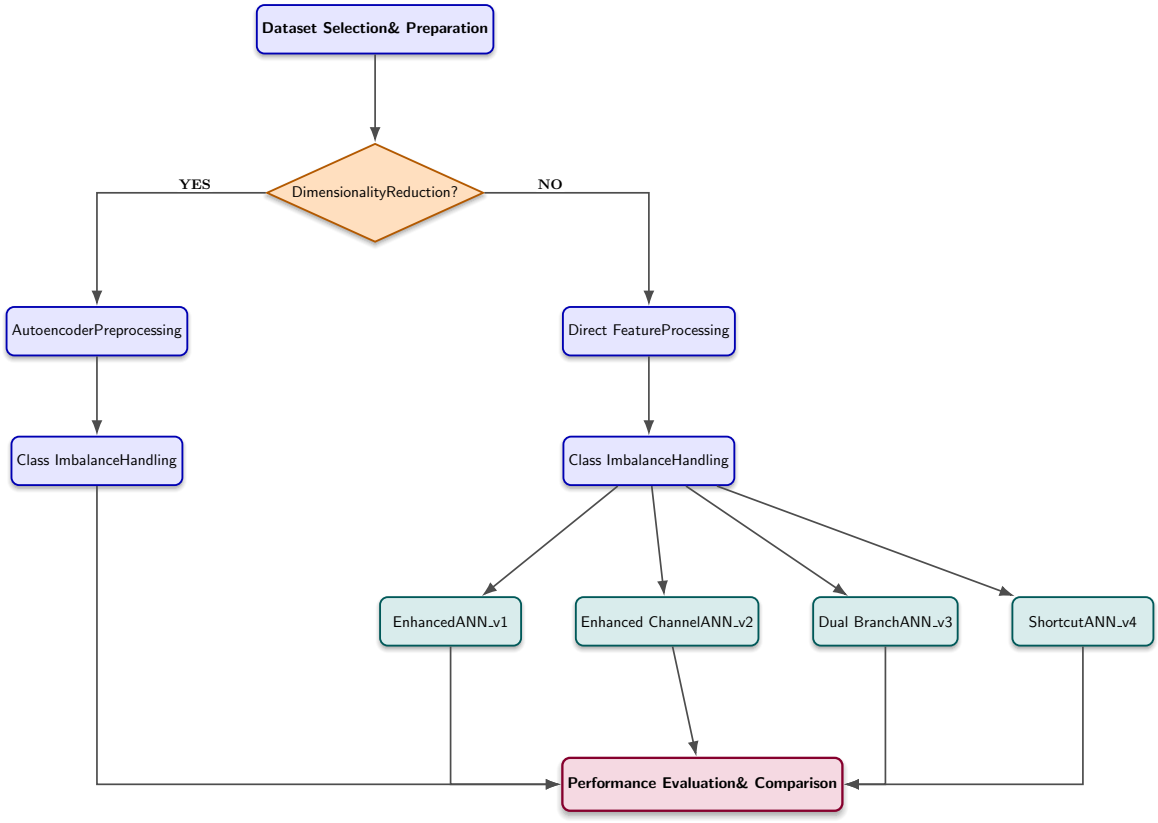

Figure 1. Methodology flowchart illustrating the parallel evaluation paths for neural network architectures with and without autoencoder dimensionality reduction.

3.1. Research Design

Our methodology encompasses six key phases:

1. **Dataset Selection and Preparation**: We utilize the WUSTL-EHMS-2020 dataset specifically designed for IoMT environments, performing initial cleaning and feature standardization.
2. **Feature Processing**: We implement two parallel processing paths:
   - **Direct Feature Processing**: Original features are standardized and used directly for model training.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

- **Autoencoder Preprocessing**: Features undergo dimensionality reduction through an autoencoder network before being fed to classification models.

3. **Class Imbalance Handling**: We implement and compare three distinct strategies:
   - Synthetic Minority Over-sampling Technique (SMOTE)
   - Weighted loss function approach
   - Hybrid over-under sampling method
4. **Neural Network Architecture Design**: We implement five distinct ANN architectures:
   - Standard ANN (baseline)
   - Enhanced Channel ANN (ANN_v1)
   - Dual-Branch Addition ANN (ANN_v2)
   - Dual-Branch Concatenation ANN (ANN_v3)
   - Shortcut Connection ANN (ANN_v4)
5. **Model Training and Validation**: Each architecture is trained with consistent hyperparameters across multiple class balancing configurations, using evaluation at regular intervals to assess performance.
6. **Performance Evaluation**: Models are evaluated using multiple metrics (AUC, Accuracy, Precision, Recall, and F1-score) to provide a comprehensive assessment of their detection capabilities.

This structured approach enables systematic comparison of different architectural designs and preprocessing strategies, facilitating identification of optimal configurations for IoMT intrusion detection.

3.2. Dataset Description

The WUSTL-EHMS-2020 dataset was specifically developed for IoMT cybersecurity research using an Enhanced Healthcare Monitoring System (EHMS) testbed [16,20]. This testbed captures both network flow metrics and patient biometric data in real-time, creating a realistic representation of IoMT environments.

The dataset incorporates two types of man-in-the-middle attacks:

- **Spoofing attacks**: Intercept communications between gateway and server, potentially exposing confidential patient information.
- **Data injection attacks**: Alter packets in transit, compromising data integrity.

The dataset comprises 44 features: 35 network flow metrics, 8 patient biometric indicators, and 1 label feature (1 for anomalous samples, 0 for normal samples). Samples containing MAC addresses associated with the attacker's laptop are labeled as attacks. After preprocessing and feature selection, our final input vector contains 34 dimensions, consisting of 26 network features and 8 biometric features.

Table 2 presents the statistical distribution of the dataset, highlighting the significant class imbalance that must be addressed.

Table 2. Dataset Statistical Information.

Measurement	Value
Size	4.4 MB
Normal samples	14,272 (87.5%)
Attack samples	2,046 (12.5%)
Total number of samples	16,318

3.3. Autoencoder for Dimensionality Reduction

A critical component of our methodology is the exploration of autoencoder networks for dimensionality reduction in preprocessing. Autoencoders are neural networks designed to learn efficient encodings of input data in an unsupervised manner. Figure 2 illustrates our autoencoder architecture.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

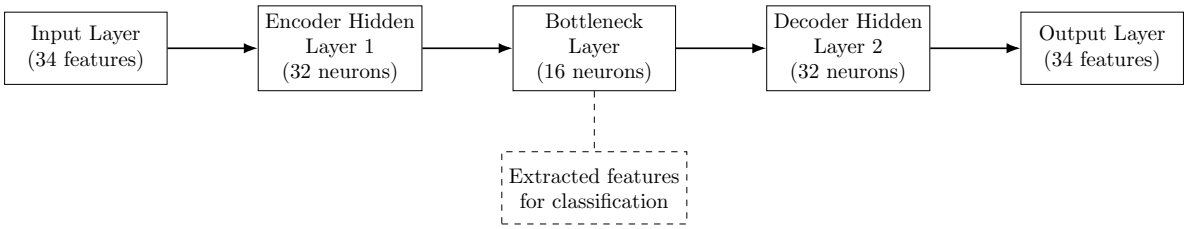

Figure 2. Autoencoder architecture for dimensionality reduction, converting 34-dimensional input to 16-dimensional representation.

The autoencoder consists of:

- **Encoder**: Compresses the 34-dimensional input features to a 16-dimensional bottleneck representation
- **Decoder**: Reconstructs the original 34-dimensional features from the bottleneck representation

During training, the autoencoder minimizes reconstruction error, learning to preserve the most important information while discarding noise and redundancy. After training, we discard the decoder and use only the encoder portion to transform input features before feeding them to classification models.

3.4. Class Imbalance Strategies

Class imbalance is a recurrent challenge in machine learning, particularly pronounced in domains such as healthcare, fraud detection, and security where minority classes often represent critical cases. In our dataset, attack samples constitute only 12.5% of the total, creating a significant imbalance that can lead to biased models favoring the majority class. To address this challenge, we implement and compare three distinct approaches:

3.4.1. Synthetic Minority Over-Sampling Technique (SMOTE)

This method generates synthetic samples in the feature space of the minority class, creating new instances that are not exact copies but share characteristics with existing minority samples. SMOTE operates by selecting a minority class instance and its k-nearest neighbors, then generating new samples along the line segments connecting the selected instance to its neighbors [31]. This approach increases the minority class representation while introducing greater variability than simple duplication.

3.4.2. Hybrid Over-Under Sampling

This strategy combines under-sampling of the majority class with over-sampling of the minority class. Specifically, we select 50% of the negative (normal) samples and then duplicate the positive (attack) samples to achieve an even distribution between classes. This balanced approach reduces computational burden while maintaining sufficient representation of normal network behavior [32].

3.4.3. Weighted Cross-Entropy Loss Function

Rather than altering the dataset distribution, this method maintains the original class distribution but assigns different weights to each class in the cross-entropy loss function. The weight for each class is inversely proportional to its frequency in the training set, effectively penalizing misclassification of minority class samples more heavily. The weighted cross-entropy loss function is formulated as:

$$L_{CE}^{w}(y, \hat{y}) = -\sum_{i=1}^{N} [w_1 y_i \log(\hat{y}_i) + w_0 (1 - y_i) \log(1 - \hat{y}_i)] \quad (1)$$

where y_i is the true label, $\hat{y}_i$ is the predicted probability, and w_1 and w_0 are the weights for the positive (attack) and negative (normal) classes, respectively. In our implementation, we use weights of $w_0 = 0.15$ and $w_1 = 0.8$, assigning approximately 5.3 times more importance to correctly classifying attack instances compared to normal instances.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

This approach preserves the natural data distribution while still addressing the learning bias [33], which is particularly valuable in intrusion detection where maintaining the realistic ratio of normal to attack traffic is important for model generalization.

3.5. Neural Network Architectures

In this research, we implement and evaluate five distinct neural network architectures for intrusion detection, each designed to explore different architectural strategies.

3.5.1. Standard ANN (ANN)

Our baseline architecture consists of a deep multi-layer perceptron with eight layers processing 34-dimensional input features:

- **Input Layer**: 34 features (most relevant network and biometric parameters)
- **Hidden Layers**: Seven fully-connected layers with dimensions [40, 40, 20, 10, 10, 10, 10]
- **Output Layer**: 2 neurons for binary classification
- **Activation**: ReLU for all hidden layers

This architecture features a relatively deep structure with a moderate number of parameters, creating a progressive dimensionality reduction from 40 neurons in the initial hidden layers down to 10 neurons in the deeper layers.

3.5.2. Enhanced Channel ANN (ANN_v1)

This architecture enhances the standard ANN by significantly increasing channel width while maintaining the same depth:

- **Input Layer**: 34 features
- **Hidden Layers**: Seven fully-connected layers with dimensions [256, 256, 128, 64, 64, 64, 64]
- **Output Layer**: 2 neurons for binary classification
- **Activation**: ReLU for all hidden layers

The substantial increase in channel width (up to 6.4 times wider than the standard ANN) provides greater representational capacity, allowing the network to learn more complex feature interactions and potentially improve classification performance.

3.5.3. Dual-Branch Models (ANN_v2 and ANN_v3)

Our dual-branch architectures explicitly model the multimodal nature of the IoMT data by processing network and biometric features through separate pathways before fusion (see Figure 3). Both models follow the same preprocessing approach:

- **Data Splitting**: The 34-dimensional input is divided into network metrics (first 26 features) and biometric parameters (remaining 8 features)
- **Specialized Processing**: Each feature type is processed through dedicated network branches
- **Different Fusion Mechanisms**: The two models differ in how they combine branch outputs

Dual-Branch Addition ANN (ANN_v2)

This architecture processes the split data through parallel branches and combines them through scaled addition:

- **Network Branch**: Two fully-connected layers [256, 256] with ReLU activation process the 26 network features
- **Biometric Branch**: Two fully-connected layers [256, 256] with ReLU activation process the 8 biometric features
- **Fusion Mechanism**: Element-wise addition of branch outputs multiplied by 0.5 (averaging operation)
- **Shared Layers**: Three layers [128, 64, 64] with ReLU activation and dropout (0.4)

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

- **Output Layer**: 2 neurons for binary classification

The addition operation forces the model to learn complementary representations in both branches that can be effectively combined through summation. The scaling factor of 0.5 maintains a consistent magnitude in the fused representation.

Dual-Branch Concatenation ANN (ANN_v3)

This architecture maintains the same parallel branch structure but employs concatenation for feature fusion:

- **Network Branch**: Two fully-connected layers [256, 256] with ReLU activation process the 26 network features
- **Biometric Branch**: Two fully-connected layers [256, 256] with ReLU activation process the 8 biometric features
- **Fusion Mechanism**: Concatenation of branch outputs, resulting in a 512-dimensional feature vector
- **Shared Layers**: Three layers [256, 128, 64] with ReLU activation and dropout (0.4)
- **Output Layer**: 2 neurons for binary classification

The concatenation approach preserves all features from both branches without mixing them, allowing subsequent layers to learn more complex relationships between different feature types. The first shared layer after concatenation is wider (256 neurons vs. 128 in ANN_v2) to accommodate the larger input dimension.

The input handling approach for these dual-branch models leverages domain knowledge about the fundamental difference between network traffic metrics and patient biometric signals. By processing these distinct data types through specialized branches before fusion, the models can learn modality-specific patterns that might be obscured in a unified architecture.

3.5.4. Shortcut Connection ANN (ANN_v4)

The Shortcut Connection ANN architecture incorporates residual-like skip connections that allow information to bypass certain layers, facilitating gradient flow during backpropagation as shown in Figure 4. Unlike the dual-branch models, this architecture processes the complete 34-dimensional feature vector without domain-specific splitting:

- **Input Layer**: 34 features (combined network and biometric parameters)
- **Hidden Layers**: Seven fully-connected layers with dimensions [256, 256, 128, 64, 64, 64, 64]
- **Shortcut Connections**: Four identity shortcuts creating residual blocks
  - Layer 1 output added to Layer 2 output
  - Layer 4 output added to Layer 5 output
  - Layer 5 output (with previous shortcut) added to Layer 6 output
  - Layer 6 output (with previous shortcut) added to Layer 7 output
- **Output Layer**: 2 neurons for binary classification
- **Activation**: ReLU for all hidden layers

This architecture maintains a similar parameter count to the Enhanced Channel ANN (ANN_v1) but introduces shortcut connections that create multiple paths for gradient flow. The successive addition of features through these shortcuts allows later layers to refine representations while maintaining access to earlier features, potentially improving the network's ability to learn complex patterns at different levels of abstraction.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

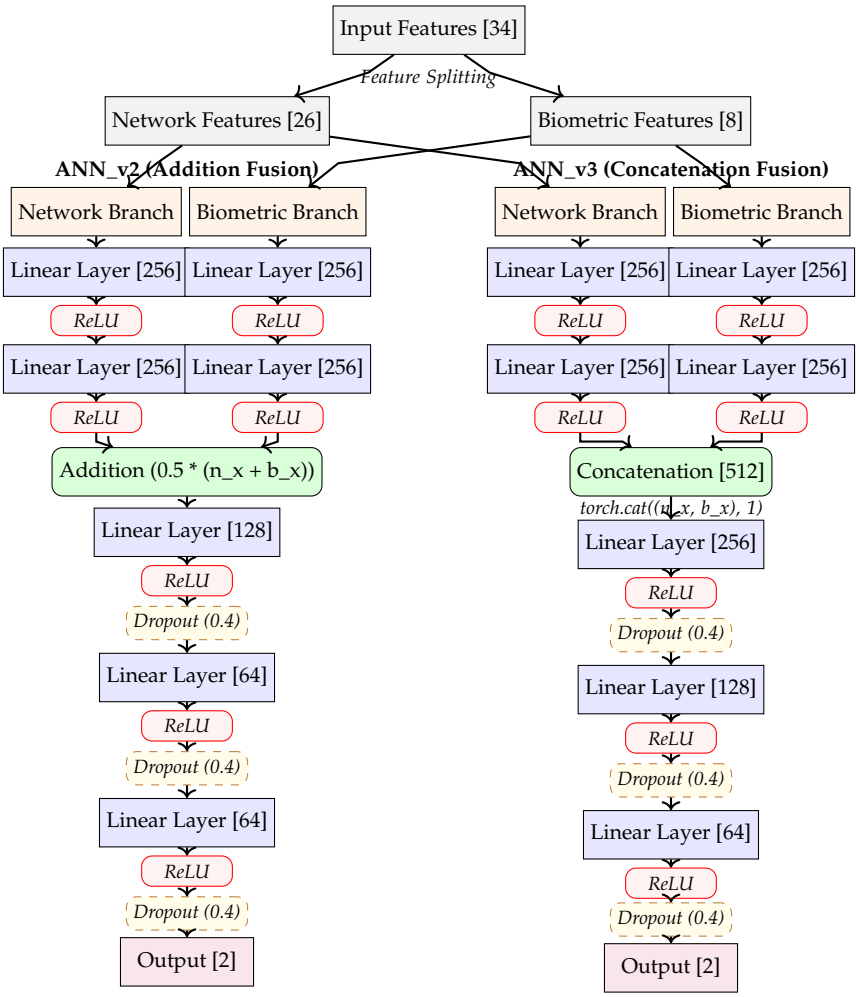

Figure 3. Architecture diagram of the dual-branch ANN models (ANN_v2 and ANN_v3), illustrating feature splitting, parallel processing paths, and different fusion mechanisms. The left model (ANN_v2) uses element-wise addition to combine features from both branches, while the right model (ANN_v3) uses concatenation to preserve all information from both branches. Both architectures employ dropout regularization after each shared layer to prevent overfitting.

3.6. Training and Hyperparameter Settings

All neural network models are trained with consistent hyperparameter settings to enable fair comparison, as summarized in Table 3. The learning rate is set to 1×10^{-3}, with a batch size of 64 and weight decay of 5×10^{-4} for regularization. We employ the AdamW optimizer, which provides adaptive learning rates with improved weight decay regularization.

For addressing class imbalance using the weighted loss approach, we modify the standard cross-entropy loss by applying class weights of [0.15, 0.8] for normal and attack classes, respectively. This assigns approximately 5.3 times more importance to correctly classifying attack instances compared to normal instances.

Training is conducted for 200-500 epochs depending on the model architecture, with evaluation performed every 50 epochs to monitor performance metrics including AUC, accuracy, precision, recall, and F1-score. This regular evaluation enables tracking of model convergence and helps identify the best-performing model checkpoint.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

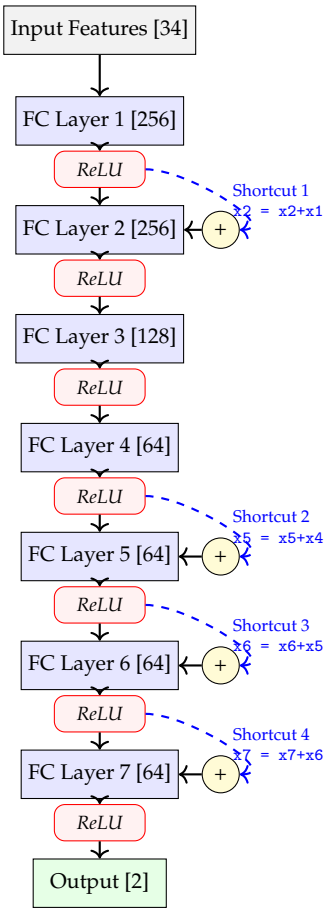

Figure 4. Architecture diagram of ANN_v4, showing ReLU activations and residual-like skip connections. The network features eight layers with increasing abstraction and four shortcut connections that help gradient flow during training by creating alternative paths for information flow. Each shortcut connection corresponds to an addition operation in the forward pass of the model.

Table 3. Hyperparameter Settings.

Hyperparameter	Value
Learning rate	1×10^{-3}
Batch size	64
Weight decay	5×10^{-4}
Optimizer	AdamW
Loss function	Cross Entropy (standard or weighted)
Evaluation interval	50 epochs
Maximum epochs	200-500 (architecture dependent)

3.7. Evaluation Metrics

To comprehensively assess model performance in the imbalanced dataset context of intrusion detection, we employ five complementary evaluation metrics:

- **Area Under the ROC Curve (AUC)**: Measures the model's discrimination capability across all possible classification thresholds. AUC represents the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance. Mathematically:

$$\text{AUC} = \int_0^1 \text{TPR}(t) \cdot \text{FPR}(t)\, dt \tag{2}$$

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

where TPR is the true positive rate and FPR is the false positive rate at threshold t. AUC values range from 0.5 (random classification) to 1.0 (perfect classification). This metric is particularly valuable for imbalanced datasets as it is insensitive to class distribution.

- **Accuracy (ACC)**: The proportion of correctly classified instances among all instances:

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3}$$

While intuitive, accuracy can be misleading in imbalanced datasets, as high accuracy can be achieved by simply classifying all instances as the majority class.

- **Precision (PR)**: The proportion of true positive predictions among all positive predictions:

$$\text{PR} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4}$$

High precision indicates a low false positive rate, which is particularly important in intrusion detection systems where false alarms can lead to alert fatigue and reduced trust in the system.

- **Recall (RC)**: Also known as sensitivity or true positive rate, recall measures the proportion of actual positives that are correctly identified:

$$\text{RC} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

High recall indicates that the model successfully captures most attack instances, which is critical in security applications where missing an attack (false negative) can have severe consequences.

- **F1-score (F1)**: The harmonic mean of precision and recall, providing a balance between these two potentially competing metrics:

$$\text{F1} = 2 \cdot \frac{\text{PR} \cdot \text{RC}}{\text{PR} + \text{RC}} \tag{6}$$

F1-score ranges from 0 to 1, with higher values indicating better performance. This metric is particularly useful when seeking a balance between precision and recall.

In the context of intrusion detection, these metrics offer complementary perspectives on model performance. High precision is important to avoid overwhelming security administrators with false alarms, while high recall is essential to ensure that actual attacks are not missed. The F1-score helps balance these requirements, while AUC provides a threshold-independent assessment of discriminative ability. We report all these metrics using the macro-averaging approach, which calculates metrics independently for each class and then takes the average, treating all classes equally regardless of their frequency in the dataset.

By examining these metrics collectively rather than focusing on any single measure, we gain a more comprehensive understanding of each model's strengths and weaknesses in detecting intrusions in IoMT environments.

4. Results and Discussion

We evaluate all model architectures using five key performance metrics: Area Under the ROC Curve (AUC), Accuracy (ACC), Precision (PR), Recall (RC), and F1-score (F1). The comprehensive results are presented in Table 4.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

Table 4. Performance Results of Different Neural Network Architectures and Class Balancing Methods.

Model	Class Balancing Method	AUC	ACC	PR	RC	F1
ANN	SMOTE	0.8491	0.8753	0.7427	0.8491	0.7783
ANN_v1	SMOTE	0.8544	0.8983	0.7750	0.8544	0.8062
ANN_v2	SMOTE	0.8766	0.8955	0.7721	0.8766	0.8096
ANN_v3	SMOTE	0.8740	0.9032	0.7851	0.8740	0.8182
ANN_v4	SMOTE	0.8554	0.9035	0.7839	0.8554	0.8129
AE+ANN_v4	SMOTE	0.7485	0.9308	**0.9383**	0.7485	0.8085
ANN	Hybrid	0.8518	0.9179	0.8132	0.8518	0.8308
ANN_v1	Hybrid	0.8577	0.9213	0.8201	0.8577	0.8373
ANN_v2	Hybrid	0.8750	0.9323	0.8437	0.8750	0.8582
ANN_v3	Hybrid	0.8671	0.9203	0.8163	0.8671	0.8387
ANN_v4	Hybrid	0.8739	0.9151	0.8047	0.8739	0.8335
AE+ANN_v4	Hybrid	0.7925	0.9059	0.7942	0.7925	0.7934
ANN	Weighted cross-entropy Loss	0.8588	0.9145	0.8049	0.8588	0.8283
ANN_v1	Weighted cross-entropy Loss	0.8559	0.9197	0.8168	0.8559	0.8345
ANN_v2	Weighted cross-entropy Loss	**0.8786**	**0.9403**	0.8650	**0.8786**	**0.8716**
ANN_v3	Weighted cross-entropy Loss	0.8467	0.8161	0.6938	0.8467	0.7216
ANN_v4	Weighted cross-entropy Loss	0.8534	0.8382	0.7096	0.8534	0.7431
AE+ANN_v4	Weighted cross-entropy Loss	0.7463	0.9200	0.8705	0.7463	0.7909

4.1. Impact of Neural Network Architecture

The results demonstrate a clear progression in performance as we move from simple ELM architectures to more sophisticated neural network designs:

- **Standard vs. Enhanced ANNs**: The enhancement through increased channel numbers (ANN_v1) consistently improves performance, confirming that greater parametrization enables better feature learning for this task.
- **Dual-Branch Architectures**: The dual-branch models (ANN_v2 and ANN_v3) consistently achieve the highest performance across all balancing methods. The addition-based combination (ANN_v2) generally outperforms the concatenation approach (ANN_v3), suggesting that the summation of features from parallel branches provides more effective feature integration for intrusion detection.
- **Shortcut Connections**: The ANN_v4 model with shortcut connections shows comparable performance to ANN_v1, indicating that for this particular task and dataset size, shortcut connections do not provide substantial additional benefits over simply increasing channel numbers.

Consequently, the dual-branch architecture with addition operations (ANN_v2) emerges as the most effective design, achieving the highest performance metrics across different balancing methods.

4.2. Effectiveness of Class Balancing Methods

Our comparison of three class balancing approaches reveals important insights:

- **SMOTE**: While SMOTE improves model performance compared to no balancing (not shown), it generally yields lower accuracy and precision compared to the other balancing methods. However, it maintains reasonably good recall, indicating its ability to identify attack instances.
- **Hybrid Over-Under Sampling**: This approach consistently outperforms SMOTE across all architectures, achieving better balance between precision and recall. The improved performance suggests that selective under-sampling of majority class instances combined with minority class duplication provides an effective balance for intrusion detection.
- **Weighted Cross-entropy Loss Function**: This method yields the highest overall performance, particularly when combined with the ANN_v2 architecture (0.9403 accuracy, 0.8716 F1-score). It demonstrates superior precision compared to other methods while maintaining competitive recall. This suggests that preserving the original data distribution while adjusting the learning objective is most effective for this task.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

The superior performance of the weighted cross-entropy loss approach indicates that maintaining the natural distribution of network traffic while adjusting the learning process is more effective than artificially altering the dataset distribution.

4.3. Impact of Dimensionality Reduction

The autoencoder-based dimensionality reduction (AE+ANN_v4) shows mixed results:

- With SMOTE, the AE+ANN_v4 model achieves the highest precision (0.9383) among all SMOTE-based models, but with significantly lower recall (0.7485) compared to other ANN architectures.
- With the weighted cross-entropy loss function, the AE+ANN_v4 model shows similar trends: high precision (0.8705) but reduced recall (0.7463), resulting in lower overall F1-score (0.7909) compared to other ANN architectures.

These results suggest that while dimensionality reduction through autoencoders can improve precision by creating a more compact feature representation, it typically results in information loss that negatively impacts recall. For intrusion detection in IoMT environments, where detecting all potential attacks is critical, this trade-off may not be desirable.

4.4. Key Findings and Practical Implications

Based on our comprehensive evaluation, we derive several key findings with practical implications for IoMT intrusion detection:

1. **Feature Normalization**: Our experiments confirm that proper feature normalization is crucial for neural network performance in intrusion detection tasks. Standardization ensures consistent scaling across diverse network and biometric features, facilitating more effective learning.
2. **Architectural Considerations**: Dual-branch neural network architectures with addition operations (ANN_v2) consistently outperform other designs except in precision which may be enhanced with experts reviewing alerts or post-filters to avoid false positives, suggesting that parallel processing paths with feature integration through addition is particularly effective for capturing the complex patterns indicative of network intrusions.
3. **Class Balancing Strategy**: Weighted cross-entropy loss functions provide the most effective approach to addressing class imbalance for intrusion detection, outperforming both SMOTE and hybrid sampling strategies across most architectures. This suggests that maintaining the natural distribution of network traffic data while adjusting the learning objective is preferable to artificially altering the dataset distribution.
4. **Dimensionality Reduction Trade-offs**: While autoencoders can simplify models through dimensionality reduction, the associated information loss typically reduces recall, which is particularly problematic for security applications where missing attack instances (false negatives) can have serious consequences.
5. **Optimal Configuration**: The combination of ANN_v2 architecture with weighted loss function emerges as the most effective configuration for IoMT intrusion detection, achieving 94.03% accuracy and 0.8716 F1-score. This configuration offers an excellent balance between precision and recall, making it well-suited for real-world deployment.

4.5. Comparative Analysis with Previous Work

In [11], the application of ELM architectures for intrusion detection in IoMT environments was explored using the same dataset. That study established baseline performance metrics for ELM models with varying hidden layer sizes (64, 128, and 256 nodes) and demonstrated that ELM (256) with SMOTE could achieve an AUC of 0.7789 and an F1-score of 0.7223. While promising, these results indicated limitations in the ELM's ability to capture the complex patterns necessary for optimal intrusion detection.

The current study represents a significant advancement over this previous work in several key aspects:

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

- **Performance Improvement**: Our dual-branch ANN architecture with addition operations (ANN_v2) combined with weighted loss function achieves an AUC of 0.8786 and an F1-score of 0.8716, representing relative improvements of 12.8% and 20.7% respectively over the best ELM model from [11].
- **Architectural Sophistication**: Moving beyond the single hidden layer constraint of ELM, our current work explores multi-layer architectures with various connectivity patterns, demonstrating that architectural design choices significantly impact detection performance.
- **Dimensionality Reduction Analysis**: While [11] work focused on direct classification of input features, this study provides critical insights into the trade-offs associated with autoencoder preprocessing, revealing that the information loss during dimensionality reduction compromises recall—a crucial metric for security applications.

When compared with other recent studies using the same dataset, our work offers distinctive contributions. Hady et al. [20] achieved an accuracy of 92.46% with SVM and an AUC of 92.98% with ANN by integrating both medical and network data. While their approach of combining different data types is complementary to ours, our best model achieves superior accuracy (94.03%) through architectural innovation.

Mohammed et al. [30] reported impressive results using an ensemble-based approach with random over-sampling, achieving 99.80% accuracy and a 0.9980 F1-score. However, their use of random over-sampling (direct duplication of minority samples) potentially introduces overfitting risks, as the model may memorize specific attack instances rather than learning generalizable patterns. In contrast, our approach employs more sophisticated class balancing techniques and achieves robust performance without ensemble methods, offering a more streamlined deployment path.

Table 5 summarizes these comparisons, highlighting the progression from ELM-based approaches to our current advanced ANN architectures.

Table 5. Comparative Analysis of IoMT Intrusion Detection Approaches on WUSTL-EHMS-2020 Dataset.

Study	Approach	Acc	F1	AUC	PR	RC
Cherif A. [11]	ELM (256) + SMOTE	0.8444	0.7223	0.7789	0.6949	0.7789
Cherif A. [11]	ELM (256) + Weighted cross-entropy Loss	0.9305	0.8037	0.7404	0.9518	0.7404
Hady et al. [20]	SVM with SMOTE	0.9246	Not reported	0.8237	Not reported	Not reported
Hady et al. [20]	ANN with SMOTE	0.9040	Not reported	0.9342	Not reported	Not reported
Mohammed et al. [30]	Ensemble with random over-sampling	0.9980	0.9980	Not reported	0.9980	0.9980
Current study	ANN_v2 + Weighted Loss	0.9403	0.8716	0.8786	0.8650	0.8786
Current study	AE+ANN_v4 + SMOTE	0.9308	0.8085	0.7485	0.9383	0.7485

This comparative analysis highlights the evolutionary progression in IoMT intrusion detection techniques. While the previous ELM-based work established important baselines, the current study's advanced ANN architectures offer substantial performance improvements. The analysis of autoencoder preprocessing provides crucial insights into the limitations of dimensionality reduction in security applications, where the preservation of all potentially relevant features may be more important than computational efficiency. These findings contribute valuable knowledge to guide future development of intrusion detection systems for healthcare environments, where the balance between detection accuracy and computational efficiency must be carefully optimized.

5. Conclusions

This study conducted a comprehensive evaluation of advanced neural network architectures and autoencoder preprocessing for intrusion detection in Internet of Medical Things (IoMT) environments. Our systematic investigation has yielded several significant findings with important implications for the design and implementation of security systems in healthcare contexts.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

The results demonstrate that architectural design significantly impacts intrusion detection performance. Our dual-branch neural network with addition operations (ANN_v2) combined with weighted cross-entropy loss function achieved superior performance (0.9403 accuracy, 0.8786 AUC, and 0.8716 F1-score), substantially outperforming conventional architectures. This finding underscores the value of specialized network designs that explicitly account for the heterogeneous nature of IoMT data, processing network and biometric features through separate pathways before integration.

Our analysis of dimensionality reduction through autoencoders revealed an important trade-off: while autoencoder preprocessing improved precision (up to 93.83%), it consistently reduced recall (down to 74.85%). This precision-recall trade-off is particularly problematic in security-critical applications like intrusion detection, where missed attacks (false negatives) can have severe consequences. The results suggest that preserving the full feature space is preferable for IoMT intrusion detection, as the information loss during compression appears to disproportionately affect the model's ability to identify attack instances.

Among class imbalance mitigation strategies, weighted loss functions consistently outperformed both SMOTE and hybrid sampling approaches across most architectures. This indicates that maintaining the natural distribution of network traffic while adjusting the learning objective provides more effective model training than artificially altering the dataset distribution. This approach better preserves the realistic context in which intrusion detection systems must operate, where normal traffic significantly outnumbers attack instances.

Several limitations of this study present opportunities for future research. First, while our models achieve high performance on the WUSTL-EHMS-2020 dataset, validation across multiple IoMT datasets would strengthen the generalizability of our findings. Second, the current work focused primarily on binary classification (normal vs. attack); future research should extend to multi-class classification to distinguish between different attack types, enabling more targeted security responses.

Acknowledgments: The authors would like to thank Prince Sultan University for their support.

References

1. Islam, S.R.; Kwak, D.; Kabir, M.H.; Hossain, M.; Kwak, K.S. The Internet of Things for health care: a comprehensive survey. *IEEE Access* **2015**, *3*, 678–708.
2. Alrawi, O.; Lever, C.; Antonakakis, M.; Monrose, F. Security analysis of IoT devices. *ACM Transactions on Privacy and Security (TOPS)* **2019**, *22*, 1–36.
3. Osama, M.; Ateya, A.A.; Sayed, M.S.; Hammad, M.; Pławiak, P.; Abd El-Latif, A.A.; Elsayed, R.A. Internet of medical things and healthcare 4.0: Trends, requirements, challenges, and research directions. *Sensors* **2023**, *23*, 7435.
4. Zarpelão, B.B.; Miani, R.S.; Kawakani, C.T.; de Alvarenga, S.C. A survey of intrusion detection in Internet of Things. *Journal of Network and Computer Applications* **2017**, *84*, 25–37.
5. Butun, I.; Morgera, S.D.; Sankar, R. A survey of intrusion detection systems in wireless sensor networks. *IEEE communications surveys & tutorials* **2013**, *16*, 266–282.
6. Mujahid, M.; Mirdad, A.R.; Alamri, F.S.; Ara, A.; Khan, A. Software defined network intrusion system to detect malicious attacks in computer Internet of Things security using deep extractor supervised random forest technique. *PeerJ Computer Science* **2025**, *11*, e3103.
7. Farhan, S.; Mubashir, J.; Haq, Y.U.; Mahmood, T.; Rehman, A. Enhancing network security: an intrusion detection system using residual network-based convolutional neural network. *Cluster Computing* **2025**, *28*, 251.
8. Alrayes, F.S.; Zakariah, M.; Amin, S.U.; Khan, Z.I.; Alqurni, J.S. CNN Channel Attention Intrusion Detection System Using NSL-KDD Dataset. *Computers, Materials & Continua* **2024**, *79*.
9. Mitchell, R.; Chen, I.R. A survey of intrusion detection techniques for cyber-physical systems. *ACM Computing Surveys (CSUR)* **2014**, *46*, 1–29.
10. Farnaaz, N.; Jabbar, M. Random forest modeling for network intrusion detection system. *Procedia Computer Science* **2016**, *89*, 213–217.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

11. Cherif, A. Intrusion Detection for Internet of Medical Things (IoMT) using Extreme Learning Machine. In Proceedings of the 2025 2nd International Conference on Advanced Innovations in Smart Cities (ICAISC), 2025, pp. 1–7. https://doi.org/10.1109/ICAISC64594.2025.10959678.
12. Yamashita, T.; Hirasawa, K.; Hu, J.; Murata, J. Multi-branch structure of layered neural networks. In Proceedings of the Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02., 2002, Vol. 1, pp. 243–247 vol.1. https://doi.org/10.1109/ICONIP.2002.1202170.
13. Geirhos, R.; Jacobsen, J.H.; Michaelis, C.; Zemel, R.; Brendel, W.; Bethge, M.; Wichmann, F.A. Shortcut learning in deep neural networks. *Nature Machine Intelligence* **2020**, *2*, 665–673. https://doi.org/10.1038/s42256-020-00257-z.
14. Diro, A.A.; Chilamkurti, N. Distributed attack detection scheme using deep learning approach for Internet of Things. *Future Generation Computer Systems* **2018**, *82*, 761–768.
15. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507.
16. Hady, A.A. WUSTL-EHMS-2020 . https://www.cse.wustl.edu/~jain/ehms/index.html, 2020. [Online; accessed 20-October-2024].
17. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research* **2002**, *16*, 321–357.
18. Information.; Computer Science University of California, I. KDD Cup 1999 Data . http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html, Oct. 2007. [Online; accessed 19-October-2024].
19. Zhang, J.; Zulkernine, M.; Haque, A. Random-Forests-Based Network Intrusion Detection Systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **2008**, *38*, 649–659. https://doi.org/10.1109/TSMCC.2008.923876.
20. Hady, A.A.; Ghubaish, A.; Salman, T.; Unal, D.; Jain, R. Intrusion Detection System for Healthcare Systems Using Medical and Network Data: A Comparison Study. *IEEE Access* **2020**, *8*, 106576–106584. https://doi.org/10.1109/ACCESS.2020.3000421.
21. Li, Y.; Xia, J.; Zhang, S.; Yan, J.; Ai, X.; Dai, K. An efficient intrusion detection system based on support vector machines and gradually feature removal method. *Expert Systems with Applications* **2012**, *39*, 424–430. https://doi.org/https://doi.org/10.1016/j.eswa.2011.07.032.
22. Tesfahun, A.; Bhaskari, D.L. Intrusion Detection Using Random Forests Classifier with SMOTE and Feature Reduction. In Proceedings of the 2013 International Conference on Cloud & Ubiquitous Computing & Emerging Technologies, 2013, pp. 127–132. https://doi.org/10.1109/CUBE.2013.31.
23. Shah, B.; Trivedi, B.H. Reducing Features of KDD CUP 1999 Dataset for Anomaly Detection Using Back Propagation Neural Network. In Proceedings of the 2015 Fifth International Conference on Advanced Computing & Communication Technologies, 2015, pp. 247–251. https://doi.org/10.1109/ACCT.2015.131.
24. ZAIB, M.H. NSL KDD Dataset. https://www.kaggle.com/datasets/hassan06/nslkdd, 2024. [Online; accessed 19-July-2024].
25. Kale, R.; Lu, Z.; Fok, K.W.; Thing, V.L.L. A Hybrid Deep Learning Anomaly Detection Framework for Intrusion Detection. In Proceedings of the 2022 IEEE 8th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS), 2022, pp. 137–142. https://doi.org/10.1109/BigDataSecurityHPSCIDS54978.2022.00034.
26. Albulayhi, K.; Abu Al-Haija, Q.; Alsuhibany, S.A.; Jillepalli, A.A.; Ashrafuzzaman, M.; Sheldon, F.T. IoT Intrusion Detection Using Machine Learning with a Novel High Performing Feature Selection Method. *Applied Sciences* **2022**, *12*, 5015. https://doi.org/10.3390/app12105015.
27. Iwendi, C.; Anajemba, J.H.; Biamba, C.; Ngabo, D. Security of Things Intrusion Detection System for Smart Healthcare. *Electronics* **2021**, *10*. https://doi.org/10.3390/electronics10121375.
28. Nayak, J.; Meher, S.K.; Souri, A.; Naik, B.; Vimal, S. Extreme learning machine and bayesian optimization-driven intelligent framework for IoMT cyber-attack detection. *J. Supercomput.* **2022**, *78*, 14866–14891.
29. Nour, M. ToN_IoT Datasets. https://research.unsw.edu.au/projects/toniot-datasets, 2024. [Online; accessed 19-July-2024].
30. Alani, M.M.; Mashatan, A.; Miri, A. Explainable Ensemble-Based Detection of Cyber Attacks on Internet of Medical Things. In Proceedings of the 2023 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), 2023, pp. 0609–0614. https://doi.org/10.1109/DASC/PiCom/CBDCom/Cy59711.2023.10361448.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2025 doi:10.20944/preprints202510.1779.v1

31. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research* **2002**, *16*, 321–357.
32. Liu, X.Y.; Wu, J.; Zhou, Z.H. Exploratory undersampling for class-imbalance learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **2020**, *39*, 539–550.
33. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Doll'ar, P. Focal loss for dense object detection. In Proceedings of the Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2980–2988.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.